Advanced Transmission Electron Microscopy Investigation of Nano-clustering in Gd-doped GaN

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m
(Dr. rer. nat.)
im Fach Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
der Humboldt-Universität zu Berlin

von
Herrn M.Eng. Mingjian Wu
Hunan, China

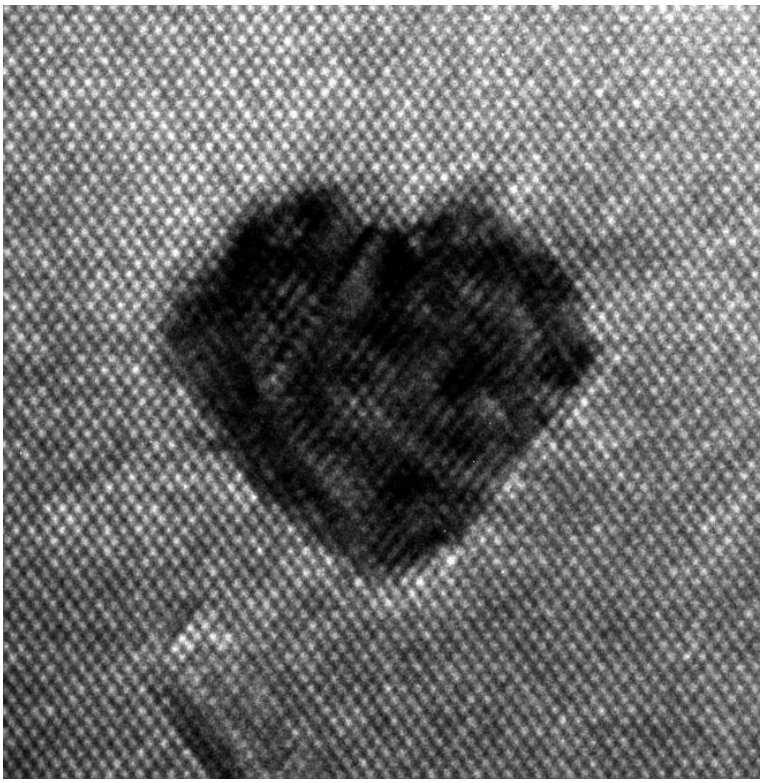
Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Stefan Hecht, Ph.D.

Gutachter:
1. Prof. Dr. Henning Riechert
2. Prof. Dr. Thomas Schroeder
3. Prof. Dr. Philippe Vennéguès

eingereicht am: 19.12.2013
Tag der mündlichen Prüfung: 29.4.2014

“Seeing is believing.” ?
“Sehen ist Glauben.” ?



This work is dedicated to my parents and family.
Diese Arbeit ist meinen Eltern und Familie gewidmet.

Cover image:
A high-resolution transmission electron micrograph
of an InAs nano-cluster embedded in Si.
Cluster hight: about 8 nm.
Corresponding study can be found in Chapter 5.1.

Abstract

Many physical properties of semiconductor compounds and their heterostructure depend on the phase purity, i.e., on the homogeneity in chemical composition and crystal structure. Gd-doped GaN (GaN:Gd) is one of the so-called "dilute magnetic semiconductor" materials, whose magnetic properties are significantly affected by phase separation (usually in the form of structurally nano-clusters formation), because of the generally existing miscibility gap in the material systems. However, the structural information about Gd atoms in GaN:Gd is not yet clear, due to the extremely low Gd concentration in the material, which challenged most microscopic and spectroscopic techniques. The central goal of this dissertation is (1) to clarify the distribution of Gd atoms in GaN:Gd with Gd concentration in the range between 10^{16}–10^{19} cm^{-3} by means of advanced (scanning) transmission electron microscopy [(S)TEM]; and based on that, (2) to understand the mechanisms that control such distribution.

The investigation of embedded nano-objects in crystals is highly challenging. To enable the detection and the quantitative analysis, it requires a combination of experimental techniques, with which any of them provide only partial information about the object. In case of GaN:Gd, the large difference in the atomic radius of Gd and Ga may induce strong lattice distortion that can, in principle, be detected and analyzed.

In this dissertation, we discuss in detail the application and limitations of (S)TEM imaging techniques, the quantitative analysis of local lattice distortion [geometric phase analysis (GPA)] and modeling methods [based on valence force-field (VFF) and density-functional theory (DFT)] dedicated to observe and analyze nano-clusters embedded in semiconductor epilayers. Besides, two case studies of semiconductor material systems that contain apparently observable nano-clusters are considered to apply the methods and to help investigation and understanding nano-clustering in GaN:Gd. In one case, the interface character and strain state of intentionally grown InAs nano-clusters (or, quantum dots) embedded in Si are analyzed by high-resolution TEM (HRTEM) and GPA. In the other case, the formation and phase transformation of Bi-containing clusters in annealed $GaAs_{1-x}Bi_x$ epilayers are investigated by combination of (S)TEM imaging techniques.

Finally, we are able to identify the occurrence of GdN clusters in GaN:Gd samples and to determine their atomic structure. Strain contrast imaging in conjunction with two-beam dynamic contrast simulation unambiguously identifies the occurrence of small, platelet-shaped GdN clusters. These clusters are nearly uniform in size with their broader face parallel to the GaN (0001) basal plane. The result is confirmed by dark-field STEM Z-contrast imaging. The strong local lattice distortion (displacement field) induced by the clusters is recorded by HRTEM images and quantitatively analyzed by GPA. By comparing the displacement fields which are analyzed experimentally with these fields that are derived from energetically favored models based on either VFF or DFT, we conclude that the clusters are bilayer GdN with platelet diameter of only few Gd atoms; their internal structure is close to rocksalt GdN. This atomic structure model enables our discussion about the energetics of the clusters by DFT calculations in conjunction with the classical Frenkel-Kontorova model. The results indicate that the driving force for the formation of observed platelet in specific size is a compromise between the gain in cohesive energy and the penalty from interfacial strain energy due to lattice mismatch between the GdN cluster and GaN host.

These findings suggest that further studies to understand the magnetic properties of GaN:Gd that should take into account the fact of nano-clustering even in the extremely low Gd concentration. The way how we combined existing conventional routine techniques to explore unknown questions at the edge of limit extends our ability to verify or falsify our beliefs, philosophically.

Keywords: nano-clustering, interface, strain, (S)TEM, GPA, GaN:Gd, InAs/Si, Ga(As, Bi)

Zusammenfassung

Viele physikalische Eigenschaften von Halbleiterverbindungen und deren Heterostrukturen hängen von der Phasenreinheit, d.h. von der Homogenität der chemischen Zusammensetzung sowie der Kristallstruktur, ab. Gd dotiertes GaN (GaN:Gd) zählt zu den Vertretern der sogenannten „verdünnten magnetischen Halbleiter", dessen magnetische Eigenschaften maßgeblich auf Grund der im Materialsystem generell auftretenden Mischungslücke von Phasenseparation (gewöhnlich in Form struktureller Nano-Cluster-Bildung) betroffen ist. Die strukturelle Aufklärung über die räumliche Verteilung der Gd-Atome in GaN:Gd ist jedoch bisher ungeklärt, da die extrem geringe Gd-Konzentration im Material eine Herausforderung für die meisten mikroskopischen und spektroskopischen Methoden darstellt. Das zentrale Ziel der vorliegenden Arbeit besteht einerseits darin, die Verteilung von Gd in GaN:Gd mit Gd-Konzentrationen von 10^{16}–10^{19} cm^{-3} mittels fortgeschrittener (Raster-) Transmissionselektronenmikroskopie [(S)TEM] zu bestimmen. Darauf basierend wird zum anderen das Verständnis des Mechanismus, der diese Verteilungen bewirkt, entwickelt.

Die Untersuchung eingebetteter Nano-Objekte in Kristallen stellt eine große Herausforderung dar. Um die Beobachtung und Analyse zu ermöglichen, wird eine Kombination aus experimentellen Techniken benötigt, von denen jede lediglich eine Teilinformation über das Objekt beleuchtet. Im Fall von GaN:Gd kann der große Unterschied in den Atomradien von Gd und Ga starke Gitterverzerrungen hervorrufen, die prinzipiell beobachtet und analysiert werden können.

In dieser Doktorarbeit diskutieren wir detailliert die Anwendung und die Grenzen von (S)TEM-Abbildungsmethoden, von quantitativen Analysen der lokalen Gitterverzerrung [Analyse der geometrischen Phase (GPA)] und von Modellierungsmethoden [basierend auf Valenzkraftfeld (VFF) und Dichtefunktionaltheorie (DFT)], um Nano-Cluster in epitaktischen Halbleiterschichten zu beobachten und zu analysieren. Außerdem werden Fallstudien zweier Materialsysteme betrachtet, die offensichtlich Nano-Cluster aufweisen, um die TEM Methoden anzuwenden und um die Untersuchung sowie das Verständnis von Nano-Clustern in GaN:Gd zu unterstützen. In einem Fall werden Grenzflächencharakter und Dehnungszustand von gezielt gewachsenen InAs Nano-Clustern (oder Quantenpunkten), die in Si eingebettet sind, mit Hilfe von hochauflösender TEM (HRTEM) und GPA analysiert. Im zweiten Fall werden die Bildung und Phasentransformation von Bi-haltigen Clustern in ausgeheilten, epitaktisch gewchsenen $GaAs_{1-x}Bi_x$-Schichten mit einer Kombination von (S)TEM-Abbildungsmethoden untersucht.

Schließlich sind wir in der Lage, die in GaN:Gd Proben auftretenden GdN-Cluster zu identifizieren und ihre atomare Struktur zu bestimmen. Dehnungskontrastabbildungen im Zusammenhang mit Kontrastberechnungen für den dynamischen Zweistrahlfall belegen eindeutig das Auftreten von kleinen, plättchenförmigen GdN- Clustern. Diese Cluster weisen nahezu gleiche Abmessungen auf und liegen mit der ausgedehnten Fläche parallel zu den GaN(0001)-Basalebenen. Dieses Ergebnis wird durch Dunkelfeld-STEM-Abbildungen, die für die Kernladungszahl Z empfindlich sind, bestätigt. Die starke, lokale Gitterdehnung (Verzerrungsfeld), die durch die Cluster hervorgerufen wird, ist in HRTEM-Aufnahmen abgebildet und mit GPA-Analysen quantitativ ausgewertet worden. Durch den Vergleich von Verzerrungsfeldern, die experimentell ermittelt worden sind, mit theroretischen Feldern, die aus energetisch bevorzugten Modellen (basierend auf VFF- oder DFT-Berechnungen) folgen, schließen wir auf Cluster aus zweilagigen GdN-Plättchen mit einem Durchmesser von wenigen Gd-Atomen. Ihre interne Struktur entspricht etwa der NaCl-Phase des GdN. Dieses atomare Strukturmodell erlaubt unsere Diskussion der Energieverhältnisse der Cluster anhand von DFT-Berechnungen in Verbindung mit dem klassischen Frenkel-Kontorova-Modell. Die Ergebnisse implizieren, dass die treibende

Kraft für die beobachtete Plättchengröße ein Gleichgewicht zwischen der Zunahme von kohäsiver Energie und der Einschränkung durch die Dehnungsenergie an der Grenzfläche zwischen GdN-Cluster und GaN-Wirtsgitter aufgrund der Gitterfehlanpassung ist.

Diese Ergebnisse legen nahe, dass weitere Studien zum Verständnis magnetischer Eigenschaften von GaN:Gd Nanocluster selbst im Fall sehr geringer Gd- Konzentrationen in Betracht ziehen sollten. Der Weg, wie wir existierende, konventionelle Routinemethoden zur Beantwortung unbekannter Fragen an der Grenze experimenteller Möglichkeiten kombinieren, erweitert unsere Fähigkeiten, unsere Annahmen zu bestätigen oder zu widerlegen.

Stichwörte: Nano-Clustering, Grenzfläche, Dehnung, (S)TEM, GPA, GaN:Gd, InAs/Si, Ga(As, Bi)

Abbreviations

ADA	abrupt displacement approximation
ADF	annual dark-field
BF	bright-field
COLC	center of Laue circle
DF	dark-field
DFT	density function theory
DL	dislocation loop
DMS	dilute magnetic semiconductor
FOLZ	first-order Laue zone
FK	Frank-Kontorova (model)
GPA	geometric phase analysis
HAADF	high-angle annual dark-field
HOLZ	high-oder Laue zone
HRSTEM	high-resolution scanning transmission electron microscopy
HRTEM	high-resolution transmission electron microscopy
LT	low temperature
MBE	molecular beam epitaxy
MD	misfit dislocation
QD	quantum dot
RH(rh)	rhomboheral (structure)
SF	stacking fault
STEM	scanning transmission electron microscopy/microscope
TD	threading dislocation
TDS	thermal diffuse scattering
TEM	transmission electron microscopy/microscope
UFF	universal force field
VFF	valence force field
WPOA	weak phase object approximation
WZ(wz)	wurtzite (structure)
ZB(zb)	zincblende (structure)
ZOLZ	zeroth-order Laue zone

Contents

Chapter 1.

Introduction

An essential conceptual philosophy to realize multi-functional materials is the principle of combination. In a simple scheme, structurally combining a material building block of property B into a host of property A can create, in principle, a composite with novel property "A and B". This combination has been proven a successful route, especially in conductors (metal alloys) and insulators (ceramic composites), by millions of man-made structural and functional composite materials that have been applied to benefit our daily life. In the world of semiconductor materials, it is much more fascinating. Because semiconductor stands in between insulators and conductors so that its properties are easily affected and varied by any impurities: doping of only very tiny fraction of elements in semiconductors will greatly alter, and/or integrate their functionalities.

Controlling the distribution of dopant atoms in a semiconductor is the key factor for tailoring its electronic, vibrational, optical, and magnetic properties. For host-"dopant" systems that have a miscibility gap, the control of "dopant" distribution is particularly challenging because phase separation may occur through spinodal decomposition[1–3], with uncontrolled consequences for the doped material. Such instabilities can often be circumvented using molecular beam epitaxy (MBE) and *in-situ* doping under non-equilibrium conditions. This approach opens the door to growing chemically immiscible "artificial alloys". The dilute magnetic semiconductors (DMS) are an important example of such materials[4,5].

DMS are promising materials for future spintronic applications, which make use of both the spin and the charge of electrons[6]. In real DMS material systems, the miscibility gap between the semiconductor host and the incorporated transition-metal or rare-earth element makes it difficult to preserve the "dilute" state and usually leads to spinodal decomposition phenomena[7,8]. As a result, nano-clustering has been widely reported in DMS materials[9–11], in which few percent of magnetic elements are incorporated, even when they are grown far away from thermodynamic equilibrium with techniques like MBE. *Ab initio* studies have also revealed a particularly strong tendency of DMS to form non-random alloys[12–15]. As predicted, depending on the incorporated concentration and growth parameters, the magnetic element-rich phase could decompose in forms of small clusters, columns or complex network structures[5]. In addition, the solubility of transition-metal or rear-earth elements in group III-nitride semiconductor materials is believed to be smaller than in group III-arsenide material systems[8]. Therefore the nano-scale decomposition is supposed to play a crucial role for DMS materials.

Controlling the composition, stability, and ultimately the performance of these artificial alloys requires a fundamental understanding of the non-equilibrium processes that govern their growth. This understanding begins with an accurate description, at the atomic level, of the structure of the artificial alloys.

1.1. Background and motivation of the work

Gadolinium-doped gallium nitride (GaN:Gd) is a very unusual DMS material, which attracted lots of attention since the reports by Dhar *et al.*[16,17]. It is ferromagnetically ordered above room temperature even at very low Gd concentrations of 10^{15} cm^{-3}. Even more strikingly, the magnetic moment per Gd has been measured to be several thousand Bohr magnetons[16]. The origin of these unusual magnetic properties is not yet clear. Many theoretical models have been developed[18–27] based on the assumption that the Gd dopants are uniformly dispersed in GaN:Gd, notwithstanding the large miscibility gap between GdN and GaN[28]. However, this assumption has little if any experimental support. The knowledge about the details of the material structure, especially the distribution of the Gd atoms in the GaN matrix, is the key to further understand the properties of this material. So far, there is no such experimental information available, because the direct detection of Gd "dopant" atoms and their spatial distribution at such low concentrations is extremely challenging, due to the following reasons.

Using most spectroscopic methods, the interpretation of material structure requires appropriate physical model (and simulations). The results are usually not unique. Moreover, they yield only partial information about material structures. For example, previous dedicated X-ray absorption[29] and X-ray magnetic dichroism studies[30] on MBE grown GaN:Gd samples with Gd concentration in the range of 10^{17}–10^{19} cm^{-3} have only demonstrated that Gd atoms are on Ga substitutional sites, without further information about whether they are clustered or not.

On the other side, microscopic structure probing relies on the basic probe-specimen interaction which is limited to a certain resolution and sensitivity. In some reports using synchrotron X-ray diffraction, the authors claimed that second phase rocksalt GdN is detected[31,32] in MBE grown GaN:Gd samples with Gd concentration above 0.3 % (about 1.3×10^{20} cm^{-3}). Below this concentration they were not able to detect it and they believed Gd atoms dilutely dispersed in GaN[31,32]. This is maybe the main reasons to the aforementioned general assumption that Gd atoms are dilutely dispersed in GaN in low Gd doping concentration samples. Up to now, no convincing data have been reported about the distribution of Gd in samples with Gd concentration lower than 10^{20} cm^{-3} due to the detection limit for most spectroscopic and/or microscopic techniques.

The ultra-high spatial resolution of transmission electron microscopy (TEM) and scanning TEM (STEM) enables the investigation of very localized structure features down to the detection of single dopant atoms[33–39]. Neglecting the fact that these studies rely on high performance microscope, further more, all the fancy examples reported up to now are rare and extreme cases which require either a high ratio of atomic Z-number between imaged atom(s) and the matrix ($Z^{Gd}/Z^{C} = 10.67$ in Ref. 33, $Z^{Au}/Z^{Si} = 5.64$ in Ref. 34,37), or extremely thin specimen (monolayer graphene[36,38,39], monolayer BN[35] or carbon nanotube[33]). In addition, it has been shown[37] that for embedded clusters or atoms in a matrix, the channeling effect from atomic rows greatly reduces the scattering signal, which preclude the doped atoms or clusters from being observed. Therefore, in the case of Gd-doped GaN ($Z^{Gd}/Z^{Ga} = 2.06$) where Gd atoms are embedded in a 3-D GaN matrix, it is extremely challenging to observe the distribution of Gd atoms directly (see also more discussions in section 4.4 on page 67), not even to mention the extremely low Gd concentration and the existence of other disturbing factors like plenty of crystal defects in the epi-layer.

It might be almost impossible to detect single Gd atoms in the GaN matrix, nev-

ertheless, if they were clustered, we could probably detect them by the local strain (or displacement) fields produced by the large Gd atoms (atomic radius: $r_{Gd} = 233$ pm, $r_{Ga} = 136$ pm and $r_N = 56$ pm). In principle, we can image the strain field induced by Gd-related clusters using strain contrast imaging techniques and compare with simulations from potential models. In higher resolution (with limited field of view), although the Gd concentration is extremely low in the samples, we should still have chance (of about one thousandth to one hundredth, see chapter 6.1 in page 101) to find high-resolution TEM (HRTEM) images at thin enough regions that might contain Gd-related cluster to observe the local lattice distortions. In the latter case, we can extract local displacement information by performing quantitative geometric phase analysis (GPA), which has demonstrated to measure the lattice distortion with high accuracy (see Ref. 40 and also in section 3.4 in page 38). Because of the inherent limitation by the principle of the GPA method, the application of GPA to the analysis of nano-structures requires at least distinctable lattice pattern(s) or fringe(s) from the nano-object that is being analyzed. There are already a bunch of examples of nano-objects analyzed by GPA, like heterostructure interfaces[41–43], quantum well superlattices[44,45], quantum dots[46,47], 3D nano-particles[48,49], nano-scale semiconductor device structures[50,51], etc. However, the validity to apply to clusters consist of few tens of atoms is not guaranteed[52,53]. It therefore required for a thorough discussion, experimentally and theoretically, about the limitations before apply for quantitative data interpretation.

Motivated by establishing a detailed structural study of GaN:Gd samples, the central goal of this thesis is (1) to clarify the distribution of Gd atoms and their detailed local atomic structure in low-Gd-concentration (lower than 10^{20} cm^{-3}) GaN:Gd samples using complementary (S)TEM imaging techniques and analysis methods; and based on that (2) to understand the mechanism that control such kind of distribution. Encompassing this goal, the (S)TEM imaging methods dedicate to the problem as well as the limitations and application of analysis method (i.e. GPA) are explored in-depth. Besides, two case studies of semiconductor nano-clusters, whose size and density apparently detectable in (S)TEM, are investigated for a better understanding of structures of clusters and the phenomena of clustering in semiconductor epilayers.

1.2. Structure of the thesis

In **Chapter 2**, fundamental knowledge to understand nano-clusters' structure and cluster formation in semiconductor materials is covered. It is considered from two aspects: (1) the microstructure aspect of nano-clusters embedded in semiconductor epilayers that includes crystal structure, elastic properties, interface character and extended defects; (2) the theoretical aspect of clustering which summarizes basic ideas about diffusion, clustering mechanisms and the discussion of kinetic limitations of clustering.

Chapter 3 covers the theoretical and practical aspects of (S)TEM, the central experimental technique. This chapter starts with a short description of the construction of electron microscopes. Then it follows the detailed theoretical description of electron-crystal interaction for structure detection: from kinematic approach to dynamic approach of electron diffraction. This consists of the basis to quantitative interpretation of TEM images. Different modes of image formation in electron microscopes is summarized. It opens the door to contrast simulation for a quantitative image interpretation. Finally, the GPA method is introduced and extensively discussed in detail with examples. It covers the limitations and application: about the accuracy, resolution, sources of artifacts and a

practical guide line.

In **Chapter 4**, the practical side — using TEM to observe and quantitatively analyze nano-clusters in semiconductor epilayers — is discussed in detail with studied examples. Firstly, the interpretation of various images contrast for different types of nano-clusters are covered (coherent, semi-coherent and incoherent). This provide the comprehensive information to actively exploit the apropriate method to explore an unknown structure. In the following, the methods [based on valence force-field (VFF) method and on density functional theory (DFT)] of structure modeling for image simulation and quantitative image interpretation are introduced and compared. Lastly, the detectability of several modeled Gd and GdN cluster structures are discussed with image contrast simulations. It provides the practical guideline to ultimately detect the small GdN clusters in GaN:Gd.

Chapter 5 consists of two case studies about nano-clusters, whose size and density are apparently detectable in TEM, to help understanding the microstructure of nano-clusters and clustering phenomena in GaN:Gd. In the first case, InAs quantum dots buried in defect free Si matrix are analyzed in detail by HRTEM and GPA strain analysis. With this static picture, we emphasize on the interface structure and the importance of local strain on the cluster shape. In the second case, the formation and structure transformation of Bi-containing cluster in annealed epitaxial $GaAs_{1-x}Bi_x$ thin films are studied. Based on static snapshots of TEM observations and knowledge about microstructure evolution, a dynamic picture of cluster formation and structure transformation is deduced.

In **Chapter 6**, finally, the goal to identify the occurrence of GdN clusters in GaN:Gd samples and determine their atomic structure is detailed. By combination of (S)TEM imaging techniques, quantitative analysis and contrast simulations, the occurrence of GdN platelet clusters in the GaN:Gd thin films is identified. Through comparing the displacement field between experimental results and energetically favored models based on DFT, we present the way to extract the local atomic structure model of the clusters. This atomic structure model enabled our discussion about the energetics of the clusters by DFT calculations in conjunction with the classical Frenkel-Kontorova model.

In **Chapter 7**, a general conclusion is drawn together with outlook.

Author's contributions

The author has independently carried out the original work including the discussions about limitation and application of GPA method in Chapter 2 and most of the contents presented in Chapter 4–6: all the (S)TEM related experiments and discussions, data processing and interpretation, contrast simulations as well as structure modeling using valence force-field method. The density-functional theory (DFT) calculations and the corresponding interpretations were performed by Steven C. Erwin at Naval Research Laboratory in Washington D.C. through institutional collaboration, which is greatly acknowledged by the author. The studied samples were grown and provided by third part, which are listed in Appendix C and credited separately.

Chapter 2.

Fundamentals

Semiconductors include a wide range of materials, that exhibit an energy band gap. The term "semiconductors" in this thesis is restricted to the conventional semiconductor crystals (e.g., Si and group III-V compound semiconductors). While the word "nano-cluster" refers a static picture of a solid phase object with its size in nanometer scale, we use the term "nano-clustering" to address the dynamic process of formation of nano-clusters [1]. In this chapter, some basic concepts related to nano-clustering in epitaxial semiconductor thin films will be summarized. Firstly, the microstructural aspects will be presented: from basic crystal structure, extended defects, elasticity to interface properties. Because nano-clusters in semiconductor epilayers can be considered as embedded 3D heterostructure, a lot of concepts and definitions is transferable from thin film (planar) heterostructure, whereas the interface is closed curved in the case of nano-clusters. Next, the general theoretical consideration of nano-clustering semiconductor crystals will also be introduced. These remarks are the basis for later discussions on cluster formation and their phase stability.

2.1. Microstructural aspects of semiconductor nano-clusters

2.1.1. Semiconductor crystal structure

Diamond, zincblende and wurtzite structure are the most common structures for group IV and III-V semiconductor crystals, where the atoms are tetrahedrally bonded to the nearest neighbors forming a (near) close packed stacking sequence. Diamond structure is single element face-centered cubic (*fcc*) stacking. It has two basis atoms in each primitive unit cell with positions $(000)a$ and $(\frac{1}{4}\frac{1}{4}\frac{1}{4})a$, respectively. Zincblende structure is the same as diamond but different types of atoms occupy the two positions in primitive cell. Wutzite structure is a hexagonal close-packing (*hcp*) stacking with two atom spicy. The two atoms sit in position $A(\frac{1}{3}\frac{2}{3}0)a$ and $B(\frac{1}{3}\frac{2}{3}\frac{3c}{8a})a$, with space group $P63mc(186)$. The structures are depicted in Fig 2.1. In diamond and zincblende structures, the four equivalent $\langle 111 \rangle$ directions can be regarded as the close stacking directions, whereas in wurtzite structure, the $\langle 0001 \rangle$ are the close stacking directions [2]. It is clear that diamond structure is centrosymmetric, while wurtzite structure and zincblende structure are non-centrosymmetric. When the two atoms switch their position in wurtzite structure, it will result in different polarity.

[1] Nano-clusters in some cases can be also used as quantum dots. The term "quantum dots" stress the quantum confinement in the electronic properties of such structures. In this thesis we prefer to use the term "nano-clusters" to stress their microstructural character, although in few cases (and we note in some literature) both term are used with a mixed meaning.

[2] for Miller and Miller-Bravais index conversion, see Appendix A.2

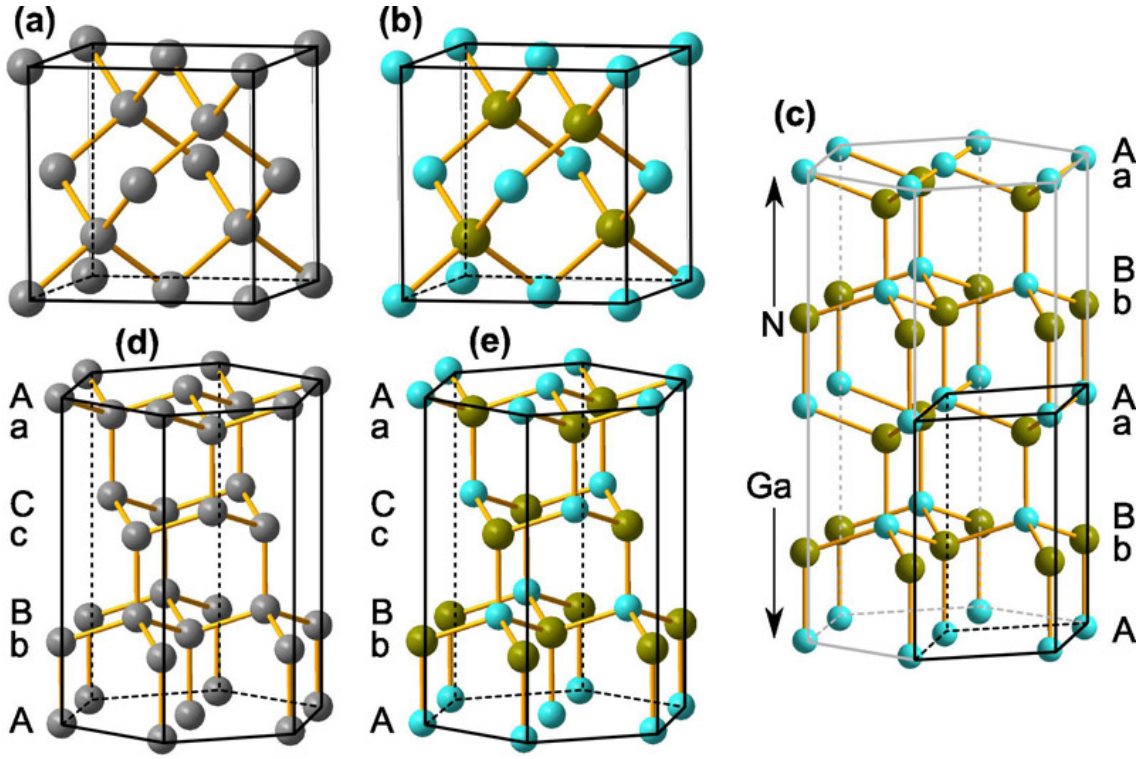


Figure 2.1: Illustration of the diamond (a), zincblende (b) and wurtzite (c) structure. For better visualize, the diamond and zincblende structure is also represented along the stacking direction in hexagons (d) and (e).

Beside the normal A and B atom sites, there are other spaces that atoms can potentially occupy. They are interstitial sites. In wurtzite GaN, the most important interstitial sites are the octahedral position (in the middle of 6 closest atoms) and tetrahedral position (in the middle of 4 closest atoms) which has the most available volume[54].

2.1.2. Extended defects

Defects in solids are the imperfections of their crystal structure. According to the dimension of the defect form, there are three categories: point defects (zero-dimension, 0-D), line defects (1-D), planar defects (2-D), and volume defects (3-D). Except for point defects, the other defects are also referred as extended defects. In epitaxial semiconductor layers, the most important extended defects are the line defects and planar defects. Nano-clusters can also be viewed as volume defects.

Dislocations Dislocations, also known as line defect, are extended in crystal in one dimension, i.e. they follows either straight of curved lines. Dislocations are created by the distortion of certain atomic planes that are distorted out of their regular position due to external or internal stresses. Dislocations are characterized by the their line direction $\mathbf{u}$ and Burgers vector $\mathbf{b}$. The Burgers vector represents the magnitude and direction of the lattice distortion of dislocation in a crystal lattice. It is geometrically defined as the displacement of an atom on the Burgers circuit that around the dislocation to its regular position, as illustrated by the arrow in Fig. 2.2. According to the relative angle between $\mathbf{u}$ and $\mathbf{b}$, three types of dislocation can be categorized, namely edge-type ($\mathbf{b} \perp \mathbf{u}$), screw-

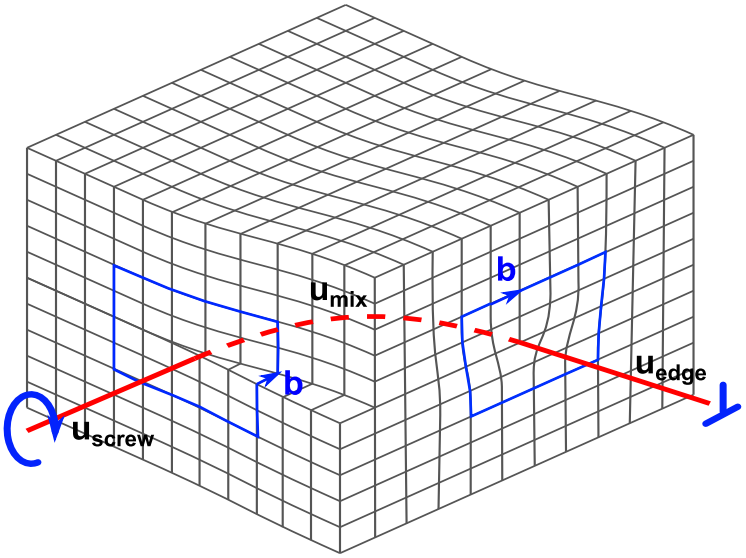


Figure 2.2: Schematic illustration of dislocations. The blue line indicate the Burgers circuit defining the vector **b**.

type ($\mathbf{b}\|\mathbf{u}$) and mixed-type [$\angle(\mathbf{b},\mathbf{u}) = (0°, 90°)$] dislocation containing of both screw and edge character.

Self-energy of dislocation The self-energy of a dislocation E_{disl} is the energy to create a unit length of dislocation. It includes the part from the core of dislocation E_{core} where there are dangling bonds, and the elastic part E_{el} from the long range elastic contribution of the media. The atomic distortion introduced by dislocation is extended in a long range around the dislocation line. The strain field can be pictured as co-centric cylinder with dislocation line as the center. For straight dislocation line in isotropic solid media, the elastic energy stored in the distorted surrounding media. The general expression of elastic energy for unit length of dislocation, using the continuum elasticity approach (refer to section 2.1.3), is[55]

$$E_{el} = \frac{Gb^2(1-\upsilon\cos^2\theta)}{4\pi(1-\upsilon)}\ln\left(\frac{R}{r_0}\right), \quad \text{or simply,} \quad E_{el} \propto b^2. \tag{2.1}$$

G is the shear modulus, ν is the Poisson ratio, b is the magnitude of the Burgers vector, R is the radius of the long range strain field centered at the dislocation line, r_0 is the radius of the core of dislocation and θ is the angle between the Burgers vector and the dislocation line, characterizing the edge or screw nature of the dislocation. It is clear that the elastic energy for unit length dislocation is proportional to square of the magnitude of Burgers vector. Therefore, the Burgers vector is an important quantity that describes dislocations.

Beside threading dislocations in thin films that tend to stretch straight, dislocations can also form closed loop, especially in heterostructure of 3D embedded nano-clusters. For a circular dislocation loop, the self-energy is well-established[56] with the following equation:

$$E_{DL}(R_{DL}) = \frac{Gb^2R_{DL}}{2(1-\nu)}\cdot\ln\frac{8\alpha R_{DL}}{b}, \tag{2.2}$$

Table 2.1: Summary of stacking fault energy [Intrinsic (I), Intrinsic type-I (I_1), Intrinsic type-II (I_2) and Extrinsic (E)] of semiconductors and some metals, in mJ/m^2.

		GaN[59]	GaAs[60]	InAs[60]	Si	Cu	Al
I	I_1	18.1	55 ± 5	30 ± 3	~ 50	$30 \sim 40$	$160 \sim 200$
	I_2	43.4					
E		68.8			~ 60		

in which R_{DL} is the radius of the dislocation loop, and $\alpha = 0.25$ accounts for the energy of dislocation core. Although semiconductor crystals are actually anisotropic, the isotropic approximation within continuum elasticity approach can still yield reasonable results for quantitative description and understanding of their microstructure[57].

Motion of dislocation Dislocations can move in respond to internal and/or external stress. They can move in two ways: glide and climb. Glide is the movement of dislocation in the plane which contain both its line and Burgers vector. Glide of many dislocation result in slip. For edge type dislocation, the Burgers vector and dislocation line define a specific plane, so it can move only in that plane. For screw type dislocation, the Burgers vector and line dislocation line is parallel, so it can move in any plane that contain the dislocation. Climb is the motion of dislocation out of the glide surface normal to the Burgers vector. Climb is achieved by absorption or release of point defect.

Stacking fault Stacking fault is a type of planar defect that usually encountered in semiconductor crystals. In *fcc* crystals two types of stacking fault could be distinguished, suggested by Frank[58], intrinsic- and extrinsic-type. The intrinsic stacking fault is equivalent to remove one stacking layer

$$\ldots ABCAB \mid ABCABC \ldots$$

or it can also treated as shear the other part of the crystal by $\frac{1}{6}[\bar{2}11]$ from the faulted plane, then $C \to A$, $A \to B$ and $B \to C$. The extrinsic stacking fault is equivalent to insert one stacking layer

$$\ldots ABCAB\overline{A}CBCABC \ldots$$

In *hcp* crystals there are three types of stacking fault, namely, intrinsic type-1 (I_1-type), intrinsic type-2 (I_2-type) and extrinsic type (E-type). The E-type in *hcp* is similar to that in *fcc*. The I_1-type can be formed by removing one of the plane and subsequently shear the remaining planes above that plane by the displacement $\frac{1}{6}[2\bar{2}03]$:

$$\ldots ABABAB \mid CACACA \ldots.$$

The I_2-type can be formed by directly shear of displacement $\frac{1}{3}[1\bar{1}00]$:

$$\ldots ABABAB \mid ACACAC \ldots.$$

The formation energy of intrinsic stacking fault is usually lower than extrinsic one.

For convenience, typical values of stacking fault energy of semiconductors (and selected metals for comparison) from literature (experimental or theoretical) are summarized in Tab. 2.1

Partial dislocation and stacking fault Perfect dislocation refers to the dislocation with a Burgers vector equal to the translational vector of the crystal, and those with Burgers vector not equal to translational vector are referred to as partial dislocations. The creation of stacking fault corresponding to a translation vector that is not equal to the perfect crystal translation vector, partial dislocation have to bounded at the periphery of stacking fault, or extend to free surface. The dissociation of perfect dislocation (because of reduction in energy) will spit to partial dislocations and stacking fault in between. The distance between dissociated dislocation providing the possibility to deduce stacking fault energy[60].

2.1.3. Continuum elasticity

Continuum elasticity neglect the fact that solids are discrete at atomic level. Nevertheless, it has been shown to provide reasonable accurate prediction even for objects like embedded quantum dots with size of only few nm scale[57]. Elasticity refers the reversible change in atomic displacement in response to the applied stress. Stress is the force per unit area that is acting on a surface of the solid. The denotation σ_{ij} refers stress along i direction applied on plane with normal j. In response to the external stress, solid material will be deformed either elastically (reversible) or plastically (irreversible). This deformation of the solid materials due to applied external stress is measured in strain, which is defined by

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{2.3}$$

where u_i are the components of distortions with a displacement at a point $\mathbf{r}(x_i, x_j, x_k)$.

Stress and strain In static equilibrium state, $\sigma_{ij} = \sigma_{ji}$. The stress is therefore symmetric and has a total number of six elements which is represented in a second rank tensor. For small amount of distortions $\partial u_i/\partial x_j$, the stress and strain are linearly related, i.e., by the Hooke's law. Because stiff rotations of an element $\omega_{ij} = \partial u_i/\partial x_j - \partial u_j/\partial x_i$ cannot give rise to stresses in the absence of internal torques. Therefore stresses depend only on the strain, and can be expressed as

$$\epsilon_{ij} = S_{ijkl}\sigma_{kl}, \quad \text{or consistently,} \quad \sigma_{ij} = C_{ijkl}\epsilon_{kl},$$

where S_{ijkl} are the elastic compliance coefficients and C_{ijkl} are the elastic stiffness coefficients. The elastic compliance coefficients S_{ijkl} is the mutually inverted tensor of elastic stiffness coefficients C_{ijkl}. The stress and strain are conventionally written as first rank vectors with six independent (reduced from the nine elements because of equivalence of three shear elements) elements according to the Voigt's notation. Then the S_{ijkl} and C_{ijkl} are of second rank matrix. The first two suffixes are abbreviated into a single one running over from 1 to 6, the last two are abbreviated in the same way, according to the replacement scheme from tensor notation to matrix notation [$11 \to 1, 22 \to 2, 33 \to 3, (23, 32) \to 4, (31, 13) \to 5$ and $(12, 21) \to 6$], the constants S_{ijkl} and C_{ijkl} is denoted as S_{mn} and C_{mn}.

Then, the equation [1.8] can be expressed in matrix form

$$\begin{bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \epsilon_{33}\\ \epsilon_{23}\\ \epsilon_{31}\\ \epsilon_{12}\end{bmatrix} = \begin{bmatrix}S_{11} & S_{12} & S_{13} & 0 & 0 & 0\\ S_{12} & S_{22} & S_{23} & 0 & 0 & 0\\ S_{13} & S_{23} & S_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & S_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & S_{55} & 0\\ 0 & 0 & 0 & 0 & 0 & S_{66}\end{bmatrix} \begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix}. \tag{2.4}$$

A similar matrix scheme form can be written for $[\sigma_{ij}] = [S_{mn}][\epsilon_{ij}]$. It is clear that $[C_{mn}] = [S_{mn}]^{-1}$, where $[S_{mn}]^{-1}$ denotes the inverse matrix of $[S_{mn}]$.

For cubic and hexagonal symmetry, independent coefficients are reduced. In cubic symmetry there are only three independent coefficients: $C_{11} = C_{22} = C_{33}$, $C_{12} = C_{13} = C_{23}$ and $C_{44} = C_{55} = C_{66}$. In hexagonal symmetry there are only six independent coefficients: $C_{11} = C_{22}$, C_{33}, C_{12}, $C_{13} = C_{23}$, $C_{44} = C_{55}$ and $C_{66} = 2(C_{11} - C_{12})$.

In elastic isotropic media, there are only two independent elastic constants. It is common to express the isotropic elastic property as shear modulus $G = C_{44} = (1/2)(C_{11} - C_{12})$ (characterizing the rigidity of materials) and Lamé constant $\lambda = C_{12}$. It is also convenient to express in other constants combination: Young's modulus (characterizing the stiffness), Bulk modulus (characterizing the compressibility of materials) and Poisson's ratio (characterizing the negative ratio of transverse to axial strain). The inter conversion of these constants can be found in Ref. 56.

Elastic anisotropy In elastic anisotropic crystals, it is more convenient to express the elastic properties by the direction dependent modulus characterizing the material response to specific stress (i.e. Young's modulus, Poisson ratio and other constants that be simply converted from these two) instead of using the mathematical expressions of stiffness tensor and/or compliance tensor. The conversion of elastic constants to direction dependent Young's modulus and Poisson ratio along the (hkl) plane normal for cubic and hexagonal crystal systems can be found in Ref. 61 and Ref. 62. For convenience, the conversion for hexagonal system is adopted:

$$E(hkl) = \frac{(\Delta + \Pi)^2}{S_{11}\Delta^2 + S_{33}\Pi^2 + (S_{44} + 2S_{13})\Delta\Pi}; \quad \Delta = h^2 + \frac{(h+2k)^2}{3}; \quad \Pi = \left(\frac{a}{c}l\right)^2 \tag{2.5}$$

and

$$\nu(hkl) = \frac{-(\Delta + \Pi)(S_{12}\Delta + S_{13}\Pi)}{S_{11}\Delta^2 + S_{33}\Pi^2 + (S_{44} + 2S_{13})\Delta\Pi}; \quad \Delta = h^2 + \frac{(h+2k)^2}{3}; \quad \Pi = \left(\frac{a}{c}l\right)^2 \tag{2.6}$$

It can be seen that within the basal plane, the elastic properties are isotropic. As an example, the direction dependent Young's modulus for Si, InAs and GaN are plotted in Fig. 2.3. The close packing direction, for Si and InAs (i.e. $\langle 111 \rangle$ directions) has the highest modulus; while for GaN (i.e. $\langle 0001 \rangle$ directions) is the smallest.

Elastic strain energy The work done by the stress component σ_{ij} acting on an elastic elemental body by differential strain increments $\mathrm{d}\epsilon_{ij}$ is $\mathrm{d}w = \sigma_{ij}\mathrm{d}\epsilon_{ij}$. The strain energy density function is the integral of $\mathrm{d}w$, which is $w = (1/2)\sigma_{ij}\epsilon_{ij}$. Then the elastic energy

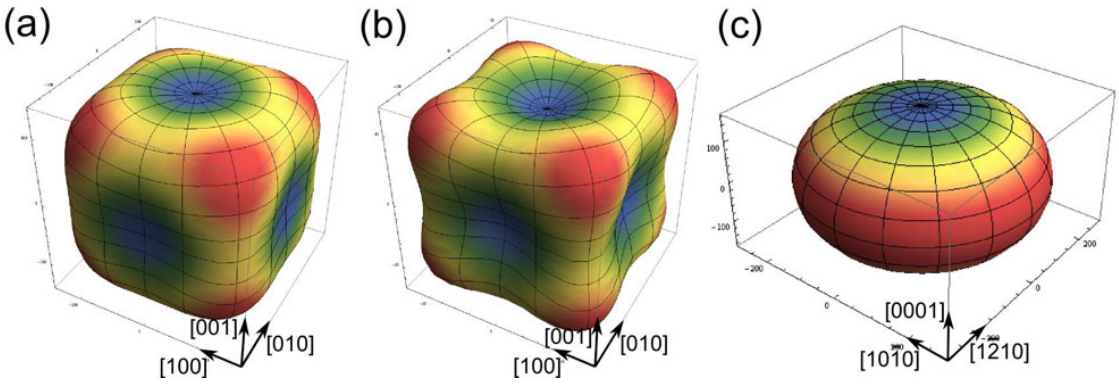


Figure 2.3: The anisotropic (direction dependent) Young's modulus of (a) Si, (b) InAs and (c) GaN (plotted along the three major crystal axis as three dimensional cube).

E_{el} stored in volume V is:

$$E_{el} = \int \frac{1}{2} \mathrm{d}V \sum_i \sum_j \sigma_{ij} \epsilon_{ij}. \tag{2.7}$$

2.1.4. Heterostructure interface

Heterostructure interface determines the orientation relationship between the substrate (matrix) and the epilayer (nano-cluster), and accordingly determines the lattice mismatches along different directions as well as the specific ways of mismatch strain relaxation. As a result, interface could considerably affect the growth (formation) behavior and the structural properties, which in turn affect the performance of devices based on such heterostructure. Although in this thesis the interfaces between nano-clusters and the matrix are more complex, the basic concept are quite similar with the simple model of infinite flat heterostructure films.

Lattice misfit A very important concept of heterostrucutres is the lattice misfit. There are various definitions in literature. The following definition is adopted in this thesis:

$$f = \frac{d_c - d_m}{d_m} \tag{2.8}$$

where d refer to the epitaxial plane spacing and the subscript c and m stand for the epilayer (nano-cluster) and substrate (matrix).

Interface coherency Depending on the atomic registry relationship across the interface, heterostructure interface is categorized into three groups, coherent, semi-coherent and incoherent interfaces, as illustrated in Fig. 2.4.

Coherent interface defines a continuous atomic lattice plane correspondence across the interface, and no misfit dislocation is contained across the hetero-interface; the misfit is accommodated by elastic strain, as illustrated in Fig. 2.4(a). Such system is very common in heterostructure with small (usually $f < 0.02$ for epilayer) lattice misfit and film thickness. In this case, the atoms cross the interface will be strained to match the substrate (matrix) lattice. The elastic strain energy (caused by this misfit stress) will monotonically increase with the radius (volume) of cluster, or the thickness of epilayer.

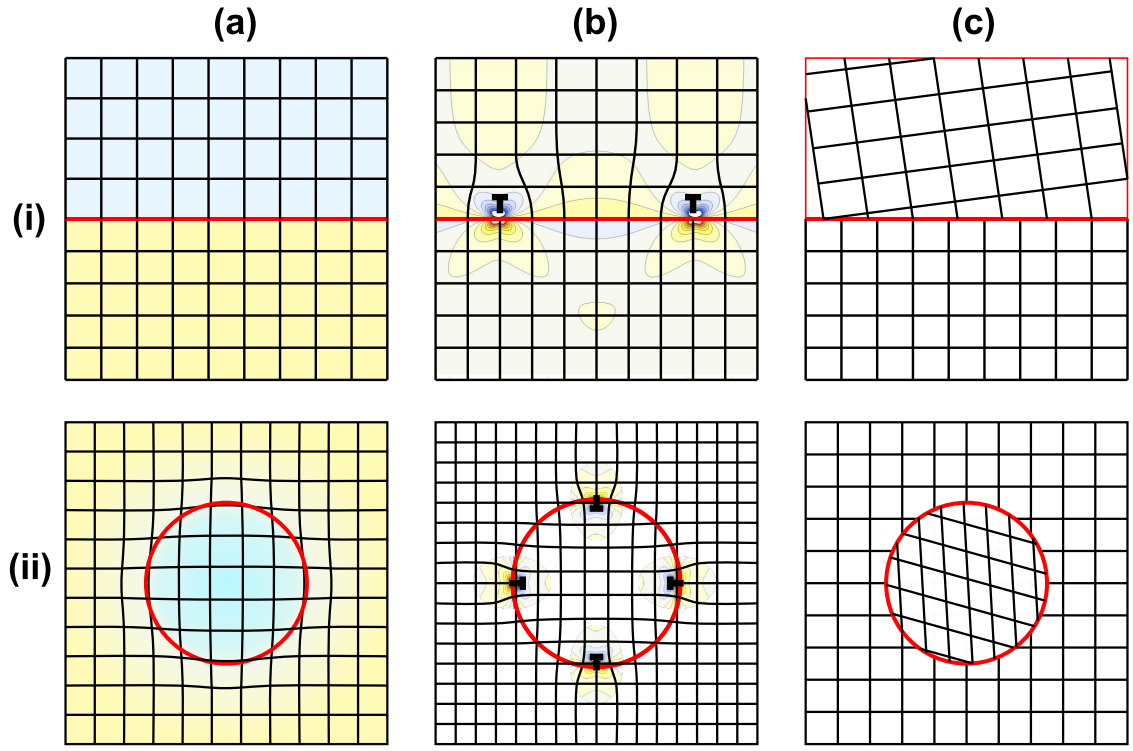


Figure 2.4: Illustration of (a) coherent, (b) semi-coherent and (c) incoherent interface of hetereo-interface. Raw (i) and (ii) are the scheme for planar interface and the analogy of cluster interface, respectively. Color depth illustrate the misfit stress distribution.

When the strain energy is too high, the coherency will break, part of the misfit strain will accommodated by the generation of misfit dislocation as depicted in Fig. 2.4(b). In case of cluster, dislocation loops are created. Such interface is semi-coherent. There is a critical volume of cluster when the interface will transform from coherent to semi-coherent. It is usually measured by the critical radius. Chaldyshev *et al.*[63] has constructed the scenario of stress relaxation in buried quantum dots based on continuum elasticity approach. They showed that the fingerprint of this scenario is the formation of specific satellite dislocation loops (SDLs) in a vicinity of the quantum dots. Configuration of the SDL appeared to be specific when compared to both dislocation loops punched by large inclusions and dislocations formed at relaxed surface islands[63].

2.2. Theoretical aspects of clustering in semiconductor epilayers

Clustering in solids is a diffusion or nucleation controlled kinetic process. In this section, the diffusion mechanisms in semiconductor crystals, (thermodynamic) driving forces and mechanisms of clustering and kinetic controlling parameters that can lead to cluster formation are briefly summarized and discussed.

2.2.1. Diffusion in semiconductor crystals

Atomic diffusion refers to the migration of atoms in space, primarily due to the thermally activated motion. In a steady state, for a system without constrains (inter-atomic interaction), a macroscopic view of diffusion is described by the Fick's laws of diffusion:

$$\text{first law:} \quad J = -D\nabla c; \qquad \text{second law:} \quad \frac{\partial c}{\partial t} = D\nabla^2 c,$$

which are results of the second law of thermodynamics, i.e., the law of entropy maximization. In the equations, J is the diffusion flux, c [or $c(\mathbf{r})$] is the (position dependent) concentration function, and D the diffusivity. The first law describes that the diffusion flux is proportional to the concentration gradient. The second law, which is derived from the first law using the mass conservation in absence of chemical reactions, predicts how diffusion causes the concentration to change with time.

With a microscopic view, diffusion in solid crystal is a random jump of atoms in the lattice. The jump rate is a temperature dependent process. Generally, the temperature dependence of diffusivity follow the Arrhenius relation:

$$D = D_0 \exp(\frac{-Q}{k_B T}), \tag{2.9}$$

in which Q is the activation energy and k_B is the Boltzmann constant.

Usually, the diffusivity of atoms in solid crystal is much lower than in liquid or gas because of the relative more condensed surrounding bonding environment. An elementary jump of the atom needs first to overcome the activation energy barrier to "stretch" the surrounding bonds. For the same type of atom diffusion, the activation energy varies greatly when the atoms diffuse by different mechanisms.

Diffusion mechanisms are recognized by two categories: defect and non-defect (i.e. self-atom) diffusion. There is a general consensus that non-defect does not play any significant role in semiconductor diffusion[64]. Defects diffusion mechanisms include vacancies, interstitials, interstitial-substitutionals (or, kick-out) and some other variants of interstitials, which usually hold relative high energy barriers. Generally (and intuitively), small foreign atoms tend to take interstitial positions and diffuse relatively faster in the interstitial space of crystal lattice. Larger atoms diffuse slower and mostly with the assistance of vacancies. Additionally, in semiconductor materials, the diffusion energy barrier also depends on the charge state of the diffusion atoms and the Fermi level of the semiconductor material itself[64], therefore making the identification and understanding of the diffusion mechanisms very complex and difficult. While the experimental data are accumulating, there is merely few progress in the understanding of the microscopic diffusion mechanisms in semiconductors over the past half century. Detailed description of the established mechanisms can be found for example, in a review by Willoughby[65], and books (or book sections) by Shaw[64,66]. Collections of experimental data about diffusion

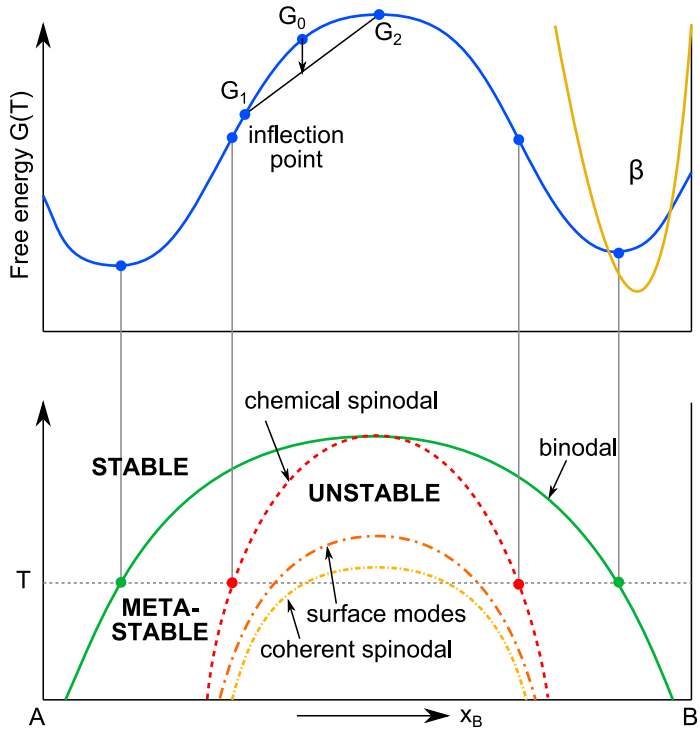


Figure 2.5: Free energy and phase diagram of a two composite A and B alloy system. Higher temperature will result in smaller miscibility gap. (See e.g., Ref. 3)

in silicon and III-V semiconductors can be found in books by Tuck[67] and Fisher[68].

Beside bulk diffusion, there are other high diffusivity channels. These channels include free surfaces[69,70], grain boundaries, and open core of dislocations (pipe diffusion)[71,72]. The diffusivity of atoms in these cases can be several orders of magnitude higher than diffusion in bulk material.

2.2.2. Mechanisms of clustering in crystals

The clustering from initially homogeneous alloy is a process of phase (trans)formation and its further evolution. At constant temperature and pressure, all phase transformations are driven by reduction in the Gibbs free energy from the original to the final state: $G = H - TS$ (H is enthalpy, T temperature and S entropy). In equilibrium the Gibbs free energy has its minimum (or, $\Delta G = 0$). Consider a binary alloy, the derivative of Gibbs free energy is

$$\Delta G = \Delta H_{mix} - T\Delta S_{mix}. \tag{2.10}$$

In an ideal alloy with zero mixing enthalpy, the free energy will always be lower in mixed case because entropy favors mixing (increase the disorder). In regular solution, enthalpy of mixing can not be neglected. There are several theoretical approaches accounting for this contribution. One of the most widely used approach especially for III-V semiconductors is the so-called quasi-chemical model[73], where the change in enthalpy of mixing can be accounted by the change of bond types between the atoms in the alloy.

A schematic sketch of Gibbs free energy of a binary alloy system with miscibility gap at temperature T and the phase diagram is show in Fig. 2.5. The curve connecting the

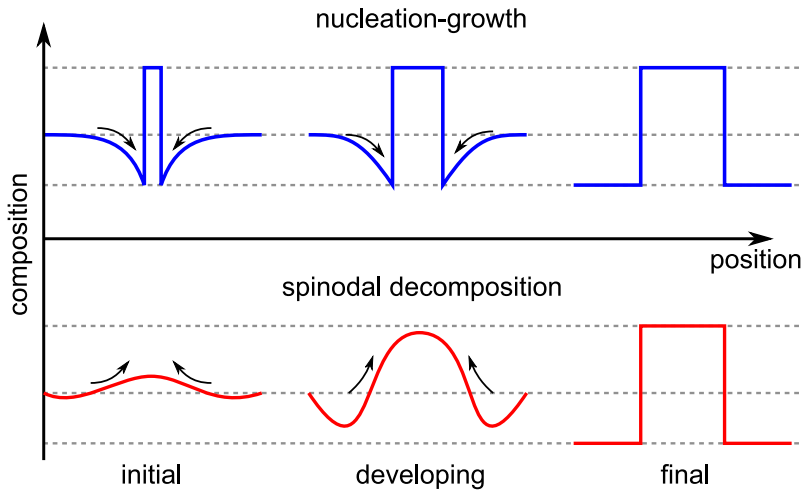


Figure 2.6: Scheme of phase separation by means of nucleation-growth mechanism and by mechanism of spinodal decomposition. (See e.g., Ref. 74)

common tangent of free energy is called the binodal curve, whereas the one in between the inflection point (i.e. where $\partial^2 f/\partial c^2 = 0$) on the free energy is called the spinodal curve. Outside the binodal curve, the system is stable against any amplitude of fluctuation in composition, because the entropy of mixing favors the intermixing state (or, the chemical potential drives a downhill diffusion). The "downhill" diffusion refers the net diffusion of atoms flux along the downgrade of composition gradient. In the spinodal region, there is no energy barrier for phase separation (diffusion barrier is not accounted). Any infinitesimal fluctuation in the composition will result in a reduction of total free energy, as illustrated in Fig. 2.5, $G_0 < G_1 + G_2$. This energy different will drive an uphill diffusion, the alloy in this region will spontaneously separate into two phase, provided that diffusion is not kinetically limited. The scheme of the two type of phase separation process is illustrated in Fig. 2.6

The model of spinodal decomposition Practically, phase separation by means of spinodal decomposition occurs in binary alloys where the separated two phases are similar in crystal structure and close in lattice parameters. Because in this case, the evolving of free energy curve with temperature is a smooth function, and the free energy of the separated two phase lies in the same curve. In the theoretical framework introduced by Cahn and Hillard[75], the development of spinodal decomposition is quantitatively described by the modified diffusion equation:

$$\frac{\partial c}{\partial t} = M\frac{\partial^2 f}{\partial c^2}\nabla^2 c - 2Mk\nabla^4 c + \text{non-linear terms}, \tag{2.11}$$

where M is a positive phenomenological diffusion mobility that can be experimentally determined, k is the gradient energy coefficient, which is the product of interaction energy and the square of interaction distance. Neglecting the nonlinear terms, this diffusion equation has the following solution:

$$c - c_0 = e^{R(\beta)t}\cos(\beta t), \quad \text{with} \quad R(\beta) = -M\beta^2\left(\frac{\partial^2 f}{\partial c^2} + 2k\beta^2\right), \tag{2.12}$$

where β is the wavenumber of fluctuation. There is a simple physical interpretation of the amplitude factor: the term outside the parenthesis account for the diffusion strength, and the term inside the parenthesis account for the strength of driving force for the diffusion. Since any type of concentration fluctuation can be expressed as summation of a group of sinusoidal wavelets by Fourier expansion, these equations can, in principal, describe any type of concentration fluctuation at least in the initial stage when the linear concentration gradient still holds.

When there is only quasi-chemical interaction between atoms, the limit of unstable curve is known as chemical spinodal. If there is lattice mismatch between the two phases, the phase separation need to overcome additional coherent strain energy. This is considered by Cahn in his equation by adding the negative strain energy part:

$$R(\beta) = -M\beta^2 \left(\frac{\partial^2 f}{\partial c^2} - \frac{2\delta^2 E}{1-\nu} + 2k\beta^2 \right), \tag{2.13}$$

in which δ is the volume mismatch, E is the Young's modulus and ν is the Poisson's ratio. This coherent strain energy can limit the chemical driving force and therefore stabilize the decomposition process. When the additional strain energy is taken into account in the free energy curve, the unstable region will be smaller than the chemical spinodal. This new curve is called coherent spinodal[1]. When the phase separation is proceeded on the surface, the strain could be partly relieved because of the surface effect. The region is larger than coherent spinodal but smaller than chemical spinodal, which is referred as "surface modes of spinodal decomposition[76]".

The nucleation-growth model Between the binodal and spinodal curve (cf. Fig.2.5), the composite alloy is metastable and can decompose only after nucleation of the other phase. This process is characterized by initial formation of a small precipitate of a new phase from within the matrix, such as the β-phase depicted in Fig. 2.5. The new phase, the nucleus, has a different composition and/or structure from the parent phase, therefore introduced an interface. The free energy of the system is discontinuous, and the process is a first-order phase transformation[77].

In the classical analysis by Gibbs[77–79], the discrete atom by atom transfer across the interface of nucleation and growth is described by a continuous model. Two assumptions are made in this classical theory: (1) the new phase behave the same as bulk and (2) the interface is same as infinite plat interface. Therefore, the local change in free energy is a combined contribution from a decrease in volume free energy, due to the transfer of atoms from a less stable to a more stable phase, and an increase in the interfacial free energy, due to the increase of the area of the interface between the two phases:

$$\Delta G(R) = -\frac{4}{3}\pi R^3 \mid \Delta F_V \mid + 4\pi R^2 \gamma. \tag{2.14}$$

R is the radius of nucleus, $\mid \Delta F_V \mid$ is the absolute value of free energy difference of the two bulk phases, γ is the interface energy. The theory qualitatively explained the critical nucleus' radius: the minimum of $\Delta G(R)$ corresponds to a definite value of R. However, both assumptions become questionable when nuclei of few nanometers, small enough that the center is not in the thermodynamic limit, and when the interfaces are sharply curved, changing their free energy.

In the study of nucleation-growth, there are difficulties in reproducibility of experi-

mental data macroscopically [77] and uncertainties in theoretical prediction [77,79,80]. Hence, nucleation theory is one of the few areas of materials science in which agreement of predicted and measured rates to within several orders of magnitude is considered a success. Therefore, in the experimental side, to the best detailed analysis about interface character, clusters composition and shape, in both microscopic and statistical perspective, are the most precious information that will promote further the understanding toward the fundamental process of nucleation and growth.

2.2.3. Kinetic limitations of clustering

Clustering in epitaxial semiconductor thin films can happen *in-situ* during the growth of layer (also called self-organization of clusters), or *ex-situ* subjective to post growth process (usually thermal treatment). Because non-equilibrium epitaxial growth technique such as MBE is governed by surface phenomena, the clustering process is limited by the kinetic processes of surface adsorption and desorption and surface mass transport. At the same time, the growth rate also matters. Typical clustering systems of this type contain self-organization of ordered quantum dots arrays [81], and self-organized nanodecomposition (0D, 1D or 2D) heterostructures [5]. The other type of clustering is subjective to post-growth process (usually thermal treatment), which is closer to thermodynamic equilibrium state. Clustering requires mass transport, which is usually a diffusion rate limited process. The energy barrier defines the fastest procedure (eq. 2.9).

From the previous discussions, it is clear that the energetic consideration of interfaces and elastic strain is of great importance to understand the complex dynamic process of clustering, which requires experimentally determine (or deducing) the local interface behavior and strain state during the cluster formation or transformation process.

Chapter 3.

Transmission electron microscopy

Transmission electron microscope (TEM) and scanning (S)TEM are instruments that can provide the highest spatial resolution and information about specimens in multiple dimensions (i.e., structural, compositional, chemical as well as electronic information) in both real and reciprocal space among all modern microscopic and spectroscopic methods [1]. TEM and STEM are the main techniques used in this thesis to investigate the microstructure of nano-clusters in semiconductor epilayers. In this chapter, the setup of the microscope, the fundamental theory of electron-crystal interaction and the image formation as well as contrast simulation will be summarized. Furthermore, the quantitative HRTEM image analysis method, i.e., geometric phase analysis (GPA), which is frequently applied in the thesis will be introduced and discussed in detail: the accuracy, resolution [2] and sources of artifacts.

3.1. The transmission electron microscope

Electron for microscopy After the classical works by Lord Rayleigh at the end of 19th century[82], which pointed out the theoretical diffraction resolution is inverse proportional to the source wavelength, the goal to observe smaller object headed for the pursuit of shorter wavelength sources instead of conventional visible light from the early 20th century. Photon of X-ray is found to be an ideal source that could yield in theory atomic resolution according to the Rayleigh criteria[83]. However, the difficulties in the lens systems to converge X-ray make it not likely to image in real space until recent breakthroughs[83]. Instead, information about specimen from X-ray diffraction are usually interpret in the reciprocal space. Microscopic particles that behave both like particle and wave[84] can be used to "image" objects of atom-scaled size, when they are accelerated to high speed, because the wave length is inverse proportional to their momentum, $\lambda = h/p = h/(mv)$, in which h is the Planck constant, p, m and v are the momentum, mass and velocity of the particle. Microscopic particles like neutron and ion beam, can be easier converged than X-ray. Unfortunately, they have significant dynamic effects (multiple scattering) and strong damage to the specimen, which also limit their resolution and applications.

Electron emerge, from the mid-20th century, as an ideal source to imaging in high spatial resolution as it could be easily refracted in electromagnetic field. For electrons travelling in the microscope which under an accelerate voltage of typically above 100 kV,

[1] It is noted that the abbreviation "M" in (S)TEM stands both for microscopy (the techniques) and microscope (the apparatus) depending on the context.

[2] We use the common definition of resolution and accuracy as it is so often mentioned in the thesis. Resolution of a technique refer to the smallest detectable distance of two spatial distinct object; accuracy is the degree to which the measured results conform to the true value.

the speed of electron is of $50-99\%$ of speed of light. So the wave length must be modified considering relativistic effects:

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2m_0eV(1+eV/2m_0c^2)}},$$

in which m_0 is the stationary mass of electron. Follow Lord Rayleigh's (diffraction limited) criteria, electron microscope with accelerate voltage of 100 kV should reach subpicometer resolution. However, along the course of the development of electron microscopes, it is not the source wavelength that restricted resolution, but the lens aberrations, mainly spherical aberration and chromatic aberration, as will be mentioned later.

By analogue to the optical microscope, the electron microscope is illuminated by the electron source. The microscopes used in this thesis are JEOL JEM-3010 and JEM-2100F. The JEM-2100F can be used both in the conventional TEM mode and STEM mode. In Fig. 3.1(a), the cross-section assembly of the JEM-3010/2100F microscope is shown. The most important working parameters (which will be explained later in section 3.2) of both microscopes are summarized in Tab. 3.1.

Table 3.1: Summary of the working parameters of the microscopes used in this thesis: accelerate voltage E, relativistic electron wavelength λ, interaction constant σ, spherical aberration C_s and chromatic aberration C_c.

	E (kV)	λ (pm)	σ ($V^{-1}nm^{-1}$)	C_s (mm)	C_c (mm)
JEM-2100F	200	2.51	0.007288	0.5	1.1
JEM-3010	300	1.97	0.006526	0.7	1.2

Conventional transmission electron microscope The term "conventional" before TEM is used here to distinguish with STEM. The setup of TEM and STEM is quite similar, as will be compared in the next block, but with different working principle. The main optical parts of an electron microscope is composed by the electron gun and a series of electromagnetic lenses, which are further divided into four major part: (1) the electron gun, (2) the condenser lens system, (3) the objective lens (where the specimen located) system and projection lens system and (4) the observation stage. The electron gun is the most important part of the microscope. The JEM-3010 equipped with a LaB_6 thermal emission gun, whereas a thermal field emission gun is installed in the JEM-2100F. Compared to the LaB_6 filament, the field emission gun can produce much brighter (higher electron density) and higher coherence electron beam. The pre-specimen lens (condenser lens system) is to focus the electron beam. It can change from convergent or parallel illumination mode. The object lens is the most important lens which focuses the diffracted beam after the specimen and determines the final image/diffraction pattern. By changing the excitation strength of the intermediate lens, it can switches from diffraction mode [Fig. 3.1(b)] and image mode [Fig. 3.1(c)]. The projection lens can shift the image/pattern or change the effective camera lens.

Except for these lenses, there are some other parts like apertures and controlling coils. Coils are essential to align and/or manipulate the electron microscope. The condenser aperture is used to limit the beam spread angle in order to select beam(s) of specific angle [3] to reduce lens aberration effect and/or to enhance the beam coherence. By insert-

[3] In the electron optics, convergence/divergence angle are actually referring to the semi-solid angle

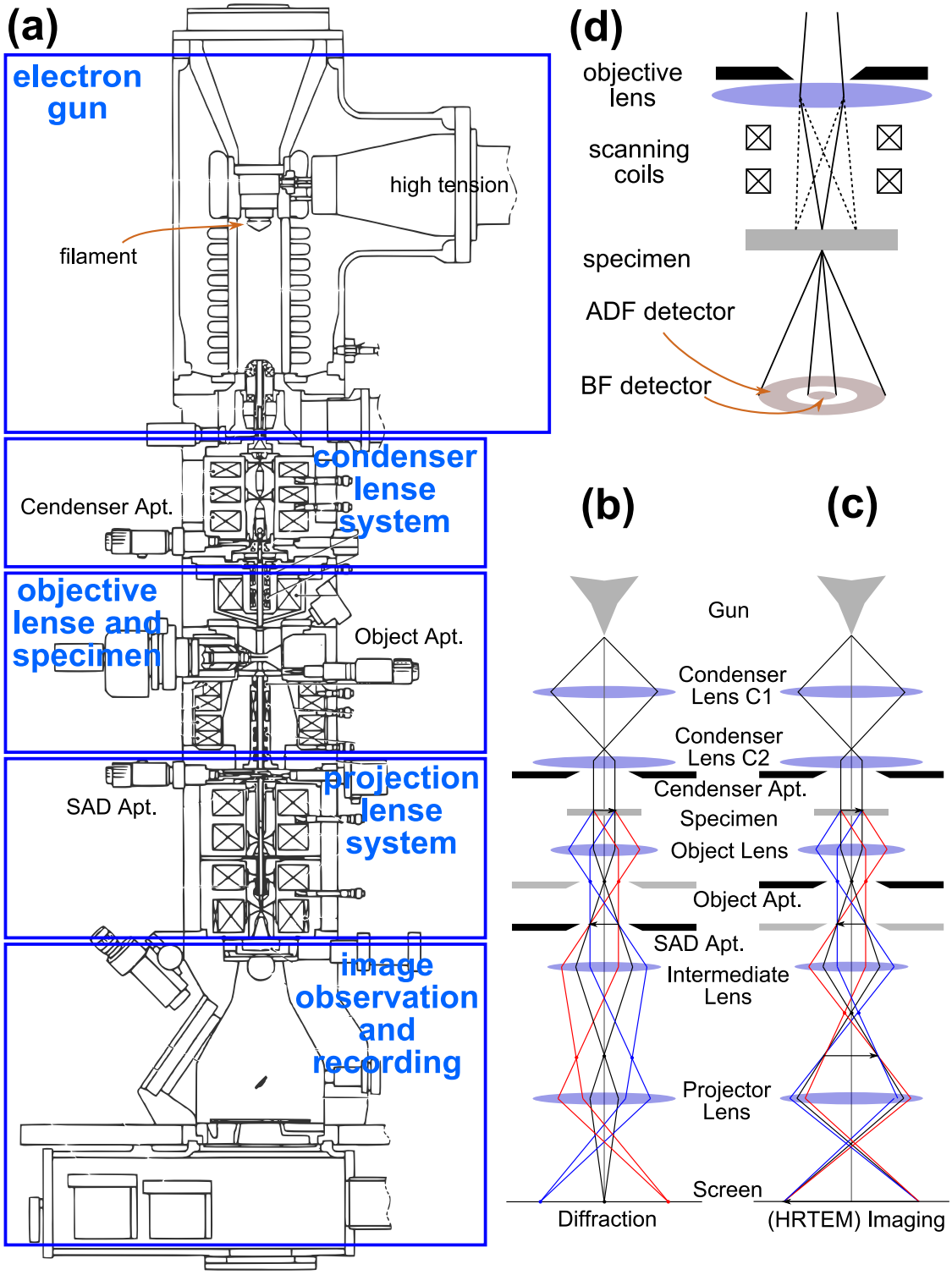


Figure 3.1: Scheme of the optics of (scanning) transmission electron microscope. Cross-section assembly of JEOL JEM-3010/2100F (a), illustrating the parts; change intermediate lens strength to switch from diffraction mode (b) and (HRTEM here) imaging mode (c); simple concept of STEM mode (d).

ing the objective aperture, specific beam can be selected for imaging. Selecting only the transmitted beam for imaging is called bright-field (BF) imaging, whereas selecting the diffracted beam for imaging is called dark-field (DF) imaging. Interference phase contrast image could be achieved by selecting multiple beams.

Scanning transmission electron microscope The most significant difference in a STEM is the electron beam (or, the probe) that is scanned over the specimen; and image is formed by collecting the signal in the time sequence. Electrons are accelerated from a source and focused into a point (the electron probe) on the specimen by a set of condenser lenses and the objective lens. In a simplified scheme [4], as shown in Fig. 3.1(d), the objective lens in STEM is *before* the specimen. An objective aperture limits the maximum angle of illumination included in the incident probe. The probe scans across the sample by a set of scan coils. The output is usually collected by detectors including BF and annular dark field (ADF) detectors. The BF detector intercepts the transmitted beam, while an ADF detector surrounds the transmitted beam and collects the scattered electrons [cf. Fig. 3.1(d)]. The inner angle of this detector can be changed with projector lenses to a variety angle range.

3.2. Electron diffraction theory

The elementary electron-crystal interaction give rise to the way how we can observe (actually interpret) objects, down to atomic scale, in crystal specimens. While the approximations for numerical solutions to the theory and the aberrations in the lens system define how "clear" we can observe, or how we can quantify, the objects in specimen.

3.2.1. Electron-crystal interaction

Electrons traveling through a specimen will interact with atoms in it by Coulomb force, or other energy excitation process. Thereafter the electrons exit at the bottom of specimen carry information of the chemical and structural information. Tow types of interaction is categorized, the elastic and inelastic scattering process. A detailed theoretical description can be found for example in Ref. 85,86.

Coulomb interaction between electron and atom yield only momentum transfer. It is elastic interaction and carries only structural information about specimen. The elastic scattered electrons contributes mostly to conventional TEM contrast including bright-field (BF), dark-field (DF) imaging and HRTEM phase contrast lattice imaging. BF/DF image contrast are sensitive to deviation of atom position from perfect crystal structure, therefore, it is a powerful technique to analyse defect structure in crystal samples. Phase contrast image emerge from the interference of scattered electron beams in different direction after interact with specimen with various phase grating. It could provide high-resolution atomic structure information about the specimen. The part of phonon scattered electrons, which are scattered in relatively high angles, is refereed as thermal diffuse scattered (TDS) electrons. The TDS scattering cross-section is proportional to the atomic Z number, therefore it is the main contribution to the HAADF-STEM [87].

During the inelastic scattering process, the kinetic energy conservation of the incident electrons is no longer satisfied. The electrons transfer energy to the specimen by mainly

[4] In JEM-2100F the TEM/STEM mode is switched through a objective minilens, which enable a fast switch without residual magnetic resonance

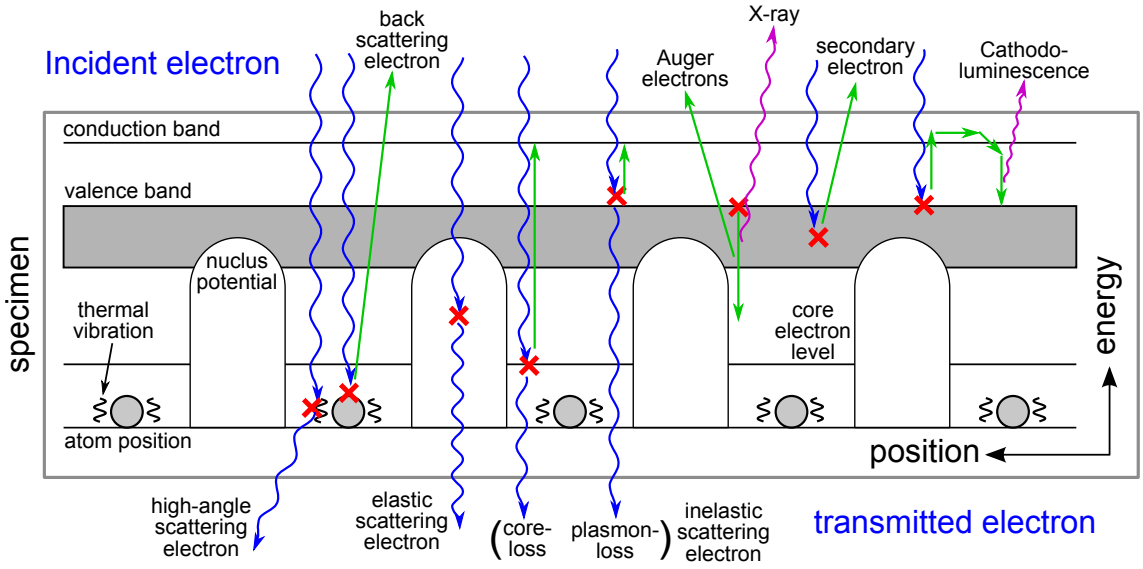


Figure 3.2: Illustration of the electron-crystal interaction. Various of signals could be used to detect the structural and chemical information about the sample. The red cross indicate the position where the interaction (energy/momentum exchange) take place.

three parts of energy contribution: phonon excitation, valance (energy band) excitation and inner shell ionization. The phonon excitation energy transfer is much smaller than 0.1 eV and thus usually can be neglected and treated the same as elastically scattered electrons in TDS[85], as already mentioned above. Electron valence excitation involve a energy transfer up to about 50 eV whereas the inner shell ionization process have a energy transfer up to 1000 eV level dependent uniquely on the atom species. When the excited electrons return to its stable state, generation of Auger electrons or X-ray is accompanied. Therefore, these signals are contributed to the analytical electron spectroscopy technique including electron energy loss (EELS), X-ray energy dispersive spectroscopy (EDS) and cathodluminescence (CL). As the electron beam in the microscope can be focus to diameter of nanometer scale, these analytical techniques can be achieved with a high spatial resolution. Various electron matter interaction and its usefulness in characterization of materials are schematically depicted in Fig. 3.2.

The main imaging techniques and simulations in this thesis utilize the elastic scattered electrons and relies on the fundamental theory of electron diffraction. In the following, the related topics of both kinematic and dynamic electron diffraction approach will be summarized.

3.2.2. Kinematic approach and electron diffraction geometry

In transmission electron microscope, the kinetic energy E_k of electron beam is given by the accelerate voltage, typically 200–300 kV, which is several orders of magnitude higher than the energy of electronic potentials in atoms. The full Hamiltonian (including electron-electron, core-core and core-electron interaction) of the specimen can therefore approximated by a potential distribution of $\phi(\mathbf{r})$ (energy of few eV). Then, the electron beam wave function is given by the stationary Schrödinger equation

$$\nabla^2\Psi + \frac{8\pi^2 m}{h^2}[E_k + \phi(\mathbf{r})]\Psi = 0 \tag{3.1}$$

The kinematic approximation states that the electron is scattered by an atom only once and only the elastic scattering is considered. This can be achieved mathematically by the first-order perturbation of the electron wave, $\Psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \psi'(\mathbf{r})$, where $\psi_0(\mathbf{r})$ is the transmitted wave and $\psi'(\mathbf{r})$ is the elastic scattered wave. Insert it into equation (3.1) by applying the Green's theorem, the high energy approximation and some rearrangements, this yield the well-known result (e.g. refer Ref. 86,88):

$$\psi'(\mathbf{r}) = A\frac{\exp(-2\pi ikr)}{r} \cdot \frac{2\pi me}{h^2} \int_\Omega \phi(\mathbf{r}_1)\exp[2\pi i(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{r}_1]\mathrm{d}\Omega. \tag{3.2}$$

It is easily recognized that the integration part is a Fourier transform of the potential distribution function. This is the fundamental equation for the kinematic theory of electron diffraction. The pre-integration term $\exp(-2\pi ikr)/r$ in the equation indicates that the scattered wave is a spherical wave. $\mathbf{k}$ is the scattering wave vector. The integration is the summation of the scattering wave vector, the strength of scattering depends on the electric potential. The spherical scattered wave is linearly related to the incident wave.

Taking $\mathbf{q} = \mathbf{k} - \mathbf{k}_0$, the integration part in equation 3.2 for potential distribution of an atom $\phi_a(\mathbf{r})$ defines the atomic scattering form factor, or the *atomic form factor*:

$$f_e(\mathbf{q}) = \frac{2\pi me}{h^2} \int_\Omega \phi_a(\mathbf{r})\exp(2\pi i\mathbf{q}\cdot\mathbf{r})\mathrm{d}\Omega \equiv \mathcal{F}[\phi_a(\mathbf{r})]. \tag{3.3}$$

In crystal, unit cell is superposition of n atoms arranged in space $\phi_{cell}(\mathbf{r}) = \sum\phi_{a,i}(\mathbf{r} - \mathbf{r}_i)$. The following equation defines the *structure factor* that characterizes the strength of scattering of a crystal unit cell

$$F(\mathbf{q}) = \sum_{i=1}^{n} f_{e,i}(\mathbf{q})\exp(2\pi i\mathbf{q}\cdot\mathbf{r}_i). \tag{3.4}$$

Clearly, the scattered wave is peaked at $\mathbf{q}\cdot\mathbf{r}_i =$ integer positions. It is easier to describe the scattering process in reciprocal space, spanned by the reciprocal basis vectors ($\mathbf{a}_1^*$, $\mathbf{a}_2^*$, $\mathbf{a}_3^*$), which fulfill $\mathbf{a}_i \cdot \mathbf{a}_j^* = 2\pi\delta_{ij}$. The points on this reciprocal lattice are specified by vectors of the type $\mathbf{q}_{hkl} = h\mathbf{a}_1^* + k\mathbf{a}_2^* + l\mathbf{a}_3^*$, where h, k, l are integers and are usually referred to as Miller Indices. Accordingly, lattice planes are denoted by (h,k,l), directions in the unit cell by $[h,k,l]$.

From equation 3.4, it can be seen that for some specific combination of $\mathbf{q}_{hkl}\cdot\mathbf{r}_i$, the structure factor can be zero. Such diffraction is referred to as kinematic forbidden diffraction. For example, zincblende structure is face-centered cubic with two different basis atom: A$(0,0,0)$, and B$(\frac{1}{4},\frac{1}{4},\frac{1}{4})$, and the structure factor is

$$F_{ZB}(hkl) = \{f_A + f_B\exp\left[2\pi i(h+k+l)/4\right]\}\{1 + e^{\pi i(h+k)} + e^{\pi i(k+l)} + e^{\pi i(l+h)}\}.$$

From the above equation, the following can be obtained:

$$\begin{aligned} &(1) \quad F_{ZB}(hkl) = 0, && \text{when } h,k,l \text{ are mixed even and odd} \\ &(2) \quad F_{ZB}(hkl) = 4(f_A - f_B) && \text{when } h+k+l = 4m+2, \quad m \in N. \end{aligned}$$

Therefore, mixed h,k,l indices [e.g. InAs(100), (110)] results in kinematic forbidden reflections. While the intensity satisfying $h+k+l = 4m+2$ [e.g. InAs(002)] are dependent on the difference of the scattering factor of the two constituent atoms. Hence it is also

referred as chemical sensitive reflections, which is very useful for the qualitative and quantitative analysis of heterostructures based on zincblende crystals[89,90].

We can construct such a sphere centered at the intersection of incident beam and specimen and with a radius of $1/\lambda$. This sphere is called the reflection sphere or, Ewald sphere. The geometric relation of $\mathbf{k}_0, \mathbf{k}$ and $\mathbf{q}_{hkl}$ are schmatically depicted in Fig. 3.3. The wavelength of electron is very short in a high-energy electron microscope, which results in a large Ewald sphere radius. So the surface of the sphere appears quite "flat". From the geometry of the triangle, it is easily inferred that

$$\frac{\sin\theta}{n\lambda} = \frac{1}{2}|\mathbf{q}_{hkl}|, \quad \text{or,} \quad 2d_{hkl}\sin\theta = n\lambda, \quad n \in N. \tag{3.5}$$

This is called the Bragg's law, in which θ is the Bragg angle. It is also written as θ_{hkl} or θ_B. The vector form in reciprocal space is easily verified in Fig. 3.3: $\mathbf{q}_{hkl} = \mathbf{k} - \mathbf{k}_0$. As convention, $\mathbf{q}_{hkl}$ is often refereed as the reciprocal lattice vector, and writes as $\mathbf{g}_{hkl}$, or simply as $\mathbf{g}$. The reciprocal lattice planes, which intersect the Ewald sphere are referred as the Laue zone. According to their scattering angle, the planes are termed zero-order Laue zone (ZOLZ), first-order Laue zone (FOLZ), second-order Laue zone (SOLZ) and the rest are higher-order Laue zone, as depicted in Fig. 3.3.

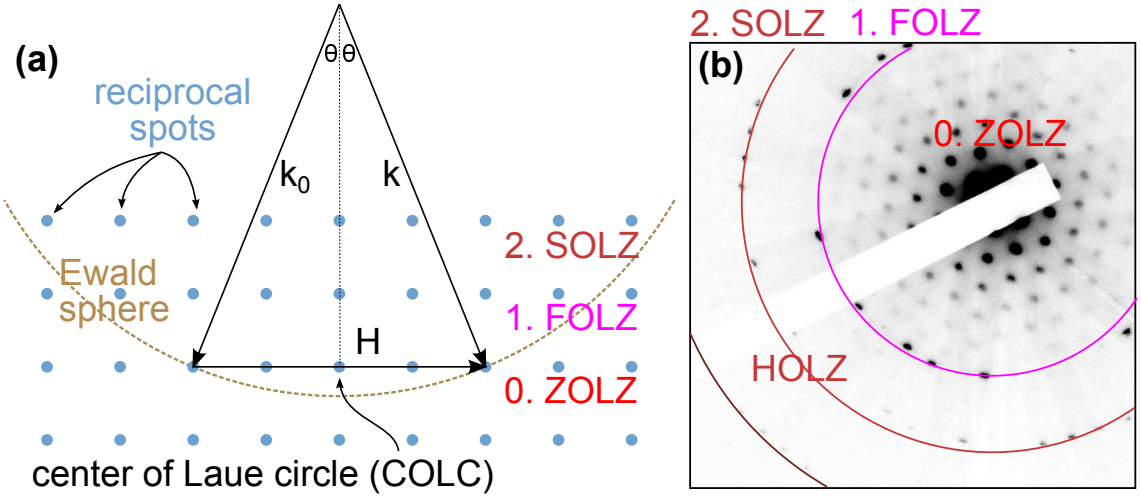


Figure 3.3: (a) Illustration of the Ewald sphere construction in reciprocal space. (b) example of the diffraction pattern taken from GaN <0001> zone axis showing the Laue zones.

The Laue construction is very convenient. Tilt of the sample is equal to tilt the reciprical spots with respect to the electron beam. In the Ewald construction, it can be viewed as moving the center of Laue circle (COLC).

Real crystals have limited size and shape, thus the Fourier transformed lattice (reciprocal lattice) does not contain infinitesimal spots, but spots with specific shape. A thin, electron transparent crystal can be regarded as a sheet in most TEM specimen. The reciprocal lattice spots of such a sheet are actually elongated rods with the longer axis parallel to the sheet normal, as illustrated in Fig. 3.4.

When the Bragg condition is not exactly satisfied, there will still be some intensity at the intersect position of the Ewald sphere. The distance from the center of g spot to the intersection of diffraction vector with the Ewald sphere is defined as the excitation errors,

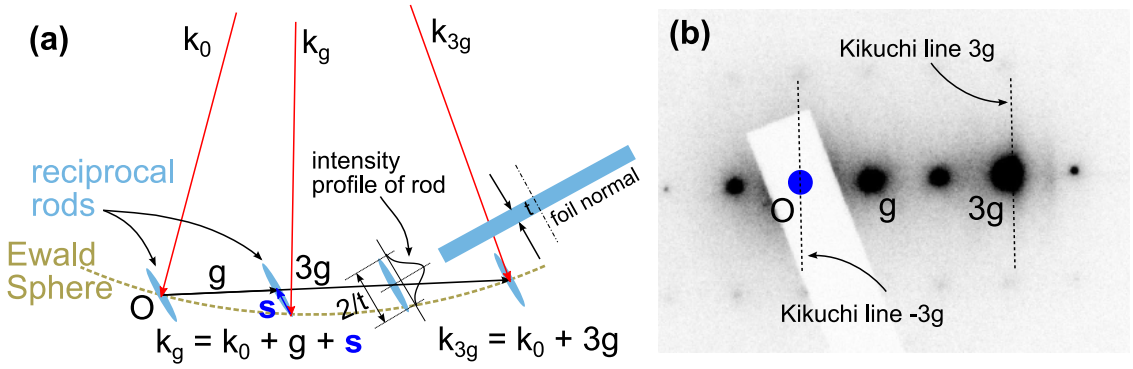


Figure 3.4: (a) Relation between reciprocal rods, thin crystal foil and the diffraction quantities. (b) An example of experimental control of s. The Kikuchi lines are used as guides to accurate control the parameters.

or deviation parameter, expressed in scalar quantity,

$$s_g = \frac{-\mathbf{g}\cdot(2\mathbf{k}+\mathbf{g})}{2|\mathbf{k}_0+\mathbf{g}|\cos\theta}. \tag{3.6}$$

s_g is a experimental parameter, which can be controlled in the experiment.

3.2.3. Dynamical approach

The kinematic approach only works when the specimen is vanishingly thin and/or far away from exact Bragg condition ($s_g \gg 0$, the condition also know as "dynamical extinction"). In practise, the multi-scattering can not be neglected therefore the dynamic approach is always taken into account in electron diffraction.

Here, several theoretical approaches are summarized. Darwin-Howie-Whelan formalism[88,91] gives the basis for understanding the contrast of crystal defects. Direct space description[88] setup the basis for multi-slice method simulation. Bloch wave method[85] can account for scattering beyond the limitation of Bragg diffraction like diffuse scattering and inelastic scattering.

Darwin-Howie-Whelan formalism

The Schrödinger equation 3.1 can be rearranged to:

$$\nabla^2\Psi + 4\pi^2 k_0^2\Psi = -4\pi^2 U(\mathbf{r})\Psi \quad \text{with} \quad U(\mathbf{r}) \equiv \frac{2meV(\mathbf{r})}{h^2} = \frac{\sigma V(\mathbf{r})}{\pi\lambda}, \tag{3.7}$$

$k_0 = 1/\lambda$ is the effective relativistic wave number that includes the acceleration by an atom nucleus, σ is called the interaction constant measuring the interaction strength of electron and specimen. $V(\mathbf{r})$ is the real space crystal potential function, which is the Fourier series of structure factor

$$V(\mathbf{r}) = \sum_{\mathbf{g}} F(\mathbf{g})\exp(2\pi i\mathbf{g}\cdot r).$$

Absorption can be implemented phenomenologically by replacing $V(\mathbf{r})$ by a complex potential $V_c(\mathbf{r}) = V(\mathbf{r}) + \mathrm{i}V'(\mathbf{r})$ and described by a Fourier series:

$$V_c(\mathbf{r}) = V(\mathbf{r}) + \mathrm{i}V'(\mathbf{r}) = F(0) + \sum_{g\neq 0} F(\mathbf{g})e^{2\pi\mathrm{i}\mathbf{g}\cdot\mathbf{r}} + \mathrm{i}\sum_{g} F'(\mathbf{g})e^{2\pi\mathrm{i}\mathbf{g}\cdot\mathbf{r}}.$$

The zero-frequency component $F(0)$ is separated from the others, which defines the mean inner potential. It is a positive number which describing the acceleration of the electron beam, known as refraction. Therefore the wave number should include the refracted acceleration voltage, K_0. Assign

$$U = \frac{2me}{h^2}\sum_{\mathbf{g}\neq 0}\phi_{\mathbf{g}}e^{2\pi\mathrm{i}\mathbf{g}\cdot\mathbf{r}} \quad \text{and} \quad U' = \frac{2me}{h^2}\sum_{\mathbf{g}}\phi'_{\mathbf{g}}e^{2\pi\mathrm{i}\mathbf{g}\cdot\mathbf{r}},$$

equation 3.7 becomes

$$\nabla^2\Psi + 4\pi^2K_0^2\Psi = -4\pi^2[U + \mathrm{i}U']\Psi. \tag{3.8}$$

The total wave function at the exit plane must fulfill the Bragg condition,

$$\Psi(\mathbf{r}) = \sum_{\mathbf{g}}\psi_{\mathbf{g}}e^{2\pi\mathrm{i}(\mathbf{k}_0+\mathbf{g})\cdot\mathbf{r}}. \tag{3.9}$$

Inserting eq. 3.9 in eq. 3.8 result in [88]

$$\frac{\mathrm{d}\psi_g}{\mathrm{d}z} - 2\pi\mathrm{i}s_g\psi_g = \mathrm{i}\pi\sum_{g'}\left[\frac{\exp(\mathrm{i}\theta_{g-g'})}{\xi_g} + \mathrm{i}\frac{\exp(\mathrm{i}\theta_{g-g'})}{\xi'_g}\right]\psi'_g. \tag{3.10}$$

s_g is the deviation parameter defined earlier; ξ_g and ξ'_g are the distinction length and absorption length, respectively, defined as

$$\frac{1}{\xi_g} \equiv \frac{\mid U_g \mid}{\mid \mathbf{k}_0 + \mathbf{g} \mid \cos\theta_g}; \quad \text{and} \quad \frac{1}{\xi'_g} \equiv \frac{\mid U'_g \mid}{\mid \mathbf{k}_0 + \mathbf{g} \mid \cos\theta_g}. \tag{3.11}$$

The exponent numerator, $\exp(\mathrm{i}\theta_{g-g'})$, is from the exponent part of Fourier expansion of the potential term $U + \mathrm{i}U'$.

Howie and Whelan derived [91] in 1961 similar equations (the eq. 3.13) to eq. 3.10 for the two-beam case. The formalism is actually similar to the dynamic theory of X-ray derived in the work of Darwin in 1914 [92,93]. The multi-beam equations 3.10 is therefore referred as the Darwin-Howie-Whelan equations.

The equations are valid for arbitrary number of diffracted beams, and describe how the amplitude of a diffracted beam changes with depth in the crystal. These changes depend on how well the Bragg condition is satisfied, and on the strength of the interaction with other diffracted beams. It is clear that the diffracted beams are dynamically coupled and absorption mediates the interaction strength between diffracted beams. Figure 3.5 shows an example of calculated beam intensities of GaN as a function of sample thickness.

We can tilt the specimen in practice to excite two strong beams, a transmitted and a diffracted beam, which is called "two-beam" condition. Under this condition, the absorption can also be neglected for the purpose of interpreting the diffraction contrast feature from defects in crystals, since it only reduce the overall intensity. The two-beam

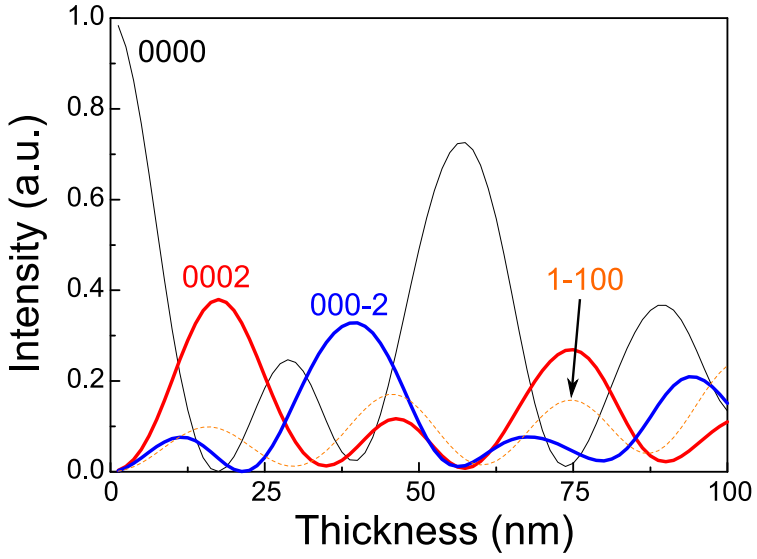


Figure 3.5: Calculated electron beam intensity of GaN under [1$\bar{2}$10] zone axis condition as function of the sample thickness. Beam intensities are dynamically coupled.

Table 3.2: Extinction distances for selected semiconductor materials under two-beam condition (unit in nm), operated under 300 kV (for 200 kV, multiply a factor of 0.785).

hkl or (*hkil*)	0002	0004	$\bar{1}$100	11$\bar{2}$0	002	004	111	220
GaN	98	–	165	88	–	–	–	–
6H-SiC (3 × *hkl*)	72	58	–	–	–	–	–	–
GaAs	–	–	–	–	> 250	111	73	51
GaP	–	–	–	–	> 250	128	72	69
Si	–	–	–	–	–	196	86	101

Howie-Whelan equations are:

$$\frac{d\psi_g}{dz} = \frac{\pi i}{\xi_0}\psi_g + \frac{\pi i}{\xi_g}\psi_0 \exp(-2\pi i s z); \tag{3.12}$$

$$\frac{d\psi_0}{dz} = \frac{\pi i}{\xi_0}\psi_0 + \frac{\pi i}{\xi_g}\psi_g \exp(+2\pi i s z). \tag{3.13}$$

s is the vector deviation parameter projected along z direction. The extinction distance ξ_g can also be expressed as

$$\xi_{hkl} = \frac{\pi V_{cell}\cos\theta_{hkl}}{\lambda F(hkl)}, \tag{3.14}$$

in which V_{cell} is the volume of the unit cell, λ is the electron wave length. It can be derived from these equations that the extinction distance ξ_g in two-beam condition is physically equal to the periodicity of the modulating intensity of diffracted beam, which is the origin of thickness fringes in images.

It is illustrated in Fig. 3.6 showing the calculated electron beam intensity as function of the sample thickness of GaN under two-beam consition. The calculation is obtained with JEMS program[94]. Extinction distance (under two-beam condition) values for typical semiconductor materials are summarized in Tab. 3.2.

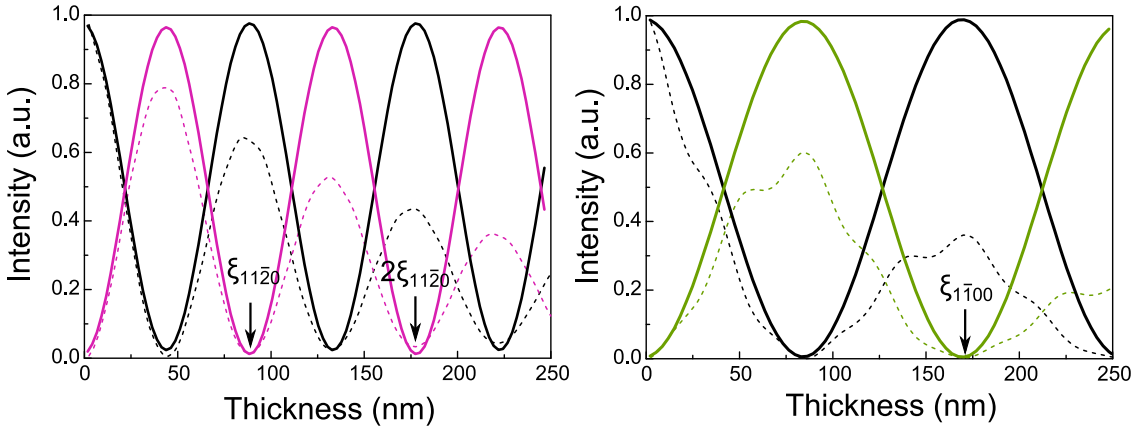


Figure 3.6: Calculated electron beam intensity as function of the sample thickness of GaN under two-beam consition. Left and right panel are for $0000/11\bar{2}0$ and $0000/1\bar{1}00$ two-beam conditions, respectively. Full/dashed lines are without/with absorption.

Direct space description and multi-slice method

The incident electron is represented by a plane wave with free space wave vector $\mathbf{k}_0$. After passing through the crystal the wave function is a modulated plane wave:

$$\Psi(\mathbf{r}) = \psi(\mathbf{r})e^{2\pi i\mathbf{k}_0\cdot\mathbf{r}}. \tag{3.15}$$

Insert it into equation 3.8, and separate the z component from $\mathbf{k}_0$ since the z component is much larger than x and y component. Using the *high energy approximation*[95] which neglects the second order derivative along the beam direction, the result is

$$\frac{\partial\psi}{\partial z} = \left[\frac{\mathrm{i}}{4\pi k_{0z}}(\nabla_{xy}^2 + i4\pi k_{xy}\cdot\nabla_{xy}) + \frac{\mathrm{i}\pi}{k_{0z}}(U + \mathrm{i}U')\right]\psi \quad \text{or,} \quad \frac{\partial\psi}{\partial z} = (\overline{\Delta} + \overline{V})\psi. \tag{3.16}$$

For a very thin crystal slice, the lateral scattering vanishes and only $\partial\psi/\partial z = \overline{V}\psi$ is left over. The solution is

$$\psi(x,y,z_0) = e^{\int_0^{z_0}\overline{V}(x,y,z)\mathrm{d}z} = e^{\int_0^{z_0}\mathrm{i}\sigma V_c(x,y,z)\mathrm{d}z} = e^{\mathrm{i}\sigma V_p}. \tag{3.17}$$

V_p is the potential projected along the electron penetrating direction, referred as projected potential. This equation clearly shows that the crystal induces a phase shift to the electron wave in a thin slice. If the phase shift is small, the exponent can be approximated by its Taylor series:

$$e^{\mathrm{i}\sigma V_p} = 1 + \mathrm{i}\sigma V_p + \cdots \tag{3.18}$$

Neglecting the higher order terms, and keeping only the terms linear to V_p leads to the weak-phase object approximation (WPOA). It is very important in interpreting the high-resolution TEM images. For most elements consisting semiconductor materials, the WPOA is a very good approximation.

When there is no crystal present, $\overline{V}$ equals zero and equation 3.16 reduced to $\partial\psi/\partial z = \overline{\Delta}\psi$. It takes the form of diffusion equation which describe how the electron propagate along the z direction. Solve the equation for small depth of propagation ϵ. It has the

character[88]:

$$e^{\overline{\Delta}\epsilon}\psi = \mathcal{F}^{-1}\left[\exp(\mathrm{i}\pi k\frac{R}{\epsilon})\right] \otimes \psi(\mathbf{R}) = \mathcal{P}_\epsilon(R) \otimes \psi(\mathbf{R}). \tag{3.19}$$

$\mathcal{P}_\epsilon(R)$ is the Fresnel propagator. It means the propagating of electron beam from one slice to the next is mathematically equivalent to a convolution with the Fresnel propagator.

Therefore, if we divide the crystal into n layers with each slice having vanishing thickness ϵ, the equation 3.16 can be expressed as

$$\psi(x,y,z_0) = e^{\overline{\Delta}\epsilon}e^{\mathrm{i}\sigma V_p^n}\cdots e^{\overline{\Delta}\epsilon}e^{\mathrm{i}\sigma V_p^2}e^{\overline{\Delta}\epsilon}e^{\mathrm{i}\sigma V_p^1}\psi(x,y,0). \tag{3.20}$$

This is the basis of multi-slice method for numerical simulation of dynamical image contrast as schematically illustrated in Fig. 3.7.

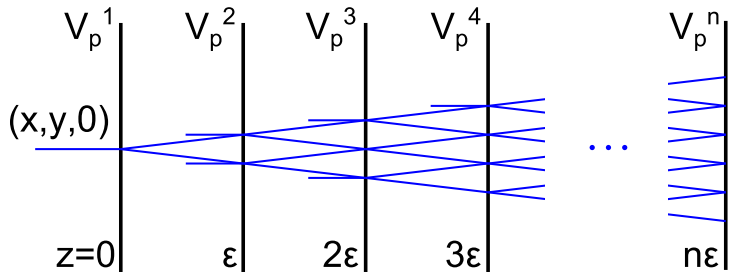


Figure 3.7: Illustration of the multi-slice method and the physical meaning.

Bloch wave method

The previous approach assumes that electrons will travel only in the directions $\mathbf{k}_0$ and $\mathbf{k}_0 + \mathbf{g}$ which fulfill the Bragg rule. While it is valid for external observer, it is not necessarily true inside the crystal. The situation of high-energy electron traveling through matter is analogue to the phonon dispersion: before entering the crystal, the electron can travel with any wave vector (while energy is restricted by the acceleration voltage, the direction is arbitrary); inside the crystal, the direction of electron is dependent on the crystal structure and on the electron energy. Such relation is called the dispersion relation. The electron dispersion relation to the wave vector is described by the Bloch wave method. In analogue to the previous section, the crystal potential is expanded as Fourier series

$$V(\mathbf{r}) = \sum_h V_h \exp(2\pi\mathrm{i}\mathbf{h}\cdot\mathbf{r}). \tag{3.21}$$

First introduced by Bloch into quantum mechanics, the electron wave inside the crystal is a Bloch wave[96], which is the product of a wave with arbitrary wave vector k, and the function $C(\mathbf{r})$ with periodicity of the lattice

$$\Psi(\mathbf{r}) = C(\mathbf{r})\exp(2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{r}) = \sum_g C_g \exp[2\pi\mathrm{i}(\mathbf{k}+\mathbf{g})\cdot\mathbf{r}]. \tag{3.22}$$

The coefficients C_g are called Bloch wave coefficients. In perfect crystal, the C_g are independent of position, while dependent on the wave vector $\mathbf{k}$. Inserting in equation 3.7

leads to an equation that must valid for any position $\mathbf{r}$, which we get[88]:

$$[k_0^2 - (\mathbf{k} + \mathbf{g})^2]C_g + \sum_{h \neq g} U_{g-h} C_h = 0. \tag{3.23}$$

The general wave solution inside the crystal consists of a superposition of $2N$ Bloch waves, each with its own amplitude $\alpha^{(j)}$:

$$\Psi(\mathbf{r}) = \sum_j \alpha^{(j)} \sum_g C_g^{(j)} \exp[2\pi \mathrm{i}(\mathbf{k}^{(j)} + \mathbf{g}) \cdot \mathbf{r}] = \sum_j \alpha^{(j)} C^{(j)}(\mathbf{r}) \exp[2\pi \mathrm{i} \mathbf{k}^{(j)} \cdot \mathbf{r}]. \tag{3.24}$$

The coefficient $\alpha^{(j)}$ is the excitation amplitude of the jth Bloch wave. There are infinite branches of Bloch wave in the k-space.

Apparently, the wave function (excitation amplitude) will take its maximum when the wave vectors $\mathbf{k}$ satisfy the Bragg condition. Other scattering like diffuse scattering and inelastic scattering can be predicted by properly treating the crystal potential energy that is in between the space of atomic nucleus. However, the calculation complexity is much higher than the previous method, especially when accounting for crystals with defects that break the periodicity of crystal lattice.

3.3. Image formation and contrast simulation

In conventional TEM, there are mass-thickness contrast, diffraction contrast and phase contrast. The mass-thickness contrast is governed by the different scattering ability of different atoms and their integration from different thickness. Diffraction contrast is powerful tool and have already well developed both in theory and practice for imaging crystal defects[97]. Phase contrast is from the interference of diffracted beams with different electron phase relationship.

3.3.1. Diffraction image formation and contrast simulation

Image formation

We describe here the diffraction contrast only for the case under two-beam condition, which is important for qualitative and quantitative defect analysis.

There are two types of diffraction contrast. The first is from the phase difference of the coupled diffracted beams. The second is the deviation of atom displacement in crystal specimen from its expected position in perfect crystal. While the first type is the reason for thickness fringes, the second type is more interesting for defect imaging and analysis, since the defect will produce locally displacement. Then the new position $\mathbf{r}'$ of atoms in the specimen is the expected position in perfect form $\mathbf{r}$ plus a displacement $\mathbf{u}(\mathbf{r})$, i.e., $\mathbf{r}' \rightarrow \mathbf{r} + \mathbf{u}(\mathbf{r})$. Inserting it back to equation 3.10, it will resulted in the two-beam Howie-Whelan equations that take the new form:

$$\frac{d\psi_g}{dz} = \frac{\pi i}{\xi_0}\psi_g + \frac{\pi i}{\xi_g}\psi_0 \exp\{-2\pi \mathrm{i}[sz + \mathbf{g}\cdot\mathbf{u}(\mathbf{r})]\} \tag{3.25}$$

$$\frac{d\psi_0}{dz} = \frac{\pi i}{\xi_0}\psi_0 + \frac{\pi i}{\xi_g}\psi_g \exp\{+2\pi \mathrm{i}[sz + \mathbf{g}\cdot\mathbf{u}(\mathbf{r})]\}. \tag{3.26}$$

The equations can be interpreted that the displacement of atoms from its expected perfect crystal position will introduce an additional phase shift to the coupled transmitted and diffracted electron wave functions. These equations are the basis of diffraction contrast formation from defect.

After propagation through the specimen, the lens system focus and project the image on the screen. In this case, the imperfection in the lens system that limit resolution is much less pronounced, because of the very small objective lens aperture. The final image is simply

$$I = \psi_g \psi_g^*, \tag{3.27}$$

where ψ_g^* is the complex conjugate of wave function ψ_g.

Diffraction contrast simulation

As according to the image contrast formation mechanism, the diffraction contrast simulation consists generally of two steps:

- deriving the analytical (or numerical) function of displacement field $u(\mathbf{r})$ for the defect structure under consideration
- solving the multi-beam equations 3.10, or two-beam equations 3.13 which depending on the experimental setup by implementing the displacement field function

$u(\mathbf{r})$

The solved exit wave function is a function of t (the specimen thickness), ξ_g (the material and instrument related constant extinction distance), s (constant deviation parameter that is controlled by experimental condition) and $u(x, y)$ (the projected displacement field function). So, the final image intensity is

$$I(x, y) = I(t, \xi_g, s, u(x, y)).$$

The deviation parameter s can be calculated by equation 3.6.

Because the displacement field function is material dependent, which is usually elastically anisotropic and the fields are continuous, the analytic or numerical calculation is very complex, and no universal general form is available. In real case, approximations are usually made to simplify the displacement field. Early attempt was accomplished by Ashby[98–100] where the analytical solution was solved for spherical inclusions in elastically isotropy media. Later, a numerical solution is also derived for elliptical inclusions. The quantitative interpretation is also achieved for more complex bodies like quantum dots on surface, by introducing the abrupt displacement approximation (ADA)[101,102].

Validity of the two-beam equations to many-beam cases The solution of the two-beam equations describe a one to one "peak to valley" dynamic coupling relation between the zeroth (transmitted) beam and selected diffraction beam. In this case, $I_0 + I_g \approx 1$ (or $I_0 + I_g \approx \exp(-zk^{\text{abs}})$ including an absorption factor k^{abs}). An exact two-beam condition cannot be always obatinable. In more general case, nevertheless, the equations can be still valid, if the one to one "peak to valley" dynamic coupling relation between the zeroth beam and selected diffraction beam exists. An example of such a relation for GaN {0002} systematic row condition and (0002) two-beam condition is shown in Fig. 3.8.

As can be clearly seen, the intensity of the zeroth and 0002 beam under two-beam condition is well coupled peak to valley; and $I_{0000} + I_{0002} \approx \exp(-zk^{\text{abs}})$. In the system row condition, $I_{0000} + I_{0002} + I_{000\bar{2}} \approx \exp(-zk^{\text{abs}})$. The peak to valley dynamic coupling relation between the zeroth beam and selected diffracted beam (0000/0002, or 0000/000$\bar{2}$) roughly holds for a thickness smaller than about 100 nm, especially when the sample thickness is less than 50 nm.

However, in zone axis condition, as can be seen in Fig. 3.5 (compare the dark, red and blue lines), the peak to valley dynamic coupling does not hold for the 0000/0002 or 0000/000$\bar{2}$ beams.

Therefore, the simulation of GaN 0002/000$\bar{2}$ dark-field contrast for *system row* condition (experimental condition) using the *two-beam* Howie-Whelan equation is still valid. Under this condition, the extinction distance should be replaced by an "apparent distinction distance". In the above illustrated case, $\xi_{0002} = 98$ nm, $\xi_{0002}^{app} = 50$ nm. Under [1$\bar{2}$10] zone axis condition, the interpretation of dark field contrast using the two beam equations is not valid. Specific problem of interpreting the dark field contrast from tiny strained GdN clusters using this method is discussed in detail in Chapter 6.

3.3.2. High-resolution transmission electron microscopy

Image formation

High-resolution TEM (HRTEM) is a very powerful tool to image the specimen down to atomic resolution. The electron beam passing through the specimen contains full (am-

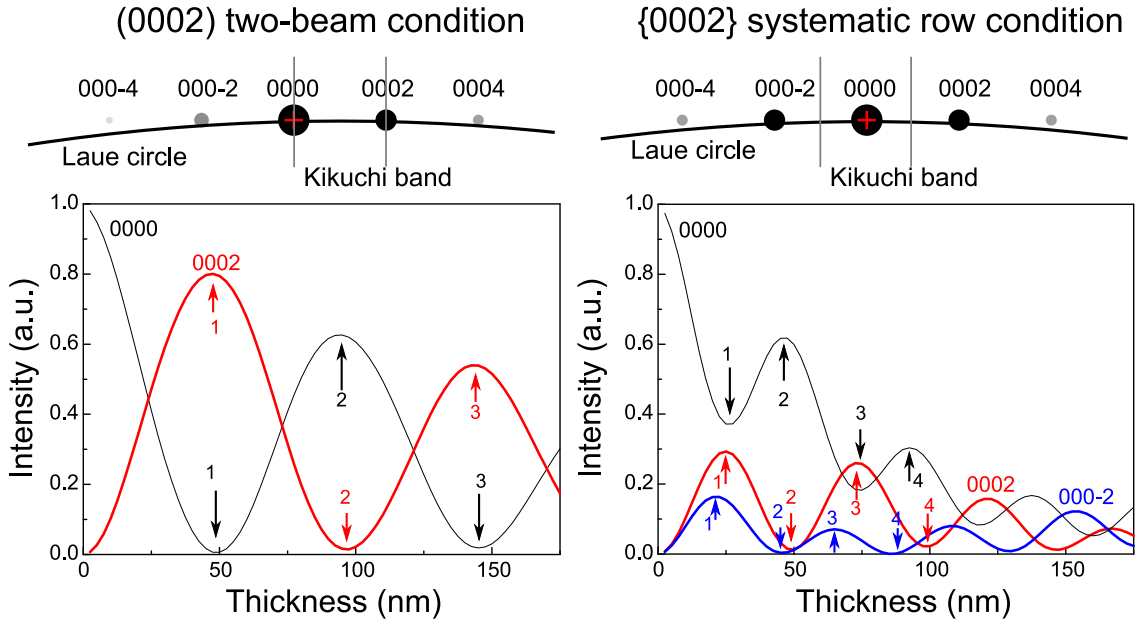


Figure 3.8: Compare the diffracted beam intensity for two-beam condition (left) and system row condition (right). The dynamic "coupling" relation between the zeroth beam and 0002/000$\bar{2}$ beam also roughly establishes when sample thickness smaller than 100 nm.

plitude and phase) information about the specimen. The exit beams will be further affected by the optic-magnetic lens system and interfere with each other. The final image is produced at the image plane. In conventional diffraction contrast imaging the electron phase information is lost. Therefore, HRTEM image is also called phase contrast image. The lens system (especially the objective lens) plays an essential role in the imaging formation since the imperfection of the lens will also produce phase shift to the electron waves.

The HRTEM image formation follows first the propagation through the specimen, then propagation through the lens system. This is also the way how numerical simulation of HRTEM image contrast is proceeded, as schematically illustrated in Fig. 3.9.

The plane wave incident electron beam passing through the specimen has been described in eq. 3.2.3. In practice the multi-slice method is better to understand electron dynamic scattering from defected crystal specimen. As the specimen can be sliced into defected slices with projected potential and perfect slices, the exit electron wave after passing trough a total specimen thickness of z_0 is then

$$\Psi(\mathbf{r}_p, z_0) = e^{\bar{\Delta}\epsilon} e^{i\sigma V_p^n} \cdots e^{\bar{\Delta}\epsilon} e^{i\sigma V_p^2} e^{\bar{\Delta}\epsilon} e^{i\sigma V_p^1} \psi(\mathbf{r}_p, 0) e^{2\pi i \mathbf{k}_0 \cdot \mathbf{r}} \tag{3.28}$$

The lens system in a microscope is not perfect. It will also affect the exit electron wave after the specimen. The most common aberrations are spherical aberration (C_s), astigmatism, comma, etc., and chromatic aberration (C_c). The influence by comma and low-order astigmatism can be minimized by proper alignment. In normal uncorrected microscope, the major contribution that will affect the final image are the first order (usually called defocus Δf) and third order (C_s) geometric aberrations. Neglecting other aberrations, the electron will get an additional phase factor $\chi(q)$ by the objective lens described by the

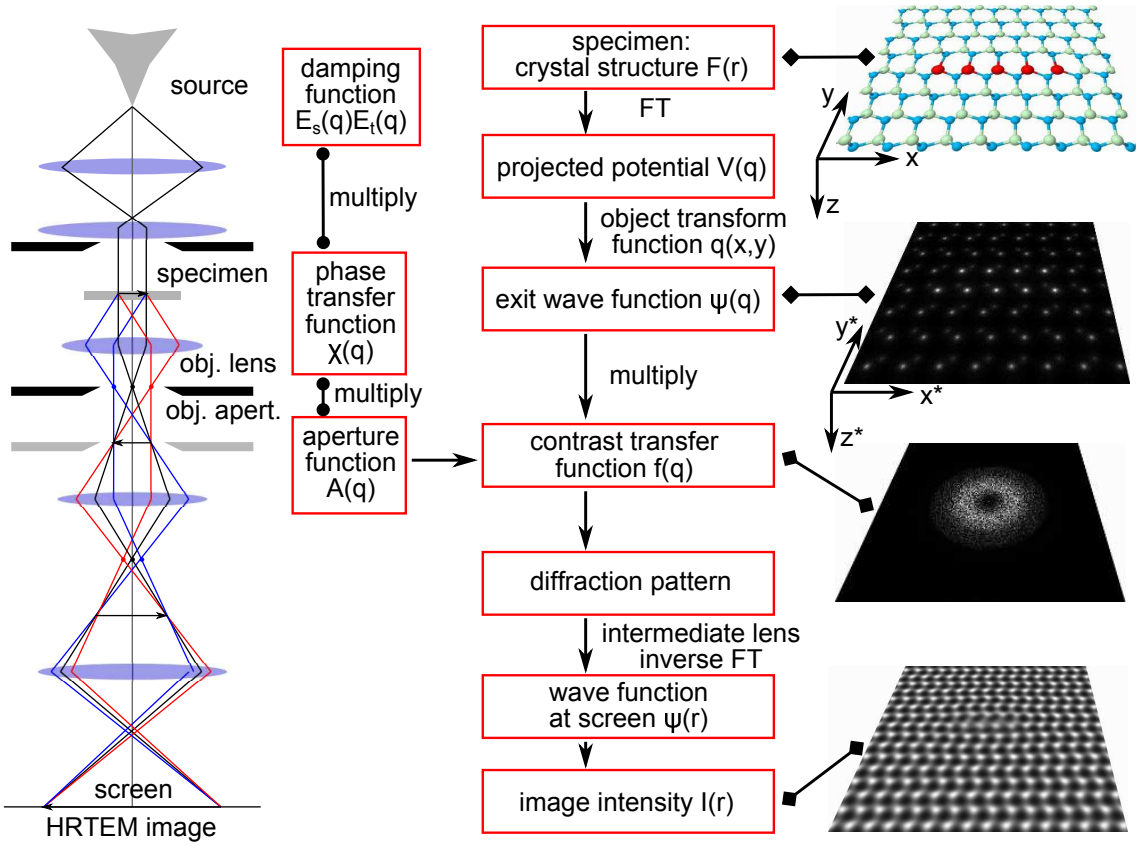


Figure 3.9: Image formation (contrast simulation) process of HRTEM images.

phase transfer function,

$$\chi(q) = \pi\left(\frac{1}{2}C_s\lambda^3 q^4 - \Delta f\lambda q^2\right). \tag{3.29}$$

Using the WPOA, if a microscope setting for which $\sin\chi(q) = -1$ over a significant range of q can be found, the corresponding image will show dark regions at the projected atom position. The defocus value is called the Scherzer defocus where the largest range of spatial frequency is transferred (the function $\sin(\chi(q, \Delta f))$ reaches the first minimum), $\Delta f = \sqrt{(3/2)C_s\lambda}$. The reciprocal quantity q at this point then defines the point resolution of a microscope. The phase transfer function is plotted for JEM-3010 in Fig. 3.10.

The partial coherence of the electron beam works as attenuation damping function to the electron function. The microscope transfer function is then

$$T(\mathbf{q}) = A(\mathbf{q})E_t(\mathbf{q})E_s(\mathbf{q})\exp[-i\chi(\mathbf{q})]. \tag{3.30}$$

$A(\mathbf{q})$ is the aperture function, $E_s(\mathbf{q})$ describe the defocus spread which introduce electron spatial incoherences and $E_t(\mathbf{q})$ describe the source spread which introduce electron temporal incoherence.

After propagating through the lens system, the final electron wave function in real space equals

$$\Psi_F(\mathbf{r}) = \Psi(\mathbf{r}, z) \otimes T(\mathbf{r}) = \mathcal{F}^{-1}[\Psi(\mathbf{q}, z)T(\mathbf{q})]. \tag{3.31}$$

We could then obtain the final image intensity $I(\mathbf{r}) = \Psi_F(\mathbf{r})\Psi_F^*(\mathbf{r})$.

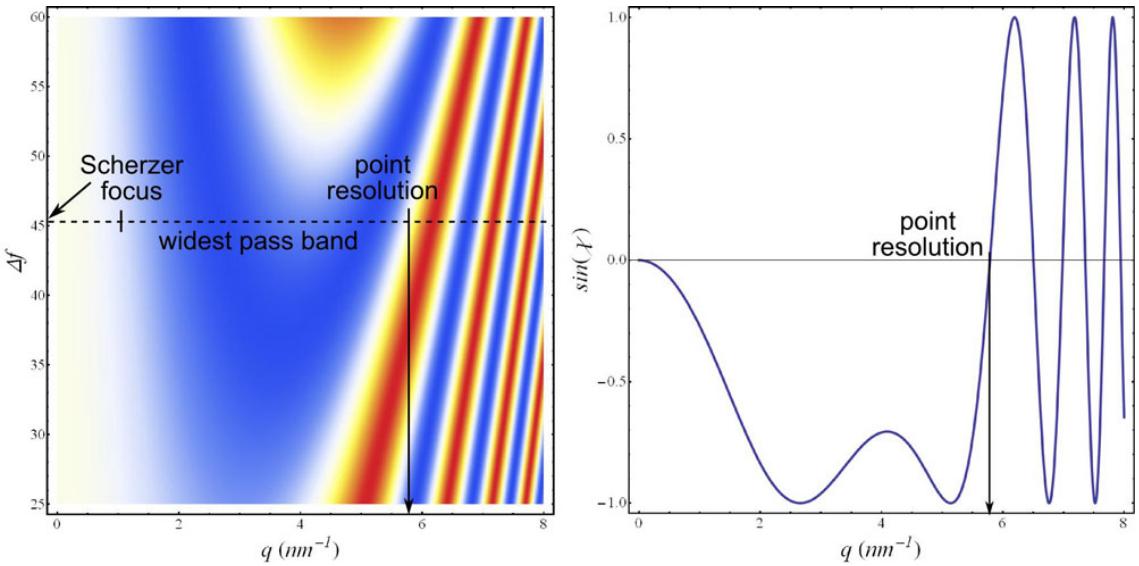


Figure 3.10: The phase transfer function for JEM-3010 ($C_s = 0.7$ mm, operated at 300 keV).

HRTEM contrast simulation

According to the contrast formation mechanism summarized in 3.3.2, and illustrated in Fig. 3.9 the HRTEM contrast simulations are performed following these steps:

- built super-cell with atomic positions of the investigated object
- slice the super-cell into slice (usually monolayer) stacks along the imaging projection direction (super-cell slice)
- calculate the projected potential of slices
- calculate the exit electron wave at the bottom of specimen according to the multi-slice scheme (equation [1.22])
- calculate the electron wave passing through the lens system according to [1.23]
- calculate the image intensity

In practice, by proper microscope alignment, the final image intensity will be reduce to only a function of specimen thickness t and defocus value of objective lens Δf. One example of the defocus-thickness map of GaN projected along the $[1\bar{2}10]$ direction is shown in Fig. 3.11. By varying this two values, the final image intensity oscillates from two positions: brighter dots corresponding to the projected atom columns (atom position) and darker dots corresponding to the projected atom columns (tunnel position).

Conventional quantitative retrieval of the atomic configuration[103] usually requires defocus series (multiple images with defocus difference known in experiment). By this way, the unknown parameters can be retrieved by appropriate algorithm[104,105], and therefore the exit wave function could be calculated from experimental images.

With only one single image, the defocus value can still be roughly obtainable by comparing the contrast transfer function to contributions from amorphous material in the image. Afterwards, thickness become the only parameter. Trough a thickness map, or by other independent thickness measurement, quantitative interpretation could also be achieved. The theoretical background of this method is provided in Ref. 106,107.

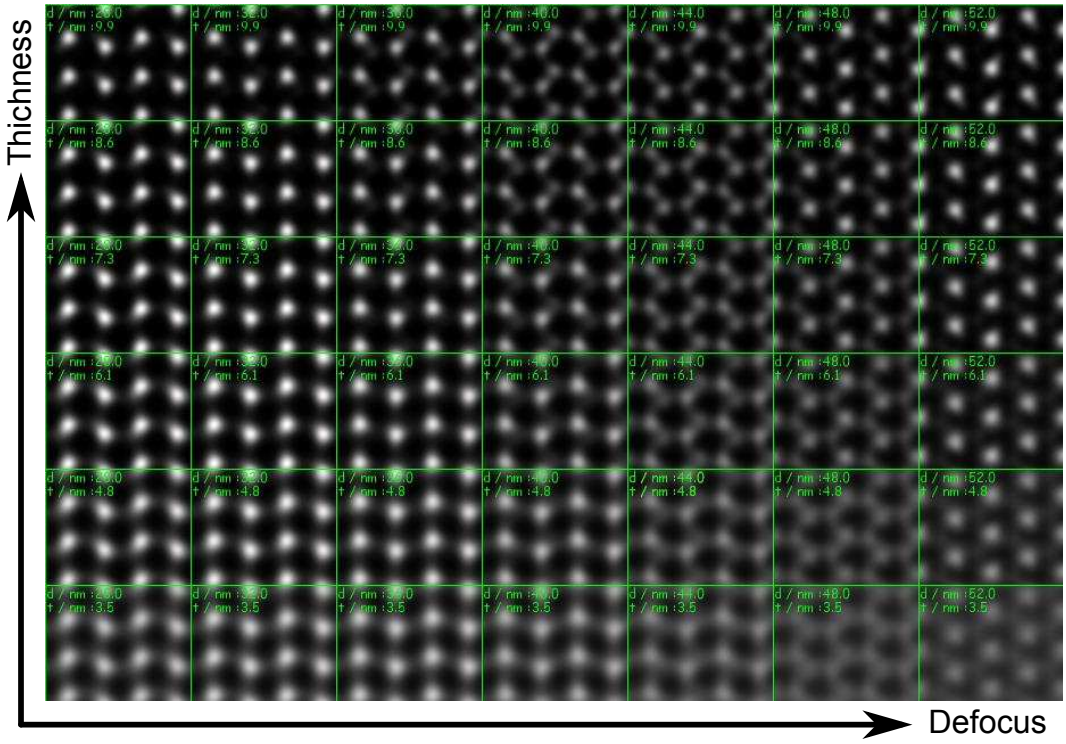


Figure 3.11: Thickness-defocus map of GaN projected along [1$\bar{2}$10] direction.

3.3.3. STEM image formation

The image formation is intuitively easy as already mentioned in Section 3.1. Various detector (BF, ADF, etc.) can be used in STEM to collect electrons for imaging. The transmitted beam, which is collected by the BF detector, are mostly contributed from the elastically scattered electrons. The electrons in the beam are in phase relative to atom position. The final image is an interference pattern of the overlapping diffraction disk collected by the BF detector. The BF-STEM imaging is therefore coherent imaging. In this case, the image intensity peak is dependent on the defocus of the objective lens, similar to HRTEM images[87].

In contrast, The high angle scattering electrons, which are collected by the ADF detector, are contributed from the thermal diffuse scattering. The higher the collecting angle (usually higher than 70 mrad[87]), the more enhancement in the atomic number (Z) dependence of the image contrast. This is often referred to as a high-angle ADF (HAADF), or Z-contrast image, which is the main technique used in this thesis with the STEM mode. On the other side, according to the Einstein model, there is no correlation between atomic vibration, therefore the exit wave are incoherent. The scenario can also be viewed by the frozen lattice approximation[108,109]: the high energy electrons is traveling so fast that they "see" the thermal vibrating atoms frozen at a certain phase. The collective atoms vibration give random phase to the electrons, therefore the final electron wave will not interference with each other. The image intensity is simply the convolution of probe distribution function and the specimen propagate function, and the intensity peak is therefore always located at atom position which does not change with the defocus setting[110]. The ADF-STEM imaging is known as incoherent imaging.

3.4. Geometric phase analysis: limitation and application

The geometric phase analysis (GPA) is an image processing technique. It can be applied to HRTEM images to quantitatively extract information about the local "lattice" distortions. In principle, it could be applied to any type of lattice/fringe image, like off-axis HRTEM fringes, HR-STEM image and the reconstructed exit wave lattice image etc. This section includes the basic principal of the method and a detailed discussion on its limitation and application.

3.4.1. Basic principal

A lattice image $I(\mathbf{r})$ can be expressed by the sum of inverse Fourier transform of the Bragg "spots", $\mathbf{g}$:

$$I(\mathbf{r}) = \sum_g H_g(\mathbf{r}) \exp(2\pi i \mathbf{g} \cdot \mathbf{r}). \tag{3.32}$$

The integration of inverse Fourier transform here becomes a summation because images are always composed of discrete pixels. Each Fourier coefficient stands in the real space for a set of lattice fringe, which can be treated as a position dependent amplitude and phase component. So the inverse Fourier transform of a single Fourier coefficient can be defined by

$$H'_g(\mathbf{r}) = A_g(\mathbf{r}) \exp[2\pi i \mathbf{g} \cdot \mathbf{r} + i P_g(\mathbf{r})]. \tag{3.33}$$

The phase component here is called the geometric phase, because it describes the position dependent lattice deviation. If there are deviations of the spacing between the lattice fringes, there will be changes to the geometric phase and it can be interpreted in two ways as illustrated in Fig. 3.12.

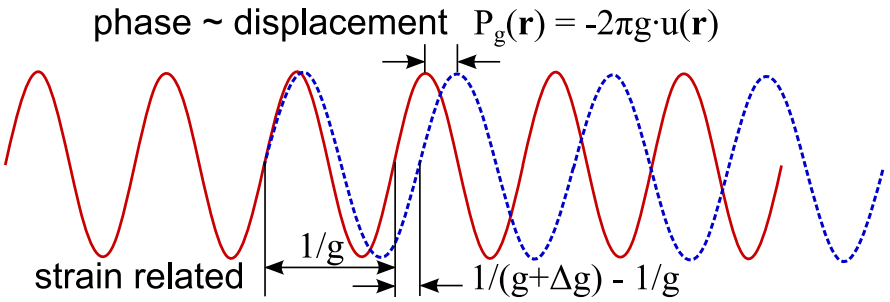


Figure 3.12: The relationship of geometric phase and displacement/strain.

In the first manner, if we introduce a position dependent variation in the change of the reciprocal lattice vector, $\mathbf{g} \to \mathbf{g} + \Delta\mathbf{g}(\mathbf{r})$, then

$$H'_g(\mathbf{r}) = A_g(\mathbf{r}) \exp[2\pi i \mathbf{g} \cdot \mathbf{r} + 2\pi i \Delta\mathbf{g}(\mathbf{r}) \cdot \mathbf{r} + i P_g]. \tag{3.34}$$

By comparing with equation 3.33, we get

$$P_g(\mathbf{r}) = 2\pi \Delta\mathbf{g}(\mathbf{r}) \cdot \mathbf{r}. \tag{3.35}$$

The strain information which is contained in $\Delta\mathbf{g}(\mathbf{r})$ can be extracted from the derivative of the geometric phase.

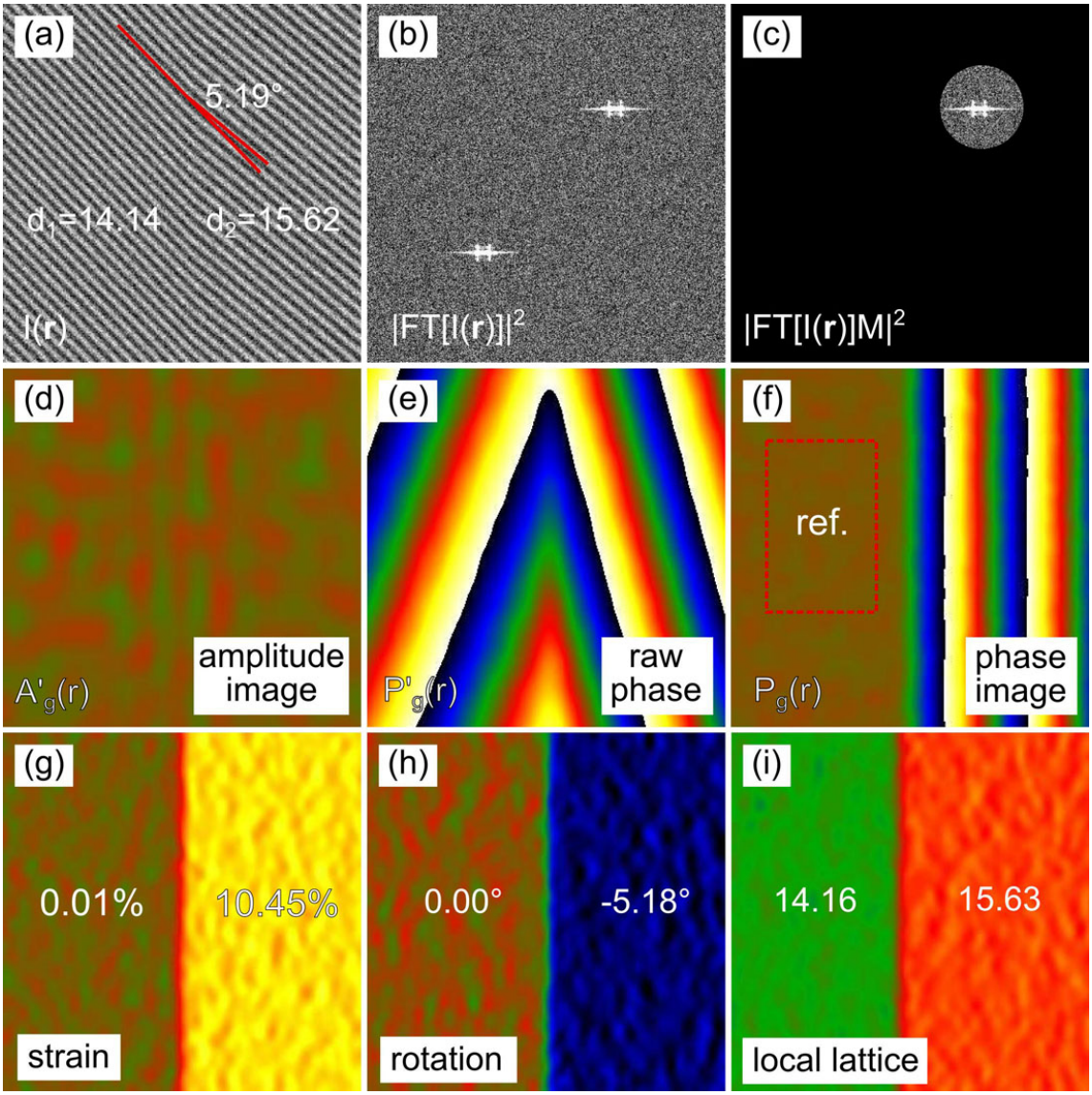


Figure 3.13: The principal procedure of GPA: (a) fringe image under analysis; (b) power spectrum of (a); (c) after applying mask on a Fourier "spot"; (d),(e) extracted raw amplitude and phase image, respectively; (f) refined phase image with respect to the marked region; (g)–(i) the derived strain, rotation and local lattice parameter maps, respectively.

Alternatively, if we introduce a displacement $\mathbf{u}$ to the position, $\mathbf{r} \rightarrow \mathbf{r} - \mathbf{u}(\mathbf{r})$, the deviation of phase can also be interpreted as the displacement of the local lattice position. Similar to the previous treatment, we get

$$\boxed{P_g(\mathbf{r}) = -2\pi \mathbf{g} \cdot \mathbf{u}(\mathbf{r}).} \tag{3.36}$$

It is obvious that if we know the exact direction and magnitude of $\mathbf{g}$, the displacement field information can be extracted directly from the geometric phase.

In practice, the principal steps of GPA is explained in Fig. 3.13, with the starting image generated with given "lattice" parameters [Fig. 3.13(a)]. The input image is 256×256 pixels with artificial 5% of Gaussian noise. Left and right side of the image has aver-

age fringe spacing 14.14 and 15.62 pixels, respectively. the right side fringes is 5.19 degrees rotated anticlockwise to the left ones. For this given fringe image, strain equals $(15.62 - 14.14)/14.14 = 10.47\%$.

First, calculate the Fourier transform of the input image [Fig. 3.13(b)]. Second, select and mask a reciprocal fringe vector $\mathbf{g}'$ (i.e., the "averaged" $\mathbf{g}$ of the whole image) with mask function M [Fig, 3.13(c)]. This mask function works as an envelope in the Fourier space on a small area around $\mathbf{g}'$. The mask should be centred at $\mathbf{g}'$ with only one main frequency and exclude the origin (zero frequency). All through the thesis, we apply the round shaped Gaussian mask. The effect of mask radius will be discussed in the next subsection. After performing inverse Fourier transform of the mask filtered image, the amplitude and phase component are extracted [Fig. 3.13(d) and (e)]. The amplitude component is an envelope to the Fourier coefficient. It will be discarded in the following GPA procedures but can be helpful to assess the quality and usability of the HRTEM image being processed. The raw phase component $\mathbf{g}'$, as in Fig. 3.13(e), contains an average effect of the masked pixels. A more accurate value of the vector $\mathbf{g}$ can be (and must be) calculated by selecting a reference area in the image, and by minimizing with respect to $\mathbf{g}$ [Fig. 3.13(f)][41,53]. In other words, the phase image is refined in such a way that the selected reference area has a zero (geometric) phase. This refinement is an important step in GPA as it returns the accurate $\mathbf{g}$ vector for further calculations. Afterwards, the strain is calculated from the derivatives of the phase image. The rotation is extracted from the directional relation of local $\mathbf{g}(\mathbf{r})$ vectors and the local lattice/fringe parameters are extracted from the magnitude of local $\mathbf{g}(\mathbf{r})$ vectors. As can be clearly seen, the extracted information by GPA is quite close to the input image, even when we have introduced a noise signal of 5% of the peak signal intensity.

As a standard procedure introduced by Hÿtch *et al.*[41], the two dimensional (2D) displacement field/strain tensor can be extracted from two non-colinear Fourier frequencies, $\mathbf{g}_1$ and $\mathbf{g}_2$. We have

$$P_{g1}(\mathbf{r}) = -2\pi \mathbf{g}_1 \cdot \mathbf{u}(\mathbf{r}) = -2\pi[g_{1x}u_x(\mathbf{r}) + g_{1y}u_y(\mathbf{r})], \tag{3.37}$$

$$P_{g2}(\mathbf{r}) = -2\pi \mathbf{g}_2 \cdot \mathbf{u}(\mathbf{r}) = -2\pi[g_{2x}u_x(\mathbf{r}) + g_{2y}u_y(\mathbf{r})]. \tag{3.38}$$

Write in matrix form,

$$\begin{bmatrix} P_{g1}(\mathbf{r}) \\ P_{g2}(\mathbf{r}) \end{bmatrix} = -2\pi \begin{bmatrix} g_{1x} & g_{1y} \\ g_{1x} & g_{1y} \end{bmatrix} \begin{bmatrix} u_x(\mathbf{r}) \\ u_y(\mathbf{r}) \end{bmatrix} = -2\pi \mathbf{G} \begin{bmatrix} u_x(\mathbf{r}) \\ u_y(\mathbf{r}) \end{bmatrix}. \tag{3.39}$$

The vectorial displacement field can be calculated by

$$\begin{bmatrix} u_x(\mathbf{r}) \\ u_y(\mathbf{r}) \end{bmatrix} = -\frac{1}{2\pi}\mathbf{G}^{-1} \begin{bmatrix} P_{g1}(\mathbf{r}) \\ P_{g2}(\mathbf{r}) \end{bmatrix}. \tag{3.40}$$

The strain matrix, which is the first derivatives of the displacements, can then be calculated by

$$\epsilon = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} \\ \epsilon_{yx} & \epsilon_{yy} \end{bmatrix} = \begin{bmatrix} \frac{\partial u_x}{\partial x} & \frac{\partial u_x}{\partial y} \\ \frac{\partial u_y}{\partial x} & \frac{\partial u_y}{\partial y} \end{bmatrix} = -\frac{1}{2\pi}\mathbf{G}^{-1} \begin{bmatrix} \frac{\partial P_{g1}(\mathbf{r})}{\partial x} & \frac{\partial P_{g1}(\mathbf{r})}{\partial y} \\ \frac{\partial P_{g2}(\mathbf{r})}{\partial x} & \frac{\partial P_{g2}(\mathbf{r})}{\partial y} \end{bmatrix}. \tag{3.41}$$

The matrix at the left hand side is the deformation matrix[111]. The strain tensor is obtained by taking the symmetric part of the above deformation matrix, and the rotation is

equal to the anti-symmetric part:

$$\varepsilon = 1/2(\epsilon + \epsilon^T), \tag{3.42}$$

$$\omega = 1/2(\epsilon - \epsilon^T), \tag{3.43}$$

where T denote the transpose of the matrix.

An example of the procedures to extract the 2D strain tensor from a HRTEM image is shown in Fig. 3.14. The input image is an experimental HRTEM micrograph of an InAs nano-cluster semi-coherently embedded in Si matrix with edge-on projection misfit dislocation loops at the interface and a stacking fault in the middle of the cluster. We demonstrate here only the procedure with GPA method. We will come back to the detailed discussion of the material aspect (strain state, interface, misfit dislocations, etc.) in Chapter 5.1. The procedure to obtain the GPA strain from the HRTEM image follows the previously described principle. An important step is to choose two reflections that are not linearly correlated. In the case of zincblende structure projected in [110] zone axis, the two {111} reflections are selected to extract the phase information and to calculate the strain. Finally we could extract the strain tensor ϵ_{xx}, ϵ_{yy} and ϵ_{xy} and the rigid body rotation ω_{xy}.

3.4.2. The mask function: the relation between resolution and accuracy

The spatial resolution of GPA method is directly defined by the inverse of the radius of mask function. Hence, the larger the mask radius, the finer the spatial resolution of the extracted displacement and strain. In terms of accuracy, the smaller the mask radius, the higher the accuracy of the measurement. This can be understood in the following way.

The mask function is an envelope in the reciprocal space on the selected main Fourier coefficient. The Fourier transform can be interpreted as "fitting" the raw image function with series of sinusoidal functions. The closer the raw image function to sinusoidal function, the sharper the Fourier coefficient in the frequency/reciprocal space. In real HRTEM image, the lattice fringe is periodic but the intensity is not necessarily in perfect sinusoidal function, thus the Fourier coefficient for the selected fringe has a spread. Therefore, the larger the radius of mask function, the closer the fitting sinusoidal series to the original fringe, or say, higher spatial resolution. This can also schematically understood with Fig. 3.15. With larger mask radius more information about the deviation of the image fringes from perfect sinusoidal function will be included, which results in a higher fluctuation of the extracted strain maps. This in turn lead to a poorer accuracy with larger mask radius.

To reach the highest possible accuracy and spatial resolution, it is suggested to use fringes above the limit of microscope point resolution. In this case, the "lattice spots" in the HRTEM image have the least number of "averaged" real atom columns. For example, GaN 0002 fringe is the "averaging" electron interference patten of 2 atom columns (i.e., a Ga and a N atom column pair), whereas 0001 fringe consists of contribution from 4 atom columns. Therefore, the analyse using 0002 fringe can provides more accurate information and higher in spatial resolution than using 0001 fringe. As illustrated in Fig. 3.16. On the other hand higher order lattice fringe presented in the power spectrum is from the fine structure of the interference pattern which already beyond the microscope point resolution.

As is already discussed previously, higher GPA resolution needs larger mask radius which lead to the loss in accuracy. This is especially obvious for lattice fringes that is

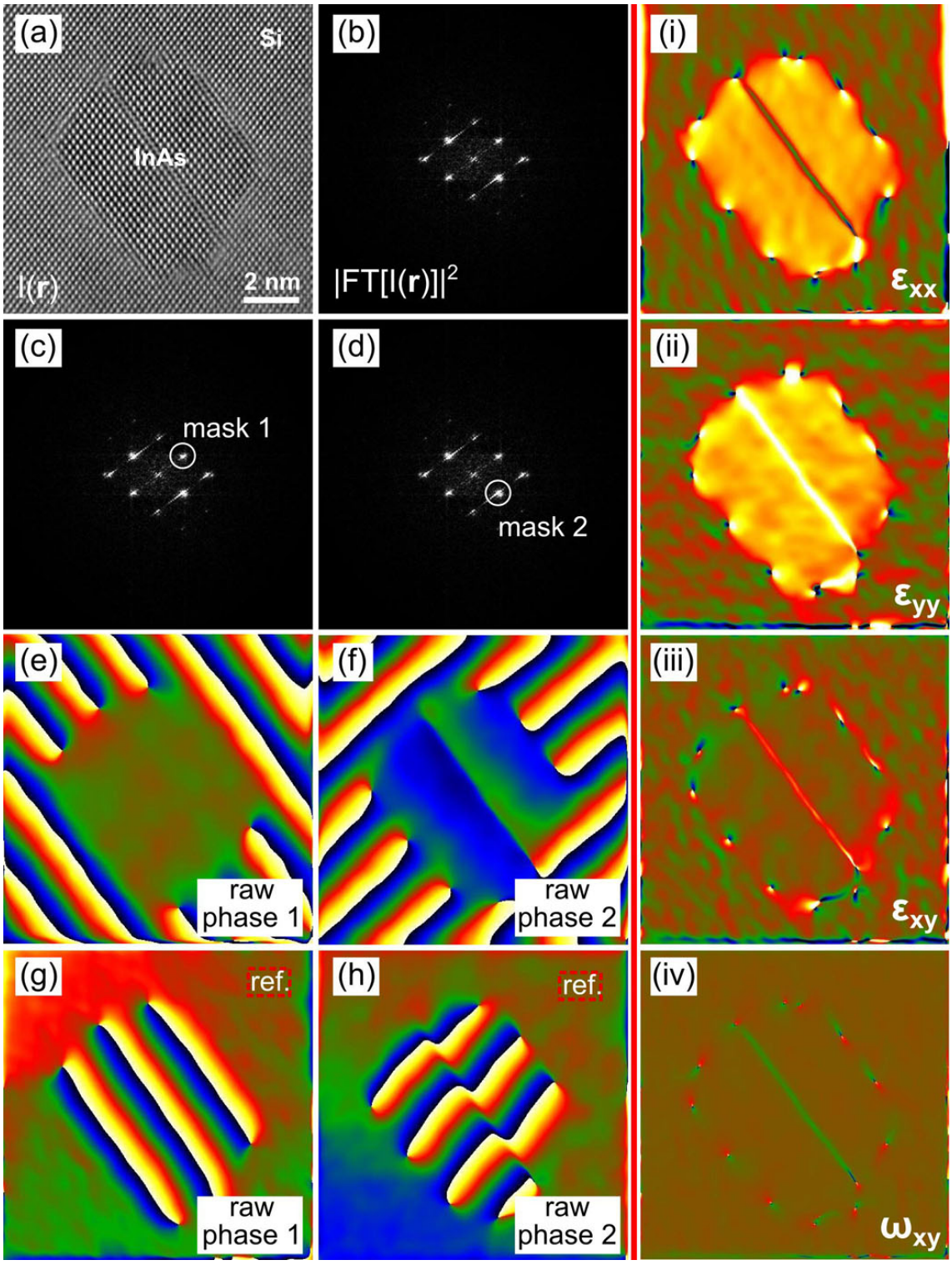


Figure 3.14: Procedure of 2D GPA strain map from real HRTEM lattice: (a) HRTEM image of a InAs cluster embedded in Si semi-coherently; (b) power spectrum of (a); (c), (e) and (g) are the masked g_1 component, the extracted raw phase P'_{g1} and refined phase P_{g1}, respectively, while (d), (f) and (h) correspond to g_2. From P_{g1} and P_{g2}, the strain tensors and rotation are obtained and are shown in (i)-(iv), respectively.

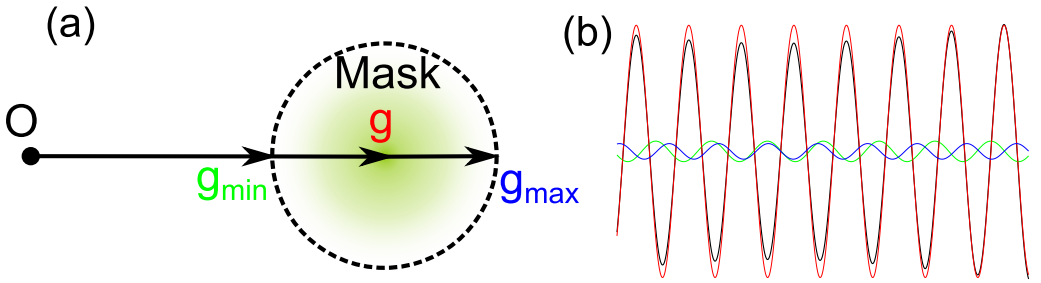


Figure 3.15: Fourier coefficient included in a given mask radius (a). (b) more Fourier components give more accurate (higher spatial resolved) description of the original fringe function (thick black).Red, green and blue lines are Fourier coefficient from the red, green and blue **g** in (a), respectively.

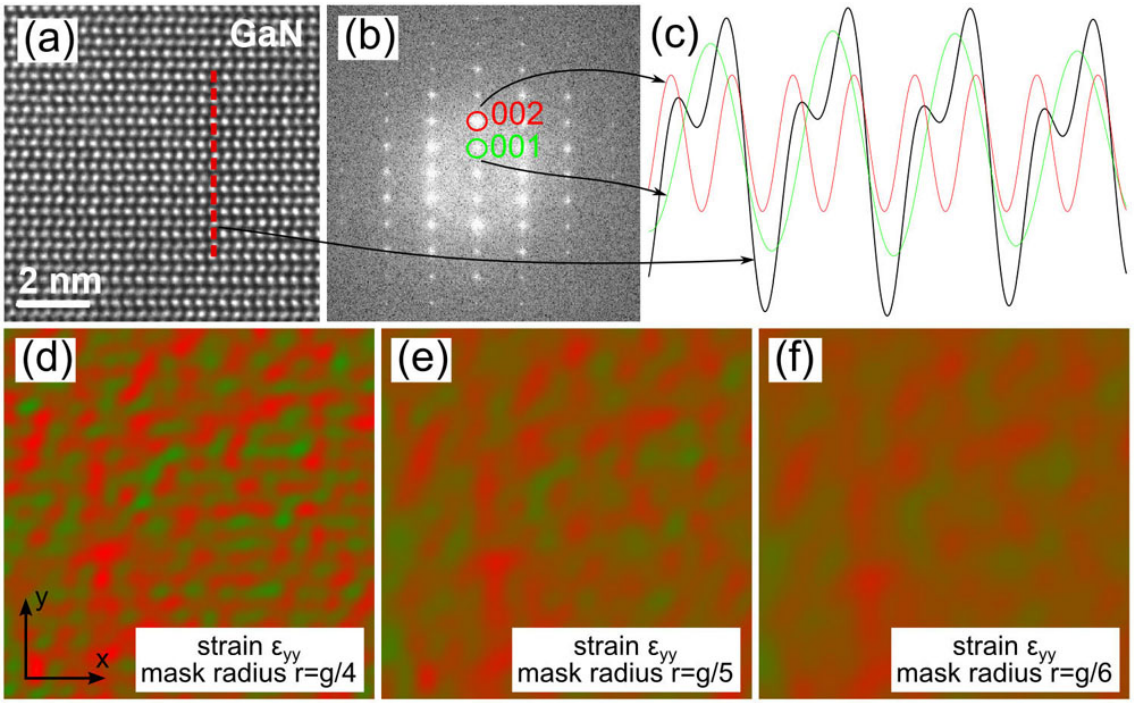


Figure 3.16: Optimizing mask radius for measuring GaN by GPA. (a) experimental HRTEM image of GaN along $[1\bar{2}10]$ zone axis; the dashed line profile is shown in (c) as the thick black line. (b) The power spectrum of (a). selected main Fourier coefficient representing 0001 fringes and 0002 fringes are shown in (c) in green and red line, respectively.

non-central symmetric. As an example show in Fig. 3.16(d)–(f), GaN[0001] direction is non-central symmetric. Larger mask size ($r = g/4$) bring more Fourier coefficient thus in term show more "fine structure" [cf. Fig. 3.16(d)] in the strain mapping instead of the strain information because of the nature of non-central symmetry (that deviate more from perfect sinusoidal function than central symmetry reflections). For the analysis of a good quality HRTEM images of GaN in our microscope, we find $r = g/5$ [cf. Fig. 3.16(e)] is a good balance between spatial resolution and accuracy. Thus this mask radius is applied for GaN later.

3.4.3. Accuracy of GPA displacement and strain measurement

As already introduced previously, two types of information can be extracted from HRTEM images: displacement and strain. While strain is more widely used in discussing the material related properties, we will show that the GPA measurement of displacement have an accuracy of one order of magnitude higher than GPA strain measurement. This are very important, especially when we measure very small value of strain in a very localized region (i.e., the study of GdN cluster produced strain in Chapter 6).

Strain is the first derivative of displacement as defined in equation 2.3 (in page 9). Because the information extraction by GPA are based on discrete numerical calculation, the numerical derivative algorithm will result accuracy degeneration, due to the error transfer in the numerical calculations. Here, we illustrate in the following with two model images where the strain is strain function is artificially assigned to represent cases of real crystal lattice image. The two model "lattice images" both have a dimension of 512 $\times$ 512 pixels, with a artificial scale defined 0.02 nm per pixel. The lattice spacing in the y direction keep constant while in the x direction is variable. The image intensity is defined as

$$I(\mathbf{r}) = \cos(2\pi \cdot \omega(x)) + \cos(2\pi \cdot \omega_0(y)), \quad \text{with} \quad \omega(x) = \begin{cases} \omega_0 = 1/15 & : x < x_0 \\ \omega(x) & : x \geq x_0 \end{cases}$$

where x_0 is a "interface". A Gaussian random noise with 10% of the maximum intensity of the images signal is added to the images to account for 10% of experimental noise level. The strain function is then $\epsilon(x) = (\frac{1}{\omega(x)} - \frac{1}{\omega_0})/(\frac{1}{\omega_0})$, which is shown by the red line in Fig. 3.17(b) and Fig. 3.18(b), respectively. This means the lattice fringe in the x direction has a "lattice spacing" of 15 pixels before the "interface" position x_0, and is variable according the strain function. The expected displacement (integration of the strain function with integration constant $C = 0$) is also shown in the same figures. The first model is capable to describe a heterostructure with constant strain separated by a interface. Similar case is presented in Chapter 5.1. The second model describes a heterostructure with non-constant strain separated by an interface.

The displacement and strain are extracted from the model lattice images following the steps described in section 3.4.1 (in page 38), and the results are presented in Fig. 3.17(c) and (e) and Fig. 3.18(c) and (e). Gaussian round mask function with a radius of 8 pixel is used in the calculation, which define a spatial resolution of $(8/512)^{-1} = 64$ pixel $=$ 1.28 nm. Afterwards, line profiles of displacement/strain from the extracted maps are taken as depicted by the arrow and shown in Fig. 3.17(d) and (f) and Fig. 3.18(d) and (f).

The line profile is used to fit with the defined displacement/strain function. The fitting parameters, residual sum of square (RSS) and adjusted R-square which characterize the exactness of the fitted data (or, the accuracy of the measurement), are summarized in Tab. 3.3.

It can be clearly seen that the RSS of fitted displacement data is an order of magnitude smaller than the RSS of the fitted strain data. The adjusted R-square of the fitted displacement data is almost equal to 1, whereas the value from the strain data fitting is deviate more from 1. This means the accuracy of displacement measurement from GPA is about an order of magnitude higher than strain measurement, independent of the lattice image model.

We note that displacement is a relative value which relies the definition of "zero displace" position. In the current lattice models, the position is apparently defined as the

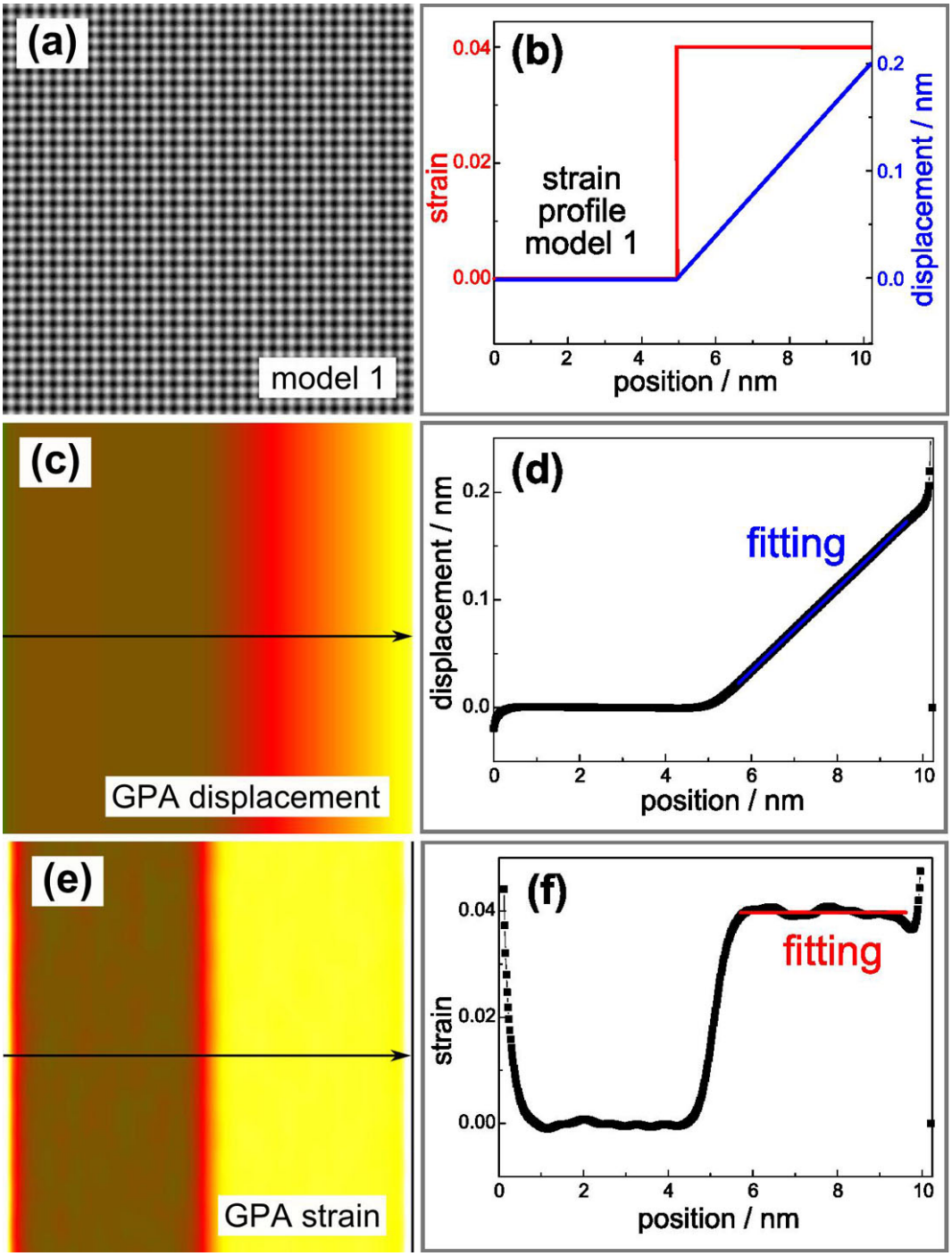


Figure 3.17: Accuracy of displacement and strain measurement with model 1. Computer generated lattice image (a) with given strain along the horizontal direction defined by the strain function as shown in (b); GPA measurement of displacement (c) and strain (e) from the model image; Extracted displacement (d) and strain (f) line profile from the corresponding images are used to fit the displacement/strain function. The fitting parameters are summarized in Tab. 3.3 characterize the accuracy of the measurements.

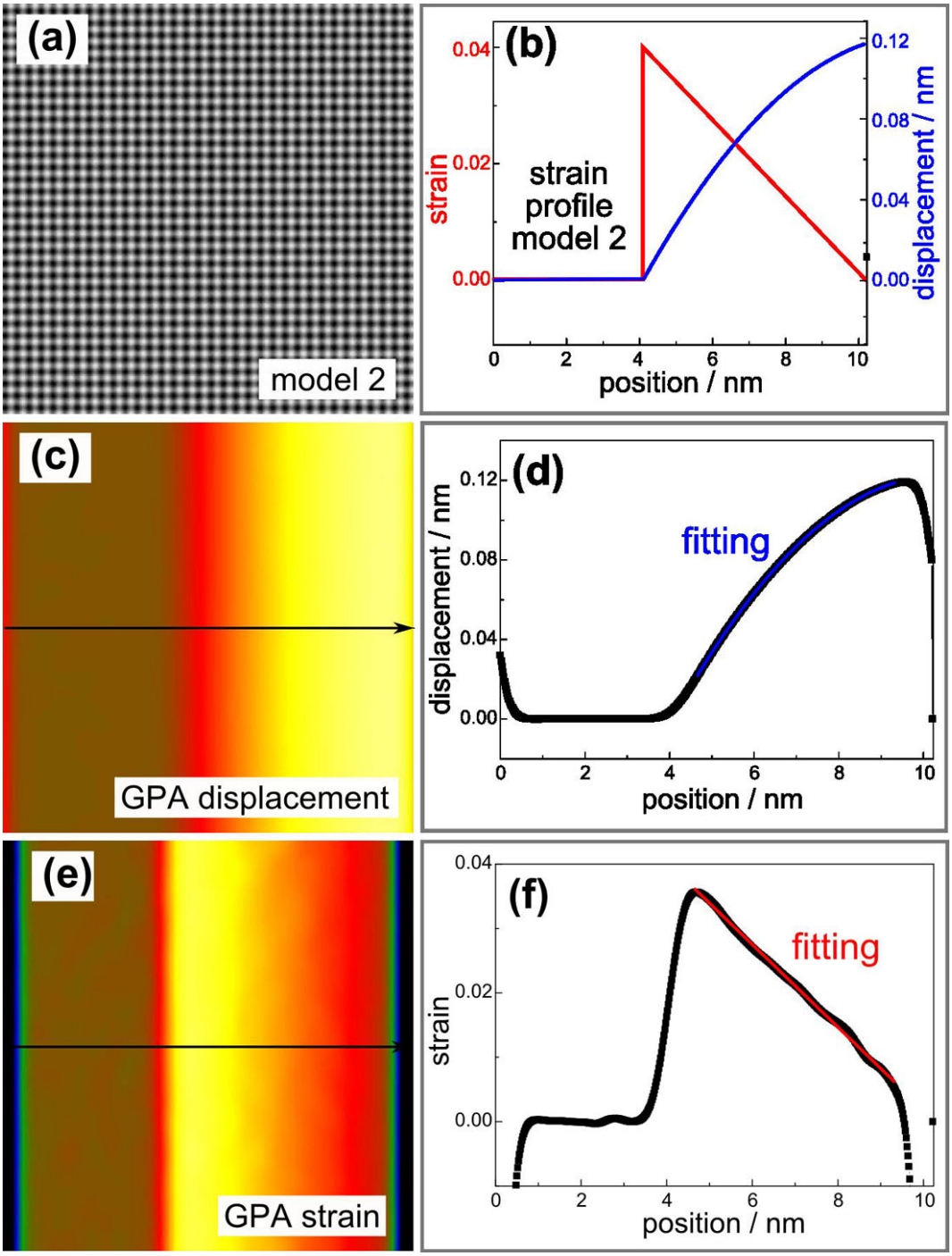


Figure 3.18: Accuracy of displacement and strain measurement with model 2. Computer generated lattice image (a) with given strain along the horizontal direction defined by the strain function as shown in (b); GPA measurement of displacement (c) and strain (e) from the model image; Extracted displacement (d) and strain (f) line profile from the corresponding images are used to fit the displacement/strain function. The fitting parameters are summarized in Tab. 3.3 characterize the accuracy of the measurements.

Table 3.3: Summary of fitting parameters of GPA measured displacement and strain from the two models as shown in Fig. 3.17 and Fig. 3.18. RSS (residual sum of square) and Adj. R^2 (adjusted R-squared) are parameters characterize the exactness of the fitting process. Smaller value of RSS and Adj. R^2 closer to 1, respetively, means higher accuracy.

	displacement		strain	
	RSS	Adj. R^2	RSS	Adj. R^2
model 1	2.811×10^{-6}	0.99999	6.945×10^{-5}	0.24145
model 2	2.662×10^{-6}	0.99999	3.173×10^{-5}	0.99819

reference lattice ($x < x_0$ region). In real lattice image analysis, it is important to define the "zero displace" position that is physical plausible.

3.4.4. Discussion on artifacts

There are already intensive theoretical discussions on the interpretation of GPA measurements and parameters that affect the accuracy of the GPA [52,53,112]. I discuss here some specific issues that emerged during the study of the thesis.

Thickness and defocus effects As described in section 3.3, the lattice spot position in HRTEM images are sensitive to the defocus and thickness change. GPA measurement accuracy is sensitive to the sample thickness change and defocus. The thickness inhomogeneity will result in the difference of electron phase at the exit plane because the difference in phase grating by the specimen. This effect will finally transferred to the image. Although very small, it can still be measured by GPA. An example is given in Fig. 3.19: a thickness change of 5 nm in a distance of 10 nm result in a measured artifact displacement of 0.021 Å/nm, linearly increasing along the measured direction.

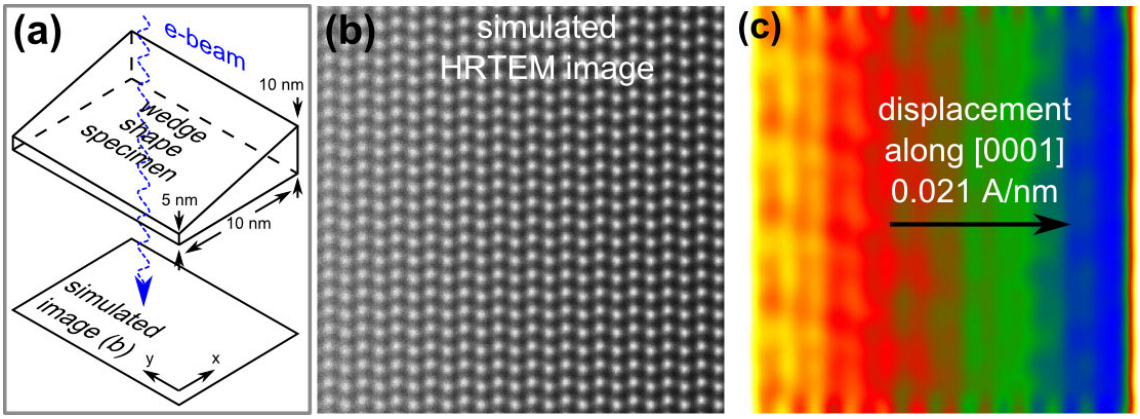


Figure 3.19: Thickness variation between 5 and 10 nm result in small shift of the spots in the HRTEM interference pattern. Measured by GPA, this shift is about 0.021 Å/nm, which is not a real displacement.

When the specimen thickness is constant, the defocus change will only shift the overall position of the lattice spots, i.e. the relationship for all the pixels between real atomic position and lattice spots kept constant. Therefore, artifacts can be minimized when the specimen has almost constant thickness. For heterostructure the strain gradient will in-

troduce inaccuracy to the measurement at the interface (see e.g., the extracted information at the "interface" region in Fig. 3.17 and Fig. 3.18), this can be partly compensated in experiment by fine tuning of the defocus to an optimal value, as suggested by Chung and Rabenberg[112],

$$\Delta f = -C_s \lambda^2 g^2, \tag{3.44}$$

where g is the magnitude of the analysed $\mathbf{g}$ vector.

Imaging magnification Because the recorded image is pixel image, for the accuracy of discrete Fourier transform, it is important to record the image fringes with enough sampling points. In this thesis, all the images are recorded in magnification that the lattice fringes to be analysed have a spacing of 10 - 25 pixels.

Raw image filtering GPA is an Fourier space analysing method. As described in section 3.4.1, the mask function is an frequency pass filter, therefore frequency pass filter like bandpass and low/high pass filters, which can enhance the image quality in real space, will have no constructive effect on the GPA results.

Wiener filter is an statistical filter which minimizing the error between the anticipated signal and the input signal. Therefore, statistical noise like thermal noise from the recorded CCD camera can be filtered by Wiener filter. Contrast features which are introduced by an amorphous material covering the specimen can also be suppressed by Wiener filter.

Projector lens distortion and CCD defects The projection lens aberration may cause image distortions (see Fig. 3.20). In the discussion from Ref. 113, Hüe proposed an post-acquisition algorithm to compensate the effect. Actually, working with our microscope setup, we find that the effect is only obvious when recording the image with large size detector (e.g., film). When images are recorded with small CCD camera, the distortion is not significant and can even be neglected. Alternatively, it can be minimized by spreading the beam as shown in Fig. 3.20(b). In fact, it is found that the projector lens distortion is negligible for systems equipped with typical CCD camera, which covers smaller area than the film camera system. While for the images captured by film, the projector lens distortion should be taken into account.

However, in the CCD camera system or during the digitization of film, the data over flow for specific pixels or bad CCD pixels do have a strong affect on the GPA analysis. An example is shown in Fig. 3.21. The recorded image of Fig. 3.21(a) has an average pixel intensity of 2000 counts, while the 6 bad pixels [magnified in Fig. 3.21(b)] are overflown and and have a intensity of about 30000. Colour lines in Fig. 3.21(b) are the iso-phase lines for the corresponding Fourier coefficient. Bad pixels lay on the fringe peak of the red line so they only introduces small artefacts in the strain measured from spot 1. For the blue line, the bad pixels lay between two fringe peaks, so it shows the strongest artefacts. After fixing the bad pixels by assign the value to average of the neighbour pixels, the artefacts disappeared.

3.4.5. Practical guidelines

In practice these rules should be followed for minimizing artefacts and errors:

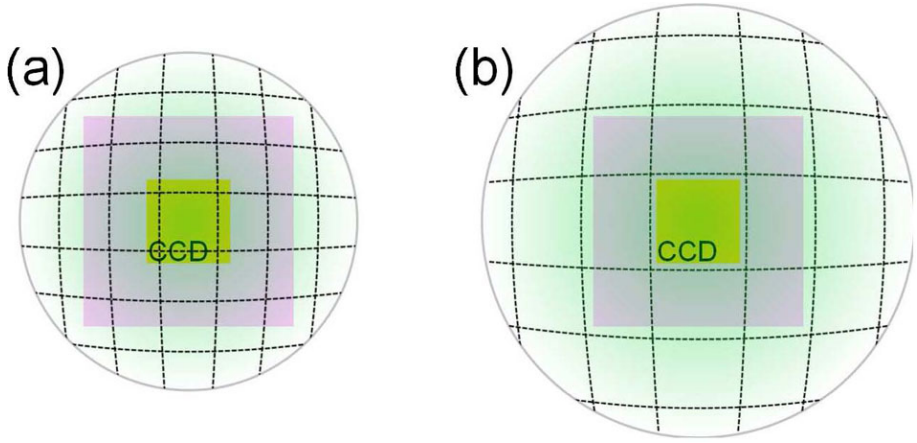


Figure 3.20: Projector lens distortion effect. For small CCD camera, the effect is not significant; for the same CCD size, the effect can be reduced if the beam is more spread.

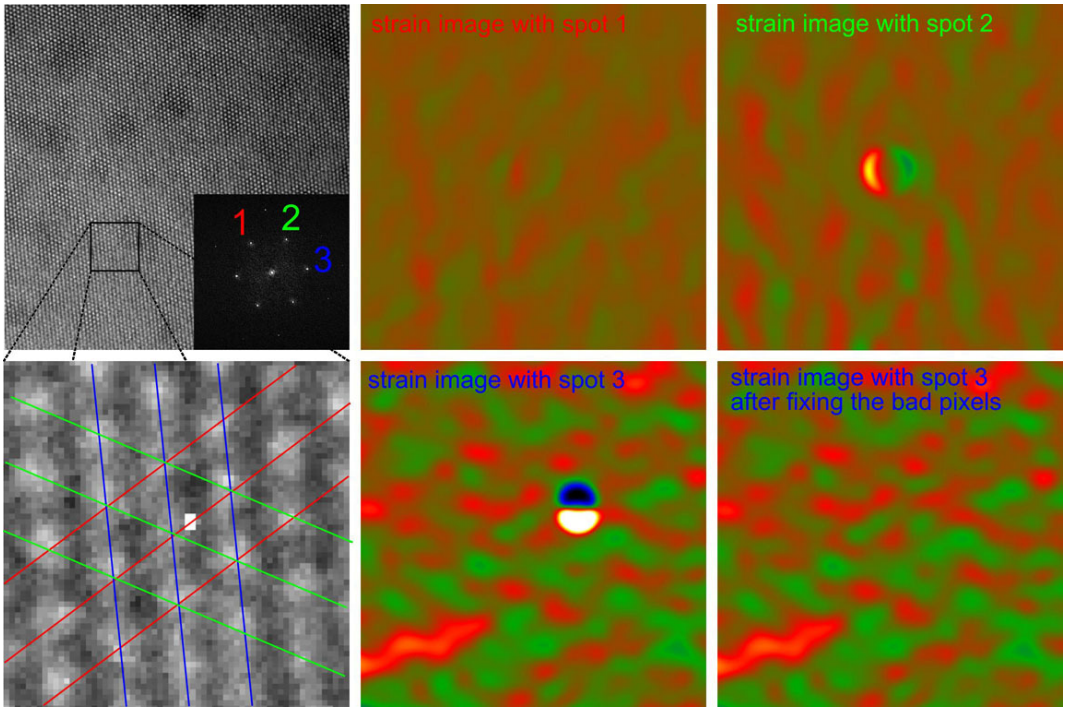


Figure 3.21: "Bad" CCD pixels with output overflow causing artefacts in the GPA measurements.

- choose conditions (thickness and defocus) where the fringe contrast is maximal and avoid regions where the fringe contrast changes rapidly;
- choose magnification that the lattice fringes are with at least 10 pixel space;
- explode the CCD during image recording long enough to suppress thermal noise;
- spread the beam before recording the image;
- analyse central-symmetric or symmetry-related reflections;

- verify of the displacement field using different reflections;
- verify the analysis at different defocus values.

Chapter 4.

Observation and analysis of nano-clusters embedded in semiconductor epilayers

The central objective to study the microstructure of nano-clusters and clustering in semiconductor epilayers relies first the observation of the nano-clusters. The detection of nano-scale objects embedded in matrix is challenging. Comprehensive analytic characterization methods are dedicated to it, including most of modern microscopic and spectroscopic methods. In this thesis, we consider (S)TEM techniques as the ideal way to observe the nano-clusters and carry out the study. In this chapter, interpretation of various image contrasts produced by nano-clusters will be discussed with examples. These examples include (1) coherent Bi-rich Ga(As, Bi) nano-clusters embedded in GaAs matrix and rhombohedral Bi clusters incoherently embedded in GaAs matrix which is later studied in Chapter 5.2, (2) semi-coherent InAs nano-clusters (or quantum dots) embedded in Si which is later studied in Chapter 5.1 and (3) theoretical models of GdN platelet cluster in GaN or single Gd substitution Ga in GaN which considerably assist our study of the structure of Gd-doped GaN. Quantitative interpretation of image contrast relies on contrast simulation of the structure that is modeled properly. Therefore structure modeling methods used in this thesis are summarized and compared next. Finally, the discussion is presented on the detectability of few atom sized GdN clusters models in GaN.

4.1. Challenges in observation of nano-clusters

(S)TEM as the ideal tool The analysis of nano-scale objects at the first place requires analytical tools that provide nanometer resolution, and information of nanometer scale accuracy. Traditional methods like X-ray scattering can in principle provide resolution or information in nanometer scale. However, the traditional approaches rely on Bragg reflections and usually fail to reveal nano-clusters, especially when they are coherently strained in the matrix lattice. The versatile imaging and diffraction modes in a (S)TEM can provide comprehensive information about nano-clusters in a thin specimen up to atomic resolution. Therefore, it comprises an ideal tool to study the nano-clusters.

In Fig. 4.1, we show an example to detect coherently strained zincblende Bi-rich Ga(As, Bi) nano-clusters embedded in $GaAs_{1-x}Bi_x$. The figure show the compared results of X-Ray diffraction (XRD) and TEM on the same $GaAs_{1-x}Bi_x$ sample before and after thermal annealing. In Fig. 4.1(a), the GaAs (004) peak from the substrate is aligned before and after thermal annealing. The $GaAs_{1-x}Bi_x$ epilayer is strained with respect to the GaAs substrate in the as-grown sample. This resulted in the shift of (004) reflection by about 480 arcsecond (mismatch shift). After thermal annealing, the mismatch shift is reduced to about 250 arcsec. In both pattern, the thickness fringes reflect the good epilayer crystal quantity and the high interface/surface quantity [83]. However, the interpretation of the changes in mismatch is not straightforward.

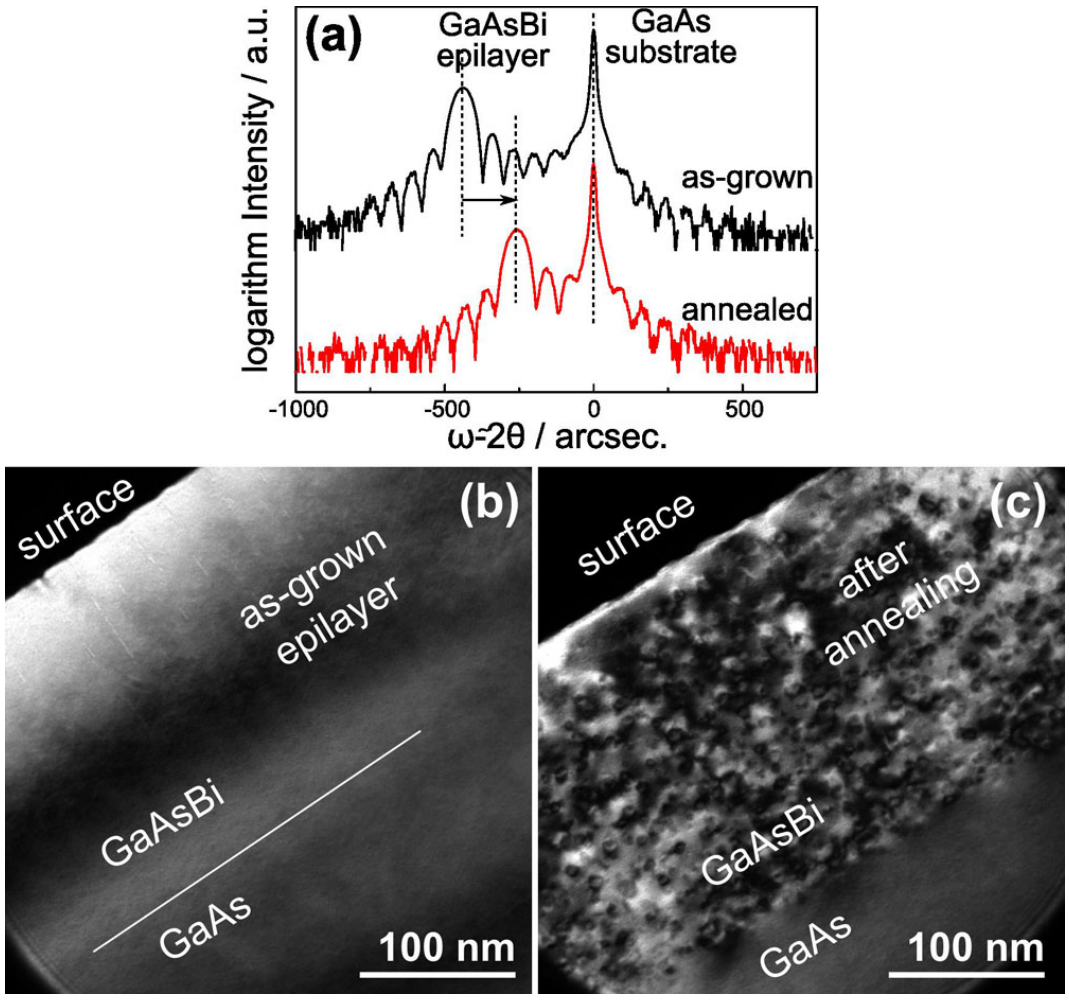


Figure 4.1: (a) ω-2θ scan of (004) reflection of GaAsBi epilayer on GaAs substrate. The strained epilayer peak before thermal annealing has a higher "mismatch" to the substrate peak than after annealing. This information cannot directly interpret as cluster formation. The same sample examined under TEM before (a) and after (c) annealing. Cluster formation is clearly revealed by the strain contrast feature, as will be discussed in detail in Section 4.2. [Courtesy XRD measurements by Janne Puustinen at Tampere University of Technology, Finland.]

In contrast to the situation with XRD, TEM can provide direct evidence of cluster formation. Compare the results as shown in Fig. 4.1(b) and (c)—the $\mathbf{g}_{004}$ strain contrast dark-field TEM micrographs—the bright-dark contrast lobs (coffee bean contrast) in the annealed sample indicate locally strained clusters of about 10 nm in diameter. This will be explained and discussed in more detailed manner in section 4.2 and in Chapter 5.2.

Requirements of TEM specimen In the study of nano-clusters using TEM, a great challenge comes from the specimen preparation. Electron transparent specimen is a prerequisite to study in a (S)TEM, that is, a maximum thickness of about 200 nm for most semiconductor materials. For the case to achieve atomic resolution and accuracy, the specimen thickness is usually optimal within several to a dozen of nanometers. On the other hand, nano-clusters are small 3D objects in matrix. When electrons passing through

a thin specimen of embedded clusters, they carry information not only about the cluster but also from the matrix. This is called the matrix effect, which made the quantitative interpretation very complex and challenging.

It is important to prepare a specimen with homogeneous thickness that is comparable to the diameter of clusters to minimize the matrix effect There is no universal methods to prepare such specimen. Our "mechanical polishing–dimpling–ion milling" routine (described in detail in Appendix B) can produce specimen of high quality, but the nano-cluster-matrix relation is spatially uncontrolled (i.e., it is not able to selective prepare a special part of the sample). Generally, this is not a problem for samples with relatively high density of clusters (case studied show in Chapter 5.1 and 5.2). In samples with low density of clusters, this problem becomes the most challenging part.

Multi-scale information about nano-clusters Despite the availability of multiple imaging and diffraction modes in (S)TEM for probing nanostructures, finding an accurate comprehensive structural solution for nanostructure is a formidable challenge because any one of the technique yields only a partial view (structural, chemical compositional, or spatial information) of the structure. The problem is further aggravated in nano-clusters because substantial fractions of atoms in the clusters being associated with surfaces both external (interface) and internal (planar defects). In addition, structural information is often needed at more than one length scale: from the atomic scale (sub nm), to the thickness of interfaces (1 nm), and to the size of nano-clusters (2-10 nm). Therefore, a combined input from multiple experimental techniques in a (S)TEM and theory is needed to generate a more comprehensive insight into the nano-clusters.

4.2. Interpretation of contrast from nano-clusters

According to the previous discussion of the interface character about nano-clusters embedded in the matrix (section 2.1.4, in page 11), the clusters can be coherent, semi-coherent or incoherent to the matrix. The difference in the interface character is accompanied with different character of lattice strain field, which can be detected and characterized by TEM. Generally, a nano-cluster embedded in matrix can be imaged [1], (1) by the strain contrast, (2) by the difference in the structure factor between cluster and matrix, (3) by the diffraction of the planes (in incoherent clusters) that do not exist in cluster or matrix, and/or (4) by the difference in the constituents atomic numbers (Z-contrast). However, there is no general methods like a "check list" that can applied to observe the embedded nano-clusters. It is always problem specific. Thus we must consider the character of the object, not only the microscopy. In the following, we will discuss how to apply the techniques to some examples of nano-clusters of various interface and strain character.

4.2.1. Image contrast of coherent nano-clusters

Coherent clusters have one to one atom registry cross the cluster-matrix interface. The lattice misfit strain is accommodated elastically. This strain field of a coherent interface is usually long-range. It will produce strain contrast.

[1] We consider only microscopic method in TEM, and will not discuss spectroscopic method here, like X-ray energy-dispersive spectroscopy, electron energy-loss spectroscopy, etc.

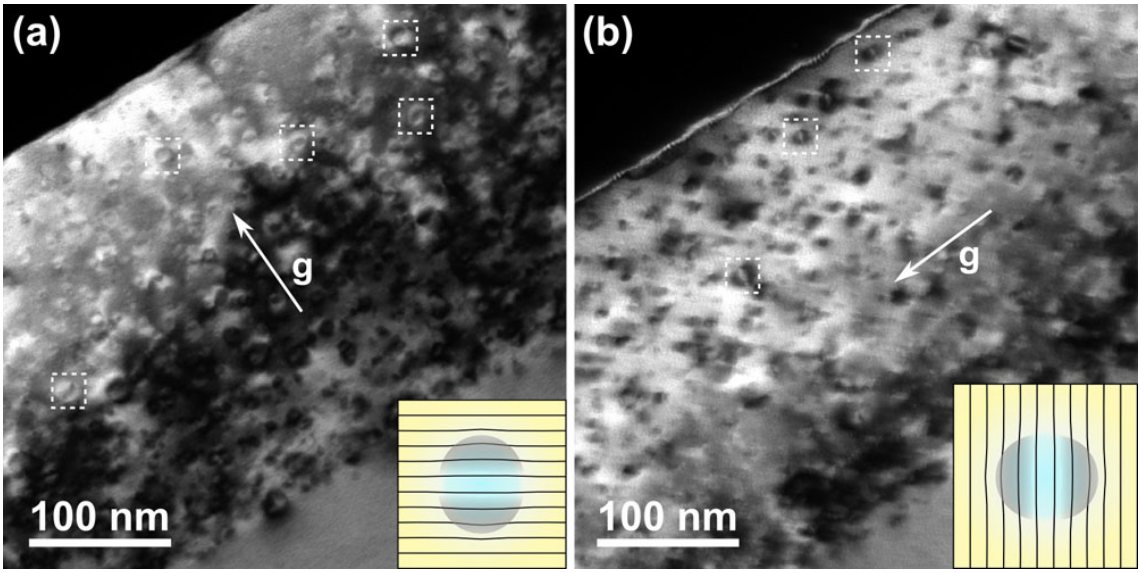


Figure 4.2: Typical "coffee bean" strain contrast images taken (marked by dashed boxes) from coherent Bi-rich clusters embedded in GaAs matrix with different diffraction vectors: (a) Imaged with $\mathbf{g}_{004}$ and (b) imaged with $\mathbf{g}_{2\bar{2}0}$. Usually, there is a "line of no contrast" in the middle of the "coffee bean" perpendicular to the diffraction vector. The inset schematically depict the formation of the contrast.

Strain contrast In Fig. 4.2 the typical "coffee bean" strain contrast is illustrated by two-beam condition dark-field image from coherent clusters. The image show a cross-section of a $GaAs_{1-x}Bi_x$ epilayer after thermal annealing where lots of coherent zincblende Bi-rich Ga(As, Bi) nano-clusters are created (details see Chapter 5.2). The "coffee bean" feature is formed by the distorted cluster and matrix lattice planes. Usually, there is a "line of no contrast" in the middle of the "coffee bean" perpendicular to the diffraction vector[100]. This is illustrated with the graphs in the insets of the figure. By selecting various specific lattice planes (diffraction vectors), the distribution of the strain field can be deduced roughly, which in turn can suggest the shape of the cluster strain field.

The "coffee bean" contrast depends on the strain field of the cluster/matrix, on the extinction distance ξ_g and experimental deviation parameter s, and on the depth of the cluster from the specimen surface. A more quantitative interpretation of the contrast can be derived from numerical integration of the dynamic Howie-Whelan equation (equation 3.26, on page 32) with appropriate structure modeling. This is not a routinely formidable task, because of the difference in the geometry of structure model and experimental conditions. Each model/condition combination is almost unique to be solved with the dynamic Howie-Whelan equations. Numerous work over the half century has been dedicated to the quantitatively understanding of strain contrast of clusters embedded in matrix, e.g. in Refs. 98–100,102,114,115.

HRTEM lattice image contrast Figure. 4.3 show a typical HRTEM lattice image of coherent Bi-rich clusters embedded in GaAs matrix. The red circles mark the position where the clusters are embedded. The lattice pattern follow that of zincblende GaAs. As can be seen in the inset of Fig. 4.3, the Fourier transform of the image show only the lattice periodicity of GaAs in [110] zone axis.

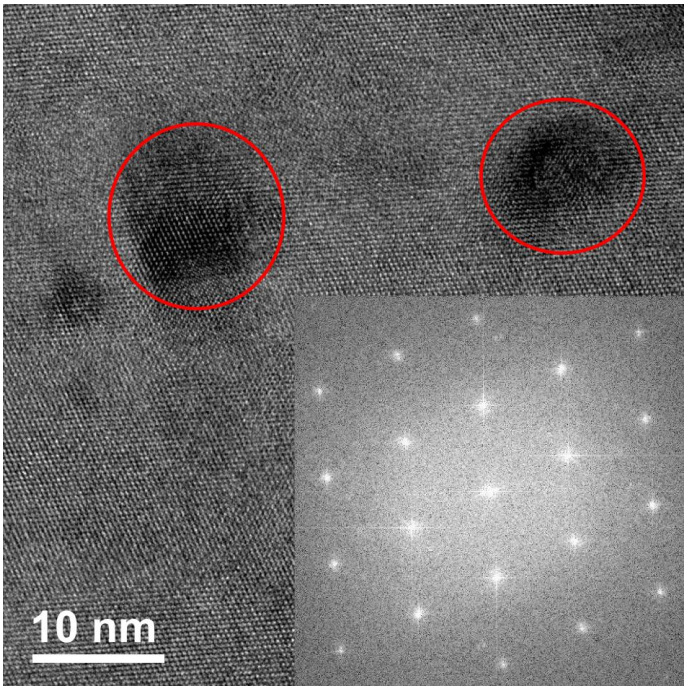


Figure 4.3: HRTEM lattice images of coherent Bi-rich clusters embedded in GaAs matrix, darker region maked by red circles are the Bi-rich Ga(As, Bi) nano-clusters. Fourier transform of the image in inset show only reflection of GaAs.

4.2.2. Image contrast of semi-coherent nano-clusters

Some semi-coherent clusters (in highly mismatched systems with dislocation arrays or loops) have epitaxial relation to the matrix; but the lattice misfit strain is (partly) relieved by the misfit dislocation loops, which have an strain field that is localized around the dislocation core[cf. Fig. 4.4(d)]. Because of the difference of the common lattice plane, there will be modulation in the coincidence in the lattice planes (or Moiré pattern). An example of semi-coherent InAs clusters burred in Si matrix is show in the following.

Diffraction contrast Figure 4.4(a) show a dark-field image of semi-coherent InAs clusters burred in Si matrix obtained with $\mathbf{g} = 004$. No detectable "coffee bean" strain contrast (as shown in Fig. 4.2) is seen in the Si matrix, whereas Moiré fringe, is seen in the clusters. One typical cluster is magnified in Fig. 4.4(b). A line profile of the intensity along the modulation direction is extracted in Fig. 4.4(c). The modulation has an periodicity of 1.51 nm. This contrast feature can be interpreted by the model as depicted in Fig. 4.4(d). The relaxed InAs clusters has different lattice distance than Si matrix. The overlapping InAs(004) planes and Si(004) planes result in contrast modulation periodicity corresponds to the least common multiple of Si(004) and InAs(004) planes (or, the coincident lattices):

$$(n+1) \times [004]a_{Si} = n \times [004]a_{InAs}, \quad n \text{ is integer.} \tag{4.1}$$

Using the bulk value of lattice parameters of Si ($a_{Si} = 5.4301$ Å) and InAs ($a_{InAs} = 6.0584$ Å), the modulation periodicity of 1.51 nm is found very close to 11 times of Si(004)

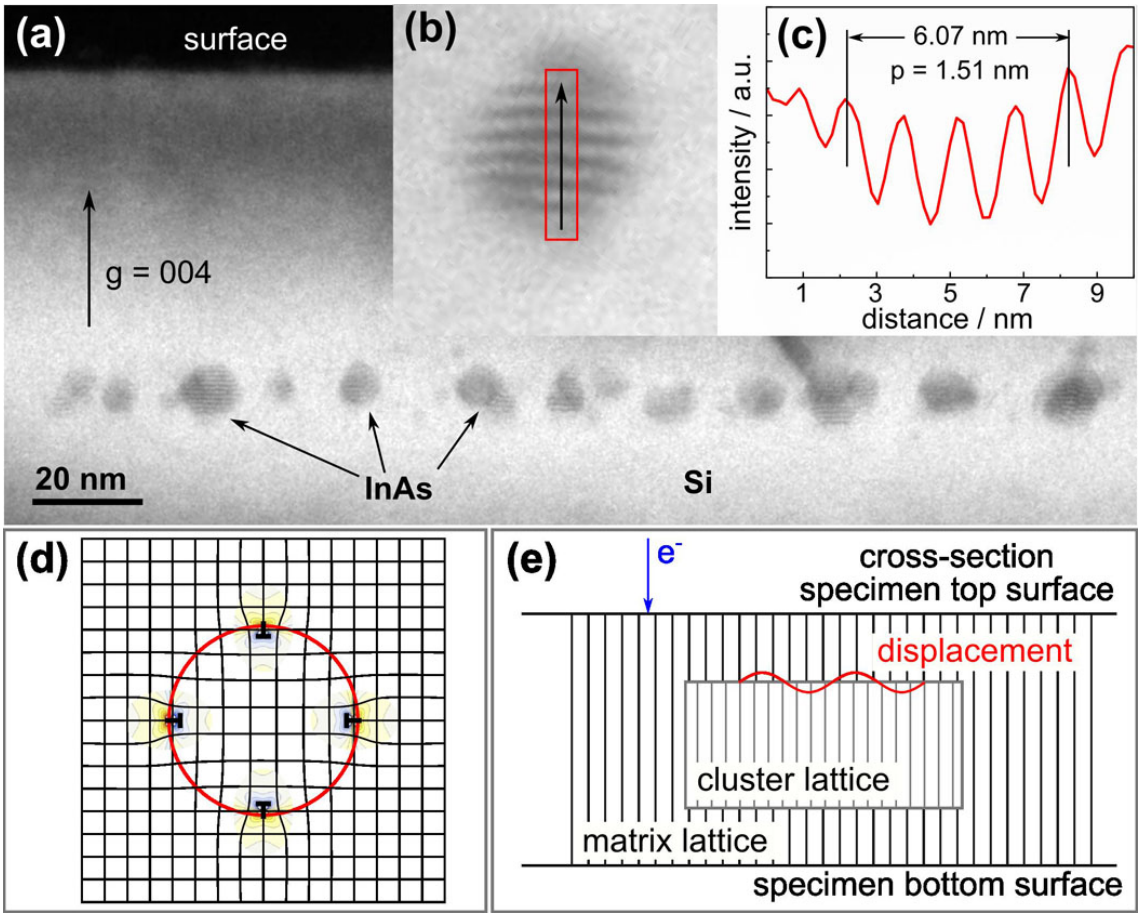


Figure 4.4: Diffraction contrast and analysis of semi-coherent InAs clusters embedded in Si matrix. (a) The $\mathbf{g} = 004$ strain contrast image show no detectable strain in Si matrix; InAs clusters burried in Si matrix have spherical faceted shape. (b) A typical cluster is magnified with modulation fringe contrast. (c) the fringes have a periodicity of 1.51 nm, corresponding to 10 InAs/ 11 Si 004 plane distance, see text. (d) Illustrate a semicoherent cluster (d) depict the modulation fringe contrast formation.

and 10 times of InAs(004), that is

$$(10+1) \times \frac{0.5431}{4} = 10 \times \frac{0.6058}{4} = 1.51 \text{ nm}, \tag{4.2}$$

This indicate the buried InAs clusters are (almost) completely relaxed from the lattice misfit strain; and indeed close to its bulk lattice parameter.

HRTEM lattice image contrast When the specimens are prepared good enough to enable some nacked cluster explode directly to the penetrating electron beam without matrix effect, the analysis become much more straightforward, like the examples studied in Chapter 5.1. When they are embeded in matrix the analysis is bit more complex.

Figure 4.5 show a typical HRTEM image of semi-coherent InAs cluster embedded in Si matrix. The modulation in the coincidence lattice planes between InAs and Si resulted in the formation of Moiré pattern, as marked in the image by the red circle. This is also revealed by the satellite spots around the main zincblend GaAs spots in the Fourier

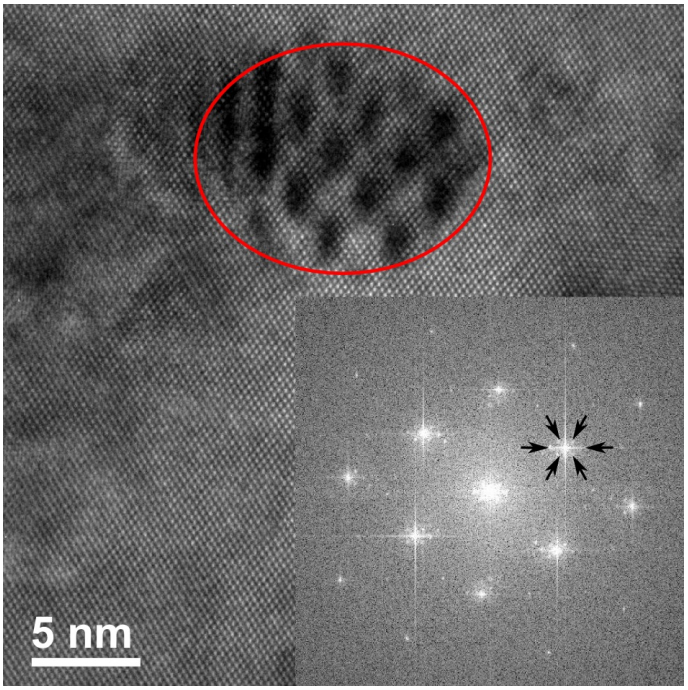


Figure 4.5: HRTEM lattice images of semi-coherent InAs cluster buried in Si matrix resulted in Moiré pattern. Fourier transform of (b) in inset show "extra reflection" of the Moiré patten.

transform of the HRTEM image, as shown in the inset of Fig. 4.5. The periodicity of the Moiré "lattice" pattern can be either measured from the direct space lattice image, or read from the reverse distance of satellite spot to main spot in the Fourier transformed image. The semi-coherent cluster and matrix follow the same crystal direction; and there are no or neglegible rotation between them. Therefore the Moiré "lattice" pattern periodicity directly follows

$$\lambda_{hkl}^{\text{Moiré}} = \frac{1}{g_{hkl}^{\text{cluster}} - g_{hkl}^{\text{Si}}} = \frac{1}{1/d_{hkl}^{\text{cluster}} - 1/d_{hkl}^{\text{Si}}} = \frac{d_{hkl}^{\text{Si}}}{1 - d_{hkl}^{\text{cluster}}/d_{hkl}^{\text{Si}}}. \tag{4.3}$$

Therefore, the lattice plane distance of the cluster can be calculated from the periodicity of Moiré pattern,

$$d_{hkl}^{\text{cluster}} = d_{hkl}^{\text{Si}} + \frac{(d_{hkl}^{\text{Si}})^2}{\lambda_{hkl}^{\text{Moiré}}}. \tag{4.4}$$

Using the lattice parameter of bulk Si ($a^{\text{Si}} = 5.4309$ Å) and InAs ($a^{\text{InAs}} = 6.0584$ Å), the results are summzried in Tab. 4.1.

Comparing the lattice plane distance of the calculated cluster and bulk InAs in the table, it is clearly see that the cluster has very close lattice plane distances to the corresponding value of bulk InAs.

4.2.3. Image contrast of incoherent nano-clusters

Incoherent clusters is characterized by the lack of interfacial coherent strain and usually do not have common lattice planes of matrix crystal. Therefore, by selecting the lattice

Table 4.1: Calculated InAs lattice planes distance from Moiré pattern. $\lambda^{\text{Moiré}}$ Moiré periodicity (measured in Fourier transform), d^{Si} lattice plane distances of bulk Si, d^{cluster} calculated lattice plane distances of embedded cluster and d^{InAs} lattice plane distances of bulk InAs. The error bar is defined by the pixel size in Fourier space.

	$\lambda^{\text{Moiré}}$ (Å)	d^{Si} (Å)	d^{cluster} (Å)	d^{InAs} (Å)
d_{111}	28.7 ± 1.15	3.14	**3.48**$\pm$0.01	**3.50**
d_{220}	18.1 ± 0.13	1.92	**2.12**$\pm$0.01	**2.14**
d_{002}	23.5 ± 0.75	2.72	**3.03**$\pm$0.01	**3.03**

planes of the matrix for DF imaging, the clusters will appear dark; or the other way around: the cluster will appear bright if selecting a reflection **g** corresponding to the cluster lattice only.

Diffraction contrast An example of rhombohedral Bi (rh-Bi) clusters incoherently embedded in GaAs matrix is shown in Fig. 4.6. The projected shape of rh-Bi clusters are revealed as marked by the white dashed boxes. The interface between cluster and matrix can be seen very sharp from the image.

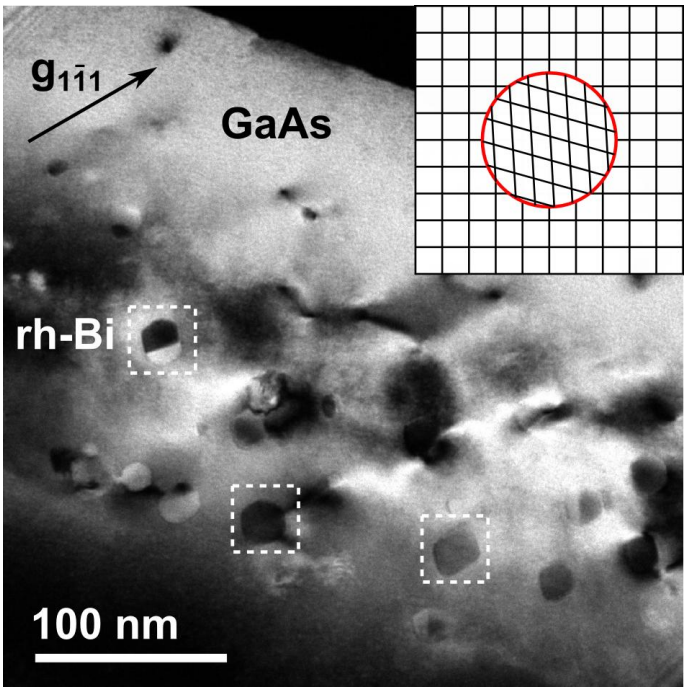


Figure 4.6: Typical diffraction contrast DF image (marked by dashed boxes) from incoherent clusters. In the sample, incoherent rh-Bi clusters formed in the GaAsBi epilayer after annealing. The image was obtained using GaAs $\mathbf{g}_{1\bar{1}1}$. The image was obtained using reflection only in GaAs, so the cluster appears dark. The inset illustrate a incoherent cluster embedded in matrix. rh-Bi, rhomboheral Bi clusters.

HRTEM lattice image contrast Incoherent cluster may have no definite crystal orientation relation. Therefore, when the matrix lattice is oriented on zone axis, the cluster lattice may not exactly on axis. In many cases, the matrix crystal overlapping the cluster crystal

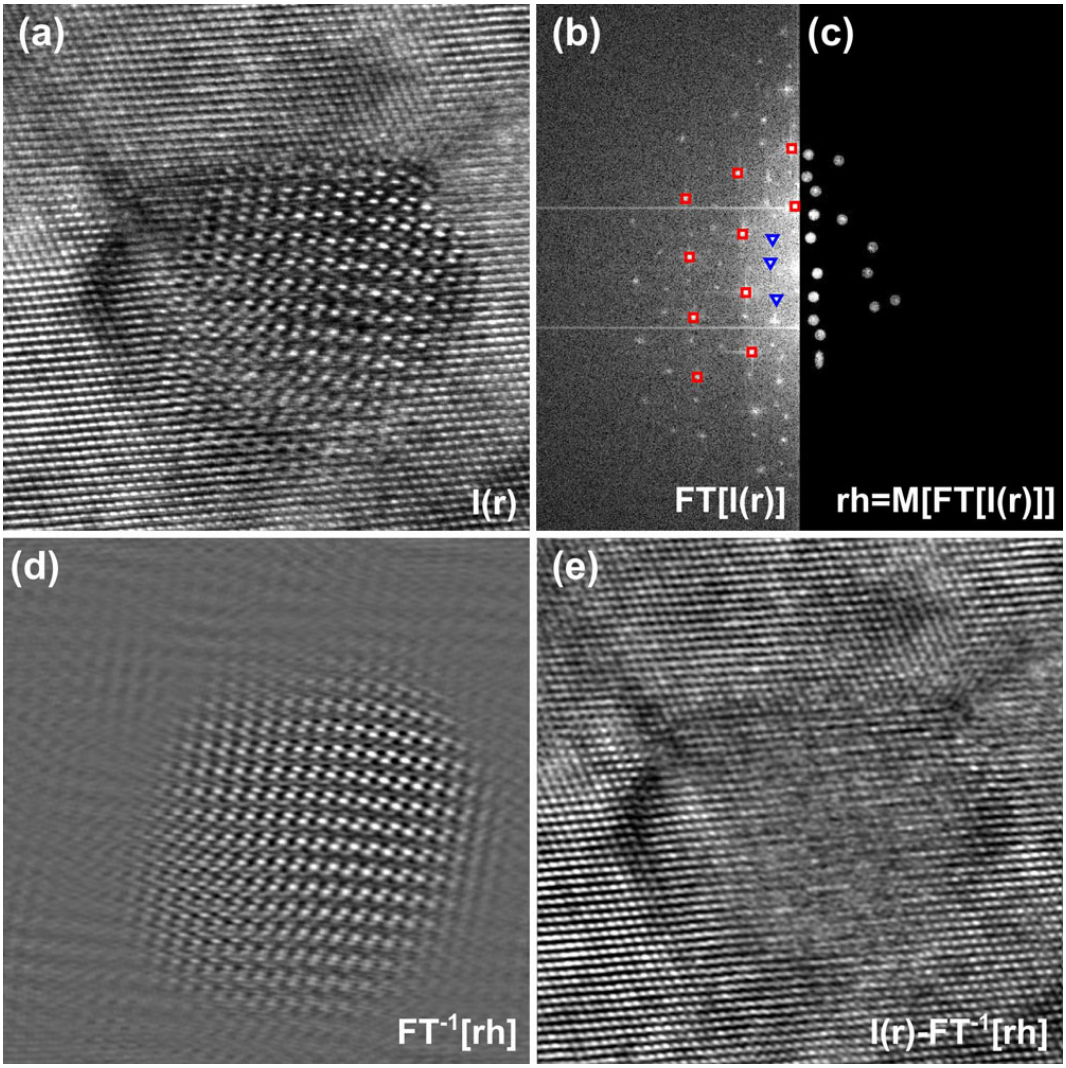


Figure 4.7: Procedures of HRTEM image analysis of embedded rh-Bi clusters. (a) acquired HRTEM image; (b) (magnified) fast Fourier transform of image (a) show "extra" Fourier components (marked by blue triangles) other than those from zinc-blend GaAs (marked by red boxes); (c) mask the "extra" Fourier components; (d) inverse Fourier transform of (c) show the lattice image of the embedded Bi cluster; (e) the lattice image of (a) with (d) subtracted reveal the degree of overlapping matrix effect.

will further complicate the analysis of incoherent clusters. Image process is necessary to remove the contribution of overlapping matrix lattice.

The image processing is proceeded in the way shown in Fig. 4.7 and explained as follows. An acquired HRTEM image is shown in Fig. 4.7(a). Complex lattice image reveals the cluster position. The fast Fourier transform of this image is presented in Fig. 4.7(b). Only half of the FFT image is shown. Besides the Fourier components from zincblende GaAs, some "extra" Fourier components are also revealed. After masking and filtering the "extra" Fourier components and then perform an inverse fast Fourier transform, the resulted lattice image is shown in Fig. 4.7(d). This filtered lattice image is mainly contributed from the rh-Bi cluster. The result of subtracting the acquired raw image by the filtered lattice image, as is shown in Fig. 4.7(e), can be regarded as an measurement on the

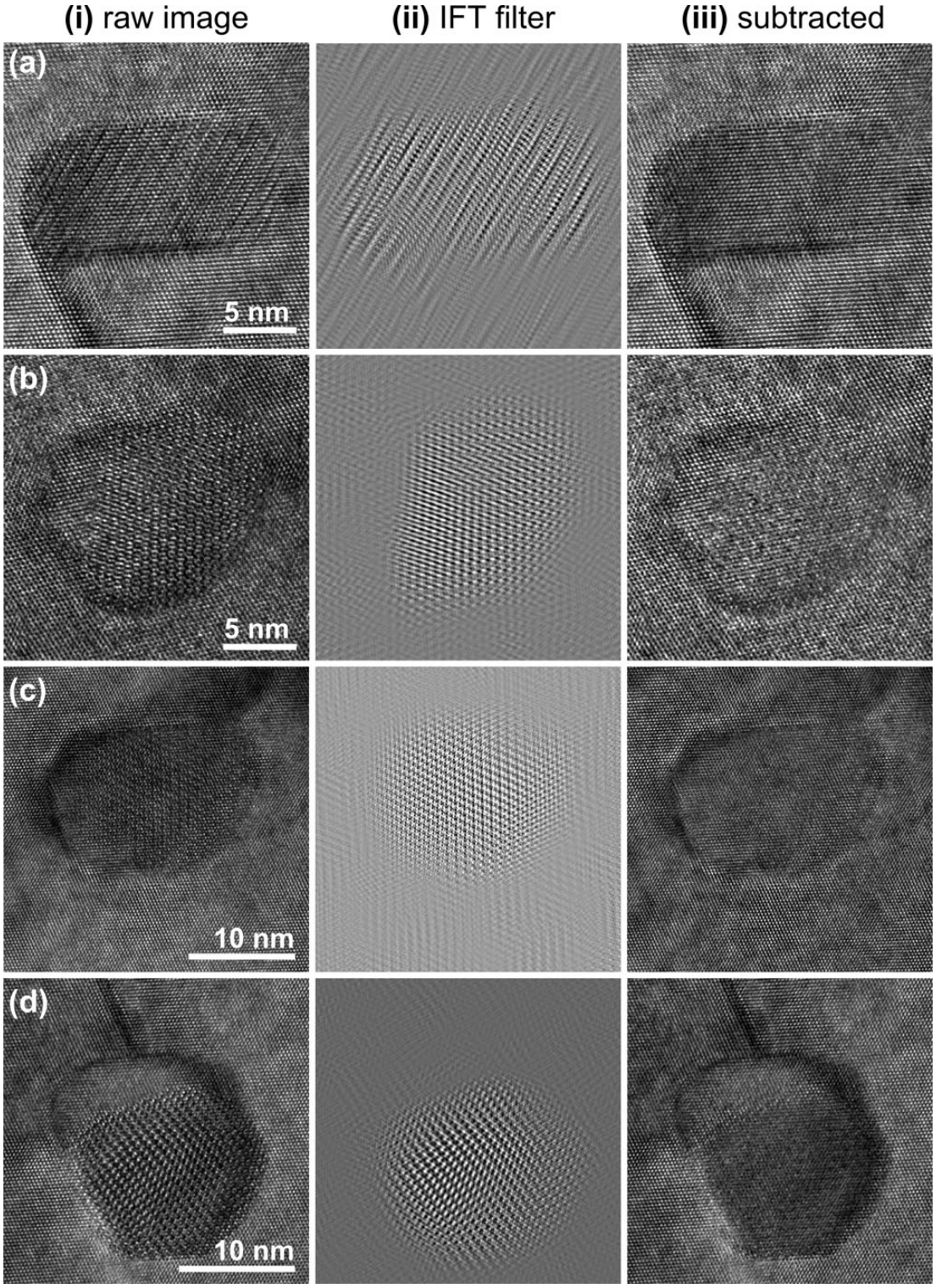


Figure 4.8: HRTEM images of embedded rh-Bi clusters and corresponding processed images. All images are acquired along GaAs [110] zone axis and aligned in crystallographic directions. The clusters show four different sets of patterns suggesting the four variant of equivalent crystal orientations.

degree of the overlapping matrix effect. This can be understood following the electron scattering theory with multi-slice method, as described in page 30.

This HRTEM image processing is applied to more rh-Bi clusters. The results are shown in Fig. 4.8. In the image, rows are different HRTEM image of rh-Bi clusters; columns are the raw acquired images and processed images. The clusters appeared in four different sets of lattice patterns. This will be discussed in more details in the corresponding case study in chapter 5.2.

4.2.4. Detection of embedded nano-clusters

When the clusters are really small, it is important to choose the informatioin that can produce the strongest contrast to reveal the existence of clusters. On the other hand, it is also important to improve experimental conditions to enhance the contrast from clusters.

Contrast from difference in structure factor Contrast can also arise from the difference in structure factor. An example of InAs clusters embedded in Si matrix is given in Fig. 4.9. The image was obtained using the (002) reflection. As discussed in section 3.2.2, the structure factor of Si for (002) reflcetion equals zero (kinematically forbidden). However, it has a finite value for zincblende InAs. Therefore, the InAs clusters embedded in Si can be revealed by the contrast from the difference in structure factor.

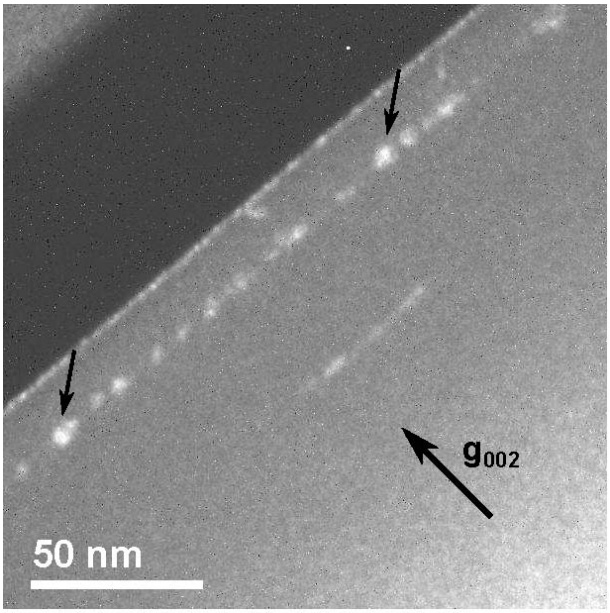


Figure 4.9: Diffraction contrast from difference in structure factor. InAs clusters embedded in Si matrix is imaged with diffraction vector $\mathbf{g}_{002}$. (002) reflection in Si is kinematically forbidden but is allowed in InAs.

However, in most cases for nano-clusters, because of the complex internal and external structures, the contrast is produced not only from one of the factors discussed above, but a combination of many factors. This make the quantitative interpretation of contrast more difficult.

Z-contrast dependence on detector inner semi-angle HAADF images, also known as Z-contrast images, could produce a contrast proportional to the atomic Z number. The

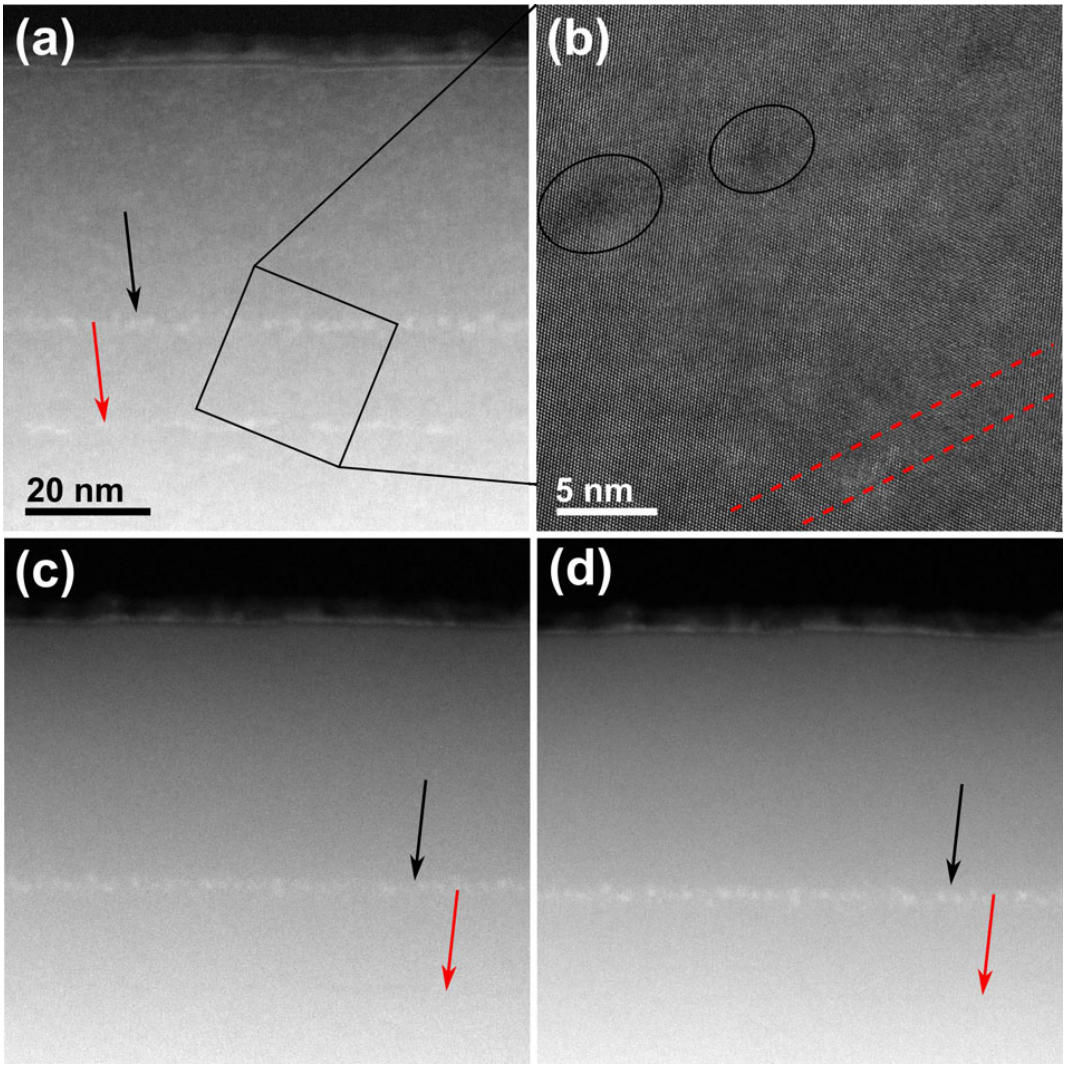


Figure 4.10: HAADF-STEM Z-contrast images reveal the existence of InAs clusters in Si matrix. (a) ADF image with detector inner collecting semi-angle 25 mrad. Both the Si homo-epitaxial buffer interface impurities (red arrow) and embedded InAs clusters appears bright. (b) HRTEM of the marked position in (a) with ambiguous contrast from clusters. (c),(d) HAADF image with detector inner collecting semi-angle 60 and 70 mrad, respectively. Only the InAs clusters with higher Z-number appear bright.

contrast is strongly depended on the inner collection semi-angle. An example of embedded InAs clusters in Si is illustrated in Fig. 4.10. The imaging region in Fig. 4.10(a), (c) and (d) are identical but acquired with different detector inner collection semi-angle (adjusted by changing the camera length).

With HRTEM [Fig. 4.10(b)], the contrast from the expected InAs clusters is quite weak, and ambiguous. When imaged with semi-angle of 25 mrad [ADF mode, Fig. 4.10(a)], the Si homoepitaxial interface, where some impurities exist, and the layer containing the InAs clusters appear both bright. With detector collection semi-angle of 60 mrad and 70 mrad, [in Fig. 4.10(c) and (d), respectively] only the expected InAs cluster appear bright. This shows that the contrast is stronger with higher inner collection semi-angle.

4.3. Structure modeling

After the general discussion on the detection of large-sized nano-clusters of different degree of coherency, we come to the main topic about Gd-doped GaN to think about how to detect small Gd-containing cluster of few tens of Gd atoms down to the possibility to detect single Gd atoms in GaN matrix. We start here with structure modeling.

As already mentioned earlier, quantitative interpretation of image contrast and thus the analysis of clusters requires contrast simulations that are based on appropriate structure models. Modeling of nanostructure can be achieved conventionally using continuum elastic approach within analytic field function description (e.g., the case study in Chapter 5.1) or with finite element method. The continuum elastic approach is only a good approximation when the embedded object is large enough (at least few tens of thousands of atoms).

However, when the object is on the atomic scale and/or the bulk elastic strain tensor are not available (e.g., the study of GdN clusters in Chapter 6), the continuum approximation is not accessible. Instead, discrete atomic structure models and structure relaxation based on more fundamental inter atomic interactions should be considered. Such description of inter atomic interaction includes, of different level of accuracy, the empirical force field methods, semi-empirical quantum mechanical methods, and *ab inito* methods. The methods follow a sequence of increase in the accuracy, but also corresponds to a increasingly higher requirement in the computation demands. The optimum geometry of atomic sized embedded objects, in case of GdN clusters in GaN (in Chapter 6), is obtained through either empirical valence force field method or first-principal method based on density function theory. The basic principle of these methods are summarized in the following. And a comparison of the displacement map measured from the results of either method is presented.

In this section, we first briefly summarize the principle of each method used in this thesis (i.e., modeling based on either VFF method or DFT calculations). Then we compare the results from both approach with the embedded GdN cluster in GaN case.

4.3.1. Valence force field (VFF) method

The search for the equilibrium crystal structure (relaxed structure) is therefore equivalent to find the minimum on the potential energy surface. The solution to find the potential energy surface is also quite complex. The force field [2] based method uses instead the empirical fit to the potential energy surface as parameterized power series function (similar to eq. 4.5). The utilized fitting function as well as the parameters for the fitting are expressed as data sheet which is commonly called a force field. The quantum mechanical effects are often insignificant to account for the motion of the relatively heavy nuclei. Its motion can therefore be expressed by the Newton's equation of motion on the potential energy surface.

The fitting function in force field to describe potential energy of a system can be expressed as a sum of valence, non-bond and cross-term interaction:

$$E_{total} = E_{valence} + E_{nonbond} + E_{crossterm}.$$

In semiconductor materials, the force between atoms are mainly Coulomb interaction, which are attributed to the valence term in total potential energy. The energy of valence

[2] "Force field" account both for the force field method and the force field parameterized functions and data.

interaction is mostly accounted for by bond stretch (E_{bond}), valence angle (E_{angle}) and dihedral angle torsion ($E_{torsion}$)[116]:

$$\begin{aligned} E_{valence} &= E_{bond} + E_{angle} + E_{torsion} \\ &= \sum_{bond}^{i} \frac{1}{2} K_i (R - R_0)^2 + \sum_{angle}^{j} \frac{1}{2} C_j (\cos\theta - \cos\theta_0)^2 + \sum_{torsion}^{n} \frac{1}{2} D_{\theta,n} [1 - d\cos(n\phi)]. \end{aligned} \quad (4.5)$$

In addition to the functional form of the potentials, a force field defines a set of parameters for each type of atom. For example, a force field would include distinct parameters for an nitrogen atom in octahedral bonding environment and tetraheral bonding environment. The typical parameter set includes values for atomic mass, van der Waals radius, and partial charge for individual atoms, and equilibrium values of bond lengths, bond angles, and dihedral angles for pairs, triplets, and quadruplets of bonded atoms, and values corresponding to the effective spring constant for each potential.

Most force field covers only limited atom species, which are empirically available, like C, O, N, Ga, Si, etc. Nevertheless, coverage of all elements of the periodic table is provided in the universal force field (UFF)[117], where part of the force field parameters are not fitted empirical but extrapolated. This is the only force field available that covers the rare-earth elements so far. It is therefore used in this thesis, as implemented in the force field module in Cerius2^{TM} package.

4.3.2. Density-functional theory (DFT)

Beside the empirical fitting, the ground state chemical and physical properties of the crystalline material can be calculated without the need of any *a priori* information. The straightforward but (almost) impossible way would be to solve the corresponding many-electron Schrödinger equation. A feasible way is to consider simpler functions, which still contain as much useful information, as possible. One possibility to calculate the electronic structure of the matter, is *ab initio* methods like the Hartree-Fock[118] approach. The Hartree-Fock method describes the stationary electronic state by a many-body wave function which is the solution of the Schrödinger equation. The nuclei of the atoms are treated as coordinated objects generating a static potential in which the electrons are moving. The complicated many-body wave function results in huge computational effort and makes the method unsuitable for large and complex systems. The main point of the density functional theory (DFT)[119,120] is to replace the many-body wave function of the electrons with a simple treatable electron density. In the Kohn-Sham approach the non-interacting electrons are moving in an protective potential of the nuclear charges. The ground state energy is found by minimizing the energy expression having the terms of the interaction with the external potential, the self interaction of the charge density and inter nuclear repulsion, which all depend only on the arbitrary electron density function.

4.3.3. Displacement map from VFF and compare to results from DFT

Atomic model The comparison of the structure relaxation results from VFF and DFT are carried out with the atomic model as depicted in Fig. 4.11. A thin sheet of monolayer GaN supercell which consists of 640–2000 atom along GaN [1$\bar{2}$10] direction is first constructed. Next, consecutive center Ga atoms are substituted by Gd atoms or Mn atoms. The supercells of various width of substitutional atoms are used as input structure for relaxation

using either VFF or DFT. The specific results of the models are referred to according the number of substitution atoms, like "x Gd" model.

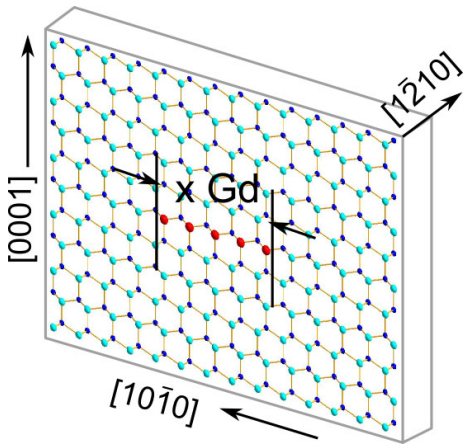


Figure 4.11: The model supercell used as input structure for the relaxation by VFF and DFT. This figure show a "5 Gd" model.

Evaluation of displacement map Starting from a relaxed structural model, the HRTEM image is first simulated with a thickness of 4 nm and a defocus value $\Delta f = -40.5$ nm (and other parameters represent the experimental condition of JEM-3010). From the resulting HRTEM image, displacement field map is then constructed by GPA, as described in Section 3.4.1 (referred to as HRTEM simulation–GPA extraction routine).

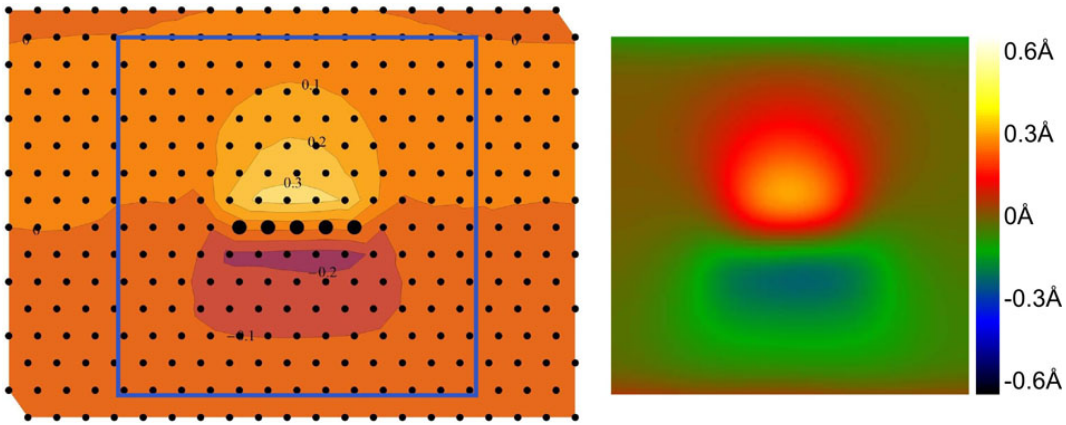


Figure 4.12: Displacement field map component along GaN [0001] direction extracted directly from the atom position in the relaxed structure (left) and from the "HRTEM simulation and GPA extraction" process (right). (the model shown here is the "5 Gd" model)

The displacement map can be also calculated directly from the relaxed atom coordinates. It is verified that the displacement map extracted from the "HRTEM simulation–GPA extraction" routine is in good agreement with the other one calculated directly from the the atom position in the relaxed structure. The results for the "5 Gd" model are show in Fig. 4.12. From this comparison, it can be seen that there is a good match of the "shape" of the displacement field and the maximum displacement value between the measured displacement field with the absolute displacement, except for the region of about 1 layer

close to the Gd atom columns. This is because the GPA measurement is a based on the Fourier analysis, the region close to the Gd atom column have an average effect from its adjacent Ga layer therefore introduced artifacts to the measurement. Therefore, the resulting maps of displacement field from modeled structures and experiments are directly comparable.

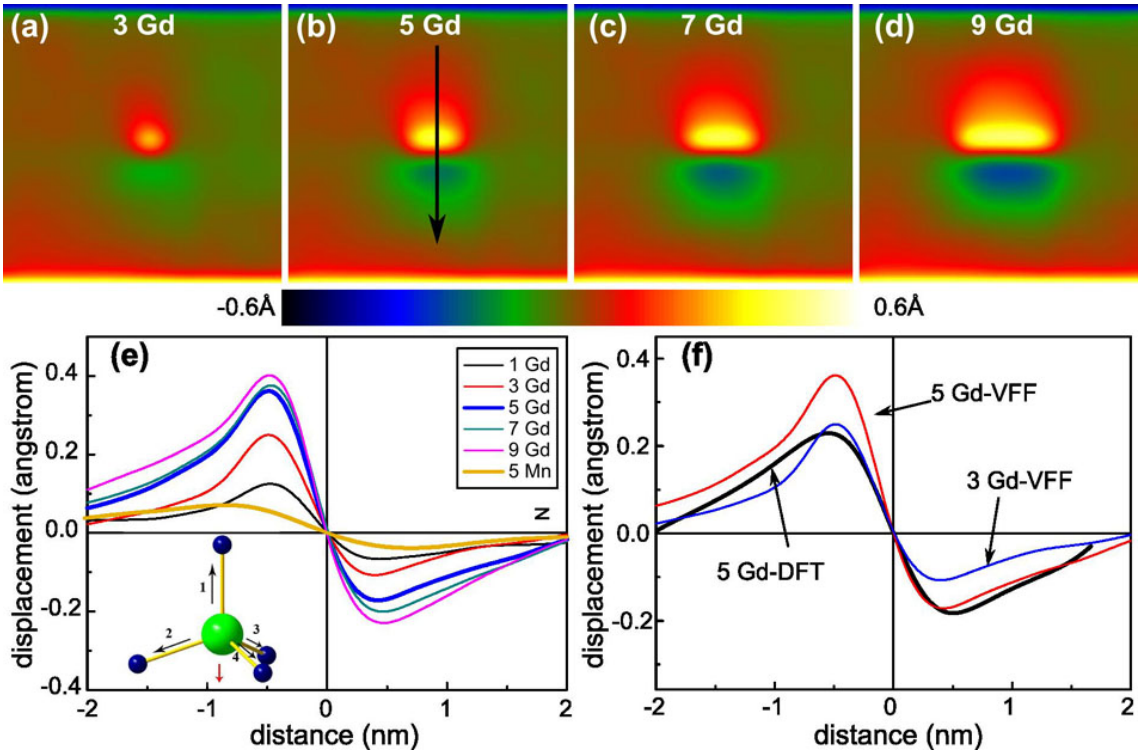


Figure 4.13: Displacement field mapping component along [0001] direction extracted from simulated structures relaxed via VFF method. The models are monolayer platelets with diameter of (a) 3 Gd atoms, (b) 5 Gd atoms, (c) 7 Gd atoms and (d) 9 Gd atoms. The line profiles in (e) compare the displacement value from different models, in (d) compare the result relaxed via DFT.

Displacement map from VFF Figure 4.13 show the displacement field mapping component along [0001] direction and corresponding line profiles extracted from simulated structure relaxed via VFF method. Strong lattice distortions along GaN [0001] direction are visible, similar to that measured in experiments. It is clearly seen from Fig. 4.13(a)-(d) that the lateral expansion of the distorted lattice increase with more Gd atoms placed in the model (1 to 9 Gd atoms), as expected.

Line profile cross the center of cluster along the [0001] direction as indicated by the arrow line in Fig. 4.13(b) is presented in Fig. 4.13(e) to describe the quantitative difference. It is evidenced that the maximum displacement also increases with increased adjacent Gd atoms in the models. Moreover, this increase is quite dramatic from 1 Gd model to 5 Gd model: from about 0.1Å to about 0.4Å; and the displacements keep almost constants for 5 Gd model to 9 Gd model. In addition, a model, in which 5 adjacent Mn atoms are used instead of Gd is also constructed to verify if the displacements in the relaxed structure are reasonable. The extracted maximum displacement is, as expected, only 0.08Å [cf. Fig. 4.13(e)], which is much smaller than that resulted from Gd case. With a closer exam-

ine on these line profiles, it is also noticeable that the displacement are quite asymmetric, i.e. the absolute displacement value are not the same at equal distance from the center of cluster. This is attributed to the artifact from the VFF method structure relaxation. As mentioned in the previous sub-section, in the universal force-field parameter database bond length, angle and torsion data are fitted from empirical data or extrapolated. While GdN in nature is in rock-salt structure, it is forced to take wurtzite structure in the models. Therefore, the contribution proportion of bond length, angle and torsion may not work properly. When the bond length contribution dominate, the four bonds connecting Gd and N will tend to be equal. Consequently, the displacement projected in $[1\bar{2}10]$ direction will be asymmetric.

Compare displacement from VFF and DFT Figure 4.13(f) shows the comparison the displacement line profile extracted from structures relaxed via VFF method and based on DFT. The asymmetric artifact from VFF method is completely removed in the results from DFT. The displacement value is slightly smaller obtained from DFT. Nevertheless, they are in a similar range if compared to the substitutional Mn case. It is therefore concluded that structure relaxation via VFF method can provide qualitative information for model matching. Structure relaxed based on DFT should able to provide more accurate structure information for model matching. However, due to the much higher computation demands, the super-cell size is limited only to few hundreds of atoms in current computation center.

4.4. Detectability of Gd in GaN

In this section, we discuss the detectability of Gd atoms in GaN by means of HRTEM phase contrast, strain contrast and Z-contrast with simulated results based on models.

4.4.1. HRTEM phase contrast

HRTEM lattice images provide the highest spatial resolution. Therefore it might resolve the presence of Gd atoms in GaN lattice. We will examine it with simulated results. Detection of Gd atoms in GaN with HRTEM phase contrast use the information that Gd atoms cause a different phase with respect normal Ga atoms to the electrons or, a perturbation signal to electrons.

Extensive HRTEM contrast simulations were performed to study the effect of introducing Gd atom(s) in the GaN lattice on the image contrast features for a quantitative HRTEM image interpretation. The HRTEM simulations presented here are based on rigid body Gd_{Ga} models, i.e. the wurtzite lattice constant take the value of bulk GaN, and Gd atoms are all placed at Ga substitution site.

Single embedded Gd substitution atom

Shown in Fig. 4.14(a) is a simulated HRTEM image with one center Ga atom column replaced by Gd atoms, simulation parameters mentioned in the figure caption. Comparing the center region of the image to the boarder region where no Gd substituted Ga atom column, it is hardly seen any qualitative different concerning the existence of the impurity Gd atom column, i.e. the spots does not change from tunnel position to atom column position, or vice versa.

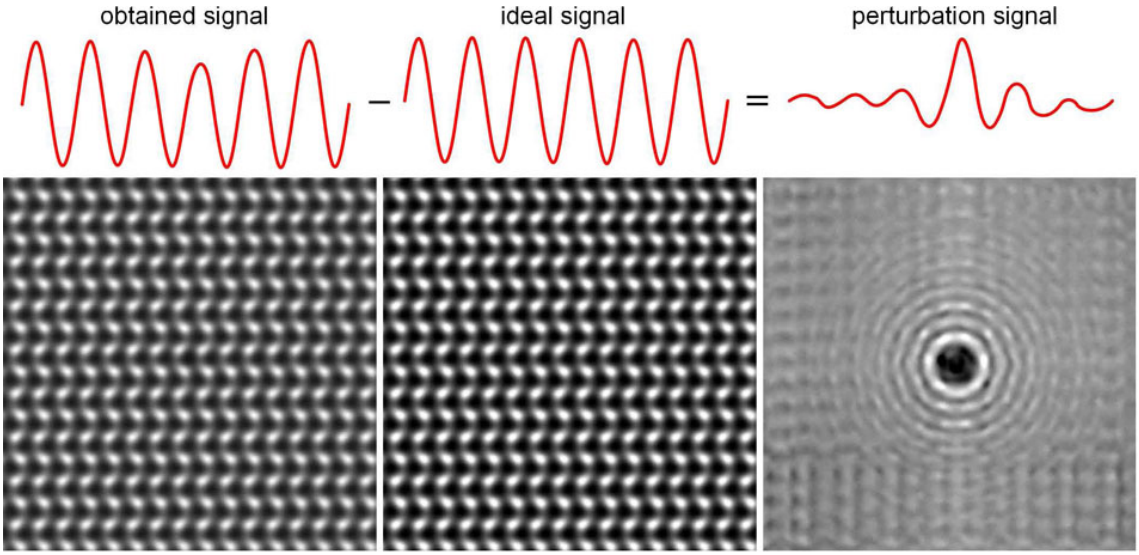


Figure 4.14: Schematic illustration of perturbation signal extraction. (a) Simulated results of $[1\bar{2}10]$ projected GaN lattice model with one center Ga atom column replaced by Gd atoms; (b) the expected ideal GaN lattice iamge from (a) by performing a radical Wiener filter; (c) subtracted perturbation signal $(a - b)$, highlighting the feature like Frensnel propagation of electron wave from the Gd impurity atom column perturbation. Simulation parameters are thickness $t = 2.5$ nm and defocus $f = -30$ nm. Image intensity are normalized for better visulization.

Actually, it will be shown later that in the thickness and defocus range close to the experimental conditions this results always hold. Quantitatively, the perturbation from the replaced Gd atom column to the electron wave can be visualized by subtracting the ideal signal from pure GaN lattice (cf. scheme in Fig.4.14). This can be simply obtained by another simulation of pure GaN super-cell with identical parameters. However concerning the feasibility of experimental process, this is technically obtained by performing a Wiener filter on the raw image, which treat the perturbation signal as part of the "noise". The result is shown in Fig. 4.14(b). Subtracting the period lattice fringes contribution, the filtered image is presented in Fig. 4.14(c). The image contrast is quite similar to a cylindrical propagated wave centred at the Gd atom column. This contrast can be mainly contributed by the Gd atoms with different scattering factor. This could be qualitatively interpreted as follows.

According to the direct space multi-slice description (eq. 3.20), the exit electron wave is the multiply of passing through a thin slice and then propagate in the slice plane. Since the multi-slice scheme is an iterative process, the discussion below will simply focus on one slice. The exit wave is then:

$$\psi(x, y, z_0) = \exp(\overline{\Delta}\epsilon) \cdot \exp[\mathrm{i}\sigma V_p(x, y)] \cdot \psi(x, y, 0), \tag{4.6}$$

where V_p is the projected potential of the simulated super-cell, expressed as

$$V_p(x, y) = V_p^{GaN}(x, y) + V_p^{Gd}(x_1, y_1). \tag{4.7}$$

The effects of the substitute Gd atom, whit a different scattering factor, on the projected potential of a perfect GaN lattice is treated as perturbation at specific position (x_1, y_1).

After multiply the propagation term,

$$\psi'(x,y,z_0) = \exp(\overline{\Delta}\epsilon) \cdot \exp[\mathrm{i}\sigma V_p^{Gd}(x_1,y_1)] \cdot \psi(x,y,0), \tag{4.8}$$

with the propagator (i.e. eq. 3.19),

$$e^{\overline{\Delta}\epsilon} \mid \psi\rangle = \mathcal{F}^{-1}\left[\exp(\mathrm{i}\pi k\frac{R}{\epsilon})\right] \otimes \psi = \mathcal{P}_\epsilon(R) \otimes \psi. \tag{4.9}$$

The back Fourier transform of the plane wave $\exp(\mathrm{i}\pi kR/\epsilon)$ is a sphere wave. Within the small slice thickness ϵ range, the perturbation wave is then approximated as a cylindrical wave centered at (x_1, y_1). The exit wave will then convolute with the microscope transfer function, and final form the image on the screen. These following transfer process alters mainly the phase of the electron wave depending on spatial position but not the amplitude, which means it does not substantially change the nature of the cylindrical propagation wave. Therefore, the information could be reflected in the final image intensity. The process of subtracting the periodic information is similar to pattern recognition and subtraction. When the main contribution from GaN lattice is filtered [cf. Fig. 4.14(b)], the subtracted image intensity [cf. Fig. 4.14(c)] is mainly contributed from Gd perturbation to the electron wave that formed the image.

However, it is also worth to mention that the perturbation image signal is only about 2–3% of the raw image signal. In the simulation, the specimen is assumed to be homogeneous thin 2.5 nm, i.e. 8 atoms in each atomic column. Simulation results (not shown) also reveal if Gd atom column is buried in the matrix (matrix effect), the perturbation signal will be smaller. Other effects that have to take into account in experimental image acquisition, like thickness inhomogeneity, surface amorphous layer scattering and CCD camera thermal noise due to short exposure time will all work as counterpart to reduce the perturbation signal.

Embedded platelet GdN cluster

When there are more Gd atoms placed in nearest neighbor Gd_{Ga} position, the perturbation is expected to be stronger. The results of the model of 5 Gd atom column arranged adjacent to form a ribbon are shown in Fig. 4.15. However, in the thickness and defocus range, the results show still no qualitative difference. In addition, Gd atom is larger than Ga, the structure relaxation should also taken into account, the results reveal still the perturbation will not produce any qualitative effect on the contrast feature, the details will be provide in the next sub-section.

It is therefore concluded that the introduction of Gd atoms in Ga substitution site with either single atom or embedded platelet geometry could not produce strong perturbation to the phase contrast HRTEM image that could be detected in experimental setup.

4.4.2. Z-contrast

Gd and Ga atoms has a Z ratio of $Z_{Gd}/Z_{Ga} = 64/31 = 2.06$. The distribution of Gd atoms in GaN may also be detected by the Z-contrast. Z-contrast imaging uses electrons that are scattered to high angles. These electrons have strong dynamic scattering phenomena. Matrix effect on the reduction of image contrast will be stronger than other imaging modes. When the crystle sample is oriented on zone axis, the channeling of the scattered

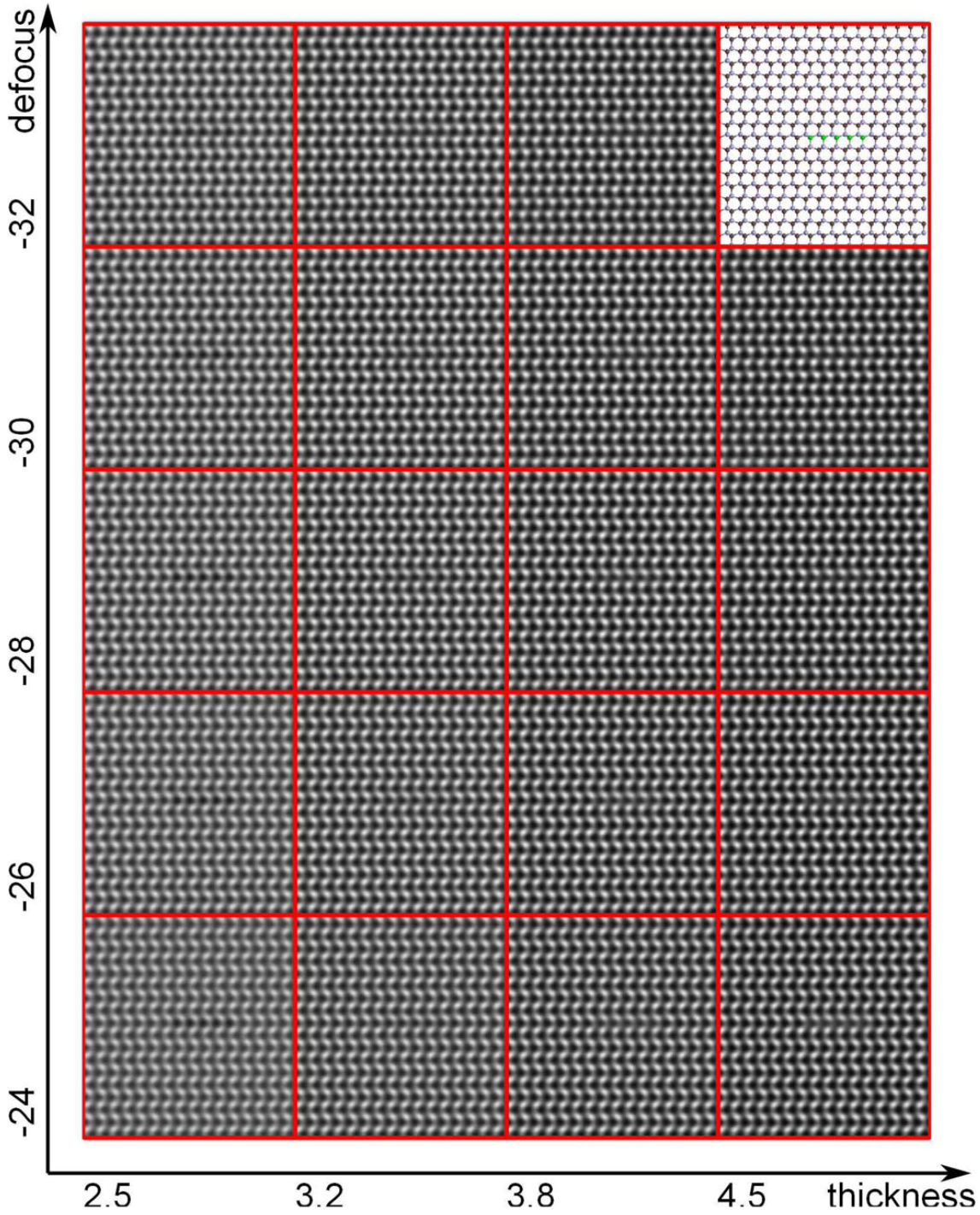


Figure 4.15: Thickness-defocus map of simulated HRTEM contrast for the rigid body model shown on topright corner, where 5 adjacent Ga atoms are simply replaced by Gd atoms. These results show that in the defocus and thickness range of the experiment condition, there is no quantitative difference when the Gd atoms are introduced on Ga substitution site, i.e. brighter spots do not shift from tunnel position to atom column position, or vice versa.

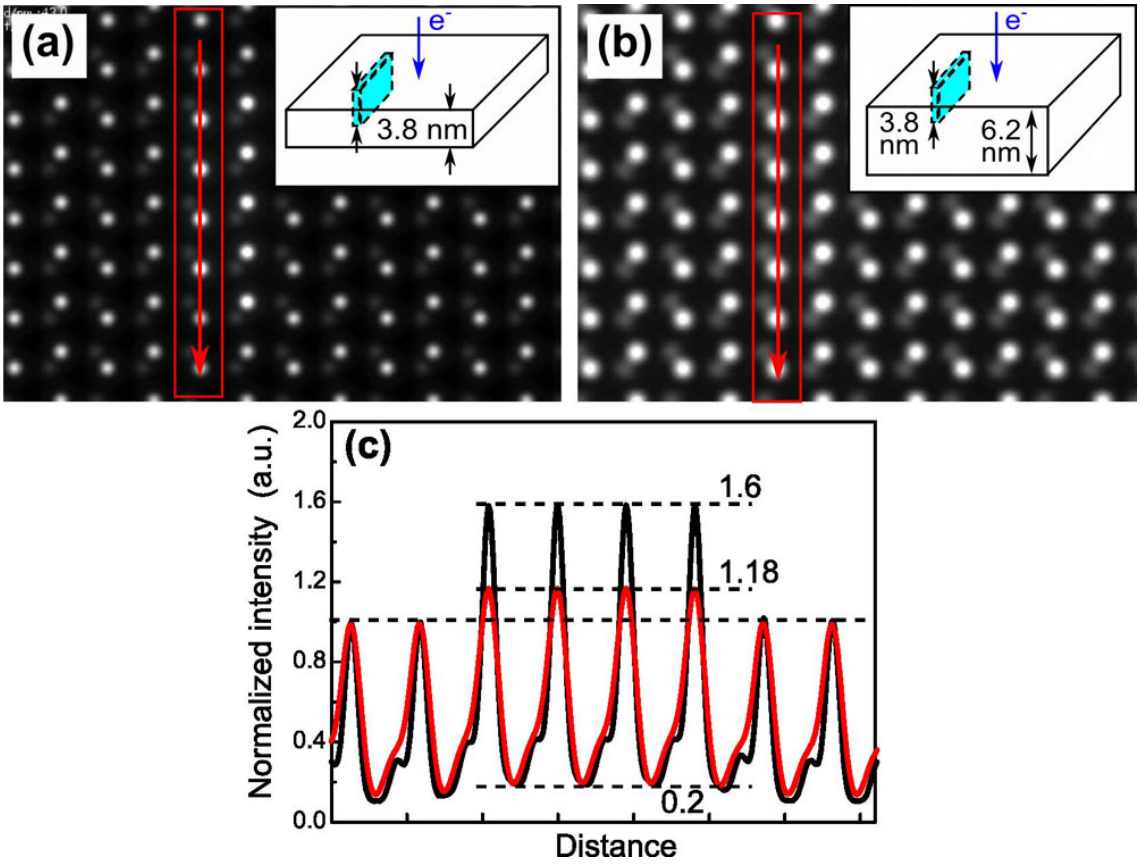


Figure 4.16: Simulated HAADF image contrast of bilayer GdN cluster model in different geometry. (a) The 3.8 nm diameter cluster penetrates through the top and bottom surface of specimen (without matrix effect); (b) the cluster is followed by 2.4 nm GaN lattice along the electron beam direction. (c) the simulated intensity profile from the images as marked in (a) and (b).

electron down along the atom column is strong. Even small amount of matrix crystal is overlapping the embedded clusters, the image contrast will result in a strong decrease.

HAADF image contrast simulation based on multi-slice method is performed to illustrate the strong reduction of signal enhancement by the matrix effect (channelling effect). The high-angle scattering electrons are simulated using bicubic potential interpolation with the frozen phonon approximation[85] at 200 K, as integrated in the recent JEMS[94] program. Simulation parameters are chosen to reasonable agree to our experimental setup (coherent uniform probe forming function with semi-convergence angle of 14 mrad, $C_s = 0.5$ mm, $\Delta f = -43$ nm, detector collection semi-angle 80 – 210 mrad). The electron probe is then scanning through a bare 3.8 nm thick GdN cluster contained foil in the cross section $[1\bar{2}10]$ zone-axis scheme [Fig. 4.16(a)]; and through a foil containing a 3.8 nm GdN layer and a 2.4 nm GaN layer, demonstrating the matrix effect on the image intensity [Fig. 4.16(b)].

The simulated results are shown in Fig. 4.16. The intensity profile (normalized to GaN atom column) across the projected GdN cluster is shown in Fig. 4.16(c). Without GaN matrix, the image intensity at Gd atom column position is quite strong. The signal enhancement is about $(1.6 - 1.0)/(1 - 0.2) = 75\%$. However, after introducing only 2.4 nm GaN lattice in the atom column results in a signal enhancement reduction sig-

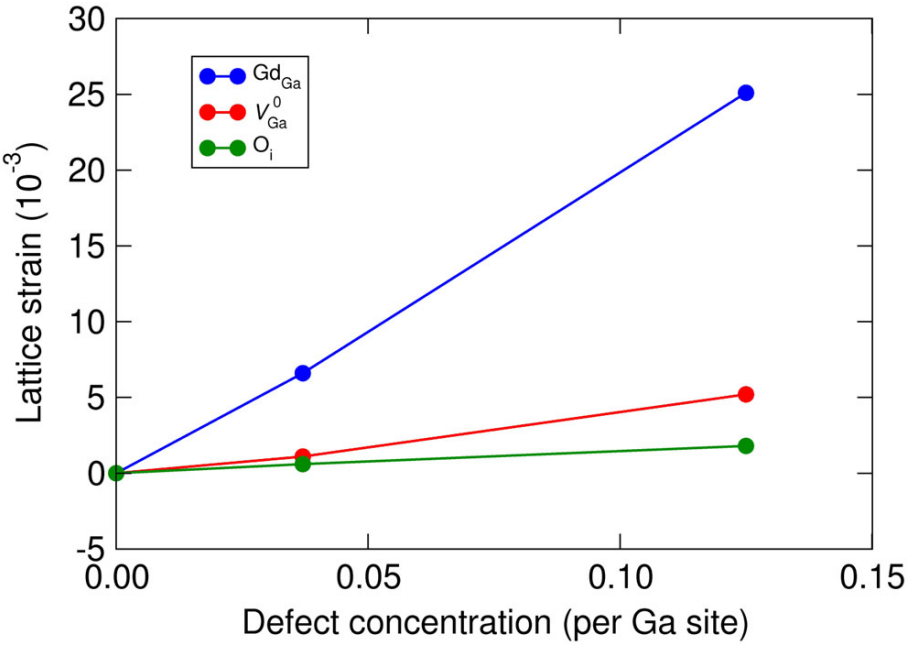


Figure 4.17: DFT calculated lattice strain as function of concentration of different disperse distributed defects. [Courtesy DFT calculations by Steven C. Erwin @ Naval Research Laboratory in Washington D.C.]

nificantly to $(1.18-1.0)/(1.0-0.2) = 22.5\%$, less than one third of the original signal enhancement.

In practise, it will be beneficial to tilt the specimen slightly off the zone axis to reduce the channeling of electrons and thus enhance the contrast from embedded clusters.

4.4.3. Strain contrast

Gd atoms are much larger than Ga atoms, the incorporation of Gd atoms in GaN lattice will induce a strong lattice distortion (strain field). Figure 4.17 shows the DFT calculated results for the strain created in bulk GaN by substitutional Gd, Ga vacancies, and oxygen interstitials. The defects are assumed to be dispersed, with concentrations given on the x-axis. The result is that only Gd can create any significant strain (about 2% strain for 10% Gd content – far higher concentration than in the experiments). The strain caused by Ga vacancies or O interstitials is an order of magnitude smaller.

Hence, the existence of GdN clusters may create strong local lattice strain that can may detected by the strain contrast images. The simulated strain contrast image of an ideal GdN cluster model will be presented here and the detectability will be discussed.

The geometry and set-up of the simulation are schematically shown in Fig. 4.18(a) and (b). Assume a GdN cube cluster of 1 nm in length, where volume corresponds to about 40 Gd atoms taking the Ga substitutional site that follow the wurtzite GaN lattice symmetry. We assume the displacement of GaN matrix lattice $R(\mathrm{r})$ induced by GdN cluster follow an exponential decay with maximum displacement equals half of the difference in bond length between Gd-N (2.49 Å) in GdN and Ga-N (1.94 Å) in GaN, i.e.,

$$\mathrm{u}(x) = \begin{cases} u_{\max}\exp(-x) & x > 0 \\ -u_{\max}\exp(x) & x < 0 \end{cases} \quad \text{with} \quad u_{\max} = \frac{1}{2}(2.49 - 1.94) = 0.27 \tag{4.10}$$

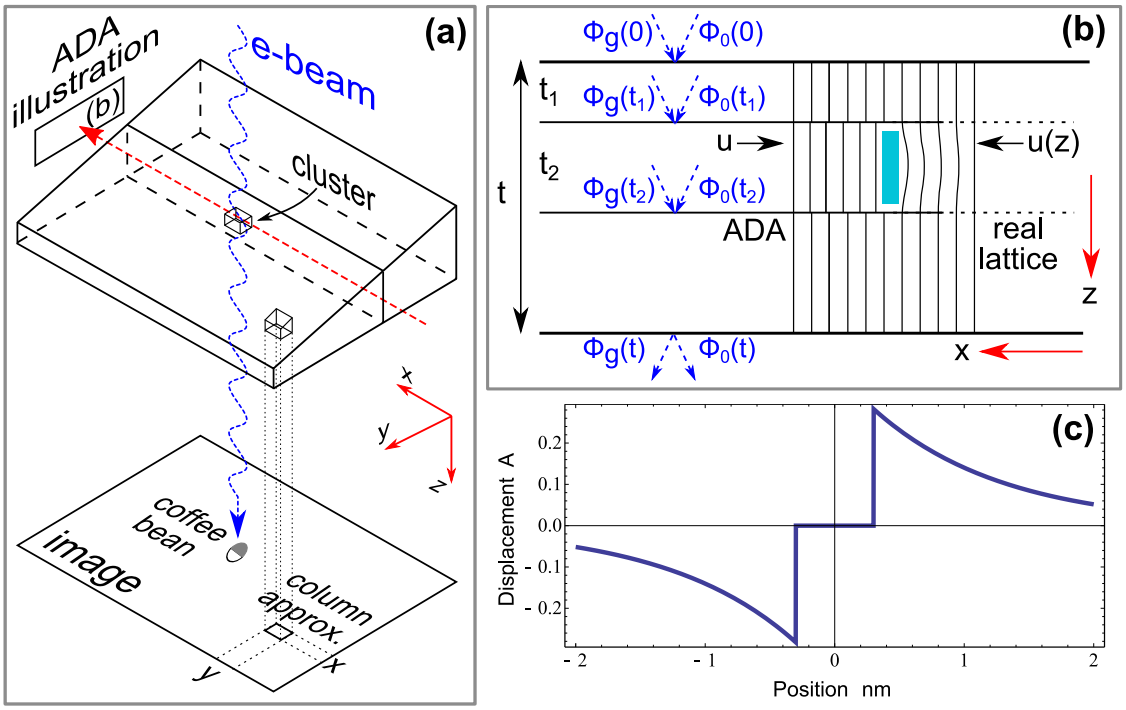


Figure 4.18: Scheme of DF strain contrast simulation. (a) Geometry of the sample. (b) In the simulation, we uses the abrupt displacement approximation (ADA), details see text. (c) the displacement profile along x direction.

This displacement profile is shown in Fig. 4.18(c). The middle part of zero value accounts for the part of cluster.

In real crystals, the distorted lattice is varying continuously over the thickness, therefore the displacement field is a depth-dependent function $u(z)$, illustrated at the right side of the blue cluster in Figure 4.18(b). Miller *et al.*[102] have demonstrated that DF contrast simulation with abrupt displacement approximation (ADA), as illustrated on the left side of the blue cluster in Figure 4.18(b), which assumes a constant displacement field over the depth, is able to reproduce contrast features of strained objects. Thus, the specimen can be considered as consisting of three parts along the electron penetrating trajectory. The middle part, having a constant displacement field, is buried within the perfect crystal. As shown in Fig. 4.18(b), the total specimen thickness is denoted as t, and the lattice slice with Gd cluster of width $t_2 = 1$ nm is buried in the matrix starting at a depth of t_1. Because of the small size of the cluster, its contribution to electron diffraction is also neglected. Then an analytical solution of ϕ_g from the two beam Howie-Whelan equation (eq.3.26 discussed in page 32) can be obtained as

$$\phi_g = \frac{i\pi}{\xi_g}\left(\int_0^{t_1} e^{-2\pi i s z} dz + e^{-2\pi i \mathbf{g}\cdot\mathbf{u}(x)} \int_{t_1}^{t_1+t_2} e^{-2\pi i s z} dz + \int_{t_1+t_2}^{t} e^{-2\pi i s z} dz \right). \tag{4.11}$$

The final image intensity $I_g = \phi_g \phi_g^*$ is thereby a function of t, t_1, t_2, s, ξ_g and $\mathbf{u}(x)$. We consider a sample thickness of 50 nm under GaN 0002 system row condition. So the extinction distance $\xi_g = 50$ nm (refer Fig. 3.8 and the discussions in page 34). The excitation error s is an experimental variable. We set $s = -0.0175$. The image intensity is only function of cluster burried depth $I_g(t_1)$.

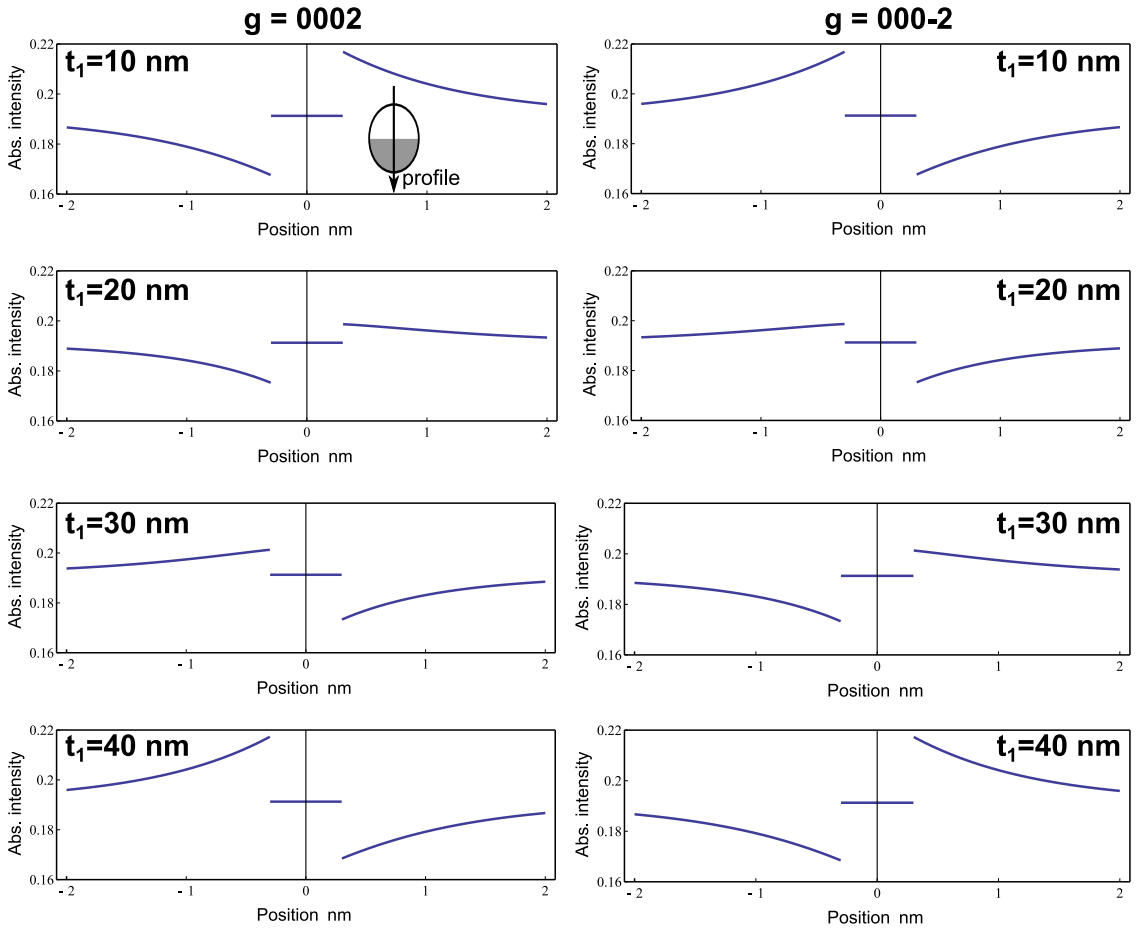


Figure 4.19: Simulated DF intensity profile of GdN cluster model. Left and right panels are calculated with $\mathbf{g} = 0002$ and $\mathbf{g} = 000\bar{2}$ conditions. The cluster buried depth are indicated in each intensity plot. Insert in the top left profile illustrate how the intensity profile is obtained over the "coffee bean" contrast feature.

The simulated results are shown in Fig. 4.19. We draw attention to two main features of the intensity plots: (1) the intensity difference at opposite side of the cluster show a contrast up to (in $t_1 = 10$ and 40 nm) about $(0.22 - 0.17)/0.17 = 30$ %; (2) the contrast inversion by using a inverted **g**-vector. This indicate the strain contrast is very sensitive to the local lattice distortions: a exponential decay in maximum lattice distortion of only 0.27 Å could produce strong contrast. On the other hand, using the $\mathbf{g}/-\mathbf{g}$ contrast inversion criteria under 0002 system row condition, it could be easy to operate and identified in experimental conditions.

4.5. Summary

In this chapter, we have discussed in detail the observation and analysis of nano-clusters embedded in semiconductor epilayers by TEM techniques. Firstly, we have discussed limitations of techniques to study embedded nano-clusters and shown in general that TEM as a suitable technique. Secondly, we have presented and discussed results from lots of examples of various imaging mode and correct way to interpret the contrast from

coherent, semi-coherent and incoherent embedded nano-clusters. We compared the results of some Gd-doped GaN models via structure modeling methods which are based on VFF and on DFT. In general, if we assume DFT methods provide the most accurate local structure, the VFF methods can provide very close results. Finally, we have discussed with simulation results about the detectability of Gd atoms in Gd-doped GaN models. We have shown that indeed the detection of single Gd atom substitution in GaN is almost impossible while there is still chance to detect them by the local strain field if they were clustered.

With all the discussions, we realized that there are no universal method to detect and study nano-scaled objects embedded in crystal matrix. It is always problem specific. One has to take into account the size of the expected cluster, the interface character, strain magnitude and Z number ratio etc. to decide the most appropriate method or technique. Therefore, before arbitrarily conclude that one can not observe what was expected, we should consider more possibilities when study unknown objects.

Chapter 5.

Case studies of nano-clustering in semiconductor epilayers

In this chapter, two case studies are carried out to demonstrate how TEM is dedicated to the investigation of the microstructure of nano-clusters and to understand the clustering in semiconductor thin films. Compared to the previous chapter, the studies in this chapter concerns more on the materials side. In the first case study, InAs nano-clusters buried in a defect free Si matrix will be studied in detail by HRTEM and GPA strain analysis. This case study of a static microstructure emphasizes the interface structure and the importance of local strain on the cluster shape. In the second case study, the formation and phase transformation of Bi-containing clusters in annealed epitaxial $GaAs_{1-x}Bi_x$ thin films will be investigated based on TEM observations. The studies of the two cases are challenging from the material science point of view, because of the complex dynamic process of atom diffusion, cluster shape evolution, dislocation nucleation and motion and phase transformation. Nevertheless, from the microscopy point of view (i.e. imaging, detection and analysis), the studies are of fundamental importance. The implementation of this chapter is deliberately considered as to prepare the related knowledge and technique for the ultimate challenging of detection and understanding the structure and formation of GdN clusters in Gd-doped in GaN that will be presented in Chapter 6.

5.1. Case study I: microstructure of InAs nano-clusters in Si

5.1.1. Introduction

Research of monolithic integration of III-V semiconductors onto Si gained tremendous attention due to the promise to realize opto-electronic or photonic devices in a single chip[121–125]. A great challenge for epitaxial growth is the lattice misfit between the Si substrate and the III-V heterostructures. This lattice misfit usually leads to processes of plastic relaxation or creation of treading dislocations after exceeding a critical dimension[121,126]. The relaxation mechanisms are only approachable with detailed microstructure information, especially the local strain state of the structure. In the case of InAs integration into Si, there is a large lattice misfit of 11.55%. This in turn result in a critical thickness of less than a monolayer for planar layer structures and a critical radius of 1.76 nm for quantum dot (nano-cluster) structure[63], respectively. Despite few early studies of embedded tiny (1–2 nm diameter) coherent InAs clusters in Si[127,128], no studies have been reported about the microstructure of embedded InAs clusters in Si which exceed the critical radius. The mechanism of misfit stress relaxation is unknown.

In this work, we present a detailed microstructure analysis of large InAs nano-clusters embedded in Si matrix by TEM analysis. We emphasize on the relation among interface structure, strain state of the InAs nano-clusters and the cluster shape. The observations

could be explained by the proposed scenario of stress relaxation derived from continuum elasticity approach.

5.1.2. Experimental

The sample growth and specifications are detailed in Appendix C.2 (page 140). Cross-sectional TEM samples oriented along the [110] and [1$\bar{1}$0] direction were prepared in a conventional manner, as is detailed in Appendix B. In order to minimize the sputtering damage, the Argon ion beam energy was reduced to about 1.2 keV for about 15 min at the final milling step. Finally, homogeneous thin foil with thickness of about (5 to 10) nm is obtained, where some of the buried InAs nano-clusters penetrate through the bottom to top of the foil. The HRTEM images acquisition and processing and the 2D strain tensors extraction by GPA have already discussed in detailed in Chapter 3.

5.1.3. Results and discussion

Strain contrast image analysis

Figure 5.1 show dark field images obtained with $\mathbf{g} = 004$ and $\mathbf{g} = 220$. The image contrast is sensitive to local strain. No detectable strain lob contrast (cf. discussions in section 4.2 in page 53) is seen in the Si matrix, indicating the Si matrix is strain-free. InAs nano-clusters buried in a depth of about 50 nm inside the Si surface is clearly revealed by the darker spots. The nano-clusterss have faceted spherical shape and diameter of about 8–15 nm.

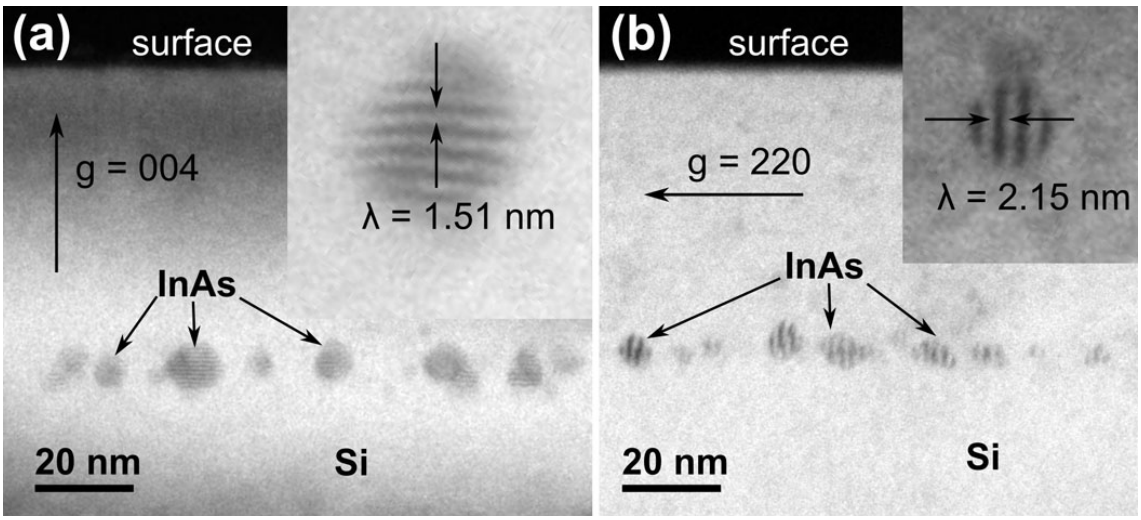


Figure 5.1: The strain contrast image show no detectable strain in Si matrix. (a) obtained with $\mathbf{g} = 004$; inset is a magnified cluster show contrast modulation with period $\lambda = 1.51$ nm. (b) obtained with $\mathbf{g} = 220$; inset is a magnified cluster show contrast modulation with period $\lambda = 2.15$ nm

The magnified cluster images clearly show contrast modulation which originated from the difference lattice plane distance of the matrix and clusters. (cf. discussions in section 4.2 in page 53) The contrast modulation periodicity along 004 direction is 1.51 nm and along 220 direction is 2.15 nm. It is easily verified using bulk InAs and Si lattice

parameter that

$$1.51 \approx (10+1) \times [004]a_{\mathrm{Si}} \approx 10 \times [004]a_{\mathrm{InAs}};$$
$$2.15 \approx (7+1) \times [220]a_{\mathrm{Si}} \approx 7 \times [220]a_{\mathrm{InAs}}.$$

This indicate the lattice parameter of the InAs clusters are close to its bulk value, i.e., the lattice misfit strain is (almost) totally relaxed.

However, the measurement of the contrast modulation period in dark field images rely on the image pixel size and integration width, which result in an accuracy of about ± 0.2 nm. Therefore the accurate strain state of the InAs clusters need further analysis.

HRTEM image analysis

Figure 5.2(a) is a HRTEM lattice image projected along the [110] direction of the embedded InAs nano-clusterss in Si matrix. Three InAs nano-clusterss are distinguishable from the contrast in this image. The two small ones have a diameter of 4–5 nm, as marked by the dashed circles in Fig. 5.2(a). The a large one has a diameter of about 12 nm measured in the center of red rectangle. The Si matrix is free of extended defects. When projected along the perpendicular $[1\bar{1}0]$ direction, the nano-clusterss are of similar shape. Therefore we conclude that the nano-clusters have polyhedron shape close to sphere. The interfaces are the {111} and {001} planes, which lie parallel to the electron beam (z) direction [1]. A stacking fault is clearly seen cross the middle of nano-clusters. This is also reflected in the streak in the diffractogram as marked by the arrow in the inset of Fig. 5.2(a). The lattice spacing of InAs and Si matrix is different, as is revealed by the splitting of diffractogram spots in the inset of Fig. 5.2(a) marked by small boxes (from InAs) and triangles (from Si). Assuming an unstrained bulk value for Si, the lattice constant of InAs deduced from the diffractogram is close to its bulk value.

After background Fourier filtering, we constructed Burgers circuits around a interface misfit dislocation and around a partial dislocation terminating the stacking fault, as shown in the magnified (and rotated) image in Figure 5.2(b). The interface misfit dislocation has the Burgers vector $\mathbf{b} = 1/2[110]$; it is a perfect edge-type misfit dislocation. The terminating dislocation has Burgers vector $\mathbf{b} = 1/3[111]$, which is a Frank partial dislocation[129].

Misfit dislocations (MDs) around the large as well as the small nano-clusterss are also revealed in the {111} Fourier filtered images as marked by the dashed circles in Fig. 5.2(c) and (d). It is evidenced that the MDs are highly localized at the interface. Since in the Si matrix it is defect free, the MDs must form complete loops. The expression for the self-energy of a dislocation loop (see Chapter 2.1.2 in page 7):

$$E_{DL} \approx \frac{Gb^2R}{2(1-\nu)} \ln \frac{8\alpha R}{b}, \quad \text{or} \quad E_{DL} \propto Gb^2 \tag{5.1}$$

show that E_{DL} is proportional to the shear modulus G and the square of Burgers vector b. It is eliminately reasonable to analyze the problem assuming the isotropic elasticity. The isotropic shear modulus and Poisson ratio are derived from the equivalence of direction dependent shear modulus surface and Poisson ratio surface. Using the elastic constants for Si and InAs[130,131], we get $G_{Si} = 67.7$ GPa, $\nu_{Si} = 0.20$ and $G_{InAs} = 30.1$ GPa, $\nu_{InAs} = 0.26$. The equivalent isotropic shear modulus of Si is more than twice compared

[1] x and y direction are defined as the $[\bar{1}10]$ and [001] directions, respectively, as shown in Fig. 5.2

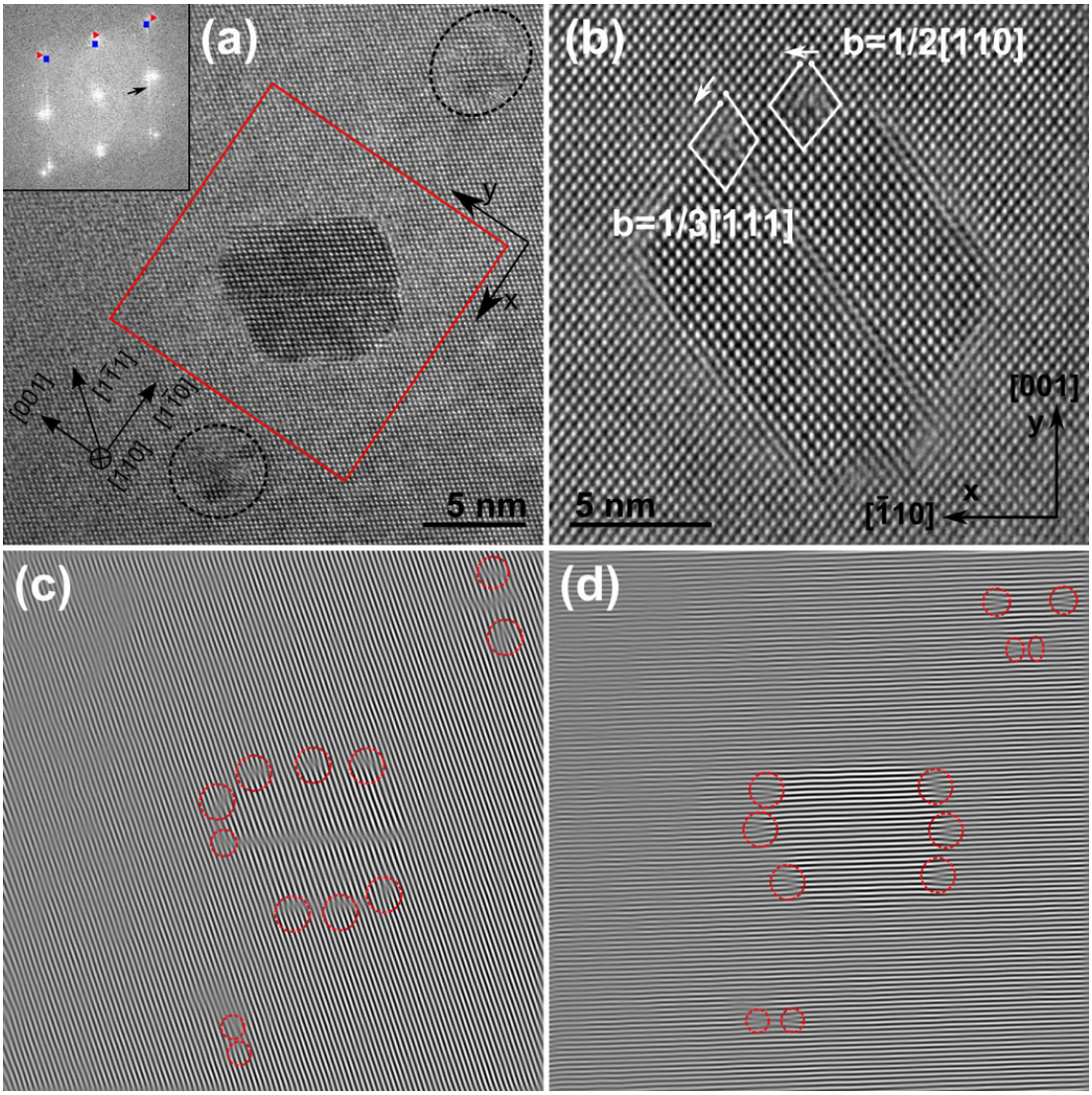


Figure 5.2: (a) HRTEM image of the embedded InAs nano-clusterss in Si. Inset is the diffractogram of the image. (b) The magnified (and rotated) region of the red rectangle in (a) show burgers circuit constructuion of the misfit dislocation $\mathbf{b} = 1/2[110]$ and frank partial dislocation $\mathbf{b} = 1/3[111]$ terminating the intrinsic stacking fault. The (111) (c) and $(1\bar{1}1)$ (d) Fourier filtered image clearly reveal the extra {111} planes as misfit dislocations as marked by red dashed circles.

to that of InAs. Therefore it is energetically favorable that the dislocation loops are restricted at the interface rather than treading into Si matrix. The scenario is similar to the "stand-off" phenomenon of misfit dislocation arrays to the elastically "softer" materials in heteroepitaxial thin films with large difference in elastic properties[132,133].

While coherently strained InAs nano-clusterss with diameter of about 10 nm buried in GaAs are reported (where the lattice misfit is 7.17%)[134], misfit dislocation loops are observed around the small nano-clusterss with diameter of (4 to 5) nm in our case. Chaldyshev and colleagues[63] have proposed a stress relaxation model. In this model, the coherent strain energy $E_{QD}(\epsilon, R)$ increases with the increase of nano-clusters radius R buried in matrix with lattice misfit of ϵ, and a critical radius of nano-clusters exist when the total

coherent strain energy equal to the residual strain energy $E^*_{QD}(\epsilon^*, R)$ plus the self-energy of creating a misfit dislocation loop $E_{DL}(R)$, or[63]

$$E_{QD}(\epsilon, R) = E^*_{QD}(\epsilon^*, R) + E_{DL}(R),$$

where

$$E_{QD}(\epsilon, R) = \frac{2(1+\nu)}{1-\nu} G\epsilon^2 \cdot \frac{4}{3}\pi R^3$$

account for the equi-axial dilatation of the buried nano-clusters, and the material conservation is satisfied via

$$(\epsilon^* - \epsilon)\frac{4}{3}\pi R^3 + b^{\perp}_{DL} S_{DL} = 0,$$

in which S_{DL} is the area of misfit dislocation loop. Using the elastic constants for Si and InAs and the coherent lattice stain of $\epsilon = 11.55\%$ (also known as lattice misfit f), the critical radius (or diameter of coherent limit) is calculated to be 1.76 nm according to the model. This suggest that misfit dislocation loops will be created to relax the misfit stress when the diameter of nano-clusterss exceed 3.54 nm. The prediction is consistent with our observations.

For the large nano-clusters, a total of 13 extra Si {111} planes counted encompass the interface with a total projected length of 36.88 nm. Based on the room temperature bulk lattice constants of InAs and Si, the difference in the number of {111} planes that could be accommodated in a length of 36.88 nm is 12.4 planes. It means in a rigid body InAs/Si interface, there will be an extra Si {111} plane covering average 12.4 InAs {111} planes. While we counted 13 InAs {111} plane in such length (of interface), if we assume that the Si matrix is free of strain, this indicates that the lattice of InAs nano-clusters is nearly fully relaxed, and only a small amount of residual tensile strain overall[135]. This is further analyzed by strain mapping. The small ones are buried in certain depth in the TEM thin specimen foil. The contrast feature is interfered by the Si matrix and consequently the strain mapping will result in artifacts. The following strain mapping is therefore focused on the large nano-clusters and the axis (x: in-plane and y: out-of-plane) are defined as depicted in Fig. 5.2(a).

Strain mapping of InAs nano-clusters embedded in Si

Figure 5.3 shows the extracted strain tensor components. The InAs nano-clusters is clearly revealed as the presence of remarkable amount of strain. At the interface the the cores of misfit dislocation loops are also clearly revealed by the stain lob pairs. Neglecting the artifact value from stacking fault, the average measured GPA strain values are attached on the mappings, which are $\epsilon^{GPA}_{xx} = (10.82 \pm 0.5)\%$, $\epsilon^{GPA}_{yy} = (12.00 \pm 0.6)\%$ and $\epsilon^{GPA}_{xy} = (0.00 \pm 0.3)\%$. No rigid body rotation is detectable. The measured GPA strain can then be interpreted as

$$\epsilon^{GPA}_{ij} = \frac{a_{loc} - a_{ref}}{a_{ref}},$$

where a_{loc} is the lattice spacing at the measured position, a_{ref} is the selected reference lattice. As mentioned earlier, the reference lattice is chosen to be in the Si side away from the nano-clusters. The lattice misfit

$$f = \frac{a^{bulk}_{InAs} - a^{bulk}_{Si}}{a^{bulk}_{Si}}$$

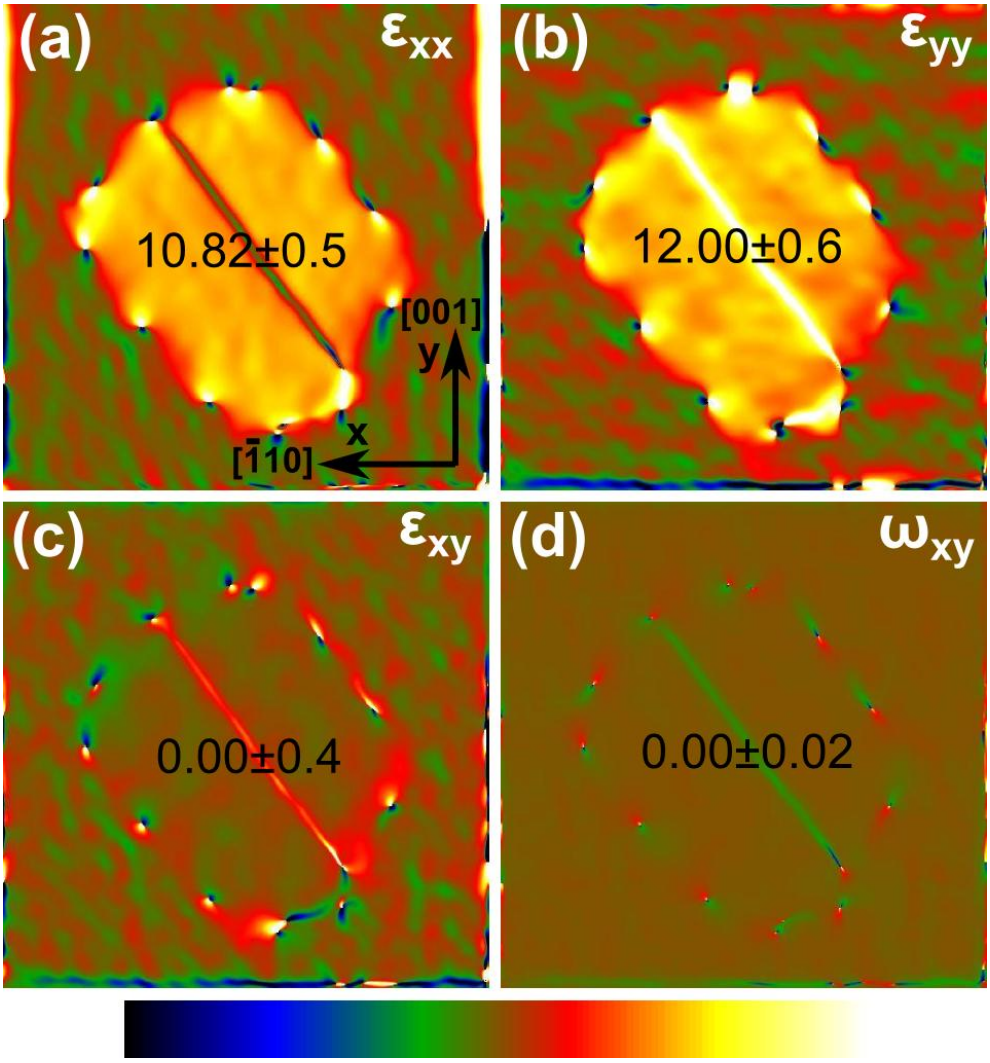


Figure 5.3: GPA strain mapping of the embedded InAs nano-cluster in Si. The principal strain component are (a) ϵ_{xx}, (b) ϵ_{yy}, the shear strain is (c) ϵ_{xy} and the rigid body rotation is (d) ω_{xy}. Color scale for strain maps (a), (b) and (c) is -20% to 20% and for rotation map (d) is -1° to -1°, anti-clockwise defines the positive rotation. The procedure to obtain the strain tensors is presented in Fig. 3.14, in Chapter 3.4.1, page 38.

for the case InAs/Si is $f = 11.55\%$. The measured GPA strain is quite close to the lattice misfit, however the "real" material strain state relies on the true materials strain which is defined as

$$\epsilon_{ij} = \frac{a_{loc} - a_{InAs}^{bulk}}{a_{InAs}^{bulk}},$$

No detectable strain gradient is present in the Si matrix, therefore it is reasonable to assume the reference Si lattice constant is grown by its bulk value: $a_{ref} = a_{Si}^{bulk}$. Then, the material strain is derived:

$$\epsilon_{ij} = \frac{\epsilon_{ij}^{GPA} - f}{1 + f}. \tag{5.2}$$

The resulting dilatation strain calculated from eq. 5.2 is shown in Fig. 5.4. Neglecting the artifact from the stacking fault, the average in-plane strain of InAs nano-clusters is

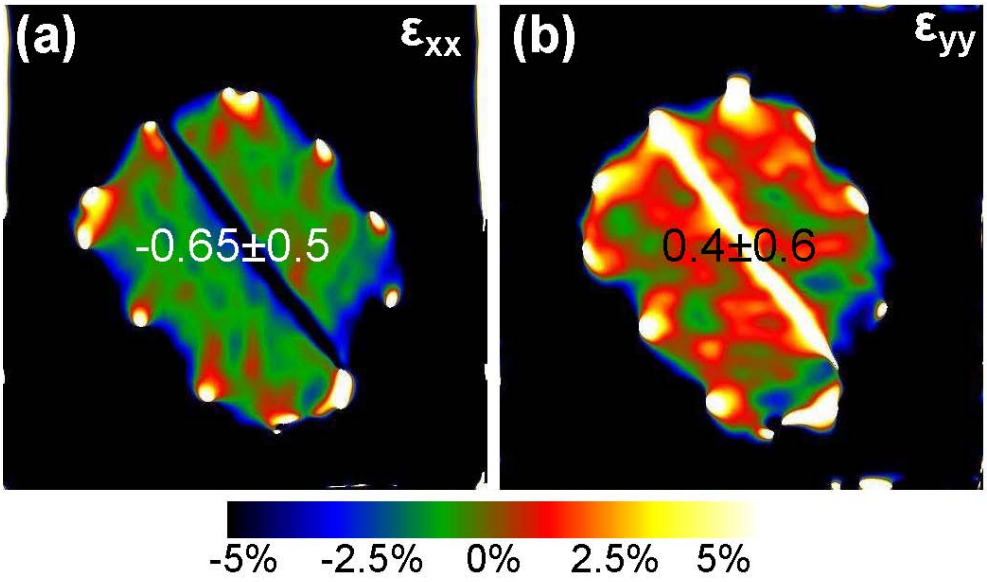


Figure 5.4: The calculated material strain (a) ϵ_{xx}^{InAs} and (b) ϵ_{yy}^{InAs}.

compressive, $\epsilon_{xx} = -0.65 \pm 0.5\%$, while the out-of-plane strain is tensile, $\epsilon_{yy} = 0.4 \pm 0.6\%$. This anisotropic in-plane and out-of-plane strain could be inherited before the nano-clusterss are covered by Si and buried.

The materials strain analysis is extended to several other nano-clusterss with different planar defects inside. The results are shown in Fig. 5.5, where the four rows are nano-clusterss without planar defects (a), with one stacking fault (b), (identical to Fig. 5.2), with two stacking fault (c) and with twinning (d). The strain value (exclude artifacts from GPA measurement) is evaluated locally and written in the corresponding images, with errors that characterize the standard deviation of the strain measurement. Concerning the in-plane strain ϵ_{xx}^{InAs} and out-of-plane strain ϵ_{yy}^{InAs}, it is clearly see that in the nano-clusters without defects inside (a) has the strongest strain anisotropy: the out-of-plane direction is strongly tensile strained. On the other hand, it is apparent that among all the nano-clusterss, this one (a) is the closest in shape to the surface lens shape InAs islands. The strain anisotropy is gradually reduced when stacking fault(s) are presented inside the nano-clusters, As can be read off row c and row d in Fig. 5.5. In case of twinning, which can be regarded as a array of stacking fault with interval of every other {111} stacking layer, the strain state is almost isotropic: under a small compression strain in both directions.

5.1.4. Summary

Interface character and strain state of InAs nano-clusters embedded in Si matrix are analysed by HRTEM lattice images and strain analysis by GPA from the corresponding HRTEM images. The embedded nano-clusters exhibit eclipse to faceted sphere in shape, as a result of incomplete to complete transformation from its initial lens shape on Si surface. The lattice misfit of 11.55% between InAs and Si are nearly fully relaxed through the creation of misfit dislocation loops around the nano-clusterss. The minimization of elastic strain energy of the misfit dislocation loops resulted in its restriction in the InAs/Si interface, leaving the Si matrix defects free. Eclipse shaped QD without internal stacking

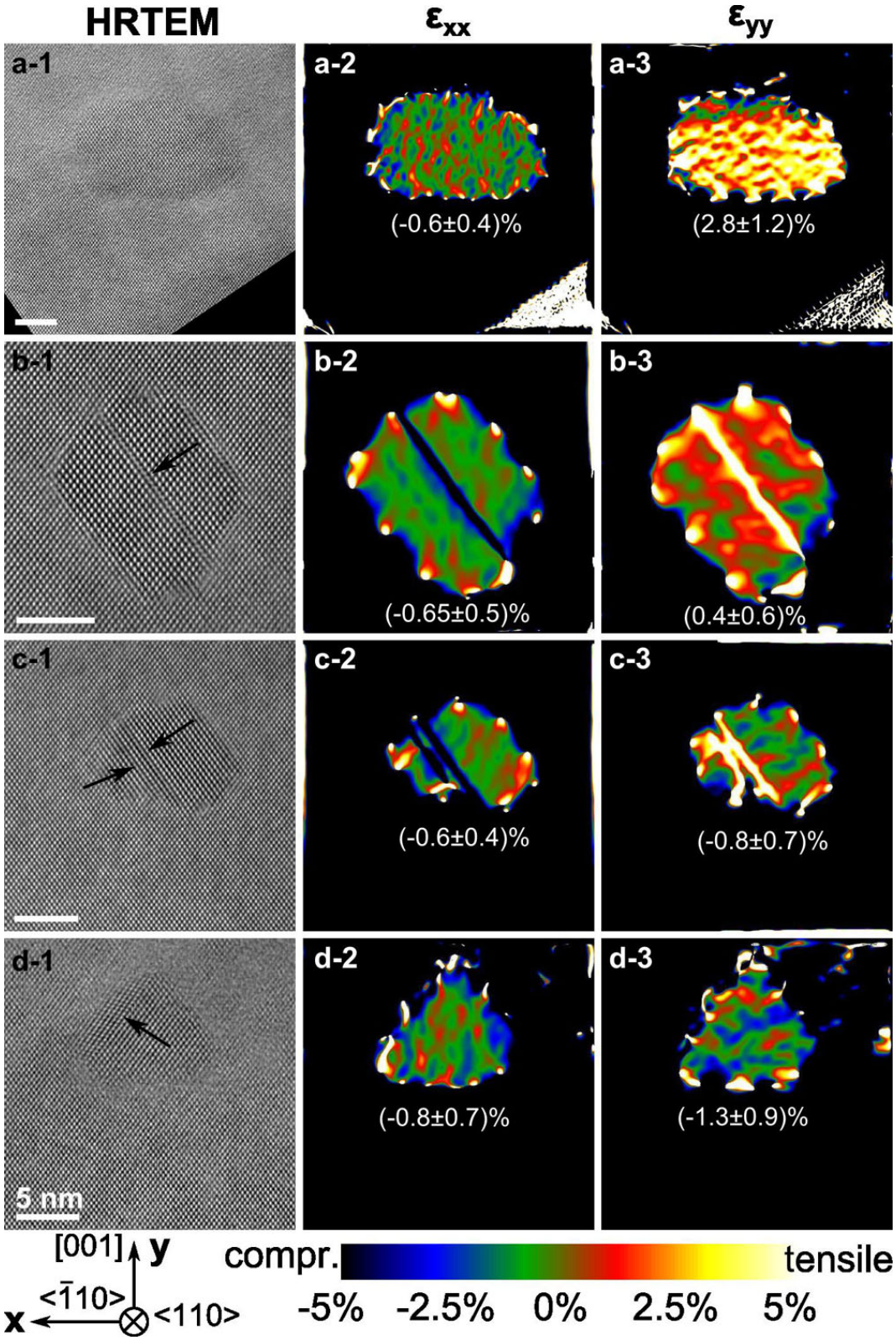


Figure 5.5: Material strain (*i*-2) ϵ_{xx}^{InAs} and (*i*-3) ϵ_{yy}^{InAs} of QD with different defects inside as show in the HRTEM images in *i*-1. (a) QD without defects inside, (b) QD with one (intrinsic type) stacking fault, (c) QD with two (intrinsic type) stacking faults and (d) QD with twinning.

fault or twin show the strongest anisotropy in strain, which is $0.6 \pm 0.4\%$ compressible in the growth plane and $2.8 \pm 1.2\%$ tensile out of the growth plane. The creation of stacking fault or twin inside the nano-clusters results in the reduction of tensile strain and therefore the interface energy turn to dominate its transform to faceted sphere shape. These results show the interplay between strain energy and interface energy in the determination of the shape of nano-clusters.

5.2. Case study II: formation and phase transformation of Bi-containing clusters in annealed GaAsBi

5.2.1. Introduction

The semiconductor alloy $GaAs_{1-x}Bi_x$ attracts increasing interest due to its potential applications in optoelectronics, such as in mid-infrared lasers, terahertz emitters and detectors, solar cells and spintronic devices[136–138]. Incorporating a small amount of Bi in the GaAs lattice leads to a strong band gap reduction, a strong enhancement of the spin-orbit splitting, and a temperature insensitive band gap[139,140]. However, due to the large difference in electron negativity, which results in a large miscibility gap, and due to the difference in atomic radius between Bi and As, which leads to a local high misfit strain, the growth conditions are limited to a very narrow window[136,138]. Creating high quality layers is therefore a great challenge.

On the other hand, semiconductor quantum dots (QDs) have been shown to be excellent candidates for photonic applications. More importantly, heterostructures based on QDs have a higher degree of freedom for epitaxial strain relaxation than planar structures[81]. This in turn enables the growth of semiconductor alloys within a wide composition range. Examples of III-V semiconductors forming QD structures have been extensively studied in systems such as InGaAs[81,134,141], InGaN[142,143] and InGaSb[144] where the QDs are grown following the Stranski-Krastanov mechanism[81,145]. In the case of GaAsBi, QDs are still not realized according the Stranski-Krastanov mechanism.

Alternatively, the spontaneous formation (or self-organization) of nano-crystals in alloys can also be driven by their intrinsic tendency to phase separate[146]. Examples of such material systems includes alloys like TiAlN[147], GeMn[9] and GaAsMn[11] etc., where there exist a miscibility gap in the material systems. In these cases, the self-organization depends strongly on the diffusion of the atoms *ex-situ* in the bulk during post-growth thermal treatment. In the case of GaAsBi, we have recently found that the Bi-containing clusters can be formed upon thermal annealing within few tens of seconds in low temperature grown epilayers, which is responsible for the reduction in lattice parameters[148]. Due to the large miscibility gap between GaAs and Bi, it is naturally reminiscent that the material system would be a good candidate to realize self-organized Bi-rich $GaAs_{1-x}Bi_x$ nano-clusters. In this respect, however, there are still no reports about the self-organization of Bi-rich nano-clusters that could, in principle, work as quantum dots.

In this study, we present a systematic investigation on the formation and the phase transformation of Bi-containing clusters in annealed $GaAs_{1-x}Bi_x$ layers by transmission electron microscopy (TEM). We propose a model accounting for the formation of zincblende Bi-rich Ga(As, Bi) nano-clusters and their phase transformation to pure rhombohedral Bi clusters. Finally, we show experimental evidences of the potential use of the zincblende GaAsBi nano-clusters as QD emitters.

5.2.2. Experimental

The samples were grown by solid-source molecular beam epitaxy on semi-insulating GaAs(001) substrates. After depositing a 130–140 nm GaAs buffer layer at 580°C, the 240-270 nm thick $GaAs_{1-x}Bi_x$ layer was grown at either low (220°C) or high (315°C) temperatures. Growth rates were 0.4–0.5 μm/h and the As/Ga flux ratios were close to the stoichiometric value, as required for efficient Bi incorporation. After growth, the samples

Table 5.1: Summary of the $GaAs_{1-x}Bi_x$ samples investigated in this work. Bi content (x), growth temperature (T_g), annealing temperature (T_a) and annealing time (t_a).

sample	A1	A2	A3	B	C
Bi (x)	$x = 0.015$	$x = 0.015$	$x = 0.015$	$x = \mathbf{0.047}$	$x = 0.013$
T_g	220°C	220°C	220°C	220°C	**315**°C
T_a	600°C	600°C	**800**°C	600°C	600°C
t_a	60 s	**120** s	60 s	60 s	60 s

are annealed in a rapid thermal annealing oven at 600°C, 700°C and 800°C for 60 s or 120 s. A summary of the investigated samples is presented in Tab. 5.1. Samples A1, A2 and A3 correspond to the same as-grown wafer (referred to as sample A hereafter) but annealed at different temperatures and/or periods. The Bi content mentioned in Tab. 5.1 is estimated from the lattice parameter extracted from X-ray diffraction (XRD) measurements on the as-grown samples after assuming Vegard's law and the literature value for the GaBi lattice constant ($a_{GaBi} = 6.33$ Å[136,140]). XRD investigations of the as-grown and annealed samples A and C are presented in[148]. We have also checked the as-grown sample by TEM. The results turned out that there were no signs of cluster formation in the as-grown samples.

Cross-section TEM specimens oriented with the foil normal along the [110] and $[1\bar{1}0]$ directions were prepared in a conventional manner, as is detailed in Appendix B. The samples are investigated either on a JEOL 2100F microscope operating at 200 kV or a JEOL 3010 microscope operating at 300 kV. The high-angle annular dark-field (HAADF) images are acquired on the JEOL 2100F system with a probe semi-angle of 14 mrad, and using a detector collection semi-angle of 80–210 mrad, ensuring the Z-contrast imaging condition[149].

The statistics on the cluster size were obtained from the analysis of the TEM micrographs using DigitalMicrograph™. In particular, we measure the diameter of the features corresponding to the clusters in the different TEM images (the contrast associated to the clusters changes depending on the imaging mode) and then plot the data as histograms.

5.2.3. Results

Clusters in sample A1

Figure 5.6(a) is a $\mathbf{g}_{002}$ dark-field TEM (DFTEM) image of sample A1. Here, the image contrast is sensitive to the crystal structure and in particular to the difference in the chemical content of group III or V element in the zincblende (zb) structure[89,150]. As is clearly seen, the presence of clusters is revealed by the dark dots in the $GaAs_{1-x}Bi_x$ layer. The size distribution of the clusters, obtained by measuring the diameter of the dark dots in the images [cf. inset in Fig. 5.6(a)] shows two peaks. There are small clusters of about 5 nm in diameter (marked by circles in the image) and large clusters of about 12 nm in diameter (marked by squares). The small clusters exhibit a relatively sharp periphery to the matrix while the large clusters show a cloudlike interface, fading to the matrix. As will be discussed below, the analysis of the contrast suggests that (1) the small clusters are single phase with a sharp interface to the matrix; (2) the large clusters have a concentration gradient at their periphery.

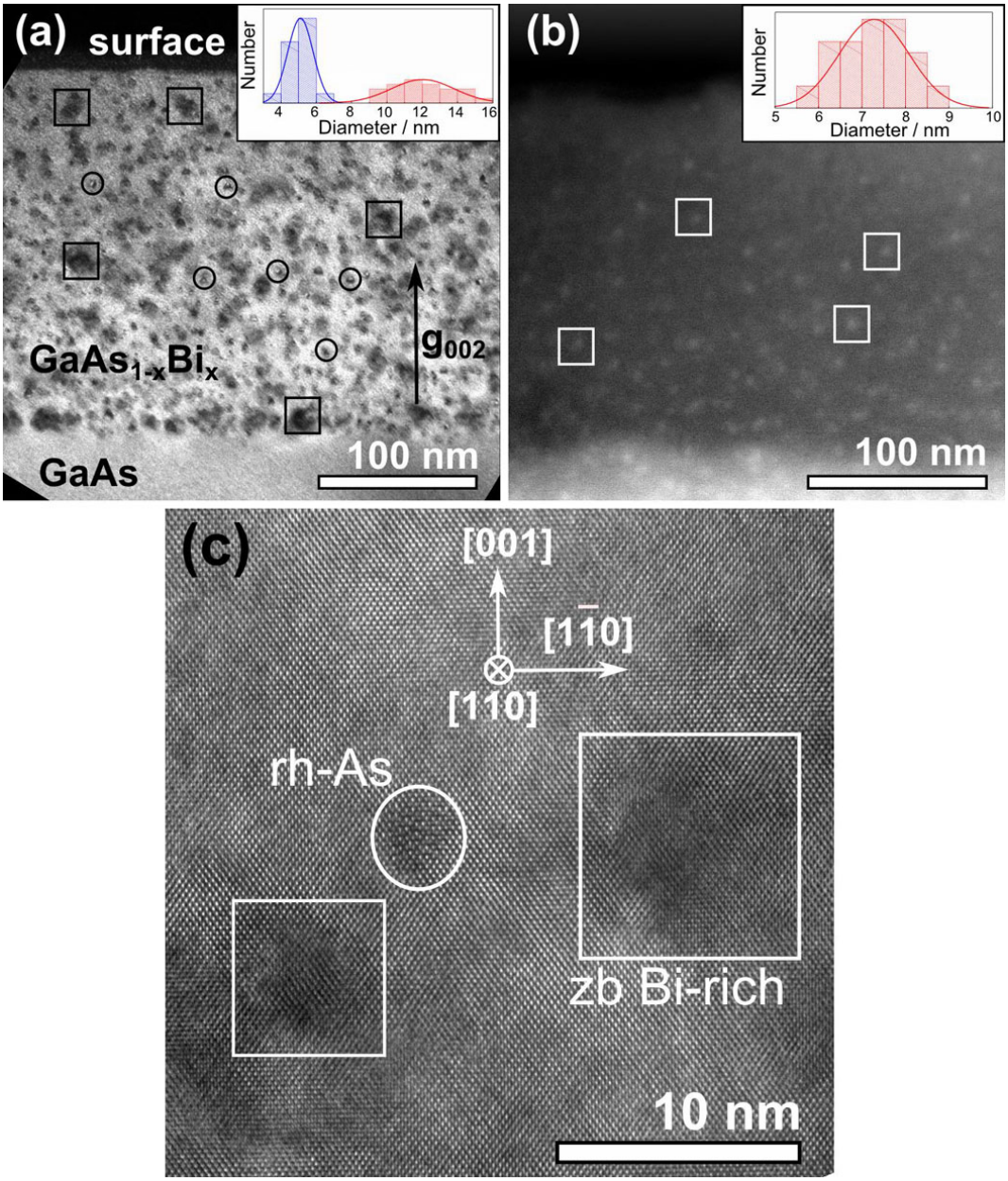


Figure 5.6: Clusters in sample A1. (a) imaged with the $\mathbf{g}_{002}$ chemically sensitive reflection; (b) HAADF-STEM i.e., Z-contrast image (c) HRTEM image. The insets plotted in (a) and (b) show the size distribution of the clusters. The boxes mark the zincblende (zb) Bi-rich Ga(As, Bi) clusters while the circles indicate the rhombohedral (rh) As clusters.

Figure 5.6(b) is an HAADF-STEM image of sample A1. In this case, the signal intensity is proportional to the atomic Z number in the projected atomic column, hence, it is also known as Z-contrast image. In this image, the contrast arising from thickness variation of the wedge-shaped specimen is already subtracted. Therefore, the brighter spots marked by the boxes can be directly interpreted as Bi-containing clusters ($Z_{\mathrm{Bi}} = 83$, $Z_{\mathrm{Ga}} = 31$ and $Z_{\mathrm{As}} = 33$). In this case, the cluster size distribution exhibits only one peak at about 7.6 nm. The small clusters observed in Fig. 5.6(a) are invisible here, indicating that only the large clusters are Bi-containing. Taking into account the difference in the

contrast formation mechanisms in DFTEM and Z-contrast images, the size difference in the observed Bi-containing clusters by this two imaging techniques reflects that there is a radical Bi composition gradient of about 4.5 nm in these clusters. This is further confirmed by high resolution TEM (HRTEM) lattice image analysis.

HRTEM lattice image reveal the presence of Moiré-like patterns with a clear interface to the zincblende matrix in the small clusters, indicating a different lattice distance and/or crystal structure of the small clusters with respect to the matrix. An example of this kind of cluster is given in Fig. 5.6(c)(marked by the circle). These small clusters are identified as As clusters in their single crystal rhombohedra (rh) structure (referred as rh-As hereafter). Similar clusters are found in annealed low temperature grown GaAs (LT-GaAs) epilayers[151] and have been extensively studied. On the other hand, the large clusters revealed by the darker regions in this image (marked by the boxes), retain the zb lattice character and are coherent to the matrix. In this case, there is no clear interface between the cluster and the matrix. Furthermore, their size agrees with the size of the large clusters found in the $\mathbf{g}_{002}$ DFTEM images.

Based on the TEM observation and on the micrographs' contrast analysis, we conclude that (1) the small clusters are rh-As clusters; and (2) the large clusters are coherent $GaAs_{1-x}Bi_x$ clusters retaining the zb structure, which are referred to as zb Bi-rich Ga(As, Bi) cluster hereafter. Since the formation mechanism of the rh-As clusters upon annealing is well-known from LT-GaAs, in this work we focus on the investigation of the Bi-containing clusters.

Clusters in sample A2

Sample A2 is grown and annealed under the same conditions as sample A1 except that the annealing time is longer, i.e., 120 s. Figure 5.7 shows a chemically sensitive $\mathbf{g}_{002}$ DFTEM image of the sample. The clusters, revealed by the dark spots in the image, are clearly seen. Again, the cluster size distribution shows two peaks, similar to sample A1. Compared to sample A1, the size of the small clusters remains the same, about 5 nm, whereas the average size of the large clusters increased by 2 nm to about 14 nm. In contrast, the periphery of the large clusters is sharper than in sample A1, indicating a sharper chemical interface of the large clusters to the matrix. HRTEM image analysis of this sample (image not shown here) reveal similar results to those from sample A1 [Fig. 5.6(c)]: (1) the large clusters remain coherent to the GaAs matrix; (2) the small clusters show Moiré patterns. Therefore, the small clusters in this sample are attributed to rh-As and the large ones are zb Bi-rich Ga(As, Bi) clusters. Additionally, the cluster size statistics indicate the zb Bi-rich Ga(As, Bi) clusters increased in diameter with longer annealing time under 600°C.

In samples A1 and A2, few clusters are also occasionally found on the surface (i.e., at the epilayer free surface, *not* at the prepared TEM specimen surface). A good example of such surface clusters is shown in Fig. 5.8, where the specimen is thin and the lattice planes are clearly resolved because of the lack of the overlapping matrix effect that could produce Moiré patterns. In this case, distinct lattice planes corresponding to rh-As and to rhombohedral single crystal Bi[152] (rh-Bi) are clearly resolved, as marked in Fig. 5.8. As observed, the crystal rh-Bi and rh-As are not intermixed, leaving an inclined interface, as marked by the red dashed line. In the power spectrum shown in the inset of Fig. 5.8, lattice planes arising from three different materials are identified: zb-GaAs, rh-As and rh-Bi. The orientation relation follows roughly $As(003) \parallel GaAs(1\bar{1}1)$ and $As(102) \parallel GaAs(2\bar{2}0)$

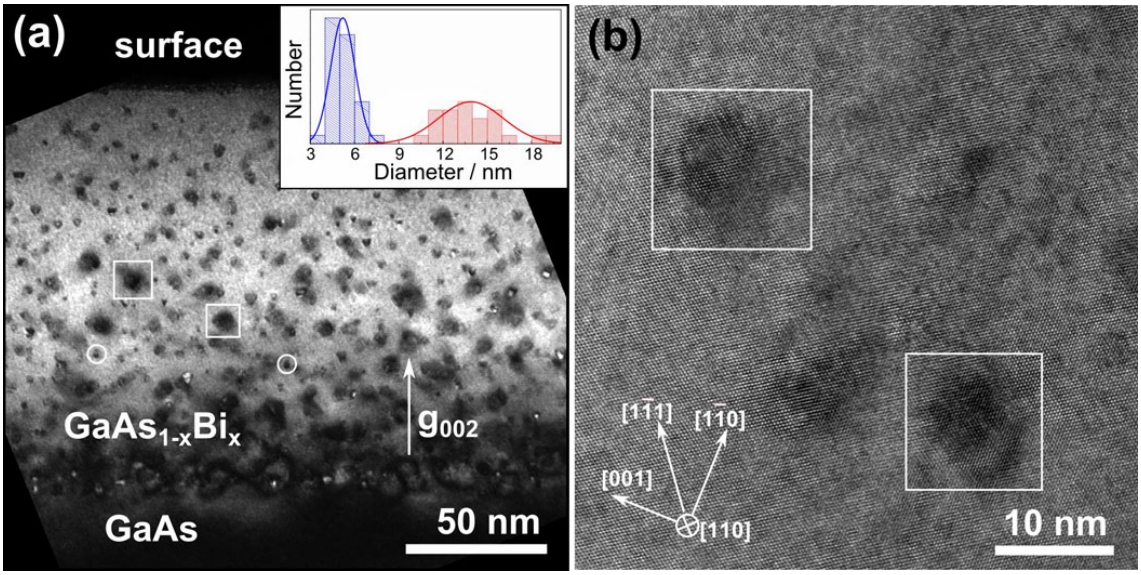


Figure 5.7: (b) Chemically sensitive $\mathbf{g}_{002}$ dark-field TEM image of sample A2. The inset shows the cluster size distribution. The boxes mark the zb Bi-rich Ga(As, Bi) clusters and the circles indicate the rh-As clusters. (b) HRTEM micrograph of sample A2. Note that because of the wedge-shaped specimen (involving a dramatic change in the foil thickness from some areas to another), the contrast of the bottom GaAs layer appears different to that in Fig. 5.6.

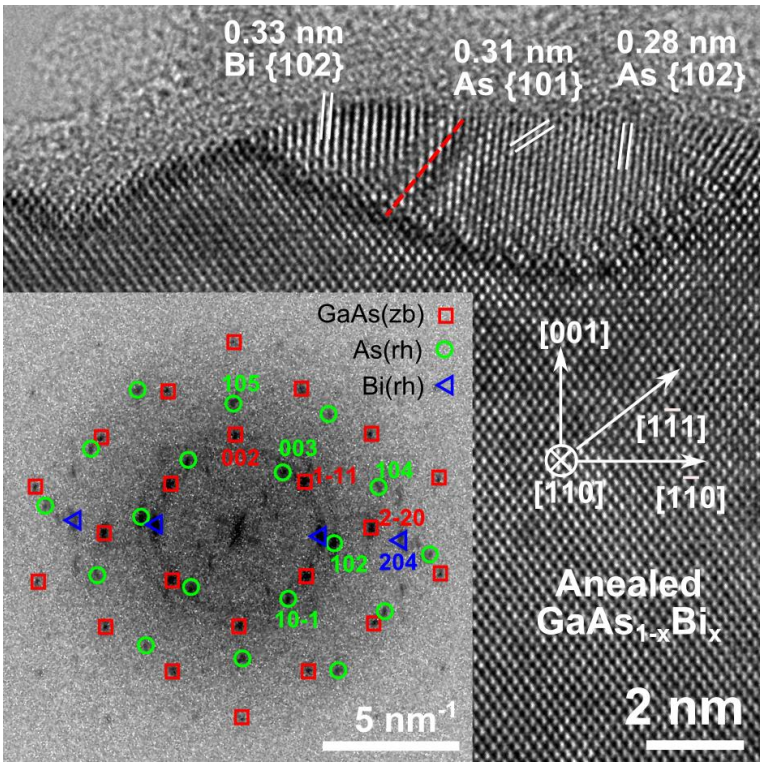


Figure 5.8: HRTEM image of a surface cluster without the overlapping matrix. The inset is the (magnified) power spectrum of the HRTEM image. Three set of lattice patterns are identified: zincblende GaAs, rombohedral pure As and rombohedral pure Bi.

[2], as has been reported earlier for As clusters in LT-GaAs[154]. Similarly in rh-Bi, with lattice parameter larger than that of rh-As, only the {102} planes are visible, which are roughly parallel to GaAs{220}, as expected. Other planes from rh-Bi are maybe out of the GaAs[110] zone axis and thus irresolvable in this case. This surface cluster provides direct crystal structure information and can be used as a reference for our analysis of the embedded cluster in the layer. We note, however, that surface cluster formation may be different to that in the bulk, because of the different free volume in each case. Additionally, the number of surface clusters is much less than the number of clusters in the layer. Therefore we will focus our discussion on the clusters *in* the epilayer.

Clusters in sample A3

Figure 5.9 presents an overview of the clusters found in sample A3. This sample is annealed for 60 s at a higher temperature (800°C) compared to A1 (600°C). The size of the clusters here varies between 6 nm and 22 nm, and the size statistics [cf. inset in Fig. 5.9(a) and (b)] does not follow that observed in samples A1 and A2. The clusters have a well-defined and sharp interface to the matrix, furthermore the clusters have coarsened (compared to samples A1 and A2, annealed at 600°C), leaving large areas of matrix free of any precipitate. Apparently, all the clusters in sample A3 contain Bi, as revealed by the bright motifs in the Z-contrast images [cf. Fig. 5.9(b)]; and most of them resembling the diamond-like morphology of the clusters in the bright-field image [cf. Fig. 5.9(b)]. Besides, the chemical interface between the cluster and the matrix is sharper here than in the Bi-rich Ga(As, Bi) clusters found in sample A1 and A2, as revealed by the sharper contrast in both the bright-field TEM image and in the Z-contrast image [Fig. 5.9(a) and (b)] and by the similar size statistics (cf. the corresponding insets), regardless of the TEM imaging mode.

Figure 5.10 shows HRTEM lattice images of several clusters evidencing the different types of structures found in sample A3. We observe that some clusters with a diameter smaller than about 10 nm remain coherent to the matrix and tend to be spherical in shape [see an example marked by box at bottom left in Fig. 5.10(b)]. Since the shape and size of these clusters are also reproduced in the Z-contrast image, they are attributed to zb Bi-rich Ga(As, Bi) clusters. On the other hand, we find that most of the clusters with diameter larger than about 16 nm show Moiré patterns [an example of Moiré patterns is shown in Fig. 5.10(a)], indicating a different crystal structure, and/or crystal orientation of the cluster with respect to the GaAs matrix. Interestingly, the patterns seem different from cluster to cluster [for example, the left and right one in Fig. 5.10(a)], although we identify four basic sets of patterns, as will be discussed later. The clusters with Moiré patterns are attributed to rh-Bi clusters, which are incoherent to the GaAs lattice. These clusters tend to have a faceted interface to the zincblende {111} and {220} planes. Finally, there are some clusters with a predominant zb structure partly exhibiting local Moiré contrast features, reflecting the presence of nanometer-sized rh-Bi crystals. Furthermore, the rh-Bi part also tends to orient in such a way that the rh-Bi/zb Ga(As,Bi) interface is parallel to the zincblende {111} planes [cf. the red dashed line in the magnified inset in Fig. 5.10(b)]. As will be discussed later, these clusters can be considered as snapshots of an intermediate stage in the phase transformation from zb Bi-rich Ga(As, Bi) to rh-Bi, and thus they are referred to as transient clusters hereafter. Additionally, we find that in

[2] We use the Miller indexed hexagonal representation for the rhombohedral structure; for the conversion, we refer to: Ref. 153, which is briefly summarized in Appendix A.2

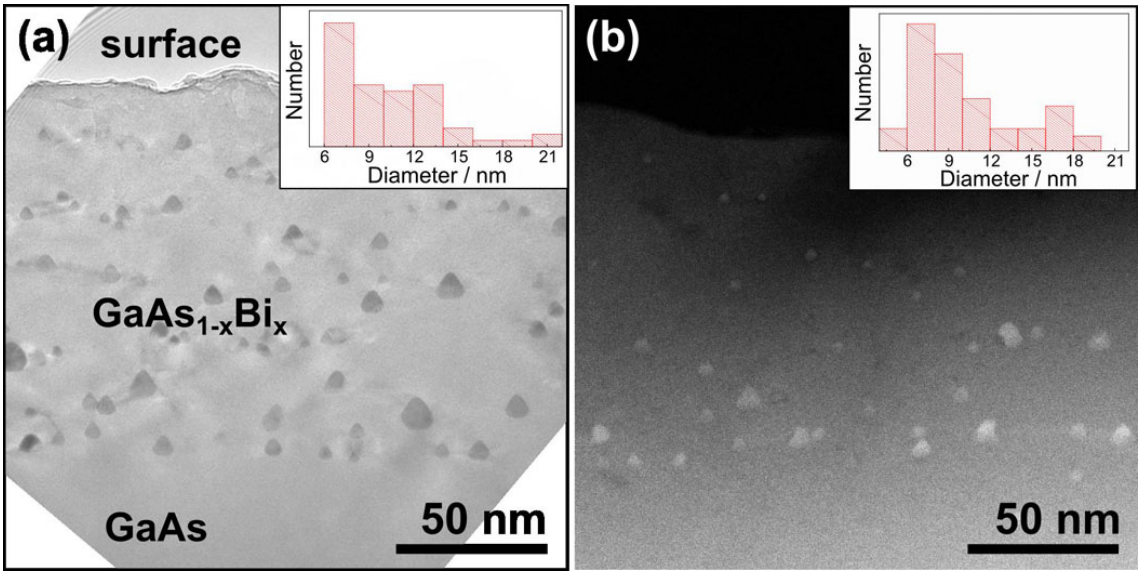


Figure 5.9: Overview of the clusters in sample A3. (a) Bright field TEM image clearly reveal the shape of the clusters; the clusters have a well-defined and sharp interface to the matrix; (b) Z-contrast image revealing that all the clusters within the film contain Bi ($Z_{\mathrm{Bi}} = 83$). The insets show the cluster size statistics. Note that (a) and (b) are taken from regions with different specimen thickness, therefore the cluster density appears different in each case.

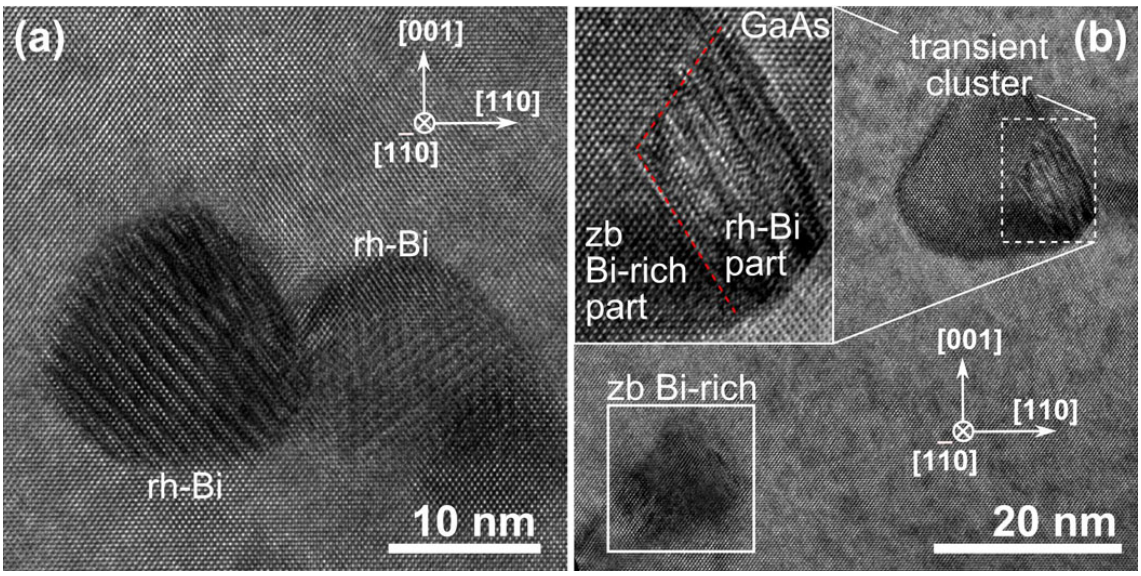


Figure 5.10: HRTEM micrographs showing clusters with different types of structures in sample A3. (a) the observed Moiré contrast feature suggests that the Bi-containing clusters are rh-Bi. (b) the small cluster retaining the zb lattice pattern is a zb Bi-rich Ga(As, Bi) cluster (bottom left) while the larger one (top right) showing a partly Moiré pattern (magnified in the inset) is considered a transient cluster from zb Bi-rich Ga(As, Bi) to rh-Bi.

this sample the rh-As clusters (cf. samples A1 and A2) precipitate to the film free surface forming clusters of 10–50 nm diameter, and even large islands of 100–200 nm (images not shown here).

Clusters in sample B

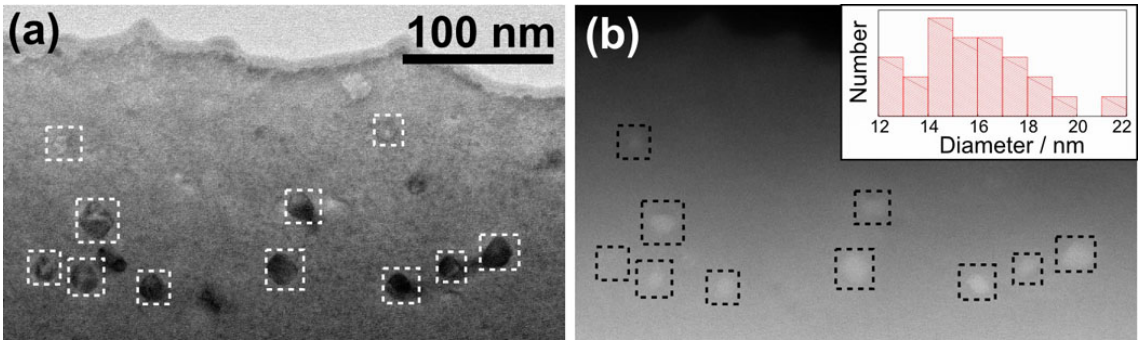


Figure 5.11: Overview of the clusters in sample B. (a) bright field image and (b) Z-contrast image of the same specimen region reveals that all clusters contain Bi. The inset in (b) is a histogram of the diameter of the bright spots in Z-contrast images.

Figure 5.11 provides an overview of the clusters in sample B. This sample has a higher Bi content of 4.6%, and is annealed at 600°C for 60 s. The clusters found in this sample have a diameter ranging between 12 and 22 nm, as revealed by the size statistics in the inset of Fig. 5.11(b), obtained from estimations of the diameter of the bright spots in Z-contrast images. Comparing bright-field and Z-contrast images [Fig. 5.11(a) and (b)], it is clear that Bi is present in all the clusters.

HRTEM images of representative clusters in sample B are shown in Fig. 5.12(a) and (b). Both images were taken at a specimen thickness of about 15 nm. Hence, there is almost no GaAs matrix overlap that would produce Moiré patterns [cf. Fig. 5.10(a)], and therefore the interpretation of the lattice fringes is straightforward. In the cluster displayed in Fig. 5.12(a) only the rh-Bi{102} lattice planes are resolvable. Nano-beam diffraction shown in the inset clearly reveals that rh-Bi{102} is roughly parallel to GaAs($2\bar{2}0$). This orientation relation is identical to that extracted from Fig. 5.8. The lattice image in Fig. 5.12(b) also corresponds to a rh-Bi cluster. In this image, the rh-Bi{102} planes are roughly parallel to GaAs(220). Therefore, Fig. 5.12(a) and (b) can be regarded as two perpendicular projections of rh-Bi. Both images are out of the low index zone axis of rh-Bi.

Results of sample C

Figure 5.13 is a $\mathbf{g}_{002}$ DFTEM image of sample C. This sample was grown under higher temperature (315°C) and contains nearly the same amount of Bi as sample A1 (about 1.3 %). The sample is annealed under the same conditions as sample A1, i.e., at 600°C for 60 s. Contrastingly, the image contrast reveals a homogeneous epilayer after annealing, except for only a few signs of surface clusters (see Fig. 5.13); no cluster formation in the layer is detected.

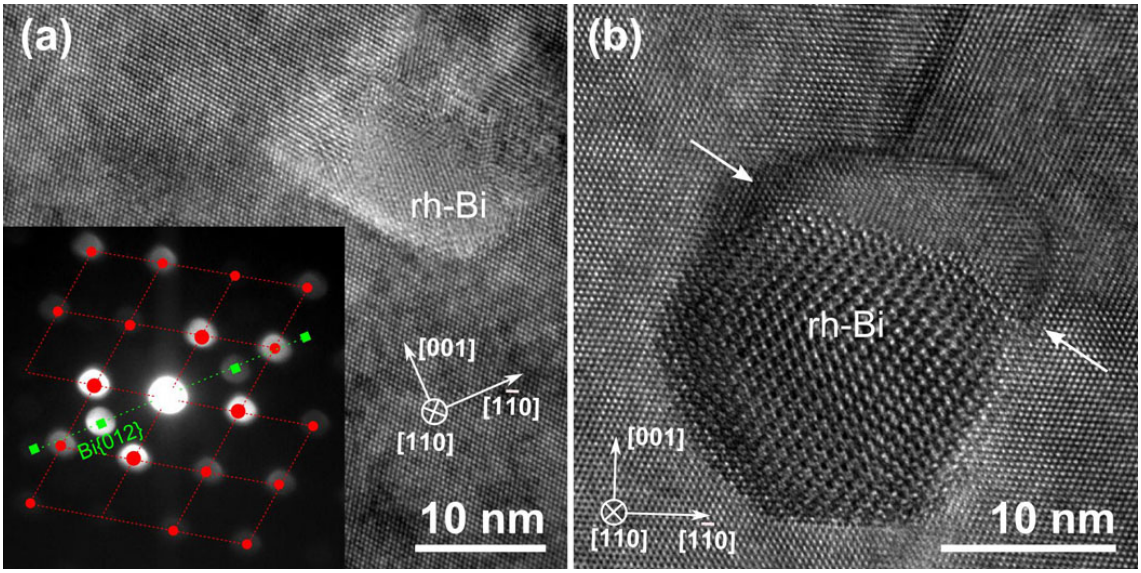


Figure 5.12: HRTEM lattice images of sample B, resolving the lattice planes of rh-Bi. (a) In the displayed rh-Bi cluster, only the rh-Bi{102} planes are resolvable. The inset shows a nano-beam diffraction pattern of the imaged cluster obtained with a probe size of 10 nm: besides the spots associated to zincblende GaAs, strong rh-Bi{102} reflections are also visible. (b) lattice image of another rh-Bi cluster. The pair of white arrows indicate the sharp interface to zb Bi-rich Ga(As, Bi).

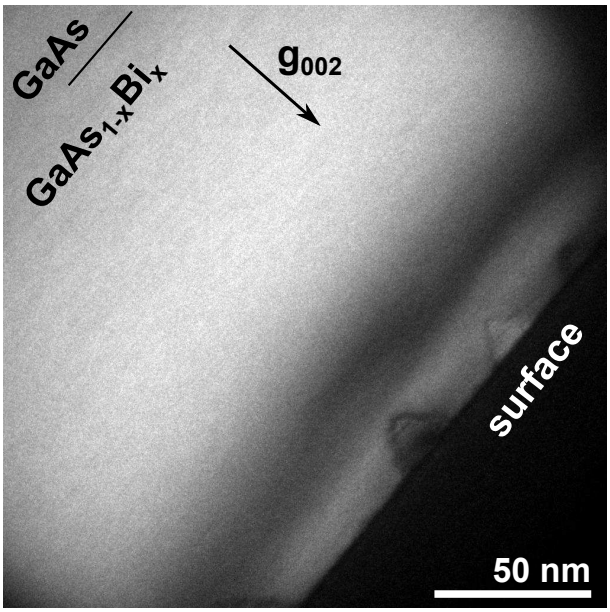


Figure 5.13: Chemically sensitive $\mathbf{g}_{002}$ dark-field image of sample C showing no clusters in the Bi containing layer, except for a few surface clusters.

5.2.4. Discussions

Beside the identified nano-clusters, dislocation loops are also found in samples A3 and B. Since it is not the main concern in this study, the results will be presented and discussed in Appendix D.

So far, we have described the clusters found in the low temperature grown samples (i.e., A1, A2, A3 and B), where several different types of clusters with various sizes and crystallographic structures have been identified. The main results are summarized in Tab.5.2.

Table 5.2: Summary of the different clusters (crystal structure and size) identified in the annealed $GaAs_{1-x}Bi_x$ layers. zb (zincblende structure); rh (rhombohedra structure); DL (dislocation loops). Note that no clusters are detected in sample C.

sample	A1	A2	A3	B
Bi (x)	$x = 0.015$	$x = 0.015$	$x = 0.015$	$x = \mathbf{0.047}$
T_a	600°C	600°C	**800**°C	600°C
t_a	60 s	**120** s	60 s	60 s
cluster type	As(rh); $GaAs_{1-x}Bi_x$(zb)	As(rh); $GaAs_{1-x}Bi_x$(zb)	$GaAs_{1-x}Bi_x$(zb); Bi(rh); As(rh)	Bi(rh)
cluster diameter	As(rh): 5 ± 1 nm; (zb): 12 ± 3 nm	As(rh): 5 ± 1 nm; (zb): 14 ± 3 nm	(zb): 10 ± 3 nm; Bi(rh): 16 ± 5 nm	rh: 17 ± 5 nm
note	no DLs	no DLs	few DLs	DLs

Formation of zb Bi-rich Ga(As, Bi) clusters

Cluster formation requires a thermodynamic driving force, together with a diffusion pathway for efficient atom transport. It is suggested that there is a general chemical driving force for phase separation in group III-V semiconductor alloys[146], which in turn will affect the material microstructure. Examples of "chemically-driven" microstructures include the development of composition fluctuations and/or cluster formation[155,156]. In the case of $GaAs_{1-x}Bi_x$, theoretical studies[140] have shown that Ga-Bi atoms are bonded very weakly in the zincblende structure. Furthermore, the equilibrium phase diagram[157] of Bi–GaAs shows that the materials system is 100% immiscible under equilibrium conditions for temperatures below about 220°C. This means that Bi in GaAs has a strong chemical driving force to phase separate.

Regarding the possible diffusion pathways for Bi clustering, it is well-known that LT-GaAs films contain a high density of point defects, typically Ga vacancies and As anti-sites[158,159]. Upon annealing, point defects, especially vacancies can diffuse fast and activate a vacancy-assisted mechanism for the diffusion of Bi. At this respect, note that in the $GaAs_{1-x}Bi_x$ sample grown at higher temperature of 315°C, where the point defects concentration is expected to be much lower than that grown at 220°C, no clusters are detected after post growth annealing at 600°C for 60 s (sample C, cf. Fig. 5.13). Although the exact role of point defects assisting Bi-diffusion is still unknown and requires further studies, this result indicates that the diffusion of Bi atoms towards the Bi-rich clusters have to be related to the presence of point defects.

Phase transformation of zb Bi-rich Ga(As, Bi) to rh-Bi

As already mentioned, we observed that depending on the annealing conditions, some zb Bi-rich Ga(As, Bi) clusters undergo a phase transformation towards pure rh-Bi clusters. As deduced from the TEM results, the phase transformation initiates after the local accumulation of Bi atoms in the zb Bi-rich Ga(As, Bi) clusters and proceeds through the

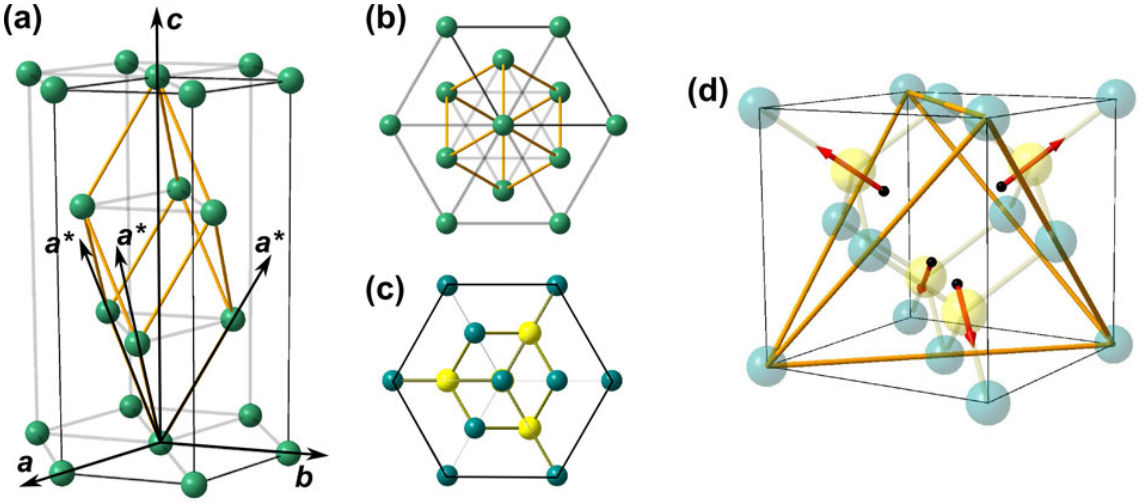


Figure 5.14: (a) Crystal structure of rh-Bi: the a, b and c vectors define the unit cell in in hexagonal representation. The center yellow "bonds" indicate the rhombohedral primitive cell where the a^* vectors indicate the basis in rhombohedral representation. (b) The rh-Bi structure projected along the c [001] direction and (c) the GaAs structure projected along the [111] direction. (d) Illustration of the four equivalent GaAs{111} planes.

nucleation of pure rh-Bi nano-crystals at energy favored sites (heterogeneous nucleation). We observe that rh-Bi and zb GaAs share the crystal orientation rh-Bi{102}||GaAs{220} and that the rh-Bi part in the transient clusters tends to have interfaces parallel to the zb {111} planes. This suggests that the nucleation of rh-Bi starts on zincblende {111} planes. Rhombohedral Bi [Fig. 5.14(a)] has one (001) basal plane, with six-fold symmetry and that contains 6 equivalent {102} planes [Fig. 5.14(b)]. On the other hand, cubic (zb) GaAs has 4 equivalent six-fold symmetry {111} planes [Fig. 5.14(d)], each containing 6 {220} planes [Fig. 5.14(c)]. Therefore, there are four equivalent ways to achieve the crystal orientation relation of rh-Bi{102}||GaAs{220}. This is consistent with our experimental observations (cf. Fig. 5.10 and Fig. 5.12) where basically four sets of (out of zone axis) lattice images or Moiré patterns were recognized.

Furthermore, our experimental observations reveal that the rh-As clusters and the rh-Bi clusters do not intermix (Fig. 5.8). This is different to the clustering behavior in the LT-Ga(As, Sb) material system, where As atoms and the incorporated Sb tend to precipitate and co-alloy to form rh-AsSb clusters[160]. Indeed, the different behavior of these material systems arises because of the completely different miscibility for As-Sb (100% miscible) and for As-Bi (100% immiscible). As deduced from the results on samples A1, A2 and A3, it is clear that the rh-As clusters and the zb Bi-rich Ga(As, Bi) clusters did not show indications of any interaction.

Model of formation and phase transformation of Bi-containing clusters

In the light of these results, we propose a model accounting for the formation of the zb Bi-rich Ga(As, Bi) clusters (stages F1 and F2) and its subsequent phase transformation (stages T1, T2 and T3) towards rh-Bi clusters, as is schematically shown in Fig. 5.15. Post-growth annealing launches the point defect-assisted (bulk) diffusion of the Bi atoms. Due to the immiscibility of Bi in GaAs, there is a chemical driving force for phase separation, which in turn will promote the spontaneous up-hill diffusion of the Bi atoms (spinodal

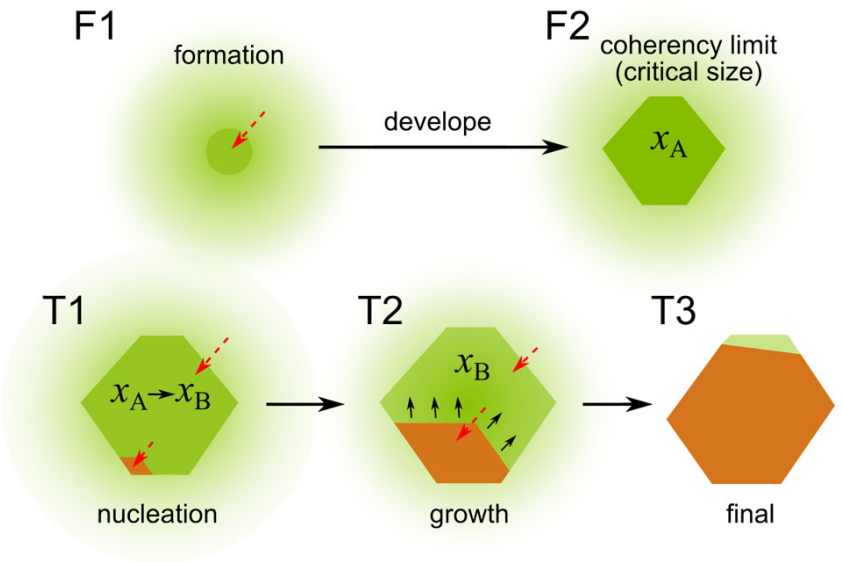


Figure 5.15: Scheme of the Bi-rich Ga(As, Bi) cluster formation stages F1 and F2 and of the phase transformation stages T1, T2 and T3. Green color stands for the zincblende $GaAs_{1-x}Bi_x$ phase, deeper green color means higher x value; brown color stands for the rhombohedral Bi phase. x_A and x_B indicate the Bi content in the clusters. The dashed red arrows illustrate the Bi mass transport; in stages T1 and T2 this results in a dynamic equilibrium Bi content.

decomposition) towards Bi-rich areas. This leads to the local accumulation of Bi atoms, hence to spatial fluctuations of x in $GaAs_{1-x}Bi_x$, and in particular to a Bi concentration gradient at the periphery of the clusters (the stage F1 corresponds to the observed cloud-like coherent clusters in Fig. 5.6). Longer annealing times will result in the further development of the up-hill diffusion until consumption of the Bi atoms in the concentration gradient region at the cluster periphery (cf. the clusters in Fig. 5.7 with a larger size). With higher annealing temperature, the Bi atoms accumulate into the clusters which triggers the phase transformation to rh-Bi. It initiates with the creation of a rh-Bi nucleus in/on zincblende {111} planes. Provided there are enough Bi atoms in the neighborhood, the rh-Bi nucleus will continue growing until total consumption of the Bi atoms (stages T2 and T3). The previous scheme suggest that the matrix behaves as a reservoir of Bi atoms, from which they diffuse up-hill into the clusters. Thus, it is assumed that, provided the Bi atoms have enough energy, the up-hill diffusion will continue until total consumption of the Bi atoms in the matrix, which will become Bi depleted.

During the thermal annealing treatment, the total free energy of the Ga(As, Bi) system as a function of time during the zb Bi-rich formation and development is a continuous smooth curve, which is a typical nature of spinodal systems[161]. The first order phase transformation of zb Bi-rich to rh-Bi results in a discontinuity in the free energy[161]. This discontinuous character in the free energy will in turn lead to two stable amounts of Bi in the different clusters: x_A in stage F2 and x_B in the transformation stages T1–T3 (Fig. 5.15), explained as follows. (1) The Bi content x_A is controlled by the local strain energy balance: more Bi entering the zb Bi-rich clusters (driven by the chemical force) will elevate the local strain energy thus reduce the rate of Bi atoms diffusing into the clusters. (2) The Bi content x_B is determined by the dynamic balance of Bi transport in the zb Bi-rich part of the transient clusters between up-hill diffuse *in* and phase transformation diffuse *out* (as

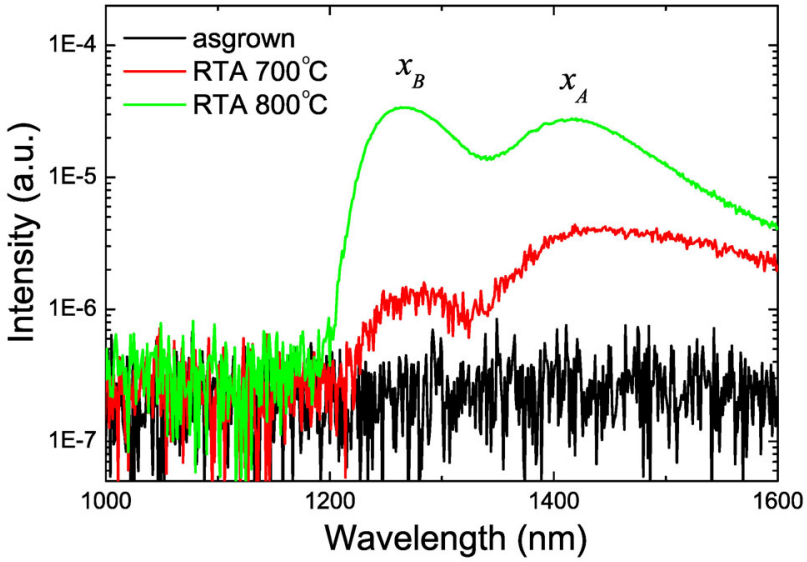


Figure 5.16: Photoluminescence of sample A1/A2 before annealing (black line), after annealing at 700°C (red line) and at 800°C (sample A2, green line) measured at 20K.

illustrated by the dashed red arrows in Fig. 5.15). Because of the critical phenomenon of free energy in the first-order phase transformation, x_A must be higher than x_B.

Structure property relation

The zb Bi-rich Ga(As, Bi) clusters of can be considered as an ideal ensemble to realize quantum dot emitters. Preliminary photoluminescence (PL) measured at 20 K shows two broad peaks at about 1.3 and 1.4 μm for samples annealed at 700°C and 800°C. The intensity of the peaks increases with increased annealing temperature. Similar results, i.e., a double peak PL emission at about 1.3 and 1.4 μm in annealed of LT-Ga(As, Bi) samples, have been recently reported by Butkute *et al.*[162]. The author speculated about the existence of several Ga(As, Bi) phases but did not provide any concrete structural information. From our detailed analysis of the clusters microstructure, we suggest that the double-peak PL emission could correlate with the existence of two types of zb Bi-rich Ga(As, Bi) "quantum dots": one ensemble with a Bi content x_A, and another one with a Bi content x_B, as already discussed.

The bi-stable Bi content configuration in the clusters (i.e., at the formation and transformation stages) could therefore result in double peaks in PL, taking into consideration the band gap reduction in Ga(As, Bi) upon Bi incorporation (about 90 meV/% of Bi[139,140]). Additionally, higher temperatures annealing promotes a faster zb Bi-rich Ga(As, Bi) to rh-Bi transformation, hence with high T_a there would be more clusters in the x_B state than in the x_A state. This is consistent with our PL results, at least qualitatively, since we find that the relative intensity of the PL peaks associated to the Ga(As, Bi) clusters with Bi contents x_B and x_A peaks, respectively, increases with increased annealing temperature.

5.2.5. Summary

In summary, systematically investigations have been performed about the microstructure of annealed $GaAs_{1-x}Bi_x$ epilayers using TEM. Bi-containing clusters only appear in low temperature grown and subsequently annealed samples. Depending on Bi content and annealing condition, the clusters show various structures with relative homogeneous size in the range of 5–20 nm. Clusters smaller than about 12 nm are zb Bi-rich and remain coherent with GaAs matrix, and larger ones begin to transform to rh-Bi clusters. We found that: (1) the formation of the zb Bi-rich clusters is driven by the intrinsic tendency of the alloy to phase separate and is mediated by the native point defects present in the low temperature grown epilayers; (2) the phase transformation from zb Bi-rich to rh-Bi nucleate in zincblend {111} planes and grow until consumption of Bi. We have proposed a model for the zb Bi-rich cluster formation and its phase transformation to rh-Bi.

Chapter 6.

Nano-clustering of GdN in epitaxial GaN:Gd

The results of the goal to clarify the distribution of Gd atoms in GaN and to understand of the formation mechanism are presented in this chapter. Firstly, we present the proofs of the occurrence of GdN clusters in the samples. The interpretation of the proofs is only possible when we put together the combination of (S)TEM imaging, processing (GPA) and simulation techniques which are discussed previously. Secondly, we extract the atomistic structure of the clusters by comparing the displacement field between experimental measurements and simulated models. Finally, the formation mechanism of the clusters is discussed based on DFT calculations in conjunction to the Frenkel-Kontorova model.

6.1. Practical challenges

Besides the challenges in the study of 3D embedded objects and the chances by using TEM (Chapter 4), the greatest difficulty come from the conflict between high spatial resolution (limited field of view) in the TEM and the extremely low Gd concentration in a sample. For example, in a GaN:Gd sample with Gd concentration of about 5×10^{16} cm^{-3}, there are GdN clusters each with about 40 Gd atoms and diameter of 1–2 nm. If we assume all Gd atoms are in such cluster and they are randomly distributed in the sample, the volume density of the cluster is $(5 \times 10^{16})/40 = 1.25 \times 10^{15}$ cm^{-3}. The possibility to "see" one cluster in a typical 1k$\times$1k CCD captured HRTEM image (about 20 nm$\times$20 nm) of a 2 nm thick specimen is about 1% (one image out of 100 images of consecutive region). This also means to obtain one image of GdN cluster in thin edge of specimen, it is necessary to move across an average distance of 100 image, i.e., 2μm. This also means a requirements for high quality sample preparation.

6.2. Experimental methodology

6.2.1. Sample preparation

Gd-doped GaN epilayers were grown by MBE using the approach previously described by Dhar and others[16,17,163]. The epilayers were grown either along the hexagonal [0001] direction (Ga-polar C-plane GaN) on GaN(0001) or 6H-SiC(0001) substrates, or along the [1$\bar{1}$00] direction (M-plane) on $LiAlO_2$(001) substrates. All GaN epilayers were intentionally doped by Gd with concentrations ranging from 10^{16} to 10^{19} cm^{-3} as determined by secondary ion mass spectrometry. We confirmed that the samples showed the same magnetic properties previously reported by Dhar *et al.*[16,17,163]

The growth conditions and specifications of the samples studied are detailed in Appendix C.1 (page 139). All the as-grown samples were examined carefully by (S)TEM.

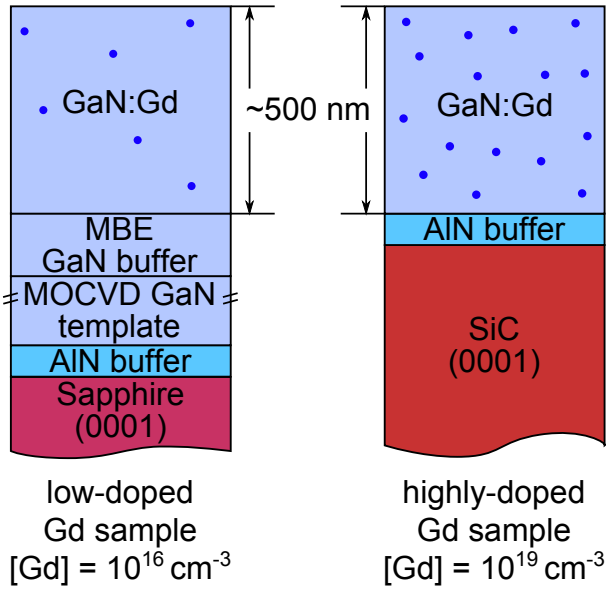


Figure 6.1: Schematic diagram of the GaN:Gd samples under investigation.

The results presented in this chapter are mostly based on two typical samples which have low and high doping concentration levels (referred as low-doped or highly-doped sample thereafter). The nominal structures of this two sample are depicted schematically in Fig. 6.1. Results that are different from other samples will be specified correspondingly.

6.2.2. Rapid thermal annealing

Ex-situ annealing of the sample with Gd concentration of 3×10^{19} cm^{-3} were carried out to target temperature of 760°C, 880°C and 1000°C (typical growth temperature is about 800°C), respectively. The annealing experiments were carried out in a programmable rapid thermal annealing (RTA) furnace with nitrogen air flow. The sample is sliced to small pieces and heated to target temperature within 40 s and kept the target temperature for 20 min. After that the sample is rapidly cooled down to room temperature with liquid nitrogen flow. The annealed sample slices were then prepared for TEM observations.

The results about *ex-situ* RTA samples will be denoted as "annealed" or "RTA processed" in the context. Otherwise, the results are referring to that of as grown samples.

6.2.3. (S)TEM imaging, analysis and contrast simulation

Cross-sectional TEM samples were prepared in a conventional manner, as is detailed in Appendix B. Conventional-TEM as well as HRTEM are performed on the JEOL 3010 microscope operating at 300 kV . The STEM observations are carried out on the JEOL 2100F microscope. HAADF-STEM images are acquired with an inner collecting angle of 70 mrads ensuring incoherent imaging.

Routine HRTEM image analysis and GPA displacement and strain extraction are according our previous discussions in Chapter 3.4 (page 38). In the case of GaN HRTEM images taken along $[1\bar{2}10]$ zone axis, the **g** vectors used to calculate the geometric phases

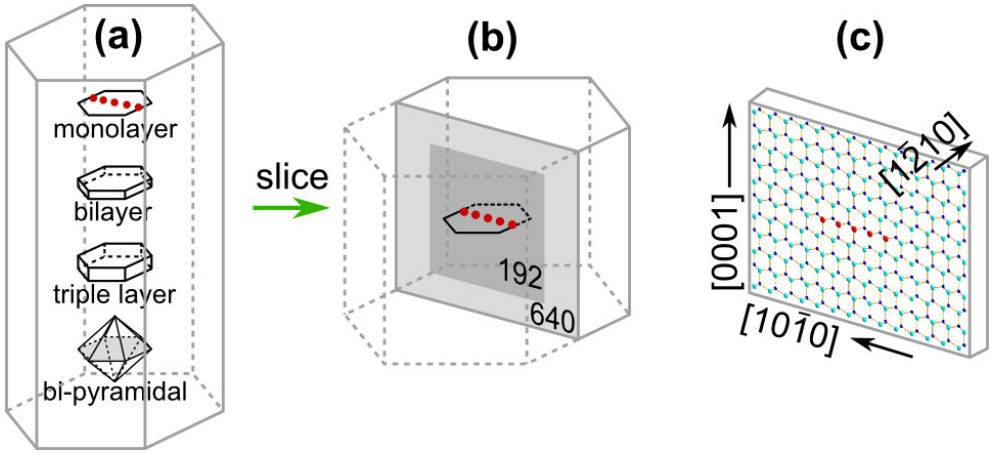


Figure 6.2: Schematic approach to modeling of GdN clusters in GaN host. (a) Various platelet models, as described in the text, oriented in the basal plane of the host GaN. (b) The supercells are defined as slicing one layer thick cell normal to $[1\bar{2}10]$ from the center of cluster. (c) A cutout from the large supercell of GaN in which the monolayer cluster is embedded. Red atoms are Gd, green atoms are Ga, blue atoms are N.

are $\mathbf{g}_{0002}$ and $\mathbf{g}_{10\bar{1}0}$, the displacement field is reduced to scalar quantities

$$\begin{aligned} u_{0002}(\mathbf{r}) &= -\frac{P_{0002}(\mathbf{r})}{2\pi \cdot g_{0002}}, \\ u_{10\bar{1}0}(\mathbf{r}) &= -\frac{P_{10\bar{1}0}(\mathbf{r})}{2\pi \cdot g_{10\bar{1}0}}, \end{aligned}$$

where $g_{0002} = c/2$ and $g_{10\bar{1}0} = \sqrt{3}a/2$ (a and c are GaN bulk lattice constants). Then the displacement is expressed as

$$u_{0002}(\mathbf{r}) = -\frac{P_{0002}(\mathbf{r})}{2\pi \cdot (c/2)} = -\frac{c \cdot P_{0002}(\mathbf{r})}{4\pi} \tag{6.1}$$

$$u_{10\bar{1}0}(\mathbf{r}) = -\frac{P_{10\bar{1}0}(\mathbf{r})}{2\pi \cdot (\sqrt{3}a/2)} = -\frac{\sqrt{3}a \cdot P_{10\bar{1}0}(\mathbf{r})}{4\pi}. \tag{6.2}$$

Dynamic strain contrast simulation is implemented in home-made Mathematica™ scripts by analytically solving the two-beam Howie-Whelan equations[91]. Details about the approximations used and the geometry of simulated models will be described in the corresponding subsection together with the results.

6.2.4. Theoretical approach and structural models

First-principles total-energy calculations were used to determine equilibrium geometries and relative energies of GdN clusters embedded in the bulk GaN lattice. Total energies and forces were calculated within the Perdew-Burke-Ernzerhof (PBE) generalized-gradient approximation to DFT using projector-augmented-wave potentials, as implemented in VASP[164,165]. The DFT calculations and interpretations were done by Steven C. Erwin at Naval Research Laboratory in Washington D.C. through collaboration.

Figure 6.2 illustrates our general approach to model GdN clusters in GaN. Our earlier

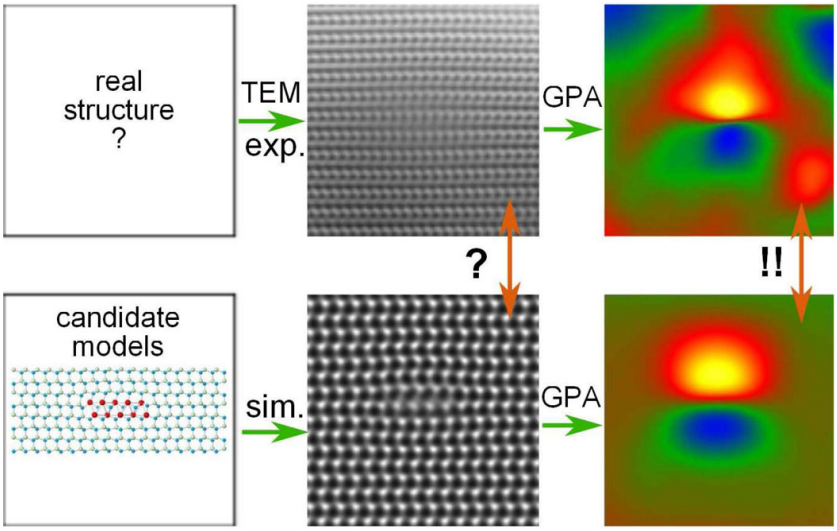


Figure 6.3: Scheme of structure model evaluation. Atomic coordinates of energy favored configuration are used to simulation HRTEM contrast and subsequent extraction of displacement field map. The real structure is then determined from the best matched displacement map.

HRTEM images suggested that Gd clusters have a platelet shape with thickness of 1–2 monolayers (MLs). Hence we considered cluster models formed by substituting Gd atoms at adjacent Ga sites within a single atom layer in the hexagonal basal plane. The example shown in Fig. 6.2(b) and (c) is for a monolayer cluster with five adjacent Gd atoms. The width of this cluster, about 1.5 nm, is in good agreement with the observed width in TEM. In a similar way, we can construct platelets with two layers (bilayer model) and three layers (triple layer model), as well as a bipyramidal shape. For simplicity, the clusters were assumed to be periodic along the direction of the electron beam, $[1\bar{2}10]$, with the periodicity of bulk GaN.

Gd has a larger atomic radius than Ga and therefore clusters defined by substitution on adjacent sites will relax outward to relieve local strain. To study this relaxation accurately we embedded the clusters in large GaN supercells containing up to 640 atoms [the grey plane in Fig. 6.2(b)]; a small cutout is illustrated in the darker grey plane of Fig. 6.2(b) and the atomistic model in Fig. 6.2(c). All atomic positions were relaxed until the largest force component on all atoms was smaller than 0.05 eV/Å. In all calculations, we used supercells sufficiently large that the periodicity had a negligibly small effect on the local strain.

6.2.5. Experimental evaluation of structural models

Conventional structure determination by quantitative HRTEM contrast analysis is based on the comparison between an experimental image and a simulated image obtained from a model of the atomistic structure[103]. Our earlier simulations revealed that the replacement of few Gd atoms in GaN matrix does not strongly affect the HRTEM phase contrast. However, even small amounts of Gd create a local lattice distortion large enough to be detected experimentally. The quantitative characterization of these displacements, using HRTEM and GPA, has previously been demonstrated.[40]. In the present study we will therefore compare the resulting displacement field map to simulated maps based on our

candidate models.

This evaluation procedure is illustrated in Fig. 6.3. Starting from a DFT-relaxed structural model, we first simulated the HRTEM image with parameters reflecting the experimental conditions. The key parameters for HRTEM contrast simulation—specimen thickness and objective lens defocus—were obtained as follows. The specimen foil thickness should be comparable to the diameter of GdN platelet [i.e. (2 to 4) nm] since the surrounding distorted atom columns is clearly observable in the projection. In the simulation, the slices are simply stacked until the desired thickness, which in turn means the clusters in simulation are penetrating from the top to bottom surface of the specimen. The defocus is estimated by fitting the contrast transfer function[106,107] to be $\Delta f = -40.5$ nm. From the resulting HRTEM image we then constructed the displacement field map using the method of GPA, as described in Chapter 3.4 (page 38). It is important to check that the periodic boundary conditions used in the DFT calculations do not affect this displacement map. We indeed confirmed that our supercells were sufficiently large that the map is negligibly affected.

6.3. Determination of GdN clusters in Gd-doped GaN

6.3.1. Results I: strain contrast imaging

Figure 6.4 shows typical dark-field TEM images of the low-doped sample, where images (a)-(d) are taken near the $[1\bar{2}10]$ zone axis and image (e) is taken near the $[10\bar{1}0]$ zone axis, respectively. Sparsely distributed, faint coffee bean contrast features are observed under condition $\mathbf{g} = 0002$ in Fig. 6.4(a), while a higher magnified image of the area marked by the dashed rectangle shown in Fig. 6.4(b) emphasizes the dark/bright character of the coffee bean contrast (indicated by white dashed circles). Similar strain contrast is discussed previously in our case studies in Chapter 4.2 (page 53). Contrast inversion is achieved by inverting the diffraction vector $\mathbf{g}$ as it is characteristic for strain contrast caused by local lattice distortions [cf. Fig. 6.4(c)]. In case of using the diffraction vector $\mathbf{g} = 10\bar{1}0$ [Fig. 6.4(d)], the strain contrast is much less pronounced pointing to a strain field that is predominantly along the [0001] direction. If the dark-field imaging is carried out close to the $[10\bar{1}0]$ projection, i.e. perpendicular to the $[1\bar{2}10]$ zone axis, the size of the coffee bean contrast is quite similar [Fig. 6.4(e)]. Additionally, it is evident from the images that the coffee bean contrast has only been found in the Gd doped epi-layer indicating that the lattice distortions are attributed to the incorporation of Gd atoms. Based on these results, we deduce that Gd forms ultra thin platelet-like GdN clusters lying parallel to the GaN basal plane and being coherently strained.

In Fig. 6.5, typical dark-field images are shown of the highly-doped sample. The characteristic coffee bean contrast, as it is observed in the low-doped case, is also visible here with similar size in the two perpendicular projection directions [cf. Fig. 6.5(a) and (b)]. Likewise, the coffee bean contrast feature is hardly seen when imaged with diffraction vector $\mathbf{g} = 10\bar{1}0$, [Fig. 6.5(d)]. Additionally, treading dislocation (TD) line bending is occasionally observed [Fig. 6.5(e)] closely related to small dislocation loops (marked by dashed circle). This observation will be discussed later. Considering that the two samples have Gd concentration differences of 3 orders of magnitude, the strain field width and extension along the [0001] direction are almost in the same range, i.e. (2 to 4) nm and (10 to 12) nm, respectively. However the density of the clusters is much higher. The term “strain width" is defined as its lateral size in the platelet plane and “strain extension" is

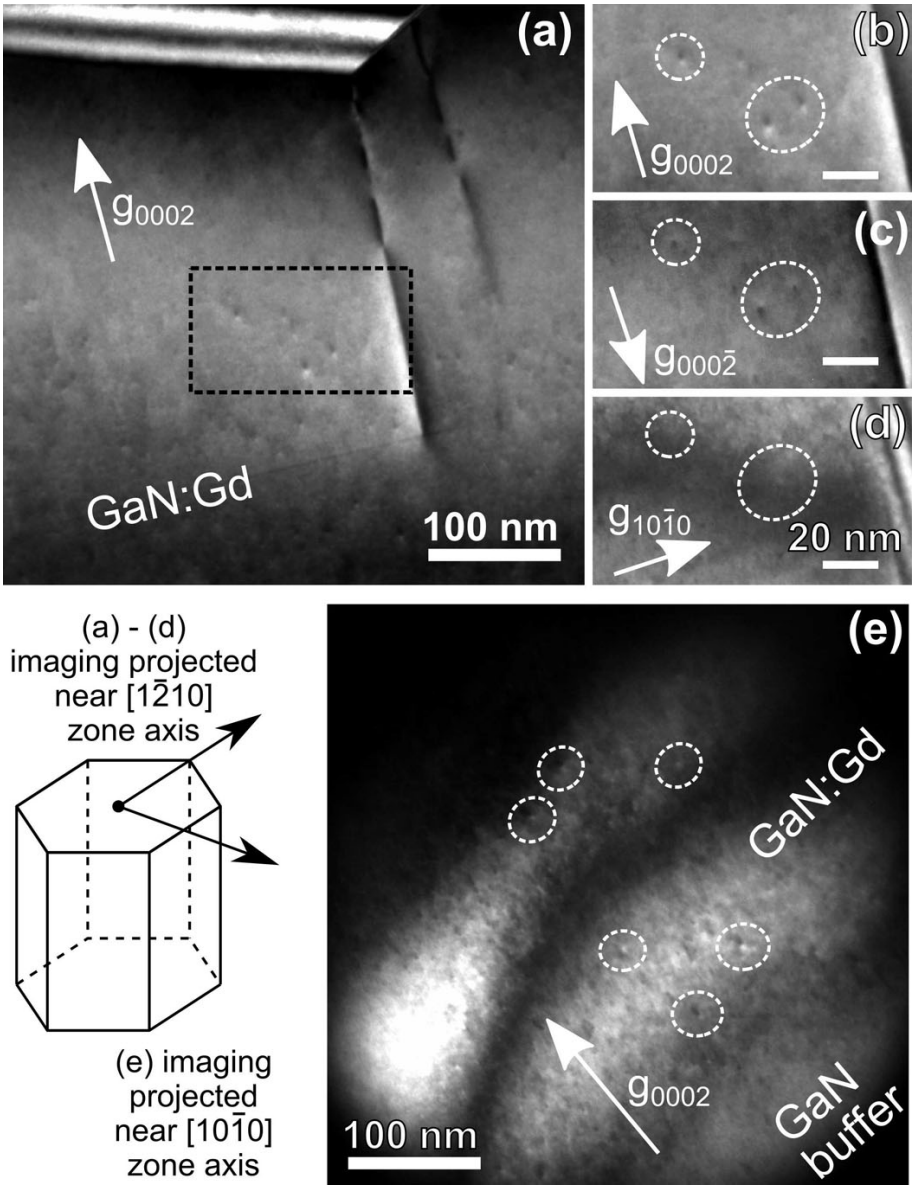


Figure 6.4: Dark-field TEM images showing the coffee bean strain contrast from the low-doped sample. The micrographs in (a)-(d) are taken near [1$\bar{2}$10] zone axis and the micrograph (e) is taken near [10$\bar{1}$0] zone axis, as illustrated by the schematic diagram. (b) Magnified view of the squared area in (a); image series in (b) to (d) display the same area, (b) and (c) are taken under $\mathbf{g} = 0002$ systematic row condition; (d) imaged under weak-beam condition.

defined in the perpendicular direction. It is a length of the line profile across the intensity maximum and minimum [dashed line in Fig. 6.5(b)] with a cut-off set to 20 % of the signal [Fig. 6.5(c)].

Normally, the quantitative interpretation of strain contrast relies on the dynamical contrast simulation which requires a prior knowledge about the displacement field of the structure. This could be obtained by an appropriate structure modeling based on the cluster configuration details or directly from HRTEM structure images. In this part the displacement field is obtained by quantitative measurement from HRTEM images.

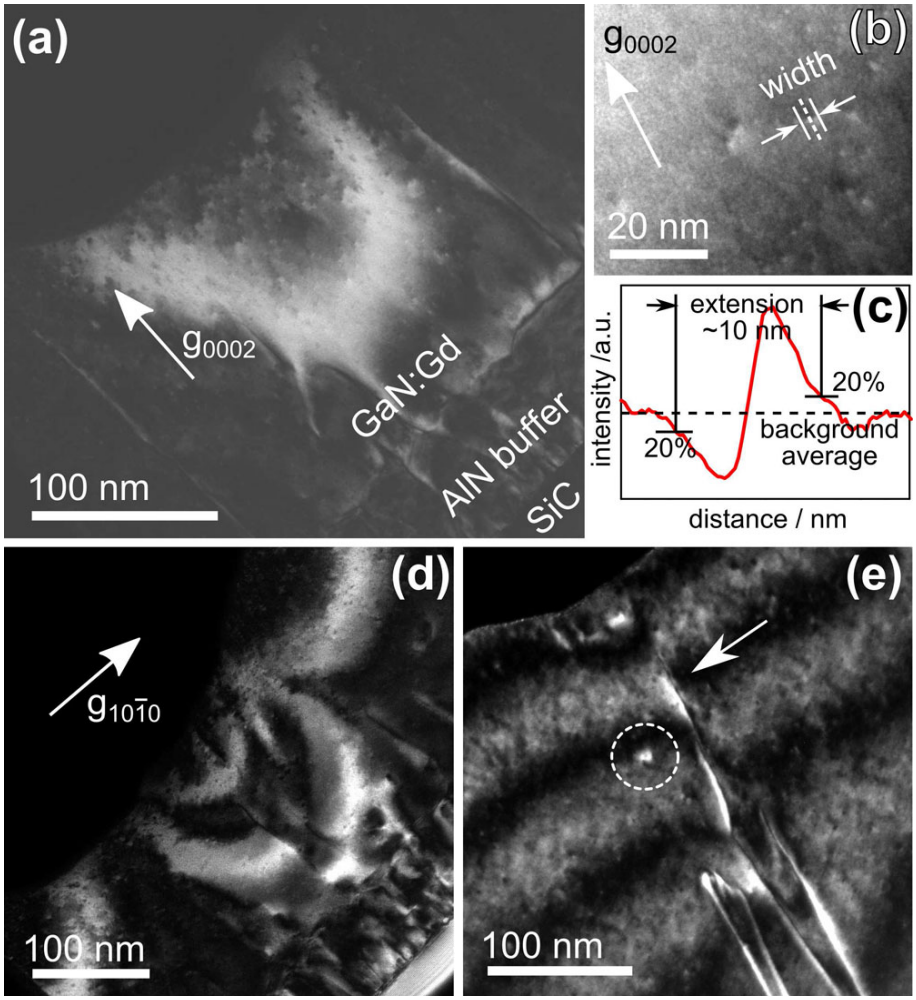


Figure 6.5: Dark-field TEM images of the highly-doped sample. The imaging **g** vectors are indicated in each image, where (a), (d) and (e) are taken near [1$\bar{2}$10] zone axis and (b) is taken near [10$\bar{1}$0] zone axis; graph (c) shows an intensity profile along the direction marked by the white dashed line in (b) defining the strain extension; the weak-beam dark-field image of treading dislocations in (e) emphasizes a bending of dislocation line (arrowed) and a pinned dislocation loop (white dashed circle).

6.3.2. Results II: HRTEM imaging and quantitative analysis

HRTEM imaging and contrast analysis are carried out to quantify the lattice distortions and to extract further information about the size and shape of the clusters. Fig. 6.6(a) is a $(1\bar{2}10)$ cross-section HRTEM image taken in a very thin area from the low-doped sample that displays lattice distortions along the basal planes. Fig. 6.6(b) is the corresponding fast Fourier transform showing the diffractogram. Applying two sets of filter masks on the 0002 and 10$\bar{1}$0 spot, the relative Bragg filtered lattice fringe images are given in Fig. 6.6(c) and (d). The 0002 lattice fringe image smears out in the center emphasizing the deviation of the basal planes from the regular periodicity, while a similar feature is not found in the 10$\bar{1}$0 Bragg filtered image. This result is consistent with the information obtained from strain contrast imaging (that strain is predominately along [0001] direction). In order to quantify the lattice deviation, GPA displacement field mapping is performed. Inserting the GaN bulk lattice constants into equations 6.1 and 6.2, the 2D

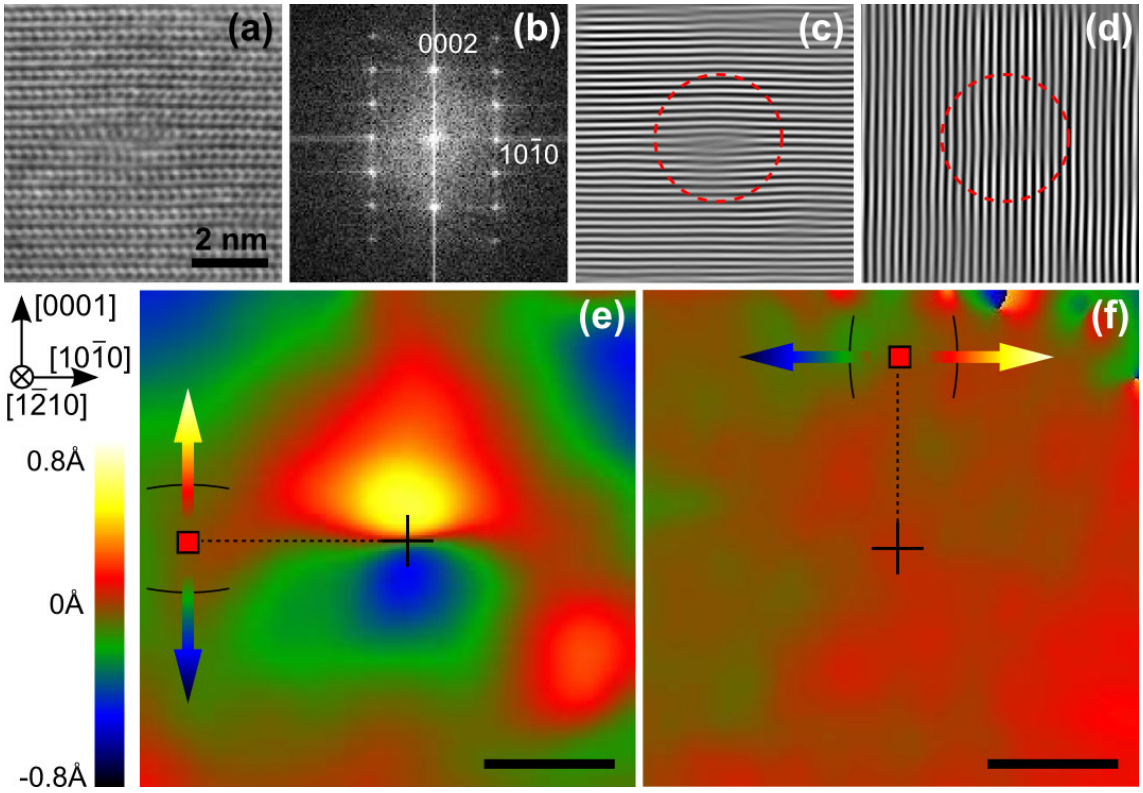


Figure 6.6: HRTEM image of low-doped sample (a) with lattice distortions and analysis (b)-(f). (b) is the fast Fourier transform of image (a), (c) is the 0002 Bragg filtered image, and (d) is 10$\bar{1}$0 Bragg filtered image. The GPA displacement field mapping of component along [0001] (e) and [10$\bar{1}$0] direction (f) calculated from (a). The red box in the mapping indicate the zero displacement position that reflects the displacement at center of cluster. The bent lines illustrate the lattice plane bend direction and the color arrow depict the strength of distortion.

displacement field $\mathbf{u}(\mathbf{r})$ is calculated from the extracted phase images $P_g(\mathbf{r})$. Figure 6.6(e) is the displacement component along the GaN [0001] representing the basal plane distortions. Strong displacements of the basal planes are indicated. The absolute maximum value obtained from the mapping is 0.53 Å. The displacement field is symmetric, i.e., the absolute value at equal distance from the center is almost identical. This is attributed to 1 to 2 monolayers of a coherent platelet-like GdN cluster with a lateral width of about 2 nm. It is supposed that part of periphery of the cluster penetrates from the top to the bottom surface in the cross-sectional specimen, therefore we could observe the projection of surrounding distorted GaN atom columns. Because of the strong surface relaxation effects in this case, the extension of lattice plane distortion is smaller, as expected, than that measured for the buried ones in strain contrast images. In Fig. 6.6(f) the displacement component along the GaN [10$\bar{1}$0] direction is stated, representing the prismatic plane distortions. No detectable displacement is identified, which is consistent with the strain contrast imaging result.

Figure 6.7(a) displays a HRTEM image of the highly-doped sample and the displacement field mapping of the component in [0001] direction [Fig. 6.7(b)]. The displacement map reveals a similar result as the low-doped case: (1) the width of the distortion is slightly larger of about 3 nm, (2) the displacement maps show a symmetric distortion

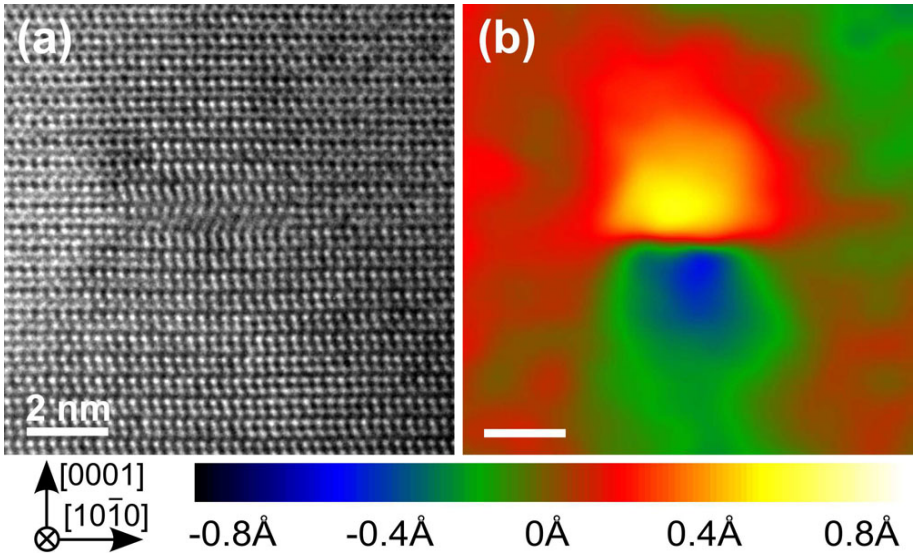


Figure 6.7: HRTEM images (a) of the highly-doped sample showing basal plane distortions and the corresponding calculated displacement field mapping of component in [0001] direction (b).

along the [0001] direction, and (3) the maximum displacement value is in a same range of 0.7 Å.

6.3.3. Results III: strain contrast simulation

After obtaining the displacement field mapping, the simulation of dark-field strain contrast image and quantitative interpretation become possible. By introducing a displacement field $\mathbf{u}(\mathbf{r})$ into the two-beam Howie-Whelan equations, one gets (see previous discussions in Chapter 4.4.3 in page 72):

$$\frac{d\phi_g}{dz} = \frac{\pi i}{\xi_0}\phi_g + \frac{\pi i}{\xi_g}\phi_0 \exp\{-2\pi i[sz + \mathbf{g}\cdot\mathbf{u}(\mathbf{r})]\}, \tag{6.3}$$

$$\frac{d\phi_0}{dz} = \frac{\pi i}{\xi_0}\phi_0 + \frac{\pi i}{\xi_g}\phi_g \exp\{+2\pi i[sz + \mathbf{g}\cdot\mathbf{u}(\mathbf{r})]\}, \tag{6.4}$$

in which ϕ_0 and ϕ_g are the amplitude of transmission and diffraction electron wave, z the depth of specimen, ξ the extinction coefficient, and s the excitation parameter. The equations are based on the dynamic electron diffraction theory using the column approximation[91,97]. Absorption is neglected since it only reduces the overall intensity.

The geometry and set-up of the simulation are schematically shown in Fig. 6.8(a) and (b). In real crystals, the distorted lattice is varying continuously over the thickness, therefore the displacement field is a depth-dependent function $u(z)$, illustrated at the right side of the blue cluster in Fig. 6.8(b). Miller *et al.*[102] have demonstrated that DF contrast simulation with abrupt displacement approximation (ADA), as illustrated on the left side of the blue cluster in Fig. 6.8(b), which assumes a constant displacement field over the depth, is able to reproduce contrast features of strained objects. Thus, the specimen can be considered as consisting of three parts along the electron penetrating trajectory. The middle part, having a constant displacement field (calculated previously by GPA from HRTEM image in Fig. 6.6), is buried within the perfect crystal. As shown

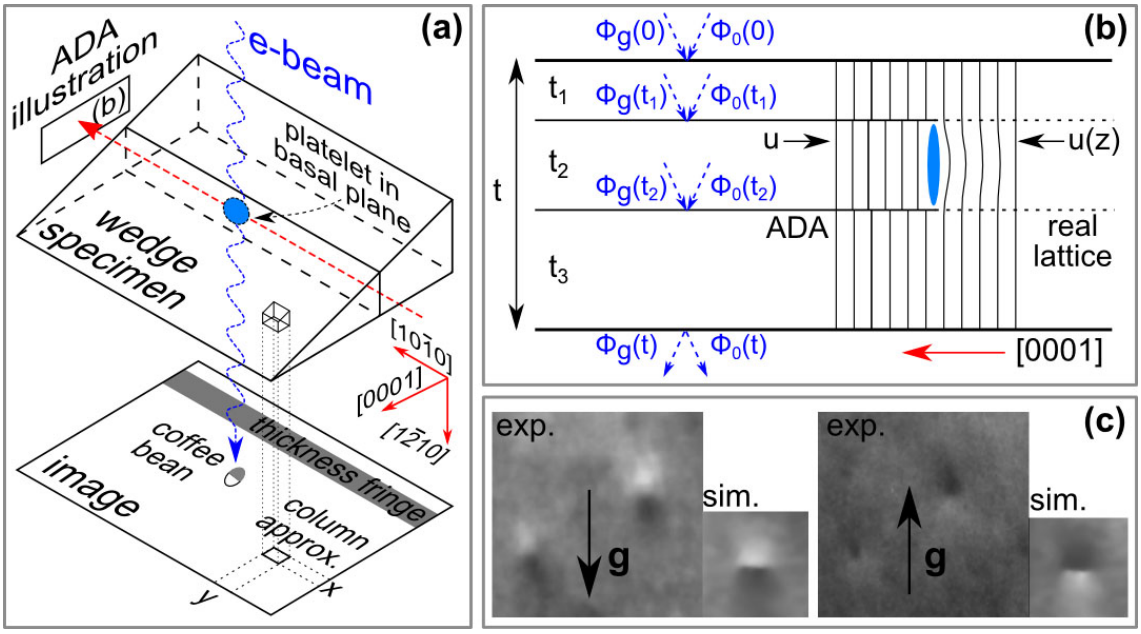


Figure 6.8: Schematic diagram of the contrast simulation method (a) and (b), and comparison of simulated and experimental results (c). Details about abrupt displacement approximation (ADA) is given in the text.

in Fig. 6.8(b), the total specimen thickness is denoted as t, and the lattice slice with Gd cluster of width t_2 is buried in the matrix starting at a depth of t_1. Because of the small size of the cluster, its contribution to electron diffraction is also neglected. Then an analytical solution of ϕ_g can be obtained as

$$\phi_g = \frac{i\pi}{\xi_g}\left(\int_0^{t_1} e^{-2\pi isz}dz + e^{-2\pi i\mathbf{g}\cdot\mathbf{u}(\mathbf{r})}\int_{t_1}^{t_1+t_2} e^{-2\pi isz}dz + \int_{t_1+t_2}^{t} e^{-2\pi isz}dz\right). \tag{6.5}$$

The final image intensity $I_g = \phi_g\phi_g^*$ is thereby a function of t, t_1, t_2, s, ξ_g and $\mathbf{u}(\mathbf{r})$. We consider only the relative contrast (normalized intensity) rather than the absolute intensity, thus the linear factor extinction distance ξ_g can be of any positive arbitrary number in the calculation. The excitation error s is an experimental variable.

In our experimental set up, $s = -0.0175\,\mathrm{nm}^{-1}$ for tilting GaN out of the $[1\bar{2}10]$ zone axis around the c axis by 5° Since under diffraction contrast imaging condition, only the component of $\mathbf{u}(\mathbf{r})$ projected on $\mathbf{g}$ will contribute to the final intensity of contrast, the term $\mathbf{g}\cdot\mathbf{u}(\mathbf{r})$ can be reduced to scalar product, $g_{0002}\cdot u_{0002} = P_{0002}(r)/2\pi$ and $g_{000\bar{2}}\cdot u_{0002} = -P_{0002}(r)/2\pi$. The parameters t, t_1, t_2 are free variables. Under these conditions, the strain contrast can be numerically calculated. In Fig. 6.8(c) the simulated intensity is presented for images with parameters of $t = 100$ nm, $t_1 = 49$ nm and $t_2 = 2$ nm, which correspond to a platelet cluster with a 2 nm width buried in the middle of a 100 nm thick specimen. In terms of the DF contrast character ($\mathbf{g}/-\mathbf{g}$ contrast inversion), the simulated contrast is in excellent agreement with the experiment results in both images with $\mathbf{g} = 0002$ and $\mathbf{g} = 000\bar{2}$ diffraction vector. Thus we conclude that the observed lattice distortion in HRTEM is of the same origin as that of the dark-field (strain contrast) images, i.e. induced by the coherently strained platelet GdN clusters.

6.3.4. Results IV: correlation of cluster density and size statistics

Based on the result that the coffee bean contrast is related to the embedded GdN clusters, we can estimate the average cluster distance and cluster diameter under the assumption that the clusters are randomly distributed within the GaN matrix. In that case, the mean distance is proportional to the inverse cluster number density $\langle\rho\rangle$, which is given by the number of clusters N per volume V:

$$\langle l\rangle = \sqrt[3]{\frac{1}{\langle\rho\rangle}} = \sqrt[3]{\frac{V}{N}} = \sqrt[3]{\frac{At}{N}}, \tag{6.6}$$

where A is the observed specimen area and t the TEM foil thickness. Additionally, if we take into account the Gd concentration $[Gd]$ that is obtained from SIMS measurement (Appendix C.1), and assume that all Gd atoms are incorporated in clusters, the average number of Gd atoms in each cluster is given by

$$\langle n\rangle = \frac{[Gd]V}{N} = \frac{[Gd]At}{N}. \tag{6.7}$$

If we further assume that the Gd atoms occupy Ga sites in the cluster forming a rigid two monolayers GdN hexagonal platelet with wurtzite structure, the average number of Gd atoms $\langle n\rangle$ and the average (flat to flat) diameter of that hexagon $\langle D\rangle$ could be derived from simple geometry (in wurtzite crystals, the (0001) planes are the close-packed planes with closest site distance a, the lattice constant). The total number of Ga site n_{Ga} in a hexagon plane with (flat to flat) diameter of d is equal to the area of the whole hexagon area divided by the unit area occupied by a single atom hexagon (the Wigner-Seitz cell): $n_{Ga} = (\sqrt{3}d^2/2)/(\sqrt{3}a^2/2) = d^2/a^2$. Therefore, two monolayer rigid GdN wurtzite platelet of (flat to flat) diameter $\langle D\rangle$ contains $\langle n\rangle = 2 \times \langle D\rangle^2/a^2$ Gd atoms.), which is expressed as

$$\langle D\rangle = \frac{a}{2}\sqrt{2\langle n\rangle}, \tag{6.8}$$

where a is the GaN lattice parameter. The quantities for calculation and results are summarized in Table 6.1. The measured lateral cluster width of about 2 nm for low-doped sample and 3–4 nm for the highly-doped sample are in reasonable agreement with the estimated platelet diameter.

6.3.5. Results V: Z-contrast imaging

Figure 6.9(a) presents a low magnified off-axis HAADF-STEM image taken from the highly-doped sample. Some slightly brighter features in an elliptical shape are visible, as indicated by dashed circles. The HAADF-STEM technique, which collects electrons scattered to high angles, produces contrast by the difference in the atomic number. The signal intensity is proportional to the atomic number, Z, it is therefore usually referred as Z-contrast imaging. The brighter feature must thus be the scattering signal from the Gd atoms having the higher Z number. It reproduces the projected shape of the clusters. The wider face of the clusters is parallel to GaN basal planes, and the apparent projected average distance of these clusters is quite similar to the DF imaging observations (cf. Fig. 6.4 and 6.5).

For crystalline specimens, the channeling effect plays an important role in the final collected image intensity[37,110]. As the GdN clusters are coherently embedded in

Table 6.1: Results of estimated average cluster number per volumn $\langle\rho\rangle$, average cluster distance $\langle l\rangle$, average number atoms in a cluster $\langle n\rangle$ and size (diameter) of cluster $\langle D\rangle$ from equations 6.6, 6.7 and 6.8. $[Gd]$ is the Gd concentration of samples derived from SIMS measurement; N is number of clusters counted in the dark-field image; A is the imaged area and t is the estimated sample thickness.

sample	low-doped	highly-doped
$[Gd]$ (cm^{-3})	5×10^{16}	3×10^{19}
N	9	38
A (nm×nm)	300×200	100×100
t (nm)	120 ± 20	40 ± 20
$\langle\rho\rangle$ (cm^{-3})	$(1.3 \pm 0.2) \times 10^{15}$	$(1.27 \pm 0.63) \times 10^{17}$
$\langle l\rangle$ (nm)	93 ± 6	22 ± 4
$\langle n\rangle$	40 ± 6	315 ± 158
$\langle D\rangle$ (nm)	1.4 ± 0.12	3.9 ± 1.1

the GaN matrix and buried at a certain depth, the channeling effect will thus reduce the enhanced scattering from heavier Gd atoms when imaging along the projected atom columns, especially when the specimen is relatively thick. Actually, in on-axis experiments, no detectable signal enhancement were found in our TEM foils. Therefore, it is necessary to tilt the sample slightly off the zone axis to reduce the channeling effect and enable contrast from the small GdN clusters. Figure 6.9(b) illustrates schematically the geometry of the experiment: the specimen is tilted about 3° out of the $[1\bar{2}10]$ zone axis towards $[10\bar{1}0]$ direction by maintaining the 0002 planes in Bragg condition.

A magnified image of the dashed box in Fig. 6.9(a) is shown in Fig. 6.9(c). The 0002 lattice planes are visible inspite of the relative large sample thickness. The fringes of the layers in the center are remarkably brighter. The normalized signal line profile along this fringe and along the adjacent layer, as indicated by the red and blue line in Fig. 6.9(c), are extracted and presented in Fig. 6.9(d) in red and blue lines, respectively. A signal enhancement of around 15% well above the noise level [the grey band in Fig. 6.9(d)] is revealed with a width of about 1.7 nm. All these results are consistent with the previous analysis. Therefore, our Z-contrast STEM results confirm the occurrence of Gd containing platelet clusters with a width of about 1.7 nm coherently incorporated along the GaN basal planes.

6.3.6. Discussion: more on the lattice distortion

The crystal structure of thermodynamically stable GdN is rock-salt (octahedrally bonded, six nearest neighbours) with lattice parameter $a = 4.99 \pm 0.2$ Å[166]. The Gd-N bonding length in this case is $d_{RS} = 2.49$ Å. In wurtzite GaN, Ga atom is tetrahedrally bonded with four nearest neighbours, and the bonding length is $d_T = 1.94$ Å as depicted in Fig. 6.10. This means the coherent platelet GdN clusters in wurtzite GaN will be under compressive strain and have the trend to form octahedral bonds to the surrounding N atoms. As a consequence the incorporated Gd atoms will relax towards the octahedral interstitial position and finally find a balance between octahedral and tetrahedral position [cf. positions i_O and s_{Ga} in Fig. 6.10] where the surrounding lattice is distorted. In a rigid body model, the bonding length of GaN octahedral interstitial position to the nearest N atom is $d_O = 2.25$ Å. The difference in bonding lengths for Gd with respect to the two

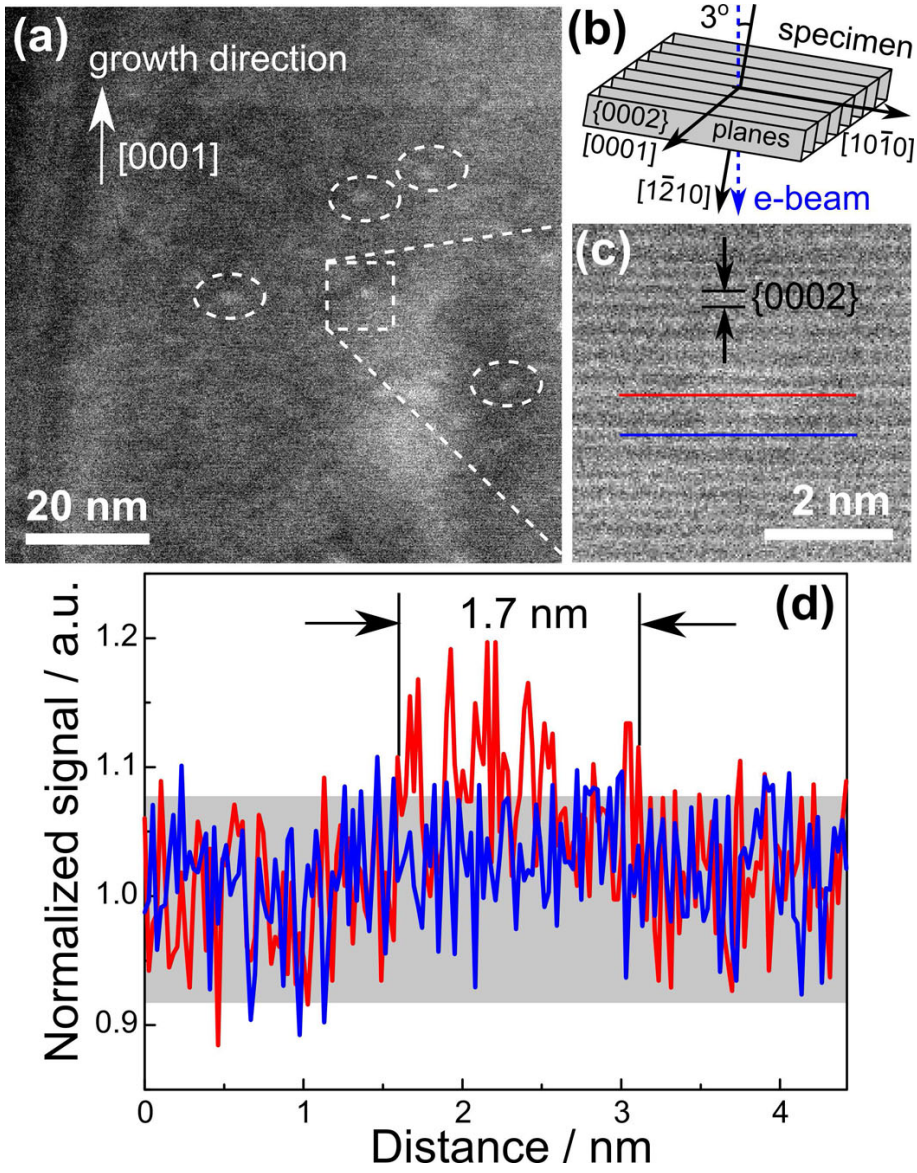


Figure 6.9: (a) Z-contrast STEM image of the highly-doped sample, image contrast enhaced for better visulization; (b) schematic illustration of HAADF-STEM (Z-contrast) imaging geometry, the specimen is 3° tilted off the $[1\bar{2}10]$ zone axis; (c) a magnified image of the platelet feature in (a); (d) line profiles of normalized signals taken from positions indicated by the color lines in (c).

possible sites is $d_{RS} - d_T = 0.55$ Å and $d_{RS} - d_O = 0.24$ Å, respectively. Then, two monolayers of GdN incorporated in the GaN lattice will introduce a maximum displacement of $2\times(0.24 - 0.55)$ Å, where we define the center of platelet at zero displacement. We introduce a lattice relaxation factor $k < 1$ to account for the lattice relaxation that will reduce the maximum displacement because of the compressive stress. Then the maximum displacement should be in the range of $k \times 2\times(0.24 - 0.55)$ Å. Therefore our measured maximum displacement of basal plane of (0.53 to 0.7) Å is in good agreement for such a two monolayer GdN platelet.

Now we estimate the stress that the distorted GaN matrix enforces on the GdN platelet. Using the elastic moduli of GaN[167], the Young's modulus of GaN along (0001)

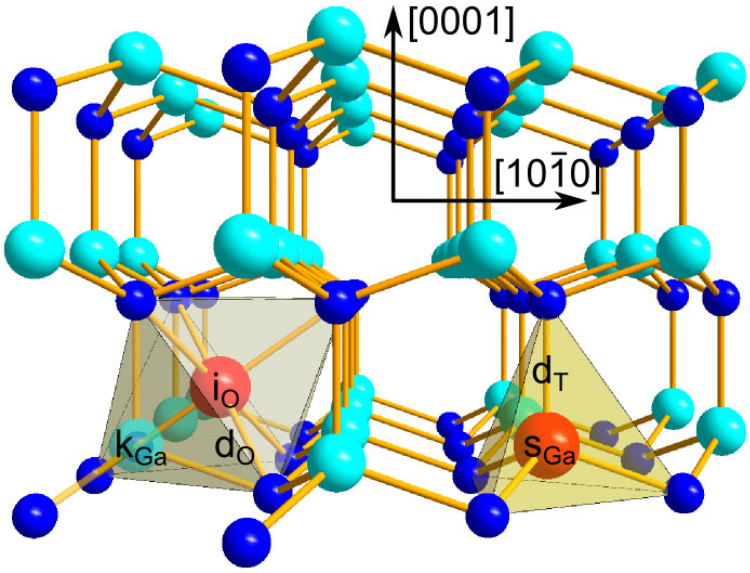


Figure 6.10: Perspective rigid-body model of different lattice sites for Gd atoms and bonding length in GaN. i_O is octahedral interstitial position, s_{Ga} is Ga substitutional site, k_{Ga} is the nearest Ga atom to i_O with distance of 1.94 Å that might be kicked if Gd enter i_O position. In the rigid body model, the bonding length to the nearest N for different site are d_O = 2.25 Å and d_T = 1.94 Å, respectively.

plane normal is $E(0001) = 161$ GPa, according to the equation of $E(hkil)$ derived in Ref.[62]. The measured maximum displacement of $d_c = (0.53 - 0.7)$ Å at the adjacent GaN layer corresponds to a local strain of $\epsilon_{001} = d_c/(c/2) = (19 - 27)$ %. Then the local stress can be estimated to be $\sigma_{001} = E(001)\epsilon_{001} = (31 - 43)$ GPa. The surrounding distorted GaN lattice enforces a compressive stress on the GdN cluster. *Ab initio* studies have predicted a phase transition of rock-salt GdN under hydrostatic pressure of (19 to 68) GPa (depending on the approximations used for calculation) to a metastable wurtzite structure[168]. Our findings are in reasonable agreement to this prediction. We will come back to this point with DFT calculation results in Chapter 6.4.1.

6.3.7. Discussion: interaction of GdN cluster with other defects

The occurrence of GdN clusters is furthermore reflected by the defect structure of the epitaxial films. As mentioned earlier, the trace of a curved TD found in the highly-doped sample, marked by the arrow in Fig. 6.5(e) (also here in Fig. 6.11), is attributed to the result of the interaction of the TD with the GdN cluster, similar to that happened during strengthening described by Orowan[56]. According to this Orowan mechanism, the glide of dislocations is blocked by rigid obstacles where the dislocation lines are pinned and the line segments in between the obstacle bow out toward the direction of motion. When the glide force is sufficiently high, the dislocation can overcome the obstacle and leaves behind a dislocation loop around. In Fig. 6.5(e), a small dislocation loop is observed (indicated by the dashed circle) which is located closely to a bent treading dislocation line. We conclude therefore that this loop is a result of the interaction of TD with the GdN cluster. The glide of TDs is driven by thermal stress emerging during cooling down the sample from growth temperature due to the mismatch in the thermal expansion coefficient between GaN layer and AlN/SiC substrate. This fact indicates that the cluster has

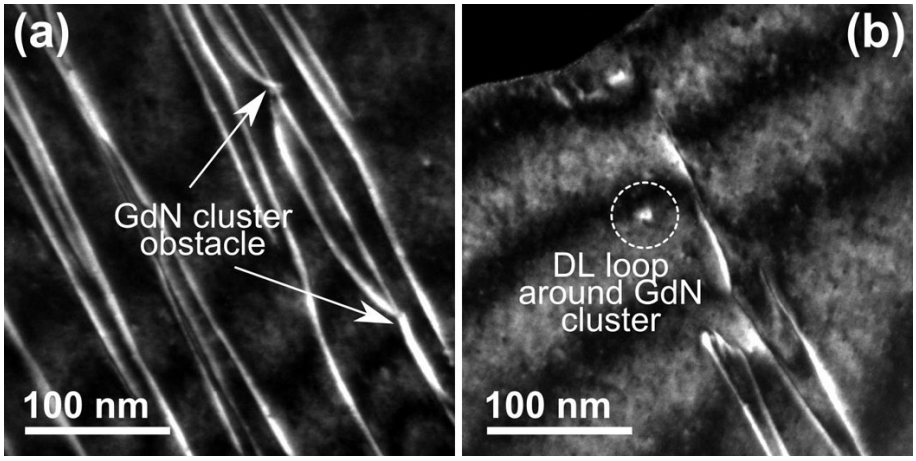


Figure 6.11: Interaction of the GdN cluster with the motion of dislocations. (a) and (b) are observed "snapshots" according to the Orowan mechanism.

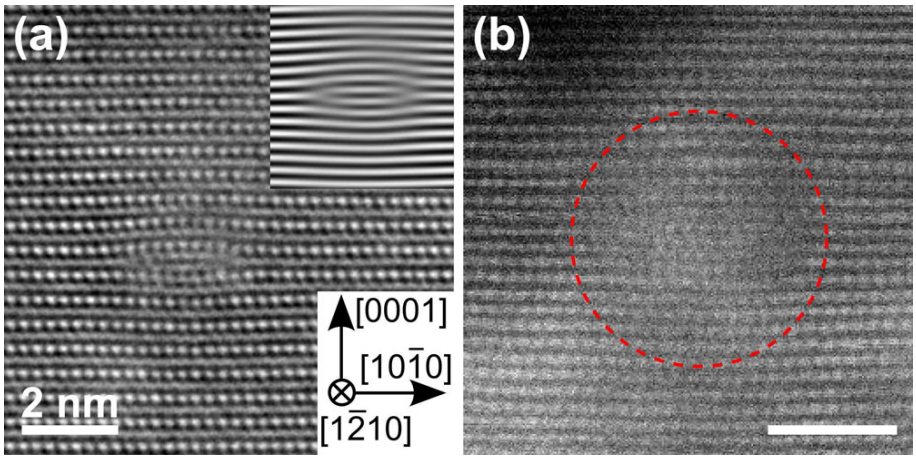


Figure 6.12: $(1\bar{2}10)$ HRTEM image showing a stacking fault loop found in the highly-doped samples; inset: Bragg filtered image demonstrating the "extra" basal plane. (b) Corresponding HAADF-STEM image with such extra plane in the center, but no detectable signal enhancement.

been formed earlier before cooling down. In this sense, we expect a hardening effect in samples doped with Gd comparing to pure GaN.

6.3.8. Discussion: stacking fault loops

Additionally, in highly-doped samples (including other highly-doped samples without in-situ annealing in growth chamber after growth), some small stacking fault (SF) loops are detected. An example is given in Fig. 6.12(a). The 0002 Bragg filtered lattice fringe image (see inset) clearly reveals an "extra" basal plane. The width of the extra plane is typically (2 to 3) nm. With Z-contrast imaging [Fig. 6.12(b)], we could hardly detect any signal difference. This kind of loop is attributed to an agglomeration of interstitial point defects potentially generated by the diffusion of Gd atoms in the GaN matrix. Similar SF loop features are also found in plasma-assisted MBE grown GaN layers[169,170] and rare-earth implanted GaN layers[171]. Based on the experimental results on Er depth profiles

in GaN thin layers through thermal diffusion, a activation energy barrier for diffusion has been deduced to be (1 ± 0.4) eV[172,173]. The authors claim that such low diffusion barrier is explainable by an interstitial-assisted mechanism. We suppose an analogue scenario with extrinsic oxygen interstitial-assisted mechanism for Gd diffusion in GaN. On one hand, oxygen in GaN is unavoidable for our growth condition, and on the other hand, it can be assumed that Er and Gd behave similarly in GaN because of their comparable atomic size and electric properties. In highly-doped samples, there is higher probability for the Gd diffusing atoms being close to each other, where the strain field produced by Gd atoms will overlap and work as sink to attract the interstitial defects and consequently form stacking fault loops. In the low-doped sample, the probability for diffusing Gd atoms being close to each other is much lower, and therefore, the SF loops are formed hardly.

6.4. Atomistic structure and energetics of the GdN clusters

6.4.1. Crystal structure of the GdN clusters

We found that the GdN platelet clusters are always oriented with the broader face parallel to the GaN basal (0001) *C*-plane, regardless of whether the samples were grown along *C*-plane or *M*-plane orientation. Figure 6.13 shows a typical dark-field strain contrast micrograph illustrating the cluster orientation for both cases. We previously established that the bright-dark "coffee bean" contrast indicates that the predominant strain field is along the [0001] direction. This in turn implies that the broader face of the GdN platelet cluster is always parallel to GaN basal plane, i.e., perpendicular to the growth direction in *M*-plane samples. The basal planes in GaN are elastically softer than the prismatic planes, and hence this behavior is expected. The same phenomenon governs the formation of so-called Guinier-Preston (GP) zones in metal alloys, where the precipitates form plate or needle shapes in the elastically softer direction of the matrix in order to minimize the elastic strain energy [3,161].

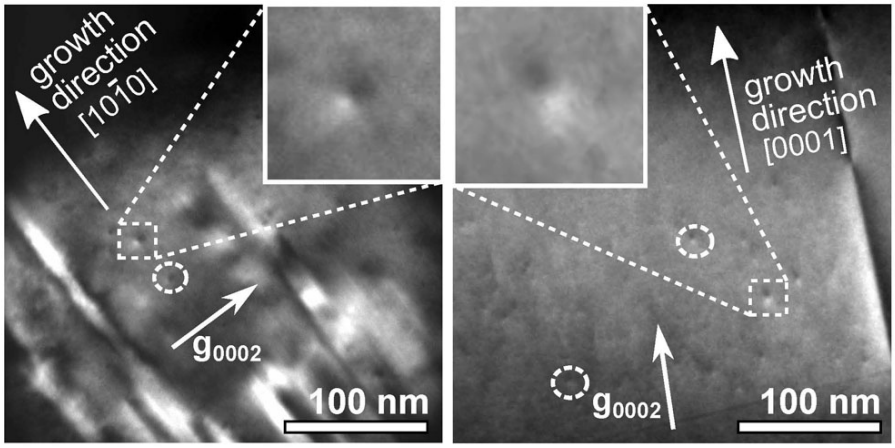


Figure 6.13: Dark-field image of an m-plane GaN:Gd sample (left). The white dashed circles indicate the characteristic dark-bright "coffee bean" strain contrast produced by the GdN platelet clusters, which is 90° rotated compared to the c-plane sample (right), see insets. This indicates that GdN clusters habit in the basal plane regardless of the crystal growth direction.

In general we found that all cluster models, regardless of their shape, relaxed within DFT to have an internal structure closer to distorted rocksalt GdN (octahedral coordination) than to wurtzite GdN (tetrahedral coordination). An example of a such a relaxed bilayer platelet is shown in Fig. 6.14. Comparing the resulting arrangement of Gd and N atoms [Fig. 6.14(b)] to bulk rocksalt GdN [Fig. 6.14(c)] it is clear that the relaxed structure is locally close to that of rocksalt. This is physically sensible because rocksalt is the most favorable crystal structure of GdN [168].

6.4.2. Atomistic structure of the GdN clusters

Now we look into the atomistic crystal structure of GdN clusters embedded in GaN. Figure 6.15 shows the resulting HRTEM images and displacement field maps from experiment and for four cluster models: monolayer, bilayer, trilayer, and bipyramid (see

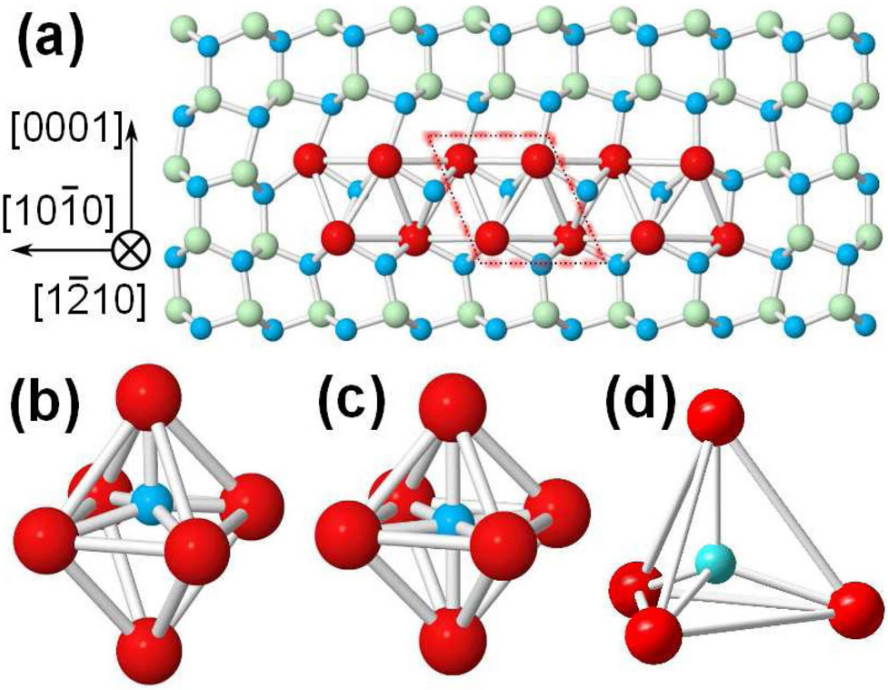


Figure 6.14: (a) Fully relaxed theoretical structure of a 6×2 platelet cluster of GdN embedded in GaN. (b) Detailed view of the internal structure of a GdN cluster defined within the dotted outline of panel (a). The bond angles and distances are close to those of rocksalt GdN. (c) and (d) Corresponding cutout from a bulk GdN crystal with ideal rocksalt structure and from an artificial wutzite GdN, respectively, for comparison to panel (b).

scheme in Fig. 6.2 in page 103). It is immediately clear from the experimental images that the component of the displacement along $[10\bar{1}0]$ is much weaker than along [0001], indicating that the largest distortions occur in the hexagonal basal planes. In view of this we can rule out the bipyramid model, because it shows equally large distortions in both the horizontal basal planes and vertical prismatic planes. For the other cluster models the displacement maps are qualitatively consistent with experiment, with $u_{0001} > u_{10\bar{1}0}$. In all cases the displacement along [0001] is approximately equal above and below the central basal plane (center of GdN cluster). With increasing platelet thickness the displacements in the $[10\bar{1}0]$ direction become correspondingly larger.

In order to quantitatively differentiate among the three models, we show in Fig. 6.16 line profiles of the experimental and simulated displacement field maps along the [0001] direction. The maximum displacement of u_{0001} (red crosses in Fig. 6.16) occurs at nearly identical distances for the monolayer and bilayer model, and match very closely the location of the experimental maximum; on the other hand, for the trilayer model this maximum is shifted outward. The magnitude of the maximum displacement varies from model to model; the best match to experiment occurs for the bilayer model (0.40 Å compared to 0.53 ± 0.15 Å). Hence we judge that the bilayer model offers the closest match to the experimental displacement maps. The discrepancy in the magnitude of the displacement cannot be simply attributed to the assumed unit periodicity along the TEM beam direction, because a more realistic model would likely lead to smaller displacements, not larger. We found this discrepancy in the analysis of all observed GdN clusters. Nevertheless we tentatively conclude that the experimental results are best explained by a

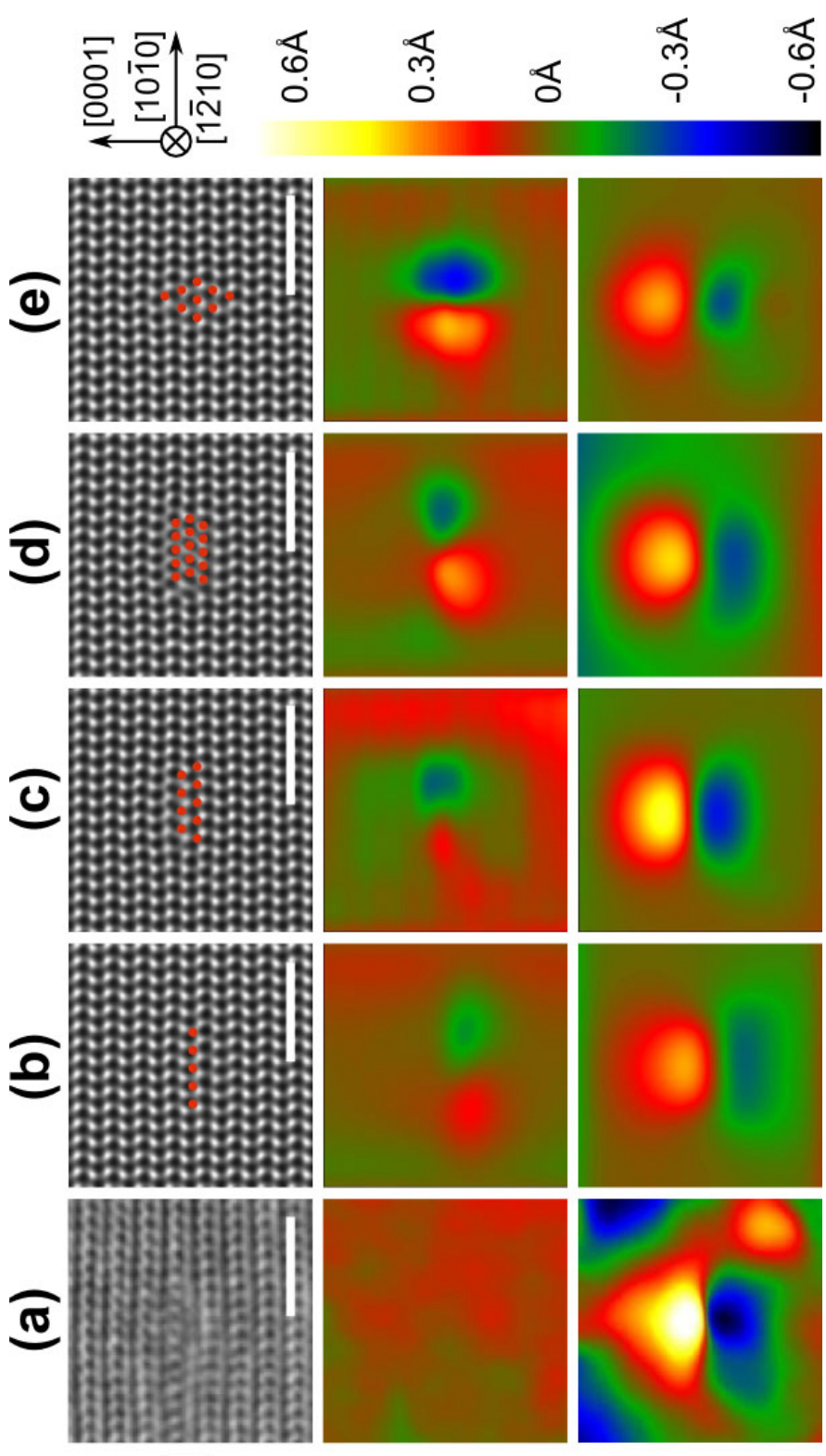


Figure 6.15: Experimental evaluation of structure models. Rows are the experimental and simulated HRTEM images, the displacement field map component in [10$\bar{1}$0] and [0001] directions, respectively. Columns are results of experiment, monolayer model, bilayer model, trilayer model and bipyramid model, respectively. Small red circles in the simulated HRTEM images represent Gd atom position. Scale bar is 2 nm.

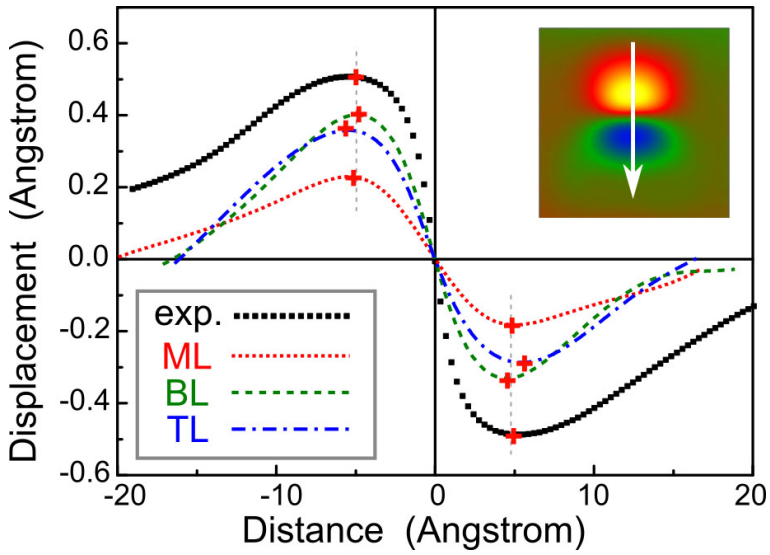


Figure 6.16: Line profile of displacement value of u_{0001} component from experiment and four cluster models: monolayer, bilayer, and trilayer, as indicated in upper inset.

coherently strained GdN bilayer platelet cluster with a distorted rocksalt crystal structure.

6.4.3. Energetics of GdN platelet clusters

We turn now to the question of why these clusters have a particular, well-defined size. Our previous experimental results established that the distribution of GdN cluster sizes is quite homogeneous[174]. To gain further experimental insight, we carried out *ex-situ* annealing experiments up to 1000°C for 20 min on a sample with Gd concentration of 3×10^{19} cm^{-3} (the typical growth temperature is about 800°C). Subsequent TEM imaging revealed no significant change in the lattice distortions or cluster size; that is, coarsening of the clusters did not occur. This means that the observed cluster configuration is stable once formed and that coarsening is negligible at least up to our annealing temperature and time—already much higher and longer than used for the growth itself. This suggests that the observed cluster size represents a minimum-energy configuration. Our theoretical modeling, to which we now turn, confirms this.

We used DFT to determine the optimal size of GdN platelet clusters embedded in GaN. To make the calculations feasible we restricted our attention to $N \times 2$ clusters and considered the range $N = 2$ to 9. The clusters were embedded in GaN supercells containing 384 atoms, with approximate dimensions 67 Å $\times$ 21 Å, sufficiently large to make the interaction across supercells negligible. The clusters were fully relaxed and the formation energy was defined in the usual way,

$$E_f = E_t - n_{\mathrm{Gd}}\,\mu_{\mathrm{Gd}} - n_{\mathrm{Ga}}\,\mu_{\mathrm{Ga}} - n_{\mathrm{N}}\,\mu_{\mathrm{N}}, \tag{6.9}$$

where E_t is the total energy of the supercell containing n_{Gd} atoms of Gd, n_{Ga} atoms of Ga, and n_{N} atoms of N. Assuming thermodynamic equilibrium with bulk GdN and bulk

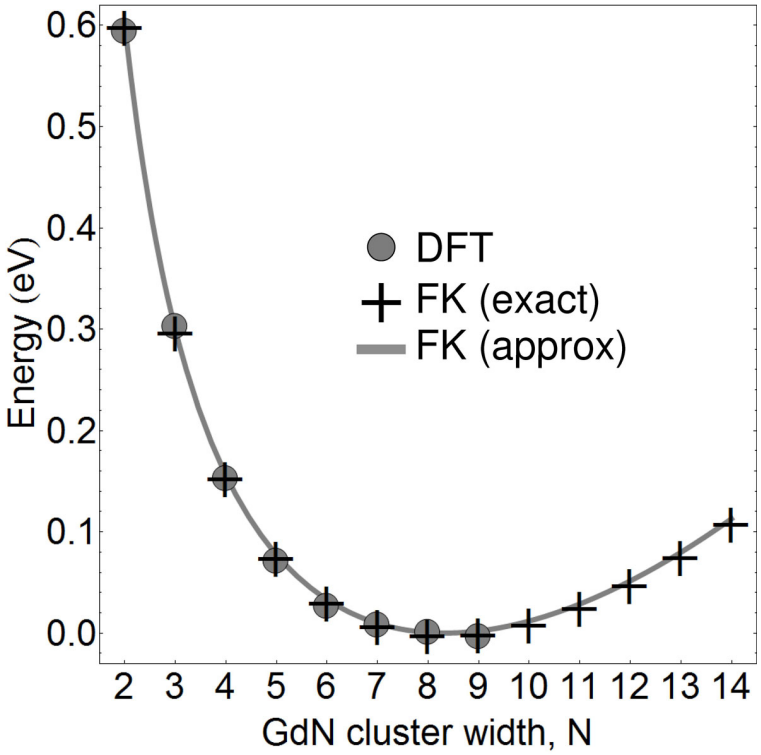


Figure 6.17: Theoretical formation energies per Gd for relaxed $N\times 2$ platelet clusters of GdN in GaN. Formation energies from density-functional theory (circles) are given by Eq. 6.13. Energies for the Frenkel-Kontorova model are given either by Eq. 6.14 evaluated with the exact strain field of Eq. 6.19 (crosses) or by the approximate continuum solution of Eq. 6.16 (curve). [Courtesy DFT calculations by Steven C. Erwin @ Naval Research Laboratory in Washington D.C.]

GaN, the chemical potentials must satisfy the constraints

$$\mu_{\mathrm{Gd}} + \mu_{\mathrm{N}} = \Delta H_f^{\mathrm{GdN}} \tag{6.10}$$

$$\mu_{\mathrm{Ga}} + \mu_{\mathrm{N}} = \Delta H_f^{\mathrm{GaN}}, \tag{6.11}$$

where $\Delta H_f^{\mathrm{GdN}}$ and $\Delta H_f^{\mathrm{GaN}}$ are the heats of formation of GdN and GaN. An additional constraint,

$$n_{\mathrm{Gd}} + n_{\mathrm{Ga}} = n_{\mathrm{N}}, \tag{6.12}$$

arises because each Gd atom takes the place of exactly one Ga atom. These constraints imply that the formation energy is independent of the chemical potentials and can be written simply as

$$E_f = E_t - n_{\mathrm{Gd}}\,\Delta H_f^{\mathrm{GdN}} - n_{\mathrm{Ga}}\,\Delta H_f^{\mathrm{GaN}}. \tag{6.13}$$

To determine the optimal value of N we reason as follows. Consider a system consisting of a fixed number $2M$ of Gd atoms in a fixed volume of GaN. The Gd atoms are arranged into clusters of size $N\times 2$. The total number of such clusters is therefore M/N. The formation energy of each individual cluster is $E_f(N)$ and hence the formation energy of the system is $(M/N)E_f(N)$. Thus, minimizing the formation energy of the system with fixed M is equivalent to minimizing the formation energy per Gd atom, $E_f(N)/N$.

Figure 6.17 shows the fully relaxed DFT formation energies per Gd for values of N

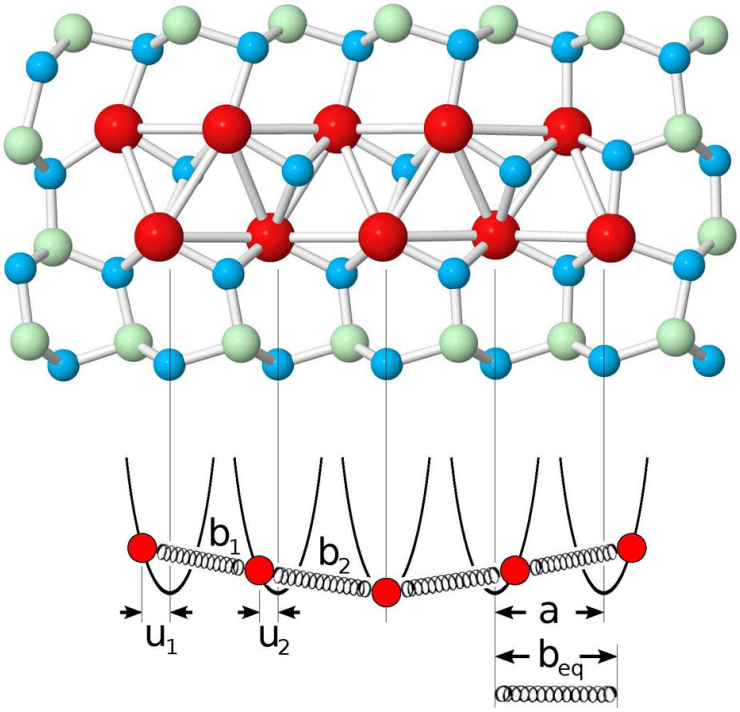


Figure 6.18: One-dimensional Frenkel-Kontorova model representing the most important contributions to the formation energies of GdN bilayer platelet clusters in GaN. See text for definitions. [Courtesy Frenkel-Kontorova modeling by Steven C. Erwin @ Naval Research Laboratory in Washington D.C.]

in the range 2 to 9. The energies decrease with increasing N and then flatten out at the upper end of the range. Hence from these results alone it is unclear where exactly the minimum is located.

To help determining the minimum, and to better understand more generally the origin of these results, we turn to a simple one-dimensional Frenkel-Kontorova (FK) model[175]. The FK model captures the competition between the energy gain from forming larger GdN clusters (which arises because GdN has a larger cohesive energy[176,177] than GaN) and the energy penalty due to the lattice mismatch between GdN and GaN.

The model is illustrated schematically in Fig. 6.18 for the example of a 5×2 platelet. The model only treats one dimension, the width N of the cluster. The Gd atoms are represented by five points (red circles). Their positions deviate from the ideal GaN lattice sites (thin black lines) because of the lattice mismatch between GdN and GaN. The energy penalty arises from two sources: the deviations u_i of the Gd positions from ideal GaN lattice sites, and the deviations of the GdN bond lengths b_j from their ideal equilibrium value b_{eq}. Two spring constants, k and K, give the energy associated with these strains. The energy gain arises because $N-1$ new bonds are formed (E_b) when a cluster is created. Thus the FK total energy is

$$U_N = \left[\sum_{i=1}^{N} \frac{1}{2} k u_i^2 + \sum_{j=1}^{N-1} \frac{1}{2} K (b_j - b_{\mathrm{eq}})^2\right] - (N-1)E_b. \tag{6.14}$$

One possible way to use the FK model is to first determine the model parameters

individually from experiment, DFT calculations, or a combination of both. Instead, we choose to treat the five FK parameters (a, b_{eq}, k, K, E_b) as fitting parameters to be determined by a least-squares fit to the DFT formation energies. Prior to performing this fit we must first solve Eq. 6.14 for the displacement field u_i that minimizes U_N.

We begin by considering a very simple limit, $k \ll K$, which not only gives insight into the exact solution but also turns out to be an extremely good approximation for the GaN:GdN system. In this limit the bond lengths all assume their equilibrium value: $b_j = b_{\mathrm{eq}}$. This immediately gives the result that the displacements are linear in the distance from the center of the cluster:

$$u_i = \left[i - \frac{N+1}{2} \right] fa, \tag{6.15}$$

where $f = (b_{\mathrm{eq}} - a)/a$ is the lattice mismatch. Using this solution we can perform the summation in Eq. 6.14 analytically to obtain

$$U_N^{\mathrm{approx}} = \frac{N(N^2-1)kf^2a^2}{24} - (N-1)E_b. \tag{6.16}$$

This approximate expression correctly exhibits the competition mentioned earlier: as N increases, the internal bond-energy gain (the second term) is eventually overwhelmed by the strain-energy penalty (the first term).

The exact solution for the displacement field can be obtained most easily from the condition that in equilibrium the force F_i on every atom is zero. For the interior atoms,

$$F_i = 0 = K[-(2+\kappa)u_i + u_{i+1} + u_{i-1}] \tag{6.17}$$

where the ratio of spring constants is denoted as $\kappa = k/K$. The force on the two end atoms has a similar form but with an extra term fa. In the limit of large N these finite-difference equations correspond to a differential equation whose solution is

$$u(x) = A \sinh\left(\frac{x}{a/\sqrt{\kappa}}\right). \tag{6.18}$$

For arbitrary N we solve the system of linear force equations to obtain an exact expression for u_i. Because the equations are linear the displacements are rational functions,

$$u_i = (P_i/Q_N)fa, \tag{6.19}$$

where P_i and Q_N are polynomials in κ with integer coefficients. They are most naturally expressed as recurrence relations. The denominator has the form

$$Q_N = (2+\kappa)Q_{N-2} - Q_{N-4} \tag{6.20}$$

where the first four values are $Q_0 = 2$, $Q_1 = 1$, $Q_2 = 2+\kappa$, $Q_3 = 1+\kappa$. The form of the numerator depends on the atom index i. For the atom adjacent to the central atom, $P_i = 1$. For the atom at the right end of the cluster, P_i has the same recurrence relation as Q_N but with different initial values: $P_0 = -1$, $P_1 = 0$, $P_2 = 1$, $P_3 = 1$. Taking as an example the $N = 5$ cluster shown in Fig. 6.18, the atom at the right end has the equilibrium displacement

$$u_5 = \frac{2+\kappa}{1+3\kappa+\kappa^2} fa, \tag{6.21}$$

which indeed reduces to $u_5 = 2fa$ when $\kappa \to 0$, as stated above.

We now use the exact displacement field to evaluate the exact FK total energy U_N and, by fitting to the DFT formation energies, to determine the FK parameters. Two fits are shown in Fig. 6.17, one using the exact U_N and one using the approximate continuum result U_N^{approx}. Both reproduce the DFT energies extremely accurately throughout the entire range. Moreover, the FK results now make clear that the optimal bilayer platelet size is $N = 8$, which is in reasonable accord with our conclusions from the TEM displacement field maps.

The parameters obtained for the FK model also provide insight into the relative magnitudes of the different energy terms, and hence to the physical origin of the optimal GdN cluster size. In particular, the fit to DFT energies gives the ratio of spring constants as $k/K = 0.0004$. This very small value explains why the continuum approximation U_N^{approx} is nearly as accurate as the exact energies. More importantly, it reveals that the GdN largely maintains its own lattice constant and, as a consequence, that only the first term and last terms in Eq. 6.14 are significant. The first term can be regarded as representing the interfacial strain between the GdN and GaN. The energy penalty from this strain eventually overwhelms the energy gained by forming new internal GdN bonds. The crossover between these competing contributions ultimately determines the equilibrium size of GdN bilayer platelet clusters in GaN.

6.5. Summary

The occurrence of small platelet-like GdN clusters in GaN:Gd samples with Gd concentration in the range of 10^{16} cm^{-3} to 10^{19} cm^{-3} have been identified by a combination of several (S)TEM techniques and simulations. With growing Gd content by three orders of magnitude, the number density of clusters is increased and the average cluster distance is reduced concomitantly. Astonishingly, the mean cluster diameter grows only from about 2 nm to 4 nm, although the formation process does not seem to be kinetically restricted. Our results pointed to the critical role of the generally existing phase separation trend in real dilute magnetic semiconductor systems, and the formation of three dimensional clusters as a result of nano-decomposition, even in the case of very low concentrations of the dilute element and/or growth conditions which can be far from thermodynamic equilibrium.

We used TEM and DFT to determine the typical size, atomistic structure, formation mechanism and energetics of GdN clusters in Gd-doped GaN, a magnetic semiconductor grown by MBE. We find that the GdN clusters have a slightly distorted rocksalt structure, and are oriented as bilayer platelets with respect to the GaN basal planes. The results of DFT total-energy calculations, in conjunction with a Frenkel-Kontorova model of the platelet energetics, predict a platelet width of approximately eight atoms—quite close to the size obtained by comparing displacement maps from HRTEM images to those from atomistic models. The consistency between the total-energy predictions and the displacement-map predictions leads to the conclusion that the GdN clusters form as a compromise between a gain in internal cohesive energy and a penalty from interfacial strain energy due to lattice mismatch between the GdN cluster and GaN host. We recommend that future theoretical investigations into the magnetic semiconductor properties of Gd-doped GaN take these findings into account.

Chapter 7.

Conclusion and outlook

The central theme of the work described in the thesis is to clarify the distribution and understand the mechanism controlling the distribution of Gd atoms in Gd-doped GaN with Gd concentration in the range of 10^{16}–10^{19} cm^{-3}. **The objective is threefold**: (1) to identify whether Gd-related clusters exist or not in GaN:Gd samples using transmission electron microscopy (TEM); (2) to determine the local cluster structure on the atomic level; (3) to understand the formation mechanism of such clusters based on the structure information.

Unlike other material systems with miscibility gap that show clustering or phase separation in a concentration range of few percent, the current topic has its characteristics: (1) an extremely huge miscibility gap[28] and (2) the extremely low Gd concentration range[16,17]. The identification and determination of the clusters is highly challenging. It requires on the experimental side a solid knowledge about the contrast formation mechanism in TEM to correctly interpret the results, as presented in Chapter 3–4. And it requires on the material side a deep understanding in the microstructure and its evolution, interaction during the dynamic process of epitaxial growth and/or post growth treatment. This is discussed in Chapter 2. In conjunction with these fundamental knowledge and practical discussions, cases studies of clusters in other group III-V semiconductor materials, of which the size and density are apparently detectable, are investigated in Chapter 5 in order to help better understanding of the materials and techniques.

In spite of the difficulties, the occurrence of GdN clusters is finally identified with fidelity and their detailed structure is determined. The insight has been obtained in Chapter 6 that the formation of the clusters is a compromise between the gain in cohesive energy from forming large GdN clusters and the penalty from interfacial strain energy due to lattice mismatch between the GdN cluster and GaN host. Therefore, in general, we evaluate that **the objectives have been achieved**, as concluded in the following:

- We have discussed in detail the limitations and application of GPA method to the study of embedded nano-clusters. The accuracy of displacement measurement is of an order of magnitude higher than the measurement of strain. The discussions of artifacts of GPA have lead to a practical guide line of the experimental procedures and analysis.

- We have studied in detail the observation and analysis of nano-clusters embedded in semiconductor epilayers with TEM techniques. The results serve as a guide line to study embedded nano-clusters in general.

  - We have presented and discussed results from examples of various imaging modes and suggested the appropriate way to interpret the contrast from coherent, semi-coherent and incoherent embedded nano-clusters.

  - We compared the results of specific Gd-doped GaN models via structure modeling methods which are based on VFF and on DFT. In general, if we assume DFT methods provide the most accurate local structure, the VFF methods can provide very close results.
  - We have discussed the detectability of Gd atoms in Gd-doped GaN models. We have shown that indeed the detection of single Gd atom in GaN is almost impossible while there is still chance to detect them by the local strain field if they were clustered.

- In the case study of InAs nano-clusters embedded in Si, we have shown a static view illustrating the analysis of interface character by HRTEM imaging and analysis techniques and the local strain by GPA and. These results reveal the interplay between strain energy and interface energy in the determination of the shape of clusters.

- In the other case study, we have performed a systematic TEM study about formation and phase transformation of Bi-containing clusters in annealed GaAsBi epilayers. The results showed "snapshots" of different stages of cluster development and provide valuable information to understand the formation of Bi-rich Ga(As, Bi) clusters and their phase transformation to rh-Bi clusters.

- Finally, the occurrence of small GdN clusters have been identified in GaN:Gd thin film samples with Gd concentration in the range of 10^{16} cm^{-3} to 10^{19} cm^{-3} by combination of complementary TEM imaging techniques, simulations and modeling. The formation and the corresponding energetics of the homogeneous GdN clusters in GaN:Gd samples is studied.
  - Dark-field strain contrast imaging identified coffee bean contrast feature in the Gd contained layer, suggesting platelet object that produces strong lattice distortion along GaN [0001] direction.
  - With HRTEM imaging and quantitative analysis by GPA, the displacement of the distorted lattice have quantitatively determined.
  - Dynamic strain contrast calculations taking the measured displacement field as an input parameter attained an excellent agreement with the experimental diffraction contrast results.
  - The appearance of platelet-shaped GdN clusters is also verified by chemically sensitive Z-contrast scanning TEM imaging with off zone-axis geometry.
  - The statistical analysis of cluster number density indicates that most of the Gd atoms are incorporated in the GdN clusters. With growing Gd content by three orders of magnitude, the number density of clusters is increased and the average cluster distance is reduced concomitantly. The mean cluster diameter grows only slightly.
  - Comparing the maps of displacement field between experiments and modeled structures, the local atomic configuration is achieved: the clusters are of platelet distorted rocksalt structure with the broader face parallel to the (0001) basal planes of the wurtzite GaN matrix.

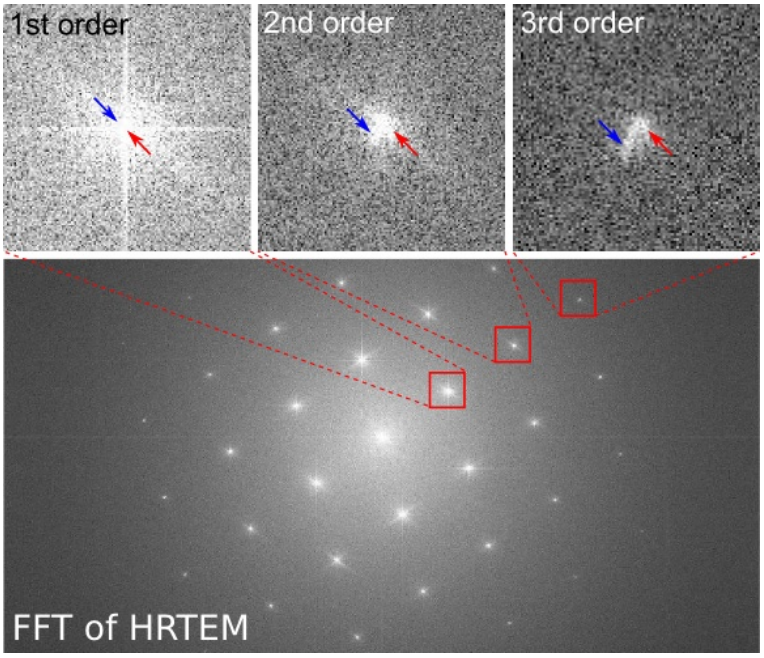


Figure 7.1: Possibility of combining higher order Bragg reflections to measure strain to higher accuracy. The Fourier transform of a cross-section HRTEM image of a strained epilayer which contain an average of 0.5% psudomorphic elastic strain (measured by X-ray diffraction) compared to the substrate. The split of the Bragg spots between substrate and epilayer (pointed by the red and blue arrows) become more and more obvious. This information can in principle be used to refine the geometric phase retrial and therefore provide higher accuracy of displacement and strain measurement.

- We used DFT in conjunction with Frenkel-Kontorova model to explore the minimum energy configuration of GdN cluster configuration. The results predicted that the bilayer platelet with lateral size is in reasonable agreement with our conclusions from the analysis of TEM results.
- The good agreement among the DFT results, the approximate and the exact Frenkel-Kontorova model revealed that the formation of homogeneous GdN clusters is a compromise between the gain in cohesive energy from forming large GdN clusters and the penalty from interfacial strain energy due to lattice mismatch between the GdN cluster and GaN host.

Outlook

At this point, it is appropriate to point out some questions that could become the focus for further studies, or guidelines for similar studies.

The investigation of nano-objects embedded in a matrix relies on certain techniques. The development of techniques is always important for further practical studies. In terms of the GPA method, the strain measurement results are still quite noisy. The accuracy of strain measurement by normal GPA algorithm, which uses only the low order Bragg reflections, can be about 0.5% (absolute scale), which is below the requirements in material science. We noticed in lots of our studied images, the reflections of the higher order contain more delicate variations of phase, as is illustrated in Fig. 7.1. This is very similar to

the the so-called phase multiplication phenomenon, well known in optical interferometry[178] Furthermore, it has been demonstrated recently the enhancement of strain and displacement measurement by "multiple phase Frouier" approach on psudoperiodic lattice image acquired by atomic force microscopy[179]. Similarly, the use of multiple Bragg reflections integrating to the GPA algorithm may allowing for more sensitive measurements of displacements and strains.

Although (S)TEM could provide multiple information of specimen at multiple spatial scale and phase space, the studies of nano-clustering using TEM currently are still limited to the snapshots of the final product of nano-clusters. One of the most important fundamental process of cluster formation, namely, the diffusion process is still missing. Therefore, it would be very interesting to investigate such process either theoretically or experimentally. One approachable way so far of studying the dynamics of cluster formation would be to design temperature and time series of annealing samples and collect more data on the composition and size information of clusters of each snapshot using TEM. Based on the data, we can then modeling the possible basic diffusion process with established equations and then extract the fundamental physical information in terms of fitting parameters. For such kind of study, a model system is preferred with considerable cluster density and detectable size that can form within reasonable processing temperature and time. In this sense, the case study of Bi-containing clusters in GaAs would be ideal choice to carry out such studies. The Z-number ratio, atomic size difference are very high which make the observation more straight forward. It would be even more fantastic to be able to observe the cluster formation dynamics *in-situ* in a microscope.

Concerning our findings of GdN platelet clusters in Gd-doped GaN, we recommend that future theoretical investigations into the magnetic semiconductor properties of the material take these findings into account.

Appendices

Appendix A.

Miscellanenous

A.1. Molecular beam epitaxy

Molecular beam epitaxy (MBE) is one of several methods of depositing crystal materials. It was invented in the late 1960s at Bell Telephone Laboratories by J. R. Arthur and Alfred Y. Cho[180]. After the development over the past half century, MBE is now widely used especially in laboratories in the manufacture of semiconductor devices. Comparing with other epitaxial growth techniques, e.g. liquid phase epitaxy and vapor phase epitaxy, MBE shows unique advantages, most importantly the precise control of thickness (accuracy down to monolayer), ability to produce extremely abrupt interface and the progressive smoothing of the growing surface for most substrate orientations.

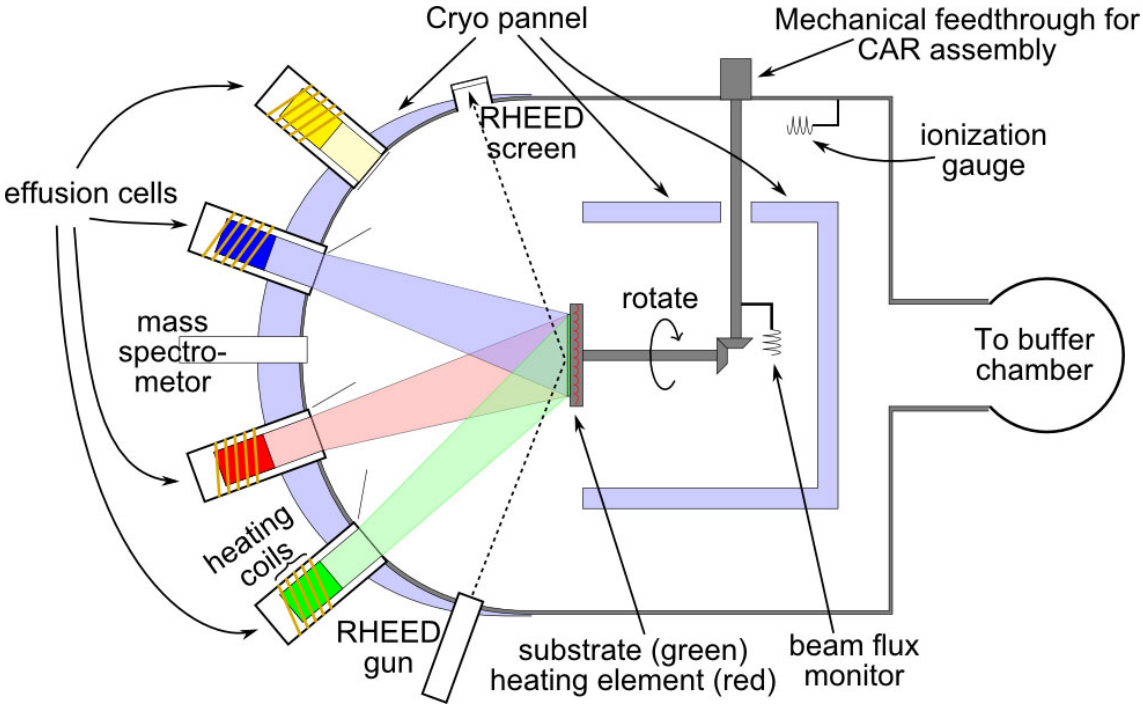


Figure A.1: Schematic illustration of the molecular beam epitaxy chamber.

Figure A.1 schematically illustrates a typical MBE growth chamber. The main feature of the MBE growth chamber is it operates under an ultra high vacuum (UHV) at about 10^{-10}–10^{-8} Pa with the help of both an ion pump and a turbo-molecular pump. It is normally equipped with a certain number of effusion cells that contain a heater to evaporate source material, allowing the growth of a variety of materials. Because of UHV, the pressure is low enough so that the atoms or molecules go from their source to the substrate

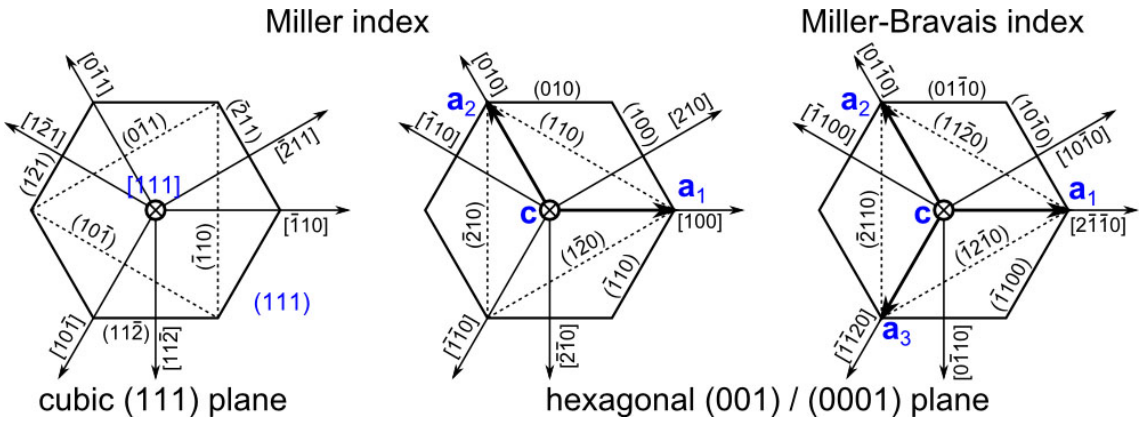


Figure A.2: Miller index and Miller-Bravais index

without collision (i.e. high value of mean free path): they form molecular beam. The growth in MBE is generally conducted far from the thermodynamic equilibrium state and governed mainly by the kinetics of surface process[181]. Reflective high energy electron diffraction (RHEED) can be utilized[182] to monitor the *in-situ* evolution during the growth process.

The basic principle of epitaxial growth is that atoms on a clean surface are free to move around until they find a correct position on the crystal lattice surface to bond. The most important kinetic process is the adsorption and desorption of ad-atoms on the surface, and the surface diffusion of ad-atoms. For the production of one atomic layer, ultrapure elements are delivered to the substrate as a molecular beam made by the effusion cell. The atoms or molecules then deposit on the surface bonding with ad-atoms. Beam fluxes of each effusion cell may be turned on and off rapidly with a shutter, which enables to precisely control the layer compositions.

A.2. Miller and Miller-Bravais indices conversion

The Miller index for indexing the lattice planes is defined via a point in the reciprocal lattice, or as the inverse intercepts along the lattice vectors, denoted as (hkl). The direction, denoted as $[uvw]$ is defined as a vector starts at origin and ends at $u\mathbf{a}_1 + v\mathbf{a}_2 + w\mathbf{a}_3$, where $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ are the basis vectors. In cubic system, it is very convenient because the plane normal is with the same direction index, for example, (110) plane normal is $[110]$.

However, in hexagonal crystal system along the [001] direction (or, c-axis) or rhombohedral crystal system along the [111] direction, the six-fold symmetry for the stacking layer is invisible in the Miller index. By introducing a redundant index $i = -h - k$, the Miller-Bravais index (hkil) inherited the nature of orthogonal to the last dimension while emphasized the in-plane six-fold symmetry. For example, (100) and $(1\bar{1}0)$ are the symmetric planes but not able to see in the index, whereas the Miller-Bravais index $(10\bar{1}0)$ and $(1\bar{1}00)$ the nature of the six-fold symmetry is visible. The two indexing system is schematically depicted in Fig. A.2.

For hexagonal unit cell, the Miller index and Miller-Bravais index can be inter con-

Table A.1: Equivalent Miller index and Miller-Bravais index for typical directions and planes

Miller index	Miller-Bravais index	Mineralogy term
[100], [110] and [010]	$[2\bar{1}\bar{1}0]$, $[11\bar{2}0]$ and $[\bar{1}2\bar{1}0]$	directions
[210], [120] and $[\bar{1}10]$	$[10\bar{1}0]$, $[01\bar{1}0]$ and $[\bar{1}100]$	
[101], [111] and [011]	$[2\bar{1}\bar{1}3]$, $[11\bar{2}3]$ and $[\bar{1}2\bar{1}3]$	
[211], [121] and $[\bar{1}11]$	$[10\bar{1}1]$, $[01\bar{1}1]$ and $[\bar{1}101]$	
[001] and $[00\bar{1}]$	[0001] and $[000\bar{1}]$	
$(2\bar{1}0)$, (110) and $(\bar{1}20)$	$(2\bar{1}\bar{1}0)$, $(11\bar{2}0)$ and $(\bar{1}2\bar{1}0)$	a-plane / 1st-order prism
(100), (010) and $(\bar{1}10)$	$(10\bar{1}0)$, $(01\bar{1}0)$ and $(\bar{1}100)$	m-plane / 2nd-order prism
(001) and $(00\bar{1})$	(0001) and $(000\bar{1})$	c-plane / basal

verted by the following relations, for directions:

$$[uvw] \rightarrow [(2u-v),(-u+2v),(-u-v),3w], \tag{A.1}$$

$$[uvtw] \rightarrow [(u-t),(v-t),w], \tag{A.2}$$

and for planes:

$$(hkl) \rightarrow \text{plane normal:}[hk(-h-k)l], \tag{A.3}$$

$$(hkil) \rightarrow (hkl). \tag{A.4}$$

The most frequently referred directions and planes are summarized in Table A.1 for both Miller index and Miller-Bravais index; their special names usually appear in mineralogy is also included.

In a more general case, the Miller-Bravais index can also be applied to rhombohedral crystal system when it is referenced in hexagonal coordinate system, the details are outlined by Frank[153].

A.3. Crystal structure of rhombohedral As and Bi

The ground state single crystal arsenic and bismuth is rhombohedral (rh) structure with space group R3m(160). The crystal structure in both hexagonal and rhobohedral representation is schematically show in Fig. A.3. In addition, the crystal structure, lattice parameters and the distances of major low index planes (in hexagonal representation) are summarized in Tab. A.2. For convenience, the corresponding parameters of zinc-blend GaAs is also included in the table.

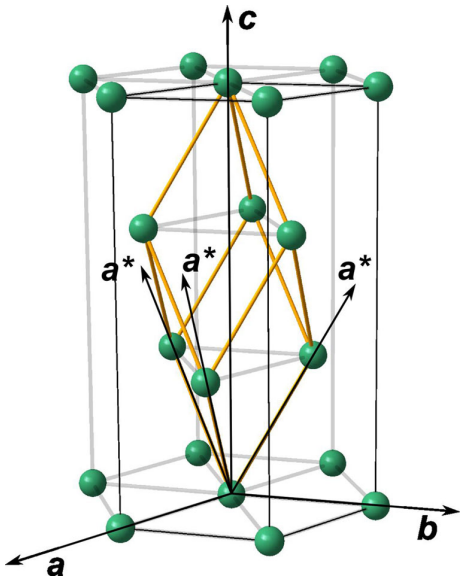


Figure A.3: Crystal structure of rhombohedral arsenic and bismuth, in hexagonal (lattice parameters $a = b \neq c$, $\alpha = 90°$, $\beta = \gamma = 120°$) and rhombohedral ($a = b = c = a^*$, $\alpha = \beta = \gamma = \alpha^* \neq 90°$) representation.

Table A.2: Crystal structure, lattice parameters (room temperature) and major low index plane distances (in real space and reciprocal space) of rhombohedral As, Bi (in hexagonal representation), and zinc-blend GaAs

		As[152]	Bi[152]	GaAs[183]
space group		R3m(160)	R3m(160)	F-43m(216)
lattice parameter (cub.)	a (Å)			5.654
lattice parameters (hex.)	a (Å)	3.762	4.546	
$\alpha = 120°$, $\beta = \gamma = 90°$	c (Å)	10.547	11.862	
lattice parameters	a (Å)	4.132	4.745	
rhombohedral	α (°)	54.13	57.23	
lattice plane distances in hexagonal representation (Å/Å^{-1})	{003}/{0003}	3.52/0.284	3.95/0.253	
	{101}/{10$\bar{1}$1}	3.11/0.321	3.95/0.253	
	{012}/{01$\bar{1}$2}	2.77/0.360	3.28/0.305	
	{104}/{10$\bar{1}$4}	2.05/0.488	2.37/0.422	
	{110}/{11$\bar{2}$0}	1.88/0.532	2.27/0.440	
	{015}/{01$\bar{1}$5}	1.77/0.565	2.03/0.492	
lattice plane distances (Å/Å^{-1})	{111}			3.26/0.306
	{002}			2.83/0.354
	{220}			2.00/0.500

Appendix B.

TEM specimen preparation

TEM specimen preparation is of essential importance to the TEM investigation. The availability and quality of TEM images strongly rely on a good specimen preparation. The major challenge in specimen preparation is to obtain electron transparent thin foil of specimen while minimizing damage and artifacts. For the studies of nano-scaled clusters buried in semiconductor thin films carried out in this thesis, a broad homogeneous region with thickness comparable to the radius of clusters is highly expected.

Two types of TEM specimen, plan view and cross-section are prepared. Illustrated in Fig B.2 is the geometry of these two types of specimen.

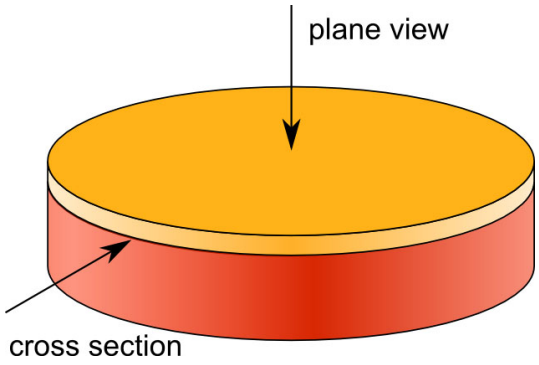


Figure B.1: Scheme of geometry of plane view and cross section specimen out of a as-grown thin film wafer.

B.1. Cross-section specimen preparation

In this thesis, the cross section epilayer specimens are prepared in the following procedure, as schematically depicted in Fig. B.1, and illustrated with photographs in Fig. B.3. In a first step, the wafer is fixed by low melt point wax on glass table and then cut into slice of about 2 mm width with the precision diamond saw [Fig. B.1(a) and Fig. B.3(a)]. Two pieces of cut slice is glued together face to face with G1 epoxy and together with appropriate sized dummy slice glued on the back sides [Fig. B.1(b) and Fig. B.3(b)]. Appropriate amount of epoxy and careful operation is crucial to ensure a very thin and homogeneous slit between two pieces without bubble. The slices sandwich is pressed by a clip stage which is then placed on hot plate at about 120–140) °C for 35–45 min until the epoxy is completely solidified (the color turn to dark brown)[Fig. B.3(c)]. Afterwards, the glued sandwich is inserted in a copper tube (inner diameter $d = 2.2$ mm, outer diameter $d = 3$ mm) of similar length filled with G1 epoxy, which is then baked on the hot plate until totally harden of epoxy [Fig. B.1(c)]. The cylinder is fixed on glass table and cut into slices of thickness about 500 μm [Fig. B.1(d) and Fig. B.3(d)–(e)].

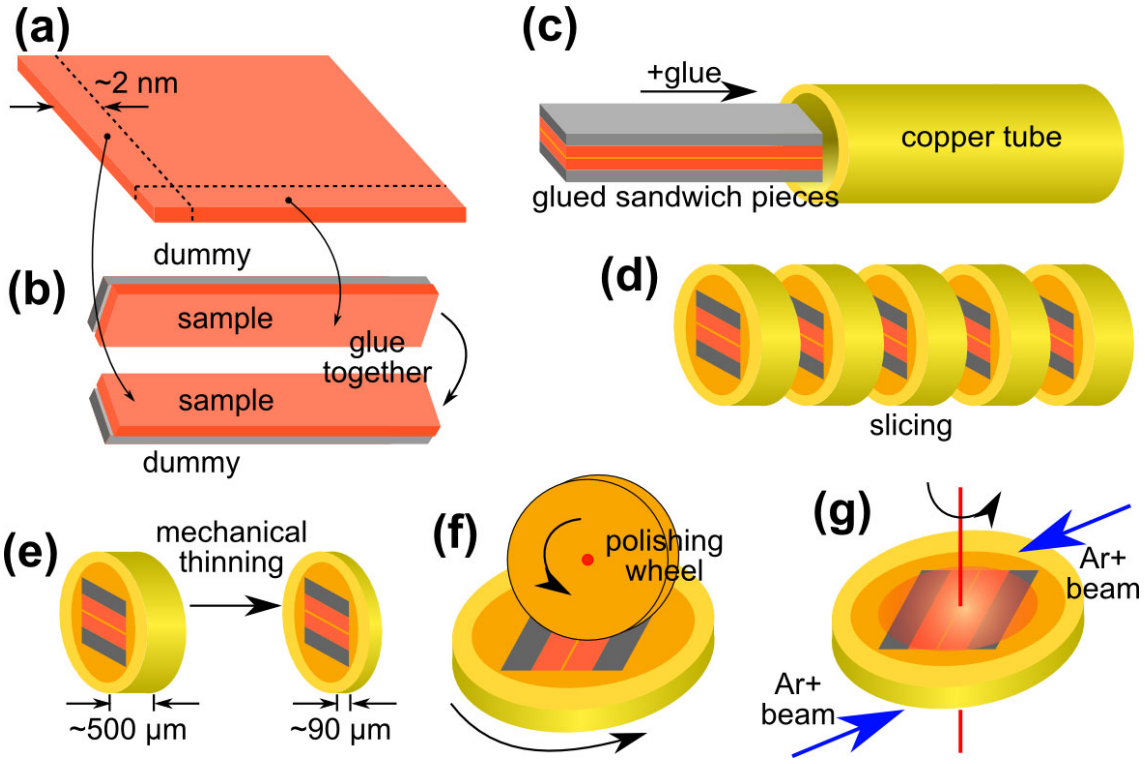


Figure B.2: Scheme of cross-section sample preparation.

The next step is the slice thinning by mechanical grinding against grind paper (SiC grind paper is efficiency enough for soft substrates like Si or GaAs, diamond grind paper is needed for hard materials like sapphire or SiC) down to a thickness of about 90 μm [Fig. B.1(e)]. The Gatan^{TM} 794 grinding set is used. The slice is fixed on either metal or glass cylinder which is then embedded in the grinder [Fig. B.3(f)–(g)]. Grind the surface by progressive finner paper from 15 μm to 3 μm until the surface is flat. Then polish the surface using fine diamond past of 1 μm or $\frac{1}{4}$ μm until no scratch line can be observed under optical microscope (magnification $\times 50$) [Fig. B.3(h)–(i)]. Then flip the slice on the cylinder and grind the other side until a thickness of about 90 μm [Fig. B.3(j)]. A commercial Gatan^{TM} dimple machine is used to further thin the center of slice into a concave [Fig. B.3(k)]. The principal is illustrated in Fig. B.1(f): a polishing wheel is spinning while the slice rotate at a constant velocity. Initially, coarser dimpling slurry (6 μm) and relative slow spin velocity is applied until the center reaches about 30 μm thick. Subsequently fine slurry (1 μm) and high spin velocity is required to ensure a smooth surface and the center of specimen reaches a thickness of about 20 μm. For transparent substrate, the thickness is measured using optical microscope by shift the focus plane. It's important to take into account the refractory index of materials to deduce the real thickness. For opaque substrate, especially for silicon, the dimpling is operated until it become faint transparent with orange color.

The final step is perforation of the dimpled sample by means of ion milling. The specimen after dimpling is transferred with holder [Fig. B.3(l)] to a Gatan precision ion polishing system (PIPS) equipped with two ion guns generating inert (Ar^{+}) gas beams. The two guns are set to shot the specimen from both sides with a incident angle of 3° to 4° [Fig. B.1(g) and Fig. B.3(m)]. Before start, it is always better to check if the ion beam is cen-

Figure B.3: Photographs show the detailed steps of cross-section TEM sample preparation.

tered and the beam energy is optimized, according to the instruction of the machine. The beam accelerate energy is set to 3–4 KeV and current 15-35 μA dependent on the materials. The ion milling process is monitored lively with TV rate camera. When transparent specimen (e.g. sapphire) is thin enough, there will be color rings (so called Newton rings) that formed by the interference of light reflected from specimen top and bottom surface. When rings appear, the thickness is then in the range of light wave length. The accelerate energy is then reduced to about 1.5–2.5 KeV and current 1–8 μA to minimize sputtering damage and at the same time ensure relative broad areas of homogeneous thickness for study. For materials sensitive to bombardment, like $LiAlO_2$, a larger incident beam angle and higher energy is applied first to perforate the specimen fast as possible, then a low energy and small incident angle is applied to remove the amorphous layers at the surface. Typically, the ion milling takes 2 to 3 hours for soft substrates and may last over 5 hours for hard materials.

B.2. Plan view sample preparation

Plan view epilayer specimens are prepared in the following procedure, as schematically depicted in Fig. B.4.

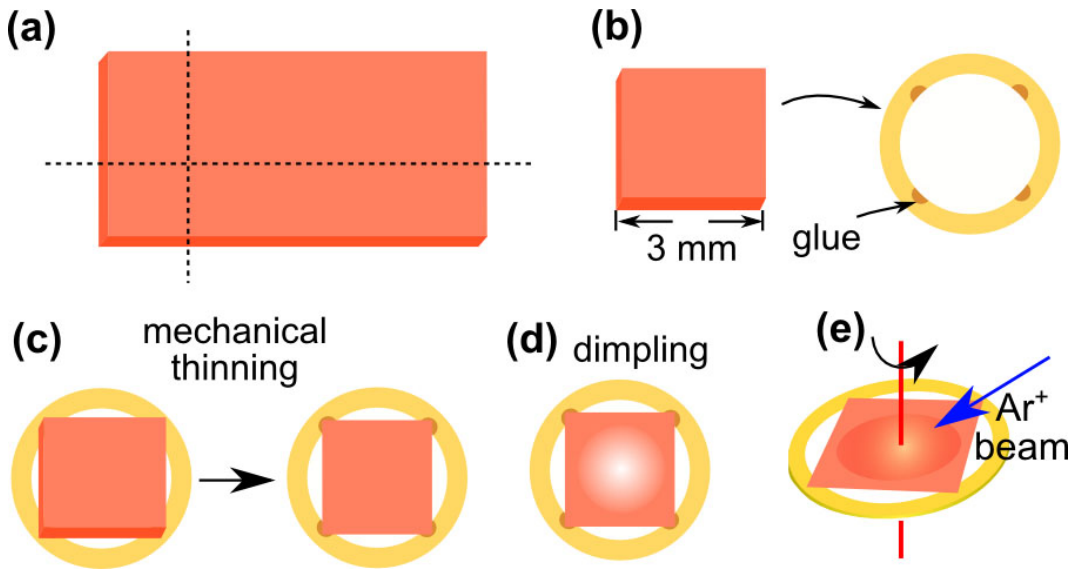


Figure B.4: Scheme of cross-section sample preparation.

Most of the cautious operation are the same with cross-section sample preparation. Firstly, epilayer is cut to 3×3 mm square. Then, glue it onto copper ring. Note the top surface of epilayer is glued onto the ring. Next, mechanic polish the (back side) of the epilayer until about 100 μm. Afterwards, dimple the sample with rolling copper wheel using diamond slurry. Finally, perforate the specimen using ion milling.

Appendix C.

Specifications of the studied samples

C.1. Gd-doped GaN samples

The samples were grown by reactive molecular beam epitaxy (MBE) along the GaN [0001] direction (c-plane samples) or GaN [$10\bar{1}0$] direction (m-plane samples). C-plane samples were realized either on SiC(0001) substrates or on GaN(0001)/Al_2O_3 templates, whereas m-plane layers were deposited on $LiAlO_2$(001) substrates. GaN:Gd films were deposited under similar growth conditions as used for high quality pure GaN layers with this orientation[184,185]. The substrate temperatures were near the Ga-desorption temperature determined during the Ga flash-off procedure, as summarized in Tab. C.1. The Ga and Gd cell temperatures were 1005°C and 900°C, respectively. The chamber pressure was kept at 1.2×10^{-5} Torr during growth. Structural properties before TEM studies were characterized by x-ray diffractometry (XRD). Rutherford backscattering spectroscopy (RBS) and particle-induced x-ray emission (PIXE) using a 2MeV He^+ beam were carried out to determine the impurities content. "Overall" or, average Gd concentration and its composition depth profiles were performed by secondary ions mass spectrometry (SIMS). The parameters of sample growth are summarized in Tab. C.1.

Table C.1: GaN:Gd samples specification: sample number, substrate, Gd concentration as determined by SIMS measuerement, growth temperature and other parameters. Samples are provied by Rocío Ranchal, Paul-Drude-Institut für Festkörperelektronik, Berlin

sample	substrate	[Gd]/cm^{-3}	T_{growth}	comment
M3113	4H-SiC(0001)	9×10^{16}	840°C	co-doped with Si
M3122	6H-SiC(0001)	9×10^{19}	840°C	AlN buffer; in-situ annealed 10 min at 840°C after growth
M3152	$LiAlO_2$(001)	5×10^{16}	–	high density of basal SF
M3187	6H-SiC(0001)	3×10^{19}	810°C	AlN buffer; in-situ annealed 20 min at 830°C after growth
M3198	GaN template	5×10^{16}	810°C	very high layer quality
M3221	GaN template	9×10^{19}	810°C	very high layer quality

It is noted that XRD measurement on all these samples did not able to reveal any sign of secondary phase, precipitate or clustering. Besides, The low Gd doping level of the studied samples is beyond the sensitivity of RBS and PIXE[163]. Nevertheless, the occurrence of small platelet GdN clusters in all these samples are determined by TEM imaging techniques and simulations as will be presented in the following section. Although samples are grown on different substrates, it seems to have no detectable affects on the distribution of Gd atoms (from SIMS depth profile). Difference in substrates do affects the density of extended defects, for instance threading dislocation (TD) density.

Nevertheless, we did not find any correlation between the distribution of GdN clusters and the existing TDs, as will be discussed. Therefore, it is not expected that difference in substrate will eventually affect the data interpretation in the following.

C.2. InAs/Si samples

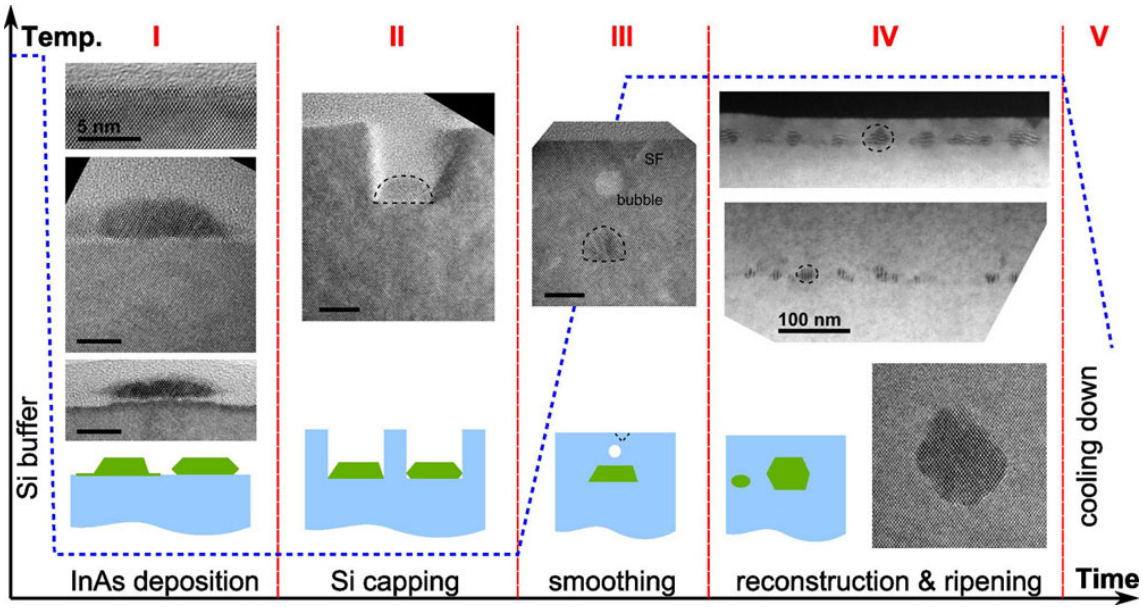


Figure C.1: Scheme of the sample growth steps. (I) InAs deposit under lowered temperature on Si(111) surface and formed flat island; (II) covering of Si under low temperature; (III) elevate substrate temperature to smooth the capping Si; (IV) keep elevated temperature for structure optimization; (V) cooling down.

The samples were grown by molecular beam epitaxy on either Si(001) or 5° off-cut Si(001) substrate. The growth sequence could be summarized in the following five stages, as schematically depicted in Fig. C.1. Initially, a silicon buffer layer is deposited at 600°C prior the deposition of InAs. The temperature is ramped down to 400°C for InAs deposition. Afterwards, a 10–50 nm thick Si capping layer is deposited at the same temperature. Next, the temperature rise again [post growth annealing (PGA)] to smooth the deposited Si capping layer, at the same time, the InAs islands experience morphological and structural changes. Finally the growth chamber cool down to room temperature.

All the samples studied are summarized in Tab. C.2.

C.3. GaAsBi/GaAs samples

The samples were grown by solid-source molecular beam epitaxy on semi-insulating GaAs(100) substrates. Conventional effusion cells were used for Ga and Bi and a two-zone cracker source for As_2. The substrates were heated for 10 min at 620°C to remove the native oxide, after which a $\sim$ 150 nm thick GaAs buffer layer was grown at 580°C. Then, the growth was interrupted and the temperature was ramped down for the growth of the target layer, which comprised 270 nm of GaAsBi. Samples were grown at substrate temperatures (T_g) of 220°C, 270°C and 315°C. During the oxide removal and the buffer

Table C.2: InAs/Si samples specifications: sample number, substrate, nominal InAs coverage (MLs) and post growth annealing (PGA). Samples are provied by Tariq Al-Zoubi, University of Kassel, Germany.

sample	substrate	MLs	PGA	comment
A1367	Si(001)	2	600°C	
A1561	Si(001)5°	2	600°C	
A1564	Si(001)5°	4	600°C	high quality sample preparation
A1565	Si(001)5°	1	600°C	coherent InAs cluster diameter 2–4 nm
A1573	Si(001)	1	600°C	few SF in Si capping layer
A1574	Si(001)	4	600°C	no InAs clusters at all
A1601	Si(001)	2	600°C	10 nm thick Si capping
A1602	Si(001)	2	NO	voids in Si capping where InAs located
A1706	Si(001)	2	–	only surface islands, no capping
A1709	Si(001)	2	NO	Si capping on A1706, voids like A1602
A1710	Si(001)	2	600°C	PGA of A1709
A1716	Si(001)	2	600°C	50 nm Si capping, SF and bubbles in Si
A1717	Si(001)	2	800°C	50 nm high quality Si capping
A1763	Si(001)	2	600°C	temp. up before Si capping, no InAs clusters
A1764	Si(001)	2	600°C	some bubble in Si capping

growth, the temperature was monitored with an optical pyrometer. Since the pyrometer range is limited to $T_s > 400°C$, the temperatures for the GaAsBi layers were monitored with a thermocouple. The thermocouple readings were calibrated by linearly extrapolating the relationship between the pyrometer values and the thermocouple values to the low temperature range. The growth rate was 0.5 mum/h for the GaAsBi estimated by layer thickness measurements performed by XRD. The atomic As/Ga flux ratio was 1.6 for GaAsBi. The flux ratios were determined from ion gauge readings. The nominal Bi flux was the same for all samples. The as-grown samples were cut in 4×4 mm chips, which were annealed in a RTA oven between 600°C and 800°C for 60 s or 120 s in a flowing nitrogen environment. A GaAs proximity cap was used to prevent desorption of As.

The samples studied are summarized in Tab.C.3.

Table C.3: GaAsBi samples specifications: sample number, growth temperature (T_g), target Bi concentration (Bi), As/Ga flux (As/Ga), annealing temperature (T_a) and annealing period (t_a). Samples are provied by Janne Puustinen, Tampere University of Technology, Finland.

sample	T_g	Bi(%)	As/Ga	T_a	t_a	comment
AsN3554.AG	220	1.46	1.6	–	–	no annealing
AsN3554.600.60	220	1.46	1.6	600	60	
AsN3554.600.120	220	1.46	1.6	600	120	
AsN3554.800.60	220	1.46	1.6	800	60	
AsN3556.600.60	315	1.5	1.6	600	60	
AsN3814.600.60	220	4.62	1.2	600	60	
AsN3835.AG	315	$4.2 + 1.6$	–	–	–	double GaAsBi layer

Appendix D.

Dislocation loops associated to Ga(As, Bi) nano-clusters

D.1. Discolation loops in sample A3 and B

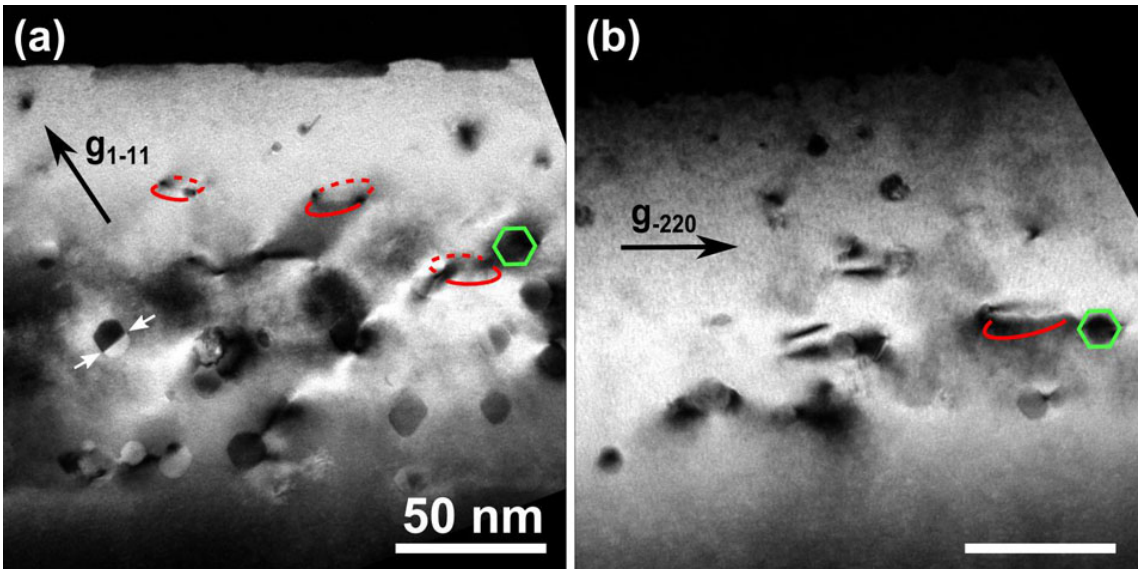


Figure D.1: Diffraction contrast images of sample B taken with (a) $\mathbf{g}_{1\bar{1}1}$ and (b) $\mathbf{g}_{\bar{2}20}$, respectively. The contrast from projected dislocation loops satellite to the clusters are observed: red continue (dash) line sketch the missing part of the loop outside (inside) the paper/foil plane; blue line mark the cluster position.

In sample A3 and B, besides the rh-Bi clusters and transient clusters, we find some dislocation loops (DLs). They are identified by the strain contrast images as shown in Fig. D.1(a) and (b). The contrast strip, as marked above the red lines by the side of the cluster in Fig. D.1(b), is produced by the strain field of a dislocation loop that is buried in the specimen foil. In this cross-section geometry, we only observe a side view along the radius direction of a half loop. While in Fig. D.1(a), a much thinner position, the part of loop inside and outside the foil plane are already removed by preparation, leaving only the edge-on part of the DL in the cross-section view. The red full (dashed) curves schematically illustrate an inclined view of the loop outside (inside) the foil/paper normal that is removed by cross-section specimen preparation. Unfortunately, the full loop is not best resolvable in the geometry of the cross-section specimen. Based on the observation, we deduce that the DLs plane normal is parallel to the [001] growth direction and is perpendicular to the cross-section specimen foil normal. The DL radius can be deduced

from the distance of the edge-on pairs of the dislocation line, which results between 12–18 nm. Such DL were referred as prismatic dislocation loop,[186] which is not usually seen in zincblend structure. Nevertheless, similar contrast of prismatic dislocation loop attached to clusters is also found in annealed LT-GaAs:Sb layers through plan-view and cross-section TEM[63,160,187], where the authors suggested the prismatic dislocation loop had a Burgers vector of $\mathbf{b} = 1/2[001]a_{GaAs}$. The author also suggested the creation of such prismatic dislocation loops were related to strain relieve.

D.2. Strain energy of zb Bi-rich clusters and DLs

A direct interpretation of the Bi content in a cluster of critical size from Z-contrast images is challenging because of the unavoidable overlapping matrix effect. We can nevertheless estimate the Bi content based on experimental observations via strain energy criteria, explained as follows. Because of the lattice misfit between GaAs and (the theoretically predicted) GaBi, the local strain of the clusters will increase with more Bi that diffuse into to the cluster. Beyond a critical size, plastic deformation (formation of dislocation loops) is energetically more favorable. This is similar to the case of misfit dislocation formation in strained epitaxial thin films. Indeed, we observed in sample A3 and B dislocation loops satellite to the clusters. Such satellite dislocations are also observed in other systems showing clusters[63,188]. Hence, with the observed dislocation loop and cluster radius, we can deduce the Bi content in clusters of critical size.

Considering the large volume difference between GaAs matrix and $GaAs_{1-x}Bi_x$ cluster, we assume the matrix is rigid as bulk GaAs. The stress that the $GaAs_{1-x}Bi_x$ clusters suffer can be imagined as a free $GaAs_{1-x}Bi_x$ cluster squeezed in a GaAs hole of the same amount of atoms. Therefore, interpolated from the lattice parameters of GaAs (5.65 Å) and (theoretical value of) GaBi (6.33 Å[136,140]) and assuming the Vegard's law, the interior strain ϵ of $GaAs_{1-x}Bi_x$ is function of x:

$$\epsilon(x) = \frac{a_{GaAs} - a_{GaAsBi}(x)}{a_{GaAsBi}(x)} = \frac{x(a_{GaAs} - a_{GaBi})}{(1-x)a_{GaAs} + xa_{GaBi}}. \tag{D.1}$$

The total elastic energy of a Bi-rich cluster with radius r and strain $\epsilon(x)$ (assuming isotropic elastically) is function of x and r :

$$E_{el}(x,r) = \left(\frac{4}{3}\pi r^3\right) \times \frac{1}{2}E\epsilon^2(x), \tag{D.2}$$

where E is the "equivalent" Young's modulus (see Chapter 2.1.3 in page 9). Regarding the dislocation loop energy E_{DL}, we use the well-established expression for the self-energy of a dislocation loop of the radius R_{DL} and Burgers vector $\mathbf{b}$ (b being the magnitude of $\mathbf{b}$)[189],

$$E_{DL}(R_{DL}) = \frac{Gb^2R_{DL}}{2(1-\nu)} \cdot \ln\frac{8\alpha R_{DL}}{b}, \tag{D.3}$$

in which G is the shear modulus, ν is the Poisson ratio, and $\alpha = 0.25$ accounts for the energy of dislocation core[56]. At the critical cluster size, the elastic energy should equal to the self energy of creating a dislocation loop:

$$E_{el}(x,r) = E_{DL}(R_{DL}), \tag{D.4}$$

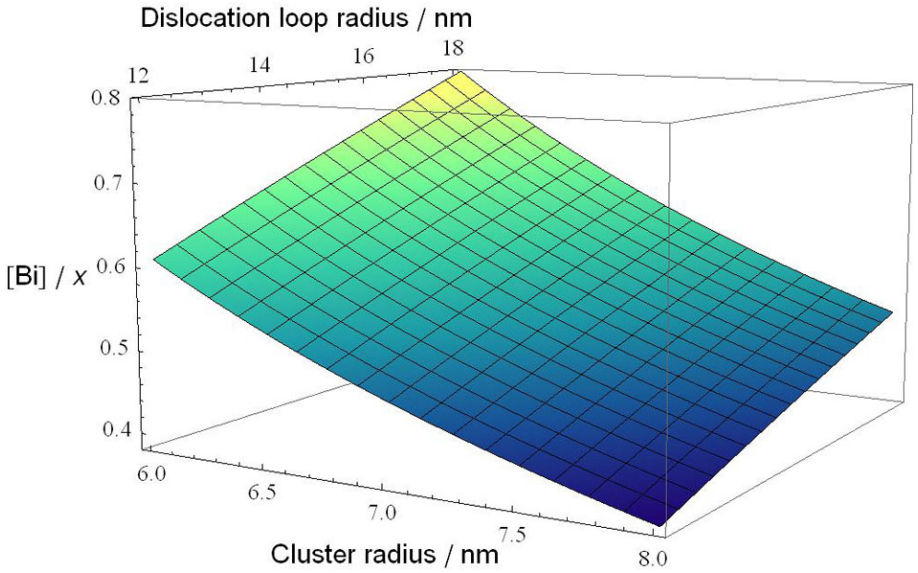


Figure D.2: Bi content calculated from soluving the energy balance equation with cluster radius r and dislocation loops radius R_{DL} from TEM observation as input parameters.

The equation defines a implicit function $x(r, R_{DL})$. Because of the lack of data on the elastic constants for $GaAs_{1-x}Bi_x$, we use that for GaAs instead. Inserting the elastic constants[183] ($C_{11} = 244.6$ GPa, $C_{12} = 59.11$ GPa, $C_{44} = 59.6$ GPa, resulting in isotropic equivalent Young's modulus, shear modulus and Poisson ratio[61] $E = 116.4$ GPa, $G = 47.8$ GPa and $\nu = 0.22$, respectively), the Bi content x in cluster of critical size can be solved from eq. (D.4) as function of cluster size r and dislocation loop radius R_{DL}.

In our TEM observation, the DLs radius ranges 12–18 nm and the radius of cluster of critical size may between 6 and 8 nm. The Bi content x as function of r and R_{DL} solved from the energy balance equation is plotted in Fig. D.2. According to the meaning of x in $GaAs_{1-x}Bi_x$, it should be in the range of $[0, 1]$, as it is apparently revealed (actually in the range $[0.4, 0.78]$) in Fig. D.2. However, the roots of the quadratic equation of x (from the energy balance equation, Eq. D.4), with parameters r and R_{DL} in the observed range, does not necessarily guarantee that x will in the range. The fact that it indeed in the range suggests that the above mentioned energy balance equation and related assumptions to simplify the calculations are physically reasonable. With the calculations, we therefore derived the critical value of Bi content in zb Bi-rich clusters is between 40 and 78 %. Thus the zb Bi-rich clusters will contain a Bi content which is much higher than the value of incorporated Bi into GaAs layers ($x = 22$ %) reported experimentally[137].

Bibliography

[1] Cahn, J. W. *On spinodal decomposition*. Acta metallurgica **9** 795 (1961).

[2] Doherty, R. D. *Diffusive phase transformations in the solid state*. In R. W. Cahn and P. Haasen (eds.), *Physical Metallurgy*, chap. 15, page 1363. North-holland, 4th ed. (1996).

[3] Porter, D. A.; Easteriling, K. E. and Sherif, M. Y. *Phase transformations in metals and alloys*. CRC Press (2008).

[4] Dietl, T. *A ten-year perspective on dilute magnetic semiconductors and oxides*. Nature materials **9** 965 (2010).

[5] Katayama-Yoshida, H.; Sato, K.; Fukushima, T.; Toyoda, M.; Kizaki, H.; Dinh, V. A. and Dederichs, P. H. *Theory of ferromagnetic semiconductors*. Physica Status Solidi (a) **204** 15 (2007).

[6] Ohno, H. *Making Nonmagnetic Semiconductors Ferromagnetic*. Science **281** 951 (1998).

[7] Dietl, T. *Self-organized growth controlled by charge states of magnetic impurities*. Nature materials **5** 673 (2006).

[8] Bonanni, A. *Ferromagnetic nitride-based semiconductors doped with transition metals and rare earths*. Semiconductor Science and Technology **22** R41 (2007).

[9] Bougeard, D.; Ahlers, S.; Trampert, A.; Sircar, N. and Abstreiter, G. *Clustering in a Precipitate-Free GeMn Magnetic Semiconductor*. Physical Review Letters **97** 237202 (2006).

[10] Dhar, S.; Brandt, O.; Trampert, A.; Däweritz, L.; Friedland, K. J.; Ploog, K. H.; Keller, J.; Beschoten, B. and Güntherodt, G. *Origin of high-temperature ferromagnetism in (Ga,Mn)N layers grown on 4H-SiC(0001) by reactive molecular beam epitaxy*. Applied Physics Letters **82** 2077 (2003).

[11] Moreno, M.; Trampert, A.; Jenichen, B.; Däweritz, L. and Ploog, K. H. *Correlation of structure and magnetism in GaAs with embedded Mn(Ga)As magnetic nanoclusters*. Journal of Applied Physics **92** 4672 (2002).

[12] van Schilfgaarde, M. and Mryasov, O. N. *Anomalous exchange interactions in III-V dilute magnetic semiconductors*. Physical Review B **63** 233205 (2001).

[13] Sato, K.; Katayama-Yoshida, H. and Dederichs, P. H. *High Curie Temperature and Nano-Scale Spinodal Decomposition Phase in Dilute Magnetic Semiconductors*. Japanese Journal of Applied Physics **44** L948 (2005).

[14] Mahadevan, P.; Osorio-Guillén, J. M. and Zunger, A. *Origin of transition metal clustering tendencies in GaAs based dilute magnetic semiconductors*. Applied Physics Letters **86** 172504 (2005).

[15] Gonzalez Szwacki, N.; Majewski, J. A. and Dietl, T. *Aggregation and magnetism of Cr, Mn, and Fe cations in GaN*. Physical Review B **83** 184417 (2011).

[16] Dhar, S.; Brandt, O.; Ramsteiner, M.; Sapega, V. F. and Ploog, K. H. *Colossal Magnetic Moment of Gd in GaN*. Physical Review Letters **94** 037205 (2005).

[17] Dhar, S.; Pérez, L.; Brandt, O.; Trampert, A.; Ploog, K. H.; Keller, J. and Beschoten, B. *Gd-doped GaN: A very dilute ferromagnetic semiconductor with a Curie temperature above 300K*. Physical Review B **72** 245203 (2005).

[18] Liu, L.; Yu, P. Y.; Ma, Z. and Mao, S. S. *Ferromagnetism in GaN:Gd: A Density Functional Theory Study*. Physical Review Letters **100** 127203 (2008).

[19] Thiess, A.; Dederichs, P. H.; Zeller, R.; Blügel, S. and Lambrecht, W. R. L. *Superparamagnetism in Gd-doped GaN induced by Ga-vacancy clustering*. Physical Review B **86** 180401 (2012).

[20] Liu, Z.; Yi, X.; Wang, J.; Kang, J.; Melton, A. G.; Shi, Y.; Lu, N.; Wang, J.; Li, J. and Ferguson, I. *Ferromagnetism and its stability in n-type Gd-doped GaN: First-principles calculation*. Applied Physics Letters **100** 232408 (2012).

[21] Mitra, C. and Lambrecht, W. R. L. *Interstitial-nitrogen- and oxygen-induced magnetism in Gd-doped GaN*. Physical Review B **80** 081202(R) (2009).

[22] Mishra, J.; Dhar, S. and Brandt, O. *High defect concentration in GaN:Gd layers grown by reactive molecular beam epitaxy*. Solid State Communications **150** 2370 (2010).

[23] Mishra, J. K.; Singh, B. P. and Dhar, S. *Optical detection of the defects associated with the magnetic properties observed in GaN:Gd layers grown by reactive molecular beam epitaxy*. Applied Physics Letters **99** 072119 (2011).

[24] Zhou, Y. K.; Choi, S. W.; Emura, S.; Hasegawa, S. and Asahi, H. *Large magnetization in high Gd concentration GaGdN and Si-doped GaGdN grown at low temperatures*. Applied Physics Letters **92** 062505 (2008).

[25] Davies, R. P.; Gila, B. P.; Abernathy, C. R.; Pearton, S. J. and Stanton, C. J. *Defect-enhanced ferromagnetism in Gd- and Si-coimplanted GaN*. Applied Physics Letters **96** 212502 (2010).

[26] Roever, M.; Malindretos, J.; Bedoya-Pinto, A.; Rizzi, A.; Rauch, C. and Tuomisto, F. *Tracking defect-induced ferromagnetism in GaN:Gd*. Physical Review B **84** 081201 (2011).

[27] Cheiwchanchamnangij, T. and Lambrecht, W. R. L. *Calculated x-ray linear dichroism spectra for Gd-doped GaN*. Physical Review B **84** 205119 (2011).

[28] Filhol, J. S.; Jones, R.; Shaw, M. J. and Briddon, P. R. *Structure and electrical activity of rare-earth dopants in GaN*. Applied Physics Letters **84** 2841 (2004).

[29] Martínez-Criado, G.; Sancho-Juan, O.; Garro, N.; Sans, J. A.; Cantarero, A.; Susini, J.; Roever, M.; Mai, D.-D.; Bedoya-Pinto, A.; Malindretos, J. and Rizzi, A. *X-ray absorption in GaGdN: A study of local structure*. Applied Physics Letters **93** 021916 (2008).

[30] Ney, A.; Kammermeier, T.; Manuel, E.; Ney, V.; Dhar, S.; Ploog, K. H.; Wilhelm, F. and Rogalev, A. *Element specific investigations of the structural and magnetic properties of Gd:GaN*. Applied Physics Letters **90** 252515 (2007).

[31] Bedoya-Pinto, A.; Malindretos, J.; Roever, M.; Mai, D. D. and Rizzi, A. *Variable range hopping transport in ferromagnetic GaGdN epitaxial layers*. Physical Review B **80** 195208 (2009).

[32] Ney, A.; Kammermeier, T.; Ollefs, K.; Ney, V.; Ye, S.; Dhar, S.; Ploog, K. H.; Röver, M.; Malindretos, J.; Rizzi, A.; Wilhelm, F. and Rogalev, A. *Gd-doped GaN studied with element specificity: Very small polarization of Ga, paramagnetism of Gd and the formation of magnetic clusters*. Journal of Magnetism and Magnetic Materials **322** 1162 (2010).

[33] Suenaga, K.; Tence, M.; Mory, C.; Colliex, C.; Kato, H.; Okazaki, T.; Shinohara, H.; Hirahara, K.; Bandow, S. and Iijima, S. *Element-selective single atom imaging*. Science **290** 2280 (2000).

[34] Allen, J. E.; Hemesath, E. R.; Perea, D. E.; Lensch-Falk, J. L.; Li, Z. Y.; Yin, F.; Gass, M. H.; Wang, P.; Bleloch, A. L.; Palmer, R. E. and Lauhon, L. J. *High-resolution detection of Au catalyst atoms in Si nanowires*. Nature nanotechnology **3** 168 (2008).

[35] Krivanek, O. L.; Chisholm, M. F.; Nicolosi, V.; Pennycook, T. J.; Corbin, G. J.; Dellby, N.; Murfitt, M. F.; Own, C. S.; Szilagyi, Z. S.; Oxley, M. P.; Pantelides, S. T. and Pennycook, S. J. *Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy*. Nature **464** 571 (2010).

[36] Meyer, J. C.; Kurasch, S.; Park, H. J.; Skakalova, V.; Künzel, D.; Gross, A.; Chuvilin, A.; Algara-Siller, G.; Roth, S.; Iwasaki, T.; Starke, U.; Smet, J. H. and Kaiser, U. *Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy*. Nature materials **10** 209 (2011).

[37] Bar-Sadan, M.; Barthel, J.; Shtrikman, H. and Houben, L. *Direct Imaging of Single Au Atoms Within GaAs Nanowires*. Nano letters **12** 2352 (2012).

[38] Zhou, W.; Kapetanakis, M. D.; Prange, M. P.; Pantelides, S. T.; Pennycook, S. J. and Idrobo, J. C. *Direct Determination of the Chemical Bonding of Individual Impurities in Graphene*. Physical Review Letters **109** 206803 (2012).

[39] Zhou, W.; Oxley, M. P.; Lupini, A. R.; Krivanek, O. L.; Pennycook, S. J. and Idrobo, J. C. *Single atom microscopy*. Microscopy and Microanalysis **18** 1342 (2012).

[40] Hÿtch, M. J.; Putaux, J. L. and Penisson, J. M. *Measurement of the displacement field of dislocations to 0.03Å by electron microscopy*. Nature **423** 270 (2003).

[41] Hÿtch, M. J.; Snoeck, E. and Kilaas, R. *Quantitative measurement of displacement and strain fields from HREM micrographs*. Ultramicroscopy **74** 131 (1998).

[42] Tai, C. W. and Lereah, Y. *Quantitative strain-field measurement of 1:1 B-site cation ordered domains and antiphase boundaries in Pb(Sc0.5Ta0.5)O3 ceramics by high-resolution transmission electron microscopy*. Applied Physics Letters **94** 051908 (2009).

[43] Molina, S. I.; Galindo, P. L.; Gonzalez, L.; Ripalda, J. M.; Varela, M. and Pennycook, S. J. *Exploring semiconductor quantum dots and wires by high resolution electron microscopy*. Journal of Physics: Conference Series **209** 012004 (2010).

[44] Gargallo-Caballero, R.; Luna, E.; Ishikawa, F. and Trampert, A. *Strain-induced composition limitation in nitrogen d-doped (In,Ga)As/GaAs quantum wells*. Applied Physics Letters **100** 171906 (2012).

[45] Wölz, M.; Lähnemann, J.; Brandt, O.; Kaganer, V. M.; Ramsteiner, M.; Pfueller, C.; Hauswald, C.; Huang, C. N.; Geelhaar, L. and Riechert, H. *Correlation between In content and emission wavelength of In(x)Ga(1-x)N/GaN nanowire heterostructures*. Nanotechnology **23** 455203 (2012).

[46] Guerrero, E.; Galindo, P. L.; Yanez, A.; Pizarro, J.; Guerrero-Lebrero, M. P. and Molina, S. I. *Accuracy assessment of strain mapping from Z-contrast images of strained nanostructures*. Applied Physics Letters **95** 143126 (2009).

[47] Reyes, D. F.; González, D.; Batiman, F.; Dominguez, L.; Hunter, C. J.; Guerrero, E.; Rolden, M. A.; Mayoral, A.; David, J. P. R. and Sales, D. L. *Photoluminescence Enhancement of InAs(Bi) Quantum Dots by Bi Clustering*. Applied Physics Express **6** 042103 (2013).

[48] Douin, J.; Donnadieu, P. and Houdellier, F. *Elastic strain around needle-shaped particles embedded in Al matrix*. Acta Materialia **58** 5782 (2010).

[49] Rosner, H.; Koch, C. T. and Wilde, G. *Strain mapping along Al–Pb interfaces*. Acta Materialia **58** 162 (2010).

[50] Hüe, F.; Hÿtch, M.; Bender, H.; Houdellier, F. and Claverie, A. *Direct Mapping of Strain in a Strained Silicon Transistor by High-Resolution Electron Microscopy*. Physical Review Letters **100** 28 (2008).

[51] Chung, J.; Lian, G. and Rabenberg, L. *Local strain measurement in a strain-engineered complementary metal-oxide-semiconductor device by geometrical phase analysis in the transmission electron microscope*. Applied Physics Letters **93** 081909 (2008).

[52] Hÿtch, M. J. and Plamann, T. *Imaging conditions for reliable measurement of displacement and strain in high-resolution electron microscopy*. Ultramicroscopy **87** 199 (2001).

[53] Rouviere, J. L. and Sarigiannidou, E. *Theoretical discussions on the geometrical phase analysis*. Ultramicroscopy **106** 1 (2005).

[54] Bogusławski, P.; Briggs, E. L. and Bernholc, J. *Native defects in gallium nitride*. Physical Review B **51** 17255 (1995).

[55] Hull, D. and Bacon, D. J. *Introduction to Dislocations*. Pergamon Press, 3rd ed. (1984).

[56] Hirth, J. P. and Lothe, J. *Theory of dislocations*. Krieger publishing company, Malabar, Florida (1982).

[57] Shchukin, V. A. and Bimberg, D. *Spontaneous ordering of nanostructures on crystal surfaces*. Review of Modern Physics **71** 1125 (1999).

[58] Frank, F. C. *Crystal dislocations. –Elementary concepts and definitions*. Philosophical Magazine **42** 809 (1951).

[59] Stampfl, C. and Van de Walle, C. *Energetics and electronic structure of stacking faults in AlN, GaN, and InN*. Physical Review B **57** R15052 (1998).

[60] Gottschalk, H.; Patzer, G. and Alexander, H. *Stacking fault energy and ionicity of cubic III–V compounds*. Physica status solidi (a) **207** 207 (1978).

[61] Turley, J. and Sines, G. *The anisotropy of Young's modulus, shear modulus and Poisson's ratio in cubic materials*. Journal of physics D **4** 264 (1971).

[62] Zhang, J.-M.; Zhang, Y.; Xu, K.-W. and Ji, V. *Anisotropic elasticity in hexagonal crystals*. Thin Solid Films **515** 7020 (2007).

[63] Chaldyshev, V.; Bert, N.; Kolesnikova, a. and Romanov, a. *Stress relaxation scenario for buried quantum dots*. Physical Review B **79** 233304 (2009).

[64] Shaw, D. *Diffusion in Semiconductors*. In *Surveys in High Energy Physics*, pages 121–135 (2007).

[65] Willoughby, A. F. W. *Atomic diffusion in semiconductors*. Reports on Progress in Physics **41** 1665 (1978).

[66] Shaw, D. *Atomic Diffusion in Semiconductors*. Springer US (1973).

[67] Tuck, B. *Atomic Diffusion in Three-five Semiconductors*. Taylor & Francis (1988).

[68] Fisher, D. J. *Diffusion in GaAs and Other III-V Semiconductors: 10 Years of Research*. In *Defect and Diffusion Forum, Vol. 157-159*, page 520. Trans Tech Publications, Ltd. (1998).

[69] Zywietz, T.; Neugebauer, J. and Scheffler, M. *Adatom diffusion at GaN (0001) and (000-1) surfaces*. Applied Physics Letters **73** 487 (1998).

[70] Neugebauer, J.; Zywietz, T.; Scheffler, M.; Northrup, J.; Chen, H. and Feenstra, R. *Adatom Kinetics On and Below the Surface: The Existence of a New Diffusion Channel*. Physical Review Letters **90** 056101 (2003).

[71] Freundlich, A.; Leycuras, A.; Grenet, J. C. and Grattepain, C. *Defect-related Si diffusion in GaAs on Si*. Applied Physics Letters **53** 2635 (1988).

[72] Legros, M.; Dehm, G.; Arzt, E. and Balk, T. J. *Observation of giant diffusivity along dislocation cores*. Science **319** 1646 (2008).

[73] Sher, A.; Schilfgaarde, M. V.; Chen, A. and Chen, W. *Quasichemical approximation in binary alloys*. Physical Review B **36** 4279 (1987).

[74] Cahn, J. *Spinodal Decomposition*. Trans. Met. Soc. AIME **242** 166 (1968).

[75] Cahn, J. W. and Hilliard, J. E. *Free Energy of a Nonuniform System. I. Interfacial Free Energy*. The Journal of Chemical Physics **28** 258 (1958).

[76] Tang, M. and Karma, A. *Surface Modes of Coherent Spinodal Decomposition*. Physical Review Letters **108** 1 (2012).

[77] Oxtoby, D. W. *Nucleation of first-order phase transitions*. Accounts of Chemical Research **31** 91 (1998).

[78] Gibbs, J. W. *The collected Works*. Longmans- Green, New York, London, Toronto (1928).

[79] Schmelzer, E. J. W. P. *Nucleation Theory and Applications*. Wiley-VCH Verlag GmbH, Weinheim, FRG (2005).

[80] Chen, L.-Q. *Phase-Field Models for Microstructure Evolution*. Annual Review of Materials Research **32** 113 (2002).

[81] Bimberg, D.; Grundmann, M. and Ledentsov, N. N. *Quantum Dot Heterostructures*. Wiley, Berlin (1999).

[82] Rayleigh, L. *Wave theory of light*. In J. W. Strutt (ed.), *Scientific papers by Lord Rayleigh Vol.3*, pages 47–189. Cambridge University Press, Cambridge (1920).

[83] Als-Nielsen, J. and McMorrow, D. *Elements of modern X-ray physics*. Wiley (2011).

[84] De Broglie, L. A. *A General View of my Scientific Works*. In P. Goodman (ed.), *Fifty year of Electron Diffraction*, pages 5–6. Reidel Publishing Company, Dordrecht (1981).

[85] Wang, Z. L. *Elastic and inelastic scattering in eletron diffraction and imaging*. Plenum Press (1995).

[86] Cowley, J. M. *Diffraction physics*. North Holland, 3rd ed. (1995).

[87] Pennycook, S. J. and Nellist, P. D. *ScanningTransmission Electron Microscopy — imaging and analysis*. Springer (2011).

[88] De Graef, M. *Introduction to Conventional Transmission Electron Microscopy*. Cambridge University Press, Cambridge (2002).

[89] Bithell, E. G. and Stobbs, W. M. *Composition determination in the GaAs/(Al, Ga)As system using contrast in dark-field transmission electron microscope images*. Philosophical Magazine A **60** 39 (1989).

[90] Cerva, H. *Transmission electron microscopy of heteroepitaxial layer structures*. Applied Surface Science **50** 19 (1991).

[91] Howie, A. and Whelan, M. J. *Diffraction Contrast of Electron Microscope Images of Crystal Lattice Defects. II. The Development of a Dynamical Theory*. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **263** 217 (1961).

[92] Darwin, C. G. *The Theory of X-Ray Reflexion*. Philosophical Magazine Series 6 **27** 315 (1914).

[93] Darwin, C. G. *The theory of X-ray reflexion. Part II*. Philosophical Magazine Series 6 **27** 675 (1914).

[94] Stadelmann, P. A. *EMS - a software package for electron difitraction analysis and hrem image simulation in materials science*. Ultramicroscopy **21** 131 (1987).

[95] van Dyck, D. *The importance of backscattering in high-energy electron diffraction calculations*. Physica Status Solidi (B) **77** 301 (1976).

[96] Bloch, F. *Bemerkung zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfaehigket*. Z. Phys. **57** 545 (1929).

[97] Williams, D. B. and Carter, C. B. *Transmission eletron microscopy — A textbook for materials science*. Plenum Press (1996).

[98] Ashby, M. F. and Smith, G. C. *Direct electron microscopy of thin foils of internally oxidized dilute copper alloys*. Philosophical Magazine page 298 (1960).

[99] Ashby, M. F. and Brown, L. M. *On diifraction contrast from inclusions*. Philosophical Magazine page 1649 (1963).

[100] Ashby, M. F. and Brown, L. M. *Diffraction contrast from spherically symmetrical coherency strains*. Philosophical Magazine page 1083 (1963).

[101] Liu, C.; Miller, P. D. and Henstrom, W. L. *Transmission electron microscopy of semiconductor quantum dots*. Journal of Microscopy **199** 130 (2000).

[102] Miller, P. D.; Liu, C. P. and Gibson, J. M. *TEM measurement of strain in coherent quantum heterostructures*. Ultramicroscopy **84** 225 (2000).

[103] Urban, K. W. *Is science prepared for atomic-resolution electron microscopy?*. Nature insight **8** 260 (2009).

[104] Coene, W.; Thust, a.; Op de Beeck, M. and Van Dyck, D. *Maximum-likelihood method for focus-variation image reconstruction in high resolution transmission electron microscopy*. Ultramicroscopy **64** 109 (1996).

[105] De Backer, A.; Van Aert, S. and Van Dyck, D. *High precision measurements of atom column positions using model-based exit wave reconstruction*. Ultramicroscopy **111** 1475 (2011).

[106] Frank, J. *a practical resolution criteria in optics and electron microscopy*. Optik **43** 25 (1975).

[107] Zou, X.; Sundberg, M.; Larine, M. and Hovmöller, S. *Structure projection retrieval by image processing of HREM images taken under non-optimum defocus conditions*. Ultramicroscopy **62** 103 (1996).

[108] Hillyard, S. and Silcox, J. *Thickness effects in ADF STEM zone axis images*. Ultramicroscopy **52** 325 (1993).

[109] Hillyard, S.; Loane, R. F. and Silcox, J. *Annular dark-field imaging: resolution and thickness effects*. Ultramicroscopy **49** 14 (1993).

[110] Pennycook, S. and Jesson, D. *High-resolution Z-contrast imaging of crystals*. Ultramicroscopy **37** 14 (1991).

[111] Nye, J. F. *Physical Properties of Crystals: Their Representation by Tensors and Matrices*. Clarendon Press, Oxford (1957).

[112] Chung, J. and Rabenberg, L. *Effects of strain gradients on strain measurements using geometrical phase analysis in the transmission electron microscope*. Ultramicroscopy **108** 1595 (2008).

[113] Hüe, F.; Johnson, C. L.; Lartigue-Korinek, S.; Wang, G.; Buseck, P. R. and Hÿtch, M. J. *Calibration of projector lens distortions*. Journal of electron microscopy **54** 181 (2005).

[114] Scheerschmidt, K. and Werner, P. *Characterization of structure and composition of quantum dots by transmission electron microscopy*. In *Nano-optoelectronics*, pages 67–98 (2002).

[115] Beanland, R. *Dark field transmission electron microscope images of III-V quantum dot structures*. Ultramicroscopy **102** 115 (2005).

[116] Blom, C. and Altona, C. *Application of self-consistent-field ab initio calculations to organic molecules*. Molecular Physics **34** 177 (1977).

[117] Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A. and Skiff, W. M. *UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations*. Journal of American Chemistry Society **114** 10024 (1992).

[118] Fock, V. *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems*. Zeitschrift der Physik **62** 795 (1930).

[119] Kohn, W. and Sham, L. J. *Quantum Density Oscillations in an Inhomogeneous Electron Gas*. Physical Review **137** A1697 (1965).

[120] Kohn, W. and Sham, L. J. *Self-Consistent Equations Including Exchange and Correlation Effects*. Physical Review **140** A1133 (1965).

[121] Li, T.; Mastro, M. and Dadgar, A. (eds.). *III-V compound semiconductors integration with silicon-based microelectronics*. CRC Press (2011).

[122] Tanabe, K.; Watanabe, K. and Arakawa, Y. *III-V/Si hybrid photonic devices by direct fusion bonding*. Scientific reports **2** 349 (2012).

[123] Mi, Z. and Chang, Y.-L. *III-V compound semiconductor nanostructures on silicon: epitaxial growth, properties, and applications in light emitting diodes and lasers*. Journal of Nanophotonics **3** 031602 (2009).

[124] Yoon, J.; Jo, S.; Chun, I. S.; Jung, I.; Kim, H.-S.; Meitl, M.; Menard, E.; Li, X.; Coleman, J. J.; Paik, U. and Rogers, J. a. *GaAs photovoltaics and optoelectronics using releasable multilayer epitaxial assemblies*. Nature **465** 329 (2010).

[125] Ko, H.; Takei, K.; Kapadia, R.; Chuang, S.; Fang, H.; Leu, P. W.; Ganapathi, K.; Plis, E.; Kim, H. S.; Chen, S.-Y.; Madsen, M.; Ford, A. C.; Chueh, Y.-L.; Krishna, S.; Salahuddin, S. and Javey, A. *Ultrathin compound semiconductor on insulator layers for high-performance nanoscale transistors*. Nature **468** 286 (2010).

[126] Zhao, Z.; Yoon, T.; Feng, W.; Li, B.; Kim, J.; Liu, J.; Hulko, O.; Xie, Y.; Kim, H.; Kim, K.; Kim, H.; Wang, K.; Ratsch, C.; Caflisch, R.; Ryu, D. and Russell, T. *The challenges in guided self-assembly of Ge and InAs quantum dots on Si*. Thin Solid Films **508** 195 (2006).

[127] Petrov, V. N.; Polyakov, N. K.; Egorov, V. A.; Cirlin, G. E.; Zakharov, N. D.; Werner, P.; Ustinov, V. M.; Denisov, D. V.; Ledentsov, N. N. and Alferov, Z. I. *Study of Multilayer Structures with InAs Nanoobjects in a Silicon Matrix*. Low-dimensional system **34** 810 (2000).

[128] Zakharov, N. D.; Werner, P.; Gosele, U.; Heitz, R.; Bimberg, D.; Ledentsov, N. N.; Ustinov, V. M.; Volovik, B. V.; Alferov, Z. I.; Polyakov, N. K.; Petrov, V. N.; Egorov, V. a. and Cirlin, G. E. *Structure and optical properties of Si/InAs/Si layers grown by molecular beam epitaxy on Si substrate*. Applied Physics Letters **76** 2677 (2000).

[129] Hull, D. and Bacon, D. *Introduction to dislocations*. Elsevier, 5th ed. (2011).

[130] Nikanorov, S. P.; Burenkov, Y. A. and Stepanov, A. V. *No Title*. Sov. Phys. Solid State **13** 2516 (1971).

[131] Burenkov, Y. A.; Davydov, S. Y. and Nikanorov, S. P. *No Title*. Sov. Phys. Solid State **17** 1446 (1975).

[132] Mader, W. and Knauss, D. *Equilibrium position of misfit dislocations at planar interfaces*. Acta Metallurgica et Materialia **40** S207 (1992).

[133] Gutkin, M. Y. and Romanov, a. E. *On the stand-off positions of misfit dislocations*. Physica Status Solidi (a) **144** 39 (1994).

[134] Grundmann, M.; Stier, O. and Bimberg, D. *InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, and electronic structure*. Physical Review B **52** 11969 (1995).

[135] Trampert, A. *Heteroepitaxy of dissimilar materials : effect of interface structure on strain and defect formation*. Physica E **13** 1119 (2002).

[136] Tixier, S.; Adamcyk, M.; Tiedje, T.; Francoeur, S.; Mascarenhas, A.; Wei, P. and Schiettekatte, F. *Molecular beam epitaxy growth of GaAs1-xBix*. Applied Physics Letters **82** 2245 (2003).

[137] Lewis, R. B.; Masnadi-Shirazi, M. and Tiedje, T. *Growth of high Bi concentration GaAs1-xBix by molecular beam epitaxy*. Applied Physics Letters **101** 082112 (2012).

[138] Lu, X.; Beaton, D. A.; Lewis, R. B.; Tiedje, T. and Whitwick, M. B. *Effect of molecular beam epitaxy growth conditions on the Bi content of GaAs(1-x)Bi(x)*. Applied Physics Letters **92** 192110 (2008).

[139] Francoeur, S.; Seong, M.-J.; Mascarenhas, A.; Tixier, S.; Adamcyk, M. and Tiedje, T. *Band gap of GaAs(1-x)Bi(x), 0<x<3.6%*. Applied Physics Letters **82** 3874 (2003).

[140] Janotti, A.; Wei, S.-H. and Zhang, S. *Theoretical study of the effects of isovalent coalloying of Bi and N in GaAs*. Physical Review B **65** 115203 (2002).

[141] De Giorgi, M.; Taurino, A.; Passaseo, A.; Catalano, M. and Cingolani, R. *Interpretation of phase and strain contrast of TEM images of InxGa1-xAs/GaAs quantum dots*. Physical Review B **63** 245302 (2001).

[142] Founta, S.; Coraux, J.; Jalabert, D.; Bougerol, C.; Rol, F.; Mariette, H.; Renevier, H.; Daudin, B.; Oliver, R. A.; Humphreys, C. J.; Noakes, T. C. Q. and Bailey, P. *Anisotropic strain relaxation in a-plane GaN quantum dots*. Journal of Applied Physics **101** 063541 (2007).

[143] Daudin, B. *Polar and nonpolar GaN quantum dots*. Journal of Physics: Condensed Matter **20** 473201 (2008).

[144] Deguffroy, N.; Tasco, V.; Baranov, A. N.; Tournié, E.; Satpati, B.; Trampert, A.; Dunaevskii, M. S.; Titkov, A. and Ramonda, M. *Molecular-beam epitaxy of InSb/GaSb quantum dots*. Journal of Applied Physics **101** 124309 (2007).

[145] Springholz, G. and Bauer, G. *Molecular beam epitaxy of IV–VI semiconductor hetero- and nano-structures*. Physica Status Solidi (B) **244** 2752 (2007).

[146] Stringfellow, G. *Spinodal decomposition and clustering in III/V alloys*. Journal of Electronic Materials **11** 903 (1982).

[147] Rachbauer, R.; Gengler, J. J.; Voevodin, A. A.; Resch, K. and Mayrhofer, P. H. *Temperature driven evolution of thermal, electrical, and optical properties of Ti-Al-N coatings*. Acta materialia **60** 2091 (2012).

[148] Puustinen, J.; Wu, M.; Luna, E.; Schramm, A.; Laukkanen, P.; Laitinen, M.; Sajavaara, T. and Guina, M. *Variation of lattice constant and cluster formation in GaAsBi*. Journal of Applied Physics **114** 243504 (2013).

[149] Pennycook, S. J. and Nellist, P. D. *Z-contrast scanning transmission electron microscopy*. In D. Rickerby; G. Valdrè and U. Valdrè (eds.), *Impact of Electron Scanning Probe Microscopy on Materials Research*. Kluwer Academic Publishers, The Netherlands (1999).

[150] Luna, E.; Ishikawa, F.; Batista, P. D. and Trampert, A. *Indium distribution at the interfaces of (Ga,In)(N,As)/GaAs quantum wells*. Applied Physics Letters **92** 141913 (2008).

[151] Chaldyshev, V. *Two-dimensional organization of As clusters in GaAs*. Materials Science and Engineering: B **88** 195 (2002).

[152] Schiferl, D. and Barrett, C. S. *The crystal structure of arsenic at 4.2, 78 and 299°K*. Journal of Applied Crystallography **2** 30 (1969).

[153] Frank, F. C. *On Miller–Bravais indices and four-dimensional vectors*. Acta Crystallographica **18** 862 (1965).

[154] Bert, N. A. and Chaldyshev, V. V. *Changes in the Moire patterns in electron-microscope images of As clusters in LT-GaAs as their size decreases*. Semiconductors **30** 988 (1996).

[155] Kong, X.; Trampert, A. and Ploog, K. H. *Composition fluctuations and clustering in (Ga,In)(N,As)/GaAs(001) heterostructures studied by analytical transmission electron microscopy*. Journal of Physics and Chemistry of Solids **69** 335 (2008).

[156] Norman, A. G.; France, R. and Ptak, A. J. *Atomic ordering and phase separation in MBE GaAs(1-x)Bi(x)*. Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures **29** 03C121 (2011).

[157] Ferro, R.; Saccone, A. and Delfino, S. *Arsenic-Bismuth-Gallium*. In *Ternary Alloys*, pages 221–225. VCH 9 (1994).

[158] Liu, X.; Prasad, A.; Nishio, J.; Weber, E. R.; Liliental-Weber, Z. and Walukiewicz, W. *Native point defects in low-temperature-grown GaAs*. Applied Physics Letters **67** 279 (1995).

[159] Gebauer, J.; Krause-Rehberg, R.; Eichler, S.; Luysberg, M.; Sohn, H. and Weber, E. R. *Ga vacancies in low-temperature-grown GaAs identified by slow positrons*. Applied Physics Letters **71** 638 (1997).

[160] Chaldyshev, V. V.; Kolesnikova, A. L.; Bert, N. A. and Romanov, A. E. *Investigation of dislocation loops associated with As–Sb nanoclusters in GaAs*. Journal of Applied Physics **97** 024309 (2005).

[161] Khachaturyan, A. G. *Theory of structural transformation in solids*. John Wiley Sons, New York (1983).

[162] Butkute, R.; Pacebutas, V.; Cechavicius, B.; Adomavicius, R.; Koroliov, A. and Krotkus, A. *Thermal annealing effect on the properties of GaBiAs*. Physica Status Solidi (C) **9** 1614 (2012).

[163] Ranchal, R.; Yadav, B. S. and Trampert, A. *Ferromagnetism at room temperature of c- and m-plane GaN:Gd films grown on different substrates by reactive molecular beam epitaxy*. Journal of Physics D: Applied Physics **46** 075003 (2013).

[164] Kresse, G. and Hafner, J. *Ab initio molecular dynamics for liquid metals*. Physical Review B **47** 558 (1993).

[165] Kresse, G. and Furthmüller, J. *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*. Physical Review B **54** 11169 (1996).

[166] Scarpulla, M. A.; Gallinat, C. S.; Mack, S.; Speck, J. S. and Gossard, A. C. *GdN (111) heteroepitaxy on GaN (0001) by N2 plasma and NH3 molecular beam epitaxy*. Journal of Crystal Growth **311** 1239 (2009).

[167] Schwarz, R. B.; Khachaturyan, K. and Weber, E. R. *Elastic moduli of gallium nitride*. Applied Physics Letters **70** 1122 (1997).

[168] Abdelouahed, S. and Alouani, M. *Calculated electronic properties and structural phase transitions of GdN pnictide under hydrostatic pressure*. Physical Review B **76** 214409 (2007).

[169] Lee, N. E.; Powell, R. C.; Kim, Y. W. and Greene, J. E. *Molecular beam epitaxy of GaN(0001) utilizing NH3 and/or NH+x ions: Growth kinetics and defect structure*. Journal of Vacuum Science and Technology A **13** 2293 (1995).

[170] Kirchner, V.; Heinke, H.; Birkle, U.; Einfeldt, S.; Hommel, D.; Selke, H. and Ryder, P. L. *Ion-induced crystal damage during plasma-assisted MBE growth of GaN layers*. Physical Review B **58** 15749 (1998).

[171] Gloux, F.; Ruterana, P.; Lorenz, K. and Alves, E. *A comparative structural investigation of GaN implanted with rare earth ions at room temperature and 500C*. Materials Science and Engineering: B **146** 204 (2008).

[172] Chen, C.-C.; Ting, Y.-S.; Lee, C.-C.; Chi, G.-C.; Chakraborty, P.; Chini, T.; Chuang, H.-W.; Tsang, J.-S.; Kuo, C.-T.; Tsai, W.-C.; Chen, S.-H. and Chyi, J.-I. *Er diffusion into gallium nitride*. Solid-State Electronics **47** 529 (2003).

[173] Ting, Y.-S.; Chen, C.-C.; Lee, C.-C.; Chi, G.-C.; Chini, T. K.; Chakraborty, P.; Chuang, H.-W.; Tsang, J.-S.; Kuo, C.-T.; Tsai, W.-C.; Chen, S.-H. and Chyi, J.-I. *Diffusion mechanism and photoluminescence of erbium in GaN*. Optical Materials **24** 515 (2003).

[174] Wu, M. and Trampert, A. *Coherent GdN clusters in epitaxial GaN:Gd thin films determined by transmission electron microscopy*. Nanotechnology **24** 255701 (2013).

[175] Erwin, S. C.; Baski, A. A.; Whitman, L. J. and Rudd, R. E. *Frenkel-Kontorova model of vacancy-line interactions on Ga/Si (112)*. Physical Review Letters **83** 1818 (1999).

[176] Kalvoda, S.; Dolg, M.; Flad, H.-J.; Fulde, P. and Stoll, H. *Ab initio approach to cohesive properties of GdN*. Physical Review B **57** 2127 (1998).

[177] Palummo, M.; Bertoni, C. M.; Reining, L. and Finocchi, F. *The electronic structure of gallium nitride*. Physica B: Condensed Matter **185** 404 (1993).

[178] Bryngdahl, O. *Moiré and higher grating harmonics*. Journal of the Optical Society of America **65** 685 (1975).

[179] Patorski, K.; Wielgus, M.; Ekielski, M. and Kaźmierczak, P. *AFM nanomoire technique with phase multiplication*. Measurement Science and Technology **24** 035402 (2013).

[180] Cho, A. Y. and Arthur, J. R. *Molecular beam epitaxy*. Prog. Solid State Chem. **10** 157 (1975).

[181] Herman, M. A.; Richter, W. and Sitte, H. *Epitaxy: physical principles and technical implementation*. Springer-Verlag, Berlin (2004).

[182] Braun, W. *Applied RHEED*, vol. 154. Springer (1999).

[183] Burenkov, Y. A.; Burdukov, Y. M.; Davidov, S. Y. and Nikanorov, S. P. *No Title*. Sov. Phys. Solid State **15** 1175 (1973).

[184] Thamm, A.; Brandt, O.; Takemura, Y.; Trampert, A. and Ploog, K. H. *Reactive molecular-beam epitaxy of GaN layers directly on 6H–SiC(0001)*. Applied Physics Letters **75** 944 (1999).

[185] Sun, Y. J.; Brandt, O.; Jahn, U.; Liu, T. Y.; Trampert, A.; Cronenberg, S.; Dhar, S. and Ploog, K. H. *Impact of nucleation conditions on the structural and optical properties of M-plane GaN(1100) grown on r-LiAlO(2)*. Journal of Applied Physics **92** 5714 (2002).

[186] Chu, S. N. G. and Nakahara, S. *1/2[100](100) Dislocation Loops in a Zinc Blende Structure*. Applied Physics Letters **56** 434 (1990).

[187] Vasyukov, D. A.; Baidakova, M. V.; Chaldyshev, V.; Suvorova, A. A.; Preobrazhenskii, V. V.; Putyato, M. A. and Semyagin, B. R. *Structural transformations in low-temperature grown GaAs:Sb*. Journal of Physics D: Applied Physics **34** A15 (2001).

[188] Bert, N. A.; Kolesnikova, A. L.; Nevedomsky, V. N.; Preobrazhenskii, V. V.; Putyato, M. A.; Romanov, A. E.; Seleznev, V. M.; Semyagin, B. R. and Chaldyshev, V. V. *Formation of dislocation defects in the process of burying of InAs quantum dots into GaAs*. Semiconductors **43** 1387 (2009).

[189] Dundurs, J. and Salamon, N. J. *Circular prismatic dislocation loop in a two-phase material*. Physica Status Solidi (b) **50** 125 (1972).

List of Figures

List of Tables

Acknowledgments

My gratitudes goes to all the people who have helped me from my start of PhD in Berlin to the point of completion of the thesis.

First of all, I am grateful to *Prof. Henning Riechert* for providing me the opportunity to carry out the research in Paul-Drude-Institut für Festkörperelektronik in Berlin (PDI). Furthermore, thank you for all the effort to help me out with troubles during my studies.

I appreciate all my thesis reviewer for the critical reading of my thesis.

At this point, I also want to express my gratitude to *Dr. Wolfgang Braun* for forwarding my PhD application and introducing me to Prof. Henning Riechert.

My deepest appreciation goes to *Dr. Achim Trampert*, who virtually supervised my whole work. Thank you for all your patient guidance in the hands-on TEM operation, in how to organize and present scientific results and in how to carry out scientific research along in the past years. Thank you for your encouragements when I was down, for all the meaningful discussions on science and philosophy and for your support in my work and life in Berlin. Without your help, I would had not accomplished what I have now.

Secondly, I acknowledge all the people who have (substantially) contributed and were connected to the work presented in the thesis. I thank *Dr. Steven C. Erwin*, from Naval Research Laboratory in Washington D.C., for all the dedicated DFT calculations, for the guidance in scientific writing and for the great patient in commenting on the manuscript word by word. Your way of critical thinking in scientific problems really sparkled my inspiration.

Dr. Rocío Ranchal, at University of Madrid in Spain: thank you for providing the GaN:Gd samples and valuable discussions about the growth parameters and properties of the samples trough email during my study. Pity we had never meet in person.

Mr. Janne Puustinen and *Prof. Mircea Guina*, at Tampere University of Technology in Finland, thank you for providing the GaAsBi samples in my study and responsive discussions on the topic.

Mr. Tariq Al-Zoubi, *Dr. Muhamed Benyoucef* and *Prof. Johann Peter Reithmaier*, at University of Kassel, Germany, thank you for providing the InAs quantum dots samples for my study.

I appreciate the help from all the people in the TEM group. I am gratefully thank *Dr. Esperanza Luna*, for all the valuable discussions on variety of research topics and critical reading of the manuscripts and for the interesting discussions about career developments, culture and life.

Ms. Doreen Stefen, thank you for your help in the TEM specimen preparation and your illustrative photos about TEM specimen preparations. Discussions on sports and German culture is appreciated.

Mrs. Astrid Pfeiffer, thank you for the help of TEM operation and maintenance, although I could still not able to communicate with you in German, :-(.

Mr. Michael Niehle, my office pal, thank you for sharing the office room, for translating the German abstract for me, for the valuable discussions on theoretical and practical aspects about TEM and for you volunteered time and energy to help deal with troubles in my life in Berlin and study in HU-Berlin throughout these years. *Ms. Claudia Roder* and *Ms. Katja Hagmann*, former and current office pals, thank you for sharing the office room and interesting scientific and non-scientific discussions, which have enriched my life experience.

Dr. Bernd Jenichen, thank you for critical reading of my manuscripts and posters and for the interesting discussions about X-ray diffraction.

Dr. Javiar Brandel, thank you for the inspiring discussions on scientific issues and on life and culture.

I am really grateful to the professors in the Institute of Physics at Humboldt Universität zu Berlin who have guided me to the realm of solid state physics and turned my minds of material engineering to physics, an interesting way of thinking, although I was required and unwilling to participate the additional courses in the beginning.

I thank *Prof. W. Ted Masselink*, who offered me the chance to be tutor in the course of Introduction to Solid State Physics, *Prof. Igor Sokolov*, who guided me to study statistic physics and *Prof. Wolfgang Nolting*, who guided me in the study of Quantum Mechanics.

Next, My appreciation goes to the people of Administration department in PDI who had helped me with administrative chores. They are *Mrs. Petra Grasnick*, *Mr. Andreas Hartung*, *Mrs. Nadine Möller*, *Ms. Anja Holldack*. Besides, I thank *Mrs. Ilka Schuster* for help me print posters, *Mrs. Sylvia Räther* and *Mrs. Dagmar Dormeyer* for the help to borrow books and literatures from other libraries.

I thank all the PhD students in PDI who shared the knowledge of different topics of research and the information about "Promotionsverfahren". I could not really match the names and faces of you all, but the names I remembered are *Dr. Stefen Breuer*, *Dr. Friedrich Limbach*, *Dr. Martin Wölz*, *Dr. Martin Dubslaff*, *Dr. Martin Wienold*, *Dr. Karthick Perumal*, *Dr. Jonas Lähnemann*, *Ms. Maria Hilse*, *Mr. Christian Hauswald*, *Mr. Pawel Bruski*, *Ms. Antje Buchholz*, *Mr. Thomas Hentschel*, *Mr. Timo Schumann Mr. Yori Manzke*, *Mr. Thilo Krause* and *Mr. Mattia Musolino*.

My special acknowledgements is reserved for the Chinese colleagues at PDI, they are *Dr. Cunxu Gao* and his family, *Dr. Kwang-Ru Wang*, *Mr. Qian Wan*, *Dr. Jianshu Yang* and her family , *Dr. Chang-Ning Huang* and his family, *Dr. Xiang Kong* and his family, *Dr. Xiang Lü*, *Dr. Xiaojing Gong*, *Ms. Wan-Hsien Lin* and *Dr. Hong Li* and his family. I will always remember the time we spend together cooking, dinning, hiking, playing and working. It is you that made my life in Berlin not so monotonic.

Finally, I heartfully thank the support from my family and friends, especially my parents *Mrs. Kuangjun Chen* and *Mr. Guxun Wu* and my wife *Mrs. Ying Xiao*. Without the continuous love and meticulous care, I can hardly imagine how my life would be in Berlin.

Publication list

Journal articles

Mingjian Wu and Achim Trampert, *Coherent GdN clusters in epitaxial GaN:Gd thin films determined by transmission electron microscopy*, Nanotechnology **24**, 255701 (2013)
highlighted by Nanotechnology and reported by Nanotechweb.org
URL http://nanotechweb.org/cws/article/lab/54266

Mingjian Wu, Steven C. Erwin and Achim Trampert, (2013) *Coherent Clustering of GdN in Epitaxial GaN:Gd Thin Film*, MRS Proceedings, **1554**, mrss13-1554-u06-01 (2013)
doi:10.1557/opl.2013.930

J. Puustinen, **Mingjian Wu**, E. Luna, A. Schramm, P. Laukkanen, M. Laitinen, T. Sajavaara, and M. Guina, *Variation of lattice constant and cluster formation in GaAsBi*, J. Appl. Phys. **114**, 243504 (2013) doi: 10.1063/1.4851036

M. Benyoucef, T. Al-Zoubi, J. P. Reithmaier, **Mingjian Wu** and A. Trampert, *Nanostructured hybrid material based on highly mismatched III-V nanocrystals fully embedded in silicon*, Physica Status Solidi (a) **211**, 817 (2014) doi: 10.1002/pssa.201330395

Mingjian Wu, Esperanza Luna, Janne Puustinen, Mircea Guina, and Achim Trampert, *Formation and phase transformation of Bi-containing QD-like clusters in annealed GaAsBi thin films*, Nanotechnology (2014) in press

Mingjian Wu, Steven C. Erwin and Achim Trampert, *Atomic structure and energetics of GdN clusters in Gd-doped GaN*, Acta Materialia, revision submitted.

Mingjian Wu, Esperanza Luna, Janne Puustinen, Mircea Guina, and Achim Trampert, *Observation of a novel triple-period (TP)-A and -B type ordering in LT-GaAsBi*, in preparation.

Rocío Ranchal, **Mingjian Wu**, and Achim Trampert, *Correlation between structural and magnetic properties of Gd-doped GaN*, in preparation.

Mingjian Wu, T. Al-Zoubi, J. P. Reithmaier and A. Trampert, *Strain of InAs quantum dots embedded in Si(001)*, in preparation.

Conference presentations

Mingjian Wu and Achim Trampert, *Quantitative TEM analysis of local lattice distortions in epitaxial Gd:GaN layers* (late breaking poster), Microscopy Conference 2011 (MC2011), Kiel, Germany, 28 August – 02 September 2011

Mingjian Wu and Achim Trampert, *Quantitative TEM analysis of local lattice distortions in epitaxial Gd:GaN layers* (poster), CFN Summer School on Nano-Electronics, Bad Herrenalb, Germany, 11–14 September 2011

Mingjian Wu, A. Trampert, T. Al-Zoubi, M. Usman, M. Benyoucef and J. P. Reithmaier, *Microstructure of InAs nano-crystals embedded in Si* (poster), 76th Annual Conference of the DPG – the Spring Meeting of the Condensed Matter Section, Berlin, Germany, 25–30 March 2012

Mingjian Wu, A. Trampert, T. Al-Zoubi, M. Usman, M. Benyoucef and J. P. Reithmaier, *Microstructure of InAs nano-crystals embedded in Si* (poster), 2012 Heraeus Physics School on "Microelectronics for Society – More than Moore expands More Moore", Bad Honnef, Germany, 10–16 June 2012

Mingjian Wu, Steven C. Erwin and Achim Trampert, *Determination of Clustering in Dilute GaN:Gd Thin Films* (poster), International Conference on Extended Defects in Semiconductors (EDS2012), Thessaloniki, Greece, 24–29 June 2012
(Awarded the **Helmut Alexander Award** for the best poster)

Mingjian Wu, Steven C. Erwin and Achim Trampert, *Determination of Clustering in Epitaxial GaN:Gd Thin Films* (poster), 2012 "Defects in Semiconductors" of Gordon Research Conference (GRC), University of New England, Biddeford, ME, USA, 12–17 August 2012

Mingjian Wu, Steven C. Erwin and Achim Trampert, *Coherent Nano-Decomposition in Epitaxial GaN:Gd Thin Films* (oral), Spring Meeting of the Materials Research Society (MRS), San Francisco, CA, USA, 1–5 April 2013

Mingjian Wu, E. Luna, A. Trampert, J. Puustinen and M. Guina, *Formation and phase transformation of Bi-containing clusters in annealed GaAsBi thin films* (oral, invited), COST MP0805 Final Meeting: Novel Gain Materials and Devices Based on III-V-N/Bi Compounds, University of Istanbul, Istanbul, Turkey 24–26 September 2013

Mingjian Wu, Steven C. Erwin and Achim Trampert, *Nano-clustering of GdN in dilute Gd-doped GaN* (poster, invited), International Symposium on Nano-scale Transport and Technology, NTT R&D Center, Atsugi, Kanagawa, Japan 26–29 November 2013

Berlin, den April 30, 2014 Mingjian Wu

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfaßt und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Ich habe mich an keiner anderen Universität um einen Doktorsgrad beworben und besitze auch keinen entsprechenden Doktorsgrad.

Ich erkläre die Kenntnis der dem Verfahren zugrunde liegenden Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin.

Berlin, den April 30, 2014 Mingjian Wu