



HAL
open science

Quantitative Transmission Electron Microscopy Study of III-Nitrides Semiconductors Nanostructures

Maxim Korytov

► **To cite this version:**

Maxim Korytov. Quantitative Transmission Electron Microscopy Study of III-Nitrides Semiconductors Nanostructures. Condensed Matter [cond-mat]. Université Nice Sophia Antipolis, 2010. English. NNT: . tel-00520605

HAL Id: tel-00520605

<https://theses.hal.science/tel-00520605>

Submitted on 23 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE NICE-SOPHIA ANTIPOLIS - UFR Sciences
Ecole Doctorale de Sciences Fondamentales et Appliquées

T H E S E

pour obtenir le titre de

Docteur en Sciences

de l'UNIVERSITE de Nice-Sophia Antipolis

Discipline : Physique

présentée et soutenue par

Maxim KORYTOV

**Quantitative Transmission Electron Microscopy Study of
III-Nitride Semiconductor Nanostructures**

Thèse dirigée par Philippe VENNÉGUÈS

soutenue le *26.04.2010*

Jury :

M. Martin ALBRECHT	Chercheur/docteur, IKZ, Berlin	Examineur
M. Michel GENDRY	Directeur de Recherche, CNRS, Lyon	Rapporteur
M. Colin HUMPHREYS	Professeur, University of Cambridge	Examineur
M. Thomas NEISIUS	Ingénieur/docteur, CNRS, Marseille	Examineur
M. Jean-Luc ROUVIÈRE	Ingénieur/docteur, CEA, Grenoble	Rapporteur
M. Philippe VENNÉGUÈS	Ingénieur/docteur, CNRS, Valbonne	Directeur de Thèse
M. Borge VINTER	Professeur, University of Nice	Président

ACKNOWLEDGEMENTS

The present thesis has been carried out in the **Center de Recherche sur l'Hétéro-Epitaxie et ses Applications** (CRHEA) of the Centre National de la Recherche Scientifique (CNRS). I would like to thank all people from the CRHEA laboratory for their support and concern.

I thank **Michel Gendry** and **Jean-Luc Rouvière** for accepting to be a member of the jury and reporting this work. My acknowledge also to **Martin Albrecht**, **Colin Humphreys**, **Thomas Neisius** and **Borge Vinter** for their participation in my jury.

I feel sincere gratitude to **Philippe Vennéguès**, who familiarized me with various techniques of transmission electron microscopy and successfully supervised my work during these three years. Due to him I have mastered a complicated, but an exciting profession of researcher.

Julien Brault and **Thomas Huault** I want to thank you for growing the top-quality samples and for being open for discussions on any epitaxy-related topics. Special appreciations to Julien for providing me essential papers on the nitride semiconductor growth. I also thank a lot **Benjamin Damilano** for always giving competent answers to my diverse questions. I want to express my gratitude to **Mathieu Leroux**, for explaining me the finest nuances of photoluminescence, as well as for critical corrections of this manuscript and for his remarks to my thesis presentation.

I would like to thank people with whom I have pleasure to work side by side on the transmission microscope of our laboratory: **Olivier Tottereau** for introducing me into TEM samples preparation and providing a comprehensive support for any event of live and **Jean-**

Michel Chauveau for fruitful discussions on the physical background of transmission electron microscopy and numerous aspects of reliable high-resolution imaging.

I thank **Luan Nguyen** and **Stéphane Vézian** for complementary characterisation of the studied samples. Thanks to them and other people (**Christian Morhain**, **Jesús Zúñiga-Pérez**, **Denis Lefebvre**, **Sylvain Sergent**, **Aimeric Courville**, **Martin Jouk** ...) for keeping me an superb company during my stay in Valbonne. Special thanks to **Boris Poulet** for his incredible efforts in teaching me French.

I am also thankful to **Isabelle Cerutti**, **Michèle Pefferkorn** and **Anne-Marie Galliana** for their administrative support and for easy solving of everyday problems. I acknowledge **Jean-Yves Duboz** for competent direction of the CRHEA laboratory and for giving me an excellent opportunity to work in it.

This thesis gave rise to collaboration with other laboratories. Many thanks to **Jorge Budagosky** from University of Valencia for his extensive strain computations. I also express my gratitude for **Vincenzo Grillo** from National Research Centre for doing quantitative analysis of HAADF images. And I thank **Mohammed Benaissa** from Centre National pour la Recherche Scientifique et Technique in Rabat for EELS analysis, which is shown in this work.

Last but not least, I thank French national project **METSA** for providing me a unique possibility to work on the microscope facilities of CEA-Grenoble and CP2M centre in Marseille. The experimental results obtained in this thesis were considerably enriched due to the employment of these state-of-art microscopes.

RÉSUMÉ

Cette thèse est dédiée à l'étude et à la caractérisation de boîtes quantiques (BQs) GaN réalisées sur une couche épaisse d' $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$. Une BQ se comporte comme un puits de potentiel qui confine les porteurs de charges dans les trois dimensions de l'espace. Le spectre d'énergie d'exciton localisé dans une BQ dépend fortement de sa taille, qui est de l'ordre de quelques nanomètres.

La compréhension des mécanismes de croissance et des propriétés physiques et structurales des objets nanométriques, telles que des BQs, nécessite l'utilisation d'outils de caractérisation adaptés à leur faible taille. La microscopie électronique en transmission (MET), largement employée dans ce travail, est une des rares techniques qui permettent ce genre d'études.

Le premier chapitre est consacré à l'introduction des fondements physiques de la microscopie électronique en transmission et à ses modes de fonctionnement. La technique de microscopie électronique en transmission haute résolution (METHR) et les facteurs limitant la résolution du microscope sont discutés en détail. Ensuite, les aspects de microscopie quantitative, tels que la simulation des images METHR et les méthodes de mesure de contrainte de la maille atomique à partir d'images METHR sont présentés. Deux techniques complémentaires au METHR, notamment l'imagerie en contraste de Z en mode balayage (STEM-HAADF) et la spectroscopie des pertes d'énergie (EELS) sont introduites à la fin de ce chapitre.

Dans le deuxième chapitre les champs d'application des nitrures d'éléments-III et certains problèmes fondamentaux liés à leur croissance épitaxiale sont discutés. Après cela, deux techniques de la croissance de ce type d'hétérostructures - l'épitaxie par jets moléculaires (EJM) et l'épitaxie en phase vapeur d'organo-métalliques (EPVOM) - sont

présentées. Ce chapitre se termine par la description des propriétés structurales des matériaux nitrures, y compris leur structure cristalline et les propriétés élastiques.

Dans le troisième chapitre, l'adaptation de l'imagerie METHR pour l'étude des matériaux à base de GaN est présentée. Tout d'abord, le moyen d'évaluation de la composition chimique dans une hétérostructure par la mesure des contraintes de la maille atomique à partir d'images METHR est décrit. L'effet de la distorsion de la maille atomique d'une couche mince due au désaccord paramétrique est pris en compte de manière analytique. Un rapport numérique entre le paramètre de réseau local et la composition des alliages $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ et $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ est déterminé. L'influence des incertitudes des constantes élastiques ainsi que l'effet de la relaxation de surface sur la précision de la détermination de la composition sont discutés.

Ensuite, une comparaison de deux techniques de mesure de contrainte, l'analyse des phases géométriques (GPA) et la méthode de projection, est présentée. Pour réaliser le traitement des images HRTEM dans l'espace réel (méthode de projection) un script dédié a été développé. Les deux méthodes ont été appliquées à des images modèles afin d'évaluer leurs performances pour la mesure des variations rapides de contrainte et pour le traitement des images bruitées. Il a été montré que la méthode GPA peut créer des fluctuations artificielles là où la contrainte varie rapidement. La méthode de projection est capable de mesurer les variations rapides de contrainte à l'échelle atomique, mais sa précision diminue considérablement pour les images bruitées.

Les effets des conditions d'imagerie sur les mesures de contrainte ont été mis en évidence dans la dernière partie de ce chapitre. Les images METHR en axe de zone et hors axe de zone ont été simulées par le logiciel Electron Microscopy Software (EMS) en utilisant les paramètres d'imagerie typique pour un microscope JEOL 2010F et un microscope Cs-correcté Titan 80-300. Le rôle critique de l'épaisseur de l'échantillon pour l'imagerie quantitative METHR en axe de zone a été montré. Les gammes de défocalisation appropriées pour une détermination fiable des déplacements atomiques pour certaines épaisseurs d'échantillon ont été déterminées. La même étude a été faite pour des images METHR acquises hors axe de zone. Les conditions d'acquisition METHR adaptées aux mesures de contrainte ont été appliquées pour l'étude des BQs GaN/AlGaIn.

Cette étude, présentée dans le quatrième chapitre, a révélé plusieurs phénomènes originaux pour les nitrures d'éléments III. Un changement de forme des BQs de surface de pyramides parfaites à pyramides tronquées avec l'augmentation de l'épaisseur nominale de GaN déposé a été observé. Le recouvrement des BQs par une couche d'AlGaN mène à une modification de leur forme de pyramide parfaite à pyramide tronquée. Dans le même temps, le volume moyen des BQs augmente. Un comportement similaire a été révélé pour des BQs GaN recouvertes par de l'AlN.

Une séparation de phase a été observée dans les barrières AlGaN recouvrant les BQs avec formation de zones riches en Al au-dessus des BQs et de régions riches en Ga placées autour des zones riches en Al. La concentration en Al dans les zones riches en Al est d'environ 70% et elle diminue avec la distance à la BQ; la concentration en Ga varie de 55% à 65%. Les facettes latérales des zones riches en Al sont bien définies, leur forme et leur taille sont similaires à celles des BQs au-dessus desquelles elles sont placées. Une séparation de phase a également été observée dans les échantillons avec des boîtes quantiques anisotropes (quantum dashes).

Les observations en microscopie électronique en transmission à balayage en mode Z-contraste (STEM-HAADF) ont fourni une preuve indépendante du phénomène de séparation de phases dans les barrières AlGaN. De plus, l'intensité des images HAADF a été convertie en compositions locales en gallium, à l'aide de simulations du signal HAADF. La concentration moyenne en Al de 70% dans les zones riches en Al obtenue par l'analyse HAADF a été confirmée par spectroscopie de pertes d'énergie électronique. En outre, l'analyse précise du changement de l'énergie de plasmon a révélé une diminution progressive de la composition en Al avec l'augmentation de la distance à la BQ.

Pour expliquer les phénomènes observés, différents modèles basés sur les données expérimentales ont été élaborés. Nos explications sont fondées sur le principe de minimisation de l'énergie totale des BQs. La variété des formes des BQs de surface est expliquée par la compétition entre l'énergie de surface et l'énergie élastique accumulée dans les BQs. L'augmentation du volume des BQ pendant leur recouvrement est expliqué par le transport de GaN de la couche de mouillage vers la BQ, qui peut aussi être relié à une baisse de l'énergie

Résumé

élastique. Une séparation de phase dans les barrières AlGaN est aussi expliquée par la relaxation de la contrainte introduite par la présence des BQs.

Plusieurs voies, basées sur les résultats de cette étude et en vue de l'amélioration des propriétés des dispositifs optoélectroniques sont proposées. Par exemple, la fabrication de BQs sans couche de mouillage est une approche perspective pour augmenter l'efficacité de photoluminescence.

Contents

ACKNOWLEDGEMENTS.....	III
RÉSUMÉ.....	V
INTRODUCTION	1
Chapter I TRANSMISSION ELECTRON MICROSCOPY	5
1.1. High-Resolution Transmission Electron Microscopy	5
1.1.1. Evolution of Microscopy: from photons to electrons.....	5
1.1.2. Electron transmission through the matter: scattering and diffraction	7
1.1.3. Principles of transmission electron microscopy; types of TEM contrast	10
1.1.4. Optical aberrations and practical resolution	14
1.1.5. Transfer function and information limit	16
1.1.6. Methods for spatial resolution improvement	21
1.2. Quantitative High-Resolution Transmission Electron Microscopy	25
1.2.1. TEM contrast simulation	25
1.2.2. Heterostructure composition evaluation by high-resolution imaging	30
1.2.3. Methods of strain measurement	31
1.3. Complementary techniques.....	36
1.3.1. High-angle annular dark-field imaging	36
1.3.2. Electron energy loss spectroscopy	38
Bibliography.....	42
Chapter II NITRIDE SEMICONDUCTORS	45
2.1. Gallium nitride specificities and domains of applications	45
2.2. Growth of III-nitrides	47
2.2.1. Molecular beam epitaxy	47
2.2.2. Metalorganic vapour phase epitaxy.....	48
2.3. Structural properties of III-nitride materials	49
2.3.1. Crystal structure	49
2.3.2. Elastic properties	52
Bibliography.....	55

**Chapter III ADAPTATION OF HIGH-RESOLUTION TRANSMISSION ELECTRON
MICROSCOPY FOR THE STUDY OF III-NITRIDE HETEROSTRUCTURES..... 57**

3.1. Relation between III-nitride alloy composition and its local lattice parameter ... 57

3.2. Evaluation of the strain measurement techniques 61

 3.2.1. Strain distribution at the interfaces 63

 3.2.2. Noisy images treatment 69

 3.2.3. Conclusion 76

3.3. Acquisition conditions for reliable high-resolution imaging..... 78

Bibliography 92

Chapter IV STUDY OF GaN/Al_xGa_{1-x}N QUANTUM DOTS MICROSTRUCTURE.. 95

4.1. Introduction..... 95

4.2. General sample characterization 98

4.3. Morphology of GaN QDs 105

4.4. Study of the AlGaN barrier microstructure by HRTEM 111

4.5. AlGaN barrier composition evaluations by STEM-HAADF and EELS 124

4.6. Study of the QD capping 130

4.7. Study of the QDash morphology 140

4.8. Discussion 144

Bibliography 160

GENERAL CONCLUSION AND OUTLOOK..... 163

5.1. General conclusion 163

5.2. Perspectives 166

Bibliography 169

APPENDICES..... 171

APPENDIX A The relationship between three- and four-index notations 171

APPENDIX B Projection of interatomic distances along different zone axis 173

APPENDIX C Aspect ratio calculation for hexagonal pyramid 175

INTRODUCTION

This thesis is dedicated to the study of GaN quantum dots (QD) grown on AlGaN templates, which are promising for the development of high-efficient electro-optical applications. The optical properties of QD-based devices are governed by the QD morphology since the energy spectrum of localized exciton strongly depends on the QD size. The study of objects as small as QDs (their typical height is between 3 and 5 nm) requires the employment of a method, which combines atomic-scale resolution with capacities of quantitative analysis. High-resolution transmission electron microscopy (HRTEM) is the tool-of-choice for so challenging objectives.

HRTEM imaging has a wide range of applications; besides GaN, it can be employed for studying the different semiconductor systems. That is why the First Chapter is dedicated to a detailed introduction of this method. First, the background of the electron interaction with matter is given. Then, the different types of electron contrast, including high-resolution contrast, are explained. The main limitations of high-resolution microscopy and the ways of spatial resolution improvement are also described. In the same Chapter, the fundamentals for the local composition determination by means of relative atomic displacement measurement are given. A brief description of two analytical TEM techniques, which were used in this thesis, is presented at the end.

The Second Chapter is consecrated to the introduction of gallium nitride. We present the actual fields of application of this semiconductor, and some fundamental problems related with its epitaxial growth. After that, Molecular Beam Epitaxy and MetalOrganic Vapour Phase Epitaxy, the two epitaxial techniques commonly used for the GaN growth, are described. This Chapter is finished by the characterization of the structural properties of nitride materials, including their crystal structure and elastic properties.

In the Third Chapter, the adaptation of the HRTEM imaging for the study of GaN-based materials is presented. First, the principle of heterostructure composition evaluation by

means of the local lattice parameter measurement is stated. The impact of epilayer lattice distortion along the growth direction is taken into account analytically. A numerical relation between the local lattice parameter and the concentration x of $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ and $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ alloy is finally determined.

The comparison of two strain measurement techniques, geometric phase analysis (GPA) and projection method, is then presented. Both methods are applied to a set of artificially created images in order to evaluate their performance in strain measurements at sharp interfaces and in treatment of noisy images.

The last part of the Chapter contains a validation of HRTEM acquisition conditions. On-axis and off-axis HRTEM images are simulated by Electron Microscopy Software with imaging parameters typical for a JEOL 2010F and a Cs-corrected Titan 80-300 microscopes, which were further used for experimental study. Strain maps calculated by geometric phase analysis from HRTEM images are analysed in order to determine the optimum image condition for strain measurements.

The study of GaN/AlGaN QD structures is presented in the Fourth Chapter. The Chapter is separated into several parts according to the investigation logic. The examination starts by the study of QD morphology change due to the QD burying by AlGaN and AlN layers. The AlGaN barriers microstructure is examined first by HRTEM and then by scanning transmission electron microscopy in high-angle annular dark-field (STEM-HAADF) mode. Then, the determination of the AlGaN barriers composition variations by various methods, including electron energy loss spectroscopy (EELS), is presented.

The study is continued by the investigations of the initial stages of the QD capping with an AlGaN layer. Cs-corrected HRTEM and probe-corrected high-resolution scanning TEM were employed for this aim. The examination of GaN/AlGaN QDs was completed by the study of anisotropic GaN islands.

Finally, the main results of the quantitative TEM study of GaN nanostructures are discussed and possible mechanisms of observed phenomena are proposed. Our explanations are based on the principle of total energy minimization and they can be used as a start point for more serious elaborations of the presented hypothesis.

Different people contributed to this study. The QD and QDash sample growth and primary characterisation were done in CRHEA laboratory by J. Brault and T. Huault. The computations of the strain tensor components for surface and buried QDs were performed by J. A. Budagosky (University of Valencia, Spain). The EELS experiments were carried out by M. Benaissa in the Max Planck Institute for Metals Research (Stuttgart, Germany). The quantitative analysis of HAADF images was done V. Grillo (National Research Centre S3, Modena, Italy). TEM study was done by the author together with P. Vennéguès using microscope facilities of CP2M-Marseille and CEA-Grenoble, which were available in the framework of METSA-project [<http://www.metsa.cnrs.fr/>].

Chapter I TRANSMISSION ELECTRON MICROSCOPY

The first Chapter is dedicated to the introduction of transmission electron microscopy (TEM) which is the main technique employed in this thesis. In doing that, we will focus on the high-resolution mode, a technique that makes possible the visualization of the atomic structure of the studied matter. The factors determining the resolving power of the TEM will be discussed as well.

The second Section addresses the extraction of quantitative data from High-Resolution TEM (HRTEM) images. We will start by introducing the methods of TEM contrast simulation, which is indispensable for a correct interpretation of HRTEM images. Then, two methods of strain analysis will be presented.

At the end, we will describe two analytical techniques, which were used to verify and complete the results obtained by the HRTEM method: high angular annular dark-field (HAADF) and electron energy loss spectroscopy (EELS).

1.1. High-Resolution Transmission Electron Microscopy

1.1.1. Evolution of Microscopy: from photons to electrons

The human eye can only distinguish adjacent structural details if they are more than 0.08 mm apart; thus we cannot see objects thinner than a hair. A microscope (from the Greek: μικρός, *mikrós* i.e. "small" and σκοπεῖν, *skopeîn* i.e. "to see") is an instrument for viewing objects that are too small to be seen by the unaided eye.

The evolution of microscopy is many centuries old. The earliest evidence of "a magnifying device, a convex lens forming a magnified image," dates back to the "Book of Optics" published in 1021. The first multilens microscopes were developed in Holland around 1590. Fast improvement and expansion of microscopes began after Galilei invented a

device with variable distances between the object-glass and eye-glass. He also made a short-focus lens that allowed scope miniaturization.

In the twentieth century, optical microscopes reached their utmost resolution. The fact is that a light ray's diffraction on the objective lens results in the image blurring. Rayleigh criterion yields a minimum spatial resolution Δl , i.e. the size of the smallest object that the lens can resolve:

$$\Delta l = 0.6 \lambda / \beta \quad (\text{eq. I-1})$$

where λ is the wavelength of light, and β is the collection semiangle of the magnifying lens. The Δl value is around 0.2 μm for microscopes operating in the visible spectrum. This means that ordinary optical microscopes are not suitable for investigation of nanostructured materials.¹

The wavelength of X-rays varies from 0,005 to 10 nm allowing very attractive spatial resolution. Nevertheless, the quality of X-ray focusing systems limits the application of this method.

In 1924 Louis de Broglie formulated the hypothesis of wave-particle duality [1], which was confirmed experimentally three years later [2]. This inspired scientists to employ electrons as microscopic probe by using magnetic or electrostatic fields as lenses. The ultimate advantage is that the wavelength of an electron with a kinetic energy of 1keV is as small as 0.038 nm and it decreases as energy increases.

When the team of Max Knoll and Ernst Ruska, initially involved in developing a cathode ray oscilloscope [3], realized this fact, they aimed at the construction of a new electron device for the direct imaging of condensed matter. The first prototype of an electron microscope had quite a low magnification; the real TEM, with a resolving power greater than that of light was already built in 1933. Five years later a resolution of 10 nm was obtained with the first practical electron microscope built by Siemens.

From that time on, the interest in electron microscopy had increased and the method developed rapidly. At present day, the atomic structure of the majority of materials can distinctly be visualized with a spatial resolution better than 0.1 nm.

1.1.2. Electron transmission through the matter: scattering and diffraction

The electron has to be described quantum mechanically by a wave function which is a solution of the Schrödinger equation or Dirac equation. Under the typical illumination conditions of an electron microscope the incoming electron can be treated as a plane wave

$$\Psi = A_0 \exp [2\pi i(\mathbf{k}_0 \cdot \mathbf{r} + \varphi_0)] \quad (\text{eq. I-2})$$

where A_0 is the amplitude, \mathbf{k}_0 is the wave vector and φ_0 is the phase of the wave. The wave function of the electron propagating through the sample has to take into account the interaction with the matter.

The ability of atoms to scatter electrons was experimentally discovered by Thomson at the beginning of the twentieth century [2]. If the re-radiated electron has the same magnitude of the wave vector (i.e. the same energy) as the incident one, then the interaction is called elastic; all other cases are called inelastic. The probability that an electron will be elastically scattered by an isolated atom is given by the scattering cross section:

$$s = p \left(\frac{Ze}{Vq} \right)^2 \quad (\text{eq. I-3})$$

where Z is the atomic number, e is the elementary charge, V is the potential of the incoming electron and θ is the angle of scattering. As a result of the scattering, the intensity of the transmitted beam decreases, but new waves propagating in different directions arise.

In order to describe the elastic scattering of an electron in condensed matter the different scattering events have to be summed up coherently in crystals or non coherently in amorphous materials. However, an exact analytical description of the elastic scattering cross section is not available, since the interaction potential has to take into account as well the nucleus of the target atom as well as its electrons with very complex screening effects.

¹ However some specialized techniques such as Near Field Scanning Optical Microscopy or Stimulated Emission Depletion Microscopy are capable to overcome the resolution limit imposed by diffraction.

Generally one can say that the cross section depends on the atomic number Z of the target atom, i.e. the heavier the atom the more probable is a scattering event. In condensed matter, the overall probability that an electron will be deflected from its initial direction is proportional to the length of his path into the substance. Thus, the thicker areas of the specimen scatter more electrons than the thinner ones.

The angular distribution of scattering events is highly inhomogeneous due to the matter periodicity, an effect which is called diffraction. The periodic spatial arrangement of atoms dictates that the outgoing waves leave the crystals only in certain directions. The fact is that the waves elastically diffracted by a set of parallel planes interfere with each other (Figure I-1). And constructive interference occurs only if the phase shift is a multiple of 2π . The diffraction angles θ which are satisfying this criterion of constructive interference are determined by the Bragg's law:

$$2d \sin\theta = n \cdot \lambda \quad (\text{eq. I-4})$$

where θ is the angle between the wave vector of incident rays and the scattering planes, d is the spacing between those planes, λ is the wavelength of electrons, and n is an integer.

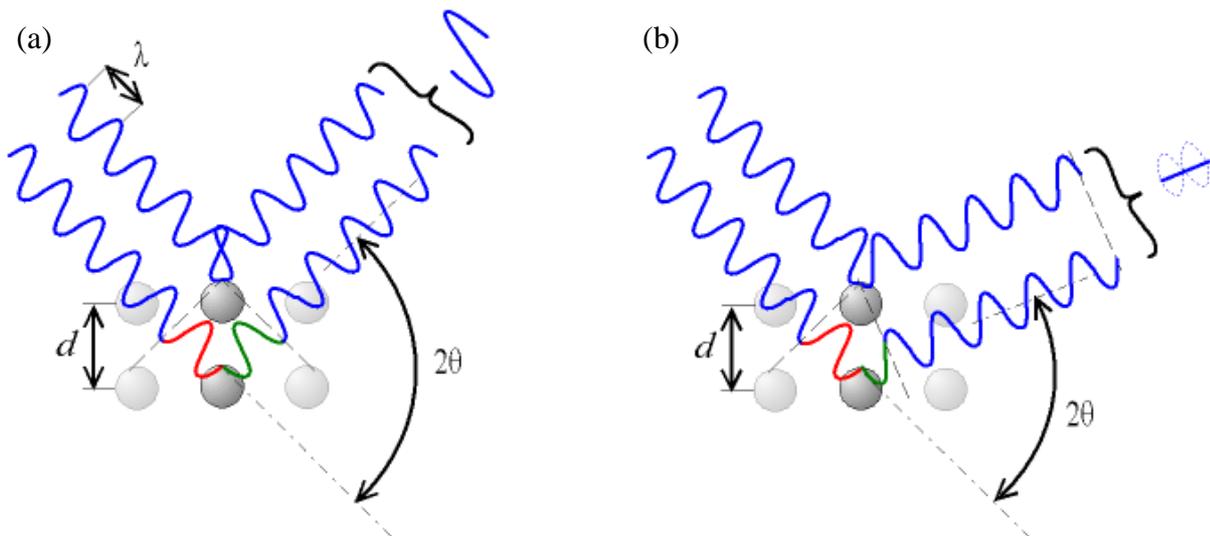


Figure I-1: Illustration of constructive (a) and destructive (b) interference: two incident waves with wavelength λ are diffracted at an angle 2θ . The red and green parts of the second wave show the path-length difference.

The same condition can be stated in terms of the wave and reciprocal lattice vectors (Laue equation):

$$\mathbf{k}_0 - \mathbf{k}_D = \mathbf{g} \quad (\text{eq. I-5})$$

where \mathbf{g} is a reciprocal lattice vector and \mathbf{k}_0 and \mathbf{k}_D are the wave vectors of the incident and diffracted beams. Those vectors have the same modulus equal to $2\pi/\lambda$, which is more than hundred times greater than the modulus of any primitive vector of the reciprocal lattice, typically few nm^{-1} long. It signifies that for an arbitrary orientation of \mathbf{k}_0 relative to the crystal lattice, several diffracted beam vectors \mathbf{k}_D satisfying Laue equation can be chosen. Hereby the incident wave always produces a set of diffracted beams.

At last, when an electron passes through the matter its phase is shifted in proportion to the specimen's thickness and the inner potential relative to the phase of the electron passed through the vacuum (Figure I-2). Thereby the interaction of electrons with crystal is quite productive. Now let us see how the electron microscope reveals the structural information.

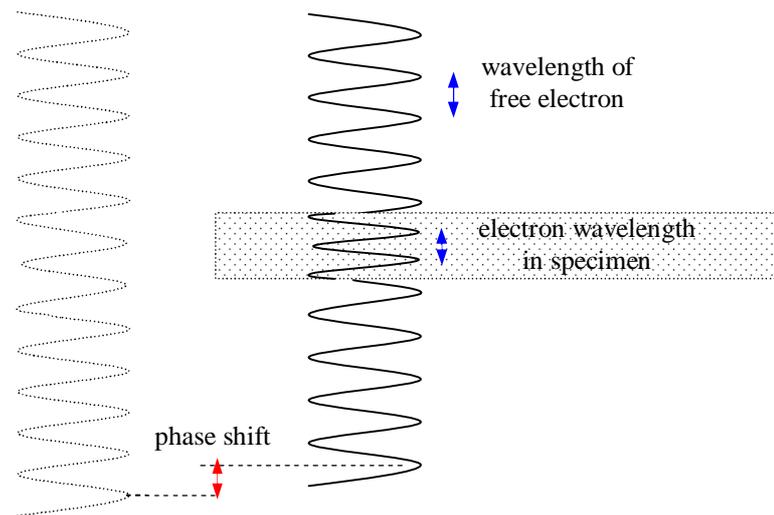


Figure I-2: The dotted line symbolizes electron propagation in vacuum. The solid line is the same for an electron travelling in the specimen. The effect of the wavelength decrease is exaggerated for better visualization.

1.1.3. Principles of transmission electron microscopy; types of TEM contrast

The incident electron wave should be highly coherent, so that the variations induced by the sample can be detected. Two fundamentally different types of electron sources are used in transmission microscopes. Thermionic emission, i.e. the electrons ejection over the potential barrier (called “work function”) is possible by heating the metals to high temperatures and application of a positive voltage. An alternative method of electron generation is field emission, the process whereby electrons are withdrawn from a sharp tip using a high electric field. In the latter case the electron beam has a higher spatial coherency and a shorter energy spread (0.3 eV as compared to 1.5 eV in the case of thermionic emission), which makes field emission guns (FEG) more favourable for high-resolution imaging, as we will see afterwards.

Then extracted electrons are accelerated to relativistic velocities by a voltage of several hundred kV applied to the anode (Figure I-3). The magnetic coils of the illuminating system perform shift, tilt and convergence of the electron beam. The specimen itself may be manipulated in three directions and tilted around two orthogonal axes.

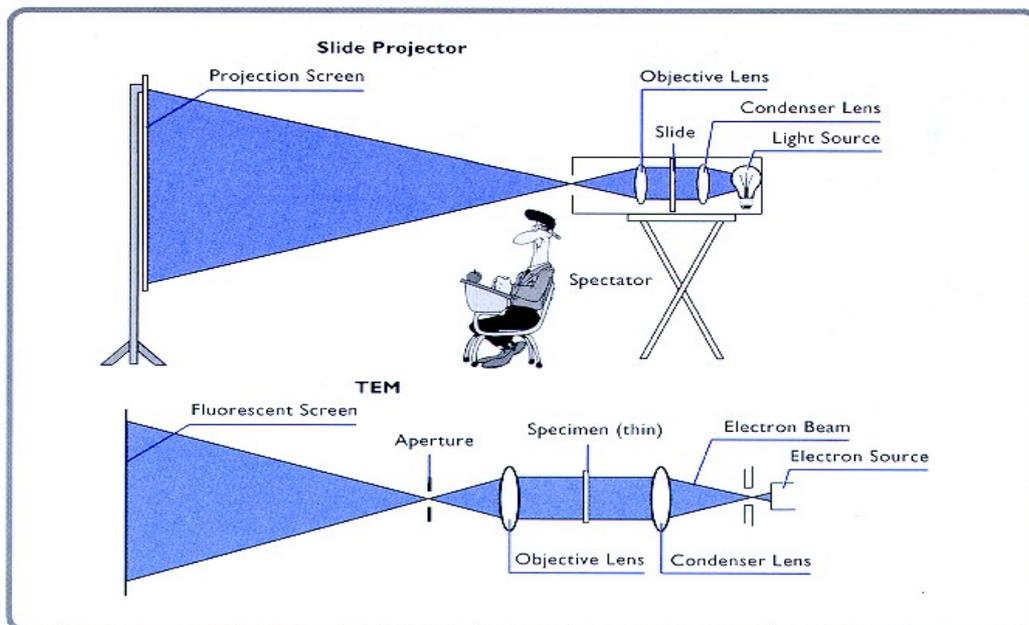


Figure I-3: Image formation in TEM, at a first approximation, is quite similar to slide projection [4]. Both procedures require a source producing photons/electrons. The condenser lens creates parallel illumination of the transparent slide or specimen. The objective lens converges the transmitted wave. The aperture selects central part of the beam which is then projected on the viewing screen.

The objective lens focuses all beams diffracted by the sample, thus forming a diffraction pattern in the back focal plane. An objective aperture selects the beams which will contribute to the TEM image formation. Magnetic lenses of the projector system just magnify the final image, which is then viewed on a fluorescent screen or a CCD camera. The final image intensity is proportional to the square of the electron wave amplitude; whereas the wave phase is lost during the acquisition. So, what is the origin of the contrast that we see on the TEM screen?

A monocrystalline sample, aligned along one of the low index zone axis produces a diffraction pattern with a great number of reflections. Let us begin with the simplest case, considering that the objective aperture cuts off all diffracted beams.

The variations in thickness or composition along the field of view will change the transmitted beam intensity: since a thicker or heavier specimen area diffracts more electrons from their initial trajectory, the beam transmitted through such areas is less intense (Figure I-4). Due to this, the final TEM image represents the composition variations superposed on the contrast caused by the sample thickness fluctuations.

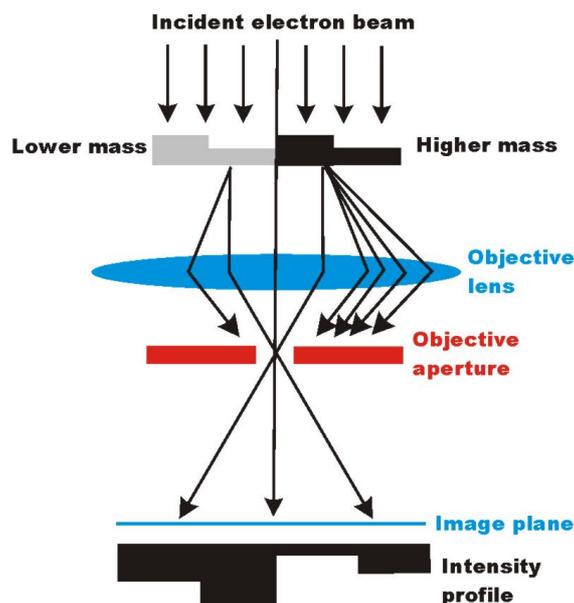


Figure I-4: The mass-thickness contrast formation: lighter or thinner zone of a monocrystalline specimen diffract less electrons than thicker or heavier one. Objective lens focuses all divergent waves, but only transmitted beam pass through the objected aperture. In the image plane, it forms an intensity profile representing mass and thickness variation in the field of view.

If the crystallographic orientation of the specimen varies, like in polycrystalline samples, zones satisfying to the Bragg's condition will diffract more electrons than others. This will change the intensity of the transmitted beam, revealing thereby the formation of diffraction contrast. Particularly, the observation of dislocations in bulk material is due to this mechanism.

Let us now assume that one diffracted beam passes through the objective aperture in addition to the transmitted beam (Figure I-5). The transmitted beam \mathbf{k}_T has the same wave vector as the incident one \mathbf{k}_0 .

$$\Psi_T = A_T \exp [2\pi i(\mathbf{k}_0 \cdot \mathbf{r} + \phi_T)] \quad (\text{eq. I-6})$$

The diffracted wave Ψ_D is described similarly, but the wave vector \mathbf{k}_D is related to \mathbf{k}_0 by equation I-5. The wave vectors of both beams have the same magnitude (i.e. the wavelength) but different directions. Thus, when converged by the objective lens, those waves will interfere. The interference image intensity equals

$$I = \Psi_\Sigma \cdot \Psi_\Sigma^*, \quad \Psi_\Sigma = \Psi_T + \Psi_D \quad (\text{eq. I-7})$$

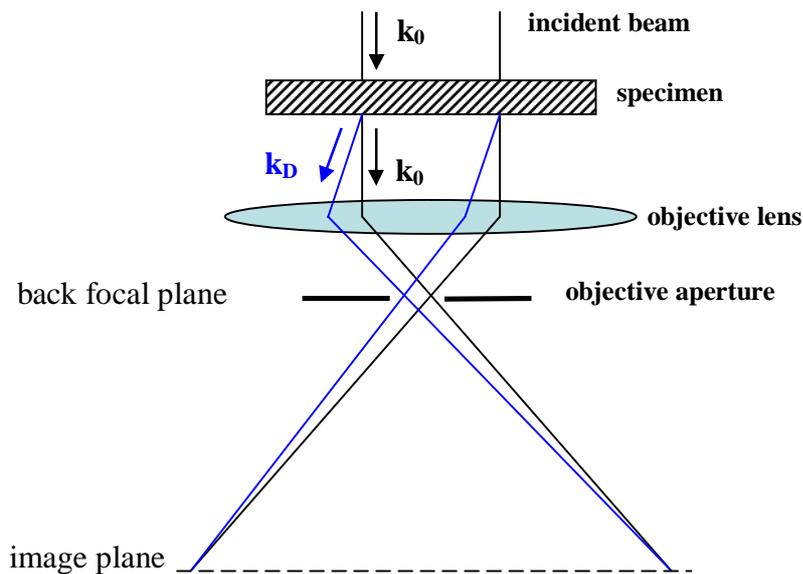


Figure I-5: Incident electron wave with wave vector \mathbf{k}_0 is diffracted in the specimen. Both transmitted \mathbf{k}_0 and diffracted \mathbf{k}_D beams are converged by the objective lens. The waves pass through the objective aperture and interfere in the image plane.

The final image intensity I contains item $\Psi_T^* \cdot \Psi_D + \Psi_T \cdot \Psi_D^*$, which is proportional to

$$\frac{\exp [2\pi i((\mathbf{k}_0 - \mathbf{k}_D) \cdot \mathbf{r} + \varphi_T - \varphi_D)] + \exp [-2\pi i((\mathbf{k}_0 - \mathbf{k}_D) \cdot \mathbf{r} + \varphi_T - \varphi_D)]}{2\cos (2\pi (\mathbf{g} \cdot \mathbf{r} + \varphi_T - \varphi_D))} \quad (\text{eq. I-8})$$

The item $\Psi_T^* \cdot \Psi_D + \Psi_T \cdot \Psi_D^*$ is responsible for the appearance of a sinusoidal oscillation in the interference image. The fringe periodicity equals $1/g$, i.e. it represents the spacing between the planes, which diffracted the beam \mathbf{k}_D . Therefore interplanar distances in the studied material can be visualised by means of the high-resolution contrast.

When a few (or all) diffracted beams interfere in the image plane, they reveal the periodicity of planes that reflected these beams. However, the overall image intensity is more complicated than just the superposition of fringes formed by each diffracted beam. The problem is that the objective lens itself affects the amplitudes and phase of the diffracted waves; and this effect strongly varies with microscope defocus. Besides, the final image is also modulated by composition variations and thickness fluctuations, if they are present in the field of view.

Therefore, no matter how fine the high-resolution image looks, one should beware of direct interpretations. Only knowledge of exact microscope settings and simulation of HRTEM images may eliminate this uncertainty. We will return to this question when we will talk about the adaptation of HRTEM for studying GaN-based heterostructures (Section 3.3). Now let us consider the TEM resolution power and discuss its main limiting factors.

1.1.4. Optical aberrations and practical resolution

Indeed, since the relativistic wavelength of the electron accelerated by 200kV potential is as small as 2.5 pm, one can expect that the electron microscope will have a spatial resolution of the same order of magnitude. In practice, until recently, TEMs' resolution was saturated around 200 pm.

The reason is, as we mentioned above, that the diffracted electron waves, which carry the structural information, are noticeably modified by the microscope itself. Magnetic lenses distort electron beams in the same way as common magnifying glasses deform the ray of light. The non-uniformities of the magnetic field inside the lens are described by aberration coefficients, similarly to light optics.

The imperfections of the objective lens have the most important impact on the spatial resolution, since that lens focuses diffracted beams and forms the primary image. Here we will briefly review the main aberrations that are of importance for HRTEM imaging.

- Azimuth nonuniformity of the magnetic field inside the objective lens causes astigmatism effect. The rays, propagating in non-parallel planes, will have foci at different points (Figure I-6). This effect introduces distortion of amount $\beta\Delta f$, where β is collection semiangle and Δf is maximum difference in defocus induced by the astigmatism. Real lenses are described by different orders of astigmatism, accordingly to the symmetry of the distortion. Fortunately, the first order astigmatism can be properly balanced by tuning the compensating field inside the stigmator.

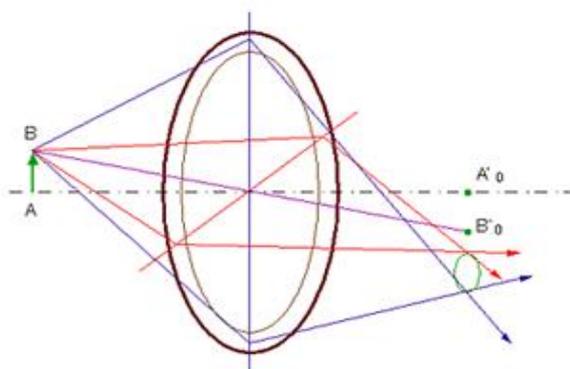


Figure I-6: Astigmatism effect: rays leaving the object B in the horizontal plane (red lines) are focused far from the rays lying in the vertical plane (blue lines) because of astigmatism. Dot-and-dash show optical axis of the lens.

- Radial nonuniformity of the objective lens field produces spherical aberration: parallel beams, which pass at different distances from the optic axis, fail to converge to the same focus (Figure I-7 (a)). Different orders of spherical aberration are also defined, but generally microscope performance is governed by the third-order aberration. As a result, a point source is imaged as a disk of minimum radius $r_s = C_3\beta^3$, where C_3 is the third-order spherical aberration coefficient. Consequently, all nonperiodic details of the object are smoothed at the same amount; thereby the microscope practical resolution decreases.
- At last, the electrons with various wavelengths are deflected by magnetic lenses at different angles (see Figure I-7 (b)). The effect, called chromatic aberration, is always present since high-tension supplies have an intrinsic energy spread. The variation of the electrons' wavelength blurs the image by an amount $r_c = C_c\beta\Delta E/E_0$, where ΔE – is the energy spread of electrons, E_0 – the initial beam energy, and C_c is the chromatic aberration coefficient.

In this way, the objective lens magnifies the image, but confuses the finest details at the same time. This problem was formulated in Scherzer theorem saying that, neither chromatic, nor spherical aberration could be corrected by the rotationally symmetric elements [5]. As a result, the resolution of conventional electron microscopes is limited to about 100 times the wavelength of the electron.

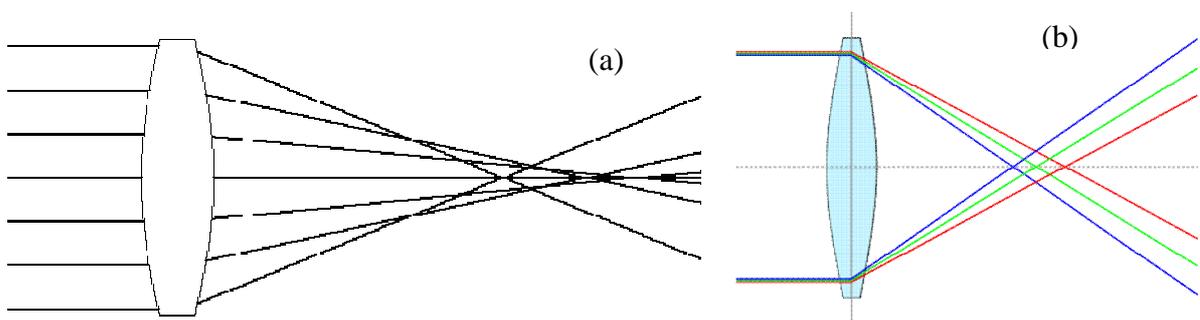


Figure I-7: Due to the spherical aberration (a) initially collimated rays are not brought to a single focus, but have different focal distances. Chromatic aberration effect (b): electrons with smaller energy (blue line) has shorter focal length, than the higher energy ones (red line)

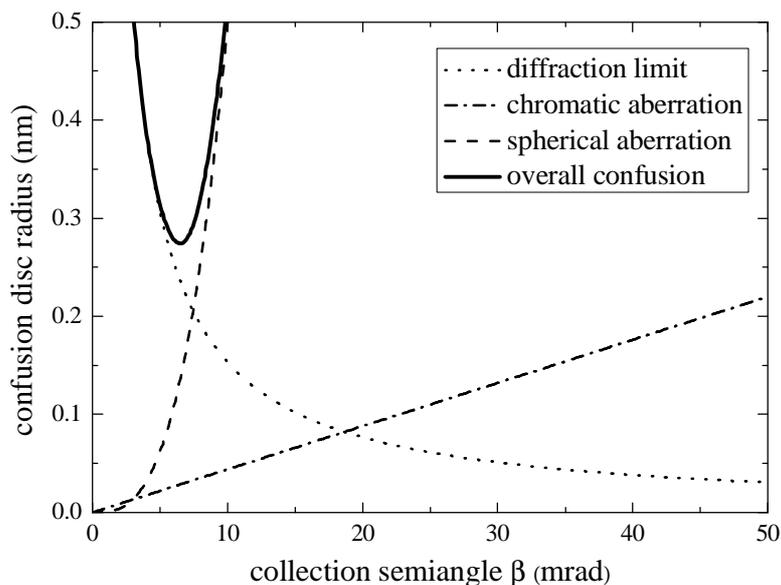


Figure I-8: Confusion disc radius as a function of the collection semiangle. The impacts of diffraction limit, spherical and chromatic aberration were plotted for JEOL 2010F with $E = 200\text{kV}$, energy spread = 0.8 eV , $C_s = 0.5\text{ mm}$, $C_c = 1.1\text{ mm}$.

This effect is illustrated on the example of the JEOL 2010F microscope. The diffraction limit, due to the Rayleigh criterion (eq. 1.1) equals $r_D = 0.6\lambda/\beta$. The total radius of the confusion disk $R = (r_s^2 + r_c^2 + r_D^2)^{1/2}$ is entirely governed by the spherical aberration (Figure I-8). And the practical resolution of the microscope is equal to the minimal radius of confusion disc $R_{\min} \approx 0.9 (C_s \lambda^3)^{1/4}$, achieved at $\beta_{\text{opt}} \approx 0.8 C_s^{-1/4} \lambda^{1/4}$.

However, sometimes images of the crystal lattice contain some periodical features separated from one another by distances smaller than R_{\min} . Are such small details distinguishable in HRTEM image still reliable?

1.1.5. Transfer function and information limit

We know that a periodic atomic lattice gives rise to strong diffraction. The diffracted beams propagate through the objective lens on slightly different paths than the transmitted one. So, once again, the role of spherical aberration is of great importance, since the electrons scattered at the highest angles (and hence the ones which pass the objective lens with the highest

distance from the optical axis) transmit the smallest periodicity in the object, which defines the microscope resolution. But the outermost beams are deflected more strongly than the ones close to the optical axis, which indicates that these beams are not properly focused in the image plane. Thereby transmission of highest spatial frequencies is hindered by the objective lens aberrations.

The transfer function formalism is introduced to describe the influence of the microscope optics on the the electron wave exiting from the sample? To describe how the specimen modifies the incident uniform electron wave we should make certain simplifications. First, we presume that diffracted beam intensities are much smaller than the direct one. Thus, we focus only on the phase changes during the wave transmission through the sample; moreover we assume that those phase changes are small. This approach is known as a weak phase-object approximation (WPOA) and it only holds true for thin specimens [6]. Using WPOA the specimen function can be expressed as

$$f(x,y) = 1 - i\sigma V_t(x,y), \quad \text{with } \sigma = \pi / \lambda E \quad (\text{eq. I-9})$$

where σ is the scattering constant, $V_t(x,y)$ is the projected potential of crystal, E and λ are the wavelength and energy of the electron wave. The image wave is given by the convolution of $f(x,y)$ with the inverse Fourier transform of the microscope transfer function. In the WPOA only the imaginary part of the transfer function $T(q)$ contributes to the image intensity and it can be written as:

$$T(q) = A(q) E(q) \sin\chi(q) \quad (\text{eq. I-10})$$

where q is a spatial frequency, $A(q)$ is the aperture function, $E(q)$ is an envelope function and $\chi(q)$ is the phase-distortion function. The aperture function shows that the objective aperture cuts off all rays diffracted above a certain angle, thus directly confining the maximum spatial frequency transmitted to the image. The phase-distortion function represents the phase difference introduced by the optical aberrations; whereas the envelope function describes their damping effect.

Let us first discuss the phase-distortion function. Every geometrical aberration contributes to the $\chi(q)$ [7], but we will consider only the main impacts:

$$\chi(q) = \pi C_1 \lambda q^2 + \frac{1}{2} \pi C_3 \lambda^3 q^4 \quad (\text{eq. I-11})$$

where C_1 is the defocus value of the objective lens, and C_3 is the third-order spherical aberration coefficient.

The behaviour of the phase-distortion for the JEOL 2010F is shown on Figure I-9 (a) (the envelope functions we will discuss bit later). The transfer function $T(q)$ oscillates within the envelope function it reverses its sign every time when $\chi(q)$ reaches a value divisible by π . Besides, certain spatial frequencies q_0 , which correspond to the zeros of $T(q)$ are not transmitted to the image.

Moreover, the transfer function shape changes with the defocus value (compare Figure I-9 (a) and (b)). Fortunately, the operating band of $T(q)$ can be maximized by setting the appropriate defocus value. The defocus $C_1^{\text{Sch}} = -(4/3C_3\lambda)^{1/2}$ proposed by Scherzer, creates the widest band for low spatial frequencies transmission [8]. The idea is that a negative value of defocus compensates, for a certain range of q , the phase distortion due to the spherical aberration (eq. I-11).

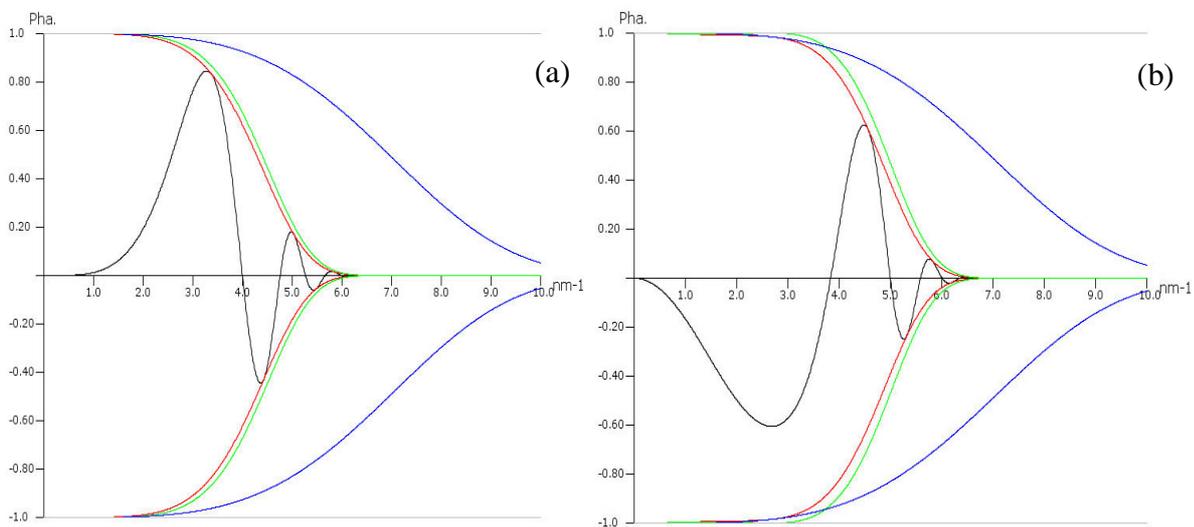


Figure I-9 : Imaginary part of the transfer function for JEOL 2010F for 2 values of defocus. Phase-distortion function causes oscillation, whereas the envelope functions (shown by colour lines) have a dumping effect. Note that the $T(q)$ shape is drastically dependent on the defocus (0 nm (a) and -23 nm (b)).

Indeed, by setting the Scherzer defocus, the first zero of $T(q)$ is pushed towards the maximum value of $q_0 = 1.5 C_s^{-1/4} \lambda^{-3/4}$ (Figure I-10). The point-to-point resolution for every microscope is always given at Scherzer defocus.

The phase-distortion function $\chi(q)$ complicates the image contrast, whereas the envelope function $E(q)$ limits the transfer of high spatial frequencies to the image. $E(q)$ can be represented as a product of single envelopes:

$$E(q) = E_s(q) E_c(q) E_d(q) E_v(q) E_D(q) \quad (\text{eq. I-12})$$

where $E_s(q)$ is the spatial envelope, describing the spatial coherency of the beam, $E_c(q)$ is the temporal envelope describing the temporal beam coherency, $E_d(q)$ is the specimen drift, $E_v(q)$ is the specimen vibration and $E_D(q)$ describes the detector performance. Specimen drift and vibration can be minimized by maintaining a stability of the working environment. The detector envelope function is due to the delocalisation of fast electrons on the CCD camera scintillator and to the limited size of CCD pixels (Nyquist criterion). The modulation transfer function can be calibrated for a given CCD camera and taken into account for the fine analysis of HRTEM images.

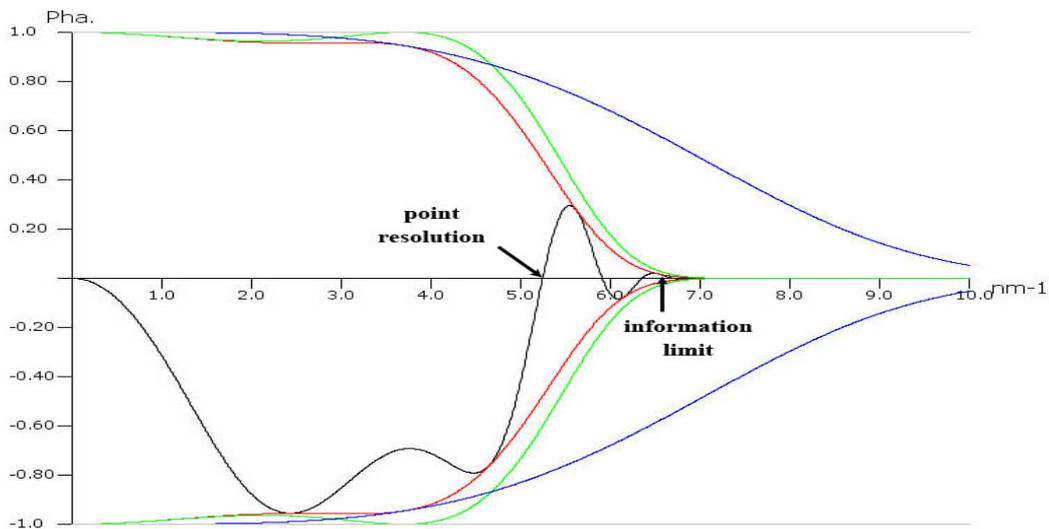


Figure I-10 : Imaginary part of the transfer function for a JEOL 2010F as a function of the spatial frequency at Scherzer defocus. The point-to-point resolution corresponds to the first zero of the transfer function. The information limit (see paragraphs below) is determined by the damping effect of the envelope function.

Spatial coherency of the beam is mainly due to the angular spread of the electron source. If we consider that the electron probe has a Gaussian distribution of intensity, the spatial envelope will be

$$E_s(q) = \exp [-(\pi\alpha/\lambda)^2 (\partial\chi(q)/\partial q)^2] \quad (\text{eq. I-13})$$

where α is the semiangle describing the Gaussian distribution of electron intensity.

The information throughput of a microscope can be maximized by setting it at Lichte defocus $\Delta f_{\text{Lichte}} = -0.75C_s(u_{\text{max}}\lambda)^2$, where u_{max} is the maximum transmitted spatial frequency [9]. At Lichte defocus the derivative of $\chi(u)$ is minimized, which decrease the damping effect of the spatial envelope. On the other hand, the use of high-coherence FEG extends the $E_s(q)$ envelope, thus making transmission of high spatial frequencies possible at various defocuses.

The temporal envelope also damps high spatial frequency due to chromatic aberration:

$$E_c(q) = \exp [-1/2 (\pi\lambda\delta)^2 q^4] \quad (\text{eq. I-14})$$

where δ is the defocus spread, which is given by the equation

$$\delta = C_c \sqrt{4\left(\frac{\Delta I}{I}\right)^2 + \left(\frac{\Delta E}{E}\right)^2 + \left(\frac{\Delta U}{U}\right)^2} \quad (\text{eq. I-15})$$

here C_c is the chromatic aberration coefficient, $\Delta I/I$ is the instability of the objective lens current, $\Delta E/E$ is accelerating tension fluctuation, and $\Delta U/U$ is the natural energy spread of the high-voltage supply. Modern transmission electron microscopes equipped with FEG have a defocus spread in the order of a few nanometres.

The spatial and temporal envelopes are shown on Figure I-11 as well as on Figures I-9 and I-10. The chromatic aberration envelope $E_c(q)$ has, de facto, the same influence than the aperture function $A(q)$, since it imposes a maximum to the transmitted spatial frequency. This maximum determines the highest resolution attainable with a microscope and is called the information limit. If we neglected any other contributions to the envelope function, than the information limit due to the chromatic aberration is

$$\rho_c = \sqrt{\frac{\pi\lambda\delta}{2}} \quad (\text{eq. I-16})$$

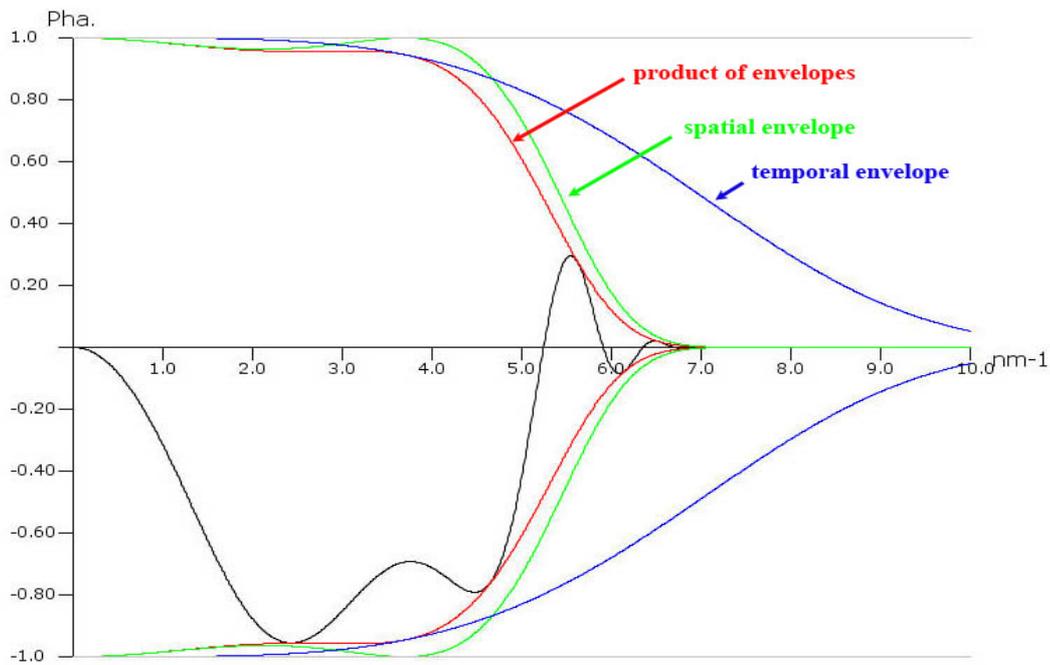


Figure I-11: Imaginary part of the transfer function for JEOL 2010F at Scherzer defocus.

This equation has a fundamental value, determining the highest possible resolution achievable on a given machine (see Figure I-10). Anyhow, the extraction and correct interpretation of information lying beyond the point resolution requires serious efforts. We will return to this subject at the beginning of the next Section, when we will talk about the simulation of high-resolution images.

1.1.6. Methods for spatial resolution improvement

We finish the introduction to the HRTEM by discussing how to improve the resolution of an electron microscope. We already saw that the spherical aberration has the most important impact on the resolving power. The minimization of spherical aberration was of great interest since the TEM development. By improving the design of the objective lenses the value of the spherical aberration coefficient could be decreased down to one half of a millimetre. But as shown by Scherzer the decrease of the magnetic field in the centre of rotationally symmetric electron lenses will always produce certain spherical aberration.

Furthermore, two fundamental approaches exist to overcome the aberration limit. The pioneering idea was to increase the accelerating voltage of the electron gun up to a few MeV, decreasing thereby the wavelength of electrons below 1pm. In that way the confusion disc radius, the phase-distortion function, the point-to-point resolution and the information limit are considerably reduced. The problem was that the optical aberration increases with the accelerating voltage, diminishing thereby the positive effect of electron wavelength reduction. However, an electron beam with an energy of 1 MeV considerably damages the specimen. Nevertheless high tension TEMs are more efficient than common microscopes; and quite a number of such machines were built. But an enormous fabrication cost makes this approach economically unattractive; only few MeV-voltage TEMs are operating nowadays [10].

The alternative way consists in the correction of optical aberrations. In light optics the spherical aberration can be minimized using combination of convex and concave lenses, having opposite values of C_s . Unfortunately all magnetic lenses are convex, thus the spherical aberration coefficient is always positive. Scherzer theorem claims that the correction of C_s is possible only by the introduction of a special correcting element.

Various prototypes were suggested, but only the concept proposed by Rose [11] turned to be realizable. The principle is to introduce multi-pole elements into the microscope column. Properly tuned internal magnetic field of the multi-pole elements will strongly deflect electron beam that passes at higher distance from the optic axis. Additional electron lenses are introduced to compensate the low-order aberrations produced by multi-poles elements.

Finally overall microscope spherical aberration (objective lens + corrector) can be tuned to zero or even to a negative value with an accuracy of a few μm . In fact, all optical aberrations up to third-order can be balanced with a high accuracy.

The realization of a C_s -corrector in 1995 broke fresh ground in the field of electron microscopy. If we look again to the problem of image blurring due to the optical aberrations described in Section 1.1.5, then we will see that now the practical resolution is twice improved (Figure I-12). The minimum confusion disc radius is governed only by the chromatic aberration $R_{\min} \approx 1.1(\lambda C_C \Delta E / E_0)^{1/2}$, where ΔE is the natural energy spread of electron source and E_0 is the beam energy. The range of optimum β is larger than in the case of a conventional microscope, with the best value $\beta_{\text{opt}} \approx 0.8(\lambda E_0 / \Delta E / C_C)^{1/2}$.

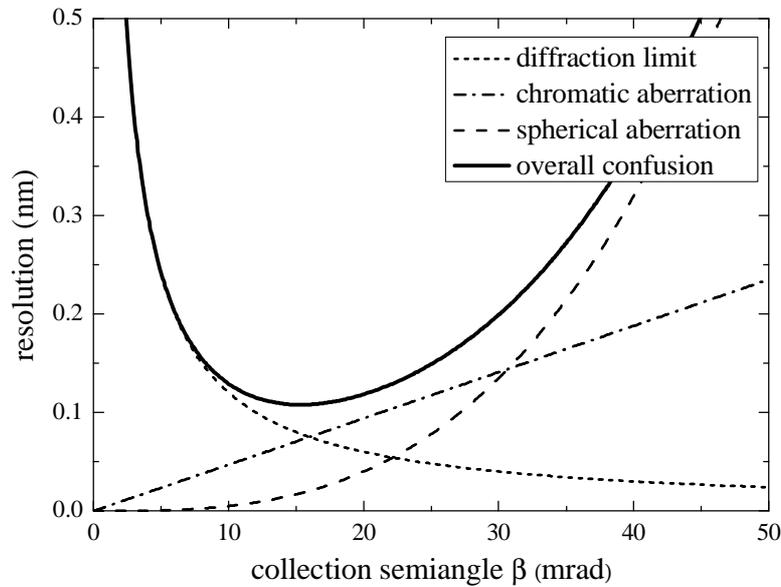


Figure I-12: Confusion disc radius as a function of collection semiangle for a Cs-corrected microscope. The impacts of diffraction limit, spherical and chromatic aberration were plotted for a Titan 80-300 with $E = 300\text{kV}$, energy spread = 0.7 eV, $C_3 = -0.005$ mm, $C_5 = 4$ mm, $C_c = 2.0$ mm. Note the difference comparing to Figure I-8.

The transfer function of the Cs-corrected microscope is smooth and simple (Figure I-13). The spatial envelope stretches over the high-frequency part due to the phase-distortion function correction. The information limit is also significantly improved, since now it is governed only by the chromatic aberration. The foremost advantage is that the point resolution approaches the information limit. In fact, “resolution” term is no more in use, since now the microscope performance is entirely determined by the information limit which mainly due to on the electron gun energy spread, chromatic aberration coefficient and microscope stability. Modern Cs-corrected microscopes, equipped with a FEG, have information limit around 0.08 nm^2 .

The further improvement of microscope performance might be achieved by the compensation of the chromatic aberration. The Cs- and Cc-correctors have the a similar principle, however in the second case the electron trajectory rectification is obtained by tuneable electric fields.

² For the time being, the ultimate information limit, as small as 50 pm, belongs to the TEAM 0.5. This 300kV microscope houses high-brightness gun with low energy spread and improved Cs corrector able to compensate fifth-order of spherical aberration. Superior machine with spherical and chromatic aberration corrector will be elaborated soon within the framework of this research project.

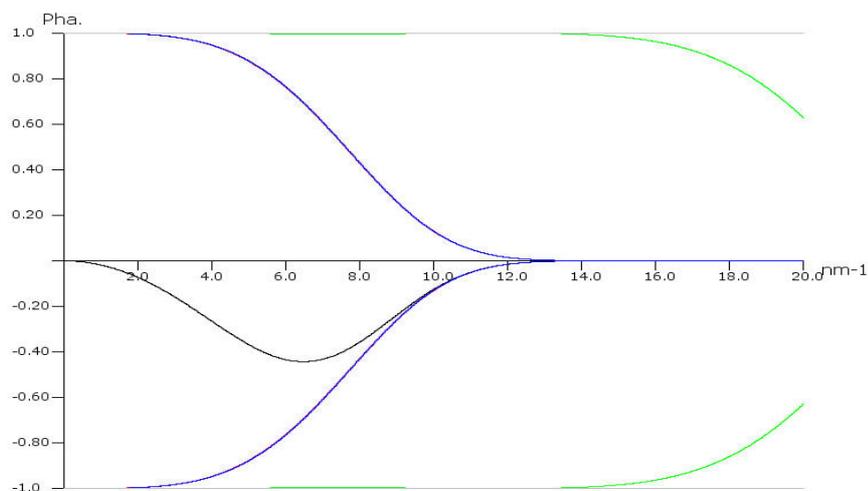


Figure I-13: Imaginary part of the transfer function for the Cs-corrected Titan 80-300 as a function of the spatial frequency. The C_3 coefficient is $-4 \mu\text{m}$, C_c is 2 mm and the energy spread 0.7 eV . Comparing to the JEOL 2010F (Figure I-9) the microscope passband is extended towards spatial frequencies allowing 80 pm resolution; note that the abscissa scale is twice greater to display high-frequency part of $T(q)$.

An alternative approach is to use a monochromator, which decreases the energy spread of the outgoing electron beam. Its mode of functioning is rather simple: transiting electrons are deflected by a magnetic field and only those having required energy pass through the exit slit. The outgoing beam is less intense, but its' energy spread is considerably reduced.

Nevertheless, machines with large spatial and temporal envelopes (considering equation I-12), are extremely sensitive to any extrinsic instabilities. Whatever the case, microscope environment should be as stable as possible. However, to achieve sub-angstrom resolution it is essential to entirely isolate sample stage from acoustic noises, ground vibrations and temperature changes. Otherwise specimen drift and vibrations will make atomic resolution unachievable. Besides, electrical supplies should provide constant voltages to electron gun and objective lens, in order to minimize the defocus spread, reducing thereby the information limit.

So, transmission electron microscopy progressed a lot since its invention in 1933. Anyhow the method is still far from being perfect and it will be certainly improved in the future. But even at the present, a non-ideal state TEM is an exclusive research tool for the characterization of materials crystal structure and obtaining unique information on the sample microstructure. In the following Section we will discuss the essential aspects of the quantitative analysis based on HRTEM.

1.2. Quantitative High-Resolution Transmission Electron Microscopy

In the first part of this Chapter we discussed the factors governing the resolving power of TEM. Here we will talk about imaging of the specimen structure, image treatment and quantitative data extraction. Though the resolution of modern microscopes (0.08 – 0.1 nm) is few times lower than the interatomic distances of most solids (0.2 – 0.5 nm), the HRTEM images do not always represent the exact atomic structure. The factors limiting the utilization of HR microscopy and the need of image simulation are discussed first. Then we will talk about heterostructure composition determination by means of local lattice parameter. After that the techniques used for the displacement measurement from the HRTEM images will be discussed.

1.2.1. TEM contrast simulation

The ultimate aim of HRTEM imaging is the local atomic structure determination. However, the wave exiting from the sample is strongly distorted by the optical aberrations. The HRTEM image thereby depends on the shape of the transfer function, which noticeably varies with the defocus (Figure I-9).

But even the wave exiting from the object does not correspond exactly to the projected potential of the studied structure. The point is that the electron wave may diffract several times even in a thin crystal. Indeed, a diffracted beam in perfect Bragg orientation is to be diffracted again by the same set of planes (Figure I-1). Consequently the exit wave function greatly depends on the crystal thickness.

As a result, finding the true atomic positions from a HRTEM image is a very intricate objective. However, it is much easier to solve an inverse problem: to calculate a HRTEM image by the convolution of the microscope transfer function with the exit wave from an assumed atomic structure. The image simulation is realized in the following way:

- At first the assumed atomic structure is modelled.
- Next the interaction of the electron wave with the inner potential of the crystal is calculated.

- Then the exit wave is distorted during the propagation towards the image plane.
- Finally the image wave interferes following the Abbe's theory and forms thereby the HRTEM image.

If the simulated HRTEM image differs noticeably from the experimental one, then, either the assumed atomic structure or the microscope characteristics should be modified. If satisfactory resemblance is achieved, than one may conclude that the atomic arrangement in the studied area is known. Anyhow it should be kept in mind that sometimes two different structures can produce similar images.

The effects of optical aberrations are taken into account by means of the microscope transfer function. But the exit wave can be calculated by different methods. The kinematical theory is the simplest approach to compute the amplitude of the scattering beam. It assumes that an electron beam passing through the sample will be diffracted only once and without losing energy. Thus, inelastic scattering and multiple diffraction are excluded. Such approach simplifies noticeably the calculations, but it fails completely for samples with thicknesses above 5 nm. The probability of multiple diffraction increases with the thickness, which necessitates the use of dynamical calculations.

The full description of any physical system evolution is given by the time dependent Schrödinger equation. The Bloch's theorem proves that the wavefunction of a particle placed into a periodic potential can be described as the product of a plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ and of a periodic function $u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}+\mathbf{R})$, where vector \mathbf{R} is the period of the potential:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}). \quad (\text{eq. I-17})$$

By the substitution of this definition into the Schrödinger equation, the latter is reformulated into the nonlinear matrix problem of finding the electron wave vectors and the Bloch wave coefficients [12].

The total calculation time of the Bloch wave method is proportional to the cube of the number of used Bloch waves. The calculation might be notably simplified, by applying the Bethe's approach. His idea is to take into account only strong reflections, by introducing the

weak diffracted beams as small perturbations to the strong ones. A simple test allows checking if enough reflections have been introduced into the calculation: one has to repeat the routine with an additional reflection and to check that the change induced in the largest eigenvalue is smaller than a given maximum [13].

The Bloch wave method is very efficient for unit cells containing a small number of atoms. The key point is that only a small number of waves are required to fully reproduce the electron state in such structures. When the lattice parameters of unit cell increases, further reflections are required to describe it; thus the calculation time increases considerably.

The multislice technique [14,15] is better adapted to calculate the contrast created by a large set of atoms. The atomic lattice is sectioned in many slices, normal to the incident beam. The incident beam interacts with the projected potential of the first slice, thereby giving rise to a number of diffracted waves. Then all these beams propagate towards the second slice potential and scatter once again. Renewed sets of waves iterate the same sequence of propagation and interaction till it leaves the specimen.

The slice thickness is a free parameter and can be diminished, improving thereby computation accuracy. The optimal way of atomic lattice cutting is when every slice is thin enough to satisfy the weak phase approximation. This criterion is fulfilled if the product of the maximum value of the projected potential on the scattering constant σ , defined by equation I-9, is smaller than 1. The scattering constant σ is inversely proportional to the electron wavelength. For microscope operating at 200 keV, σ is close to $6.26 \cdot 10^{-3}$ and for 300 keV microscope, σ equals to $5.32 \cdot 10^{-3}$.

The multislice method can only be applied to orthogonal atomic lattices. The majority of structures can be reformed to orthogonal ones by addition of further atoms; but often this procedure results in noticeable enlargement of atomic cells, which increases the calculation time.

Both techniques are illustrated by the calculation of a HRTEM image from a Si crystal using the JEOL 2010F microscope configuration. The simulation was carried on the Java version of Electron Microscopy Software (JEMS) [16]. Bloch wave algorithm uses Bethe's approach with 25 strong reflections. No sub-slicing is introduced for the multislice calculations.

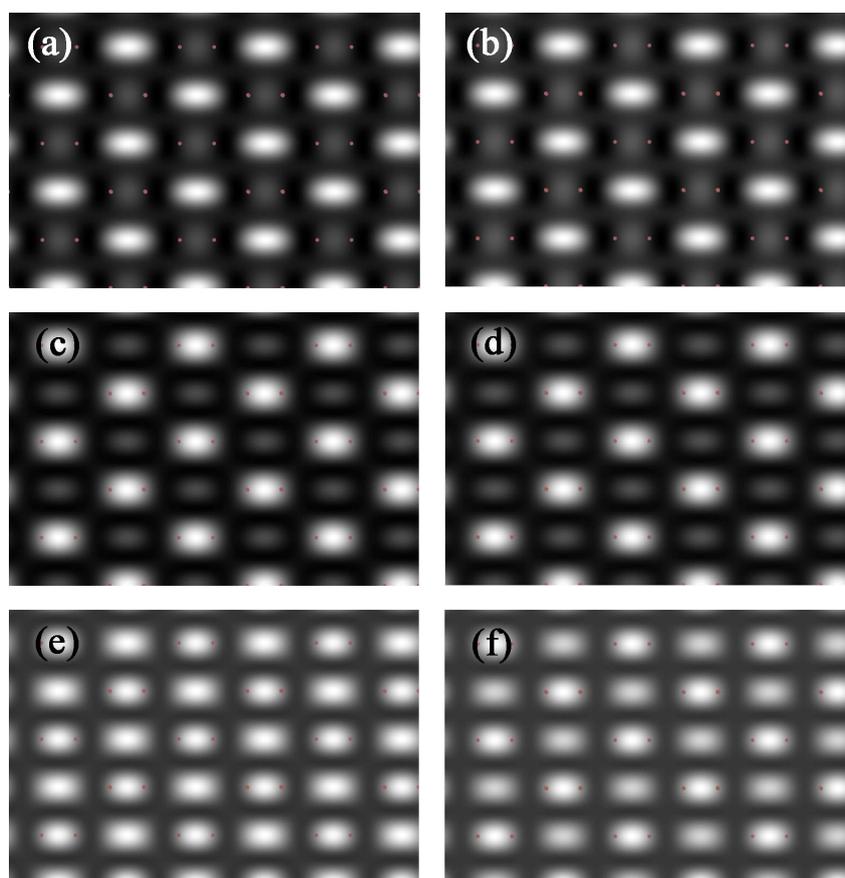


Figure I-14: A series of high-resolution TEM images of a Si crystal along the $[1,1,0]$ zone axis simulated by Blochwave (a), (c), (e) and Multislice techniques (b), (d), (f). The sample is 10 nm thick and the microscope parameters are those of the JEOL 2010F: $E=200\text{kV}$, energy spread = 0.8 eV, $C_s = 0.5\text{ mm}$ and objective aperture = 30 nm^{-1} . (a) and (b) images are taken at Scherzer defocus (-43.4 nm), (c) and (d) at zero defocus, and (e) and (f) at a defocus of -29 nm. The exact positions of atoms are indicated by the colour spots.

The projections of the distances between two neighbouring Si atoms along the $[1,1,0]$ zone axis is as small as 0.136 nm. A JEOL 2010F microscope having a point-to-point resolution of 0.194 nm is not able to resolve this distance. Thus, two atomic columns are represented by a single white spot (Figure I-14).

Both Bloch wave and multislice techniques produce rather similar images, while image contrast varies noticeably with changes of defocus. In the images acquired at Scherzer defocus (figures I-14 (a) and (b)) intensity maxima are between the atomic columns. Whereas pictures calculated at zero defocus (figures c and d) exhibit an opposite contrast. The defocus value of -29 nm (figures e and f) produces images with a double number of intensity maxima.

Besides, the image contrast changes considerably when the sample thickness changes. Consequently, HRTEM images can not be used for straight atomic position determination.

Only the combination of electron micrograph with image simulation techniques can provide reliable structural information. The best precision is obtained when a series of images with different defoci is matched with simulated data.

HRTEM images simulated using parameters of Cs-corrected machines are shown on Figure I-15. The point-to-point resolution of the Titan 80-300 with a Cs-corrector is about 0.08 nm, thus the Si doublets are instantly resolved (Figure I-15 (a)). Improving of the beam temporal coherency pushes information limit toward higher spatial frequencies, making image sharper (figures (b) and (c)). A hypothetical, aberrations-free machine produces an ultimate image (figure d).

So, HRTEM image simulation is essential to understand or even predict the contrast of experimental images. Another approach, called exit wave reconstruction, is capable to calculate back numerically the exit wave function [17]. This procedure is based on the treatment of a large focal series of images, supposing that all acquisition parameters are precisely known. Exit wave reconstruction was not used in this work, but image simulation was largely employed.

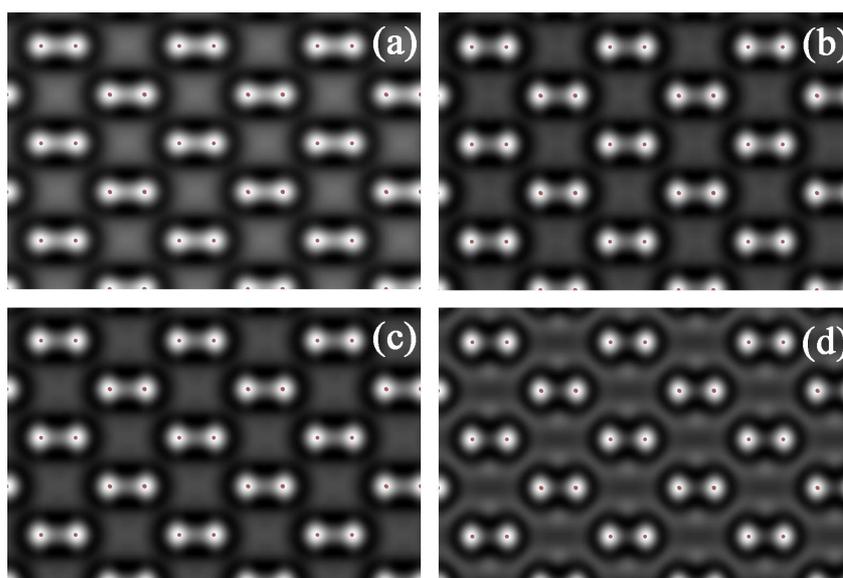


Figure I-15: High-resolution TEM images of a Si crystal simulated by the Blochwave technique for a Titan 80-300 microscope equipped with a Cs-corrector. Image (a) is for a machine operating at 300 keV with $C_3 = -5 \mu\text{m}$, $C_5 = 4 \text{ mm}$, $C_c = 2 \text{ mm}$, and energy spread = 0.7 eV. Further image improvement can be achieved either by correcting the chromatic aberration up to $C_c = 5 \mu\text{m}$ (b) or by decreasing the energy spread up to 0.3 eV (c). An ultimate image (d) was obtained on the ideal machine free of any aberrations and having an energy spread of 0.01 eV.

1.2.2. Heterostructure composition evaluation by high-resolution imaging

High-resolution transmission electron microscopy has a broad field of applications. The central objective of this thesis is an investigation of the structural properties and compositions of group-III nitride heterostructures. For doing this, we will focus on the measurement of the local lattice parameter (LLP), i. e. the distance between two adjacent atomic planes. The LLP varies with the chemical composition. The Vegard's law, which holds true for a majority of compounds, states a linear dependence between the lattice parameter a of a ternary alloy $A_xB_{1-x}C$ and its concentration x .

$$a(A_xB_{1-x}C) = a(BC) + x \cdot [a(AC) - a(BC)] \quad (\text{eq. I-18})$$

Using this equation the composition variation may be determined with a spatial resolution equivalent to an interatomic distance. It has to be noted that, despite the high lateral resolution, the LLP measured by HRTEM is always averaged through the TEM sample thickness (5-50 nm).

This approach is widely used to study compound structures, where composition variations may exist. Two main aspects giving rise to a deviation from the linearity of the Vegard's law should be considered. First, the stress field, which generally exists in heterostructures, may cause a displacement of atoms from their initial positions. However, if the atomic displacement is regular, the change of LLP can be analytically taken into account. The simplest case is the biaxial stress, which generally takes place in thin epilayers. We will produce some formulas in Section 2.3.2, when we will talk about the strain states in nitride films.

Another phenomenon that often affects the atomic arrangement is a surface relaxation effect in thin crystals. Atoms situated near free surfaces have a higher freedom to move. This results in additional atom displacements which minimizes the total stress energy. The TEM sample has to be thinned down to 5 - 50 nm, otherwise electron waves will be highly absorbed. So the number of near-surface atoms may be quite high (ratio surface/volume). The effect of atomic reorganization is non-linear and depends strongly on sample composition, thickness and geometry.

The problem can be solved numerically for a given TEM sample geometry using the Finite Element Calculations method [18]. The investigated structure is modulated by a mesh with nodes spaced at the lattice position. The elastic constants for ternary alloys are linearly interpolated between the binary endpoints. The boundary conditions determine the positions of the meshes situated at internal boundaries; the external surface atoms are not fixed.

An approximate solution is derived numerically in a series of successive iterations. The atoms are allowed to move from the initial position in any direction, and the change of the elastic energy is calculated. The system is supposed to be in equilibrium if any further atom displacement increases the total energy. Thereby, corrections for the mechanical tensions, bending and buckling behaviour are obtained.

1.2.3. Methods of strain measurement

Atomic displacement may be determined from the HRTEM images using different techniques. We will describe below two methods that were mainly employed in this thesis.

The primary approach, called **projection method** (PM), consists in the direct measurement of the distances between intensity maxima [19]. The algorithm is explained below on an example of an actual image treatment.

- A region of homogenous contrast is chosen on the HRTEM image (Figure I-16 (a)). The image is then projected perpendicularly to the direction in which the strain will be measured. Due to averaging the resulting intensity profile has a considerably improved signal-to-noise ratio.
- The strain determination requires a highly accurate localization of the maxima positions. The direct measurement of the distances between the profile maxima is, however, quite inaccurate, because of the low number of pixels per period. Fitting an analytic function is indispensable to determine the peaks' positions with a sub-pixel resolution. (Figure I-16 (b)).
- The position of the fitted function maxima are derived either numerically or analytically. Local lattice parameters are calculated as a refined distance between two

adjacent maxima (Figure I-16 (c)). The strain profile is generally normalized in a zone where the lattice parameter is known and stable.

In this thesis all HRTEM images were operated under Digital Micrograph, which is a recognized software for visualizing and processing of digital image data [20]. The projection method was realised by a specially developed script for Digital Micrograph. The script treats profiles of HRTEM images reproducing the routine described above. Different types of fitting functions were tested. The best precision of strain measurement was obtained when the peaks were fitted by Chebyshev's polynomials. Two parameters, the polynomials degree and the number of points used for fitting, can be adjusted to optimize the accuracy of the strain measurement.

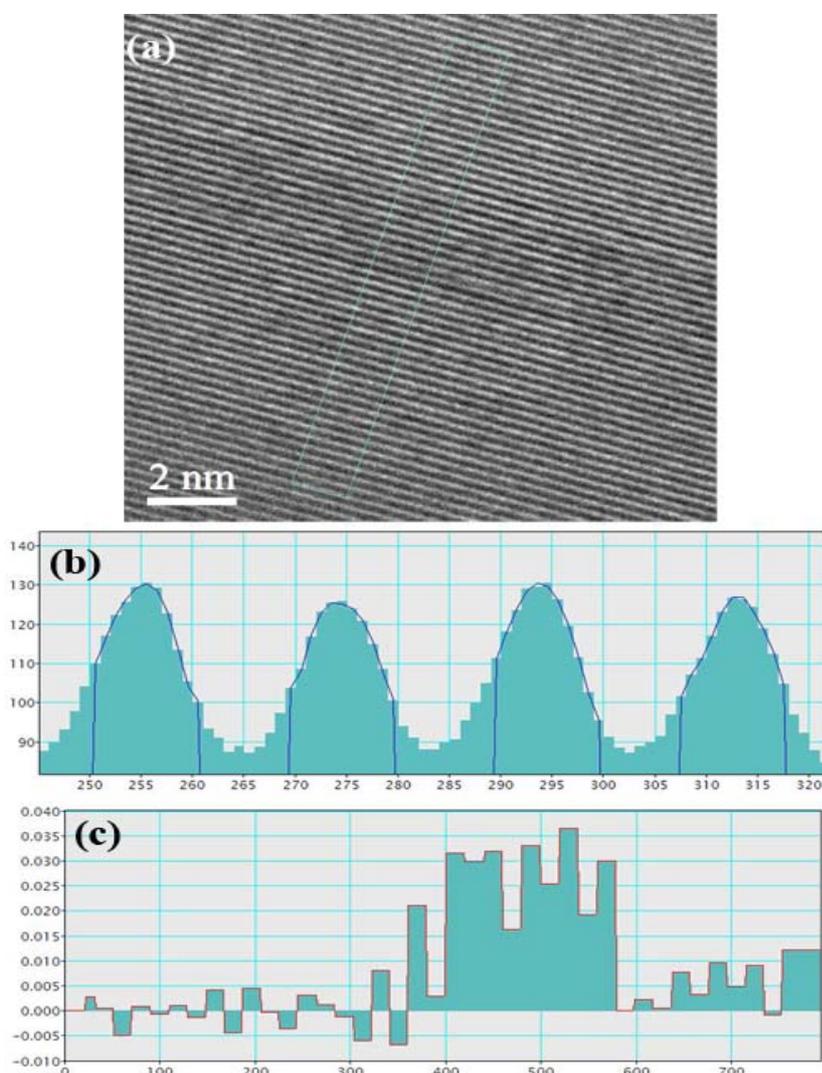


Figure I-16: HRTEM image of an InGaN quantum well embedded in a GaN matrix (a). Only the (0002) atomic planes are visualized. Intensity maxima of the averaged profile plotted in the [0001] direction are fitted by polynomials (b). Resulting strain profile normalized on the lattice parameter of the underlying GaN layer (c). The abscissa axis of the both profiles is in pixels.

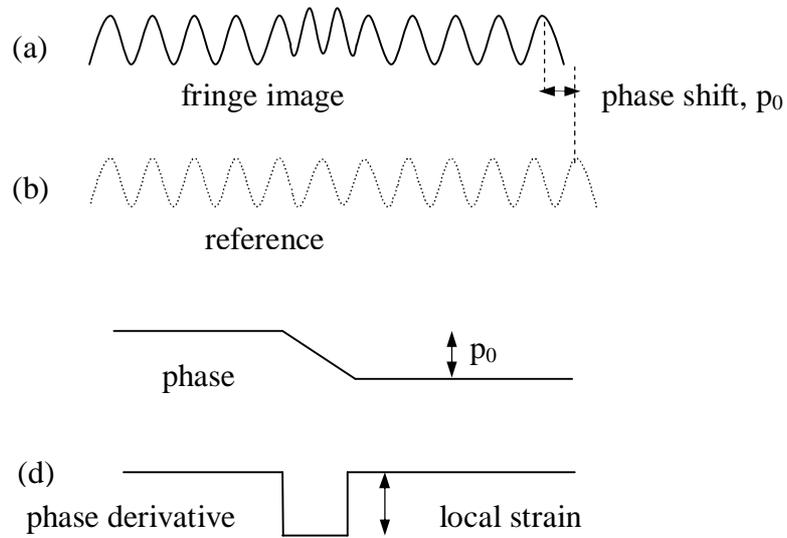


Figure I-17: Local displacements in a fringe image (a) relative to a reference image (b) gives rise to a phase shift (c). The local strain (d) can be found to be the phase derivative.

An alternative approach, called **geometric phase analysis (GPA)**, treats data in the reciprocal space [21]. The strain is determined by means of the phase shift related to a chosen Bragg's reflection. The local strain is measured with respect to a reference region contained in the same image, where the lattice parameters are known.

The relationship between the local strain and the phase in the 1D case is illustrated in Figure I-17. The local change of the periodicity of fringes relative to a reference frequency gives rise to a phase shift between them. This phase shift is proportional to the amount of periodicity difference. The local strain, thereby, can be found by the derivative of the phase.

The data processing is illustrated for the same HRTEM image as the one used for the demonstration of the projection method (Figure I-16 (a)).

- In a first step, the Fourier transform of the HRTEM image is calculated. This procedure reveals the Bragg-reflections of the waves contributing to the image formation (Figure I-18 (a)). One reflection is selected as a reciprocal lattice vector \mathbf{g} and a symmetric mask of a certain size and shape is applied around the chosen reflection.

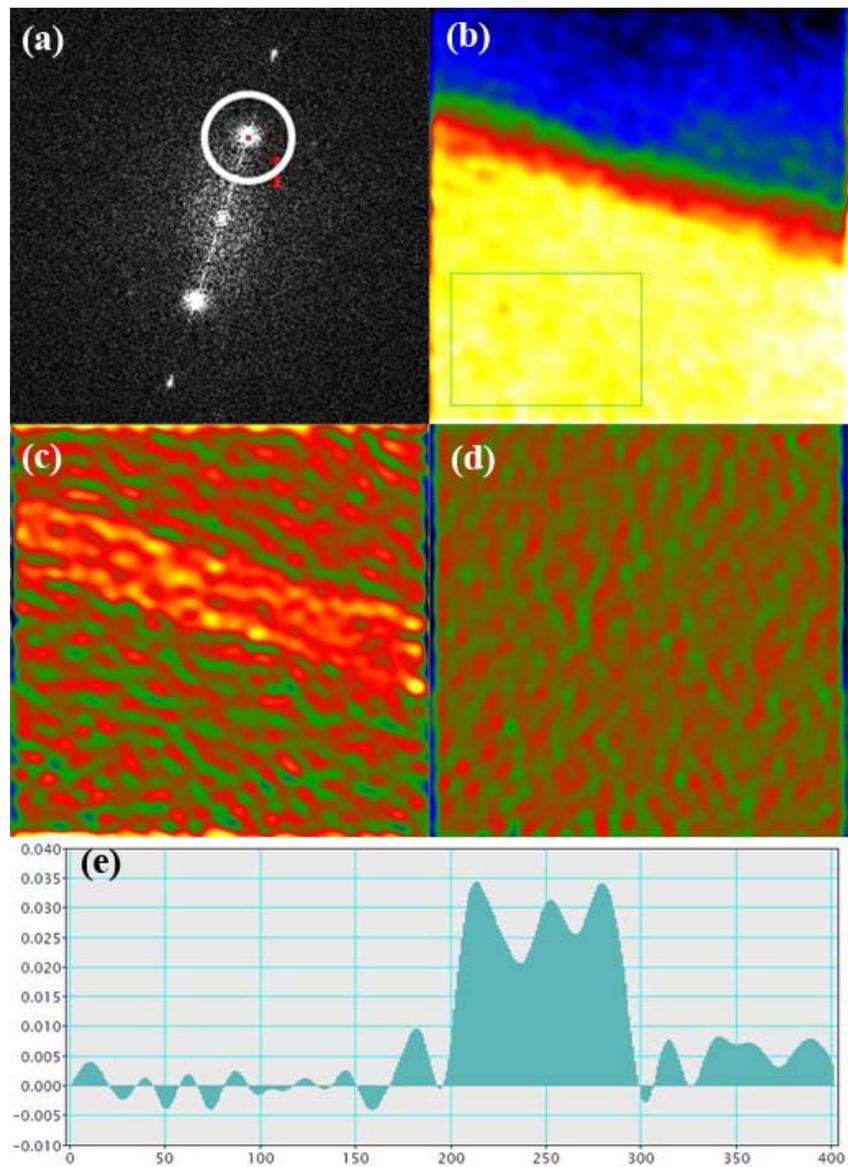


Figure I-18: The power spectrum of the HRTEM image with a mask (white circle) placed around the (0002) reflection, which is used for the phase calculation (a). The phase image; the reference zone is indicated by the green rectangle (b). The strain (c) and rotation (d) maps obtained from the phase image. A profile traced on the strain map in the [0001] direction; the abscissa axis is in pixels.

- After the application of the mask a complex inverse Fourier transformation is calculated. Further calculations deal only with the phase component. A reference region with a known lattice parameter is selected in the phase image (Figure I-18 (b)). The reciprocal vector is then refined in order to minimize the phase variation within the reference region.

- The deformation (Figure I-18 (c)) map is obtained by taking the derivative of the phase image. The map of fringes rotation (Figure I-18 (d)) can be calculated as well. All the components of the 2D-strain tensor can be determined in a similar way using two non-collinear reciprocal vectors.

The strain map profile is shown on Figure I-18 (e). It is less noisy than the one obtained by the projection method (Figure I-16 (c)). In GPA, the noise reduction is governed by the mask size: the smaller the mask the smoother the image.

However, the spatial resolution of the strain map is inversely related to the chosen mask size. Even if a noise-free image is treated with the largest possible mask, the information is averaged over two lattice periods [22]. Consequently the phase image size can be reduced without any information lost. In the given example the phase image was binned by two (note abscissa scale difference of the GPA and PM strain profiles).

A commercially available plugin GPA v 2.0 for Digital Micrograph was used in this work. It includes three standard mask sizes. The large one has radius of $g/2$, the medium $g/3$ and the small one $g/4$. In practice, the mask size is limited in the range from $g/2$ to $g/4$. The use of a mask larger than $g/2$ can create artificial fluctuations; too small masks average out details. The shape of the mask has a less influence. In this work, a cosine mask was used for experimental and simulated images treatment.

In Section 3.2 we will compare both strain measurement techniques. The GPA and projection method will be tested on artificially created lattice images. The ability of both techniques to treat noisy images will be illustrated as well.

1.3. Complementary techniques

The combination of high-resolution imaging together with strain measurement techniques is a very powerful tool for analysis of the heterostructure composition. Nevertheless, the contrast of HRTEM images is quite sensible to experimental parameters like sample thickness and or defocus of the microscope [23]. The errors in strain field measurement may be eliminated using proper imaging conditions [24].

Anyhow, the displacement of the atoms inside heterostructures is not always due to a variation in composition. For example, nonuniform strain fields around a quantum dot can not be separated from possible composition variations. Thus, the application of alternative techniques is indispensable for the validation of HRTEM results.

1.3.1. High-angle annular dark-field imaging

The data obtained by high-resolution imaging were completed by high-angle annular dark-field (HAADF) in a scanning transmission electron microscopy (STEM) mode [25]. The idea of HAADF techniques is to use the electrons which are scattered at high angles for image formation. The STEM mode consists in scanning the sample by a focused electron beam (see Figure I-19). Thereby, a narrow zone of the specimen is illuminated at each moment; the electrons which are scattered are detected by a high sensitive annular shaped detector. Since the microscope is in diffraction mode, the detection angle can be chosen by a simple variation of the camera length.

The spatial resolution of the STEM-HAADF method is mainly governed by the electron beam size. As in HRTEM, the spherical aberration of the illuminating system plays an important role, since the condenser lens, which focuses the beam, has the same imperfections as the objective lens. Accordingly, the resolution can be improved by mounting a Cs-corrector between the condenser lens and the sample. Thereby, beam diameters as small as 70 pm can be achieved [26].

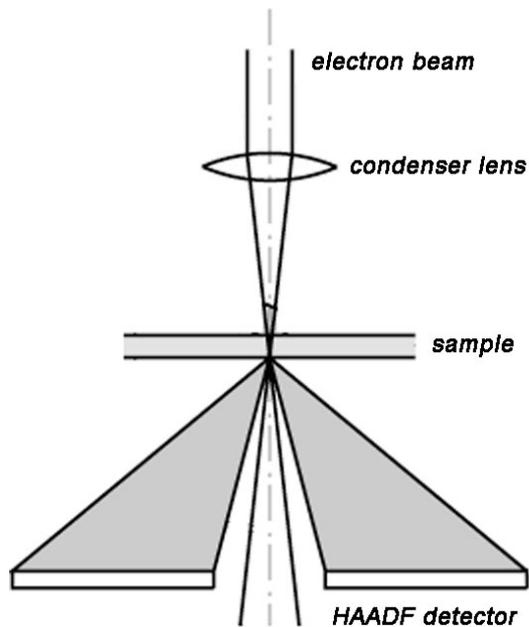


Figure I-19: Scheme of the STEM-HAADF imaging mode. The sample is scanned by a convergent electron beam. Electrons scattered to high angles are collected by a HAADF detector.

If the electron probe size is smaller than the interatomic spacings, individual atomic columns can be resolved in the image. Bright spots in the HAADF images always correspond to the atomic positions, as opposed to the HRTEM method. According to the Rutherford's law, the probability that an electron will be scattered to a high-angle is proportional to the square of the scattering atom weight Z [27]. Thus, a qualitative interpretation of HAADF images is quite straightforward: heavier atoms exhibit higher brightness (Figure I-20).

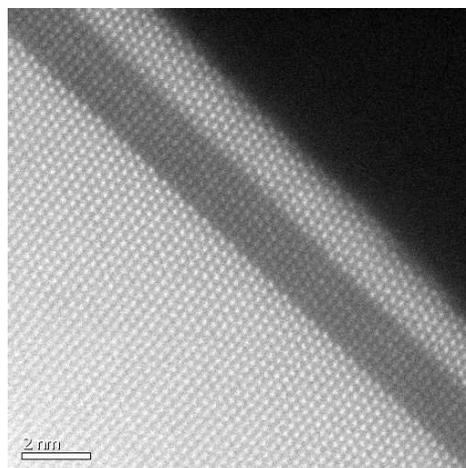


Figure I-20: Example of a HAADF image of a GaN/AlN/GaN heterostructure. The image is taken from the site of the LPN laboratory [28]

Anyhow, a quantitative analysis, i.e. the determination of the chemical composition of individual atom columns, is more complicated. The reason is that the overall intensity varies with the acquisition conditions. In particular, some elastically scattered electrons may also reach HAADF detector, changing thereby the image contrast [29]. The fraction of elastic electrons is smaller if the HAADF detector is placed closer to the sample, but the total image intensity decreases as well.

The value of HAADF signal changes with the sample thickness, which necessitates the evaluation of the thickness. Besides, if the incoming beam direction is parallel to one of the major crystalline axes, the electrons will pass along the atomic columns with increased probability. The channelling effect reduces the high angle scattering cross section.

A simple model enabling quantitative interpretation of HAADF images is presented in the third Chapter. This model implies that each atom contributes independently to the HAADF signal, neglecting thereby channelling effects. A serious analysis requires extensive calculations and HAADF image simulations [30, 31].

1.3.2. Electron energy loss spectroscopy

Electron energy loss spectroscopy (EELS) is a technique in which a substance is irradiated with electrons. The energy distribution of the transmitted electrons is measured to extract chemical and electronic information on the sample (Figure I-21).

The loss of energy is quite characteristic from the composition of the studied material. A typical electron energy loss spectrum is shown on Figure I-22. The energy resolution of the EELS spectrum is limited by the energy spread of the incoming beam. The use of a monochromator is preferable to produce a beam with a narrow range of kinetic energy.

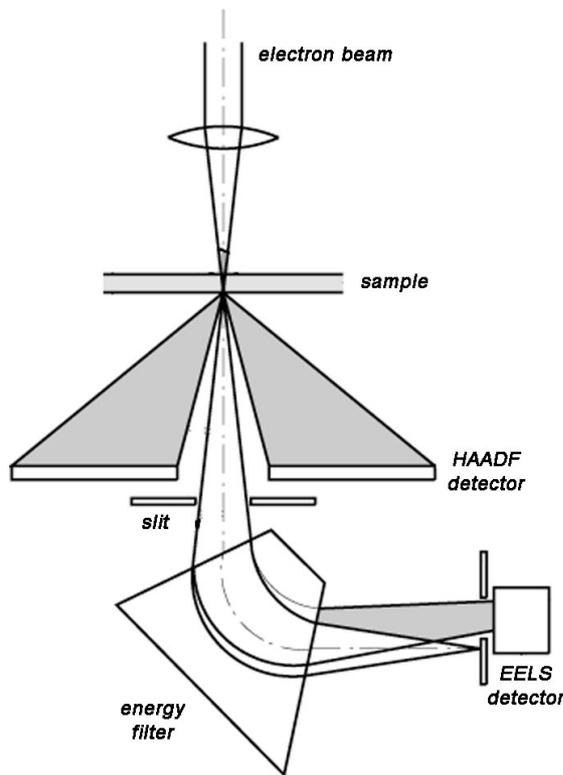


Figure I-21: Scheme of EELS combined with STEM-HAADF. The outgoing electrons, diffracted to a small angle, enter in the energy dispersive filter. The angle of deflection is proportional to the electron energy, thus the posterior EELS detector records the energy spectrum.

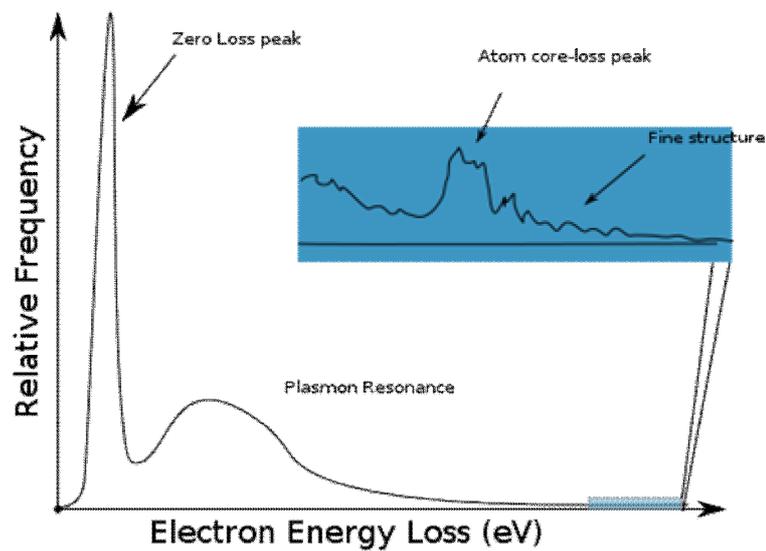


Figure I-22: EELS spectrum has a strong zero-loss peak and a plasmon resonance. The atom core-loss peaks are generally much less intense.

The predominant feature of the spectrum is the zero-loss peak, which is due to the electrons transmitted without any loss of energy, or with minor energy losses, since the energy loss due to the excitation of phonons is hidden inside the zero loss peak. The zero-loss peak width characterizes the energy distribution of the incoming electrons. Besides the specimen thickness can be evaluated by comparing the integral intensities of the zero-loss peak and of the whole spectrum [32].

Inelastic interactions include phonon excitations, inter- and intra-band transitions, plasmon excitations and inner shell ionizations. The plasmon resonance is due to the excitation of the weakly-bound electrons into collective oscillations by the electron beam. It dominates in materials with free-electron structure such as Al or Mg. The plasmons are located either in the bulk volume or in the sample surface. The bulk plasmons are localized within a volume of a few nm in size, thus providing a slightly averaged signal.

The volume plasmon energy is proportional to the square of the valence electron density; thereby the measurement of the plasmon energy loss provides an indirect chemical information. In some cases, the quantitative analysis may be quite straightforward if the plasmon energy shift is linearly proportional to the fractional concentration [33]. However, plasmon energy lies in the narrow range from 5 to 25 eV [34]. The plasmon energy shift may be as small as a few eV, which limits the use of this method on conventional microscopes (energy spread ~ 1 eV).

The low-loss part of the spectrum also includes peaks caused by inter- and intra-band transitions. An electron beam may excite a non valence electron into an unoccupied state. The corresponding energy losses can be used for the identification of the chemical components. But those peaks are generally very weak (comparing with the zero-loss and plasmon resonance) and the extraction of quantitative analysis is difficult.

The high-loss part of the spectrum represents the probability of the excitation of a core electron in an unoccupied state. Since these excitations are well distinct for all the elements and the onset of energy of the excitations for most elements are known and published, this spectral information gives a direct access to the chemical composition of the sample. In addition, slight energy shifts and the shape of the spectra can help, in combination with electronic structure calculations, to extract chemical, geometric or electronic information.

The peak intensity after the background extraction is proportional to the number of atoms which inner-shells were ionised by the electron beam. With some care, one can determine the numbers of atoms responsible for every energy loss peak. But nuances of spectrum treatment (such as background approximation, size of the window chosen for intensity measurement) have a strong influence on the outcome of the analysis. The more precise analyses require the matching of a model to the experimental data [35]. However, the intensity of the EELS signal decreases exponentially with the value of the ionisation energy. Thus, this method suits better for light elements, since heavier atoms produce too weak and too noisy signal to be analysed. In addition, the unwanted multiple loss events become more visible with an increasing of the atomic mass and the sample thickness.

The spatial resolution of EELS is determined by the diameter of the electron beam. The energy resolution is limited by the beam energy spread. Microscopes equipped with a monochromator are able to measure spectra with an energy resolution better than 0.1 eV [36].

All EELS spectra presented in this thesis were acquired in the STEM mode described above. However, EELS spectra may also be obtained using energy-filtered transmission electron microscopy (EFTEM). An EFTEM image is formed using inelastic scattered electrons of a particular energy range. A series of images acquired at successive energy ranges may be gathered in so called data cubes. The data cube is a three-dimensional array of inelastic scattered electron intensity, where two dimensions represents the spatial coordinates and the third one the electron energy. An EELS spectrum may be extracted for each image point.

Bibliography

- [1] L. de Broglie, Thesis (Paris), 1924; reprinted in *Ann. de Physique* **3**, 22 (1925).
- [2] G. P. Thomson and A. Reid *Nature* **119**, 890 (1927).
- [3] M. Knoll and E. Ruska, *Z. Physik* **78**, 318 (1932).
- [4] The figure was taken from the site of Department of Metallurgical and Materials Engineering of METU <http://www.mete.metu.edu.tr/pages/tem/TEMtext/TEMtext.html>
- [5] O. Scherzer, *Zeitschrift für Physik*, **101**, 593 (1936).
- [6] P.L. Fejes, *Acta Cryst*, **A33**, 109 (1977)
- [7] S. Uhlemann and M. Haider, *Ultramicroscopy* **72**, 109 (1998).
- [8] O. Scherzer, *J. Appl. Phys.* **20**, 20 (1949).
- [9] H. Lichte, *Ultramicroscopy* **38**, 13 (1991)
- [10] A. Tonomura, H. Kasai, O. Kamimura, T. Matsuda, K. Harada, Y. Nakayama, J. Shimoyama, K. Kishio, T. Hanaguri, K. Kitazawa, M. Sasasek and S. Okayasu, *Nature* **412**, 620 (2001).
- [11] H. H. Rose, *Optic* **85**, 19(1990).
- [12] A.J.F. Metherell, "Diffraction of Electrons by Perfect Crystals", in *Electron Microscopy in Material Science*, Ettore Majorana, (1973).
- [13] P. Stadelmann, *Image Calculation Techniques*, http://cecm.insa-lyon.fr/ciol/asu94/ICT_5.html.
- [14] J. M. Cowley and A.F. Moodie, "The Scattering of Electrons by Atoms and Crystals", *Acta Cryst.* **10**, 609-619 (1957).
- [15] K. Ishizuka and N. Uyda, *Acta Cryst.* **A33** 740-749 (1977).
- [16] <http://cimewww.epfl.ch/people/stadelmann/jemsWebSite/jems.html>.
- [17] D. van Dyck and M. de Beeck, *Proc. 12th Int. Congr. Electron Microscopy*, Seattle, p. 26 (1990).
- [18] M. M. J. Treacy and J. M. Gibson, *J. Vac. Sci. Technol.* **B4**, 1458 (1986).
- [19] R. Bierwolf, M. Hohenstein, F. Phillipp, O. Brandt, G. E. Crook and K. Ploog, *Ultramicroscopy*, **49**, 273 (1993).
- [20] <http://www.gatan.com/>.
- [21] M. J. Hÿtch, E. Snoeck and R. Kilaas, *Ultramicroscopy* **74**, 131 (1998).
- [22] J. L. Rouvière and E. Sarigiannidou, *Ultramicroscopy* **106**, 1 (2005).
- [23] P. Bayle, J. Thibault, *Microsc. Microanal. Microstruct.* **8**, 125 (1997).

-
- [24] M. J. Hÿtch, T. Plamann, *Ultramicroscopy* **87**, 199 (2001).
- [25] D. E. Jesson and S. J. Pennycook, *Proc. Roy. Soc. A* **449**, 273 (1995).
- [26] P. E. Batson *Ultramicroscopy*, **106**, 1104 (2006) .
- [27] E. Rutherford, *Phil.Mag.* **21**, 669 (1911).
- [28] <http://www.lpn.cnrs.fr/en/ELPHYSE/EPINANO.php>.
- [29] T. Walter, *J. Microsc.* **221**, 137 (2006).
- [30] A. Rosenauer, K. Gries, K. Muller, A. Pretorius, M. Schowalter, A. Avramescu, K. Engl and S. Lutgen *Ultramicroscopy* **109**, 1171–1182 (2009).
- [31] V. Grillo, *Ultramicroscopy* **109**, 1453–1464 (2009).
- [32] K. Iakoubovskii, K. Mitsubishi, Y. Nakayama and K. Furuya, *Microscopy Research and Technique* **71**, 626 (2008).
- [33] D. B. Williams and J. W. Edington *J. Microsc.* **108**, 113 (1976).
- [34] R. F. Egerton, “*Electron Energy-Loss Spectroscopy in the Electron Microscope*”, Plenum, London, (1996).
- [35] R. Door, D. Gangler, *Ultramicroscopy*, **58**, 197, (1995).
- [36] H. H. Rose, *Sci. Technol. Adv. Mater.* **9**, 014107 (2008).

Chapter II NITRIDE SEMICONDUCTORS

In this Chapter, we will discuss the GaN properties and its applications. We will address some problems related with the GaN production and introduce two techniques commonly used for GaN-based heterostructures growth. At the end, the structural and elastic properties of III-nitrides will be presented.

2.1. Gallium nitride specificities and domains of applications

Gallium nitride (GaN) is the most important semiconductor material after silicon in terms of market size. Optoelectronic GaN-based devices have been successfully adopted for many commercially attractive applications such as UV-detectors, solar cells, light emitting diodes and laser diodes [1].

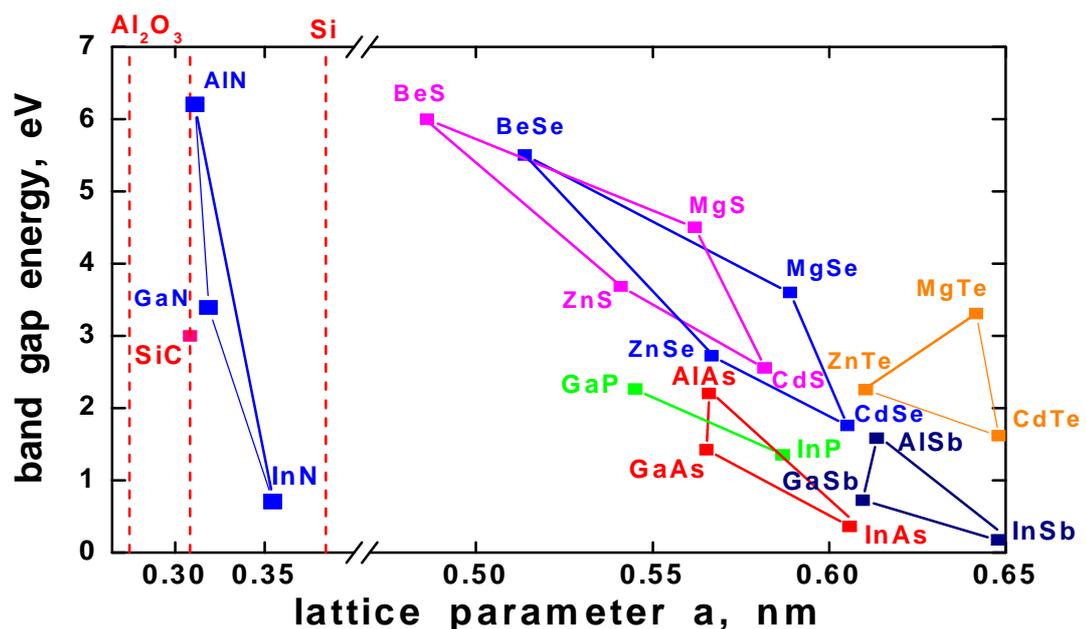


Figure II-1: Band gap energy versus lattice parameters for different semiconductors. Vertical dotted lines show the lattice constants of the common substrates.

The main advantages of GaN material, which caused its wide application in optoelectronics, are the direct electronic band structure and the high thermal stability [2]. Moreover, the band gap energy of GaN-based alloys changes in a wide range, comparing with other semiconductors (Figure II-1). In particular, the whole visible spectrum (1.6 – 3.2 eV) can be covered by varying the In fraction of $\text{In}_x\text{Ga}_{1-x}\text{N}$ alloy from 7% to 66%.

Besides, GaN is considered as a key material for the next generation of high-power and high-frequency transistors [3], and GaN-based high electron mobility transistors (HEMT) already have numerous commercial applications in wireless infrastructures.

In spite of a wide commercial use, the GaN production is confronted with certain fundamental difficulties. The principal one is the lack of a suitable substrate for epitaxial growth. Manufacture of bulk GaN crystals is too labour-consuming to be reproduced at industrial scale. Especially since the surface of produced GaN crystals is generally smaller than one inch.

The lattice mismatch of gallium nitride with commonly available substrates, such as sapphire or silicon, is larger than 10% (Figure II-1). GaN lattice parameters match well with the silicon carbide (SiC) ones, but, however, the crystalline quality of GaN grown on SiC is not sufficiently good. Besides, the SiC substrates are quite expensive. As a result, GaN is generally grown on sapphire or silicon substrates.

The other problem is that the different nitride compounds have different optimal epitaxial growth temperatures. In metalorganic vapour phase epitaxy, a technique commonly used for nitride compounds growth, GaN is grown at 1000°C, InN at 600°C, AlN at 1100°C. This may cause some side effects, such as segregation, interdiffusion or desorption in GaN-based ternary alloys.

2.2. Growth of III-nitrides

The structural quality of semiconductors has a strong impact on their optical and electrical properties. That is the reason why most of III-nitride films are produced by epitaxy, a method allowing monocrystalline layers growth. At present, two techniques are extensively used for nitride compound growth: molecular beam epitaxy (MBE) and vapour phase epitaxy (VPE).

2.2.1. Molecular beam epitaxy

The MBE principle consists in the deposition of evaporated (or sublimated) material on a crystalline substrate under an ultra high vacuum (about 10^{-8} Pa). High operation vacuum enables the use of reflection high energy electron diffraction (RHEED), which is a powerful technique for the characterization of crystalline material surfaces. Surface morphology [4], lattice parameter [5] and deposition rate [6] can be accurately measured by means of RHEED.

Thanks to RHEED, a precise control of the growing layer thickness and composition is possible in MBE [7]. The flux of the molecular beam can be adjusted by varying the effusion cell temperature. Sharp hetero-interfaces can be created by quickly cutting-off the molecular beams with a shutter.

Two different types of reactive nitrogen sources are used for the MBE growth of III-nitride films. Either nitrogen atoms are created by a radio frequency plasma source, which dissociates the N_2 molecules, or reactive nitrogen is provided by NH_3 decomposition on the heated substrate, which implies the use of higher temperatures for GaN growth to reach a sufficient dissociation efficiency: $800^\circ C$, comparing with $700-750^\circ C$ in the case of plasma assisted MBE.

The most important aspect of MBE is the slow deposition rate (typically less than $1 \mu m$ per hour). Indeed, MBE is a powerful technique for the realisation of quantum wells or quantum dots heterostructures which requires the control of the deposited thickness at the monolayer scales.

In GaN-based MBE the range of growth temperatures is limited by two factors: at high temperatures, the decomposition rate becomes faster than the deposition rate, governing an

upper limit for the substrate temperature. On the other hand, low substrate temperatures reduce surface atom mobility, leading to an increase of the defect densities. Consequently, GaN growth must be performed in a relatively narrow temperature range (770° to 810°C for ammonia-assisted MBE).

2.2.2. Metalorganic vapour phase epitaxy

Another method of crystal growth is metalorganic vapour phase epitaxy (MOVPE), also called metalorganic chemical vapour deposition (MOCVD). The initial reagents arrive to the reactor in vapour metalorganic form, such as trimethylgallium or trimethylindium. Metalorganic components are supplied to the substrate by a carrier gas, such as nitrogen or hydrogen. The formation of the epitaxial layer occurs by the organic component pyrolysis on the heated substrate surface and following chemical reactions of the released reagents. The growth pressure of the gas phase varies from 2 to 100 kPa.

MOVPE is performed at high temperatures (1000 – 1100°C for GaN), which imposes special requirements to the reactor components in order to prevent their corrosion. Cooling water must be flowing through the channels within the reactor walls to avoid overheating. The VPE reactor is usually equipped with a laser reflectometry system, which is used for the growing layer thickness evaluation.

Comparing with MBE, MOVPE technique is more suited for industrial needs. The typical growth rate is about 3-5 μm per hour. Higher growth temperatures and thicker layers gives a better crystalline quality of the grown heterostructures. Epitaxial films grown by MOVPE are quite uniform, which enables the use of large substrates, up to 8-10 inches in diameter. Besides, epitaxy can be done on several wafers simultaneously using a multiple wafers rotation system. However, very recently multi-wafer MBE systems have also been commercialized for the production of GaN-based devices.

2.3. Structural properties of III-nitride materials

2.3.1. Crystal structure

The equilibrium crystallographic phase of group-III nitrides is the wurtzite structure shown on Figure II-2. Each atom is in a sp^3 -hybridization state and is bounded to four atoms of the other type, located at the corners of a tetrahedron. The atoms arrangement can be represented as two hexagonal close-packed sublattices, shifted with respect to each other. They form a close-packing sequence of ...ABAB... type³.

The distance between neighbouring atoms in basal planes equals to the lattice constant a , the spacing between the close-packed planes composed by atoms of the same type is $c/2$. The distance between two different atoms along the c -axis is denoted as u (see Figure II-2).

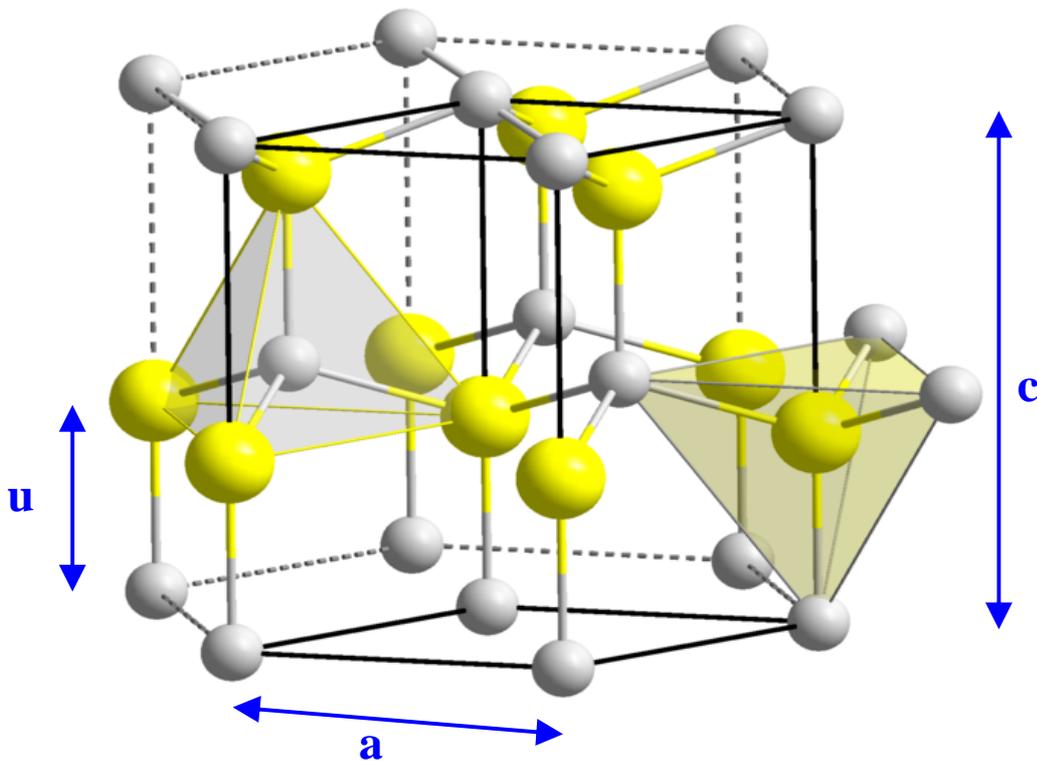


Figure II-2: Atomic structure of GaN: Ga atoms are displayed in yellow and N atoms in grey; chemical bounds are symbolized by two-colour lines. The wurtzite lattice is characterized by two lattice constants: a and c .

³ The III-V nitride semiconductors also exist in the zinc-blende phase, which is represented as ...ABCABC... sequence, but this configuration is meta-stable.

	a (nm)	c (nm)	c/a	u/c
GaN	0.3189	0.5185	1.626	0.377
AlN	0.3113	0.4982	1.600	0.382
InN	0.3538	0.5703	1.612	0.377

Table II-1: Lattice parameters and ratios for GaN, InN and AlN crystals in wurtzite phase.

The wurtzite structure belongs to the P6₃mc space group. The unit cell can be defined by two atomic positions: the nitrogen is situated in (2/3, 1/3, 0) site and the III-metal atom is placed in (2/3, 1/3, 3/8) site. Then the lattice parameters of this unit cell are related as:

$$c/a = 2\sqrt{2/3} \approx 1.633; \quad u/c = 3/8 = 0.375 \quad (\text{eq. II-1})$$

The unit cell of the actual structures is slightly distorted comparing to that described above. The experimentally measured values [8, 9, 10, 11] and their interrelations are presented in the Table II-1.

The directions in hexagonal system are generally described by four translational vectors in order to emphasize the hexagonal symmetry; we will also use this four-index notation above. The relations between three- and four-index are given in Appendix A.

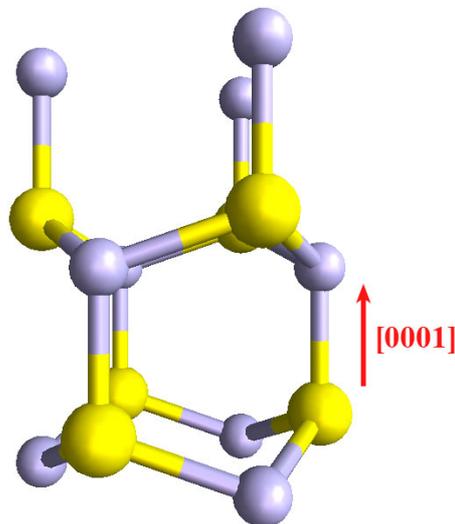


Figure II-3: The [0001] direction in wurtzite lattice is defined as pointing from Ga to N.

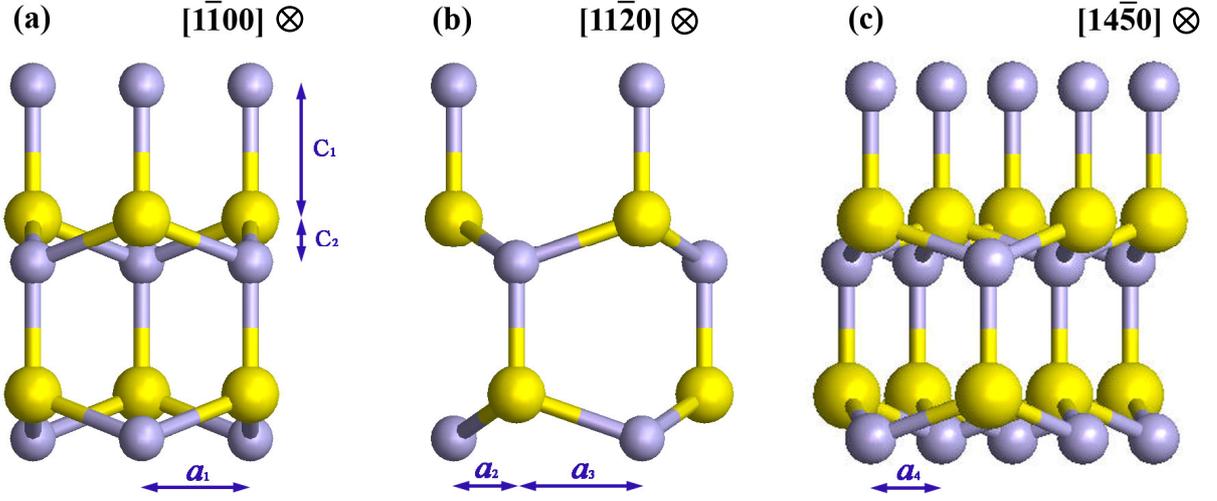


Figure II-4: Perspectives of the wurtzite lattice along $[\bar{1}100]$ (a), $[11\bar{2}0]$ (b) and $[14\bar{5}0]$ (c) directions. The blue arrows indicate the projections of the interatomic spacing on the a and c axes.

In the four-index notation, the c -axis, perpendicular to the close-packing planes, is denoted as $[0001]$. The wurtzite structure is non-centrosymmetric: the $[0001]$ and $[000\bar{1}]$ directions are not equivalent and the c -axis is polar. The $[0001]$ direction is defined as a vector pointing from the Ga (or other III-metal) to the N atom and collinear to the c -axis (Figure II-3). The crystal grown in the $[0001]$ direction is called Ga-polar; the crystal grown in the $[000\bar{1}]$ direction is called N-polar.

In this work, the samples were observed by TEM along three principal directions: $[1\bar{1}00]$, $[11\bar{2}0]$ and $[14\bar{5}0]$. The projections of the atomic structure along those directions are shown on Figure II-4.

An interatomic spacing can be resolved only if it is greater than the microscope point-to-point resolution. The calculations of the interatomic distances $a_1 - a_4$ in the different projections are done in Appendix B. Here are the analytic expressions:

$$a_1 = a/2, a_2 = a/2\sqrt{3}, a_3 = a/\sqrt{3}, a_4 = a\sqrt{3/28};$$

$$c_1 = u, c_2 = (1/2c-u) \quad (\text{eq. II-2})$$

where a and c are the wurtzite lattice constants. The numerical values of interatomic distances in GaN (based on the data from Table II-1) are:

$$a_1 = 159 \text{ pm}, a_2 = 92 \text{ pm}, a_3 = 184 \text{ pm}, a_4 = 104 \text{ pm};$$

$$c_1 = 195 \text{ pm}, c_2 = 64 \text{ pm} \quad (\text{eq. II-3})$$

We will refer to these data in the Third Chapter, when we will introduce the microscopes employed in this work and will discuss their resolvable interatomic distances for the studied structures.

2.3.2. Elastic properties

Epitaxial heterostructures are often strained due to the lattice mismatch (Figure II-1). In some cases, this deformation can be described analytically. Let us consider a continuous and atomically flat layer completely strained on a template (presuming thereby the absence of any plastic relaxations). The template is supposed to remain strainless, whereas the epitaxial film stretches or compresses in such a way to match the lateral size of the template unit cell. This effect is illustrated on the example of a silicon film deposited on a germanium substrate (see Figure II-5).

The thin layer deformation can be described using the Hooke's law. The matrix representation should be used for anisotropic materials. Anyhow the stiffness tensor of hexagonal structures has only six independent components.

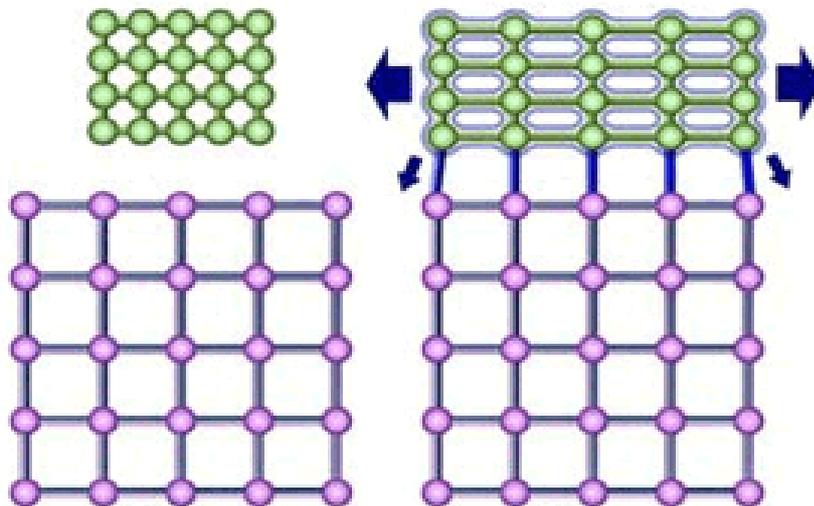


Figure II-5: Silicon (in green) has smaller lattice parameters than germanium (in violet). The blue arrows show the deformation of the Si film grown on the Ge substrate. The deformation effect is highly exaggerated.

For the wurtzite structure grown in [0001] direction the Hooke's law is written as:

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{xy} \\ \sigma_{yz} \\ \sigma_{zx} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} \begin{bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ 2\epsilon_{xy} \\ 2\epsilon_{yz} \\ 2\epsilon_{zx} \end{bmatrix} \quad (\text{eq. II-4})$$

where σ_{ij} , C_{ij} and ϵ_{ij} are the components of the stress, stiffness and strain tensors respectively; the z direction is taken along the [0001] axis. The strain tensor components are expressed in terms of lattice constants as

$$\mathbf{e}_{xx} = \mathbf{e}_{yy} = \frac{a - a_0}{a_0}; \quad \mathbf{e}_{zz} = \frac{c - c_0}{c_0} \quad (\text{eq. II-5})$$

where a_0 and c_0 are the substrate's unit cell parameters and a and c are the thin layer's unit cell parameters.

The absence of stress in the [0001] direction ($\sigma_{zz} = 0$) imposes a correlation between the strain tensor components (see eq. II-4).

$$\mathbf{e}_{zz} = -2 \frac{C_{13}}{C_{33}} \mathbf{e}_{xx} \quad (\text{eq. II-6})$$

Thereby an extension or contraction of the epitaxial layer in the z-direction ϵ_{zz} is strictly related to the lateral deformation ϵ_{xx} . The latter is due to the lattice mismatch between the epitaxial layer and the template (eq. II-5). In that way, the vertical strain is fully determined for any combination of materials.

Biaxial stress takes place in bulk heterostructures, whereas TEM samples can be as thin as 5 nm. We can go further and examine another limiting case: an extremely thin specimen, where the stress in one in-plane lateral direction is zero due to the surface relaxation effects (so called uniaxial stress). The conditions $\sigma_{zz} = 0$ and $\sigma_{yy} = 0$ imposes another correlation between ϵ_{xx} and ϵ_{zz} :

$$\mathbf{e}_{zz} = -\frac{C_{13}}{C_{33}} \mathbf{e}_{xx} \quad (\text{eq. II-7})$$

The actual TEM specimens are never completely relaxed, i.e. $\sigma_{yy} \neq 0$. In Section 3.1 an uncertainty in the composition determination due to the surface relaxation effects will be estimated.

Another origin of inaccuracy in the composition determination is that the elastic coefficients reported for group-III nitrides [12, 13, 14, 15, 16, 17, 18] vary in a large range (see Table II-2). Besides, only theoretical calculations of Wright and Kim are available for InN.

The elastic constants reported by Wright are frequently used for strain calculations [19, 20, 21]. In this thesis the Wright's data (signed in bold in Table II-2) were also utilized. Nevertheless in Section 3.1 the set of Kim's elastic constants was used to estimate an inaccuracy in the composition determination.

		c_{11}	c_{12}	c_{13}	c_{33}	c_{44}
AlN	Tsubouchi	345	125	120	395	118
	Mc Neil	411	149	99	389	125
	Deger	410	140	100	390	120
	Kim (th)	398	140	127	382	96
	Wright (th)	396	137	108	373	116
GaN	Polian	390	145	106	398	105
	Schwarz	377	160	114	209	81.4
	Deger	370	145	110	390	90
	Yamaguchi	365	135	114	381	109
	Kim (th)	396	144	100	392	91
	Wright (th)	367	135	103	405	95
InN	Kim (th)	271	124	94	200	46
	Wright (th)	233	115	92	224	48

Table II-2: Sets of elastic constants for GaN, InN and AlN crystals, all values are in GPa. The theoretical data are denoted as (th).

Bibliography

- [1] S. Nakamura and G. Fasol. “The Blue Laser Diode”, Springer, Berlin, (1997).
- [2] S. Nakamura, M. Senoh, N. Iwasa, S. Nagahama, T. Yamada, and T. Mukai, *Jpn. J. Appl. Phys.* **34**, L1332 (1995).
- [3] O. Ambacher, Growth and applications of group III-nitrides, *J. Appl. Phys.* **31**, 2653 (1998).
- [4] P. J. Dobson, A. Howie and U. Valdre, Plenum Press: New York. pp. 159–193 (1988).
- [5] W. Braun, Applied reflection high energy electron diffraction during crystal growth, Springer, Berlin, Heidelberg (1999).
- [6] J. M. van Hove, C. S. Lent, P. R. Pukite and P. I. Cohen, *J. Vac. Sci. Technol.* **B 1(3)**, 741 (1983).
- [7] N. Grandjean and J. Massies, *Appl. Phys. Lett.* **73**, 1260 (1998).
- [8] H. P. Maruska and J. J. Tietjen, *Appl. Phys. Lett.* **15**, 327 (1969).
- [9] W. M. Yim, E. J. Stofko, P. J. Zanzucchi et al, *J. Appl. Phys.* **44**, 292 (1973).
- [10] H. Schultz and K. H. Thiemann, *Solid State Com.* **23**, 815 (1975).
- [11] S. Strite and H. Morkoc, *J. Vac. Sci. Technol.* **B 10(4)**, 1237 (1992).
- [12] L. E. McNeil, M. Grimsditch and R. H. French, *J. Amer. Ceram. Soc.* **76(5)**, 1132 (1993).
- [13] K. Kim, W. R. L. Lambrecht and B. Segall, *Phys. Rev. B* **53(24)**, 16310 (1996).
- [14] A. F. Wright, *J. Appl. Phys.* **82(6)**, 2833 (1997).
- [15] A. Polian, M. Grimsditch and I. Grzegory, *J. Appl. Phys.* **79(6)**, 3343 (1997).
- [16] R. B. Schwarz, K. Khachaturyan and E. R. Weber, *Appl. Phys. Lett.* **70(9)**, 1122 (1997).
- [17] M. Yamaguchi, T. Yagi, T. Azuta et al, *Phys. Cond. Mat.* **9(1)**, 241 (1997).
- [18] C. Deger, E. Born, H. Angerer, *Appl. Phys. Lett.* **72(19)**, 2400 (1998).
- [19] D. Gerthsen, E. Hahn, B. Neubauer, V. Potin, A. Rosenauer and M. Schowalter, *phys. Stat. Sol. (c)* **0**, 1668 (2003).
- [20] E. Sarigiannidou, E. Monroy, B. Daudin, and J. L. Rouvière, *Appl. Phys. Lett.* **87**, 203112 (2005).
- [21] H. P. D. Schenk, M. Nemoz, M. Korytov, P. Vennéguès, A. D. Dräger, and A. Hangleiter, *Appl. Phys. Lett.* **93**, 081116 (2008).

Chapter III ADAPTATION OF HIGH-RESOLUTION TRANSMISSION ELECTRON MICROSCOPY FOR THE STUDY OF III-NITRIDE HETEROSTRUCTURES

The third Chapter is devoted to the adaptation of the high-resolution TEM imaging technique to the study of III-nitride heterostructures. We will start by determining a numerical relationship between the composition of a GaN-based alloy epilayer and its local lattice parameter. We will examine the influence of the difference in elastic constants and of the uncertainty in the TEM sample surface relaxation on the precision of the alloy composition measurements.

Other sources of inaccuracy in the composition determination are inherent errors related to the strain measurement techniques and to the presence of noise in HRTEM images. First, we will put aside the experimental effects and we will examine the artefacts introduced by the strain measurement techniques by applying them to specially created lattice images. Then, the ability of both techniques for the treatment of noisy images is evaluated.

Finally, the influence of the image conditions on the precision of the strain measurements is examined. HRTEM images will be simulated with imaging parameters appropriate for a JEOL 2010F and a Cs-corrected Titan 80-300 microscopes. The role of the sample thickness and of the objective lens defocus is evaluated.

3.1. Relation between III-nitride alloy composition and its local lattice parameter

We will start this Chapter by solving the problem announced above: the determination of the alloy composition by measuring its local lattice parameters. The calculations are based on the data and formulas stated in Section 2.3.2. The layer deformation is always calculated relative to the templates. As before, the layers are supposed to be continuous and atomically flat, which is true for epitaxially-grown samples. The absence of any plastic relaxations is assumed as well.

We consider a thin layer of a ternary alloy $A_xB_{1-x}N$ with a chemical composition x which is deposited on a GaN template. First, using the Vegard's law, the alloy's lattice parameters a and c are expressed as a linear function of x . The elastic coefficients C_{ij} are also supposed to change linearly with the composition x . Then, using equation II-6 (biaxial stress approximation), the relative extension/contraction ϵ_{zz} is calculated for the lattice mismatch ϵ_{xx} between the film and the template.

The calculations were performed for the principal GaN-based alloys (InGaN, AlGaN and AlInN). The sets of elastic constants reported by Wright and Kim (Table II-2) were used to illustrate the inaccuracy in the composition determination due to the difference between the published elastic data (Figure III-1).

For the all three alloys the strain ϵ_{zz} changes monotonically with the concentration x , i. e. every ϵ_{zz} is correlated with a unique value of x . Thereby the alloy composition x can be unambiguously found if ϵ_{zz} is known.

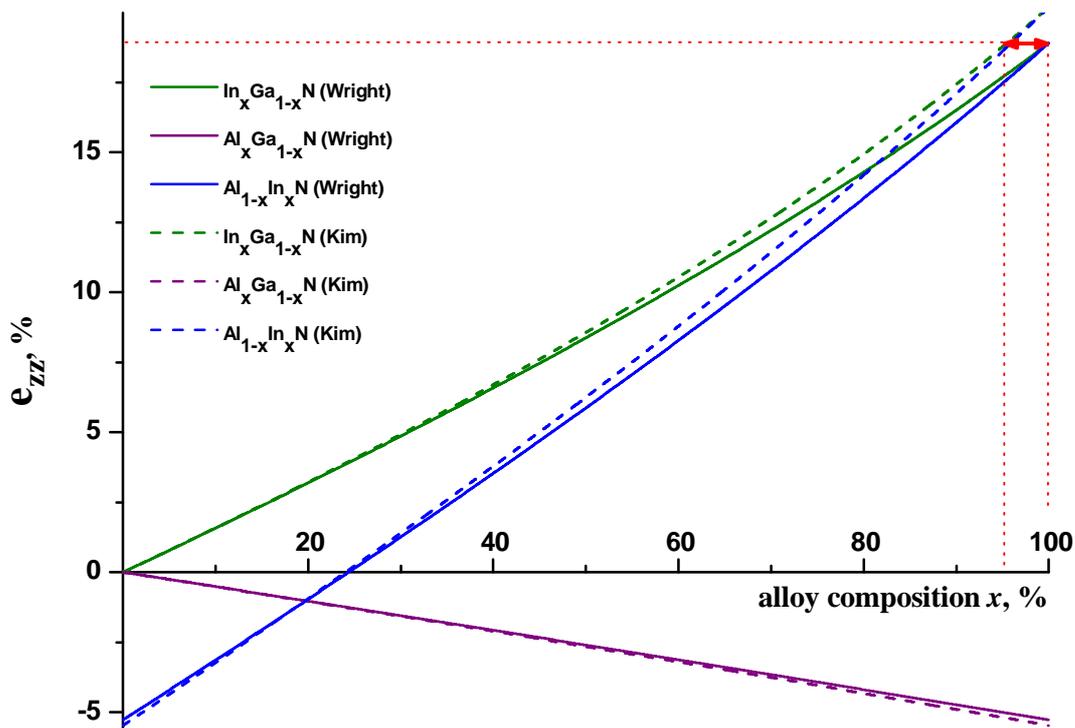


Figure III-1: Biaxial deformation of a thin film as a function of the alloy composition. The red dot lines and the arrow show the incertitude in the composition determination because of the differences between the sets of published elastic constants.

Anyhow the difference between the published elastic coefficients results in some variation of the calculated values of ε_{zz} . This variation is more noticeable in the case of InGaN and AlInN alloys. ε_{zz} of InN strained on the GaN template calculated using Wright's data is 18.9%, whereas ε_{zz} using Kim's coefficients is 20.2%.

Such a difference gives an inaccuracy in the alloy composition determination. Using different elastic coefficients a given strain value ε_{zz} can be related with different alloy compositions x (red arrow on Figure III-1). For example, a strain $\varepsilon_{zz}=18.9\%$ in $\text{In}_x\text{Ga}_{1-x}\text{N}$ corresponds to pure InN when using Wright's parameters; whereas the same strain interpreted using Kim's data gives only 95.5% of In.

Fortunately, this effect is less pronounced for smaller x . For instance, ε_{zz} of 4.02% corresponds to In concentrations of 25% (Wright) and 24.76% (Kim). Thereby, the relative mismatch in the composition determination is below 1%.

Similar calculations were made for a thin $\text{Al}_x\text{Ga}_{1-x}\text{N}$ layer grown on an $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template. The strain was plotted for biaxial (eq. III-6) and uniaxial (eq. III-7) stresses as well as for an unstressed sample (Figure III-2).

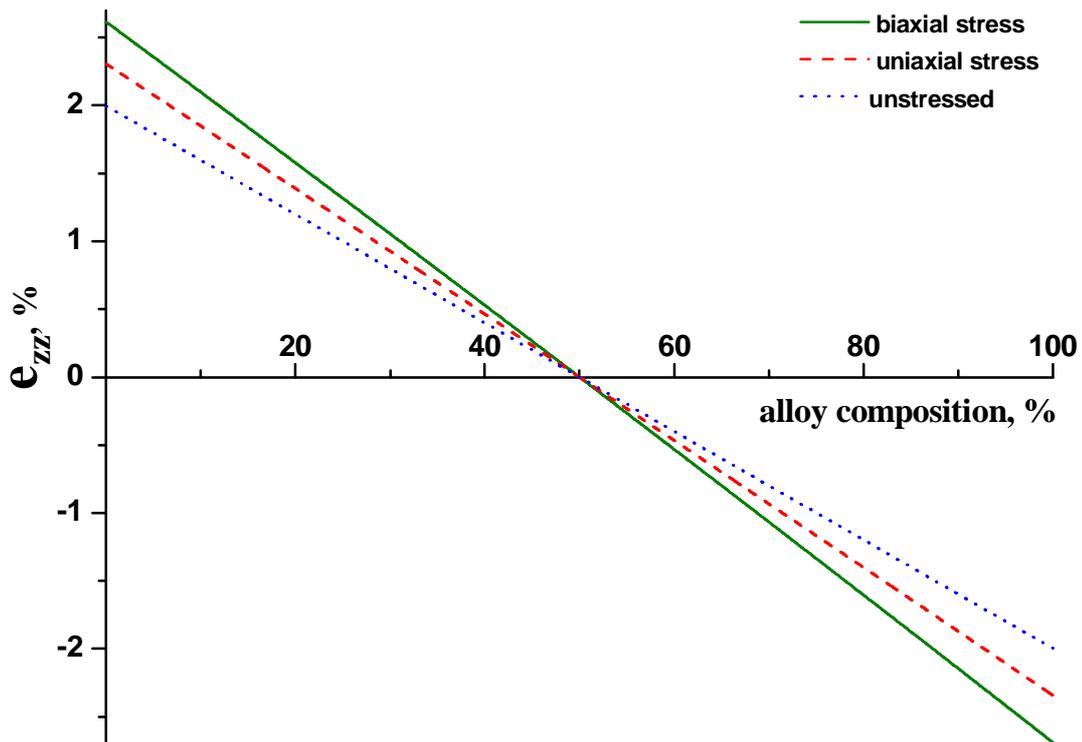


Figure III-2: Strain of a thin alloy film as a function of its composition. The blue dot line shows $\text{Al}_x\text{Ga}_{1-x}\text{N}$ lattice extension in the unstressed layer (lattice parameter c change only due to the Vegard's law).

We see that the biaxial stress amplifies the strain value comparing to that in the unstressed case. Anyhow surface relaxation effects (i.e. uniaxial stresses approximation) reduce the total amount of the strain as compared to the biaxial stress. The difference $\epsilon_{zz}^{\text{biax}} - \epsilon_{zz}^{\text{monoax}}$ increases proportionally with the deviation of x from 50%, which is the composition of the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template.

The uncertainty in the relaxation state of the TEM sample has the same effect as the difference of the elastic coefficients discussed above. The strain ϵ_{zz} variations are maximal at the opposite ends of the composition axis. For $x=0\%$ the biaxial stress produces $\epsilon_{zz}=2.61\%$ and the uniaxial one gives only $\epsilon_{zz}=2.31\%$. Whereas strain $\epsilon_{zz}=2.31\%$ corresponds to the Al concentration of 6% in the biaxial approximation. For $x=100\%$, $\epsilon_{zz}^{\text{biax}}$ is -2.69% and $\epsilon_{zz}^{\text{monoax}}$ is -2.34% . In turn, strain $\epsilon_{zz}=-2.34\%$, in the biaxial approximation, gives 93.6% of Al.

Thereby the influence of the thin foil relaxation effects is much stronger than that due to the variation of the elastic coefficients. Indeed, even for moderate concentration $x=70\%$, the difference in the stress modes can produce a mistake of 2.5% in the alloy composition. However, as the sample thickness increases, the deviation from the biaxial approximation decreases [1].

In this work both (uniaxial and biaxial) approximation were used for epilayers composition evaluation. Further precision improvements can be achieved using Finite Element Calculations [2] method, as mentioned in Section 2.3.2.

It is noteworthy that for $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ the strain variations (from -2.7 to 2.6%) are less pronounced than in the InGaN/GaN system (Figure III-1). However, $\text{In}_x\text{Ga}_{1-x}\text{N}$ with 15% - 25% of indium produces amounts of strain from 2.4% to 4%. Thus, in both cases relatively weak deformations will be measured. This requires a high accuracy in the strain determination. For instance, an inaccuracy of 5 pm in the InGaN interplanar distances measurement will introduce an error of 2% in the ϵ_{zz} determination.

3.2. Evaluation of the strain measurement techniques

The accuracy of the composition evaluation is strongly governed by the precision in the strain determination. The strain measurement precision mainly depends on the microscope stability and the TEM sample quality. Another source of inaccuracy in the strain determination is inherent errors of the strain measurement technique. In this thesis strain was measured using two techniques: the geometric phase analysis (GPA) and the projection method (PM), which were introduced in Section 1.2.3.

Here we will examine the performance of these methods in two cases, which are of particular interest for semiconductor heterostructure study: strain measurements at sharp interfaces and noisy images treatment. For doing this, several sets of lattice images representing actual HRTEM images were created (Figure III-3).

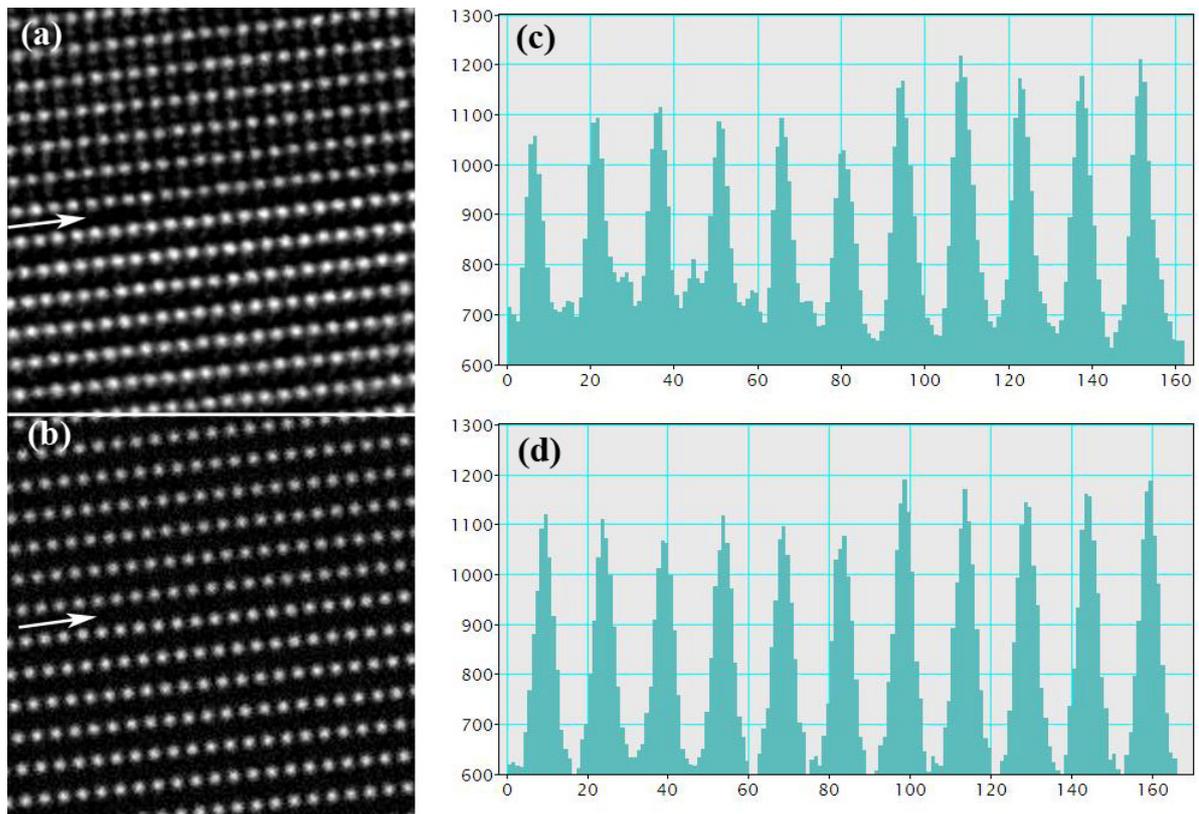


Figure III-3: Actual HRTEM image of a GaN/AlGaN interface taken along the $[1\bar{1}00]$ direction (a). Artificially created image used for strain measurement techniques evaluation (b). The white arrows show the interfaces. Single profiles traced perpendicularly to the interfaces of actual (c) and created (d) images.

A special script for Digital Micrograph was developed to generate lattice images with required parameters. The created images were similar to the projection of the wurtzite lattice along the $[1\bar{1}00]$ direction (Figure II-4 (a)). The intensity maxima were simulated by Gaussian functions. A sharp interface in the heterostructure was reproduced. Lattice parameters and image contrast could be set separately above and below the interface.

The lower part was a reference region, where the ratio of lattice parameters a and c was 1.626 (as in GaN). In the upper part, the lattice constant a was the same as below the interface, whereas the lattice parameter c could be changed, setting thereby a required amount of strain. The image pixel size was 0.017 nm, which corresponds to 14.7 pixels by period in the $[0001]$ direction, in which the deformation was measured.

The contrast of the lattice images (i. e. the amplitudes of the Gaussian functions) was matched to the actual image contrast. As we saw in Section 1.1.3, high-resolution contrast is always superposed on the mass-thickness contrast, which is due to the TEM sample thickness changes and composition variations. If the TEM sample is accurately prepared, the influence of thickness variation may be neglected for small fields of view. But composition variation always causes change of the HRTEM image contrast at the heterostructure interfaces. Besides, Fresnel fringes appear in the HRTEM image there, where the inner potential of the sample changes abruptly. The intensity of the Fresnel fringes varies with the microscope defocus value, additionally modifying the image contrast [3].

To represent this effect the lattice image contrast was set separately for the upper and lower parts of the interface. The contrast ratio is defined as the fraction of the upper part contrast to the lower part one.

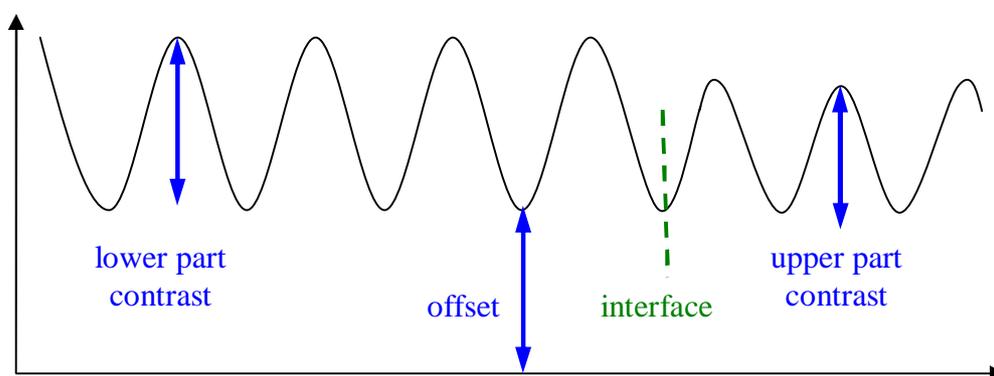


Figure III-4: Schematic profile of artificially created image. The green line shows the interface, the blue arrows define image contrast and offset parameters.

The intensity of HRTEM images acquired with Gatan CCD camera is about 1000 counts (Figure III-3 (c)). This is considered as an optimal value, since an increase of the exposure time or of the illumination intensity leads to the image degradation. An offset value was added to the whole image in order to adjust the amount of total intensity with the actual image (Figure III-4). Artificially created images were 2048x2048 pixels in size.

The GPA and PM methods are used in the standard ways described in Section 1.2.3. The zones for the strain reference are always chosen in the lower part of the interface. All profiles are traced from the reference zone to the strained area. By default the profiles are not averaged, otherwise it is specially indicated. The abscissa axis in all profiles is in pixels, the ordinate axis is either HRTEM image intensity, or amount of the lattice deformation.

3.2.1. Strain distribution at the interfaces

To perform strain measurements at sharp interfaces a lattice image with a strain of 3% and a contrast ratio of 1.05 was created. The PM treatment was realised by fitting fourth degree polynomials using seven points in each peak (Figure III-5 (a)). The strain profile showed an “atomically flat” interface between the two lattices (Figure III-5 (b)). The standard deviation of ϵ_{zz} is less than $2 \cdot 10^{-5}$ on both sides of the interface.

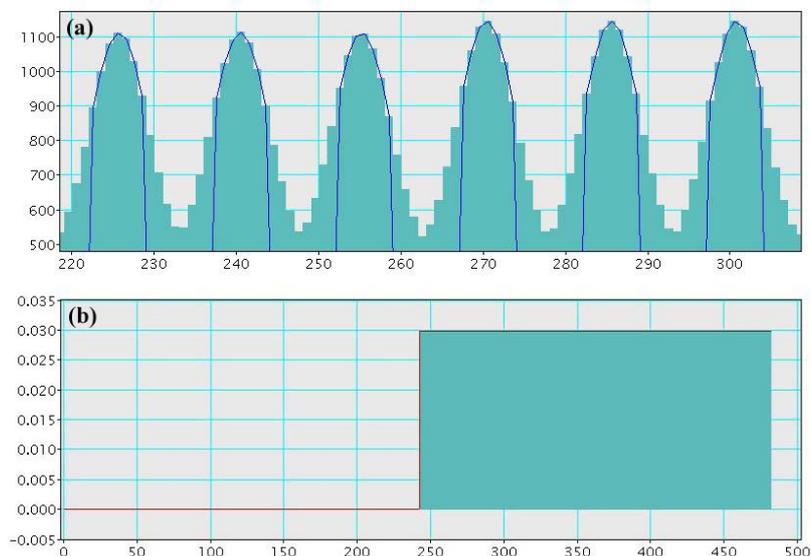


Figure III-5: Intensity maxima (a) of the treated profile are fitted by polynomials (blue lines). The strain plot obtained by PM (b) shows a sharp change of the lattice parameter.

Then the same image was treated by GPA. A cosine mask of a large size (1/30 pixels in diameter) was centred on the (0002) reflection. The deformation field was computed by derivation of the phase image. A profile was traced on the strain map perpendicular to the interface (Figure III-6).

GPA shows smooth strain change between two lattices. The interface width, measured between the levels of 0.03 and 0.27 (blue arrow on Figure III-6), is 18 pixels. Spatial resolution reduction is a typical feature of GPA, caused by the data averaging due to the use of finite mask size [4].

The strain far from the interface is quite uniform with a standard deviation $6 \cdot 10^{-5}$ for the lower part and $4 \cdot 10^{-5}$ for the upper part. But the strain fluctuations at the interface are much higher: $-1.1 \cdot 10^{-3}$ on the left side and $2.8 \cdot 10^{-3}$ on the right side (shown by the red arrows on Figure III-6).

To understand the origin of these artefacts another interface without deformation but with a contrast ratio of 1.05 was examined by both techniques. Only noises of order $2 \cdot 10^{-5}$ were revealed in the profile produced by PM (Figure III-7 (a)).

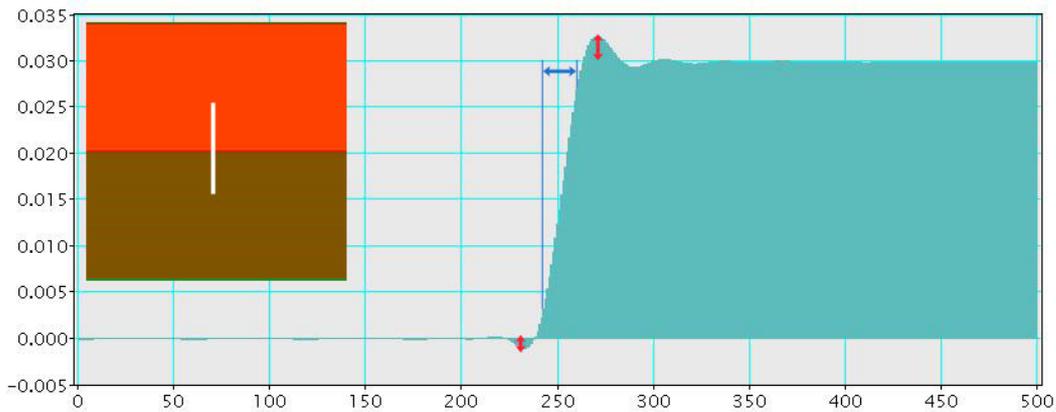


Figure III-6: Profile of the strain map done by GPA. The insert shows the colour-coded ϵ_{zz} strain map where the actual profile is marked by the white line. The blue arrow shows the interface broadening; the red arrows indicate the strain fluctuation at the interface.

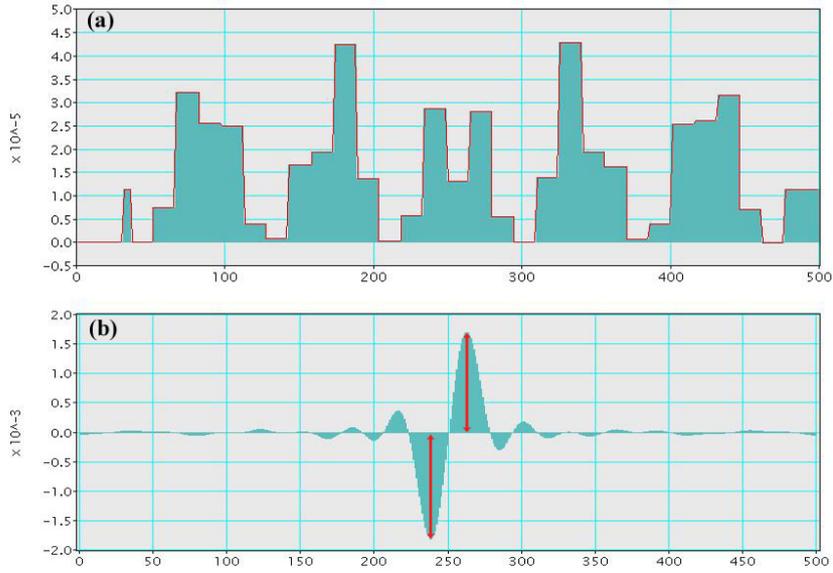


Figure III-7: The ε_{zz} strain profiles of a non-deformed lattice obtained by PM (a) and GPA (b), the red arrows indicate the strain fluctuation. Note that the ordinate axis scales differ in two orders.

At the same time the GPA profile showed strain fluctuations centred at the interface (Figure III-7 (b)). The fluctuations are inversely symmetric regarding the interface position. The amplitude of the highest fluctuations (shown by reds arrows) is $1.8 \cdot 10^{-3}$. The full width at half maximum of the largest fluctuation is 18 pixels; the total artefact size is about 180 pixels.

The same image treated with a mask of medium size (1/45 pixels in diameter) exhibits strain deviations up to $7.9 \cdot 10^{-4}$. The use of a small mask (1/60 pixels) reduces the fluctuation amplitude to $4.4 \cdot 10^{-4}$. If the contrast ratio is < 1 then all fluctuations change sign; whereas the fluctuations shapes and amplitudes remain the same.

A series of images with different contrast ratio (from 1.005 to 2) was then treated by GPA. The amount of deviation rises when the amplitude ratio increases. We found an empirical approximated relation between the contrast ratio and the amplitude of the highest fluctuation.

$$A_{fl} \approx (C-1) * 0.036 \quad (\text{eq. III-1})$$

where A_{fl} is the highest fluctuation amplitude and C is the contrast ratio.

A similar effect was observed in images with a contrast ratio = 1, where the offset was set separately for the lower and the upper parts of the image (Figure III-8).

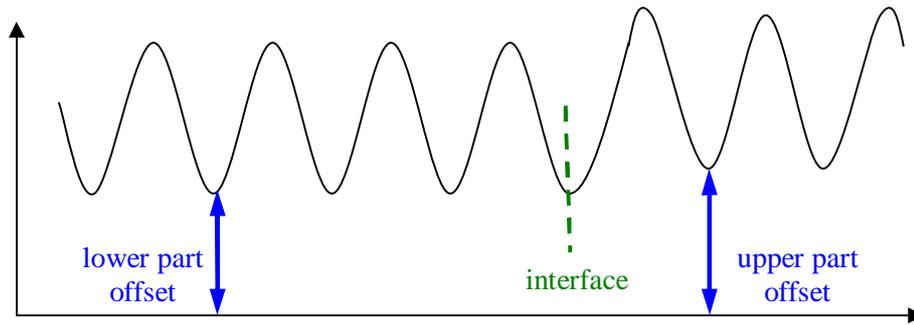


Figure III-8: Schematic profile of image used for test described below. The green line shows the interface, the blue arrows indicate change of the offset parameter.

Once again, fluctuations appeared at the interface of the GPA map, whereas PM did not reveal any essential strain deviation. The highest fluctuation amplitude can be estimated as

$$A_{fl} \approx \Delta I * 0.2 \quad (\text{eq. III-2})$$

where A_{fl} is the highest fluctuation amplitude and ΔI is the ratio of image intensity maxima on the upper and lower parts of the interface.

The following explanation is proposed to the appearance of these strain fluctuations. A sharp contrast change at an interface can be characterized by a step function. The Fourier coefficients of a step function are proportional to $1/\omega$, where ω is a spatial frequency (Figure III-9).

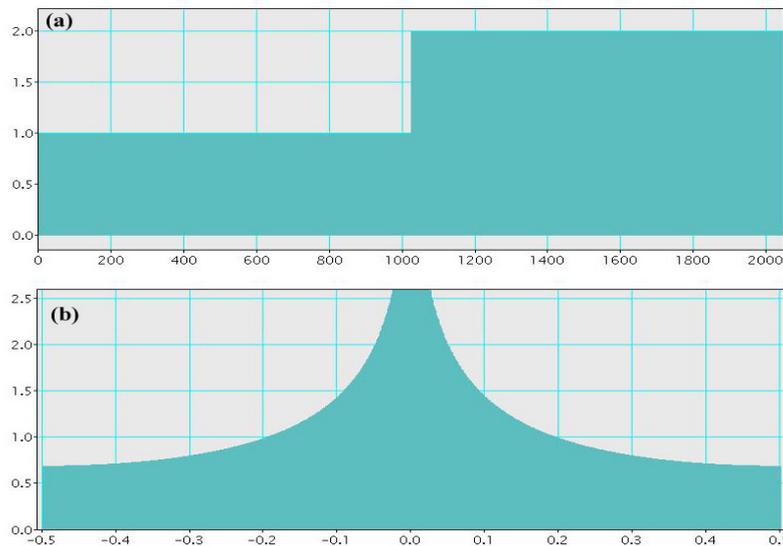


Figure III-9: The step function (a) and its Fourier transform (b).

The Fourier transform (FT) of lattice images with a contrast ratio different from 1 (Figure III-10 (b)) is a FT of lattice images with contrast ratio of 1 (Figure III-10 (a)) plus the FT of a step function.

Thus, when an image containing a contrast ratio different from 1 is treated by GPA, a part of the step function intensity passes through the mask (Figure III-10 (b)) and contributes to the phase image formation. An artificial phase shift is thereby created at the interface. When the contrast ratio rises, the Fourier coefficients of the step function increase as well. And a higher amount of artificial contribution comes to the phase image.

After that an opposite interface with 3% of deformation and with no contrast change was examined. As before, the PM reveals a sharp change of the lattice parameter (not shown). Nevertheless, the strain map obtained by GPA shows a strain deviation $\pm 1.0 \cdot 10^{-3}$ on the sides of the interface (Figure III-11). Once again, the fluctuations are inversely symmetric regarding the interface position.

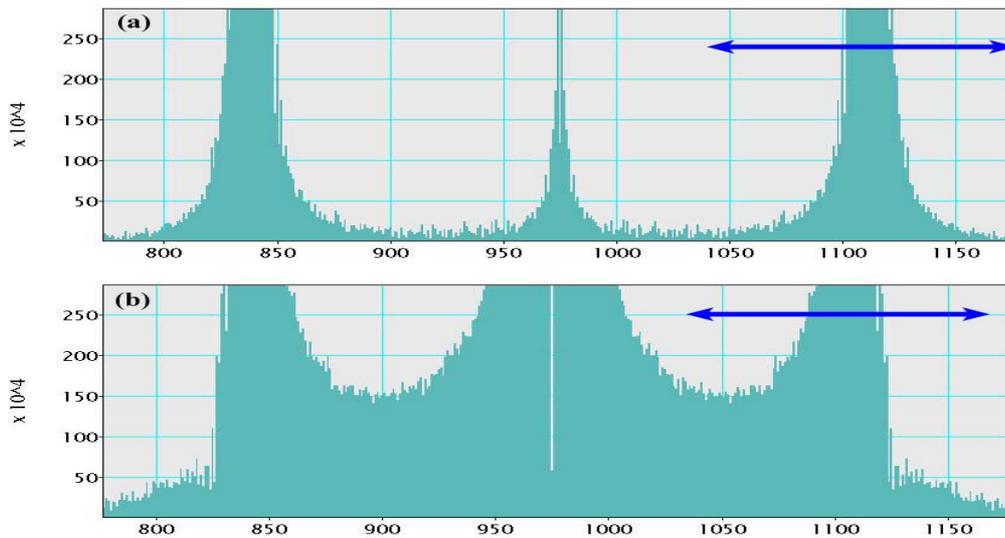


Figure III-10: Profiles of the power spectrum of an interface with a constant contrast (a) and of an interface with a contrast ratio 1.1 (b). The sizes of the masks used for the GPA strain map calculation are indicated by the blue arrows.

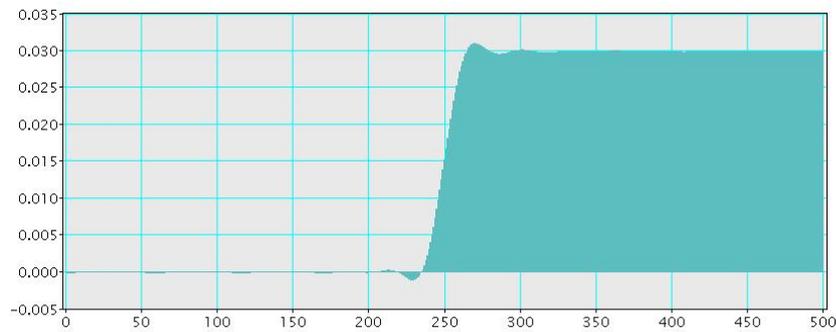


Figure III-11: Profile of the GPA strain map for the lattice with a uniform contrast.

The same image treated with a mask of medium size exhibits strain deviations up to $7.8 \cdot 10^{-4}$. The use of a small mask reduces the fluctuation amplitude to $6.9 \cdot 10^{-4}$. Similarly to the contrast effect, the strain fluctuations invert the sign for negative deformations. A series of images with different value of strain (from 0.3% to 30%) was treated by GPA. The amount of fluctuations increases almost linearly with the strain. An empirical rule is:

$$A_{fl} \approx D \cdot 0.035 \quad (\text{eq. III-3})$$

where A_{fl} is the highest fluctuation amplitude and D is the amount of strain.

Similar sets of lattice images were treated using another software, which also realises geometric phase methodology (not-shown). The treatment of lattice images with a contrast ratio different from 1 also produces strain fluctuations at the interface (Figure III-7 (b)). Whereas no rippled appearances similar to that shown on Figure III-11 were revealed at the interfaces of strain map obtained from the lattice images with uniform contrast. This signifies that this artefact type is mainly due to the particularity of the computation algorithm.

So, a change of any of the lattice characteristics (either contrast, or periodicity) produces fluctuations in the strain maps achieved by GPA. In most cases, the deviation is much weaker than the strain. Anyhow, for very small lattice deformations the fluctuations due to the contrast change can dominate the measured strain (Figure III-12).

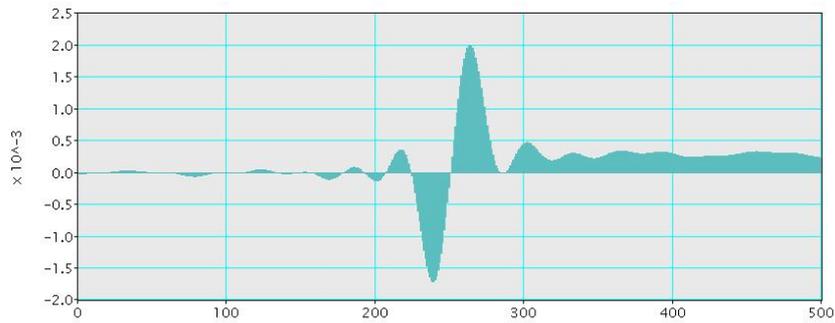


Figure III-12 ϵ_{zz} profile obtained by GPA for a lattice with 0.03% of deformation.

The strain profile shown on Figure III-12 corresponds to a 1.05 of contrast ratio in the presence of 0.03% of deformation. Note that despite the fluctuations at the interface a strain as small as 0.03% can be measured. In the case of InGaN/GaN layer, such a strain corresponds to an In concentration of about 0.2%. In practice, detection precision is governed by the noises in HRTEM images.

3.2.2. Noisy images treatment

Actual HRTEM images always contain noise, i.e. fluctuations of the image intensity. The strain measurement techniques may transform these intensity fluctuations in a certain amount of fringes displacement. Thereby, some strain fluctuations may be observed even in perfect crystals with invariable lattice parameter.

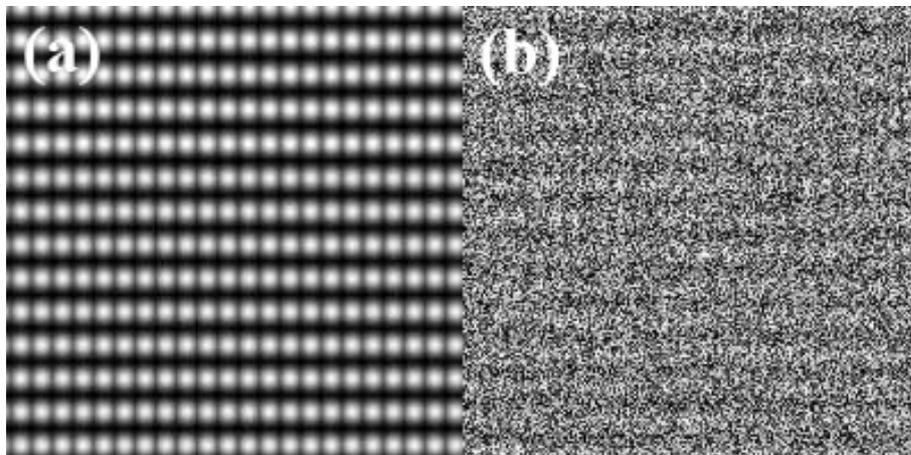


Figure III-13: The lattice images with a signal/noise ratio of 20 (a) and 0.2 (b).

To test the performance of GPA and PM in noise data processing, a series of lattice images containing different amounts of random noise was created. The initial image represented a structure with 3% of deformation and with a contrast (i. e. the signal value) of 600 counts for the lower and 630 counts for the upper parts of the interface. A random noise was added to every image pixel by summing it with an item

$$\text{noise} = (\text{RND} - 0.5) * \text{NA} \quad (\text{eq. III-4})$$

where RND is a random number in the range [0,1] and NA - is an integer determining the noise amplitude. Images with noise amplitude from 30 to 3000 were generated. Thus, the signal/noise ratio was varied from 20 to 0.2. The lattice images with the extreme amounts of noise are presented in Figure III-13.

First, the images were treated by GPA. As before a cosine mask with a large size was placed on the (0002) reflection. Examples of strain maps for the lowest and the highest signal/noise ratio are given in Figure III-14.

The noise influence was analysed on single profiles traced perpendicular to the interface. For every profile a standard deviation of ε_{zz} is calculated in the strained part of the lattice (Figure III-15). The zone for analysis was chosen quite far from the interface to avoid the effects discussed in the previous Section (Figure III-15).

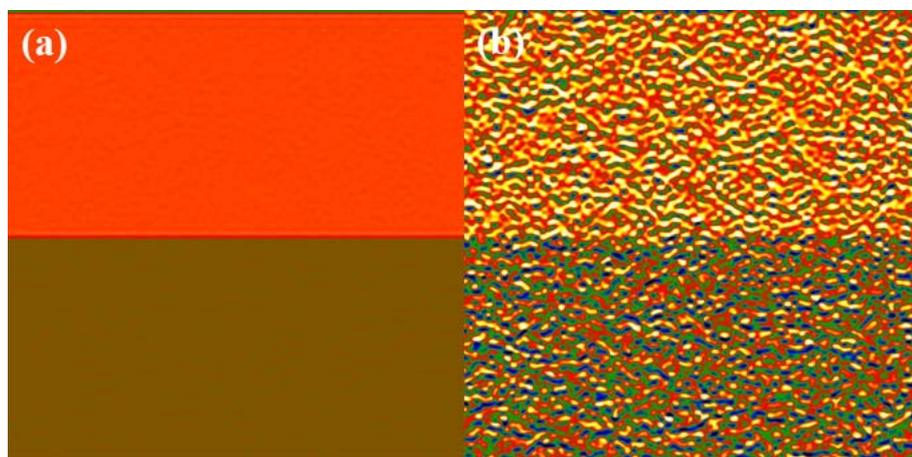


Figure III-14: Strain maps for the lattice images with a signal/noise ratio of 20 (a) and 0.2 (b).

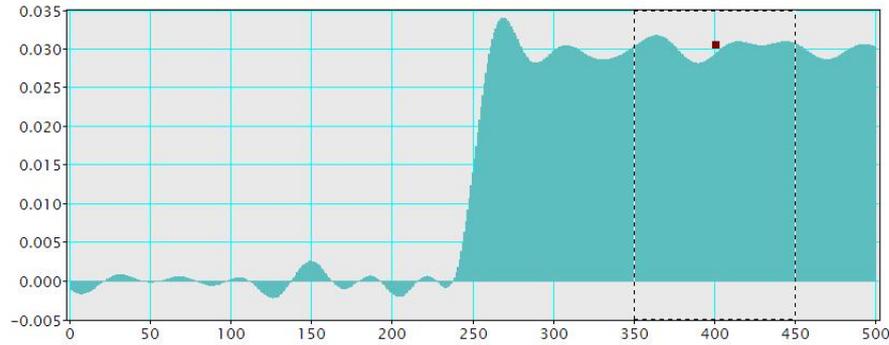


Figure III-15: GPA strain profile of lattice image with a signal/noise ratio of 6. Dashed lines show the zone used for the strain deviation calculations.

Averaged profiles were analysed as well. Indeed, since the spatial resolution of the strain map is limited by the mask size (30 pixels in this example) a single profile has the same resolution as a profile averaged by 30 lines. For GPA, the strain was averaged by 30, 150 and 750 lines. In the case of the GaN lattice this corresponds to a spatial resolution of 0.5 nm, 2.5 nm and 7.5 nm respectively. The resultants of GPA testing are given in comparison with PM (Figure III-20).

The projection method required a manual adaptation for the noisy data treatment. The first problem was that the fourth degree polynomials fit too punctually all feature of the noisy profile (Figure III-16). Because of this, the polynomial maxima do not correspond to the exact peaks' position. Consequently, the strain measurement precision drastically decreases.

This drawback was eliminated by using second order polynomials, where maxima positions were unambiguously determined. But even this approach fails when the signal/noise became < 1 (Figure III-17).

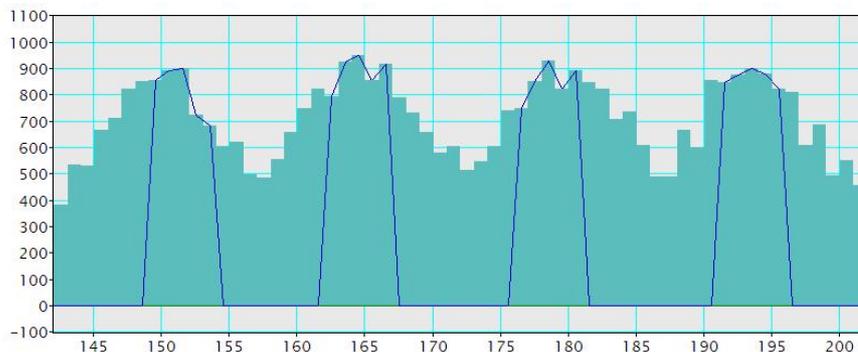


Figure III-16: Fitting of a noisy profile with fourth order polynomials.

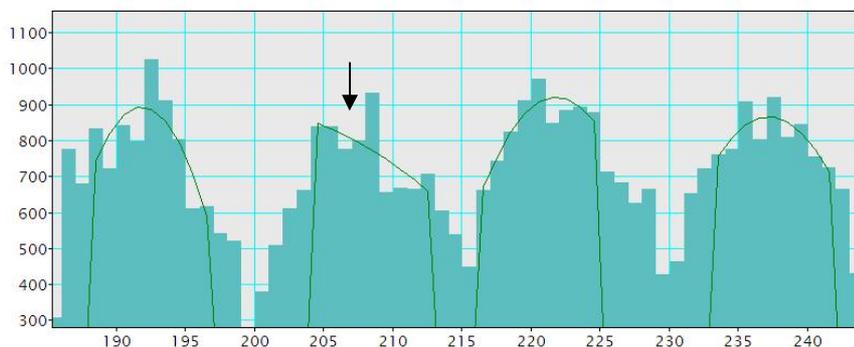


Figure III-17: Fitting of second-degree polynomials to a noisy profile by nine points. Black arrow mark pixel of maximum intensity.

The fact is that the script fits the polynomials using points situated around peak intensity maximum. In noisy profiles, the intensity of some pixel (marked, for example, by the black arrow on Figure III-17) can exceed the intensity of the central pixel of the peak. The polynomials fitted around the false position will approximate only one side of the peak, which will certainly give a wrong refined maximum position.

This problem can be solved by increasing the number of pixels used for fitting (Figure III-18). However, this number is limited by the pixel size (i.e. microscope magnification) and generally all suitable pixels are already used for fitting.

An alternative solution was realized within the PM script. The improved algorithm verifies itself the fitting quality of every peak. It compares position of refined maximum position with initial position of peak intensity maximum, which was used as a starting-point for fitting. If the difference is bigger than one pixel, than the script changes the starting-point and refits a polynomial. Then the refined maximum position is again verified.

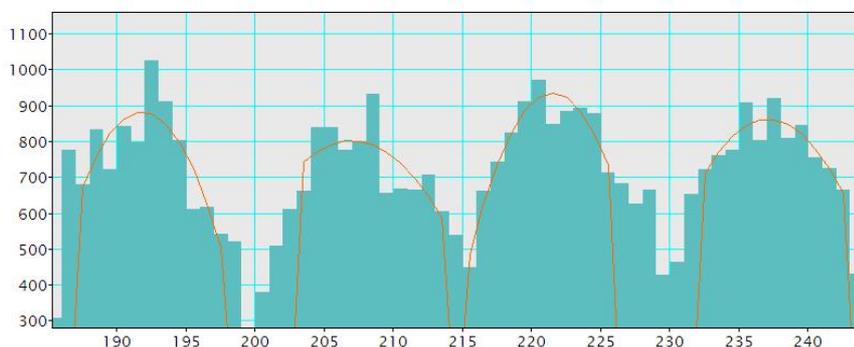


Figure III-18: Fitting of second-degree polynomials to the same profile using eleven points.

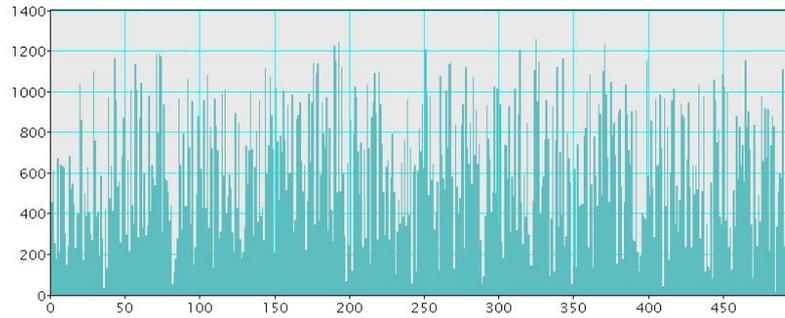


Figure III-19: Single profile of the lattice image with a signal/noise ratio = 0.6.

During the application of PM the polynomial degree and number of points used for fitting were manually adjusted to minimize the standard deviation. The empirical principle is to use for fitting all suitable points. In this example (14.7 pixels by period), the fitting with 11 points gave the best precision. Certainly, the higher the degree of the polynomial is, the better is the fitting accuracy. However, beyond a certain amount of noise, only parabolic polynomials worked sufficiently well (Figure III-18). But if the noise increases further (Figure III-19) the algorithm fails even to detect the peaks.

Before PM application, the treated profiles were averaged by 5, 30 and 150 lines. For GaN lattice this corresponds to spatial resolution of 0.083 nm 0.5 nm and 2.5 nm respectively. Further averaging increases measurements precision, but the loss of spatial resolution renders the use of PM irrelevant, since GPA gives better precision for the same spatial resolution. The standard deviation of the strain was measured in the same zone as for GPA.

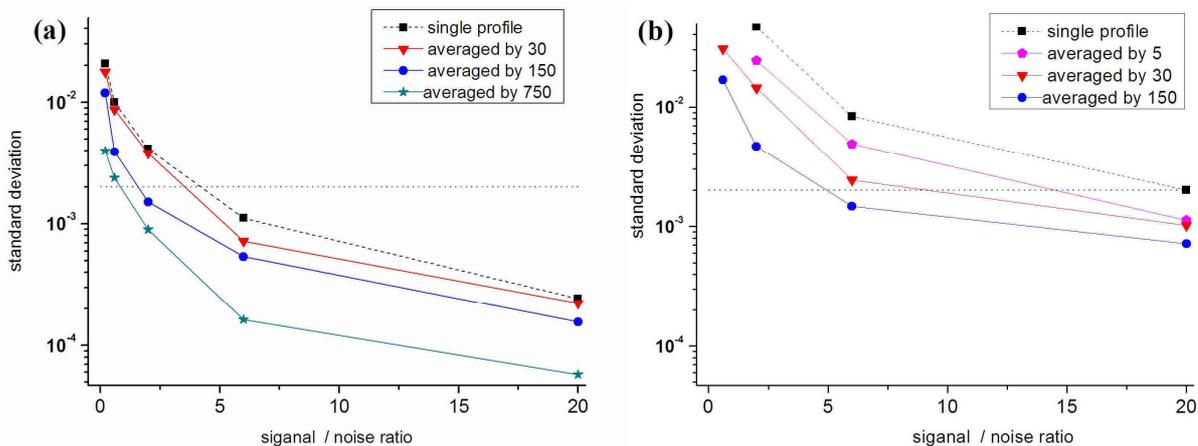


Figure III-20: Deviation of the ε_{zz} strain as a function of the signal/noise ratio for GPA (a) and PM (b). Dot line shows level of acceptability = $2 \cdot 10^{-3}$.

For both techniques the deviation of ϵ_{zz} increases with the noise and decreases with the profile averaging (Figure III-20). GPA showed a superior ability in noise images treatment even using a mask of large size. This method is capable to calculate strain even from completely degraded images, such as the one shown on Figure III-13 (b). The corresponding strain map (Figure III-14 (b)) can be interpreted, at least qualitatively. However, the strain deviation can be essentially reduced by profile averaging over a large area (Figure III-20 a).

Such an excellent behaviour is partly due to the statistical properties of the random noise. In GPA the precision of strain measurement is mainly governed by the accuracy of the reciprocal lattice vectors determination, which is done in the Fourier space. Fourier transform (FT) of the noise image equals to the FT of initial, noise-free image plus the FT of the noise. The random noise has a uniform spectral density, thus only certain amount of noise is added to every component of the HRTEM image FT (Figure III-21).

Even for the image shown on Figure III-13 (b) the intensity of the reflection used for strain calculation is much stronger than the amount of added noise (Figure III-21 (b)). Especially since in the examined case the strained area occupies half of the image surface. Thus, the reflection corresponding to the deformed area is quite intense. Thereby, the form and the intensity of the reference and signal peaks (shown by red arrows on Figure III-21 (b)) are not seriously modified comparing to the ones of noise-free image (Figure III-21 (a)).

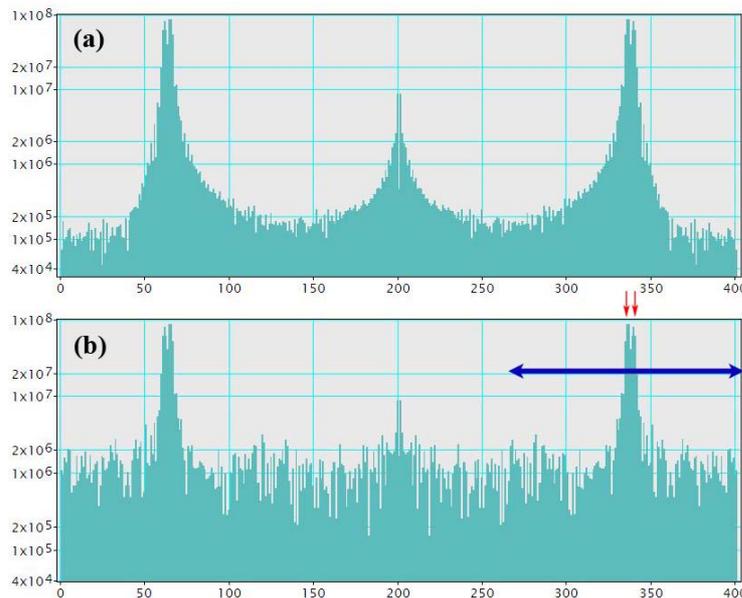


Figure III-21: Profiles of the power spectrum of a noise-free image and (a) and one with a signal/noise ratio = 0.2 (b). The blue arrow shows the size of the mask used for the GPA strain map calculation. The red arrows mark the reflections related to the deformed and reference areas.

Besides, applying the mask suppresses the noise beyond the chosen reflection. Only noise adjacent to the mask centre is transferred to the strain map. The smaller the mask size, the better the noise reduction.

The developed PM script does not use any noise filtration in order not to lose the spatial resolution of the strain. The precision of the strain measurements is mainly governed by the agreement of the fitted function maxima with the actual peaks position. The fitting of polynomials in a noise-free profile works as the development of each peak intensity as a Taylor's series around peak maximum. Consequently, the position of the polynomial maximum strictly corresponds to the position of the peak maximum, which explains the low strain deviation in noise-free images (Figure III-5 (b)).

But, an addition of a small amount of noise even to a single pixel will break this approximation. The fitted polynomial will reproduce the actual profile as precisely as possible, but the position of its maximum will not correspond any more to the peak maximum position. The strain standard deviation will drastically increase with noise (Figure III-20 b).

We should note a property of the random noise, which makes a positive influence on the PM and GPA performance. The point is that the mean value of the random noise is almost zero even in a small area. Thereby profile averaging turns to be quite effective for noise reduction. However, the actual noise is not so uniform, and its spectral density can depend on the spatial frequency. Thus, the efficiency of strain measurement techniques applied to an actual HRTEM image can be poorer than in these tests.

However, in Section 3.1 we saw that an error in the composition determination is directly related to the strain measurement precision. For example, for $\text{In}_x\text{Ga}_{1-x}\text{N}$ with In concentrations above 20%, an inaccuracy of 0.2% in ε_{zz} produces a relative error in the composition determination of less than 1%. In same Section, we saw that an uncertainty in the elastic coefficients or the surface relaxation state may also cause certain inaccuracy in the composition determination. This renders useless an ultra-precise strain measurement. Strain determination with an accuracy of about $2 \cdot 10^{-3}$ is quite sufficient for practical purposes. We will refer to this criterion in the next Section where the validation of the HRTEM image conditions for the quantitative strain analysis will be performed.

3.2.3. Conclusion

The performance and limitations of GPA and PM were estimated by applying these methods to specially created lattice images. In fact, the techniques are complementary; each has preferable fields of applications.

Projection method is able to measure strain with an atomic scale resolution. It is used to investigate the morphology of thin strained layers [5] as well as to examine clustering in random alloys [6]. Projection method is also applied to study strain distribution near the heterostructure interfaces [7] or domain boundaries [8].

However, the accuracy of the peak position determination is directly governed by the number of pixels per period. The use of a pixel size of 0.017 nm (as in the computation above) limits the field of view to 35 nm x 35 nm for a 2k by 2k pixel CCD camera.

GPA is very efficient in noisy images treatment. It shows a better precision in the strain determination than PM for a whole amplitude/noise range (Figure III-20). The noise reduction is governed by the mask diameter. However, the spatial resolution of the strain map is inversely related to the chosen precision. Even in a noise-free image the spatial resolution is limited to the doubled interplanar distance [9].

Besides, GPA can deal with image sampling as low as 5 pixels per period [10]. It means that in GaN areas up to 110 nm large can be analysed.

Another advantage of GPA is that it produces 2D strain map, whereas PM treats images profile by profile. An advanced technique of real space images treatment, called **peak pairs analysis**, maps displacement in two directions [11], but it was not considered in this work.

In total, GPA is better adapted for actual images treatment. No wonder that it has a broad range of applications: measurement of displacement fields around dislocation [12] and strain across ferroelectric domain walls [13], strain mapping in nanowires [14] and strain distribution in quantum-dot superlattices [15].

At the same time, GPA demonstrates unexpected sensitivity to the variations of the image contrast (Figure III-7 (b)). Besides, the change of the lattice periodicity in an image with a uniform contrast also produces some strain fluctuations in the GPA strain maps

(Figure III-11). Generally, the fluctuation amplitude is much smaller than the strain, but for very small lattice deformations the fluctuations can dominate over the measured strain (Figure III-12).

This characteristic leads to an interesting conclusion. The contrast of HRTEM image aimed for quantitative analysis by GPA must be as weak as possible (Figure III-22 (b)). In the ideal case you should see nothing in the field of view. It is quite contrary to the images dedicated to qualitative observation, where the contrast should be maximal (Figure III-22 a). The image contrast, in turn, is determined by the microscope setting. Once again, imaging conditions are of high importance for obtaining reliable high-resolution images.

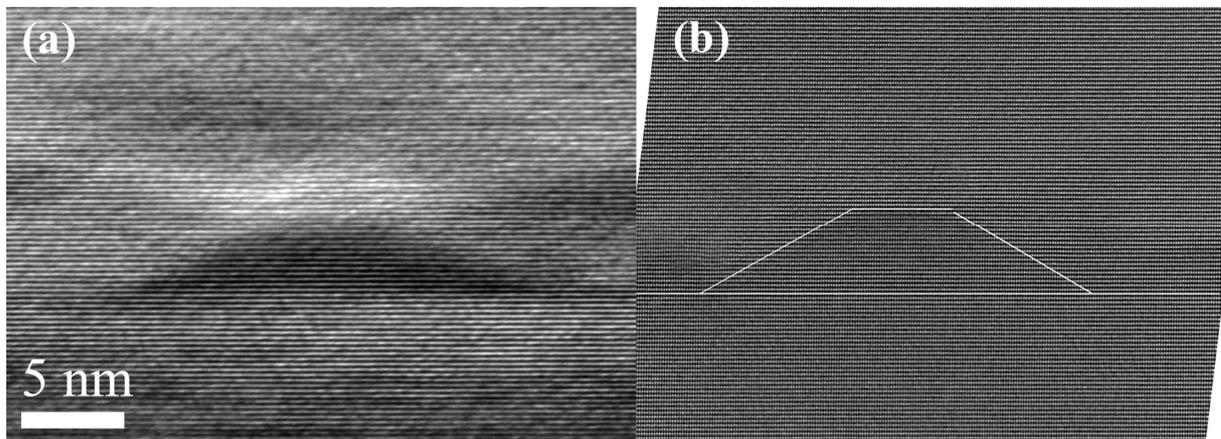


Figure III-22: HRTEM images of a GaN quantum dot in a AlGaN matrix taken with JEOL 2010F (a) and with Titan 80-300 microscope (b). The white lines display the quantum dot contours determined from a strain map (not shown).

3.3. Acquisition conditions for reliable high-resolution imaging

Imaging conditions may have an impact on the displacement field measured from HRTEM images [16]. In particular, considerable strain fluctuations may appear on hetero-interfaces in non-centrosymmetric structures [17, 18].

In this Section, we will analyse the effects of the acquisition conditions on the strain measurement in nitride heterostructures. Namely, we will determine the range of thicknesses and defocuses within which the lattice fringe displacements correspond to the actual displacements of the atomic planes.

Specimen orientations along the $[1\bar{1}00]$ and $[14\bar{5}0]$ zone axes will be compared with the two-beam conditions developed by Rosenauer et al [1] for the study of InGaN/GaN quantum wells. In this configuration, the sample is tilted by 6° around an axis parallel to the $[0001]$ direction relative to the $[11\bar{2}0]$ zone axis, which corresponds to a centre of the Laue circle of $(12, \bar{1}2, 0, L)$. The (0002) beam is centred on the optical axis and a fringe image is formed using the (0000) and (0002) beams. The excitation condition with $L = 2.2$ was found to be optimum for the imaging of $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ heterostructures [1]. A similar excitation condition was used in this study.

HRTEM images simulations are performed with the Java version of Electron Microscopy Software (JEMS) [19] using the multislice approach. The imaging parameters appropriate for the JEOL 2010F and the Cs-corrected Titan 80-300 microscopes (Table III-1) are used.

	TEM voltage (keV) Electron wave- length (pm)	chromatic aberration C_c (mm)	spherical aberration C_3 / C_5 (mm)	energy spread (eV) defocus spread (nm)	point-to-point resolution (nm)	information limit (nm)
JEOL 2010F	200 2.51	1.1	0.5 / 0	0.8 5.5	0.194	0.13
Titan 80-300	300 1.97	2.0	-0.003 / 3	0.7 4.7	0.08	0.08

Table III-1: Characteristic parameters of the microscopes used for the HRTEM image simulations.

The envelope illumination model and the Fourier method of projected potential were used in the simulations. The coefficients C_3 and C_5 were chosen similar to the typical values of spherical aberration of the Cs-corrected Titan microscope. A negative value of C_3 coefficient was aimed to compensate the impact of C_5 , in the same way as the Scherzer defocus compensates the spherical aberration in common microscopes. The beam convergence was 0.5 mrad. The pixel size in the HRTEM images was 0.023 nm, which corresponds to 11 pixels per period of GaN in the [0001] direction.

To obtain the ε_{xx} strain (i.e. in the direction perpendicular to the [0001] axis) distribution in GaN-based materials, the distance between two neighbouring atoms in the basal planes should be resolved. In the [1 $\bar{1}$ 00] projection (Figure II-4 (a)), this distance is $a_1 = 159$ pm (eq. III-3). The point-to-point resolution of the Cs-corrected Titan 80-300 microscope is sufficiently small to visualize such a small spacing.

A scheme of the atomic structure used in the simulations is shown in Figure III-23. The lattice parameters of $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ layers were chosen as a simple average of that of GaN and AlN. Ga and Al atoms with an occupancy of 0.5 are placed at the cation atomic position into AlGaN. The lattice parameters of GaN were calculated in the biaxial approximation: the parameter a was taken similar to that of AlGaN, the parameter c was calculated using the elastic constants reported by Wright (Table II-2).

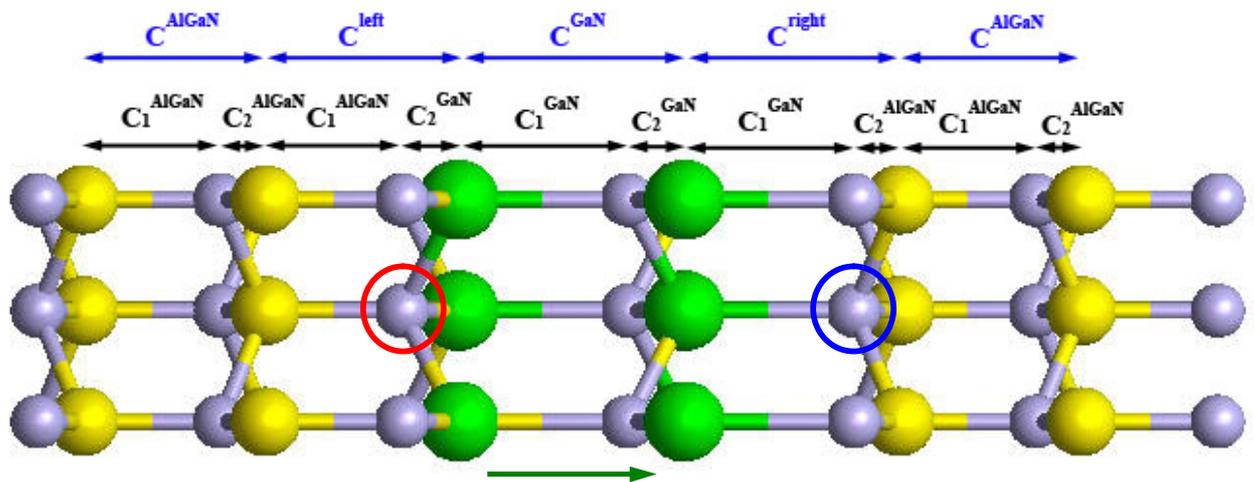


Figure III-23: Atomic structure of an AlGaN/GaN/AlGaN interface projected along the [1 $\bar{1}$ 00] zone axis. The grey spheres represent the N atoms, the green ones the Ga atoms and the yellow ones $\text{Al}_{0.5}/\text{Ga}_{0.5}$, the black arrows show interatomic distances, the blue arrows indicate the distances resolved by TEM. The green arrow shows the growth direction for a Ga-polar crystal.

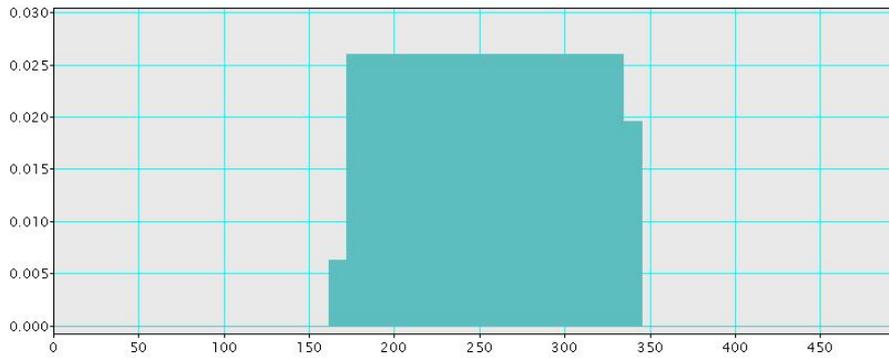


Figure III-24: Introduce strain distribution at the AlGaIn/GaN/AlGaIn interface for the structure depicted in Figure III-23. The abscissa axis corresponds to the coordinate in the growth direction (in pixels); the ordinate axis is the amount of deformation.

For a N atom lying in the AlGaIn/GaN hetero-interface (shown by the red circle), the length of the bond with a cation in AlGaIn is defined as C_1^{AlGaIn} , whereas the bond length with a Ga atom is C_2^{GaIn} . The interatomic distances for a N atom lying in the GaIn/AlGaIn interface (shown by the blue circle) are defined in the following way: C_1^{GaIn} with a Ga atom in GaIn and C_2^{AlGaIn} with a cation in AlGaIn.

The distance c_2 (64 pm for GaIn) lies beyond the point resolution of the used microscopes, thereby, two adjacent atomic columns are represented by one spot in the HRTEM images. Distances visualised in the HRTEM images are indicated by the blue arrows on Figure III-23. C_1^{AlGaIn} is 254.18 pm, whereas C_1^{GaIn} , calculated considering a biaxial lattice distortion of the GaIn layer, is 260.82 pm, which corresponds to a ε_{zz} strain of 2.61%. The interatomic distances at the interfaces are the following: $C^{\text{left}} = 255.81$ pm and $C^{\text{right}} = 259.19$ pm, which is 0.64% and 1.97% higher than the lattice parameter c of AlGaIn. Thereby, the defined interface is slightly asymmetric (Figure III-24).

Simulated images were analyzed by geometric phase analysis in a similar way as the images obtained experimentally. A large size mask ($g/2$) was placed around the (0002) reflection in the reciprocal space. The phase images were normalised with a reference zone chosen within the AlGaIn layer.

HRTEM image strongly varies with the TEM sample thickness and the objective lens defocus. First, the impact of the TEM sample thickness on the strain determined from HRTEM images taken along the $[1\bar{1}00]$ direction was highlighted for two extreme cases of

thin (Figure III-25 (a)) and thick (Figure III-26 (a)) samples. The image parameters are appropriate for the Cs-corrected Titan microscope; the objective lens defocus value was set to 6 nm.

The GPA profile obtained from a 2.2 nm thick sample accurately represents the expected strain distribution of the atomic model (Figure III-25 (c)). The strain is smaller at the left interface comparing to the right one, due to the asymmetry of the input strain profile (Figure III-24). However, small strain fluctuations are present at the interfaces: the maximal amount of deformation averaged within the GaN layer is $2.6\% \pm 0.4\%$. The amount of deformation averaged within the GaN layer is $2.6\% \pm 0.4\%$. However, small strain fluctuations are present at the interfaces: the maximal amount of deformation in the GPA profile is 2.81%, the minimal is -0.13%.

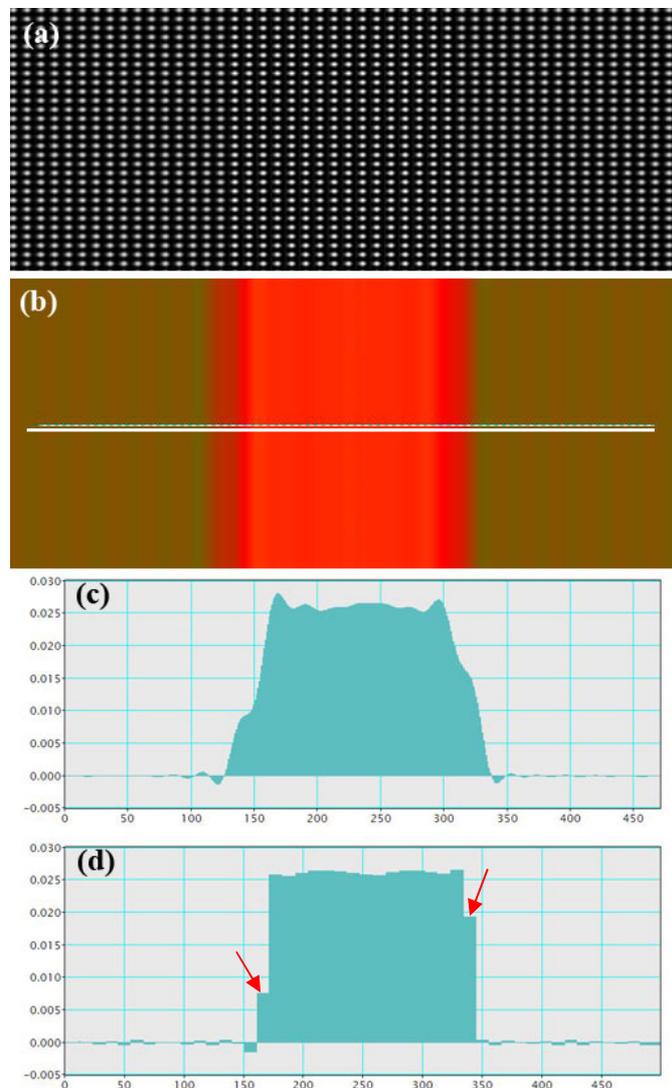


Figure III-25: Simulated HRTEM image along the $[1\bar{1}00]$ direction of a 2 nm thick sample taken with a defocus of 6 nm (a). The strain map obtained with GPA (b). The white line shows the direction of the profile (c). The strain profile calculated with the projection method (d).

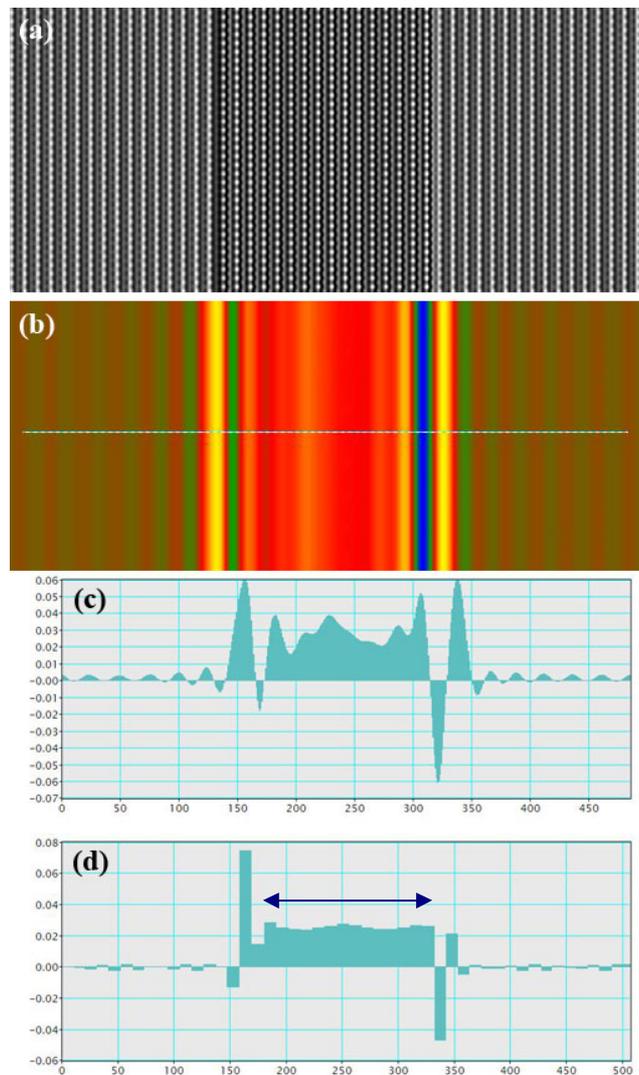


Figure III-26: A simulated HRTEM image (a) of a 35 nm thick sample taken with similar condition as the one shown on Figure III-25. The strain map obtained with GPA (b). The blue line shows the direction of the profile (c). The strain profile calculated with the projection method (d), the blue arrow shows the flat part of the profile used for strain measurements.

The strain profile calculated using the projection method for comparison (Figure III-25 (d)) is in perfect agreement with the expected strain distribution (Figure III-24). The mean ε_{zz} strain measured in the GaN layer is $2.61\% \pm 0.3\%$, the deformations at the left and right interfaces, 0.76% and 1.94% respectively, even corresponds to the initial strain asymmetry. The minimal and maximal amounts of deformation in this profile are -0.14% and 2.66% respectively.

The HRTEM image calculated with identical conditions for a 35 nm thick sample (Figure III-26 (a)) noticeably differs from the previous one. The GaN layer is clearly seen against the AlGaIn environment. The GPA strain profile (Figure III-26 (c)) exhibits numerous

strain fluctuations at both interfaces, as well as within the GaN layer. In fact, a GaN layer, thinner than 4 nm, may not be detected in the strain map due to such artefacts.

In the previous Section, we saw that GPA induces an inherent error if the HRTEM image contrast changes at interfaces (Figure III-7 (b)), whereas the projection method works correctly in this case (Figure III-7 (a)). The strain profile obtained from the HRTEM image (Figure III-26 (a)) with the projection method (Figure III-26 (d)) showed that only a few interatomic distances are distorted at the interfaces, whereas the lattice inside the GaN layer is not deformed. The averaged strain within the flat part of the profile (shown by blue arrow) is $2.597 \% \pm 0.13\%$, which is very close to the expected value (2.61%).

Once again, the projection method gives more reliable results than GPA if the HRTEM image contrast changes. However, due to the advantages in actual images treatment discussed above (Section 3.2), the strain determination with GPA was largely used in this thesis, whereas the projection method was only applied for solving particular problems. Thereby, HRTEM image conditions were adapted for the correct strain determination by GPA.

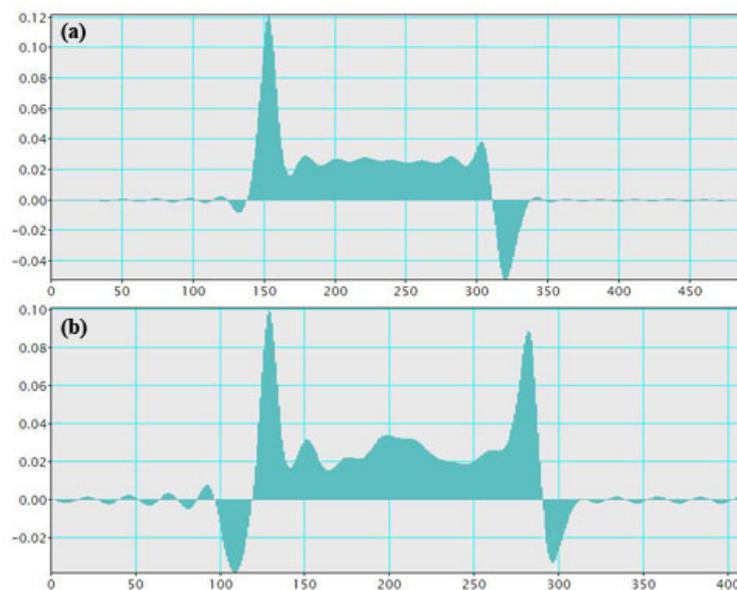


Figure III-27: GPA strain profiles obtained from simulated HRTEM images of a 35 nm thick sample taken with microscope defocuses of -9 nm (a) and -18 nm (b).

Second, to illustrate the influence of the microscope defocus on the strain measurements, a series of images of 35 nm thick sample were simulated with a defocus varying from -27 nm to 36 nm. Flat strain profiles (Figure III-27 (a)), similar to that achieved with the projection method (Figure III-26 (d)) were obtained in a narrow range (-9 nm ÷ -6 nm) of defocus. Though high strain fluctuations at the interfaces do not hamper a precise strain measurement in the flat part of the profile ($2.62\% \pm 0.14\%$), it is hard to distinguish between a reliable strain and artefacts in the actual strain map.

Moreover, the strain profile degrades rapidly with the change of defocus: a slight decrease of the defocus results in considerable increases of the strain deviation in the GaN layer (Figure III-27 (b)). As a result, no defocus values giving a satisfactory strain profile were found for a sample thickness of 35 nm.

So, the aim of the following simulations is to determine, for each sample thickness, the defocus ranges in which the deviation of the strain profile obtained with GPA from the expected values (2.61% in GaN and 0% in AlGa_{0.5}N) is smaller than 0.6%. This rather arbitrary criterion was determined empirically, as discussed below.

This problem is solved in the following sequence:

1. The atomic structure of a 4 nm thick GaN layer embedded between two 4 nm thick Al_{0.5}Ga_{0.5}N layers is created (Figure III-23). The thickness of the GaN layer corresponds to a typical QD height, whereas the AlGa_{0.5}N layers were chosen sufficiently thick to avoid edge effects. The defined strain distribution is shown on Figure III-24.
2. The HRTEM images of the created structure are simulated for given values of sample thicknesses and objective lens defocusses using the parameters of the microscopes employed in this work (Table III-1).
3. Strain maps are calculated from the HRTEM images using the GPA method. The thickness and defocus values are considered to be correct if the measured strain profile lies within the limits of tolerance.

The criterion is illustrated by the treatment of a focal series taken along the [110] zone axis with a 5 nm thick sample (Figure III-28).

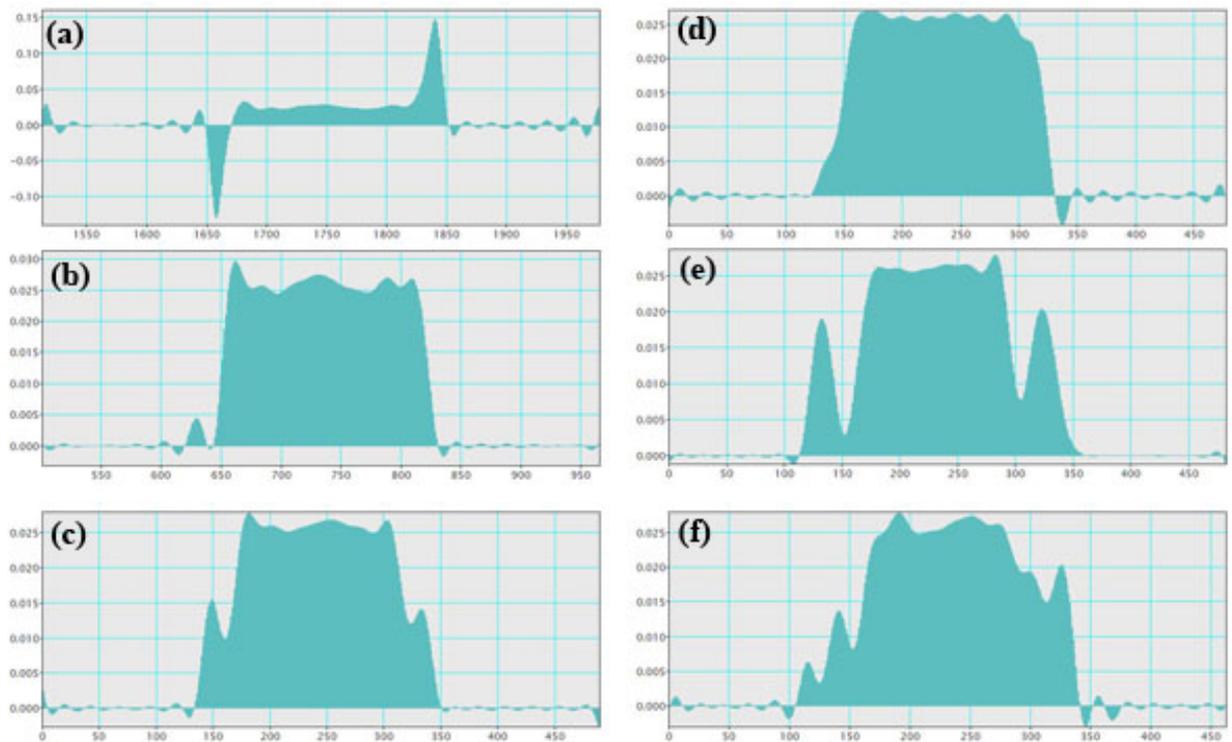


Figure III-28: GPA strain profiles obtained from simulated HRTEM images of a 5 nm thick sample with defocuses of 0 nm (a), 6 nm (b), 18 nm (c), 27 nm (d), 48 nm (e), and 60 nm (f).

The strain profile from the image acquired at a defocus zero exhibits high fluctuations at the interfaces (Figure III-28(a)). However, a slight defocusing of 6 nm makes the strain profile much more reliable (Figure III-28(b)). As the defocus value increases, small strain fluctuations remain at both interfaces (Figure III-28(c)). At a defocus value of 27 nm, the strain profile becomes again regular (Figure III-28(d)), but strong fluctuations (not shown) arise at higher defocus values.

In the defocus range between 48 nm (Figure III-28(e)) and 60 nm (Figure III-28(f)), the fluctuations in the strain profile becomes smaller. The averaged amount of strain measured in the regular strain part is within only 0.4% of deviation from the expected value. However, the flat part of the profile is almost two times narrower than the original profile. Besides, the strain behaviour at interfaces is far from the actual strain profile. Thus, we do not consider the focus range (48 nm ÷ 60 nm) as suitable for strain measurements; and accept only defocuses values between 6 nm and 27 nm. A similar treatment of a focal series for a 5 nm thick sample acquired with negative defocusing gave two ranges of confidence: -51 nm ÷ -39 nm and -12 nm ÷ -3 nm.

From the above observations, we defined a combined criterion which was used in the further analyses. Firstly, the amount of strain measured in the GaN layer should not deviate more than 0.6% from the expected value of 2.61%. Secondly, the amplitude of the strain fluctuations in AlGaN should be smaller than 0.6%. The amount of maximum deviation (0.6%) was chosen empirically, considering strain fluctuations which are present even in strain maps obtained from the thin samples (Figure III-25 (c)).

Additional claim is that the width of the regular part of the profile should be higher than 80% of the initial GaN layer thickness, which guarantees the absence of small but lengthy fluctuations at HRTEM image interfaces. We can not expect better reproducibility of strain profile width, since GPA always decreases the width of the sharp strain profiles, due to the strain smoothing (Figure III-6).

For a 10 nm thick sample, the HRTEM images taken along the $[1\bar{1}00]$ direction correctly reproduce the input strain profile only in a narrow defocus range: $-12\text{ nm} \div -6\text{ nm}$. For a 15 nm thick sample, appropriate defocus values were not found. Similar computations were performed for samples oriented along the $[14\bar{5}0]$ zone axis. Strain profile rapidly change with the microscope defocus, only certain ranges of defocus are suitable for quantitative strain analysis (Table III-2).

Thereby, the HRTEM imaging of AlGaN/GaN/AlGaN interface along the $[1\bar{1}00]$ and $[14\bar{5}0]$ zone axis works correctly only for specimens thinner than 15 nm. On the other hand, the two-beam conditions are reliable in a much larger range of sample thicknesses (Figure III-29). Correct strain profiles were obtained even for samples thicker than 40 nm.

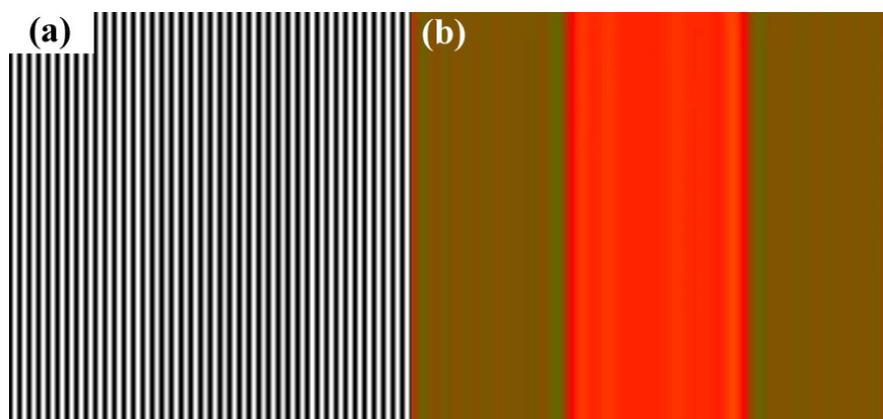


Figure III-29: Simulated fringe image of a 32.5 nm thick sample taken with a zero defocus in the two-beam conditions (a) and corresponding GPA strain map (b).

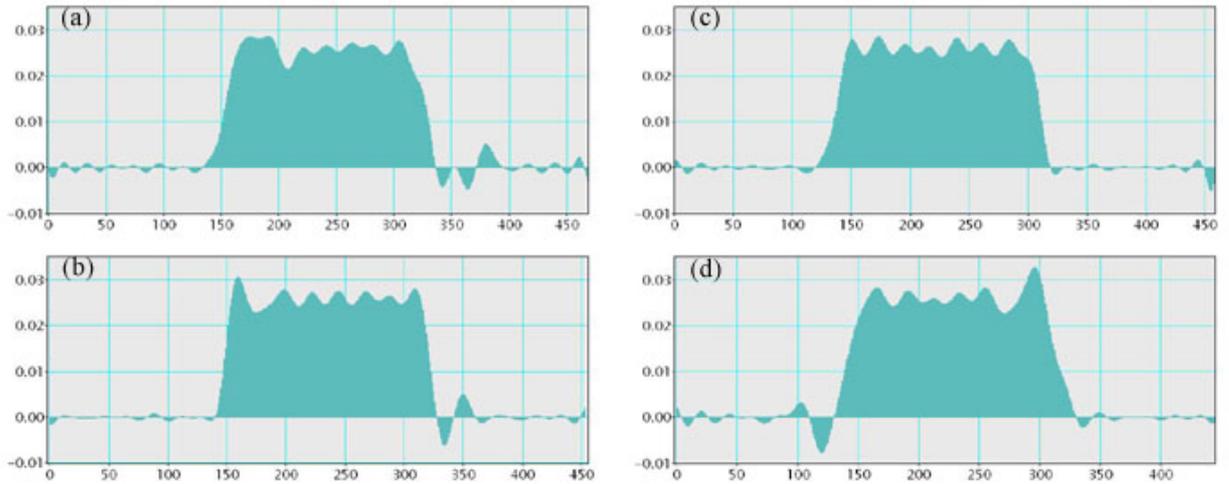


Figure III-30: Series of strain profiles determined from simulated off-axis images of a 20 nm thick sample with defocuses of -200 nm (a), -100 nm (b), 0 nm (c), 100 nm (d).

Besides, the impact of the objective lens defocus on the strain determined from HRTEM images (Table III-3) is much smaller than in the case of on-axis imaging (Table III-2). The strain profile does not exhibit any fluctuations at interfaces when defocusing increases, only the strain deviations within the flat part of the profile rises (Figure III-30).

sample orientation sample thickness	[1100]	[1450]	two-beam conditions
	5 nm	6 nm ÷ 27 nm -51 nm ÷ -39 nm -12 nm ÷ -3 nm	-57 nm ÷ -42 nm -27 nm ÷ -3 nm 3 nm ÷ 42 nm
10 nm	-12 nm ÷ -6 nm	-30 nm ÷ -27 nm -12 nm ÷ -6 nm 0 nm ÷ 24 nm	-310 nm ÷ 140 nm
15 nm	no	-3 nm ÷ 12 nm	-280 nm ÷ 120 nm
20 nm	no	no	-260 nm ÷ 90 nm
25 nm	no	no	-230 nm ÷ 70 nm

Table III-2: Defocus ranges valid for the strain measurements in GaN/AlGaIn heterostructures using a Cs-corrected Titan 80-300 microscope.

In fact, the defocus ranges indicated for off-axis conditions are not strictly defined, since no considerable strain profile degradation occurs even very strong defocusing (below -500 nm), but only strain fluctuation at interface spreads laterally.

Analogous simulations were performed using the parameters appropriate for a JEOL 2010F microscope. Defocus ranges, acceptable for quantitative HRTEM imaging for samples of different thickness, are summarized in Table III-3.

In practice, the objective lens defocus and sample thickness may vary within the field of view. Thereby, the ranges of thicknesses and defocuses assigned for quantitative HRTEM imaging should be sufficiently wide. In on-axis conditions, the ranges of confidence for the strain analysis with the JEOL 2010F microscope are quite narrow (Table III-3). In the case of the Cs-corrected Titan, the defocus of $8 \text{ nm} \pm 4 \text{ nm}$ seems to be an optimal value for the sample oriented along the $[14\bar{5}0]$ zone axis. Whereas for the sample observation along the $[1\bar{1}00]$ direction, a negative defocusing about $-9 \text{ nm} \pm 4 \text{ nm}$ is required. Both defocus values are valid only for a sample thickness below 15 nm.

sample orientation sample thickness	$[1\bar{1}00]$	$[14\bar{5}0]$	two-beam conditions
5 nm	-64 nm ÷ -52 nm -43 nm ÷ -30 nm	-67 nm ÷ -58 nm -34 nm ÷ -40 nm -6 nm ÷ -9 nm	-280 nm ÷ 140 nm
10 nm	no	-66 nm ÷ -60 nm -49 nm ÷ -39 nm	-250 nm ÷ 120 nm
15 nm	no	no	-230 nm ÷ 100 nm
20 nm	no	no	-200 nm ÷ 70 nm
25 nm	no	no	-170 nm ÷ 50 nm

Table III-3: Defocus ranges valid for strain measurements in GaN/AlGa_N heterostructures with a JEOL 2010F microscope.

In fact, if specimen thicknesses or microscope defocuses slightly differ from the approved values (Table III-2 and Table III-3), the HRTEM images do not immediately become inappropriate for quantitative treatment. In most cases, the strain profile reproduces more or less accurately the input strain distribution and the averaged amount of strain is close to the expected value. Only increased strain fluctuations at interfaces perturb the obtained GPA strain map. Thereby, the image parameters given above should be considered as an ideal guideline; and they do not indicate that on-axis imaging is unsuitable for quantitative analysis.

However, the role of the sample thickness is critically important for on-axis HRTEM imaging. Indeed, a sample of constant thickness, without any amorphous layer or any contamination, containing an ideal heterostructure with atomically flat interfaces and without any composition variation being observed with a state-of-the-art electron microscope (Cs-free, no astigmatism, noises, vibration, etc) may exhibit a very fluctuant strain map (Figure III-26 (b)). The point is that on-axis HRTEM images are formed by the interference of several (or many) diffracted beams, whose amplitude and/or phase strongly vary with the sample thickness. The contribution of the multiply diffracted beams increases with the thickness, thereby the interference image changes more violently with the defocusing. The variation of the lattice parameter and of the absorption factor strongly affects the HRTEM contrast at hetero-interfaces. In particular, the asymmetry of AlGa_N/Ga_N and Ga_N/AlGa_N interfaces governs the asymmetry of the strain fluctuations, which can be seen, for example, in Figure III-27 (a).

The practical rules for minimizing the GPA strain mapping error for the different image conditions were suggested by Hýtch and Plamann [20]. However, the effect of the lens aberrations on rapidly varying strain measurement can be minimized only in a very narrow range of optimum defocus [18, 21].

On the other hand, fringe images formed in the off-axis conditions are almost independent of the sample thickness and objective lens defocus (Table III-2 and Table III-3). The point is that the interference image contains only one spatial frequency (eq. I-8) which corresponds to the distance between the (0002) planes. A change of the diffracted wave phase due to the microscope defocusing does not have any influence on the fringes periodicity, but may only cause a relative shift of the strain profile [1]. Only the distortion of the spatial envelope E_s (eq. I-13), which occurs at high defocusing, can alter the HRTEM image.

Besides, multiple diffraction is weakened since only few diffracted beams are excited for this sample orientation.

Thereby, the two-beam conditions are essential for the HRTEM imaging with classical microscopes, since only ultra-thin samples can be truthfully examined in on-axis conditions. In the following, HRTEM images taken with the JEOL 2010F microscope in off-axis conditions are used for quantitative analysis. On the other hand, the images dedicated for qualitative observations were acquired along the $[14\bar{5}0]$ zone axis to obtain a strong chemical contrast.

The HRTEM images taken with the Cs-corrected Titan microscope along the $[1\bar{1}00]$ zone axis were obtained with a slightly negative defocus. In some cases, series of images with different defocuses were acquired and then treated by GPA. The zones for observation were always chosen close to the specimen edge, where the sample thickness is minimal and assumed to be less than 10 nm.

In the Third Chapter, high-resolution transmission electron microscopy was adapted for the study of GaN-based heterostructures. The principle of heterostructure composition evaluation is based on the measurement of the local lattice parameter by HRTEM. The interatomic distances are directly related with the composition using the Vegard's law, whereas the epilayer lattice distortion along the growth direction is taken into account analytically.

LLP can be determined either by HRTEM image treatment in real space, or by analysing the spatial frequencies in the reciprocal space. Both approaches were applied to several sets of artificially created lattice images. Real space treatment is able to measure LLP variations at atomic scale, whereas reciprocal space analysis gives a better precision in the strain determination.

HRTEM images simulations were performed to evaluate the effect of the imaging conditions on the displacement measurements. On-axis imaging reliable for strain determination is only possible within narrow ranges of objective lens defocuses and sample thicknesses. On the other hand, the two-beam conditions are almost independent of these parameters.

In the next Chapter, the application of HRTEM for a quantitative study of GaN quantum dots grown on AlGaN templates is presented.

Bibliography

- [1] A. Rosenauer, D. Gerthsen, V. Potin, *Phys. Stat. Sol (a)* **203**, 176 (2006).
- [2] M. M. J. Treacy and J. M. Gibson, *J. Vac. Sci. Technol.* **B4**, 1458 (1986).
- [3] D. R. Clarke, *Ultramicroscopy* **4**, 33 (1979).
- [4] J.L. Rouvière and E. Sarigiannidou, *Ultramicroscopy* **106**, 1 (2005).
- [5] P. H. Jouneau, A. Tardot, G. Feuillet, H. Mariette and J. Cibert, *J. Appl. Phys.*, **75**, 7310 (1994).
- [6] S. Kret, G. Maciejewski, P. Dluzewski, P. Ruterana, N. Grandjean and B. Damilano, *Materials Chemistry and Physics* **81**, 273 (2003).
- [7] M. Lamy, J. Thibault, *Ultramicroscopy* **84**, 101 (2000).
- [8] W. Yu, W. Mader, *Ultramicroscopy*, In Press, (2010).
- [9] J.L. Rouvière and E. Sarigiannidou, *Ultramicroscopy* **106**, 1 (2005).
- [10] F. Hue, M. Hÿtch, H. Bender, F. Houdellier and A. Claverie, *Phys. Rev. Lett.* **100**, 156602 (2008).
- [11] P. L. Galindo, S. Kret, A. M. Sanchez, J.-Y. Laval, A. Yanez et al, *Ultramicroscopy* **107**, 1186 (2007).
- [12] M.J. Hÿtch, J.L. Putaux, J.M. Penisson, *Nature* **423**, 270 (2003).
- [13] M.J. Hÿtch, E. Snoeck, R. Kilaas, *Ultramicroscopy* **74**, 131 (1998).
- [14] J.L. Taraci, M.J. Hÿtch, T. Clement, P. Peralta, M.R. McCartney, J. Drucker and S.T. Picraux, *Nanotechnology* **16**, 2365 (2005).
- [15] E. Sarigiannidou, E. Monroy, B. Daudin, J. L. Rouvière and A. D. Andreev, *Appl. Phys. Lett.* **87**, 203112 (2005).
- [16] P. Bayle, J. Thibault, *Microsc. Microanal. Microstruct.* **8**, 125 (1997).
- [17] K. Tillman, M. Lentzen, R. Rosenfeld, *Ultramicroscopy* **83**, 111 (2000).
- [18] E. Guerrero, P. Galindo, A. Yanez, T. Ben, S. Molina, *Microsc. Microanal.* **13**, 320 (2007).
- [19] P.A. Stadelmann, *Ultramicroscopy* **21**, 131 (1987).
- [20] M.J. Hÿtch, T. Plamann, *Ultramicroscopy* **87**, 199 (2001).

[21] J. Chung, L. Rabenberg, *Ultramicroscopy* **108**, 1595 (2008).

Chapter IV STUDY OF GaN/Al_xGa_{1-x}N QUANTUM DOTS MICROSTRUCTURE

In the Fourth Chapter a detailed study of the capping effects on GaN/AlGa_xN quantum dots (QDs) is presented. A series of samples with varying nominal thicknesses of deposited GaN was realized. The samples were first characterized by RHEED, AFM and SEM techniques; TEM was applied for the examination of surface and buried QD morphologies. Next, the microstructure of the AlGa_xN barriers was investigated by high-resolution TEM, whereas strain analysis of HRTEM images was used to determine local Al-concentrations.

The study of the barrier composition was completed by high-resolution STEM-HAADF and EELS methods which provided independent evaluations of Al concentrations. High-resolution STEM-HAADF imaging was also used for atomic scale examination of QD morphology at different stages of capping with AlGa_xN. HRTEM images done with Cs-corrected microscopes were used to determine the strain distribution around the QDs at the first stages of burying.

The observed phenomena are finally discussed and their possible mechanisms are evaluated as well.

4.1. Introduction

A quantum dot is a semiconductor cluster whose excitons are confined in three spatial dimensions, which causes a fundamental difference between QDs' and bulk material's properties. An electron or hole in a crystal has a continuous energy spectrum, whereas a confined electron may only occupy states within a discrete set of energy [1]. The energy spectrum of confined electron depends on the confining length: larger QDs have lower energy levels which are also more closely spaced.

On the other hand, the wavelength of luminescence is determined by the exciton energy. Thereby, the colour of emitted light can be tuned by changing the QD sizes.

Moreover, the combination of QD planes with properly chosen sizes may give rise to white light emission [2].

The energy spectrum of confined exciton may be also modified by a quantum-confined Stark effect [3]. Inner electric field, existing in polar heterostructures due to spontaneous and piezoelectric polarisations, changes the energy levels and increases the luminescence wavelength (red-shift). Additionally, this internal electric field separates electrons and holes, reducing thereby the recombination probability and the exciton binding energy.

As discussed in the Second Chapter, III-nitride epitaxial structures are generally deposited on highly mismatched substrates such as sapphire, silicon or silicon carbide. Consequently, epitaxial heterostructures contain a large number of dislocations. Excitons, trapped by a dislocation, recombine without light emission, which noticeably reduces the efficiency of light-emitting devices [4]. The carrier localization inside QDs may prevent their migration to dislocations [5]. The QDs are thus frequently used in nitride semiconductors for non-radiative losses suppression [6].

QDs can be obtained by various techniques, such as colloidal synthesis [7] or state-of-the-art lithography [8]. QDs may be also spontaneously formed by MBE or MOVPE techniques on highly mismatched templates via Stranski-Krastanow (SK) growth [9]. SK growth follows a two step process: initially a two-dimensional (2D) thin film is grown in a layer-by-layer mode on a mismatched crystal template. Above a certain thickness of deposited layer, the elastic energy stored in the strained epitaxial film relaxes via the layer reorganization [10]. This reconstruction results in the appearance of coherently strained islands on top of a two-dimensional wetting-layer. After that the growth continues through the QDs enlargement and coalescence.

The minimal epilayer thickness required to initiate the QD nucleation strongly depends on the lattice mismatch between the film and its template, with a greater mismatch leading to a smaller critical thickness [11]. However, QD formation is possible only if the lattice parameters of the epitaxial layer are higher than that of template. Another parameter governing the epilayer reconstruction is the surface energy, which increases due to the QD formation. The SK growth takes place only if the gain of bulk elastic energy is higher than the rise of surface energy [12].

Produced QDs may be further overgrown by a capping layer and a QD formation may be initiated again. In such a way multiple QD planes can be formed. A stacked structure emits brighter light than a single one, since luminescence intensity is directly proportional to the number of emitting QDs. However, the overgrowth may have a strong impact on the QD morphology.

At present, most of GaN QDs are grown on AlN templates, as the lattice mismatch between these materials (2.4%) is quite sufficient to initiate the QD formation. Nevertheless, AlN is an unsuitable material for electro-optical applications, since optoelectronic devices require electrical carrier injection from the barriers, whereas p-type doping of AlN still presents difficulties [13]. On the other hand, Al_{0.5}Ga_{0.5}N alloys can be highly doped up to a hole concentration above 10^{17} cm^{-3} [14] and a free electron concentration about 10^{19} cm^{-3} [15].

A UV light-emitting diode (LED) based on GaN/Al_xGa_{1-x}N QDs ($x \sim 0.1$) was already realised by MOVPE technique [16]. The QD formation was initiated using Si as an antisurfactant [17]. Electroluminescence spectrum acquired at room temperature revealed a peak at 360 nm (3.44 eV) attributed to GaN QDs luminescence.

This motivates further prospecting the use of GaN/AlGaN QDs as a base for high-performance light-emitting devices. The study of such small objects (typical QD height is about 4 nm) requires the application of transmission electron microscopy, a technique combining analytical capacities with high spatial resolution.

4.2. General sample characterization

The studied samples were grown by ammonia assisted MBE on (0001) sapphire substrates; the growth routine was identical for all the samples. First, a buffer composed of a 30 nm thick low temperature GaN layer and a 100 nm thick high temperature (900°C) AlN layer was realised. Then a 1 μm thick $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template was grown at 850°C (Figure IV-7). A GaN was then deposited in a 2D growth mode (800°C) after that the growth was interrupted in order to initiate the QD formation (the mechanism is described below). The resulting GaN QD layer was then covered by a 30 nm thick $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ barrier, after that a new series of QDs was realised. In such a way a stacked structure comprising three GaN QD planes separated by three $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ barriers was realised [18]. A fourth QD plane grown on top of the structure was uncapped for AFM imaging. After that the specimen was rapidly cooled down.

The growth of GaN QDs by ammonia assisted MBE differs from the classical SK mode described above. When a GaN layer is deposited on an $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template, a 2D growth mode continuously takes place. Above a critical thickness, the elastic energy, accumulated in the epitaxial GaN film, relaxes via misfit dislocations introduction.

Nevertheless QD nucleation can be achieved if, after the deposition of a minimum amount of GaN, the growth is interrupted and the ammonia pressure is decreased [2]. QD formation takes a certain time, depending on the GaN layer thickness and on the growth temperature. The evolution of the layer morphology is represented by the RHEED pattern changing (Figure IV-1).

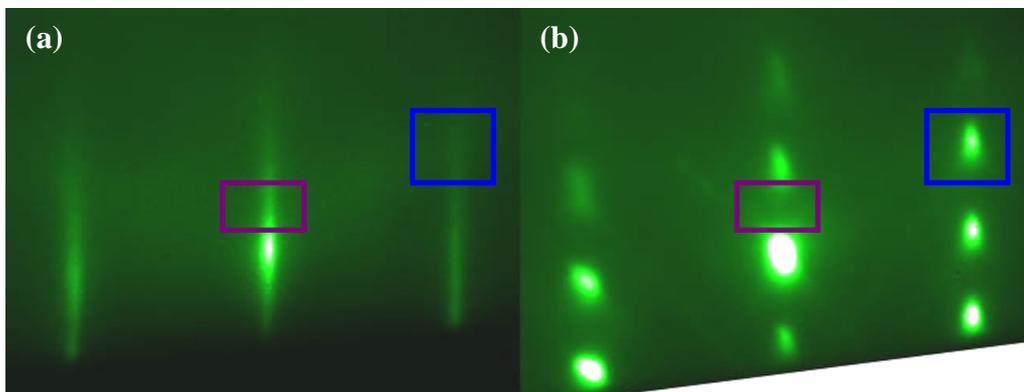


Figure IV-1: A RHEED pattern (beam energy 10 kV) from a GaN layer before (a) and after (b) the QD formation. Coloured rectangles indicate the zones used for RHEED intensity registration (shown on Figure IV-2).

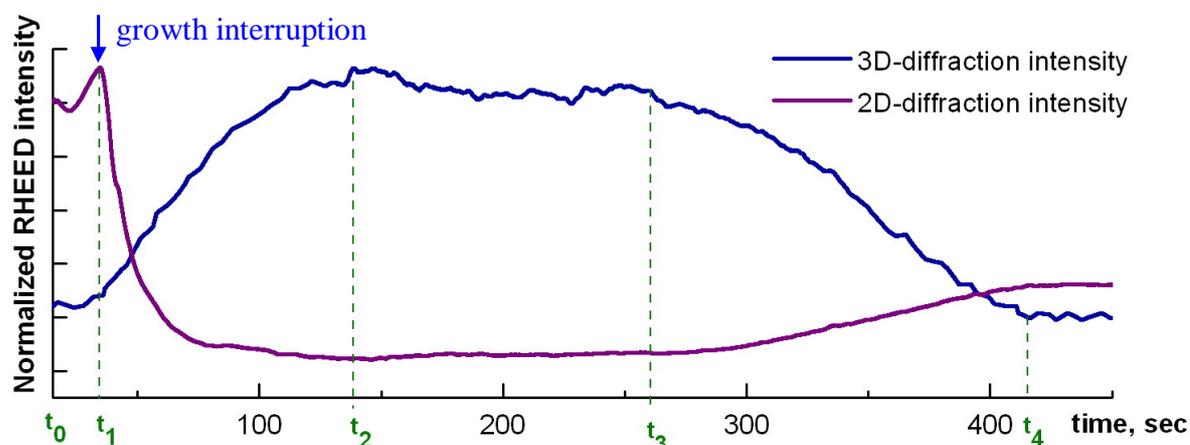


Figure IV-2: Dynamic of 2D and 3D RHEED intensities, recorded from the zones shown on Figure IV-1. The green dash lines indicate the main phases of the QDs evolution during the growth interruption as referred in the text.

Electron diffraction on a flat 2D surface produces a streaky RHEED pattern (Figure IV-1 (a)), whereas a QD plane exhibits a spotty pattern (Figure IV-1 (b)). The integral intensities of the 2D and 3D RHEED signals captured from zones shown on Figure IV-1 are presented on Figure IV-2. This plot indicates two processes in competition: reconstruction and evaporation of the GaN layer.

Four phases of GaN layer transformation can be distinguished. In a first stage (time $t_0 - t_1$), the GaN layer is deposited in a 2D growth mode. Right after the growth interruption, the 3D-diffraction intensity increases gradually (from t_1 to t_2), which signifies a progressive QD formation. After reaching a maximum, the RHEED signal remains almost constant ($t_2 - t_3$) indicating an invariance of the overall QDs surface. Then, GaN evaporation, which exists all over the growth interruption, leads to a progressive disappearance of the QD array ($t_3 - t_4$).

The minimum amount of GaN required for the QD formation on $Al_{0.5}Ga_{0.5}N$ template is 6 monolayers (ML) [18]. The critical thickness, above which plastic relaxation occurs, is about 43 ML. Anyhow the duration of the growth interruption required to accomplish the QD formation (instant t_2 on Figure IV-2) strongly depends on the GaN layer thickness. Matter removal via evaporation processes becomes dominant for a too long growth interruption, which limits the QD formation in thick GaN layer.

In a first series of samples, the amount of deposited GaN was varied. Three specimens with nominal GaN thicknesses of 8, 12, and 16 ML were fabricated (Table IV-1). The

duration of the growth interruption was identical for all four QD planes within each sample. A special sample where the growth interruptions were 2.5 times shorter than required to complete the QD formation was also made. The nominal thickness of GaN layer in this sample was 10 ML; this specimen is further referred as “unfinished sample”.

Various characterization techniques were used for the sample analysis. The composition of the AlGaN layer was controlled in situ by RHEED and confirmed by energy dispersive x-ray spectroscopy in a scanning electron microscope (SEM). According to x-ray diffraction measurements the residual deformation of the AlGaN templates is about 0.1%.

Photoluminescence experiments (PL) were done at 11 K using a frequency doubled Ar laser ($\lambda=244$ nm) [18]. Three peaks resolved in each PL spectrum are attributed to the AlGaN barrier, the GaN wetting layer and the GaN QD planes (Figure IV-3). The low temperature PL spectrum was also compared with the spectrum acquired at 300 K; the measurements were carried out with a HeCd laser working at 325 nm. The ratio of integrated PL intensity ratio between 11 K and 300 K PL is about 2.5, which is due to the increase of non-radiative recombination. However, the slight decrease of PL intensity with temperature signifies strong carrier localization in GaN QDs [19].

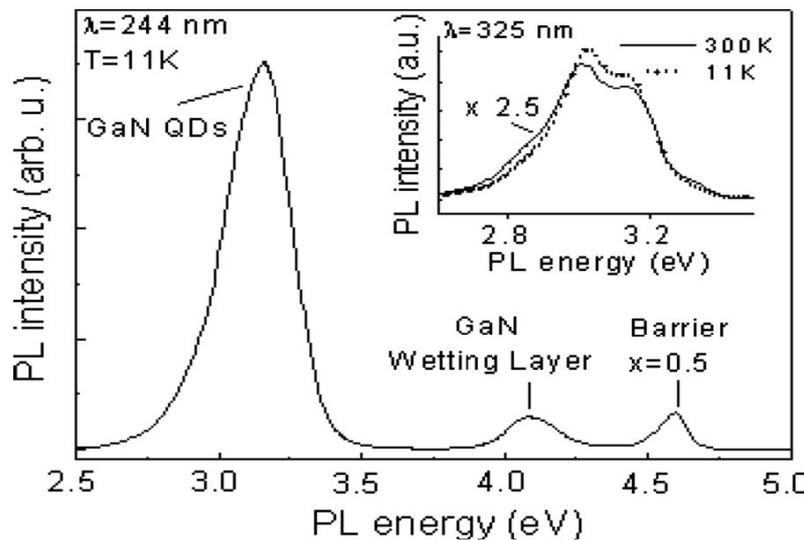


Figure IV-3: Low temperature PL spectrum of a sample with a nominal GaN thickness of 8 ML. The inset shows the comparison of low and room temperatures PL spectra.

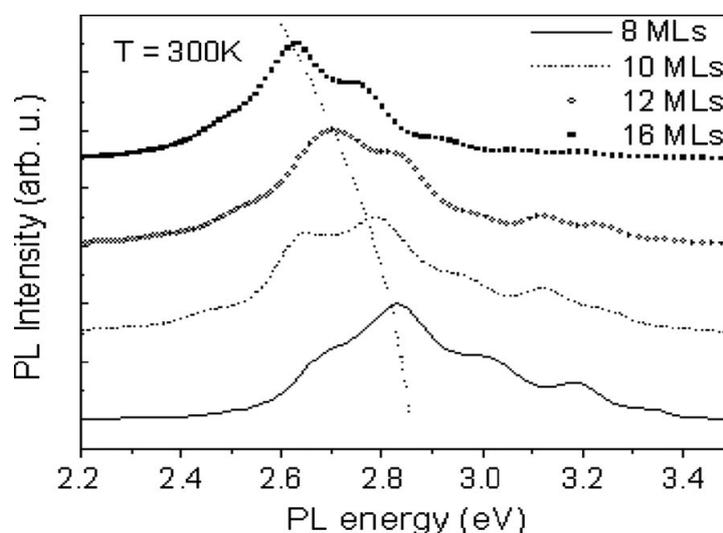


Figure IV-4: Room temperature PL spectra of the studied samples as a function of the GaN nominal thickness. The dotted line shows the PL energy maximum position.

Room temperature PL spectra of the series of samples are shown on Figure IV-4. A redshift in the PL energy maximum from 2.84 to 2.63 eV is observed by increasing the deposited GaN thickness, i.e. at lower energy than the GaN band gap (3.4 eV). This behaviour may be explained by an increase of the Stark effect which, in turn, indicates an increase of the QD height with the amount of deposited GaN.

The surface QD-plane morphology was first imaged by SEM. This observation revealed a dense and uniform distribution of QDs in all four samples (Figure IV-5).

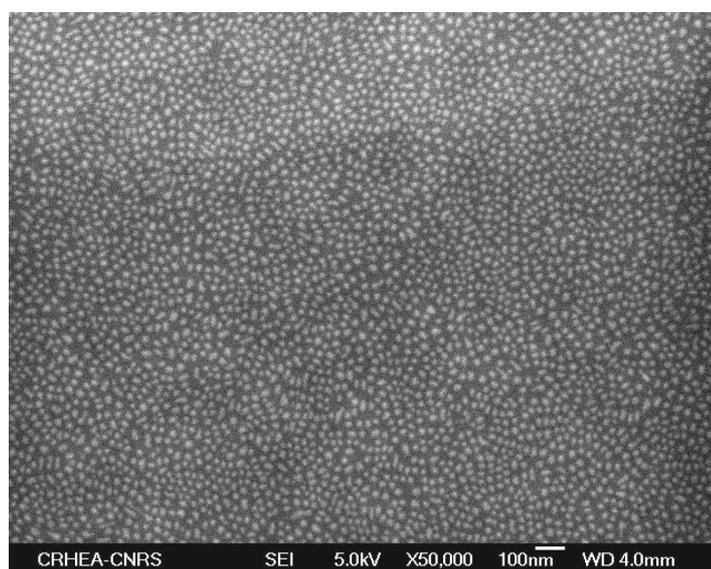


Figure IV-5: Typical SEM image of surface QDs in the sample with 16 ML of GaN.

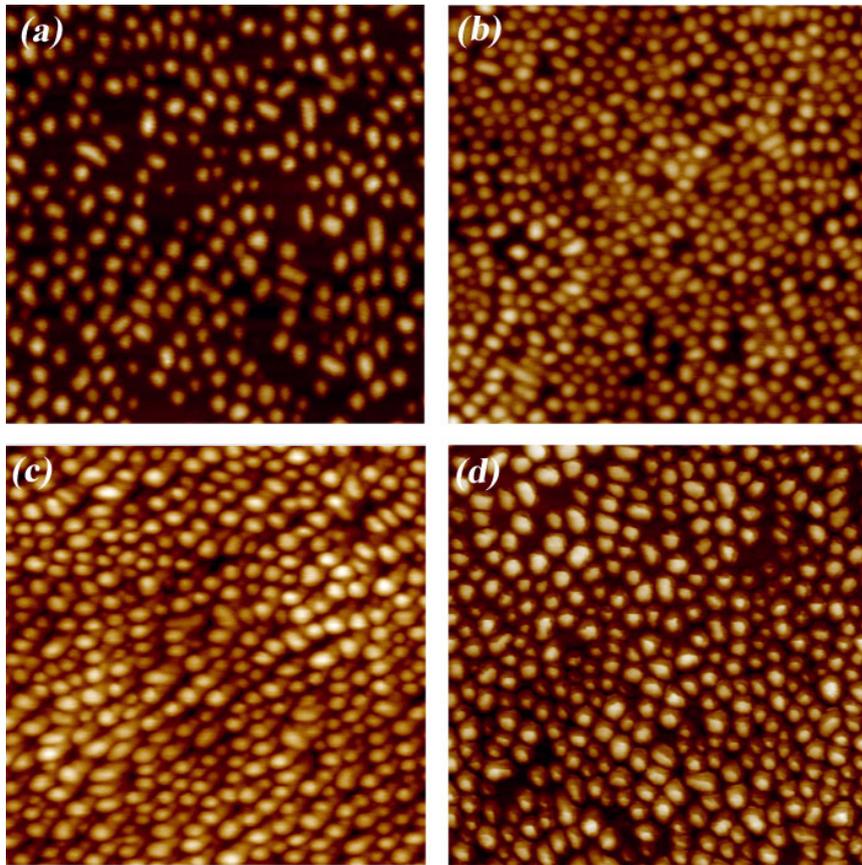


Figure IV-6: AFM $0.5 \times 0.5 \mu\text{m}^2$ images of samples with nominal thickness 8 ML (a), 12 ML (b), 16 ML (c) and unfinished sample (d) with 10 ML of GaN.

Atomic force microscopy (AFM) was then used for a more accurate characterisation of the surface QD morphology (Figure IV-6). The AFM observations underlined the homogeneity of the QD distribution in all the samples. The QD density is high even in the unfinished sample (10 ML), though here QDs are wider than in the other samples (Figure IV-6 d). The average QD densities measured by AFM are given in Table IV-1.

Sample number	N 1116	N 1136	N 1133	N 1100
Nominal GaN thickness	8 ML	12 ML	16 ML	10 ML
QD density	$1.6 \cdot 10^{11} \text{cm}^{-2}$	$2.0 \cdot 10^{11} \text{cm}^{-2}$	$1.2 \cdot 10^{11} \text{cm}^{-2}$	$1.3 \cdot 10^{11} \text{cm}^{-2}$

Table IV-1: Densities of QDs in the first series of samples.

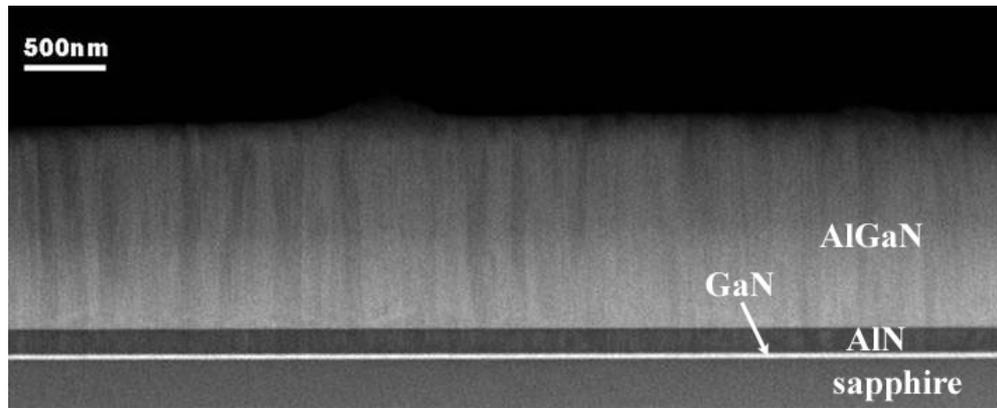


Figure IV-7: Low-magnification STEM-HAADF image displaying the sample structure.

Cross-sectional and plan-view TEM samples were prepared in a common way by mechanical polishing followed by ion milling. The initial TEM examinations concerned the Al_{0.5}Ga_{0.5}N templates (Figure IV-7).

Dark-field imaging revealed a high threading dislocation density (Figure IV-8).

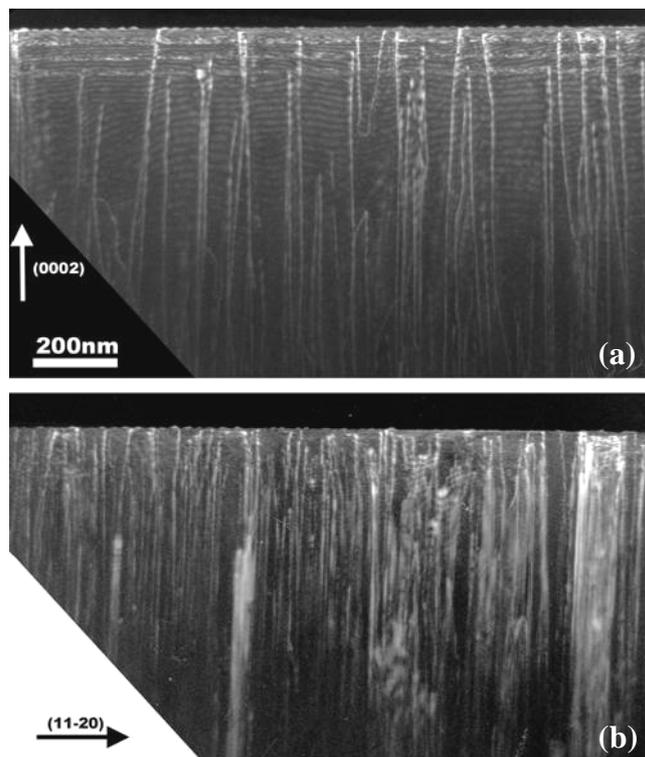


Figure IV-8: Dark-field images done close to the [110] zone axis using (0002) (a) and (11 $\bar{2}$ 0) (b) diffraction spots.

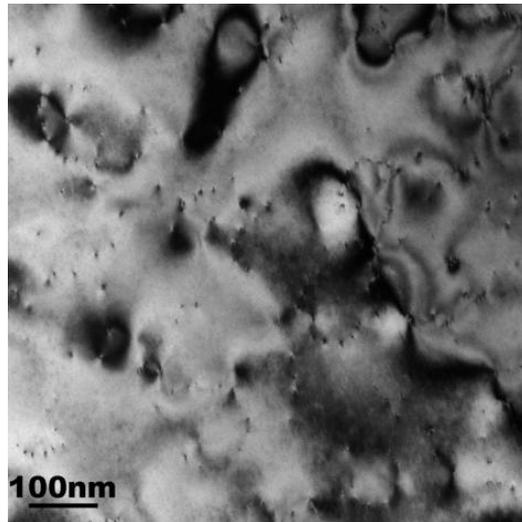


Figure IV-9: Plan-view TEM image. Dark points represent threading dislocation terminations.

The dislocation density, evaluated using plan-view images (Figure IV-9), is about $3 \times 10^{10} \text{cm}^{-2}$. This is 4-7 times smaller than the averaged QD density, which indicates that the majority of the QDs were formed in dislocation-free areas.

This is an important point, since GaN QDs show a tendency to nucleate on dislocation termination [20]. Consequently, an exciton migrating towards such QD can be trapped by a neighbouring dislocation. Anyhow this process will not play a significant role in the studied samples.

4.3. Morphology of GaN QDs

Initial TEM study was performed on a JEOL 2010F microscope. The presented images were taken along the $[14\bar{5}0]$ zone axis to obtain a strong chemical contrast.

Well-defined QD structures were revealed in all four planes of the first series of samples (Figure IV-10). Both surface and buried QDs have a hexahedral pyramidal shape, with facets parallel to $\{0001\}$ and $\{1\bar{1}03\}$ planes, as already reported for GaN QDs grown on AlN templates [21].

In all samples, buried QDs have a truncated pyramidal shape (Figure IV-11 (a)). QD sizes vary in a wide range (Table IV-2), whereas the height/diameter ratio of the majority of QDs is close to 0.17.

The shape of surface QDs depends on the amount of deposited GaN. Most of the surface QDs in the samples with 8 and 12 ML of GaN have a perfect pyramidal shape (Figure IV-11 (b) and (c)), while in the sample with 16 ML of GaN, almost all surface QDs are truncated, with a mean aspect ratio about 0.17 (Figure IV-11 (d)).

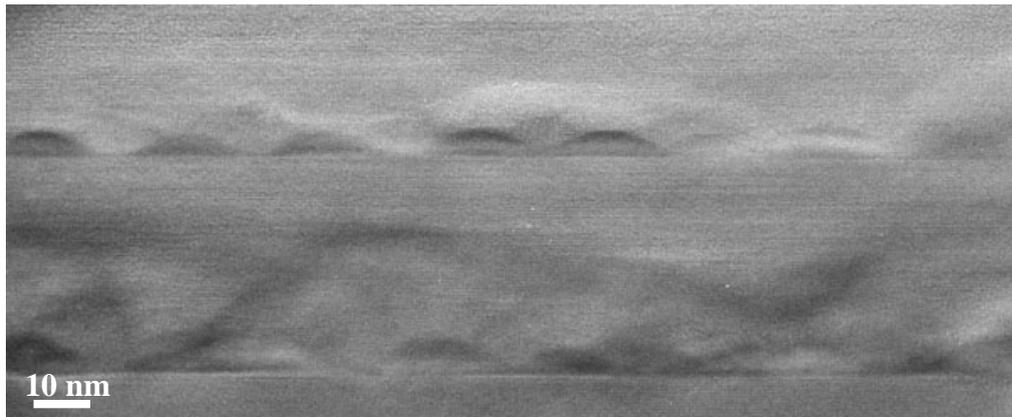


Figure IV-10 : TEM image of two buried QD planes in the sample with 12 ML of GaN.

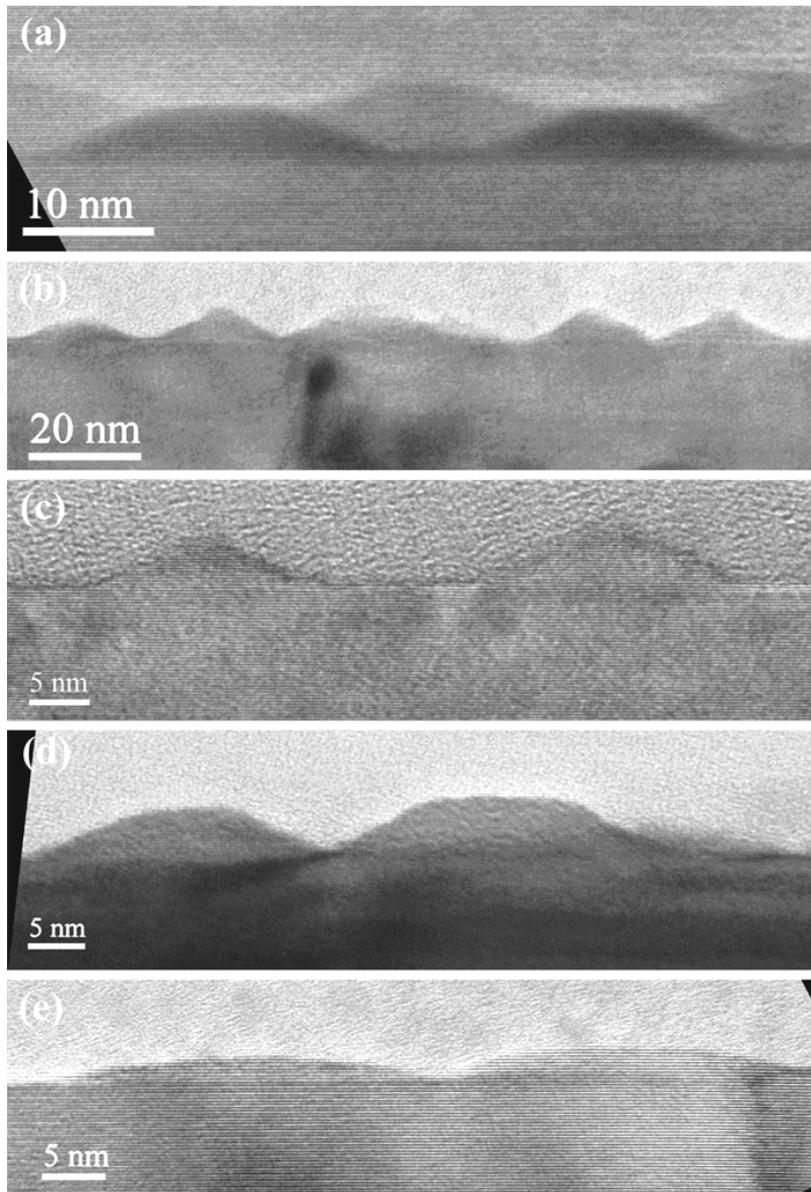


Figure IV-11: Variety of QD morphology in different samples: buried QDs in the sample with 12 ML of GaN (a), surface QDs in the sample with 8 ML of GaN (b), surface QDs in the sample with 12 ML of GaN (c), surface QDs in the sample with 16 ML of GaN (d) and surface QDs in the unfinished sample with 10 ML of GaN (e).

The removal of the QD apex due to the capping indicates a mass transport of GaN during the AlGaN barrier overgrowth. This suggestion was confirmed by the unfinished sample examination. All surface QDs are flat with an aspect ratio as low as 0.09 (Figure IV-11 (e)). However, QDs in the three buried planes have a well-defined truncated pyramidal shape, similar to that observed in the previous samples (Figure IV-11 (a)). This signifies that the QD formation initiated by the growth interruption is not stopped by the following barrier overgrowth.

To further investigate this phenomenon and evaluate the possible trends of mass transport, the sizes of the surface and buried QDs were measured from high-resolution images (Figure IV-12). During the TEM sample preparation a QD can be partially divided. For the statistic analysis we considered only those QDs which were completely embedded into AlGaIn. Data averaged over about 20–30 QDs are presented in Table IV-2. Standard deviations calculated within every data set are present as well.

The QD sizes rise with the increase of the nominal GaN thickness. The height of buried QDs is smaller than that of surface QDs in all four samples. Moreover, in samples with 8 ML and 12 ML of GaN, the mean diameter of buried QDs is larger than that of surface QDs. In the two other samples no noticeable difference of diameter was found. The aspect ratio of 0.17 appears to correspond to the equilibrium shape of the buried QDs.

Nominal GaN thickness	8 ML		12 ML		16 ML		10 ML (unfinished)	
	surface QDs	buried QDs	surface QDs	buried QDs	surface QDs	buried QDs	surface QDs	buried QDs
Height (nm)	3.86±0.65	3.05±0.58	4.82±0.61	4.22±0.55	4.83±0.45	5.2±0.59	1.92±0.28	3.25±0.32
Base diameter (nm)	16.8±2.8	19.0±3.2	21.1±3.3	25.7±3.4	28.3±1.7	26.9±3.1	20.9±1.5	21.6±2.2
Aspect ratio (height/diameter)	0.24±0.016	0.16±0.02	0.24±0.023	0.17±0.015	0.17±0.01	0.19±0.016	0.09±0.009	0.15±0.01
Volume (nm ³)	360±140	450±160	730±340	1110±370	1690±310	1620±500	390±100	690±190

Table IV-2: Morphology of surface and buried QDs as a function of the nominal GaN thickness. Height and diameter were directly measured from HRTEM images; QD height is given without the wetting layer thickness. Aspect ratio was calculated for every QD as height divided by diameter. QD volume was found for every QD using equation IV-1 and then averaged.

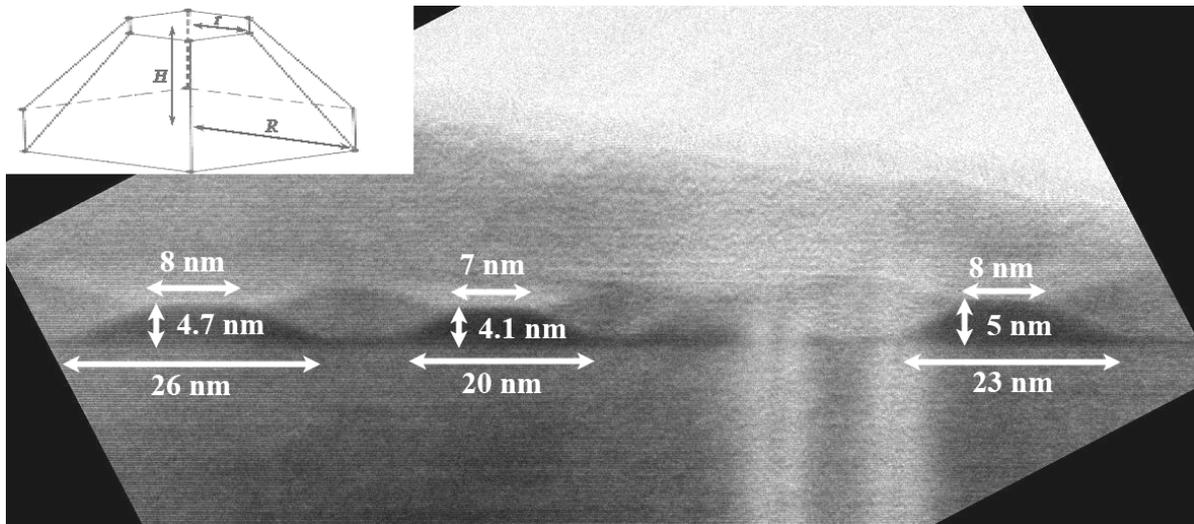


Figure IV-12: TEM image of buried QDs in the sample with 12 ML of GaN. Insert shows geometrical construction used for the QD volume evaluation.

Based on these geometrical data, the volumes of surface and buried QDs were also estimated. Truncated pyramid volume was found using the following equation

$$V = \frac{1}{3} \pi H (R^2 + Rr + r^2) \quad (\text{eq. IV-1})$$

where H is the pyramid height, R and r are the radii of large and small bases (Figure IV-12). We saw that the inhomogeneity of the QD sizes produces a statistical error on the measured height and diameter. In the case of QDs volumes measurements, standard deviations are much higher, but despite of this, a clear tendency of QD volume increase during capping is revealed (Table IV-2).

The volume increase is more pronounced in the unfinished sample with 10ML of GaN. In the same sample, the wetting layer (WL) in the surface QD plane is thicker than that of the buried one (Figure IV-13). The WL thickness measurements were based on the HRTEM image contrast variation: WL thickness in all three buried layers is 3-4 ML (as well as in all other samples). The surface WL thickness can not be determined precisely, since the contrast variations are not sufficiently high. But even qualitative observations showed a noticeable increase of the surface WL thickness (Figure IV-13). In fact, the thick GaN layer observed below the surface QD plane is not a common wetting layer as in other QD planes, but rather the rest of the initial 2D GaN layer.

No difference in the WL thickness was observed in the other samples of the series.

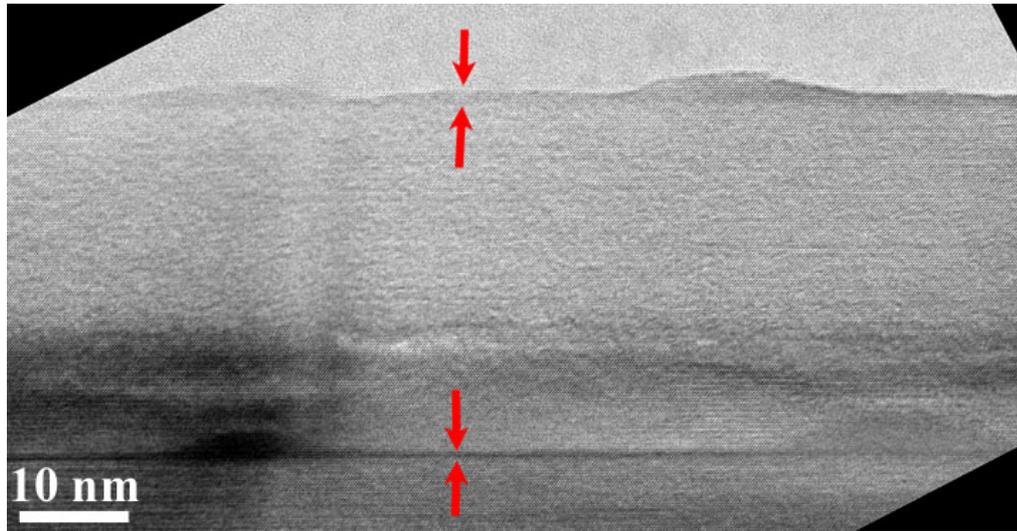


Figure IV-13: HRTEM image of third and fourth QD planes in the unfinished sample with 10 ML of GaN. The red arrows emphasize the difference of the wetting layers thicknesses.

The morphology evolution of GaN QDs grown on AlN template caused by the capping was also investigated to complete the previous study. The sample was also grown by ammonia assisted MBE on (0001) sapphire substrate. 6 ML of GaN was deposited at 800°C, after that the growth interruption was performed to initiate the QD formation. A stacking structure, comprising three buried AlN QD planes and an uncapped surface QD plane was realized (N 1316).

GaN/AlN QD shape transforms in a similar way as the GaN/AlGaIn QD: from a perfect pyramid in the surface plane (Figure IV-14 (a)) to a truncated pyramid in the internal planes (Figure IV-14 (b)). The wetting layer thickness is estimated to be about 2-3 ML.

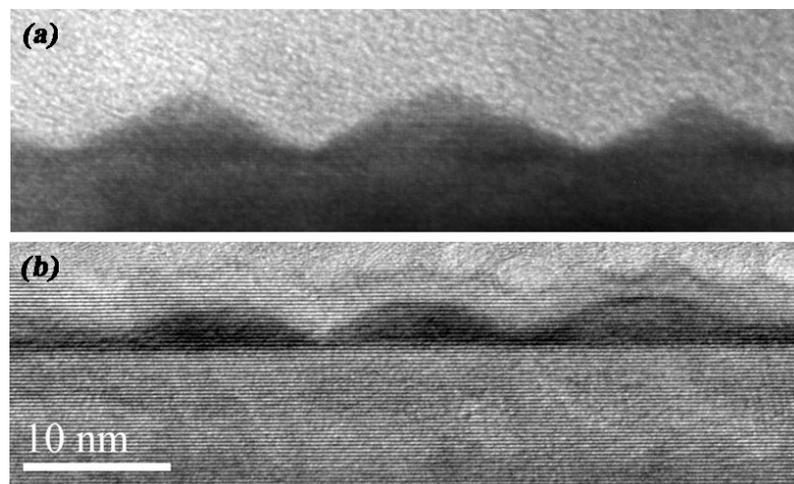


Figure IV-14: TEM images of surface (a) and buried (b) GaN QDs grown on an AlN template. Both images have the same scale.

.Nominal GaN thickness	GaN/AlN QDs	
	surface QDs	buried QDs
Height (nm)	3.26±0.77	2.6±0.41
Diameter (nm)	13.4±3.0	14.5±2.1
Aspect ratio	0.25±0.02	0.18±0.015
Volume (nm ³)	200±100	250±110

Table IV-3: Morphological parameters of GaN QDs grown on an AlN template. The height is given deducting the wetting layer thickness.

The shape modifications are also accompanied by the QD geometry sizes evolutions (Table IV-3). Statistical analysis showed a decrease of the buried QD height without noticeable increase of its diameter. The aspect ratio of buried GaN/AlN QDs (0.18) is close to that of the GaN/AlGa_N QD (0.17). The QD volume again shows a tendency to increase due to the capping process

The primary study of the GaN QDs grown on Al_{0.5}Ga_{0.5}N templates revealed many original phenomena for nitride materials [22]. Possible mechanisms will be discussed at the end of this Chapter; here we will just summarize the main results.

- The surface QD shape varies from perfect pyramid to truncated pyramid with the increase of the nominal GaN thickness. Besides, QD shape and WL thickness is governed by the duration of growth interruption. Shorter growth interruption leads to the formation of flatter surface QDs on a thicker wetting layer.
- Whatever is the nominal GaN thickness or duration of the growth interruption, buried QDs always exhibit a truncated pyramid shape with an aspect ratio close to 0.17.
- The volume of buried QDs is bigger than that of surface QDs. A similar tendency was revealed for the GaN QDs grown on AlN templates.
- The mass transport of GaN from the WL to the QDs occurs even during the AlGa_N barrier overgrowth.

4.4. Study of the AlGaN barrier microstructure by HRTEM

$\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ barriers also exhibit an unexpected behaviour. A characteristic contrast, clearly visible in Figure IV-11 (a) or Figure IV-12, exists around buried QDs. Brighter mass-thickness contrast in characteristic zones suggests that these zones are composed from lighter element, i.e. they are Al-rich; darker contrast implies Ga-rich zones.

To verify this supposition and estimate the amount of phase separation, the strain ε_{zz} was measured by GPA in the characteristic zones. HRTEM images for strain determination were acquired in two-beam conditions (Figure IV-15 (a)), which are better adapted for quantitative analysis (Section 3.3).

The strain map calculated using a small size mask (1 nm^{-1}) is shown on Figure IV-15 (b). The AlGaN barrier under the QDs, chosen as a reference region, does not exhibit any noticeable local lattice parameter (LLP) variation. GaN QDs and WL appear in red, since the lattice parameter of GaN is higher than that of $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$. The zones above the QDs (marked by white arrows on Figure IV-15) have smaller LLP, which may indicate that they are Al-rich. An elongated arc-shape zone can be distinguished between the QDs (marked by the black arrow on Figure IV-15). The averaged LLP in this zone is higher than in AlGaN barrier; thereby it can be Ga-rich.

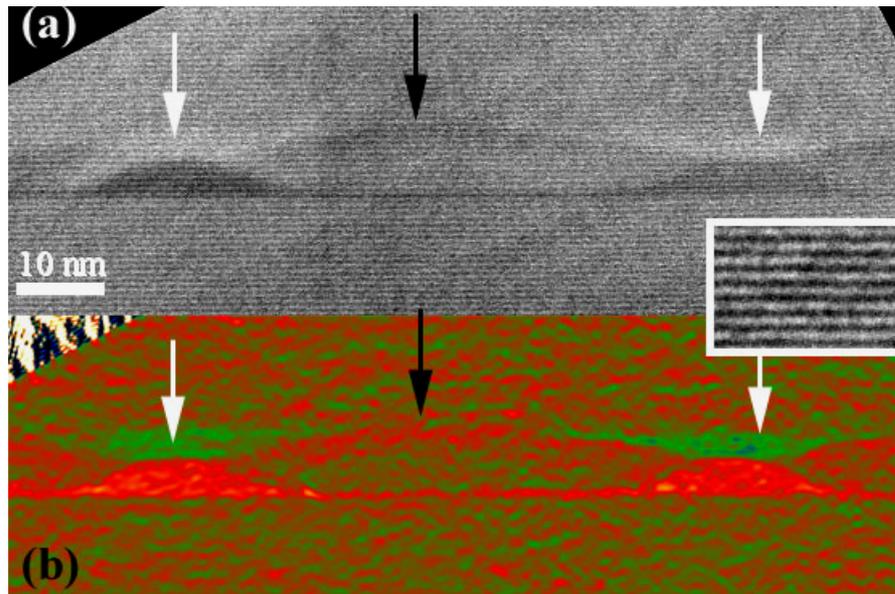


Figure IV-15: Off-axis HRTEM image of two neighbouring QDs (a). The corresponding strain map obtained using GPA (b). Arrows indicate characteristic contrasts above and between the QDs. Inset is an enlargement showing (0002) lattice fringes.

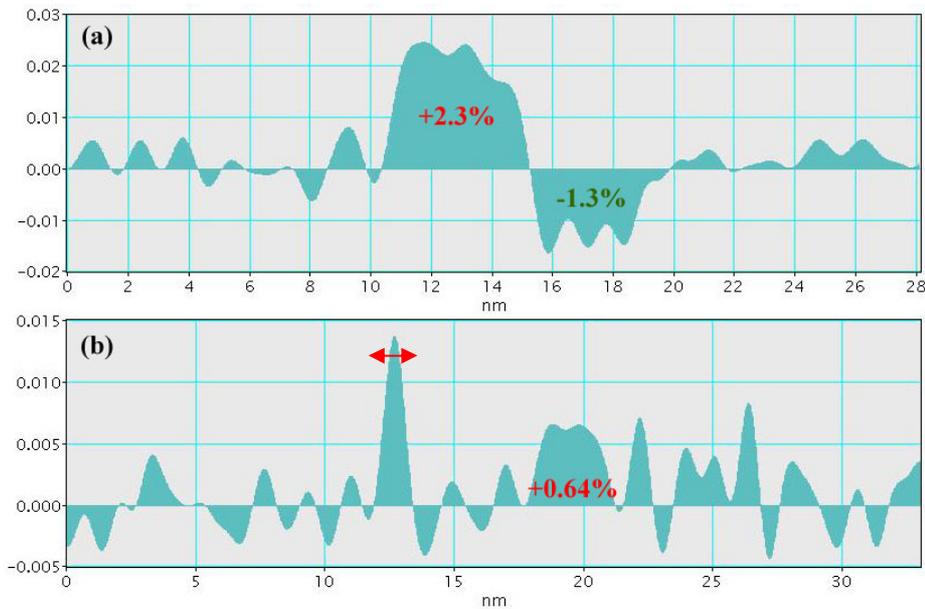


Figure IV-16: Strain profiles traced through the left QD (a) and between the two QDs (b), which are shown on Figure IV-15. The red arrow shows the wetting layer strain.

The mean value of strain within the GaN QD is $2.3\% \pm 0.2\%$ (Figure IV-16 (a)). In Section 3.1 we found that the strain ϵ_{zz} of a thin GaN film biaxially stressed on $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template is 2.61%. The same strain ϵ_{zz} calculated in the uniaxial approximation is 2.31%.

The amount of strain in QD is always smaller than in biaxially stressed film, due to the elastic energy relaxation via the QD formation. Nevertheless, in the studied samples, the ϵ_{zz} strain measured in the GaN QD is not much smaller than the value calculated theoretically for a 2D thin GaN layer. Moreover, it turns out that the measured ϵ_{zz} strain is in good agreement with the uniaxial approximation.

Complicated calculations, considering QD stress state, will provide more reliable values of strain expected within the QD [23]. Such calculations are impossible for Al- and Ga-rich zones since their exact form and local composition are unknown. However, we will use bi- and uniaxial approaches to evaluate the amount of local composition fluctuations in the AlGaN barriers. Obtained values will be verified later by HAADF and EELS measurements.

Another source of biaxial approximation inaccuracy is that QD is embedded in a bulk material with lattice mismatch. However, this effect decreases with the TEM sample thickness (Figure IV-17).

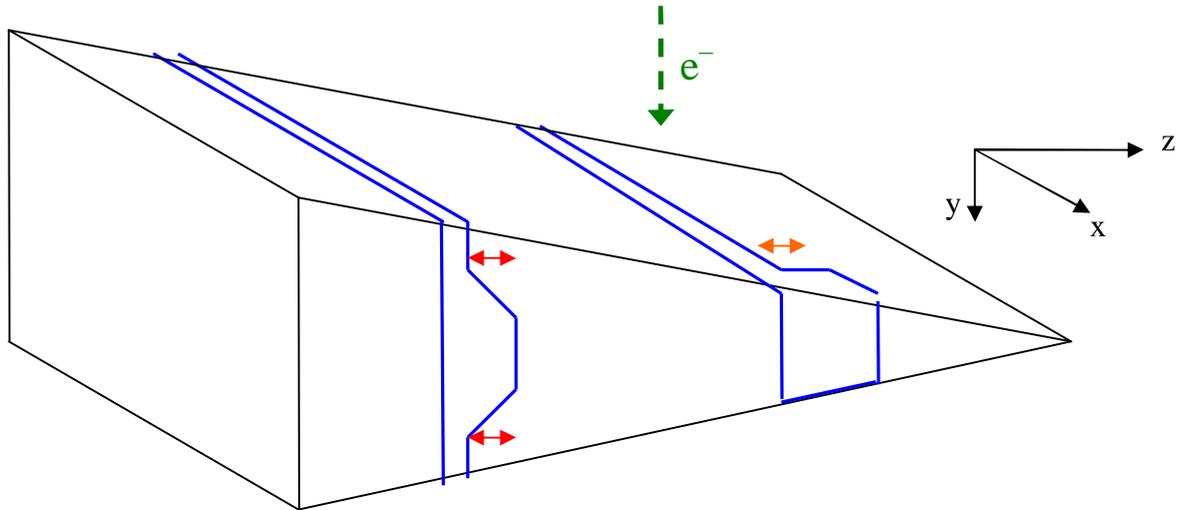


Figure IV-17: Scheme showing the cross-section of a TEM sample. Black lines indicate the sample borders, blue lines show two QDs. The coloured arrows symbolize compressing effect exerted by AlGa_N, which surrounds GaN QDs. The green dash arrow shows the direction of the electron beam.

The left GaN QD in Figure IV-17 is surrounded by AlGa_N, which has a smaller lattice parameter. This surrounding compresses the GaN QD in the z-direction (the effect is shown by the red arrows). Besides, the electrons traversing such QD in the y-direction, interact with the overlying and underlying AlGa_N. Consequently, the ϵ_{zz} measured in this area of the HRTEM image does not correspond to the LLP in the QD.

The situation is different if the QD occupies the whole sample along the y-direction (right QD on Figure IV-17). Here, the AlGa_N template does not superpose any more on the QD structure. The QD is still compressed by the neighbouring AlGa_N (shown by the orange arrow); but this effect is smaller than in the previous case. However, thin foil relaxations, may occur at a thickness of 20-25 nm (typical QD diameter), which will also decrease ϵ_{zz} (Figure III-2 in Section 3.1).

The value of strain measured in the Al-rich zone is $-1.3\% \pm 0.3\%$ (Figure IV-16 (a)), which corresponds in the biaxial approximation to $75\% \pm 5\%$ of Al. The same value of strain interpreted in the uniaxial approximation gives an Al concentration of $78\% \pm 6\%$. Similar interpretation of the Ga-rich zone (Figure IV-15 (b)), where a ϵ_{zz} strain $0.64\% \pm 0.12\%$ was measured, indicates a Ga composition about $62\% \pm 3\%$.

A large number of HRTEM images were acquired in all GaN/AlGaIn QD samples. The Al composition calculated with a biaxial approximation above different QDs varies from 69% to 77%. The Ga composition varies from 55% to 66%, though in some samples Ga-rich zones were not even detected. In fact, the size and shape of Ga-rich zones strongly depend on the distance between the QDs; consequently the Ga concentration may vary. On the contrary the sizes of the Al-rich zones are always similar to the sizes of the corresponding GaN QDs.

Strain analysis was also used for the measurement of the GaN wetting layer thickness. The WL is visible on the strain profile (marked by the red arrow in Figure IV-16 (b)), but its full width at the half-maximum (FWHM) is about 1 nm, which is the spatial resolution corresponding to the used GPA mask. The application of the large size mask (2 nm^{-1}) produces a sharper profile (Figure IV-18 (a)). However, the FWHM of the wetting layer is 0.5 nm, i.e. its thickness is still above the spatial resolution of GPA.



Figure IV-18: ϵ_{zz} strain within the GaN WL measured by GPA (a) and the projection method (b). In Figure (b) the strain scale is multiplied 10^4 times and superposed on the HRTEM image profile. Figure (c) is an enlargement of the zone shown by the red rectangle. Red arrows indicate the contrast decrease within the GaN WL.

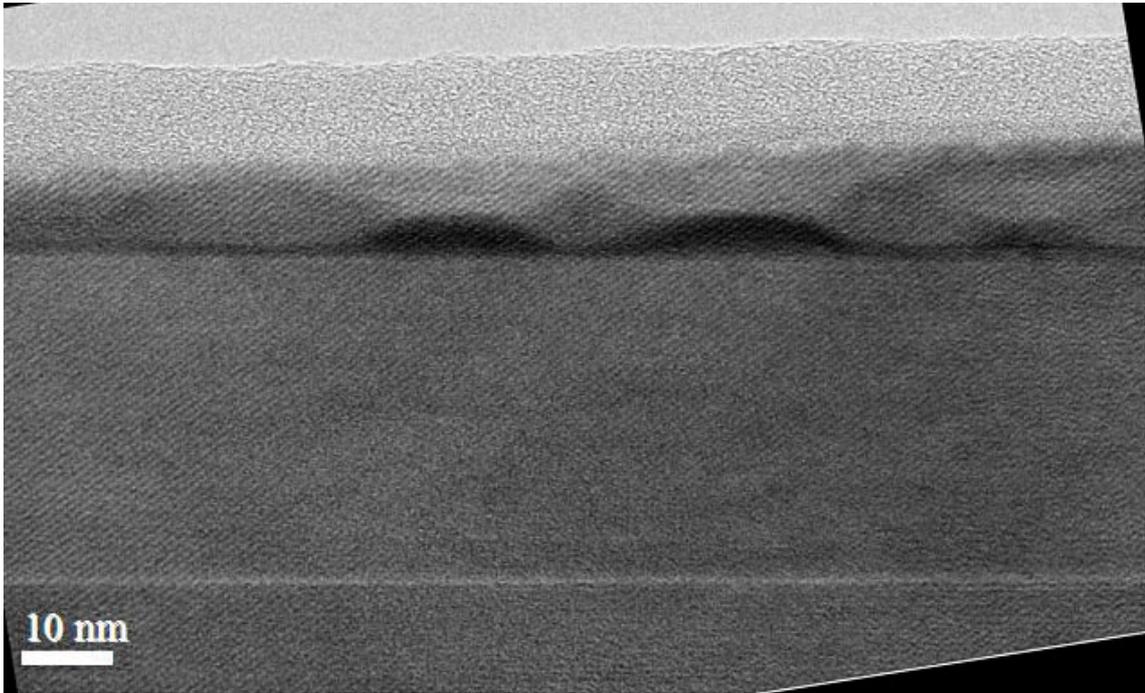


Figure IV-19: HRTEM image of an Al-rich layer underlying the first QD plane.

The profile of ε_{zz} strain obtained from the same image (Figure IV-15 (a)) by the projection method is shown on Figure IV-18 (b). The WL thickness measured from this plot is three monolayers. Interestingly, that amount of ε_{zz} strain measured within the GaN WL by the projection method ($2.3\% \pm 0.2\%$) is similar to the amount of ε_{zz} strain measured within the GaN QD by GPA. Moreover, both values are in perfect agreement with a strain ε_{zz} of a thin GaN film calculated in the uniaxial approximation.

The WL thickness determined from other HRTEM images varies between 2 ML and 3 ML, which is in opposition with the value determined by contrast variations: 3-4 ML. Indeed, in the previously analysed profile (Figure IV-18 (b)) the contrast is weaker within four intensity maxima (shown by red arrows on Figure IV-18 (c)). This discrepancy will be clarified in Section 4.5, where the results of STEM-HAADF observations are presented.

Another striking result related to AlGaN is the detection of an additional layer below the first QD plane (Figure IV-19). Brighter mass-thickness contrast suggests that this layer is Al-rich; though no intentional composition modifications of the AlGaN template were done before the first QD plane formation.

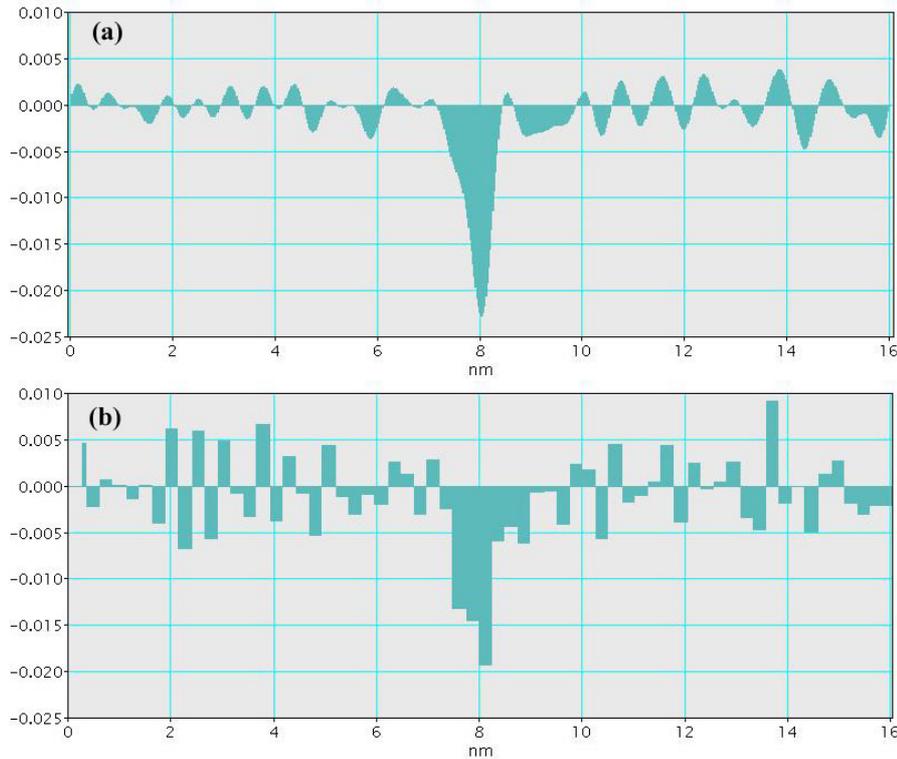


Figure IV-20: ϵ_{zz} strain within the Al-rich layer measured by GPA (a) and projection method (b).

In fact, the growth temperature is always calibrated before a QD structure realisation. For that purpose, the growth of the AlGaIn template is interrupted and a GaN layer of a given thickness is deposited above it. The duration of the GaN QD formation and evaporation is determined by RHEED pattern evolution. The temperature is then deduced from an extrapolation of the evaporation rates measured at calibrated temperatures [24].

After the growth temperature calibration, a 30 nm thick AlGaIn spacer is deposited, and followed by the GaN layer deposition and QD formation. The distance between the additional Al-rich layer and the first QD plane measured by HRTEM is about 30 nm, which indicates that this Al-rich layer is formed during the AlGaIn growth interruption.

The thickness of the Al-rich layer measured by the change of the HRTEM contrast varies from image to image within 1-1.5 nm. An interesting detail is that the Al-rich layer thickness is independent of the growth interruption duration.

The LLP of the Al-rich layer was measured by GPA and the projection method. Profiles are presented on Figure IV-20. The GPA map was obtained with a spatial resolution of 0.5 nm (large mask size), i.e. 2-3 times higher than the typical thickness of the Al-rich layer. It shows successive increase and decrease of ϵ_{zz} strain (Figure IV-20(a)). However, the

effect of strain averaging hampers the interpretation of GPA profiles at sub-nanometre scale. Strain profile calculated by the projection method (Figure IV-20(b)) indicates an Al-composition rise up to 85% in the middle of the layer.

All GaN/AlGa_N QDs samples were also examined with a Cs-corrected Titan 80-300 microscope. The employment of such a high-performance machine has two main advantages. First, due to a spatial resolution of about 0.1 nm, the nitride samples can be observed along the [1100] zone axis. As we saw in Section 3.4, the strain obtained in the [1100] projection is less affected by the image conditions comparing to the [1120] zone axis. However, the sample examination along the [1100] zone axis is too complicated with the JEOL 2010F microscope, whose point-to-point resolution (0.194 nm) is higher than the interatomic spacing to be resolved ($a_1 = 0.158$ nm).

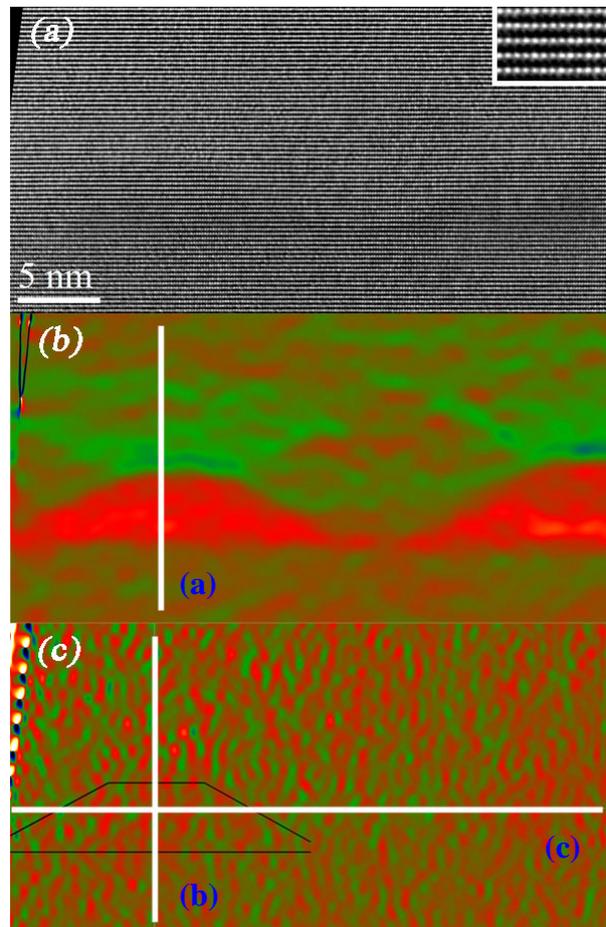


Figure IV-21: HRTEM image of two neighbouring buried QDs (a) taken along the [1100] zone axis. The inset is an enlargement of the HRTEM image. The corresponding GPA strain maps ϵ_{zz} (b) and ϵ_{xx} (c). White lines indicate the directions of the profiles shown on Figure IV-22. Black lines displays the QD contours determined from ϵ_{zz} strain map. The blue letters refer the profiles shown in Figure IV-22.

The second benefit is that, due to the high stability of the Titan 80-300, the strain measurement is more accurate, comparing to the one obtained using the JEOL 2010F. Moreover, HRTEM images, acquired using the Cs-corrected Titan 80-300 microscope have a uniform contrast (compare Figure IV-15 (a) and Figure IV-21 (a)). This is a significant advantage, especially since the GPA technique was shown to create fluctuations in strain maps because of HRTEM image contrast variations (Section 3.2.1).

Strain maps ε_{zz} and ε_{xx} were calculated by GPA using a small size mask (spatial resolution of 1 nm). Note, that the Ga-rich zone between the QDs is not revealed in the ε_{zz} strain map (Figure IV-21 (b)), whereas the Al-rich zones, placed above the both QDs, are well developed. Moreover, a characteristic distribution of the strain ε_{zz} was detected within the Al-rich zones (Figure IV-22 (a)). The strain ε_{zz} becomes weaker as the distance from the QD increases. Such behaviour may be interpreted as a decrease of the Al composition in the Al-rich zones. This guess can not be, however, generalized to all QDs, since in some HRTEM images no similar ε_{zz} strain variations were detected.

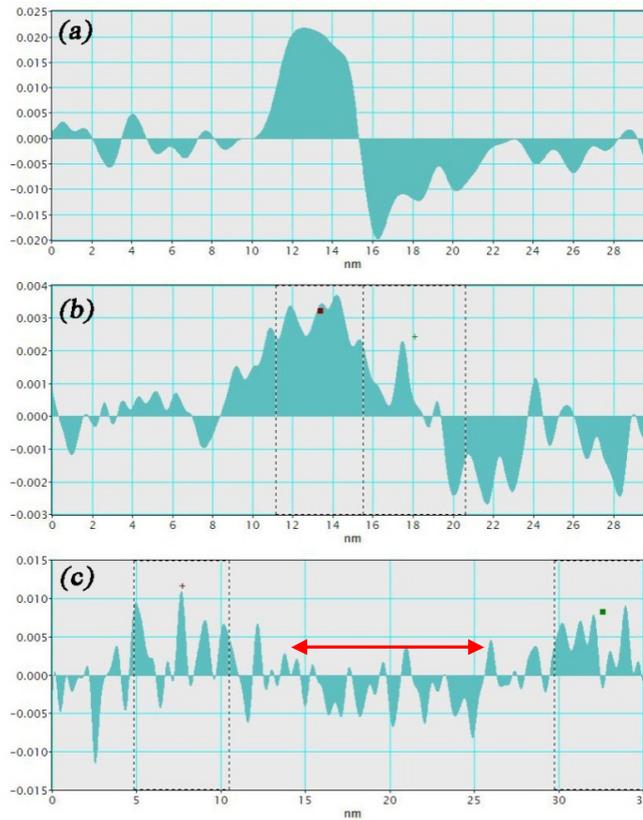


Figure IV-22: Profiles of strain maps (shown in Figure IV-21 (b) and (c)): ε_{zz} (a) and ε_{xx} (b) traced in [0001] direction. Dash lines show the positions of the GaN QD and the Al-rich zone. Profile of strain map ε_{xx} traced in [1120] direction (c). Dash lines mark the positions of the two GaN QDs, the red arrow shows the AlGaN region with slightly decreased lattice parameter.

The ε_{xx} strain map (Figure IV-21 (b)) does not exhibit any noticeable strain variations. Nevertheless the profile, traced in the [0001] direction (Figure IV-22 (b)), reveals a small increase of the ε_{xx} strain within the QD up to $0.3\% \pm 0.04\%$ (averaged value). A similar increase of LLP is visible in the profile, traced from one to another QD in the [11 $\bar{2}$ 0] direction (Figure IV-22 (c)). On the other hand, the LLP between two QDs (red arrow on Figure IV-22 (c)) shows a tendency to decrease. Low signal/noise ratio in this profile indicates that the measurements of ε_{xx} are done at the limit of the microscope performance. In fact, the detection of 0.3% deviation of $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ lattice parameter ($a_1=158$ pm, Figure II-4(a), Section 2.3) requires a precision of intensity maxima determination better than 0.5 pm.

The strain distributions obtained experimentally were compared to theoretical calculations done by the Eshelby method [25]. Buried QD was represented by a regular truncated hexagonal pyramid with an aspect ratio 0.17. No composition variations within the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ spacer were considered. The x-axis is chosen parallel to the [11 $\bar{2}$ 0] direction, the y-axis is parallel to the [1 $\bar{1}$ 00] direction, and the z-axis is parallel to the [0001] direction (Figure IV-23).

In contrast to the GPA maps, presented above, all calculated strains are referred to the lattice of unstressed material. That is the strain inside the QD is given with respect to the GaN lattice, and strain in the barrier refers to the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ lattice parameters.

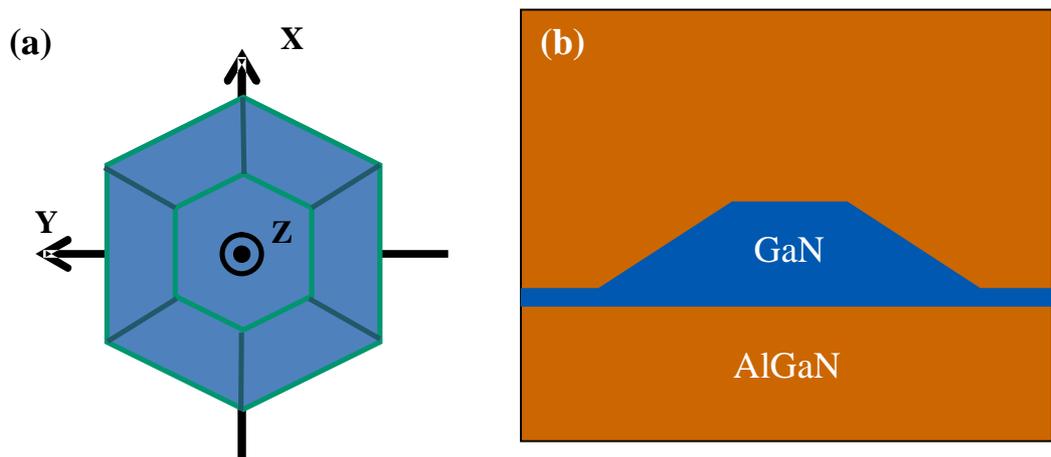


Figure IV-23: Shape of the GaN QD (a) and the AlGaN spacer structure (b) used for the strain calculations.

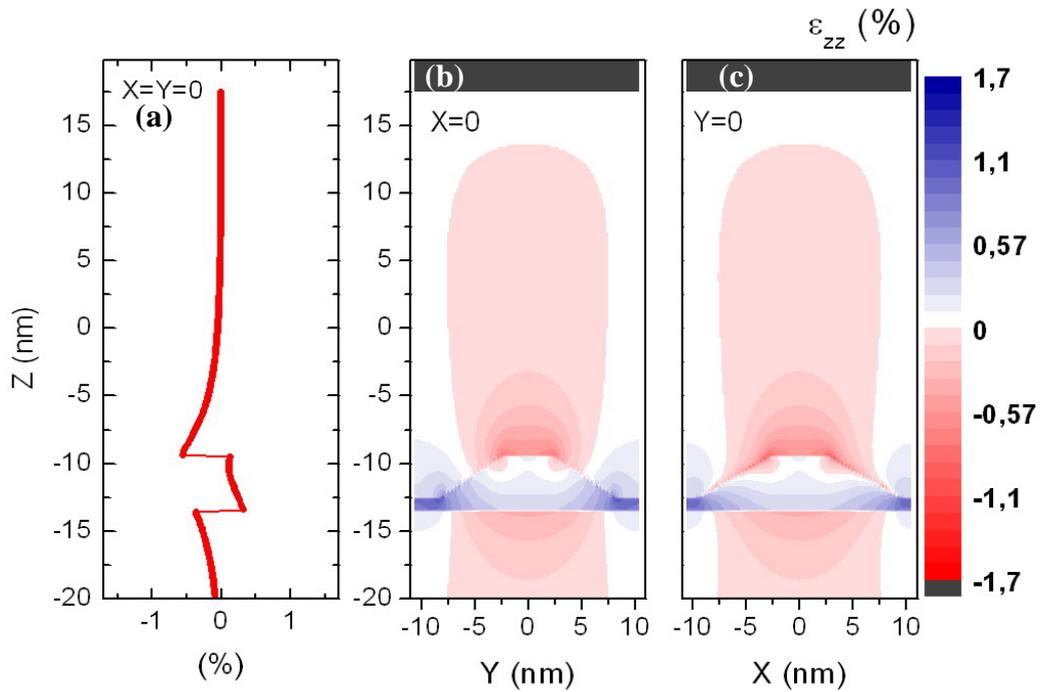


Figure IV-24: Variation of ϵ_{zz} strain into a buried QD along the z-axis (a) and sections of ϵ_{zz} strain distribution by planes $x=0$ (b) and $y=0$ (c).

The main strain tensor components ϵ_{xx} , ϵ_{yy} and ϵ_{zz} were computed in each point of the modelled QD/barrier system. The 3D strain distributions are visualized as sections of the strain tensor components by planes $x=0$ and $y=0$ (Figure IV-24 and Figure IV-25).

Actual HRTEM images, acquired along the $[1\bar{1}00]$ direction (i. e. the y-axis), correspond to the strain distribution in the plane $y=0$. However, the section of the ϵ_{zz} strain shown on Figure IV-24 (c) is not entirely identical to that obtained from TEM micrograph, since LLP measured from HRTEM image is averaged over the TEM sample thickness. Another source of differences is that the GaN QD is considered to be completely embedded into an AlGaIn matrix, as it happens in a thick TEM sample. HRTEM images, destined for strain analysis, were acquired in thin zones, where the compression effect of AlGaIn is weaker.

Consequently, the calculations show that the ϵ_{zz} strain in the QD is weaker than that in the wetting layer, which confirms that the biaxial approximation is not valid for the case of a thick TEM sample. Another model prediction is that a zone of negative ϵ_{zz} strain in the AlGaIn barrier is situated just above the QD. Moreover, the amount of strain decrease as the distance from the QD increases.

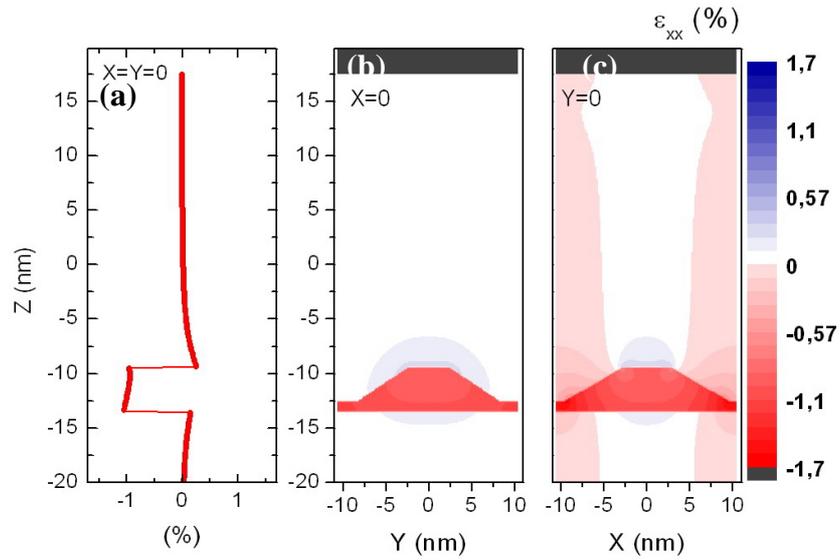


Figure IV-25: Variation of ε_{xx} strain into buried QD along the z-axis (a) and sections of ε_{xx} strain distribution by planes $x=0$ (b) and $y=0$ (c).

It has to be recalled that no composition variations in the barrier were included in the strain model. Therefore, strain variations can not be directly interpreted as a change of the alloy composition. And the hypothesis of phase separation into the AlGaIn barriers should be proved by an alternative technique (Section 4.5).

The calculated ε_{xx} strain is almost uniform within the buried QD (Figure IV-25). Zones of small compressive ε_{xx} strain in the barrier are located on each side of the QD. This result is quite similar to that revealed experimentally (Figure IV-22 (c)).

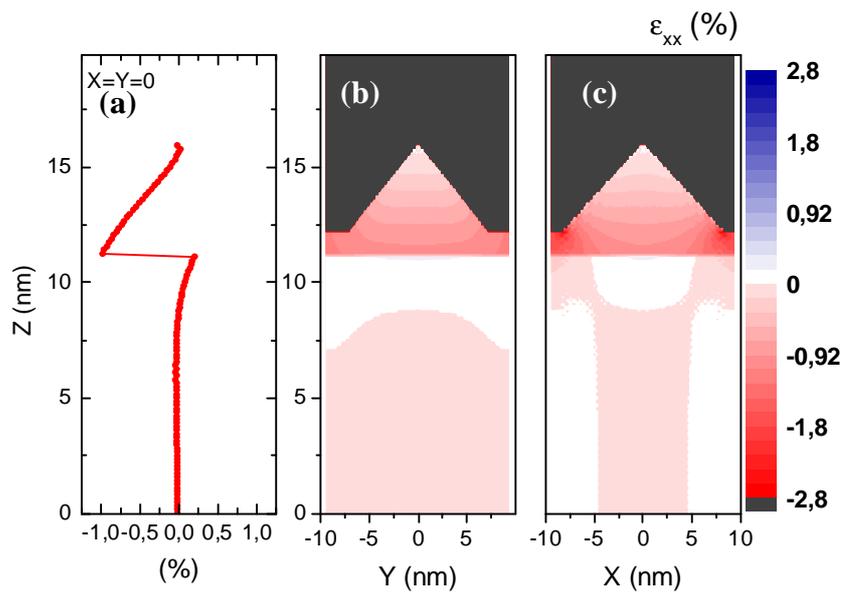


Figure IV-26: Variation of ε_{xx} strain into surface QD along the z-axis (a) and sections of ε_{xx} strain distribution by planes $x=0$ (b) and $y=0$ (c).

In the surface QD, the ϵ_{xx} strain variation is more considerable (Figure IV-26). The QD base is almost perfectly strained on the AlGaIn template, whereas the QD top is more relaxed. Such atomic reorganization inside the QD is governed by the minimization of the elastic energy stored in the GaN QD due to the lattice mismatch with AlGaIn.

The effect of the surface QD relaxation was also observed experimentally (Figure IV-27). For the GPA strain calculations, the reference region is always chosen in the unstrained AlGaIn layer lying above the QD plane. The gradual increase of the ϵ_{xx} strain (Figure IV-27(d)) indicates that the lattice parameter a inside the surface QD is higher than that of the AlGaIn template. This clearly testifies that the relaxation of elastic energy occurs due to the QD formation.

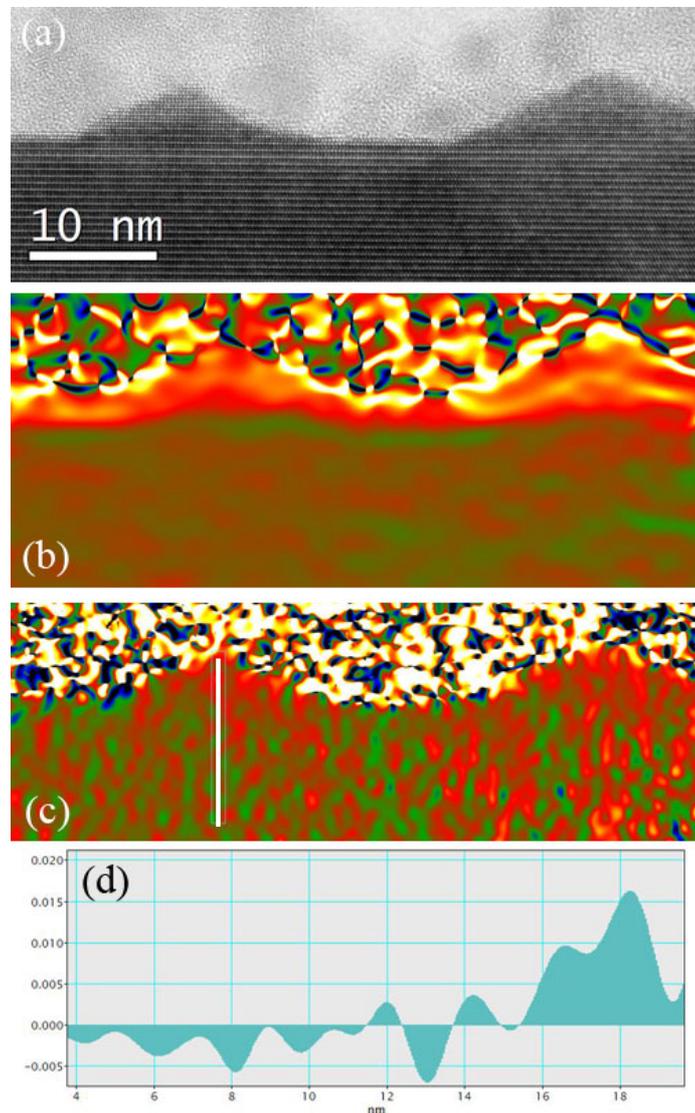


Figure IV-27: HRTEM image of two neighbouring surface QDs taken along the [1120] zone axis (a). The corresponding GPA strain maps ϵ_{zz} (b) and ϵ_{xx} (c). The white line indicates the direction of the profile (d).

Hereby, based on the quantitative information extracted from the HRTEM images by strain analysis, we can state that:

- A phase separation phenomenon was revealed into the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ barriers around every buried GaN QD. Al-rich regions are formed above the QD and Ga-rich zones are situated between the QDs. The estimated Al concentration into the Al-rich zones varies between 70% and 75% and tends to decrease, as the distance from the QD increases. The Ga composition in Ga-rich zones is about 55-65%.

However, this supposition should be independently verified (Section 4.5).

- The strain distribution determined from HRTEM images are in good agreement with theoretically calculated ε_{zz} and ε_{xx} components of the strain tensor. No considerable variations of the ε_{xx} strain around buried QDs were measured, i.e. GaN QDs are almost perfectly strained on the AlGaN template. Whereas the effect of surface QD relaxation was evidenced.

- The interruption of the AlGaN template growth results is an involuntary formation of an Al-rich layer with Al concentration up to 85%. The Al-rich layer thickness (4-6 ML) is independent of the growth interruption length.

4.5. AlGaN barrier composition evaluations by STEM-HAADF and EELS

As we saw in the previous Section, the strain distribution in AlGaN spacers observed around buried GaN QDs can not be directly linked to a composition variation. Indeed, zones of compressive strain situated above the buried QDs were already reported in GaN/AlN QD structure, where alloy demixing is excluded [23].

To examine the AlGaN barriers for alloy demixing a series of STEM-HAADF observations was performed with a probe-corrected Titan 80-300. In the initial observations the samples were tilted by a few degrees out of the $[11\bar{2}0]$ zone axis to avoid the superposition of the channelling effect over the Z-contrast image. A phase separation phenomenon was instantly evidenced in all the samples (Figure IV-28).

GaN QDs appear bright against the AlGaN background, since heavier elements produce brighter contrast in HAADF images (Section 1.3). Zones above the QDs have a darker contrast, i.e. they are Al-rich. The size and shape of the Al-rich zones are similar to the size and shape of the corresponding GaN QDs (similar conclusion was based upon strain analysis). Moreover, the Al-rich zones seem to have a truncated pyramid shape with well-definite lateral facets. The Ga-rich regions are placed between the Al-rich zones (Figure IV-15).

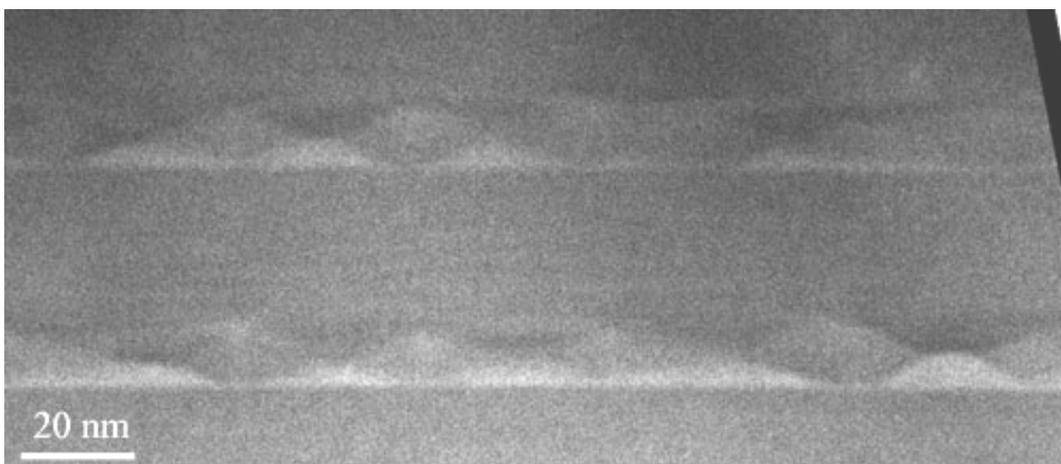


Figure IV-28: STEM-HAADF image of two QD planes embedded in an $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ matrix.

HAADF images exhibit a stronger contrast comparing to HRTEM images. Besides, the qualitative interpretation of HAADF contrast variations is more straightforward than in HRTEM. Thereby, high-resolution HAADF images taken along the $[11\bar{2}0]$ zone axis (i.e. parallel to the QD facets) were used to further investigate the QD morphology at atomic scale (Figure IV-29).

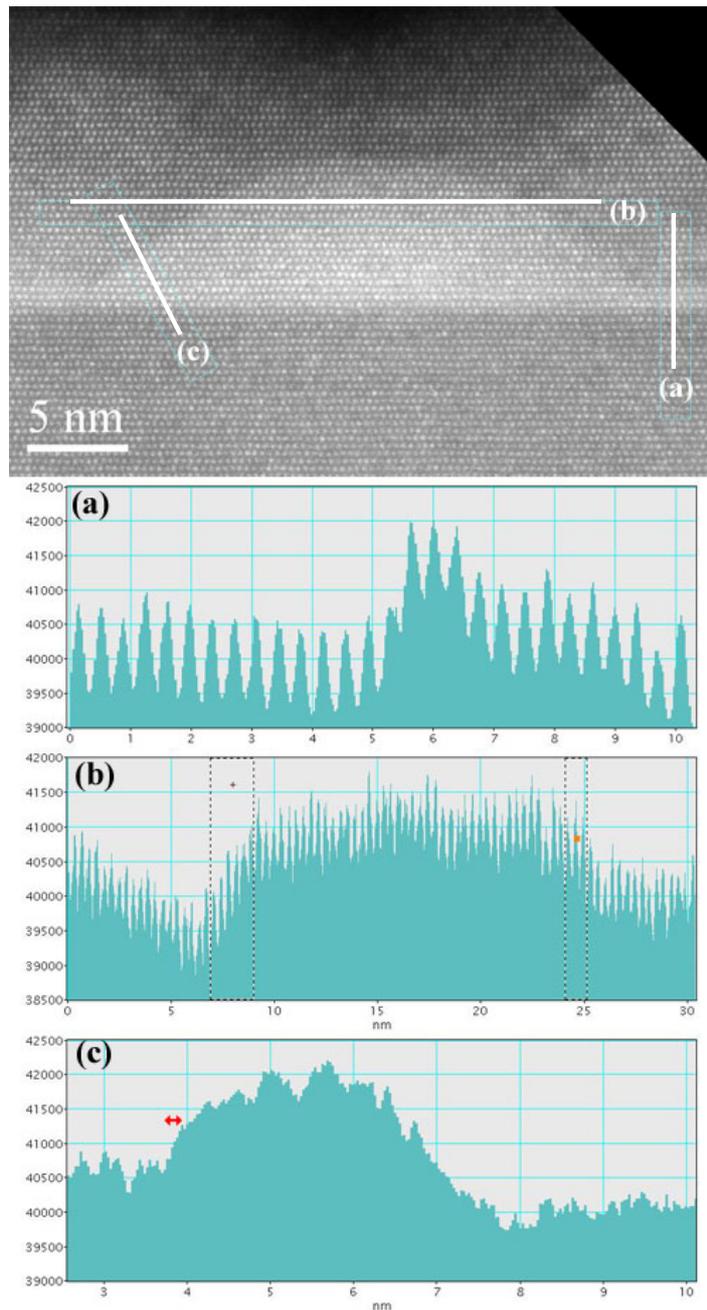


Figure IV-29: High-resolution HAADF image of a buried GaN QD. The white lines show the profiles traced in the $[0001]$ (a), $[1\bar{1}00]$ (b) and $[1\bar{1}03]$ (c) directions. Dash lines on Figure b indicate the projection of the QD facets. The red arrow shows the position of the GaN/AlGaIn interface.

First, the uncertainty in the wetting layer thickness, which is 3 ML deduced from the strain measurements or 4 ML according to HRTEM contrast variations (Figure IV-18 (c)), was eliminated. The HAADF profile (Figure IV-29 (a)) shows that the GaN wetting layer is only three atomic planes thick. Similar value was determined from other HAADF images.

Upper and lower interfaces of the wetting layer are sharp and atomically flat. Besides, the HAADF signal is constant within the wetting layer, which excludes any Al-Ga intermixing.

The possibility of Al and Ga intermixing within the QDs was examined as well. The HAADF signal does not vary across the QD (Figure IV-29 (b)), which means that its composition is uniform. Smooth change of HAADF signal between the GaN QD and the AlGa_N barrier (shown by the dash lines on Figure IV-29 (b)) is because the HAADF profile is not parallel to the [1 $\bar{1}$ 03] direction.

The profile, traced perpendicularly to the {1 $\bar{1}$ 03} facet (Figure IV-29 (c)), shows a sharp decrease of the HAADF signal at that interface (shown by the red arrow). This signifies that the QD composition is homogeneous.

In the next step, multislice HAADF simulations were performed to enable a quantitative analysis of HAADF images [26]. The HAADF intensity ratio of GaN and AlN was calculated as a function of the specimen thickness, taking into account channelling effect. The actual TEM sample thickness was interpolated using regions of known compositions: Al_{0.5}Ga_{0.5}N barriers and GaN QDs.

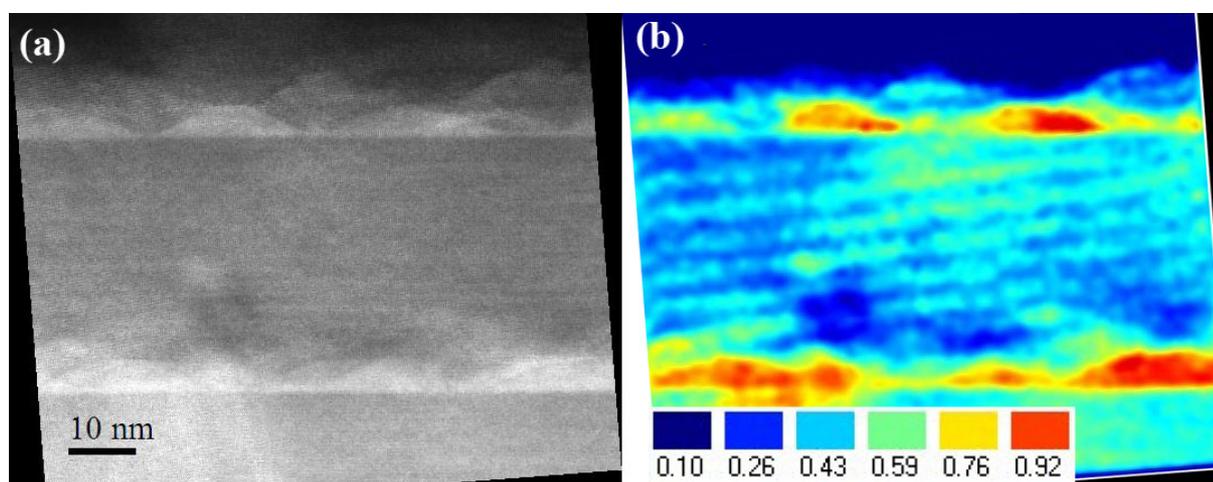


Figure IV-30: HAADF image (a) and Ga-composition map (b) obtained from it.

A Ga-composition map obtained from HAADF image (Figure IV-30 (a)) treatment is shown in Figure IV-30 (b). A precise compositional analysis is hampered by the TEM sample thickness variations, but an averaged concentration of Al in Al-rich zones varying from one QD to another may be estimated between 65% and 75%.

The barrier composition investigations were completed by electron energy-loss spectroscopy (EELS). The use of EELS for concentration determination in AlGa_xN is complicated, since the inner-shell ionisation energies of Ga lie above 1115 eV (L-edge). Consequently, the edge intensity is too weak to be used for a precise quantitative analysis.

However, since Al is a material with free-electron structure, the plasmon oscillations dominate in the low-loss part of the EELS spectrum. The volume plasmon energy is related to the local free-electron density, which varies with the composition.

The EELS experiments were carried out with a TEM equipped with a monochromator, an element indispensable for small energy shift measurements. The plasmon peak position of GaN, Al_{0.5}Ga_{0.5}N and AlN were calibrated using the spectra recorded from buffer layers (upper insert at Figure IV-31 (a)). The plasmon energy was found to change almost linearly with the composition x of Al _{x} Ga_{1- x} N alloy (upper insert at Figure IV-31 (b)) [27].

To examine the spatial distribution of the chemical composition a line-scan was taken across the area shown in the bottom-left inset of Figure IV-31 (a). Four spectra extracted from this scan are presented on Figure IV-31 (a). A variation of the plasmon energy as a function of the probe position is given on Figure IV-31 (b).

It is noteworthy that in the Al-rich zone the energy shift decreases almost linearly, as the distance from the QD increases. A similar observation was done by strain analysis of HRTEM images (Figure IV-22 (c)). The Al concentration extracted from the spectrum recorded right after the QD (plasmon energy 20.5 eV) is around 70%. The measurements accuracy, due to the peak energy determination error (~0.05 eV), is about 3-4%.

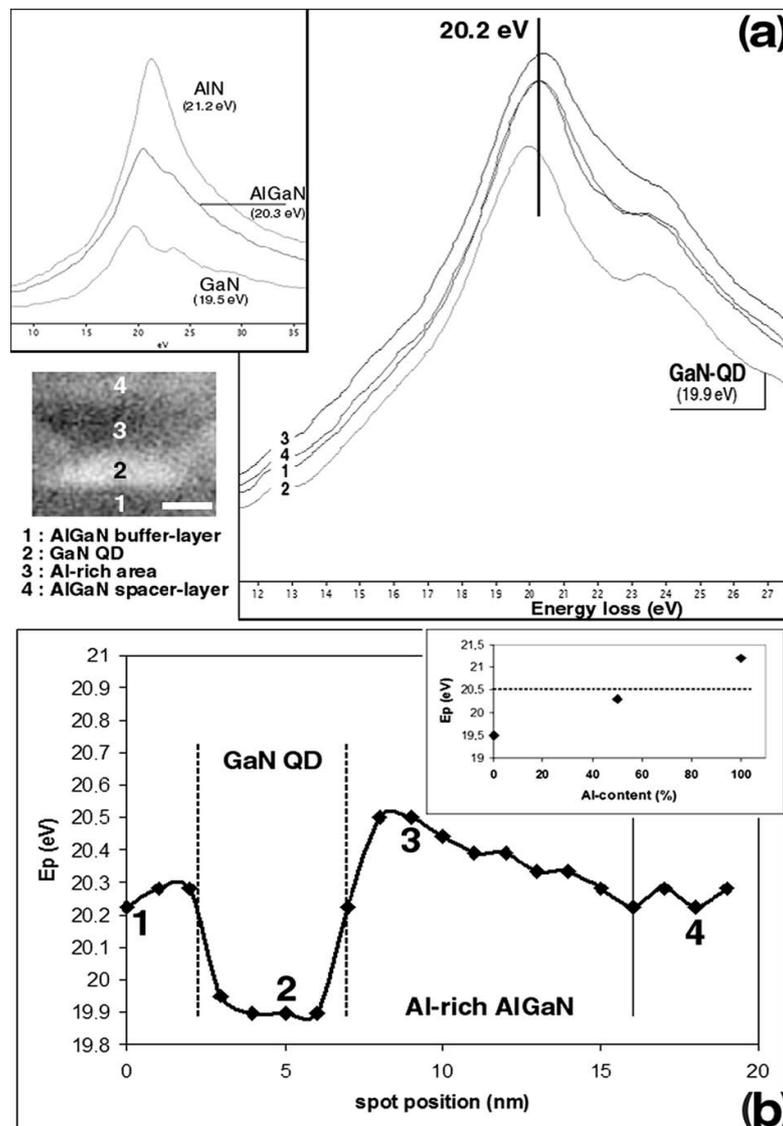


Figure IV-31: (a) Low-loss EELS spectra recorded from four different regions shown in the HAADF image (bottom-left inset). The reference spectra recorded from GaN, AlN and $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ layers are given in the upper-left inset. (b) Plasmon energy as a function of the probe position along a line-scan taken across the area shown in bottom-left inset of Figure IV-31 (a). The upper-right inset shows the plasmon energy as a function of the Al content extracted from the values reported in the upper-left inset of Figure IV-31 (a).

Several conclusions were deduced from the study of the GaN/AlGaN QDs using analytical STEM-HAADF and EELS techniques:

- A phase separation phenomenon was evidenced in the AlGaN barriers. Al-rich zones have a shape and a size similar to that of the corresponding buried QDs. Besides, the Al composition in the Al-rich zones decreases, as the distance from the QD increases.
- The wetting layer and the QDs consist in pure GaN; the possibility of Al-Ga intermixing is excluded.
- The Al concentration in the Al-rich zones, deduced from HAADF images treatment, varies between 65%-75%. The EELS plasmon peak shift analysis gives Al composition of $70\% \pm 4\%$. Both values are in good agreement with the Al concentration previously determined by the strain analysis.

4.6. Study of the QD capping

Previous study of GaN QDs grown on $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ templates revealed a number of original phenomena, which occurs due to the barrier overgrowth. In order to understand the mechanisms involved in the observed effects, a series of samples at different stages of the QDs capping with $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ was investigated.

A second series of samples was grown similarly to the first series, but only the surface QD plane was realized. In all the samples, the nominal GaN deposited thickness of GaN was 8 ML. The surface QD plane was capped with AlGaN layers of various thicknesses: 2 ML, 12 ML, and 24 ML. A sample without capping was realized as well. The peculiarity of the second series is that no AlGaN spacer was deposited after the temperature calibrations. Thereby, the QD plane was grown directly on the Al-rich layer, which involuntarily forms during the AlGaN growth interruption (Section 4.4).

As before, the surface QD morphology was first imaged by SEM. A dense distribution of QDs was observed in all the samples (Figure IV-32).

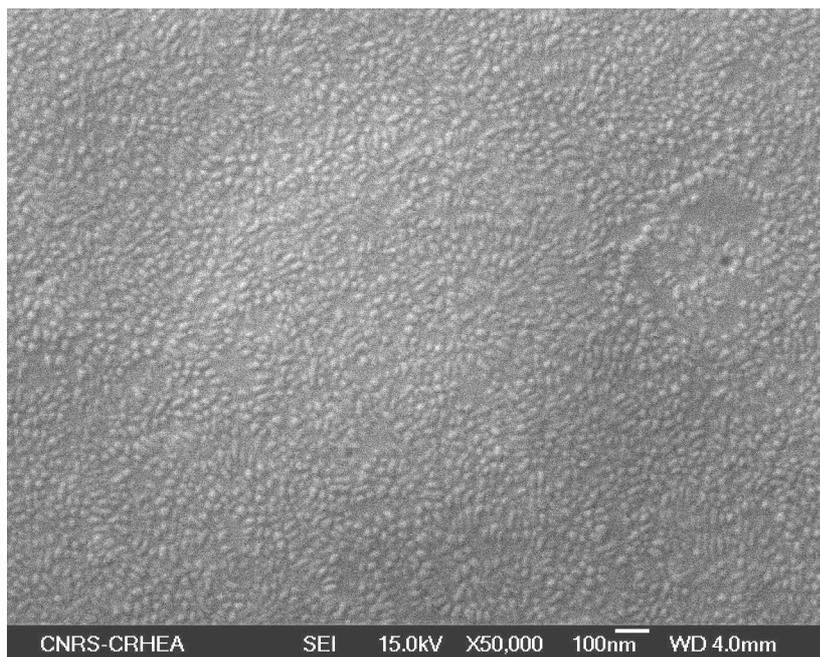


Figure IV-32: Typical SEM image of surface QDs covered with 8 ML of AlGaN.

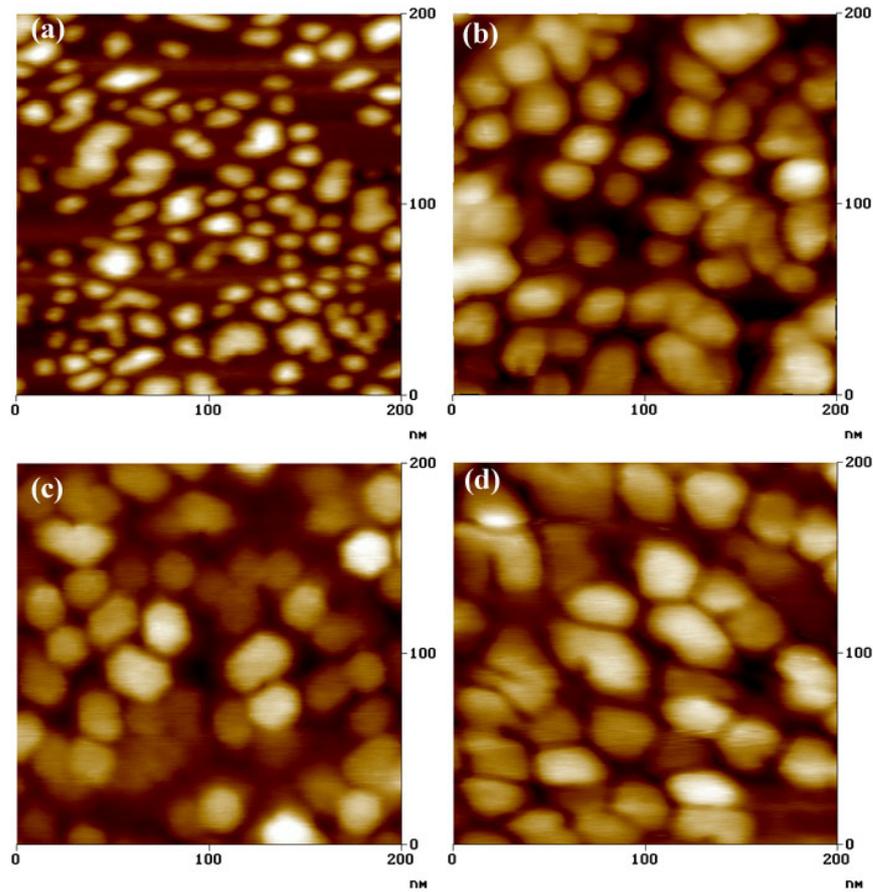


Figure IV-33: AFM images of uncapped QDs (a) and QDs progressively capped with AlGaIn layers of 2 ML (b), 12 ML (c), and 24 ML (d) thick.

Atomic force microscopy (AFM) was used to characterise the shape and size of the covered QD as a function of the capping layer thickness. (Figure IV-33).

The averaged diameter of the bump in the AlGaIn layer formed above the QD increases with the thickness of the capping layer, whereas the dispersion of the AlGaIn bumps size decreases (Table IV-4). Besides, the capping tends to regularize the QD shape: QDs covered with 3 nm of AlGaIn layer (Figure IV-33 (c)), exhibit a more regular hexagonal form than uncapped ones (Figure IV-33 (a)).

As the capping layer thickness increases, the QDs start to coalesce. However, the smoothing effect of the AlGaIn layer deposition is small: QDs remain clearly visible even after 6 nm of AlGaIn (Figure IV-33 (d)). RHEED observations (not-shown) indicate that the capping layer surface becomes flat only after the deposition of about 10 nm of AlGaIn.

The AFM images were also treated by the SPIP software [28] in order to obtain a statistics of the AlGaN bump dimensions. The overgrown QDs were detected by a watershed segmentation algorithm, which determines the positions of the local maxima in the topographic images (Figure IV-34). Some falsely chosen objects, like coalesced QDs, were manually removed from the selection.

Though the algorithm was not able to separate closely adjacent QDs (Figure IV-34 (b)), hundreds of QDs of various sizes were detected in each AFM image. The area of each object was calculated numerically and after that the object diameter was found using formula

$$d = 2\sqrt{S/p} \quad (\text{eq. IV-1})$$

where S is the object area. The averaged diameter of AlGaN bumps is given in Table IV-4.

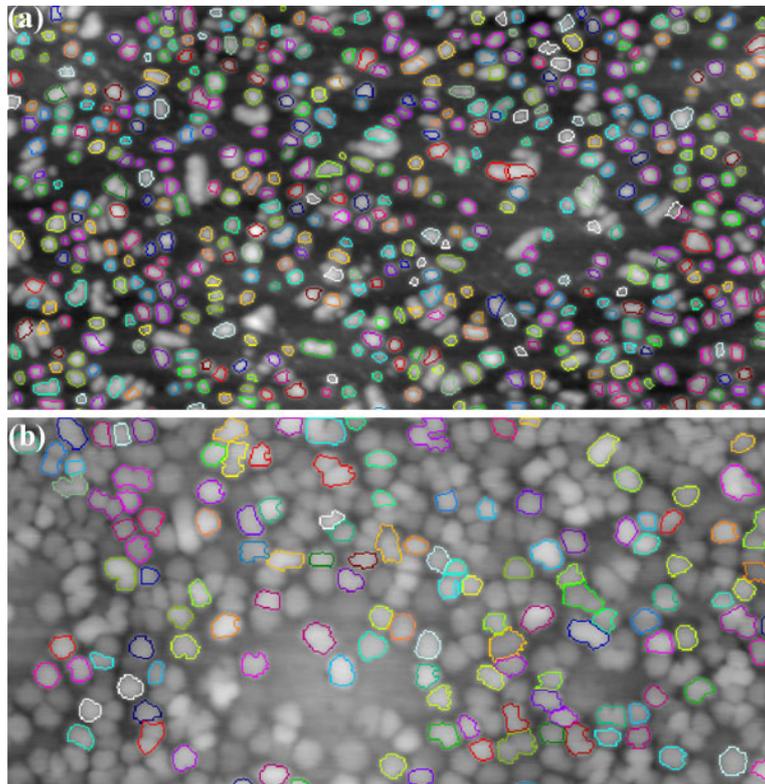


Figure IV-34: Example of detection of object in AFM image by the watershed segmentation algorithm for samples covered with 2 ML (a) and 24 ML (b) of AlGaN.

Sample number	N 1364	N 1365	N 1367	N 1382
Nominal GaN thickness	0 ML	2 ML	12 ML	24 ML
AlGaIn bump diameter, nm	16	18	22	32

Table IV-4: Diameter of AlGaIn bumps measured by AFM.

The QD morphology of the second series samples was studied by STEM-HAADF. All images were acquired along the $[11\bar{2}0]$ zone axis.

In the sample covered by 2 ML of AlGaIn, the change of the surface QD shape from perfect pyramid to truncated pyramid was observed for the majority of QDs (Figure IV-35 (a)). However, several QDs have a perfect pyramidal shape (Figure IV-35 (b)). The AlGaIn layer seems to be situated between the QDs, but it is difficult to prove that by HAADF images.

An Al-rich layer (red arrows on Figure IV-35), which involuntarily forms during the AlGaIn growth interruption, is clearly visible above each QD. In the following, we will call this layer “Al-rich layer” whereas the Al-rich zone above the QDs will be called “Al-rich zone”.

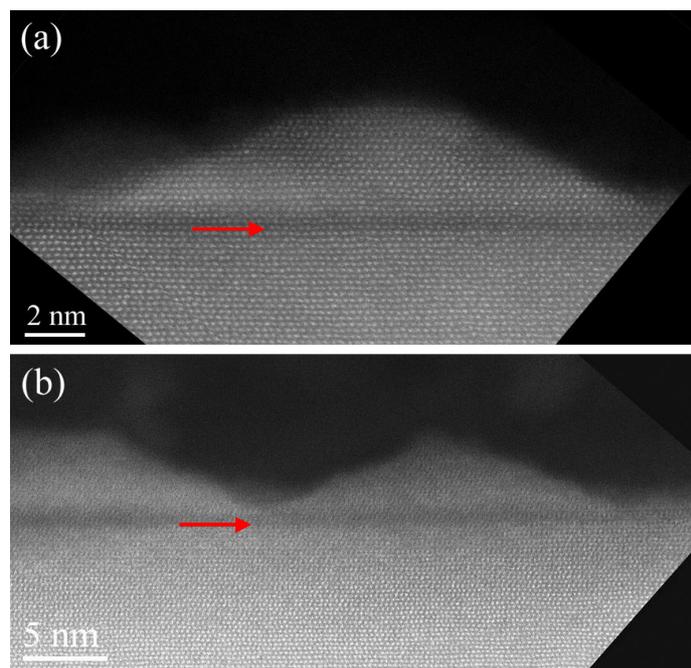


Figure IV-35: HAADF images of sample covered with 2 ML of AlGaIn (a) and (b).

All QDs in the sample with 12 ML of AlGaN have a truncated pyramidal shape. The AlGaN layer covers both the QDs and the wetting layer, but is thicker between the QDs, than above them (Figure IV-36 (a)). An important observation is that the thickness of the AlGaN layer above the QD depends on the QD height: the smaller QD is covered by a thicker AlGaN layer than the bigger one (Figure IV-36 (b)).

The averaged lateral size of the AlGaN bump in the sample with 12 ML is about 30 nm (horizontal arrow on Figure IV-36 (c)), whereas its height is about 2.7 nm (vertical arrow).

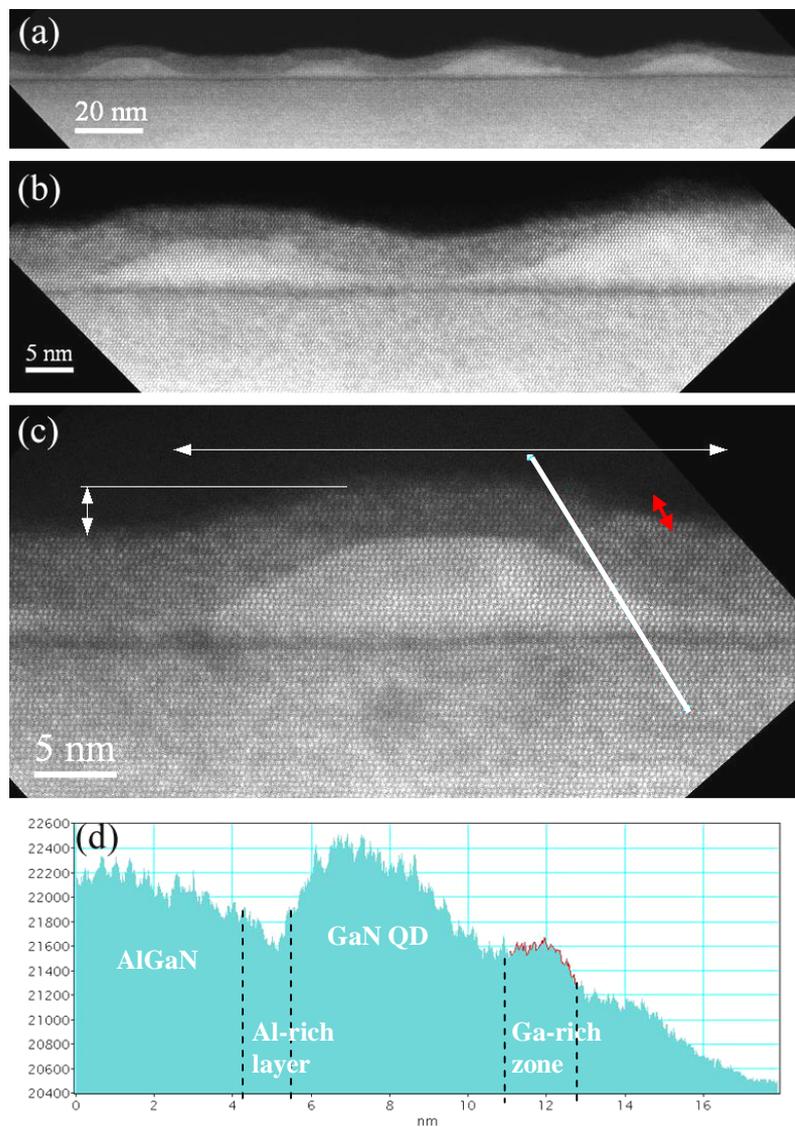


Figure IV-36: HAADF images (a), (b) and (c) of the sample covered with 12 ML of AlGaN. The White arrows indicate the dimensions of the bump in the AlGaN layer above the QD, the red arrow shows the Ga-rich region. The blue line indicates the direction of the profile traced perpendicularly to the Ga-rich zone (d). The red line on figure (d) indicates the increase of the HAADF signal intensity inside the Ga-rich zone.

The Al-rich zones have been already formed in the AlGa_N layer above the small QD, whereas the big QDs are not yet covered by AlGa_N (Figure IV-36 (b)). A Ga-rich zone is visible on the right side of the QD (shown by the red arrow on Figure IV-36 (c)). This Ga-rich “wing” links the top part of the QD lateral facets with the lateral facet of the AlGa_N bump. A profile, traced perpendicularly to the interface between the Al-rich and Ga-rich zones, (Figure IV-36 (d)) shows that the Ga-rich zone is about 2 nm thick.

Similar observations were made in the sample with 24 ML of AlGa_N (Figure IV-37 (a)). The capping layer thickness increases almost equally above the QDs and between them: the mean height of the AlGa_N bump is 2.3 nm. But its average width is about 40 nm, which is noticeably higher than in the previous sample.

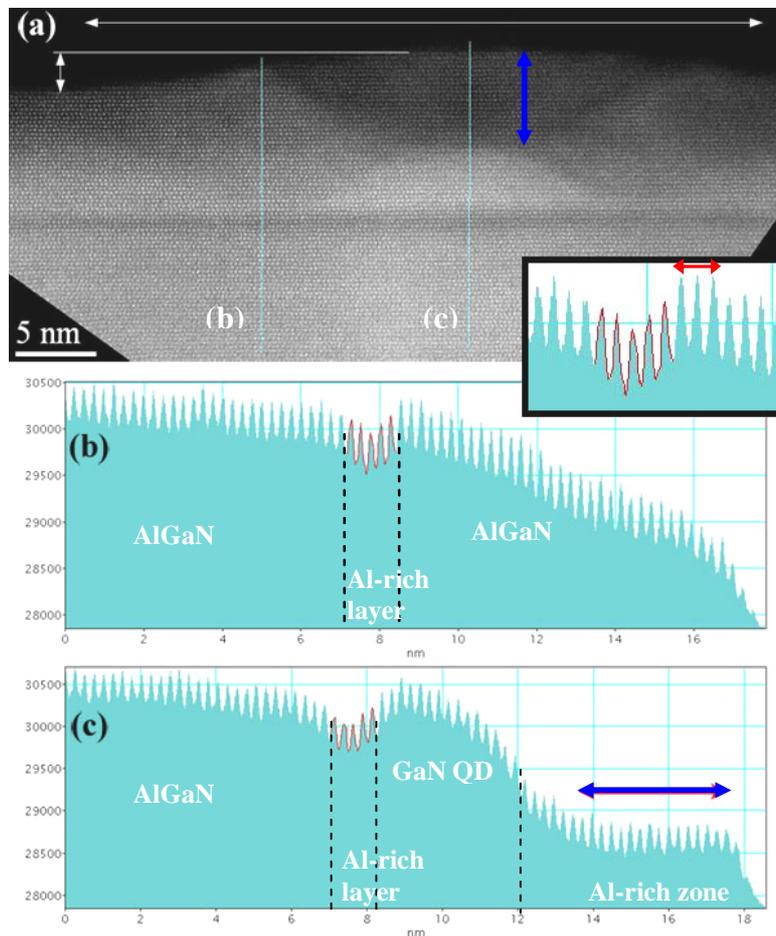


Figure IV-37: HAADF image (a) of sample covered with 24 ML of AlGa_N. The white arrows indicate the dimensions of the bump in the AlGa_N layer above the QD. The blue lines show the direction of the profiles (b) and (c). The red lines indicate the decrease of the HAADF signal inside the Al-rich layer; the blue arrows show the Al-rich zone above the QD.

As the lateral size of AlGaIn bump increases, the Al-rich zone progressively develops inside it (Figure IV-37 (a)). The Ga-rich region borders all-around the Al-rich zone and links the QD with the facets of the AlGaIn bump.

The decrease of the HAADF signal within the Al-rich layer, which lies below the QDs, is visible in the profile traced between the QDs (emphasized in red in Figure IV-37 (b)). Note that the HAADF intensity reaches its minimum at the middle of the Al-rich layer, as it was shown by the strain analysis (Figure IV-20(b)). Three following intensity maxima (shown by the red arrow in the insert of Figure IV-20(b)) correspond to the GaN wetting layer. After that, the HAADF signal linearly decreases with the TEM sample thickness.

Another profile traced inside the QD (Figure IV-37 (c)) exhibits a similar intensity decrease. But the HAADF signal is constant in the last 4 nm up to the sample edge (shown by the blue arrows in Figure IV-37 (a) and (c)). One can suggest that the sample thickness decrease is balanced by a progressive increase of Ga/Al ratio. This observation indirectly confirms the previous conclusion about a gradual decrease of Al-composition in the Al-rich zone (Figure IV-31 (b)).

The study of the QD capping was completed by TEM observations of the sample covered with 8 ML of AlGaIn (N1295). Two different mechanisms of the AlGaIn layer distribution were revealed in this sample.

In some areas, the capping layer accumulates between the QDs, wetting their lateral facets (blue arrows on Figure IV-38 (a)), but there is no AlGaIn layer on the top of the QDs (red arrows on Figure IV-38 (a)). On the other hand, a complete wetting of GaN QDs by the capping layer was observed in other places (Figure IV-38 (b)). In this case, the thickness of the AlGaIn layer above the QD is about 5 ML. Similar, or even higher AlGaIn layer thickness was observed above other QDs.

TEM contrast variations are not distinguishable in the capping layer; that is the reason why the presence of an alloy demixing in the AlGaIn layer was verified by strain analysis. A ϵ_{zz} map (Figure IV-38 (c)) was calculated using a medium size mask, which makes a suitable compromise between the precision in strain determination and the spatial resolution (0.77 nm). Zones of positive and negative strain are evidenced around the QDs, which indicates that the alloy demixing already occurred.

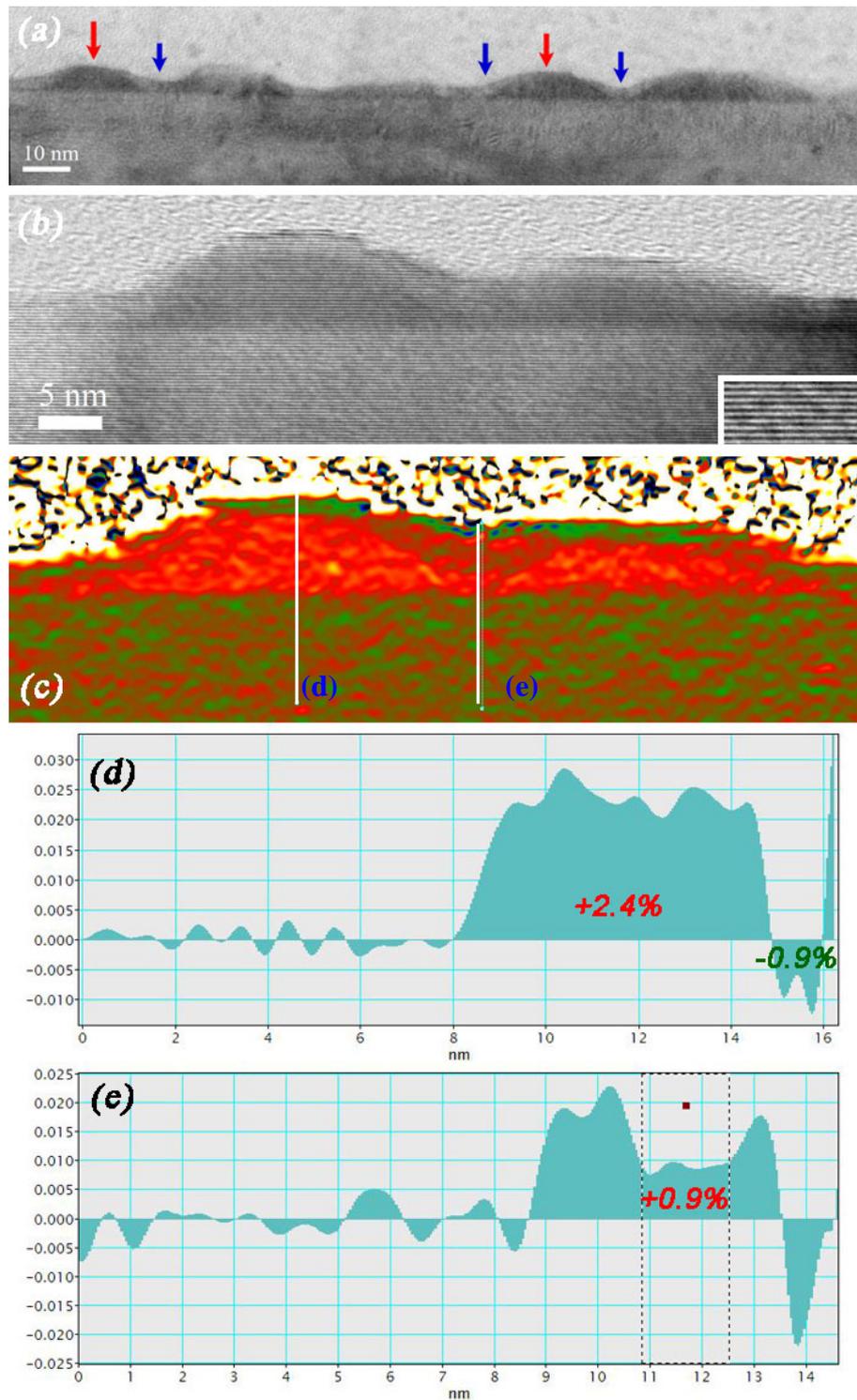


Figure IV-38: TEM image of QDs covered with 8 ML of AlGaIn (a). Coloured arrows emphasize features of the capping layer distribution. Off-axis HRTEM image of two neighbouring QDs (b), the inset is an enlargement showing (002) lattice fringes. The corresponding strain map obtained with GPA (c). The white lines show the direction of the profiles (d) and (e). Dash lines on profile (e) show the zone where the averaged strain value was calculated.

Further strain analysis carried out in the sample covered with 12 ML of AlGaIn confirmed the formation of Al-rich zone above the QD (Figure IV-39 (b)). The ϵ_{xx} map (Figure IV-39 (c)), obtained with a medium size mask, does not exhibit any strain variations above the noise level of 0.1%.

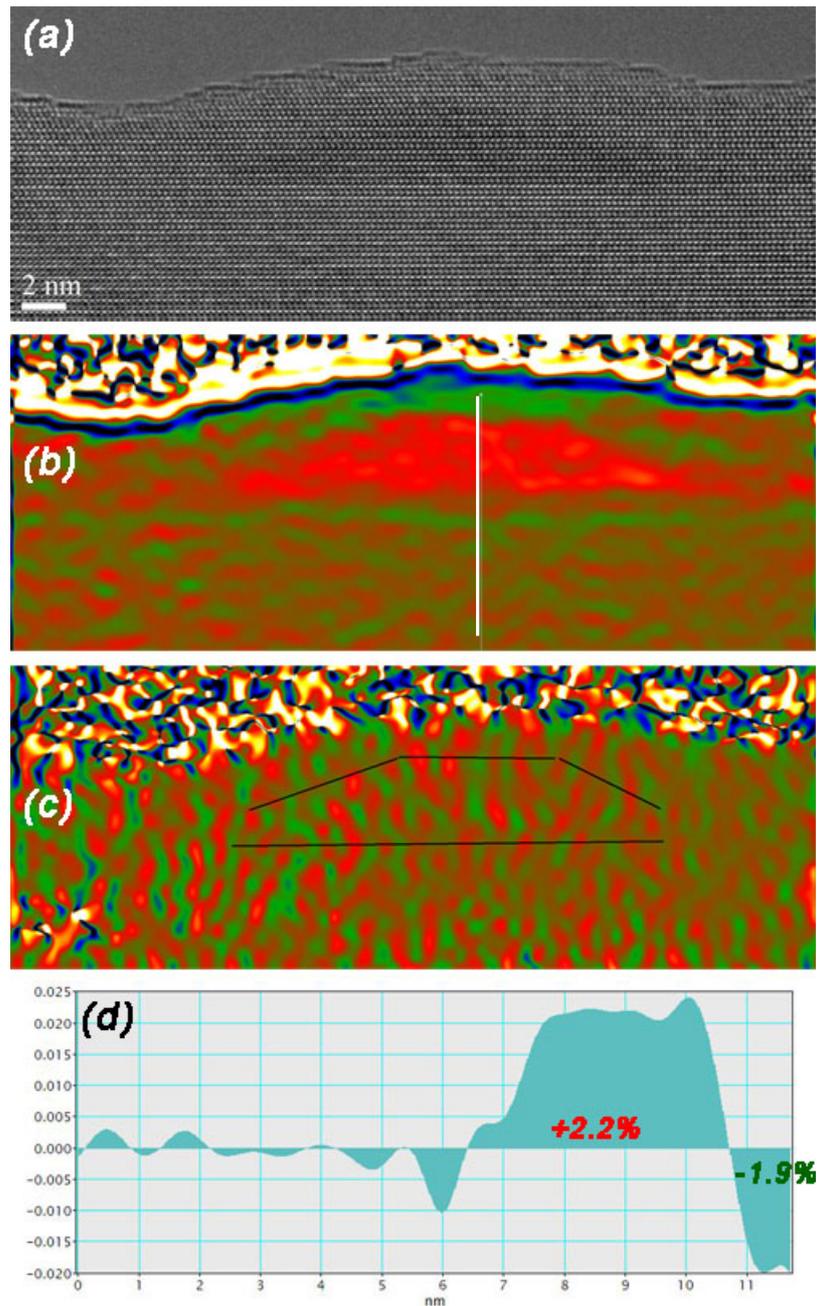


Figure IV-39: HRTEM image of a QD covered with 12 ML of AlGaIn, taken along the $[11\bar{2}0]$ zone axis with the Titan microscope (a) and corresponding ϵ_{zz} (b) and ϵ_{xx} (c) strain maps obtained with GPA (c). The black lines on figure (c) show the QD position determined by ϵ_{zz} maps. The white line on figure (b) shows the direction of the profile (d).

The following conclusions can be summarized from the study of the different stages of QDs capping with $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$:

- The change of the QD shape from perfect pyramid to truncated pyramid happens at the very first stages of the capping process.
- In the first stage of capping, the AlGaN layer grows in the area between the QDs, then it starts to wet their lateral facets. After that the AlGaN layer covers entirely the QDs forming bumps above them. The bump width increases with the AlGaN layer thickness, whereas their height slightly decreases.
- The phase separation occurs as soon as the AlGaN layer entirely wets the QDs. The Al-rich zone forms inside the AlGaN bump, whereas the Ga-rich zones link lateral facets of the QD and the AlGaN bump facets.

4.7. Study of the QDash morphology

The study of the GaN QDs was completed by investigations of anisotropic GaN islands commonly called quantum dashes (QDashes). The typical length of QDash is between 50 nm and 100 nm, whereas its width is similar to that of a quantum dot.

The QDash samples were grown by ammonia assisted MBE in a similar way as the QD structures. A GaN layer was deposited on a 1 μm thick fully relaxed $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template at a temperature of 800°C. The 2D-3D surface reorganization initiated by the growth interruption was followed using RHEED. The anisotropy of self-assembled islands was controlled by changing the ammonia pressure [29]. A stacked structure including three buried QDash planes separated by 30 nm thick $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ barriers was realized; the surface QDash plane remained uncapped. A series of samples with varying nominal thickness of deposited GaN from 7 ML to 18 ML was grown.

Photoluminescence experiments were performed at a temperature of 8 K using a frequency doubled 244 nm Ar laser. A considerable variation of PL energy from 2.9 eV to 2.05 eV within the series of samples was revealed (Figure IV-40) [29].

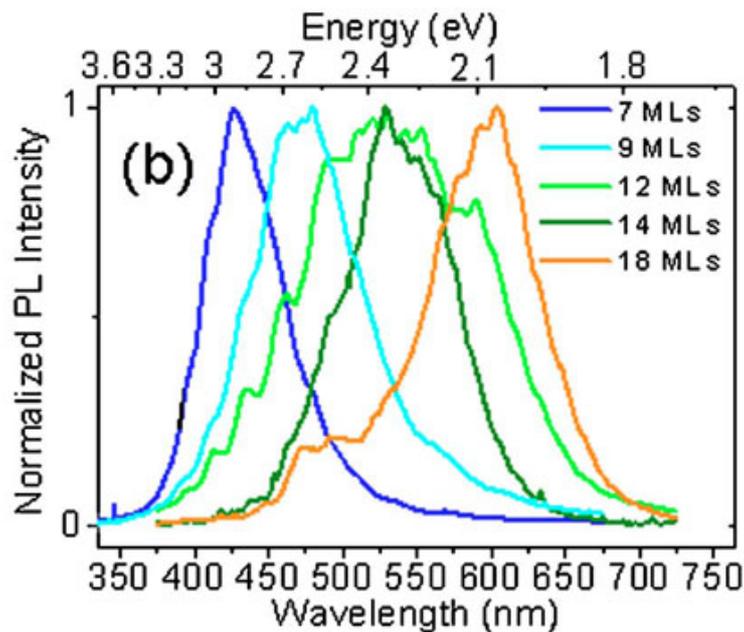


Figure IV-40: Low temperature PL spectra of the QDash samples as a function of the GaN nominal thickness.

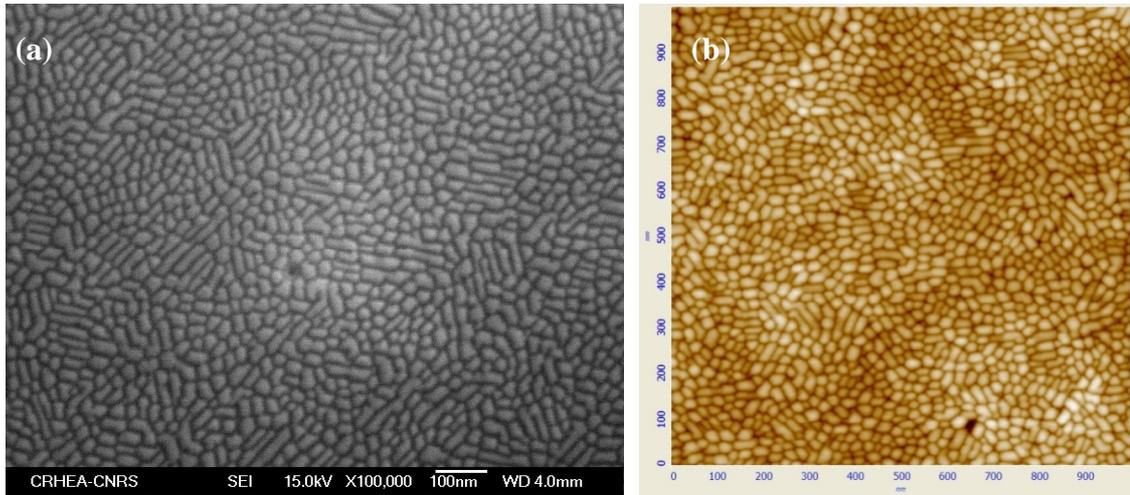


Figure IV-41: SEM (a) and AFM (b) images of the QDash sample with 12 ML of GaN

The decrease of the PL energy with the increase of the nominal GaN thickness is considerably higher than that observed in QD structures, which varies from 2.84 eV to 2.63 eV (Figure IV-4). In fact, the GaN QDash nanostructures can be used to tune the emission wavelength in a wide range of the visible spectrum. This motivates an advanced structural characterisation of QDash samples.

The surface QD-plane morphology was first examined by SEM and AFM. Both methods revealed a density of QDash about $2 \cdot 10^{10} \text{ cm}^{-2}$ in all studied samples (Figure IV-41).

Two of the samples, with 12 ML (N 1275) and 18 ML (N 1270) of nominal GaN thickness, were observed with the JEOL 2010F microscope (Figure IV-42 (a)). Both surface (not-shown) and buried (Figure IV-42 (a)) QDashes have a truncated pyramid shape. Similarly to the QD, the lateral facets of QDash are parallel to the $\{1\bar{1}03\}$ planes.

The QDash dimensions were measured from HRTEM images. The typical QDash height is about 5-6 nm (with a deduction of the wetting layer thickness), whereas their width can be higher than 100 nm. No detectable QDash height reduction due to the capping was noticed.

A phase separation in AlGaIn layer was instantly revealed with Al-rich zones placed above the QDashes and Ga-rich zones in between them (Figure IV-42 (b)). But comparing with the QD samples, the amount of ϵ_{zz} strain is considerably smaller (Figure IV-42 (c)). In the biaxial approximation it corresponds to Al concentration about 60-65%.

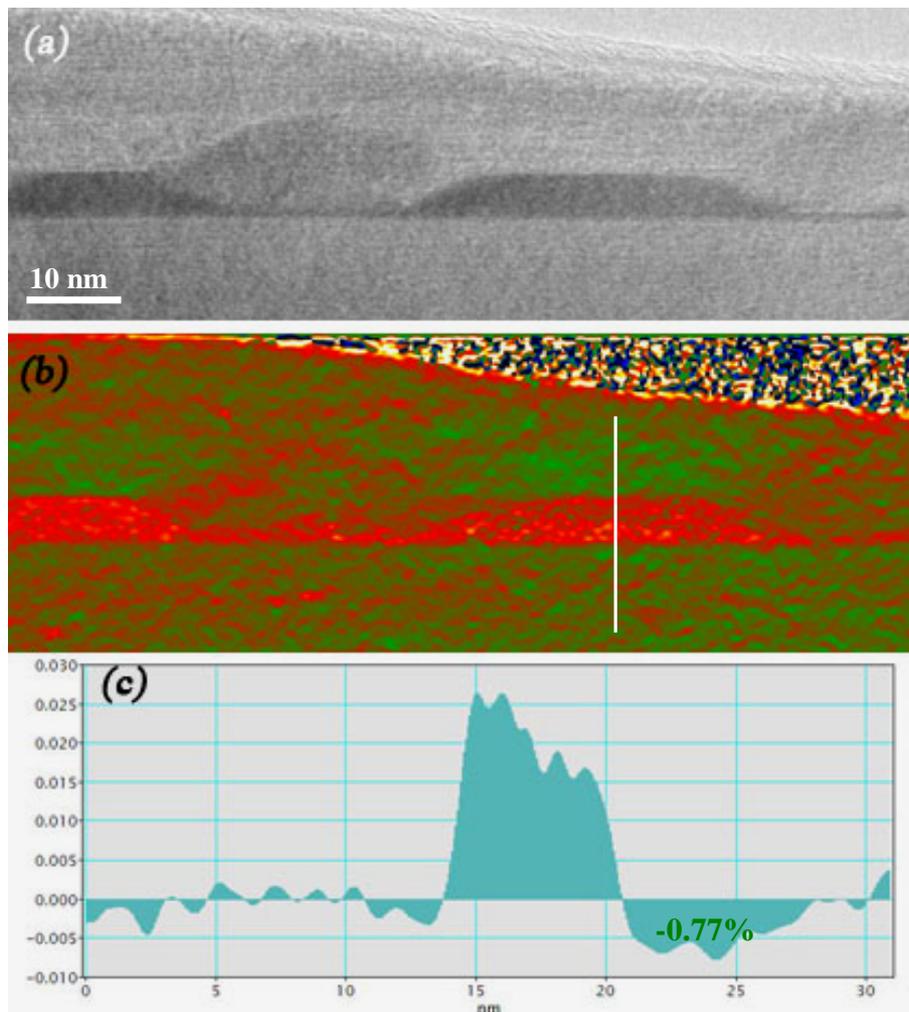


Figure IV-42: Off-axis HRTEM image of two neighbouring QDashes (a). The corresponding ϵ_{zz} strain map obtained using GPA (b). The white line shows direction of the profile (c).

The QDash facets are clearly seen in the images taken along the $[11\bar{2}0]$ direction (Figure IV-43). Moreover, the lateral facets of Al-rich zones are also visible. This is in agreement with the observations of QD samples done by STEM-HAADF (Figure IV-28).

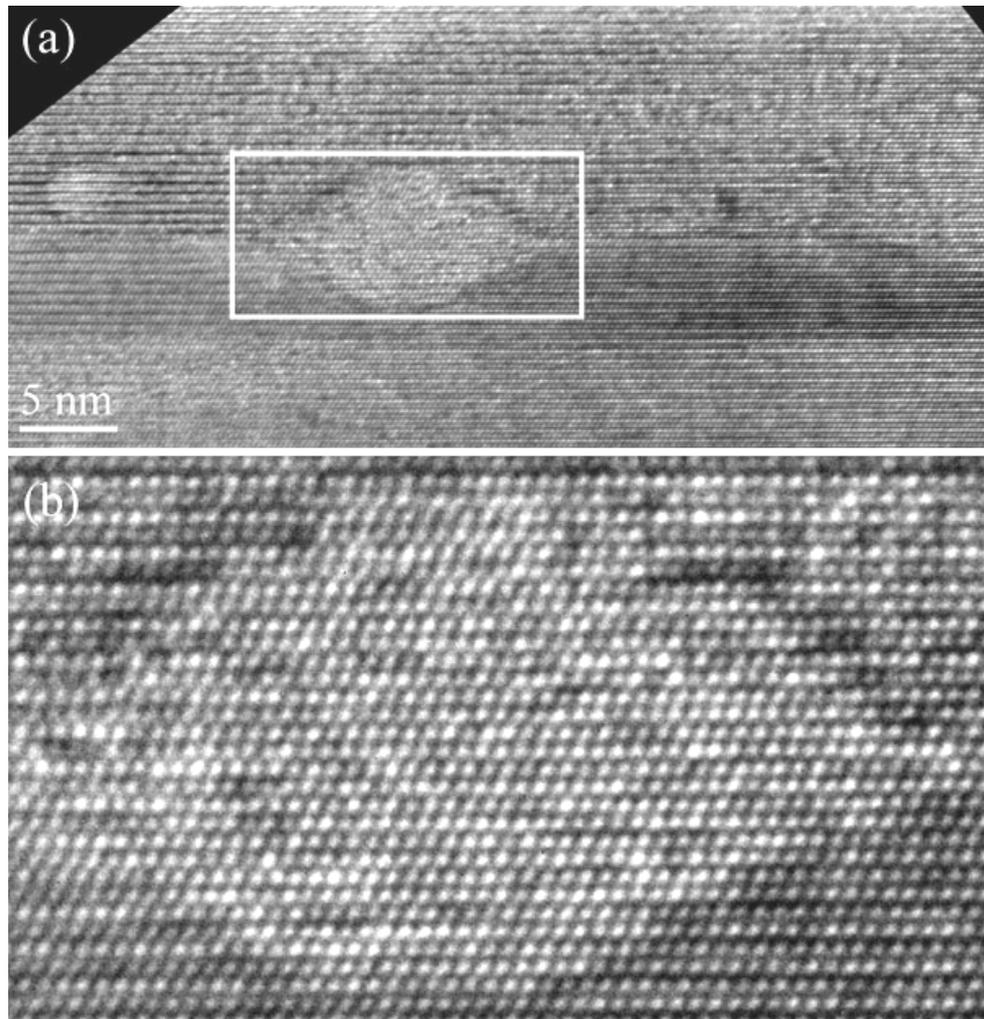


Figure IV-43: HRTEM image of neighbouring QDashes taken along the $[11\bar{2}0]$ direction (a); figure (b) is an enlargement of the zone shown by the white frame.

The study of QDash samples leads to the following conclusions:

- The QDash height and shape do not change with capping.
- An alloy demixing occurs in the AlGaN layer, though Al-composition in Al-rich zones is smaller than in the case of QDs nanostructures.

4.8. Discussion

We shall start the discussion by recalling the main observations done during the study of GaN QDs grown on AlGaN templates:

1. The shape of the surface QDs varies from perfect pyramid to truncated pyramid with the increase of the nominal GaN thickness. The surface QD shape is also governed by the duration of the growth interruption: shorter growth interruption leads to the formation of flatter surface QDs on a thicker wetting layer.
2. The buried QDs always exhibit a truncated pyramid shape with a height/diameter aspect ratio close to 0.17. A change of the QD morphology occurs at the very beginning of the capping process.
3. The averaged volume of buried QDs is bigger than the averaged volume of the surface QDs. A similar tendency was observed in GaN QDs grown on AlN templates.
4. A phase separation occurs in the AlGaN barriers at the very first stages of the capping process. Al-rich zones are formed above the buried QDs, whereas Ga-rich regions develop around the Al-rich zones. Al-rich zones have well defined lateral facets. Besides, the shape and size of the Al-rich zone is similar to that of the QD, above which phase separation occurs. The Al concentration in the Al-rich zones just above the QD is about 70% and it decreases to 50%, as the distance from the QD increases. The Ga composition in the Ga-rich zones is about 55-65%. A phase separation phenomenon was also observed in QDash samples. The Al-composition in Al-rich zones is smaller than in the case of QDs.

Below we will propose possible mechanisms of the observed phenomena, which were described above. The presented models do not pretend to explain in detail the complicated process of the QD growth, but rather to interpret some dominant tendencies. Besides, this discussion may initiate serious theoretical elaborations of the presented hypothesis.

The variety of surface QD shapes (**observation № 1**) can be explained in terms of film's free energy. As discussed above (Section 4.1) the formation of self-assembled QDs via a Stransky-Krastanow transition is governed by the maximisation of the total energy gain [12], which can be described as

$$\Delta E_{\text{tot}} = \Delta E_{\text{elastic}} - \Delta E_{\text{surf}} - \Delta E_{\text{edge}} \quad (\text{eq. IV-2})$$

where $\Delta E_{\text{elastic}}$ is the elastic energy gain due to the QD formation, ΔE_{surf} is the energy of additional QD facets formation, and ΔE_{edge} is the energy accounting for the QD edge formation [30].

The surface energy has a strong impact on the GaN/AlN QD formation [31, 32]. In particular, the GaN quantum dash structures studied in Section 4.7 were formed under high ammonia pressure conditions, which has been theoretically shown to increase the surface energy of GaN [33]. As shown in Appendix C, for a constant volume, the surface of a quantum dash is smaller than the surface of a quantum dot. Thereby the quantum dash formation is the most efficient way for the 2D GaN layer bulk strain relaxation at the lowest surface energy cost [29].

The change of the surface QD shape from perfect pyramid (samples with 8 and 12 ML of GaN) to truncated pyramid (sample with 16 ML of GaN) may also be understood from the same point of view. The energy gain due to the formation of a single QD, neglecting the edges energy ΔE_{edge} , is:

$$\Delta E_{\text{single}} = V \cdot \Delta \varepsilon_{\text{elastic}} - S \cdot \Delta \sigma_{\text{surf}} \quad (\text{eq. IV-3})$$

where V is the QD volume, S is the increment of the QD surface compared to a 2D layer, $\Delta \varepsilon_{\text{elastic}}$ is the density of elastic energy gain due to the islanding, $\Delta \sigma_{\text{surf}}$ is the increase of surface energy density due to the additional QD facets formation. The increase of nominal thickness of deposited GaN results in the increase of averaged QD size (Table IV-2). Supposing that the shape of a QD is independent on its size, we can express the QD volume and surface as

$$V = c_v \cdot d^3, \quad S = c_s \cdot d^2 \quad (\text{eq. IV-4})$$

where d is the diameter of the QD base, c_v and c_s are coefficients which consider the QD geometry. If surface QDs exhibit a closely packed structure with coverage rate ~ 1 (like

on Figure IV-6 (b) and (c)) we can assume that the average distance between the QDs λ is of the same order as their base diameter d . Then the surface QD density is

$$D \sim 1/\lambda^2 \approx 1/d^2 \quad (\text{eq. IV-5})$$

Thereby, the free energy gain per surface unit is

$$\Delta E_{\text{total}} = \Delta E_{\text{single}} \cdot D \sim (d^3 \cdot \Delta \varepsilon_{\text{elastic}} - d^2 \cdot c_s/c_v \cdot \Delta \sigma_{\text{surf}}) / d^2 \quad (\text{eq. IV-6})$$

Now, we can compare the free energy of a QD having a perfect and a truncated pyramidal shape. We suppose that the bulk energy $\Delta \varepsilon_{\text{elastic}}$ relaxes more efficiently in the truncated pyramid due to appearance of the top $\{0001\}$ facet. Moreover, we do not consider any possible change of the surface energy density $\Delta \sigma_{\text{surf}}$ due to the shape transition. However, the total energy of additional QD facets formation ΔE_{surf} is proportional to the QD shape. As shown in Appendix C, a perfect pyramid with $\{11\bar{0}3\}$ facets has a minimum c_s/c_v ratio as compared to any pyramid having a truncated shape. Thereby, the role of surface energy ΔE_{surf} is smaller for a QD having a perfect pyramidal shape than for a truncated QD, which has the same volume.

The variation of the total energy gain (eq. IV-6) with the QD size is illustrated in Figure IV-44. The energy gain ΔE_{total} increases linearly with the QD size for both types of QD shapes. But QD having a perfect pyramidal shape reaches a positive energy gain for smaller QD size (d_0). Thereby, small QDs have a perfect pyramidal shape.

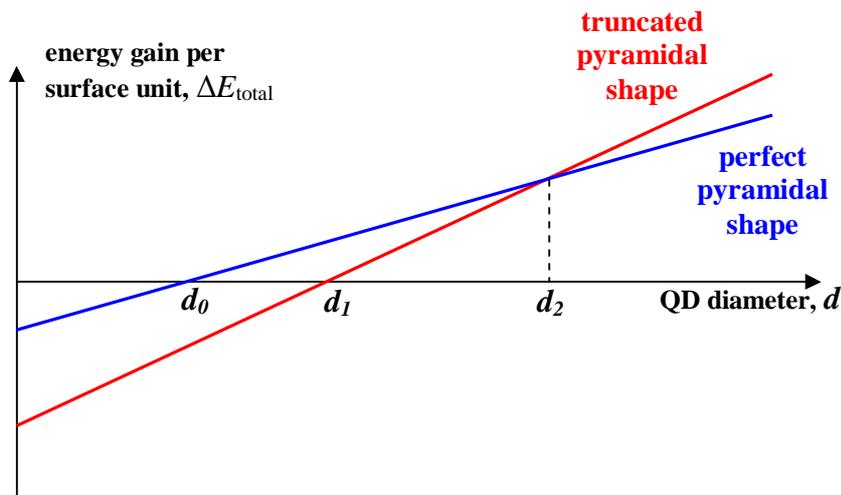


Figure IV-44: Energy gain per surface unit as a function of QD diameter for different QD shapes.

Above a critical size (d_1), QD having a truncated shape also reaches a positive energy gain, which is smaller than that of QD having a perfect pyramidal shape. However, the high dispersion of the QD sizes makes possible the co-existence of two QD types within the same sample. Above a certain QD size (d_2), the energy gain of truncated pyramidal QDs exceeds that of perfect pyramidal QDs; after that the energy gain difference will further increase with the QD size.

We may also use equation IV-6 to explain the QD size and shape in other QD structures. For example, GaN QDs grown on AlN template by ammonia assisted MBE are smaller (Table IV-3) than that grown on $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ templates (Table IV-2). Besides, the minimum thickness of the GaN layer required to initiate the GaN QD formation on an AlN template (3 ML) [2] is twice lower than that required for the GaN/AlGa_{0.5}N QD formation (6 ML) [18]. Indeed, the lattice mismatch between GaN and AlN is twice higher than between GaN and $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$, i.e. the role of the bulk energy relaxation becomes more important. Thus, the dependence of the free energy gain versus the QD size (Figure IV-44) has a steeper slope. Consequently, the critical size for the QD formation (d_0) is smaller than for GaN/AlGa_{0.5}N QDs. Small GaN/AlN QDs exhibit a perfect pyramidal shape (Figure IV-14 (a)). However, similarly to the case of GaN/AlGa_{0.5}N QDs, a shape transition is expected with the increase of the GaN/AlN QD size (above d_2).

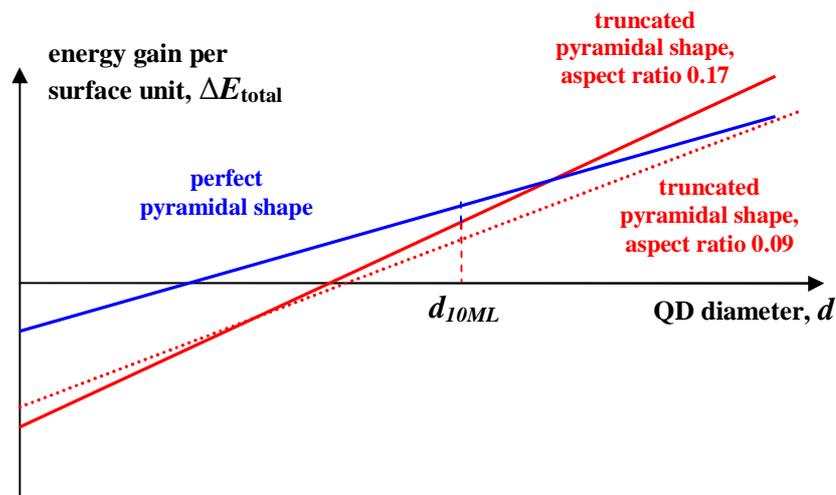


Figure IV-45: Scheme illustration of the gradual evolution of the QD shape in the unfinished sample.

On the other hand, GaN QDs grown on AlN template by plasma-assisted MBE have a truncated pyramidal shape [34]. The QD formation in plasma-assisted MBE follows a true Stranski–Krastanow growth mode and do not require a growth interruption as in the case of GaN/AlN QDs done by ammonia assisted MBE. This implies that the role of the surface energy for the 2D/3D transition is smaller in the case of plasma-assisted MBE. Thereby, surface QDs have a truncated pyramidal shape, which gives a higher gain in elastic energy at a lower cost of surface energy increase. Interestingly, the height/diameter aspect ratio reported for GaN/AlN QDs grown by plasma-assisted MBE (0.19) [32, 34] is close to that of GaN/AlGaN QDs done by ammonia assisted MBE (0.17).

Our basic argumentations suit only for a qualitative explanation of the QD shape transition. More sophisticated calculations [12] would be required to find the minimum thickness of GaN layer required to initiate the GaN QD formation or to evaluate which aspect ratio gives the maximum energy gain. However, in the framework of the same model (eq. IV-6), we can clarify the flat shape observed for the surface GaN/AlGaN QDs in the unfinished sample (10 ML of GaN). If the growth interruption would be sufficiently long, surface QD should have a shape of perfect pyramid, as in the samples with 8 ML and 12 ML of GaN. But the GaN/AlGaN QD formation is a time-expanded process and the QD shape varies gradually. As the QD shape aspect ratio changes, the energy gain progressively rises (Figure IV-45). Red dotted line corresponds to a truncated pyramidal shape with an aspect ratio of 0.09, which gives a slight increase of the surface energy, but does not allow a maximum elastic energy relaxation. In the unfinished sample, the surface QDs (Figure IV-13) were “frozen” at this intermediate stage of their development towards the equilibrium QD shape.

To explain the phenomena related with the AlGaN barrier overgrowth (**observations № 2, 3, 4**), we should first state the dynamics of the QDs burying. By doing this, we presume that the capping layer growth is also governed by the total energy minimisation (eq. IV-2).

Due to the elastic energy relaxation, the local lattice parameter a inside the surface QDs is higher than that in the wetting layer (Figure IV-26 (c)). Consequently, in the first stage of the burying, AlGaN accumulates between the QDs (Figure IV-46 (a)), where the lattice parameter a is the same as in the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template. Above a certain thickness, the AlGaN layer starts to wet the lateral QD facets (Figure IV-46 (b)), decreasing thereby the surface

energy ΔE_{surf} . Then, AlGaN completely covers the QDs forming hexagonal bumps above them (Figure IV-46 (c)).

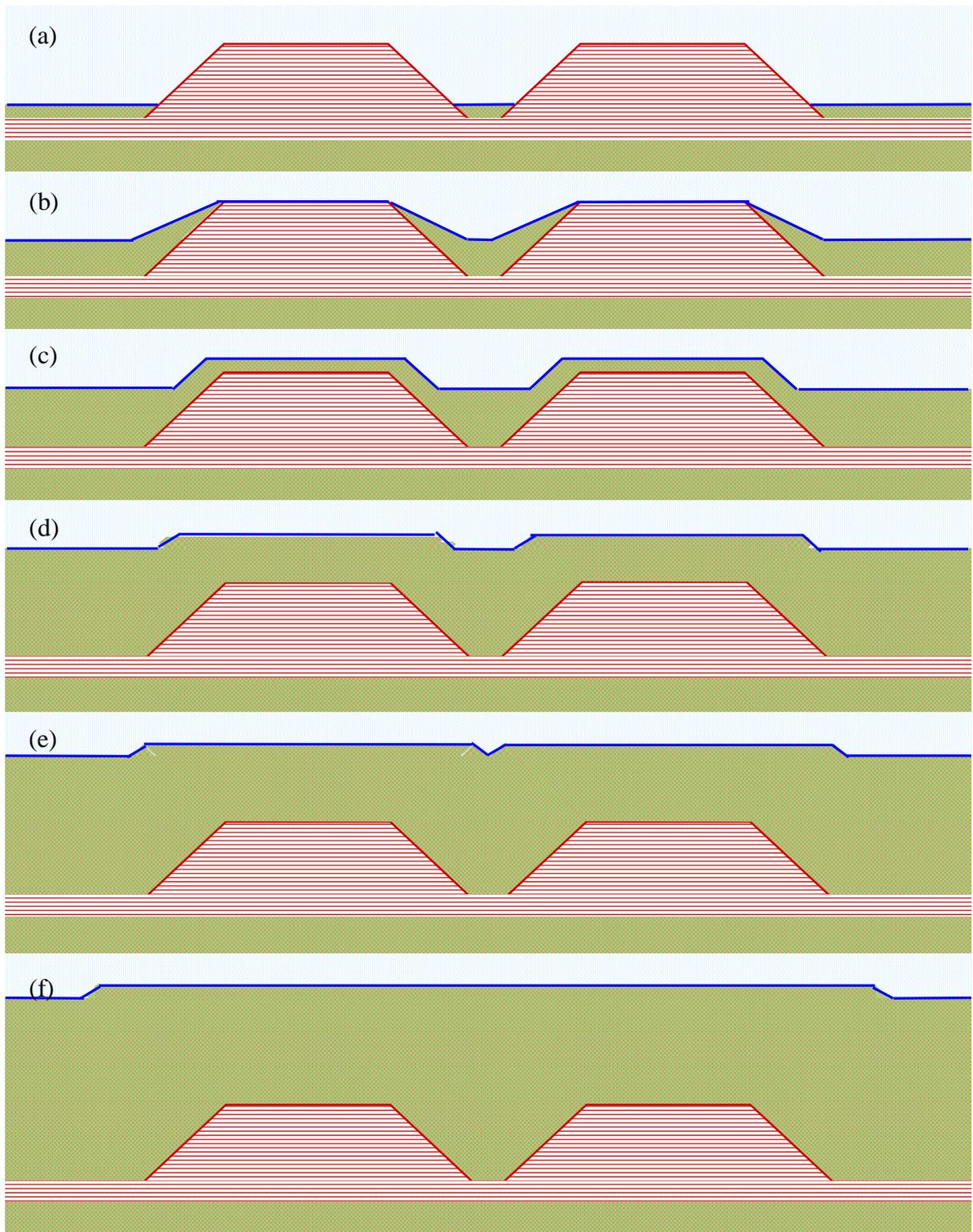


Figure IV-46: Dynamics of the QDs capping with an AlGaN layer. The GaN QDs and wetting layer are in red, AlGaN is in light brown; blue line shows the surface of the AlGaN layer.

Interestingly, the AlGaIn bumps exhibit a more regular hexagonal shape (Figure IV-33 (c)) than uncovered QDs (Figure IV-33 (a)). The regularization of the AlGaIn bump shape may also be governed by the surface energy minimisation. In the following stages of the burying, AlGaIn also tends to lie between the QDs (Figure IV-46 (d)), smoothing thereby the QD plane relief and minimising the capping layer surface. The AlGaIn bumps formed above each QD spread laterally as the capping layer thickness increases (Figure IV-46 (e)). The gradual enlargement of AlGaIn bumps results in their coalescence (Figure IV-46 (f)). RHEED observations (not-shown) indicate that the complete flattening of the AlGaIn layer surface occurs after the deposition of about 10 nm of AlGaIn.

The size of the AlGaIn bumps determined by TEM (Figure IV-36 (c) and Figure IV-37 (b)) is noticeably higher than the average bump diameter measured by AFM (Table IV-4). This may signify that the value obtained by AFM images processing underestimates the actual AlGaIn bump size.

The capping of GaN/AlN QDs grown by plasma-assisted MBE occurs in a different way and it includes only two distinguishable stages [35]. First, an AlN layer covers homogeneously the QDs and the space between them. Then, above coverage of 3-4 ML, AlN preferentially grows between the dots, gradually flattening QD plane relief. Above 18 ML of AlN coverage, the surface of the AlN layer becomes flat. No QD morphology changes due to the burying were found. Only an isotropic QD size decrease after the capping (i.e., with a conservation of the aspect ratio) was reported.

Now, we can propose a mechanism for the change of the QD morphology from a perfect pyramidal shape to a truncated pyramidal shape (**observation № 2**), which takes place at the first stages of the capping. The reduction of the QD height and the change of the QD shape due to the QD burying have been already reported for cubic semiconductors, either in the Ge/Si system [36, 37] or for III-V materials [38, 39]. In the case of InAs/GaAs, a noticeable QD height decrease was observed after the deposition of 1-3 ML of GaAs [40]. The redistribution of the InAs from the QD top was caused by the migration of Ga atoms deposited on the QD apex (Figure IV-47). Ga-rich regions formed along the QD facets represent advantageous sites for InGaAs alloy formation due to entropy gain and strain energy release [41].

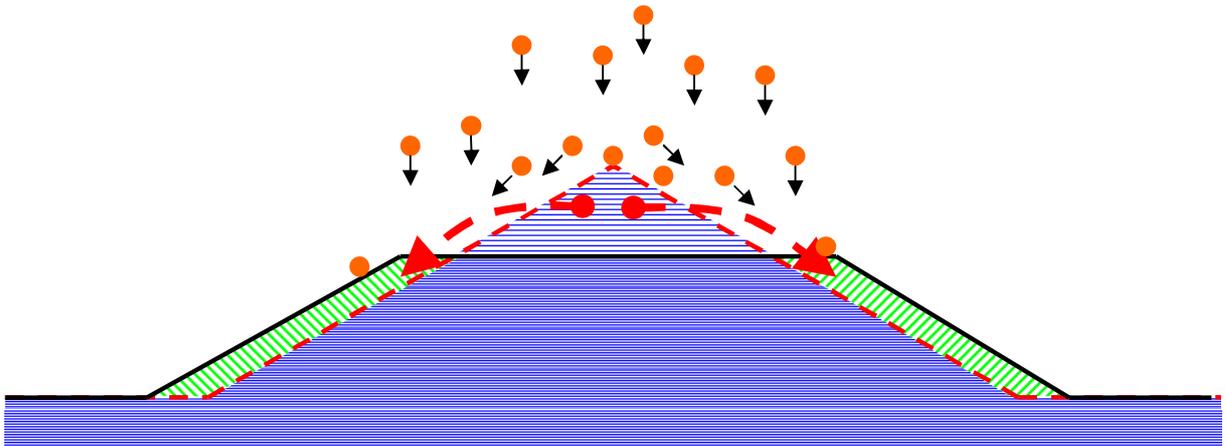


Figure IV-47: A possible mechanism for the QD apex removal. The orange circles represent atoms of the deposited material, the blue matrix is a QD and the wetting layer. The green region represents an alloy formed on the QD facets.

The GaN QD height reduction measured in the samples with 8 ML and 12 ML of GaN is 0.6 nm - 0.8 nm (Table IV-2), i.e. about 3 ML of GaN. The redistribution of GaN from the QD apex may occur in a similar way as in the case of InAs/GaAs QDs. However, the formation of an alloy on the QD facets is excluded by STEM-HAADF observations (Figure IV-29 (b) and (c)).

A trustworthy mechanism of the buried QD shape evolution should consider the surface kinetics of atoms during the AlGaIn barrier overgrowth. Here, we will just illustrate it by means of the principle of total energy minimization.

Due to the elastic energy relaxation, the local lattice parameter a reaches its maximum in the QD apex (Figure IV-26 (c)). Being covered with the capping layer, this area would have a maximal lattice mismatch with overgrown AlGaIn. Thereby, the QD apex removal eliminates the potential creation of a region of high strain. The lateral facets of QDs are advantageous sites for the deposition of Ga atoms, since their local lattice parameter a is a bit larger than in the capping layer.

Another general argument for the vanishing of the QD apex is that the surface Ga atoms are bound to a smaller number of nitrogen atoms as compared with the bulk Ga atoms. The atoms placed at the facets and the edges of the QD apex have maximum number of broken chemical bonds; thereby they are weakly fixed at their position. The interaction with atoms deposited on the QD apex during the overgrowth of the AlGaIn barrier may pull them out from the crystalline lattice.

The study of the sample with the unfinished QD formation (10 ML) revealed some tendencies of the GaN kinetics during the QD formation, which clarifies the QD volume increase due to the capping (**observation № 3**). Indeed, though the surface QDs in this sample are flattened with an aspect ratio as low as 0.09, the buried QDs are well-developed with an aspect ratio (0.15) close to the one observed in the other GaN/AlGaN QD samples (0.17). Besides, the height and volume of the buried QDs are almost 70% higher than those of the surface QDs (Table IV-2). At the same time, the wetting layer in the buried QD planes is much thinner than the 2D GaN layer lying below the surface QDs plane (Figure IV-13).

This reveals the existence of a mass transport of GaN from the wetting layer to the QDs, which can be driven by a further relaxation of the elastic energy stored in the GaN layer. We can assume that, in the other samples, the QDs development was not also “frozen” by the barrier overgrowth. The mass transport of GaN from the wetting layer may maintain the rise of the buried QDs volume at the first stages of the capping (Figure IV-48), whereas a fast sample cooling renders impossible a similar effect for the surface QDs.

Indeed, it is enough to redistribute one monolayer of GaN among all QDs (considering a typical QD density of $1.5 \times 10^{11} \text{ cm}^{-2}$) to produce an average volume augmentation of 170 nm^3 . This value is comparable with the increase of the QD volume, which varies from 90 nm^3 to 300 nm^3 depending on the sample (Table IV-2). Such a small decrease of the wetting layer thickness between the surface and buried ones should not be evidenced experimentally.

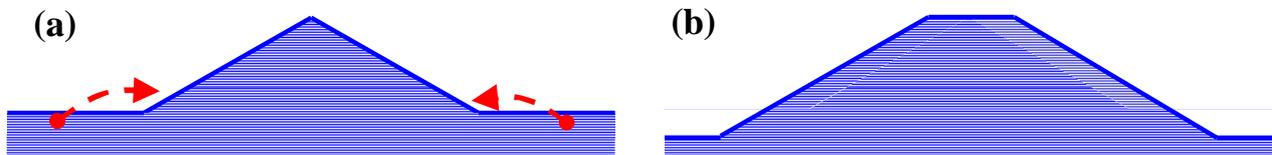


Figure IV-48: At the first stage of the barrier overgrowth, the QDs are placed on a thick wetting layer (a). The QD volume rises and the wetting layer becomes thinner according to the mass transport of GaN (b). The blue matrix is a GaN QD and the wetting layer, the red arrow show the displacement of the GaN material.

The phase separation in the AlGaN spacers (**observation № 4**) may be considered separately from the buried QD morphology change and volume increase. Firstly, because the morphology change and the volume increase occur in the GaN QDs buried by AlN layers, where an alloy demixing is excluded. Secondly, because Al-rich and Ga-rich zones were identified in the sample with 16 ML of GaN where no morphological changes are observed. Besides, the phase separation was found even in the QDash samples, whose morphology completely differs from that of the QDs.

To propose a model for the phase separation phenomenon, we should consider the atomic structure of a thin AlGaN layer, which wets a QD (Figure IV-46 (c)). At this stage, AlGaN already filled the space between the QDs. This considerably reduces the effect of a surface QD relaxation and thereby, the local lattice parameter a inside the surface QDs decreases. In a further reasoning, we consider that the lattice parameter a in the GaN QD is equal to the lattice parameter a of the AlGaN lying between the QDs (i.e. the a parameter of relaxed $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$). This supposition is in agreement with the ε_{xx} strain measurements which show that the QD and the wetting AlGaN layer are coherently strained on the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ template (Figure IV-39 (c)).

However, the local lattice parameter c in the GaN QD is larger than that in the AlGaN situated between the QDs (shown by the red arrows on Figure IV-49 (a)). Thereby, the surface atomic plane (shown by the black line) is slightly bended above the QD. The first monolayer of AlGaN grown on the top of that plane will be bended as well. Due to the displacement of the capping layer atoms situated above the QD, the adjacent chemical bonds in the AlGaN plane will be stretched (shown by the red lines). Consequently, the atoms lying above the QD will be slightly displaced towards the QD (blue arrows on Figure IV-49 (a)) in order to decrease the energy of the stretched bonds. As a result, the first atomic plane of the capping layer (the grey line on Figure IV-49 (b)) will be a bit less bended than the underlying atomic plane (black line), i.e. the displacement of the capping layer atoms weakens the strain introduced by the QD.

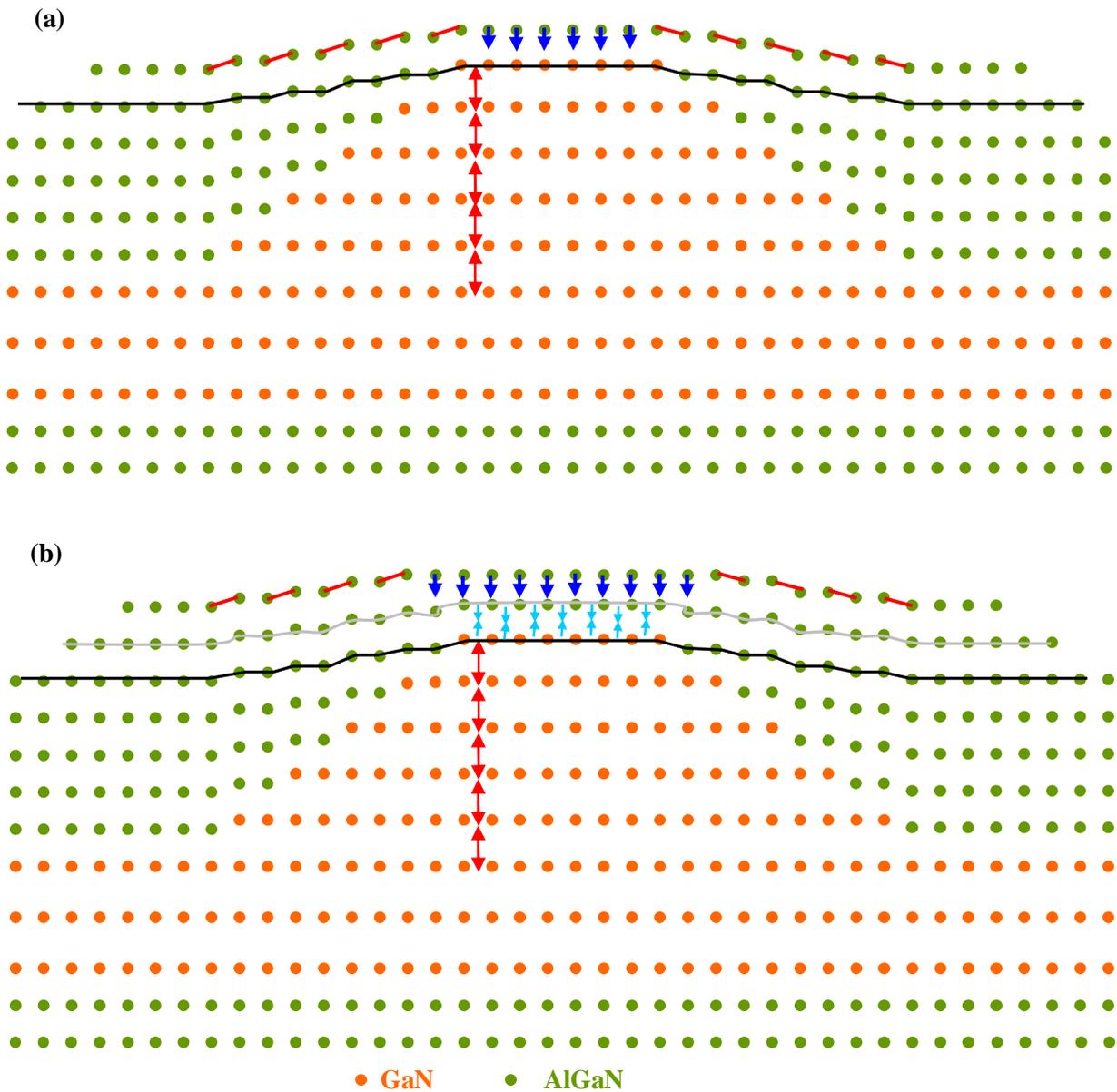


Figure IV-49: A possible atomic structure of a GaN QD and surrounding AlGaN layer. Each circle indicates an atomic position. The atomic displacements are exaggerated for a better visualization. The coloured symbols are referred in the text.

The second AlGaN monolayer, deposited above the previous one, will exhibit a similar reorganization (blue arrows on Figure IV-49 (b)). As a result, a compressive ϵ_{zz} strain will be accumulated in the central part of the AlGaN bump (shown by the light-blue arrows on Figure IV-49 (b)). The amount of deformation will decrease with the AlGaN layer thickness due to the progressive relaxation of the atomic plane bending. The final distribution of the compressive strain in a thick AlGaN layer is described in the simulated ϵ_{zz} strain tensor component (Figure IV-24 (c)).

Note also that the width of the AlGaN bump above the QD increases laterally with the AlGaN layer thickness (Figure IV-49 (b)), which is in agreement with the previous observations.

Thereby, the thin AlGaN layer lying above a {0001} QD facet is compressed (green area on Figure IV-50 (c)), whereas the AlGaN placed above the lateral QD facets is stretched (red areas on Figure IV-50 (c)). Such a strain distribution may cause a stress-driven phase separation [42] governed by the minimization of the elastic energy accumulated in the AlGaN layer. Al-rich regions will be formed in the zones of compressive strain, since the lattice parameter c of AlN is smaller than that in AlGaN whereas, the regions of tensile strain will be Ga-rich.

The stress-driven phase separation will occur in every new deposited AlGaN monolayer until a complete relaxation of the strain induced by the QD (Figure IV-50 (d)). The Al-composition in the Al-rich zones will decrease as the strain relaxes (Figure IV-50 (e)), which is in agreement with our results (Figure IV-31 (b)).

A phase separation occurs in a similar way in the QDash samples. The volume of the QDashes is larger than the volume of the QDs; the compressive effect of the overgrown AlGaN is therefore smaller in the QDash case. We can suppose that due to this, the lattice parameter a in the QDash is slightly larger than in the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ layer. The effect of the negative ε_{zz} strain existing above the QDash will be partially compensated by this positive ε_{xx} strain. This may explain why the Al composition in the Al-rich zones above the QDashes is smaller than in the case of QDs

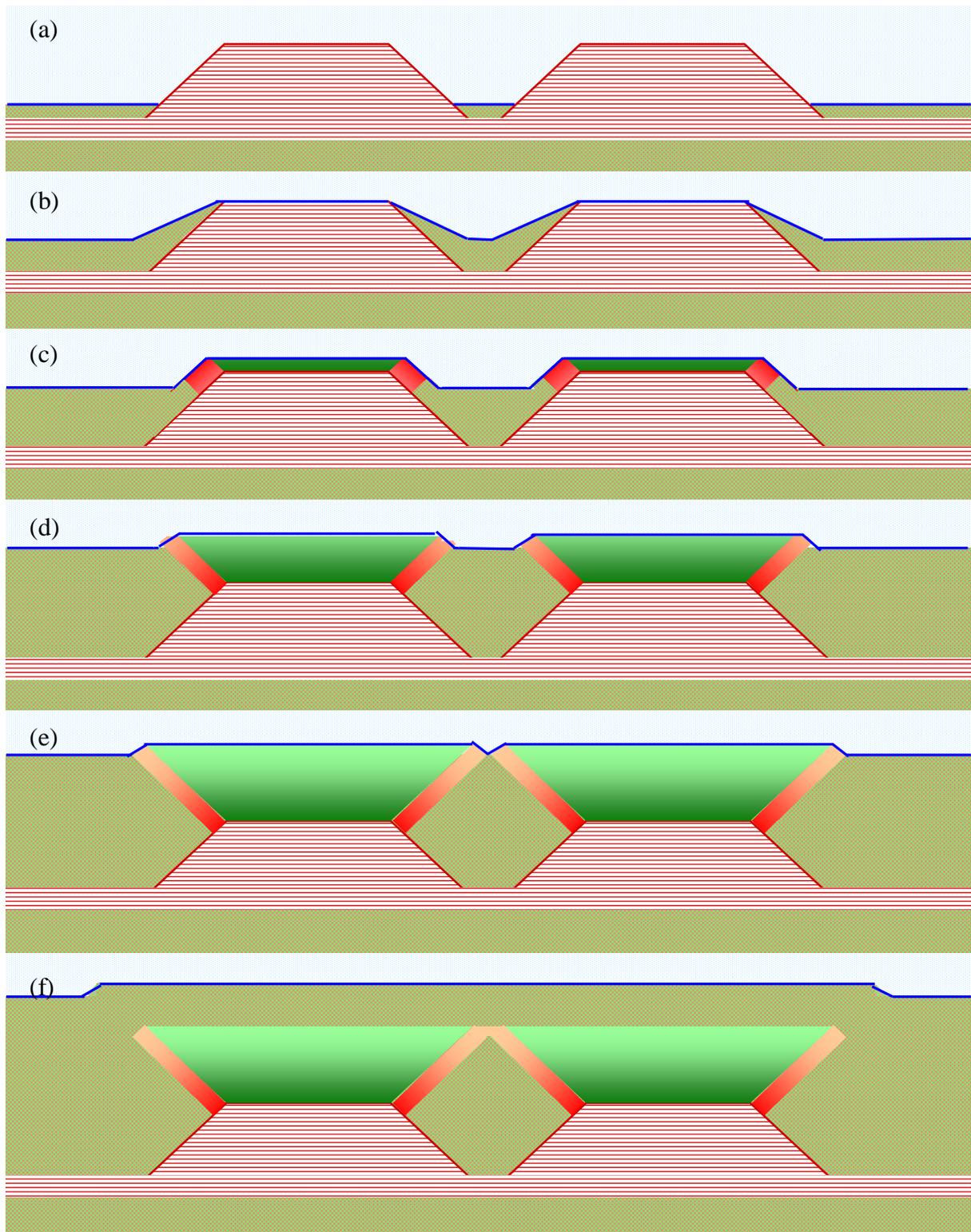


Figure IV-50: Dynamics of the phase separation process in the AlGaN capping layer. The GaN QDs and the wetting layer are in red, AlGaN is in light brown, the Al-rich zones are in green and the Ga-rich zones are in red. The blue line shows the surface of the AlGaN layer.

The size and shape of the Al-rich zones formed inside the AlGaN bumps can also be elucidated by considering the atomic structure of the AlGaN layer above the QD (Figure IV-51). The Ga-rich zones are not considered in the model for simplicity.

The atomic displacement in the AlGaN layer induced by the presence of GaN QD is localised in the close vicinity of the QD. Indeed, the atomic planes of AlGaN lying above the Al-rich zone (shown by green brace) do not exhibit any strain distribution, which is confirmed experimentally (Figure IV-15 (b)). This means that the negative ε_{zz} strain within the Al-rich zone entirely compensates the positive ε_{zz} strain caused by the GaN QD.

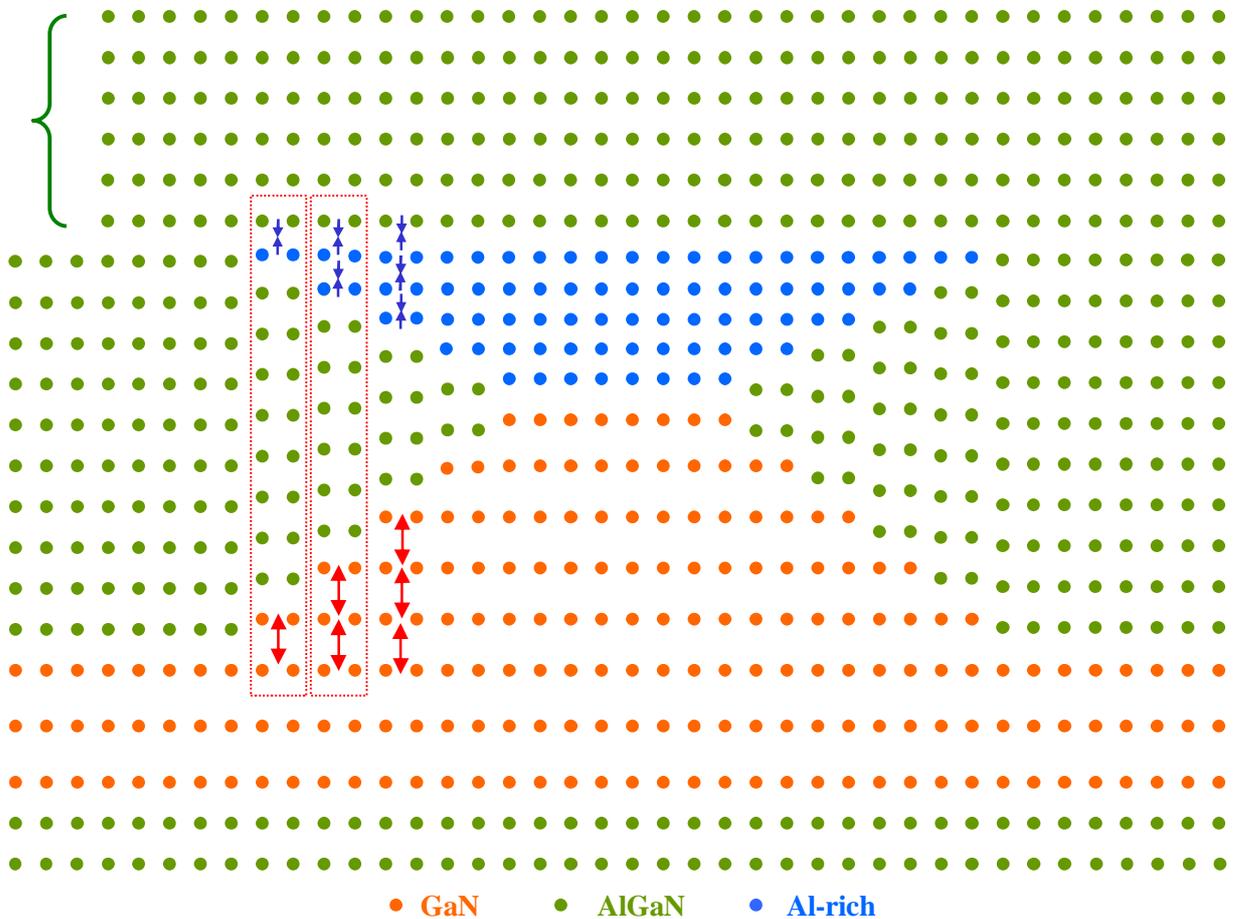


Figure IV-51: Possible atomic structure of GaN QD and surrounding barrier. Atomic displacements are exaggerated for better visualization. Each circle indicates an atomic position. The red/blue arrows show increase/decrease of interatomic distance relative to AlGaN, green brace indicates undistorted AlGaN planes.

The amount of strain induced by the QD is proportional to the local QD thickness. The atomic displacement Δ on the QD side (left rectangle on Figure IV-51) is due to the increase of only one interatomic distance from C^{AlGaN} to $C^{\text{AlGaN}} + \Delta$ (shown by the red arrow). This atomic displacement Δ can be compensated by the decrease of only one interatomic distance from C^{AlGaN} to $C^{\text{AlGaN}} - \Delta$ in the Al-rich zone (shown by the blue arrow). Two increased interatomic distances in the adjacent area (the two red arrows in the right rectangle) produce an atomic displacement of $2\cdot\Delta$, which requires the decrease of two interatomic distances in the Al-rich zone (the two blue arrows) and etc.

The displacement is maximal in the central part of the QD, thus the Al-rich zone is the thickest above the QD. In this way, the local thickness of Al-rich zone follows the local thickness of GaN QDs, forming thereby an “inverted dot”.

The involuntary formation of an Al-rich layer during the interruption of the AlGaN layer growth (Figure IV-19) seems rather paradoxical. Indeed, the Al-rich layer is 4-6 ML thick, which may signify that Ga atoms diffuse through the AlGaN matrix on distances of a few nm. On the other hand, no Ga interdiffusion was revealed during the QD formation. In particular, the upper and lower interfaces of the GaN wetting layer with AlGaN are atomically flat (Figure IV-29 (a)).

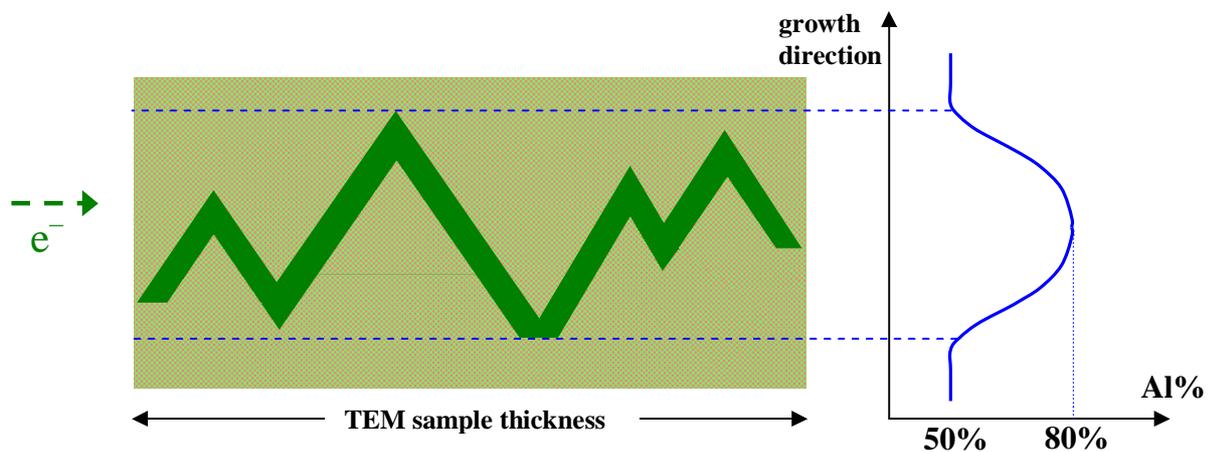


Figure IV-52: Scheme of the pseudo “Al-rich layer”. AlN monolayer is in green, AlGaN is in light brown. The profile shows the variation of Al concentration averaged along the direction of the electron beam, which is indicated by the green dash arrow.

We propose the following explanation of this observation. The formation of an AlN monolayer at the surface of AlGaN during the growth interruption is possible since the evaporation rate of GaN is much higher than that of AlN due to a greater thermodynamic stability of AlN compared to GaN [43]. According to the AFM measurements (not-shown) the mean roughness of AlGaN layer surface after the growth interruption is about 1 nm. We suppose that the following overgrowth of AlGaN does not change the morphology of the AlN monolayer formed at the sample surface (Figure IV-52).

An electron beam traversing such heterogeneous structure will interact with the AlN monolayer and the AlGaN surroundings. Thus, HRTEM image will reveal an average Al composition in this region. Due to the random variation of the local roughness, the Al concentration reaches its maximum in the centre of the pseudo “Al-rich layer” (Figure IV-52). Moreover, the roughness of the AlN monolayer does not change with the duration of the growth interruption, which explains why the thickness of the pseudo “Al-rich layer” is independent of the growth interruption length.

Bibliography

- [1] D. Bimberg, M. Grundmann, and N. N. Ledentsov, “Quantum Dot Heterostructures” JohnWiley & Sons, Chichester, (1999).
- [2] B. Damilano, N. Grandjean, F. Semond, J. Massies, and M. Leroux, *Appl. Phys. Lett.* **75**, 962 (1999).
- [3] N. Grandjean, B. Damilano, S. Dalmaso, M. Leroux, M. Laügt, and J. Massis. *J. Appl. Phys.*, **86**, 3714 (1999).
- [4] S. J. Rosner, E. C. Carr, M. J. Ludowise, G. Girolami, and H. I. Erikson, *Appl. Phys. Lett.* **70**, 420 (1997).
- [5] J. M. Gérard, O. Cabrol, and B. Sermage, *Appl. Phys. Lett.* **68**, 3123 (1996).
- [6] S. Nakamura, *Science* **281**, 956 (1998).
- [7] O. I. Micic, A.J. Nozik, *Journal of Luminescence* **70**, p. 95 (1996).
- [8] C. Delerue, M. Lannoo, “Nanostructures: Theory and Modelling”, Springer, p. 47 (2004).
- [9] I. N. Stranski and L. V. Krastanow, **146**, 797 *Abhandlungen der Mathematisch-Naturwissenschaftlichen Klasse. Akademie der Wissenschaften und der Literatur in Mainz* (1939).
- [10] J. A. Venables “Introduction to Surface and Thin Film Processes”, Cambridge: Cambridge University Press (2000).
- [11] J. W. Matthews “Epitaxial Growth” New York: Academic Press (1975).
- [12] H. Mariette, *C. R. Physique* **6**, 23 (2005).
- [13] Y. Taniyasu, M. Kasu, and T. Makimoto, *Nature (London)* **441**, 325 (2006).
- [14] S.-R. Jeon, Z. Ren, G. Cui, J. Su, M. Gherasimova, J. Hana, H.-K. Cho, and L. Zhou, *Appl. Phys. Lett.* **86**, 082107 (2005).
- [15] B. Borisov, V. Kuryatkov, Yu. Kudryavtsev, R. Asomoza, S. Nikishin, D. Y. Song, M. Holtz, and H. Temkin, *Appl. Phys. Lett.* **87**, 132106 (2005).
- [16] S. Tanaka, J.-S. Lee, P. Ramvall, and H. Okagawa, *Jpn. J. Appl. Phys.*, **42**, L885 (2003).
- [17] S. Tanaka, M. Takeuchi and Y. Aoyagi, *Jpn. J. Appl. Phys.* **39**, L831, (2000).
- [18] T. Huault, J. Brault, F. Natali, B. Damilano, D. Lefebvre, L. Nguyen, M. Leroux, and J. Massies, *Appl. Phys. Lett.* **92**, 051911 (2008).
- [19] B. Monemar, J. P. Bergman, T. Lundström, C. I. Harris, H. Amano, I. Akasaki, T. Detchprohm, K. Hiramatsu, N. Sawaki, *Solid-State Electronics*, **41**, 181, (1997).

-
- [20] J.L Rouvière, J. Simon, N. Pelekanos, B. Daudin, and G. Feuillet, *Appl. Phys. Lett.* **75**, 2632 (1999).
- [21] B. Daudin, F. Widmann, G. Feuillet, Y. Samson, M. Arlery, and J. L. Rouvière, *Phys. Rev.* **B 56**, R7069 (1997).
- [22] M. Korytov, M. Benaissa, J. Brault, T. Huault, T. Neisius, and P. Vennéguès, *Appl. Phys. Lett.* **94**, 143105 (2009).
- [23] E. Sarigiannidou, E. Monroy, B. Daudin, and J. L. Rouvière, *Appl. Phys. Lett.* **87**, 203112 (2005).
- [24] N. Grandjean, J. Massies, S. Yu. Karpov, and R. A. Talalaev, *Appl. Phys. Lett.* **74**, 1854 (1999).
- [25] J. D. Eshelby, *Proc. Roy. Soc. London Ser., A* **241**, 376 (1957).
- [26] V. Grillo, unpublished.
- [27] M. Benaissa, L. Gu, M. Korytov, T. Huault, P. A. van Aken, J. Brault, and P. Vennéguès, *Appl. Phys. Lett.* **95**, 141901 (2009).
- [28] <http://www.imagemet.com/>
- [29] J. Brault, T. Huault, F. Natali, B. Damilano, D. Lefebvre, M. Leroux, M. Korytov, and J. Massies, *J. Appl. Phys.* **105**, 033519 (2009).
- [30] P. Kratzer, Q. K. K. Liu, P. Acosta-Diaz, C. Manzano, G. Constantini, R. Songmuang, A. Rastelli, O. G. Schmidt, and K. Kern, *Phys. Rev.* **B 73**, 205347 (2006).
- [31] N. Grandjean, B. Damilano, and J. Massies, *J. Phys.: Condens. Matter* **13**, 6945 (2001).
- [32] N. Gogneau, D. Jalabert, E. Monroy, T. Shibata, M. Tanaka, and B. Daudin, *J. Appl. Phys.* **94**, 2254 (2003).
- [33] J. Elsner, M. Haugk, G. Jungnickel, and Th. Frauenheim, *Solid State Commun.* **106**, 739 (1998).
- [34] F. Widmann, B. Daudin, G. Feuillet, Y. Samson, J. L. Rouvière, and N. Pelekanos, *J. Appl. Phys.* **83**, 7618 (1998).
- [35] J. Coraux, B. Amstatt, J. A. Budagoski, E. Bellet-Amalric, J.-L. Rouvière, V. Favre-Nicolin, M. G. Proietti, H. Renevier, and B. Daudin, *Phys. Rev.* **B 74**, 195302 (2006).
- [36] A. Hesse, J. Stangl, V. Holy, T. Roch, G. Bauer, O. G. Schmidt, U. Denker, and B. Struth, *Phys. Rev.* **B 66**, 085321 (2002).
- [37] Z. Zhong, J. Stangl, F. Schaffler, and G. Bauer, *Appl. Phys. Lett.* **83**, 3695 (2003).

-
- [38] F. Ferdos, S. Wang, Y. Wei, M. Sadeghi, Q. Zhao, and A. Larsson, *J. Cryst. Growth* **251**, 145 (2003).
- [39] J. M. Ulloa, P. M. Koenraad, E. Gapihan, A. Létoublon, and N. Bertru, *Appl. Phys. Lett.* **91**, 073106 (2007).
- [40] G. Costantini, A. Rastelli, C. Manzano, P. Acosta-Diaz, R. Songmuang, G. Katsaros, O. G. Schmidt, and K. Kern, *Phys. Rev. Lett.* **96**, 226106 (2006).
- [41] U. Denker, A. Rastelli, M. Stoffel, J. Tersoff, G. Katsaros, G. Costantini, K. Kern, N. Y. Jin-Phillipp, D. E. Jesson, and O. G. Schmidt, *Phys. Rev. Lett.* **94**, 216103 (2005).
- [42] C. Priester and G. Grenet, *Phys. Rev. B* **64**, 125312 (2001).
- [43] N. Gogneau, D. Jalabert, E. Monroy, E. Sarigiannidou, J. L. Rouvière, T. Shibata, M. Tanaka, J. M. Gerard, and B. Daudin, *J. Appl. Phys.* **96**, 1104 (2004).

GENERAL CONCLUSION AND OUTLOOK

5.1. General conclusion

In this thesis, the investigation and characterisation of GaN quantum dots grown on AlGaN templates was performed. Comparing with AlN, which is the common template for the growth of GaN QDs, AlGaN is better tailored to be a basis for optoelectronic devices, due to its availability to be heavily doped. A number of publications refer to this topic [1, 2, 3, 4]; however, to our knowledge, a detailed morphological study of GaN/AlGaN QDs has never been reported.

Transmission electron microscopy was largely employed in this work for several reasons. First, this is one of the very few techniques which are capable to study the nanostructure of buried QD planes. Besides, high-resolution TEM is able to determine composition fluctuations with an atomic-scale resolution as soon as the imaging conditions have been validated. In fact, the main experimental results obtained in this work, such as the buried QD shape transition or the AlGaN barriers demixing, can not be revealed by the techniques like AFM or SEM, which are usually used for nanostructure characterization.

However, being inaccurately used, HRTEM may produce completely artificial results. A significant part of this manuscript addressed the adaptation of high-resolution TEM for the study of GaN-based semiconductors. The evaluation of heterostructure composition by HRTEM is based on the measurements of the relative atomic displacements. A numerical relation between the local lattice parameter and the concentration x of a ternary alloy was determined for $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ and $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$ structures. The impact of the uncertainties of the elastic constants as well as the effect of the surface relaxation on the accuracy of the composition determination were discussed.

Two complementary approaches are commonly used for the determination of the lattice fringe displacements from HRTEM images. GPA is a recognized technique for the

strain determination by analysing spatial frequencies in the reciprocal space. A special script for Digital Micrograph was developed for the treatment of HRTEM images in real space. The projection method algorithm was realized using the intensity maxima fitting by Chebyshev's polynomials. Both techniques were compared by applying them to several sets of artificially created lattice images. GPA was shown to create noticeable strain fluctuations at generated hetero-interfaces due to the contrast or periodicity changes. The projection method is able to measure rapid strain variations at the atomic scale, but the precision in strain determination considerably decreases in noisy images.

Extensive HRTEM image simulations using a Java version of the Electron Microscopy Software (JEMS) were performed to evaluate the effect of the imaging conditions on the strain measurements. The critical role of the TEM sample thickness for reliable HRTEM imaging along a low-index zone axis was shown. The defocus ranges for accurate atomic displacement determination at certain sample thicknesses were found. On the other hand, the strain determined from images taken in two-beam conditions [5] weakly depends on the objective lens defocus and sample thickness. High-resolution TEM imaging using such valid conditions was applied for the study of GaN QDs grown on AlGaN templates.

This study revealed several phenomena original for nitride semiconductors. The surface QD shape changes from perfect pyramidal to truncated pyramidal with the increase of the nominal thickness of the deposited GaN layer. The capping of QDs having a perfect pyramidal shape leads to a QD shape truncation and a QD volume increase. A similar behaviour was revealed for GaN QDs buried by AlN.

A phase separation was found in the AlGaN barriers with Al-rich zones formed above the QDs and Ga-rich regions placed around the Al-rich zones. The Al concentration into the Al-rich zones is about 70% and it decreases as the distance from the QD increases; Ga concentration varies from 55% to 65%. The Al-rich zones have well-defined lateral facets; their shape and size are similar to that of the QDs above which the Al-rich zones are placed. An alloy demixing was also observed in the samples with anisotropic QDs.

A precise strain measurement around buried QDs was possible thanks to the use of a Cs-corrected Titan 80-300 microscope. The ϵ_{zz} and ϵ_{xx} components of the strain tensor

determined from HRTEM images are in good agreement with theoretically calculated strain distributions. Both techniques showed the effect of surface QD relaxation.

The HAADF observations done with a probe Cs-corrected STEM microscope provided an independent proof of the phase separation phenomenon in the AlGa_N barriers. Moreover, the intensity of HAADF images was converted into local compositions by means of HAADF signal simulations. An average Al concentration in the Al-rich zone of 70% obtained by HAADF analysis was confirmed by EELS experiments. Moreover, the accurate analysis of the plasmon energy shift revealed a gradual decrease of the Al composition, with an increase of the distance from the QD top.

The examination of a sample series with varying thicknesses of the capping layer performed with a probe Cs-corrected STEM microscope revealed the dynamics of the QDs covering with AlGa_N. Moreover, we determined the different stages of the QD capping when the QD morphology changes and the phase separation occurs.

Based on the experimentally obtained data, models of the observed phenomena were proposed. Our explanations are founded on the principle of total energy minimization. The variety of surface QD shapes is explained by the interplay between the surface and the elastic energies of QD. The redistribution of the Ga_N from the QD top may be caused by the migration of atoms deposited on the QD apex during the overgrowth of the AlGa_N barrier. The increase of the buried QD volume may be due to the mass transport of Ga_N from the wetting layer, which can be driven by the further decrease of the elastic energy. A phase separation, which occurs in every new AlGa_N monolayer, is also governed by the relaxation of strain in the AlGa_N barriers introduced by the buried QDs.

5.2. Perspectives

This study can be pursued in several promising directions. First, some hypotheses done in the discussion should be proved experimentally.

The supposed difference of the in-plane lattice parameter a inside the GaN QDashes and in the AlGa_N lying between them should be testified by HRTEM. A possible dependence of the Al concentration in Al-rich zone on the QD size (the smaller QDs, the higher Al concentration) should be verified as well.

Our basic model of total energy minimization should be developed using numerical computations of elastic and surface energies. The equilibrium aspect ratio of the surface QDs can be calculated as a function of the QD size. The assumed distribution of the ϵ_{xx} and ϵ_{zz} strains in the AlGa_N layer at the first stages of the capping can be computed as well. Besides, the concentration of Al into the Al-rich zones, which corresponds to the minimisation of the elastic energy in the AlGa_N barrier, may be evaluated numerically for QD and QDash structures.

Some additional experiments, aimed to further investigate the alloy demixing in the AlGa_N layer, could also be performed. For example, a challenging objective is to minimise or to eliminate the phase separation in the AlGa_N barriers. A possible approach is the use of templates with a higher GaN content, Al_{0.4}Ga_{0.6}N for example. The decrease of the lattice mismatch between GaN and the barrier material may weaken the stress-driven phase separation. An alternative way of decreasing the lattice mismatch could be the growth of Al_{0.1}Ga_{0.9}N QDs on Al_{0.5}Ga_{0.5}N templates.

Besides, the use of a double capping method [6], which results in the reduction of the QD height, may decrease the strain distribution in the capping layer. It is also of interest to investigate the impact of the AlGa_N barrier growth temperature on the quantity of alloy demixing, presuming that a decrease of atom diffusion lengths at lower temperatures may hamper the phase separation.

From the obtained results, it is possible to prospect several new approaches for the development of light-emitting devices based on GaN/AlGaIn QDs.

The V-shaped partial demixing of the spacer layer may result in the vertical anticorrelation of QDs in adjacent QD planes [7]. This occurrence can be studied as a function of the AlGaIn barrier thickness. Such an anticorrelation may result in the decrease of the size dispersion of the QDs, therefore improving the luminescence properties of stacked structures.

The study of the QD morphology evolution after reaching the maximum of the RHEED intensity (moment t_2 on Figure IV-2) is another promising direction. We already observed that an increase of the growth interruption duration results in the decrease of the QD size and the disappearance of the wetting layer (the QD plane is marked by red arrow on Figure V-1 (b)).

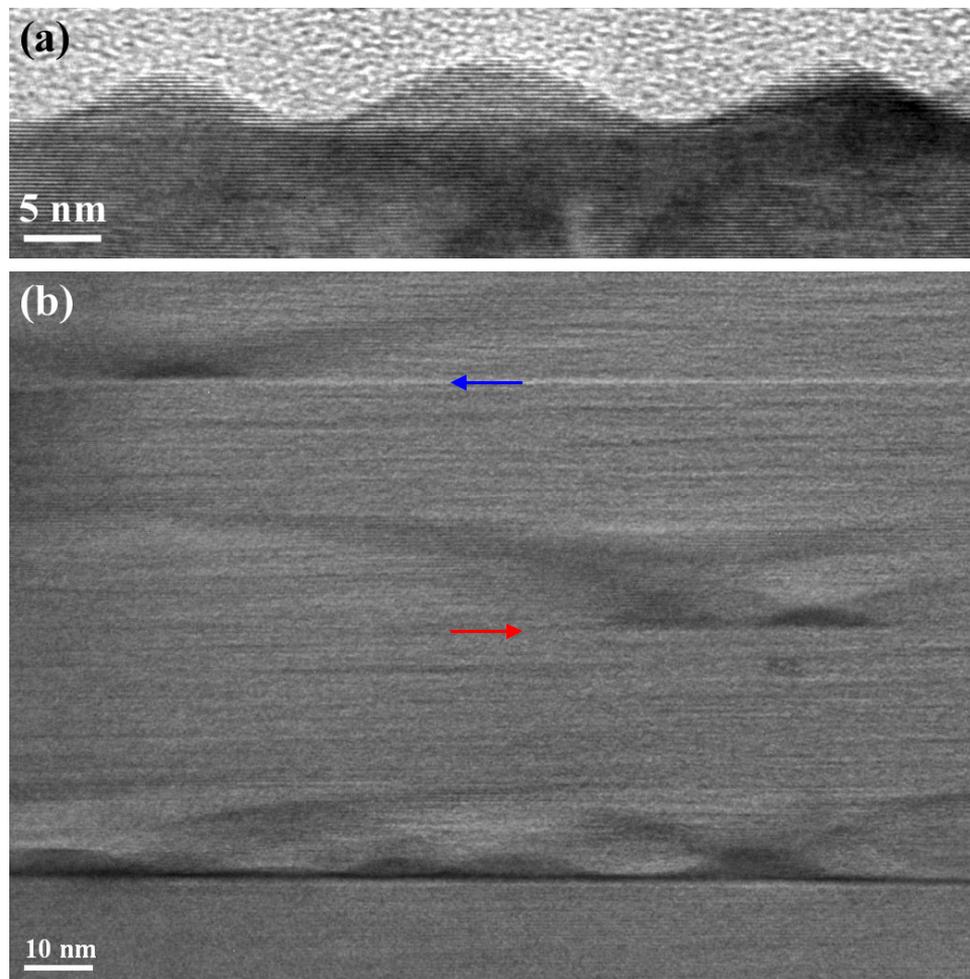


Figure V-1: TEM images of surface (a) and buried (b) GaN QDs grown with an increased duration of the growth interruption. The coloured arrows show the QD planes.

At the same time the surface QDs still keep a perfect pyramidal shape (Figure V-1 (a)). A further increase of the growth interruption duration leads to the formation of an Al-rich layer lying between the buried QD planes (marked by the blue arrow on Figure V-1 (b)).

A stacked QDs structure without any wetting layer is supposed to have improved photoluminescence properties compared with common QD structures with wetting layers due to a deeper carrier localisation inside the QDs and a better electron-hole wavefunction overlap [8, 9].

A number of phenomena unique for GaN-based semiconductors were discovered and investigated in this thesis. However, further study of GaN/AlGaIn QDs may reveal new challenging effects. This work will be certainly continued in CRHEA in collaboration with other laboratories.

Bibliography

- [1] S. Tanaka, J.-S. Lee, P. Ramvall, and H. Okagawa, *Jpn. J. Appl. Phys.*, **42**, L885 (2003).
- [2] Y.-H. Cho, H. S. Kwack, and B. J. Kwon, J. Barjon, J. Brault, B. Daudin, and Le Si Dang, *Appl. Phys. Lett.* **89**, 251914 (2006).
- [3] K. Pakula, R. Bozek, K. Surowiecka, R. Stepniewski, A. Wysmolek, and J. M. Baranowski, *Phys. Stat. Sol. (b)* **243**, No. 7 (2006).
- [4] Y. Hori, T. Andreev, E. Bellet-Amalric, O. Oda, D. Le Si Dang, and B. Daudin, *Phys. Stat. Sol. (a)* **204**, No. 1 (2007).
- [5] A. Rosenauer, D. Gerthsen, V. Potin, *Phys. Stat. Sol (a)* **203**, 176 (2006).
- [6] J. M. Ulloa, P. M. Koenraad, E. Gapihan, A. Létoublon, and N. Bertru, *Appl. Phys. Lett.* **91**, 073106 (2007).
- [7] C. Priester and G. Grenet, *Phys. Rev. B* **64**, 125312 (2001).
- [8] B. Damilano, PhD thesis, Univ. Nice-Sophia Antipolis, (2001).
- [9] Y.-R. Wu, Y.-Y. Lin, H.-H. Huang, and J. Singh, *J. Appl. Phys.*, **105**, 013117 (2009).

APPENDICES

APPENDIX A The relationship between three- and four-index notations

The standard three-index Miller notation does not completely represent the 3-fold symmetry of the hexagonal lattice system. For example, the equivalent directions are denoted as [100], [010] or $[\bar{1}\bar{1}0]$. The introduction of the fourth index is aimed to make more obvious the labelling of similar planes and directions.

The fourth collinear basis vector \mathbf{w} is chosen 120° apart from the basis vectors \mathbf{u} and \mathbf{v} lying in the basal plane (Figure VI-1).

The transfer of plane notation from three-index (UVW) to four-index (uvtw) scheme is quite simple. The additional index t is

$$t = -U - V \quad (\text{eq. VI-1})$$

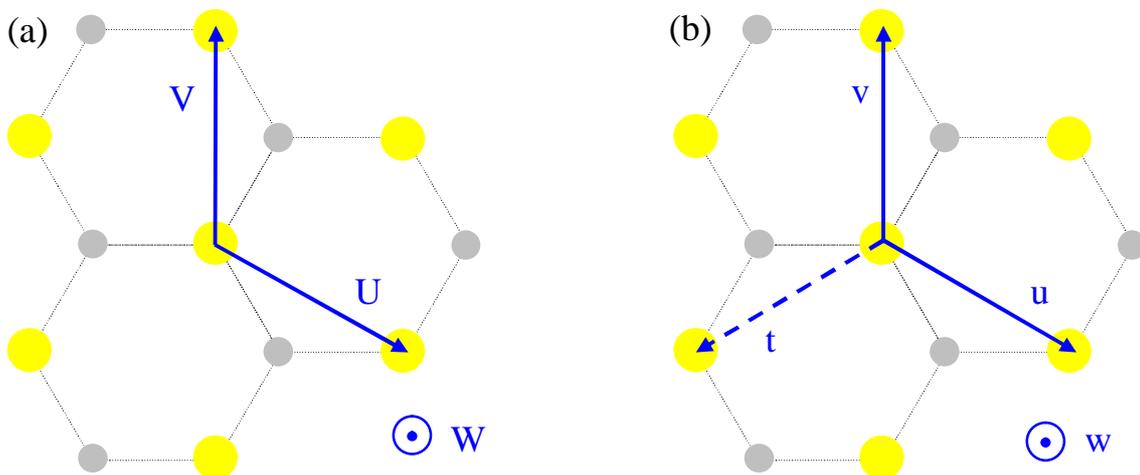


Figure VI-1: The basis vectors used for three-index (a) and four-index (b) notations. The basal plane A of Ga atoms (shown by the yellow circles) is superposed on the basal plane B of N atoms (shown by grey the circles).

The converting of directions is a bit more complicated:

$$u = 2U - V, v = -U + 2V, t = -U - V, w = 3W \quad (\text{eq. VI-2})$$

$$U = \frac{1}{3}(u - t), V = \frac{1}{3}(v - t), W = \frac{1}{3}w \quad (\text{eq. VI-3})$$

where [UVW] and [uvw] are the direction notations in three- and four-index systems.

Consequently, the equivalent directions [100], [010] and [110] are written in the four-index scheme as [2110], [1210] and [1120] respectively, which exposes their similarity.

APPENDIX B Projection of interatomic distances along different zone axis

To state analytically the interatomic distances $a_1 - a_4$ in the projections shown on Figure II-4, we should examine the atomic arrangement in the basal plane (Figure VI-2).

The distance between two adjacent atoms in the basal plane equals to the lattice constant a (shown by the green arrow on Figure VI-2). We see just a half of this distance in the lattice projection along the $\langle 1\bar{1}00 \rangle$ direction (shown by the violet arrow), i.e. $a_1 = a/2$.

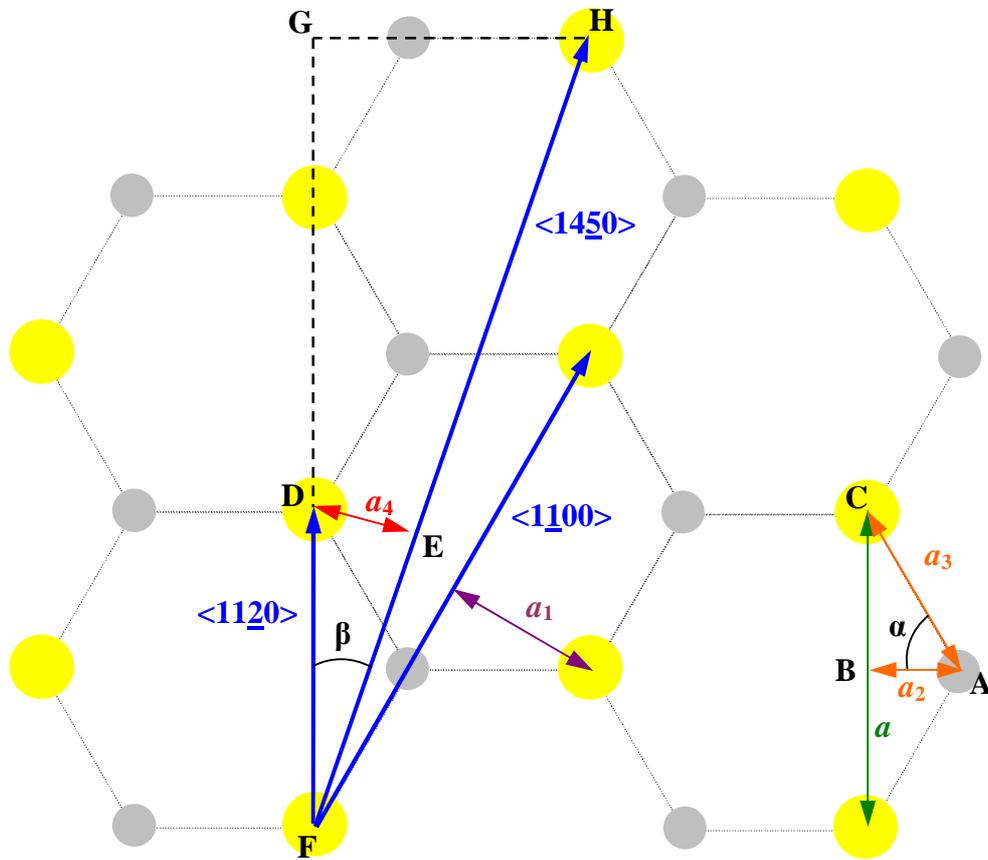


Figure VI-2: View of GaN structure along the $[0001]$ direction. The basal plane A of Ga atoms is superposed on the basal plane B of N atoms. The blue arrows show the main crystallographic axes, the coloured arrows are referred in the text.

To determine the interatomic distances a_2 and a_3 in projection along the $\langle 11\bar{2}0 \rangle$ direction (Figure II-4 (b)) we will use the geometrical construction shown on the right bottom corner of Figure VI-2. In the right-angled triangle ABC with an angle $\alpha = 60^\circ$ the cathetus a_2 and the hypotenuse a_3 are related as $a_2 = a_3/2$. Then, using Pythagorean theorem, we find that

$$a/2 = \sqrt{a_3^2 - a_2^2} = \sqrt{a_3^2 - (a_3/2)^2} = a_3\sqrt{3}/2 \quad (\text{eq. VI-4})$$

Thereby $a_2 = a/2\sqrt{3}$ and $a_3 = a/\sqrt{3}$.

The distance a_4 in the projection along the $\langle 14\bar{5}0 \rangle$ direction (red arrow on Figure VI-2) can be found from the right-angled triangle DEF. The hypotenuse DF equals a , whereas the angle β can be determined from the right-angled triangle FGH. The cathetus FG is $2.5a$, whereas cathetus HG is $1.5a_3$.

$$\text{tg } b = HG / FG = 1.5a_3 / 2.5a = \sqrt{3}/5 \quad (\text{eq. VI-5})$$

Thereby the angle β between the $\langle 11\bar{2}0 \rangle$ and the $\langle 14\bar{5}0 \rangle$ directions is 19.1° . The distance a_4 can be found as

$$a_4 = DF \sin b = a \frac{\text{tg } b}{\sqrt{1 + \text{tg}^2 b}} = a \frac{\sqrt{3}/5}{\sqrt{28}/5} = a\sqrt{3/28} \quad (\text{eq. VI-6})$$

The ratios of the interatomic distances calculated above are valid for all crystals of hexagonal lattice system.

APPENDIX C Aspect ratio calculation for hexagonal pyramid

We will start our calculations by stating the surface and the volume of the pyramid having a perfect shape (Figure VI-3). The surface of each lateral facet (for example ABC) can be found as a surface of an isosceles triangle:

$$S_{ABC} = \frac{1}{2} AE \cdot BC = \frac{1}{2} f \cdot a \quad (\text{eq. VI-7})$$

where f is the triangle height AE, a is the length of the pyramid base BC.

From the right-angled triangle ADE f can be expressed as

$$f = h / \sin g \quad (\text{eq. VI-8})$$

where h is the pyramid height and γ is the angle between the lateral facet and the pyramid base (Figure VI-3). The total surface of the perfect pyramid S_0 is a sum of the lateral facet surfaces:

$$S_0 = 6S_{ABC} = 6 \frac{h \cdot a}{2 \sin g} = 3 \frac{h \cdot a}{\sin g} \quad (\text{eq. VI-9})$$

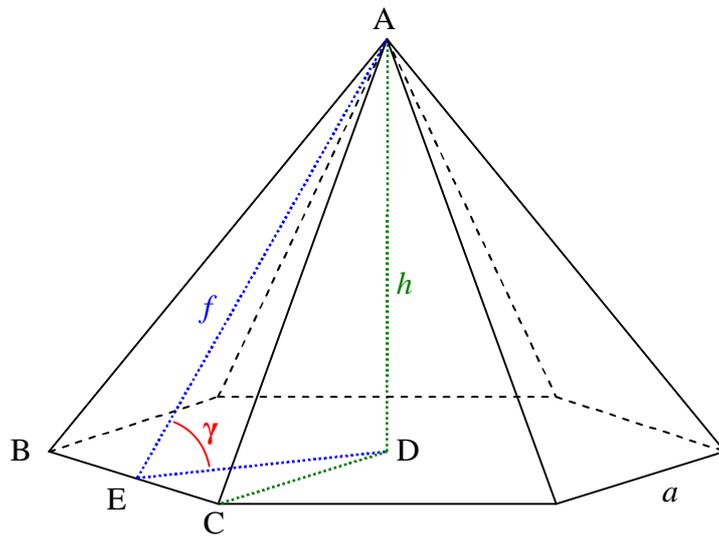


Figure VI-3: Perfect pyramid used for the aspect ratio calculation.

From the right-angled triangles AED and CDE the $\text{tg } \gamma$ is:

$$\text{tg } \gamma = \frac{AD}{ED} = \frac{AD}{\sqrt{CD^2 - EC^2}} = \frac{h}{\sqrt{a^2 - \frac{1}{4}a^2}} = \frac{2h}{\sqrt{3}a} \quad (\text{eq. VI-10})$$

Thus pyramid height h can be expressed in terms of the base length a :

$$h = \frac{\sqrt{3}}{2} a \cdot \text{tg } \gamma \quad (\text{eq. VI-11})$$

By substituting the pyramid height h to the equation VI-9 we obtain

$$S_0 = 3 \frac{a}{\sin \gamma} \frac{\sqrt{3}}{2} a \cdot \text{tg } \gamma = \frac{3\sqrt{3}}{2} \frac{a^2}{\cos \gamma} \quad (\text{eq. VI-12})$$

The perfect pyramid volume V_0 can be found using formula:

$$V_0 = \frac{1}{3} h \cdot S_{\text{base}} = \frac{1}{3} h \cdot \frac{3\sqrt{3}}{2} a^2 = \frac{1}{3} \cdot \frac{\sqrt{3}}{2} a \cdot \text{tg } \gamma \cdot \frac{3\sqrt{3}}{2} a^2 = \frac{3}{4} a^3 \cdot \text{tg } \gamma \quad (\text{eq. VI-13})$$

where S_{base} is the surface of the pyramid base, which is expressed as

$$S_{\text{base}} = 6S_{BCD} = 6 \cdot DE \cdot BC = 6 \frac{1}{2} \cdot \frac{\sqrt{3}}{2} a \cdot a = \frac{3\sqrt{3}}{2} a^2 \quad (\text{eq. VI-14})$$

The total surface of the elongated pyramid, which represents a QDash (Figure VI-4), consists of the surface of the perfect pyramid S_0 plus additional surface of lateral facets S_{add} .

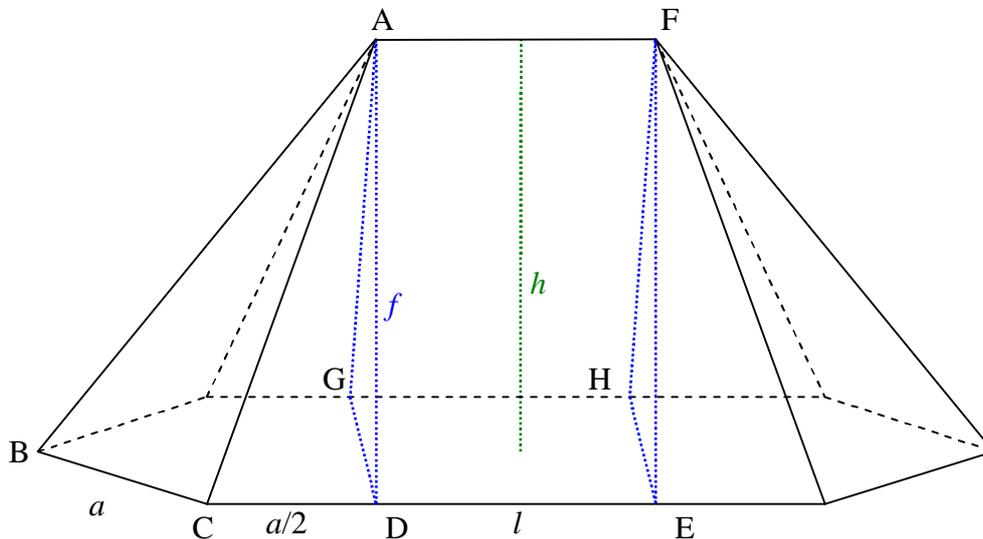


Figure VI-4: Elongated pyramid used for the aspect ratio calculation.

From the rectangle ADEF:

$$S_{\text{add}} = 2S_{ADEF} = 2f \cdot l \quad (\text{eq. VI-15})$$

Considering equation VI-7 the perfect pyramid S_0 surface can be stated as

$$S_0 = 6S_{ABC} = 6 \frac{f \cdot a}{2} = 3f \cdot a \quad (\text{eq. VI-16})$$

Thus the total surface of the elongated pyramid is

$$S_{\text{Tot}} = S_0 + S_{\text{add}} = 3f \cdot a + 2f \cdot l = S_0 \left(1 + \frac{2l}{3a}\right) \quad (\text{eq. VI-17})$$

The additional volume due to the insertion of the prism ADGFEH is

$$V_{\text{add}} = \frac{1}{2} h \cdot S_{DEGH} = \frac{1}{2} h \cdot DG \cdot DE = \frac{1}{2} h \cdot \sqrt{3} a \cdot l \quad (\text{eq. VI-18})$$

Thereby the total volume of the elongated pyramid is

$$V_{\text{Tot}} = V_0 + V_{\text{add}} = \frac{\sqrt{3}}{2} h \cdot a^2 + \frac{\sqrt{3}}{2} h \cdot a \cdot l = V_0 \left(1 + \frac{l}{a}\right) \quad (\text{eq. VI-19})$$

where V_0 is the perfect pyramid volume (eq. VI-13).

Thereby, the surface/volume ratio of the elongated pyramid $S_{\text{tot}}/V_{\text{tot}}$ is always smaller than that of the perfect pyramid (S_0/V_0) due to the factor $\frac{(1 + \frac{2l}{3a})}{(1 + \frac{l}{a})}$, which is always lesser than

one. In other words, for a similar volume, the surface of the elongated pyramid (Figure VI-4) is smaller than the surface of the perfect pyramid (Figure VI-3).

Now we will calculate the surface/volume ratio of the truncated pyramid. To express the volume and the surface of the truncated pyramid we will use the formulas derived for the perfect pyramid.

Truncated pyramid is characterized by the parameter η which is defined as a ratio of the removed apex height t to the initial pyramid height h (Figure VI-5).

Due to the similarity of the triangles ABC and ADE, the surface of each lateral facet of the truncated pyramid (trapezium BCED on Figure VI-5) is

$$S_{BCED} = S_{ABC}(1-h^2) \quad (\text{eq. VI-20})$$

where S_{ABC} is the surface of the triangle ABC (eq. VI-7).

The surface of the top hexagon DEFGHI is related with the surface of the pyramid base S_{base} (eq. VI-14) as

$$S_{\text{DEFGHI}} = S_{\text{base}} \cdot h^2 \quad (\text{eq. VI-21})$$

Thus the total surface of the truncated pyramid is

$$S_{\text{Tot}} = 6 \cdot S_{BCDE} + S_{\text{DEFGHI}} = \frac{3\sqrt{3}}{2} \frac{a^2}{\cos g} (1-h^2) + \frac{3\sqrt{3}}{2} \cdot a^2 \cdot h^2 = S_0(1-h^2 + h^2 \cos g) \quad (\text{eq. VI-22})$$

where S_0 is the surface of the perfect pyramid (eq. VI-12).

The truncated pyramid volume is just a difference of the perfect pyramid V_0 and the removed apex volume V_{apex} , which is proportional to the η^3 .

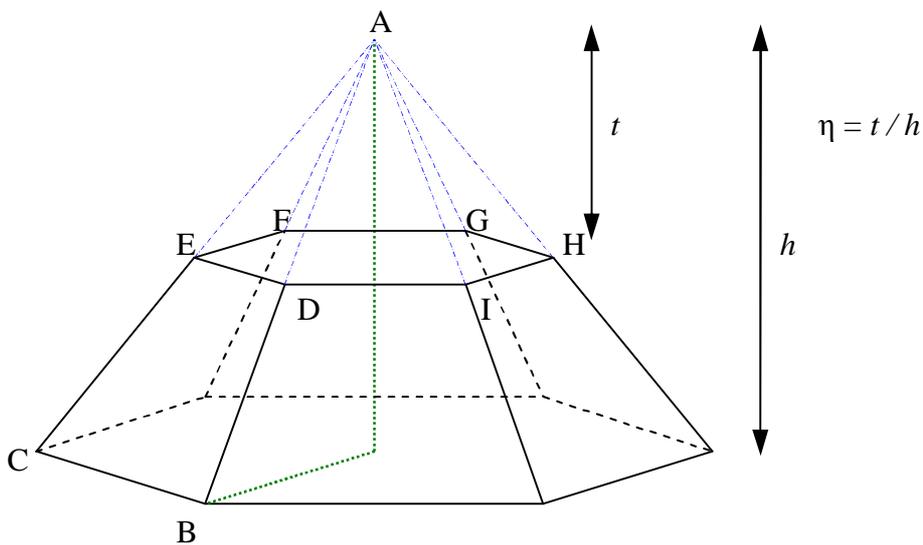


Figure VI-5: Truncated pyramid used for the aspect ratio calculation.

$$V_{Tot} = V_0 - V_{apex} = V_0 - V_0 \cdot h^3 = V_0 \cdot (1 - h^3) \quad (\text{eq. VI-23})$$

Thereby, the surface/volume ratio of the truncated pyramid can be expressed by means of the surface/volume ratio of the perfect pyramid:

$$\frac{S_{Tot}}{V_{Tot}} = \frac{S_0(1-h^2+h^2 \cos g)}{V_0 \cdot (1-h^3)} = \frac{S_0}{V_0} \cdot \frac{1-h^2(1-\cos g)}{1-h^3} = \frac{S_0}{V_0} \cdot F(h) \quad (\text{eq. VI-24})$$

The relationship between the surface/volume ratios of perfect and truncated pyramids is governed by the factor $F(\eta)$. The behaviour of this factor strongly depends on the value of $\cos g$.

In the case of the pyramid with the lateral facets parallel to the $\{1103\}$ planes (i. e. $\gamma = 32^\circ$), the $F(\eta)$ factor is almost constant for η smaller than 0.15 (Figure VI-6). It reaches local minimum with η about 0.09, but the value of $F(0.09)$ is only 0.034% smaller than $F(0)$.

For $\eta > 0.15$, the $F(\eta)$ factor rises rapidly with η . At $\eta = 0.37$, which corresponds to the pyramid with a height/diameter ratio of 0.17 (shown by the red dash line), the factor $F(0.37)$ is 3% higher than $F(0)$, which represents the perfect pyramid.

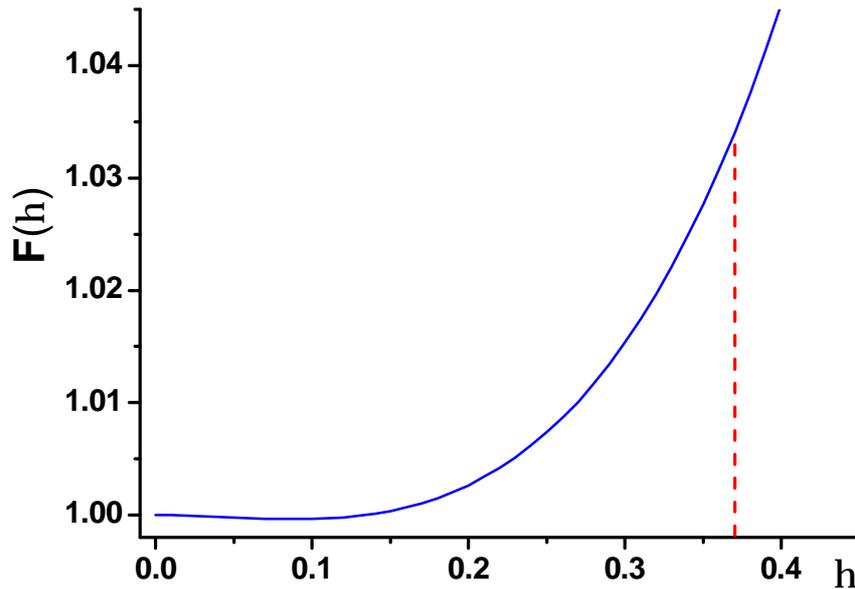


Figure VI-6: The factor $F(\eta)$ in the equation VI-24 as a function of η for $\gamma = 32^\circ$. The parameter $\eta = 0.37$, corresponding to the pyramid with height/diameter ratio of 0.17, is shown by the red dash line.

Based on the above-stated analysis, we can conclude that a perfect pyramid with $\{1\bar{1}03\}$ lateral facets has a minimum surface/volume ratio as compared to any pyramid having a truncated shape.