ABSTRACT

Atomic-resolution high-angle annular dark-field (HAADF) scanning transmission electron microscopy (STEM) has been used to the structural and compositional analysis of materials. The first part of this paper is a review of our recent HAADF-STEM investigations, which comprise physics to understand its imaging, and illustration of artifacts in images and imaging process. Next, we present our investigations of the multiple InGaN/GaN quantum wells and the strained AlGaN/GaN superlattice cladding in InGaN-based light emitting diodes or laser diodes, which have been performed by HAADF-STEM, high-resolution field-emission scanning electron microscopy, and high-resolution transmission electron microscopy. The observation of ultra-high density InGaN quantum dots is also shown.

Keywords: high-angle annular dark field scanning transmission electron microscopy, field-emission scanning electron microscopy, high-resolution transmission electron microscopy, InGaN-based light emitting diodes, laser diodes, multiple quantum wells, strain-layer superlattice cladding, quantum dots.

1. INTRODUCTION

On January 14 and 15, 2008, I presented lectures at the Department of Industrial Chemistry, Faculty of Science in Chiang Mai University. The lectures reviewed our electron microscopy investigations, which have been performed since 1961 on the structure and growth of various metallic, semiconducting, ceramic, other inorganic or organic materials and devices, using different types of electron microscopes and microscopy techniques. The electron microscopy is theory and technique for macro- and nano-structural analysis and characterization. Therefore, the lectures aimed at providing each of the audience with knowledge to find appropriate electron microscopy for her or his investigation, by showing many electron microscopy investigations. This paper is a part of my lectures and a review of our recent investigations on high-angle annular dark-field (HAADF) scanning transmission electron microscopy (STEM) and on InGaN-based light emitting diodes (LEDs) and laser diodes (LDs).

HAADF-STEM is a quite new technique...
developed in 1990's [1], more than 50 years after the invention of a transmission electron microscopy (TEM) instrument by Ruska in 1931 and the idea of scanning electron microscopy (SEM) by Ardenne in 1937. Howie [2] proposed to use a detector to high-angle range, so-called Howie detector, in order to exclude elastic coherent scattering and eliminate the phase problem. HAADF-STEM is used to characterize the structure of grain boundaries and precipitates as well as perfect crystals [1]. It has become a quite popular technique since Kawasaki first installed a HAADF detector in a commercial STEM instrument with Schottky field emission gun in 1997, which was reported by James et al. in 1998 [3]. He succeeded in taking atomic resolution HAADF-STEM images [3-6]. The electron beams run in reverse between STEM and conventional transmission electron microscopes (CTEM) (Figure 1). The bright field (BF) STEM imaging is exactly the same with the CTEM imaging according to the reciprocity rule [7], but HAADF-STEM is completely different from CTEM. We have been engaged in HAADF-STEM investigations since 1998 [4-6, 8-37]. The first part of this paper is a review from our investigations on HAADF-STEM imaging.

GaN-base light emitting diodes (LEDs) and laser diodes (LDs) [38] have been widely used as light sources over a range of visible and ultraviolet wavelength. This is due to exceptionally high photoemission efficiency from $\text{In}_{x}\text{Ga}_{1-x}\text{N}$ quantum wells (QWs) despite not a few of defects involved. Next, we present our recent investigations of GaN-based LEDs and LDs, which were performed using HAADF-STEM as well as field emission (FE) SEM, cathodoluminescence SEM and high-resolution (HR) TEM [26-37,39,40].

Figure 1. Ray diagram of scanning transmission electron microscope (STEM) and conventional transmission electron microscope (CTEM). The ray runs downward in STEM and upwards in CTEM.

Figure 2. Coherency, incoherency, and partial time- or spatial-coherency of classic particles. (a, b) Coherency. Two ballistic missiles 1 are coherent with each other since they are launched with the same velocity (energy) from the same pad and run in the same orbit. (c) Spatial incoherency. Missile 2 that is launched from a different pad with the same energy is incoherent with missile 1. Missile 3 hitting a given target area is in partial spatial-coherency with missile 1. (d) Time incoherency. Missile 2 that is launched with different energy from the same pad is incoherent with missile 1 (because time of arrival is different). Missile 3 hitting the target area is in partial time-coherency with missile 1.
2. HIGH-ANGLE ANNULAR DARK FIELD (HAADF) SCANNING TRANSMISSION ELECTRON MICROSCOPY (STEM)

2.1 Physics of HAADF-STEM Imaging

First of all, we explain the terms of ‘coherency’ and ‘incoherency’ [32]. It is a fundamental concept of quantum physics that a quantum particle also behaves as wave. As well known, the light waves passed through different holes interfere with each other to show Young’s fringes. The interference also occurs between the electron waves scattered by different atoms (exactly, electrostatic potential around atoms). It is caused only between elastically scattered waves or between inelastically scattered waves with the same wavelength. The former gives Bragg reflections and the latter Kikuchi pattern. Here, we call the scattering that causes Bragg reflections ‘Bragg scattering’. The ‘elastic scattering’ is no energy-loss scattering. According to quantum mechanism, the interference never occurs between waves of different electrons [41]. Thus, the interference occurs only between the waves of the same wavelength incorporated with ‘a single electron’. The electron is observed in terms of the intensity $I$, which gives a probability of the appearance. The observed electron does not indicate the position $r$ where it always locates, but indicates only a place where it is accidentally captured. For the visualization of the probable position of an electron, we need to detect many electrons that are captured at a region assigned with the same interference condition. These electrons are ‘coherent’ with each other. They show a physically valuable intensity distribution (such as interference fringes or diffraction spots) as a result of the cooperation. Thus, ‘interference’ is a term for the intra-electron relation of waves, while the term of ‘coherency’ or ‘incoherency’ is for the inter-electron relation between the waves that belong to different electrons. In other words, the interference is described as $\Phi_j = \sum_i \phi_{ij}$, where $\phi_{ij}$ is the wave of electron $j$ scattered by atom $i$, while the coherency and incoherency are described as $I = \sum_j I_j = \sum_j |\Phi_j|^2$, where $I_j = |\Phi_j|^2$ is the interference intensity for the electron $j$. The coherency is in the case of that all $\Phi$ are the same, that is, $I = n |\Phi|^2$ where $n$ is the number of electrons. Coherent electrons must have the same birthplace (the source or the resource for the scattered electrons) and the same energy. A field-emission gun can supply a coherent electron beam because it provides the same source at the top of the tip and gives the same energy to the emitting electrons. The spatial or time incoherency and partial coherency can be simply understood by schemata in Figure 2 which are shown for the classic particles by analogy with ballistic missiles. For the quantum particles such as electrons, we use the term of ‘partial coherency’ when the source positions or energies are different within a permissible limit given by the Rayleigh criterion.

All atoms in the specimen vibrate thermally around their equilibrium positions. The electrons whose waves are elastically scattered from the atoms at their thermal equilibrium positions keep both time coherency and spatial coherency with each other. They give rise to Bragg reflections. Actually each atom is differently displaced from the equilibrium position at different moments and does elastically scatter electrons. Thus, the waves of different electrons scattered by this atom are partially incoherent with each other. The frequency of the thermal lattice vibration $\sim 10^{13}$ Hz is the fifth power of 10 times as small as a frequency of incident electron wave $\sim 10^{18}$ Hz, so that the displacements are observable using the electron beam. The scattering from the displaced atoms, thus, gives diffuse intensity distribution, which is thermal diffuse scattering (TDS). We use the
Einstein model which assumes atoms moving independently with each other and the mean squared displacement of them. This leads to the Debye-Waller factor \[42,43\]. The lattice vibration may be quantized as phonon. The incident electrons interact with the phonons, losing or gaining energy \(\Delta E\), during the collision. The energy of phonons is in the order of \(10^{-1}\) eV or less (~infrared radiation) and that of the incident electrons is \(10^5\) eV so that \(\Delta E\) is negligibly small. In any case, the Bragg scattering is elastic, coherent scattering, while TDS is treated as elastic, incoherent scattering. Then, their intensities are given by the formulae shown in Figure 3.

In atomically resolved HAADF-STEM, the contrast of an atom column (along the incident electron beam probe) may be simply obtained as the integrated intensity for a time exposed to the probe, since the atoms in the column vibrate thermally and scatter electrons incoherently as TDS. This is one of the reasons why HAADF-STEM imaging is incoherent, and the contrast may be called ‘longitudinally incoherent’. Another reason of the incoherent imaging is ascribed to scanning of the convergent beam. The electrons scattered from an atomic column might be different with those scattered from different atomic columns because of time difference. Therefore columns separated laterally are imaged incoherently, that is, STEM image is also transversally incoherent.

The above-argument about ‘coherency’ is based on quantum physics and very strict \[32\]. However, according to classic wave

\[
I_x^{BS}(s) = |f_x(s)|^2 \exp[-2M_x(s)]
\]

\[
I_x^{TDS}(s) = |f_x(s)|^2 \{1 - \exp[-2M_x(s)]\}
\]

\(f_x(s)\) : Atomic structure factor for atom \(X\).
\(M_x(s)\) : Debye-Waller factor for atom \(X\).
\(s = (\sin \theta / \lambda)\) corresponding to the scattering angle \(2\theta\).

TDS cross-section for a detector range:

\[
\sigma_x = \left( \frac{4\pi(m/m_0)}{2\pi/\lambda} \right)^2 \int_{\text{detector}} |f_x(s)|^2 \{1 - \exp[-2M_x(s)]\} ds^2
\]

\(x\) : element.
\(\theta\) : scattering angle.
\(\lambda\) : electron wavelength.

**Figure 3.** Intensities of Bragg scattering (BS) and thermal diffuse scattering (TDS), and differential cross section of TDS. (a) Ewald construction for convergent electron beams with semiangle \(\alpha=6\) and 12 mrad. (b) Intensities of BS and TDS for a single N, Ga and In atoms. (c) Differential cross sections or intensities detected in a range of 30-190 mrad (shown in (b)) of BS and TDS for a single N, Ga and In atoms.
The waves that can interfere with each other used to be defined as being coherent. When different objects (for example, atoms) are illuminated with a plane wave, the scattered waves from the objects, which are considered to be resources, are coherent because they have a definite phase relationships. A perfect incoherent image results from a self-luminous object. Therefore, the incoherent imaging can be obtained when the convergent beam illuminates only a single object, because it emits the scattered wave as a self-luminous resource. In case of atomic-resolved electron microscopy, whether it is coherent or incoherent imaging depends on whether the Airy discs of neighboring atomic columns have permanent and definite phase relationships or not, applying the concept of Lord Rayleigh [1]. This is another and prevailing answer for why (HAADF) STEM is incoherent imaging. This definition is correlated with the transversally incoherent imaging due to the beam scanning mentioned above.

The incident electrons are scattered by atoms in the specimen as Bragg scattering and TDS. As seen in Figure 3a, reciprocal points in the zero-order Laue zone (ZOLZ) cross with one of Ewald spheres within a scattering angle of $2\alpha$, where $\alpha$ is a semiangle of the probe, and strong Bragg reflections (as the convergent beam electron diffraction) are caused, accordingly [17]. At higher angles ($\gg 2\alpha$), reciprocal points in ZOLZ and low higher-order Laue zone are not on any Ewald sphere, which means that no strong Bragg reflections occur. Furthermore, the intensity of the Bragg scattering reduces greatly with increasing scattering angle as shown in Figure 3b, comparing with TDS intensity. The incident electron probe shown in Figure 1 scans over the surface of the specimen in STEM mode. The HAADF-STEM image is the focal signal, as a function of the probe position, by electrons scattered in a high-angle range and captured on the annular detector. The differential cross section $s$ or the intensity of

**Formalism of HAADF STEM imaging**

Probes $P(R, R_o) = \int_{\Omega} \exp \{ i \mathbf{k} \cdot (R-R_o) + iW(k) \} \, dk$.

'Probes' mean within semiangle $\alpha$.

$W$: Lens aberration function or transfer function

$W(k) = \pi \lambda \left| k_\parallel \right|^2 \left( \Delta f + \frac{1}{2} G_\parallel \lambda^2 \left| k_\parallel \right|^2 \right)$

Image intensity $I^{HAADF}(R_o) = \sum_i \int_0^\infty \sigma_i \left| \sum_i \mathcal{A} (R_i - R_o, z) \right|^2 \, dz$

$\sigma$: Cross section for high-angle TDS (Einstein model)

$\mathcal{A}$: Convergent wave function

$\mathcal{A} (R_i - R_o, z) = \sum_i \psi (k) \mathcal{T} (R_i, k) \exp \{ i (K + \gamma) z \} \exp (-\mu z) \times \exp (-i k_\parallel R_o + iW(k)) \, dk$.

$\gamma$: Excitation amplitude

$\mathcal{T}$: 2-D Bloch wave

$K$: Transverse energy

$\mu$: Absorption coefficient

Figure 4. Formalism of HAADF-STEM imaging. $R_o$: the center of the beam probe. $k_\parallel$: transverse component of each plane wave having a wave length $\lambda$ within the incident electron beam. $\mathcal{A} (R_i - R_o, z)$: the wave function at a depth of $z$ in $i$-column, due to the probe located at the surface $(R_o, 0)$. $t$: specimen thickness.
TDS recorded with the detector is given by an equation shown in Figure 3. As shown in Figure 3c, the TDS cross section is approximately proportional to the square of the atomic number $Z$, which comes from the Rutherford scattering for the electron. TDS is an origin of unfavorable background, particularly in a high angle range, in electron diffraction and X-ray diffraction, but it is actively used in HAADF-STEM imaging. Bragg reflections scarcely influence HAADF-STEM imaging. We have confirmed this by simulation of the high-angle DF and middle-angle DF images, which explained the observed images of Si crystals [12].

Based on this principle, one can write STEM image simulation programs. We developed a program for HAADF-STEM simulations based on the Bethe method [8, 9]. The probe function and the image intensity can be given by the equations shown in Figure 4. Then, we proposed an algorithm for first calculations of Bragg scattering images and TDS images [12], which is based on the Bethe method using two kinds of optical potentials [43] and available to a routine work. The important optical parameters in HAADF-STEM imaging are the semiangle of the electron probe $\alpha$, the defocus of the probe forming lens $\Delta f$ and the angle range of the annular detector $D$ as well as the spherical aberration coefficient of the lens $C_s$ and accelerating voltage $V$ of the electron microscope used. These effects were examined and discussed in detail in our previous papers [8-20]. A typical example of the effects is artifacts in HAADF-STEM images, which are dealt with in the next section.

### 2.2 Artifacts in HAADF-STEM Images

It was believed that HAADF-STEM images exhibit strong atomic number ($Z$) contrast and can be directly inverted to the object without the need for image simulation, because the phase problem is eliminated [1]. However, we found experimentally and theoretically that bright spots appear on channels between the atomic columns in [110] Si images [11,12] and extraordinary intensity of spot contrasts occurs on light columns in [001] SrTiO$_3$ images [13], when the incident probe is focused to have strong subsidiary tails around a sharp central peak. Figure 5 reproduces some of simulated images of [110] Si crystals [19]. 'Total' means STEM images to be observed at given imaging conditions shown by parameters; detection range $D$, defocus value $\Delta f$ and thickness $t$. The images calculated for $D=60\sim160$ mrad are of HAADF-STEM and the images for $D=18\sim48$ mrad are of middle angle (MA) ADF-STEM. The Si-dumbbells are completely resolved but the artificial spots appear in the HAADF-STEM images at $\Delta f = -70$ nm. It was explained from wave field calculations that this is caused by the sharp probe but having strong subsidiary maximum, which is shown in Figure 6a [12,19]. The reason is that when the incident beam is located on channels between the atomic columns, the subsidiary tail forms strong wave fields along the neighboring columns to emit TDS. These artifacts might mislead the structure analysis unless one finds them from image simulation or predicts them from calculation of the probe function. The probe at $\Delta f = -30$ nm is so wide (having no appreciable subpeaks) that cannot distinctly resolve Si-dumbbells.

The shape of probe depends on $C_s$, $\Delta f$ and $\alpha$, as seen from formulae in Figure 4. An optimum imaging condition, which makes the effect of diffraction aberration and spherical aberration minimum, is $\alpha_{opt} = 2^{1/2}(\lambda C_s)^{1/4}$ and $\Delta f_{opt} = -(-\lambda C_s)^{1/2}$, and provides the most quality HAADF-STEM image. Probe functions and simulated images
Figure 5. Simulated high-angle ADF images (a) and middle-angle ADF images (b) of [110] Si crystals 20 nm and 60 nm thick, focused at $\Delta f = -30, -50$ and $-70$ nm using a semiangle of $\alpha=12$ mrad with a probe forming lens of $C_s=1.0$ mm at 200 keV. A defocus of $\Delta f = -50$ nm corresponds to the optimum focus of this lens. Images due to Bragg scattering (BS) and TDS were calculated separately, and then the total image which means an image to be observed was evaluated. CPU time for any one BS image is about 5 min on a PC (CPU of Intel Pentium 4 (2001) 2 GHz/400 MHz System Bus, 256 KB Cache). The CPU time for any one TDS image is twice as long as that for the BS image because it is necessary to calculate the intensities using two kinds of optical potentials. Additional calculation of an image at different thickness or focus takes about 11 s. Then, these HAADF and MAADF image simulations were done within 20 min [(5 min + 11 s $\times$ 5) + (10 min + 11 s $\times$ 5)].
around the optimum conditions are shown in Figure 6b and Figure 7[19], respectively. As far as the lens of Cs as large as 1.0 mm is employed, the appearance of the artificial spots is inevitable.

Cs-correctors have now been developed [44,45]. (A recent STEM instrument installed at Institute for Chemical Research, Kyoto University, has achieved a corrected value of Cs of -0.025 mm and a resolution power R of ~0.09 nm for α=23 mrad, for the lens with Cs=0.5 nm and R=~0.14 nm for α=11 mrad *). In a previous paper, we discussed the resolution of the ideal lens in terms of the contrast transfer function [17]. The Si dumbbells are completely resolved in an image calculated for the lens of Cs=0.1 mm, as shown in Figure 7b. However, a small deviation from the optimum focus, particularly to over-focus side, forms very low contrast image, whose Michelson visibility, defined as 
\[C = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}},\]
is as small as 0.133 for the [110] Si 60 nm thick, while C=0.566 at the optimum focus. This occurs with the illumination of larger area, which is caused by focusing the probe with the large semiangle above or below the surface of the crystal (see Figure 6b). The Cs correctors allow large α for high resolution, but may cause unfavorable Bragg scattering in HAADF-STEM images. Certainly, the Bragg scattering for the 0.1-mm Cs lens influences the resultant images more than that for the 1.0-mm Cs lens does in the same detection range of D=25~45 mm, as seen in Figure 7. However, HAADF-STEM generally uses the detector in high-angle range, and the strong incident probe due to use of the large semiangle allows us to make the detection range much higher, which would eliminate the coherent scattering influence.

Concerning the dependence of D, we showed experimentally and theoretically that the contribution of the Bragg scattering is so small that can be neglected in HAADF-STEM image, but that it cannot be neglected in MAADF STEM image [11]. In Figure 5, the simulations of Bragg scattering images and TDS images are shown, separately, together with the total images, for D=60~160 mrad and 18~48 mrad (and also in Fig 7, the simulations for D=25~45 mrad). The multiplicity of intensity indicated in figures shows that the Bragg scattering, which forms very complicated image contrast depending on focus and thickness, contributes more greatly to the total images as D is settled in smaller angle range.

Each atom ejects TDS electrons, depending on the dynamical wave field of the incident electrons in the crystal. The dynamical wave field is represented by the equation \(A'\) in Figure 4. It oscillates and attenuates in the crystal. Since the contrast of the HAADF-STEM image is given by a simple sum of intensities of TDS electrons collected on the detector, it monotonously increases with increasing thickness t [11], as seen in Figures 5 and 7. The influence of defocusing conditions and thickness on the experimental HAADF-STEM image was also illustrated using ceramic materials [16].

The lattice distortion around precipitates or substitutes influences the induced wave field in the crystal and causes unexpected image intensity from the simple Z-contrast, accordingly. The effect of the lattice distortion on the HAADF-STEM images was examined using SrTiO3. It should be taken into account for the image simulation of quantitative composition analysis [10,13] (see Figure 12). It was also found that a small tilt of the crystal zone axis with respect to the coma-axis of the probe-forming lens causes asymmetric dumbbells in the [110] Si image [14].

* Private communication from Prof. H. Kurata, Institute for Chemical Research, Kyoto University.
Therefore, special care to the beam alignment or the image simulation is needed for quantitative HAADF-STEM applications such as the determination of crystallographic polarity or the compositional analysis.

Thus, HAADF-STEM image contrast does not always depend on the number or atomic number of atoms in columns, or does not always show Z-contrast. The image simulation allows correct interpretation and quantification of atomic-scale HAADF-STEM images.

2.3 HAADF-STEM Image Processing

An experimental raw HAADF-STEM image can be processed by noise filtering through filtering and reconstructing of its diffractogram, which is a way generally used for HRTEM investigation. The processed image is served to compare with the calculated image. Recently, Rečnik et al. [46] have developed a new method for processing distorted HR-STEM images which is based on finding the displaced vertices in the experimental image and warping to geometrically correct reference grid of the object. We applied this method to analysis of the ZnO:Sb structure [20, 22].

HAADF-STEM image intensity \( I(R) \) may be given by convolution between the probe function \( P(R) \) and object function (projected atom structure) \( O(R) \), that is,

\[
I(R) = \int P(R') O(R-R') dR' .
\]

Figure 6. Various probes. (a) Calculated probe functions of \( \alpha=12 \) mrad focused at \( \Delta f = -70, -50 \) and \(-30 \) nm with the lens of \( C_s=1.0 \) mm at 200 keV. (b) Calculated probe functions of \( \alpha=10 \) mrad focused at \( \Delta f = -50, -40 \) and \(-60 \) nm with a lens of \( C_s=1.0 \) mm at 200 keV (left) and of \( \alpha=18 \) mrad focused at \( \Delta f = -16, -6 \) and \(-26 \) nm with a lens of \( C_s=0.1 \) mm (right). The one sides of the probes are presented. The optimum condition is \( \alpha_{\text{opt}} =10 \) mrad and \( \Delta f_{\text{opt}} = -50 \) mm for \( C_s=1.0 \) mm, and \( \alpha_{\text{opt}} =18 \) mrad and \( \Delta f_{\text{opt}} = -16 \) mm for \( C_s=0.1 \) mm. Calculations were made using a formula indicated in Figure 4.
Figure 7. Simulated middle-angle ADF images of [110] Si crystals 20 nm and 60 nm thick, focused at $\Delta f = -40$, -50 and -60 nm using $\alpha=10$ mrad with the lens of $C_s=1.0$ mm at 200 keV (a) and at $\Delta f = -6$, -16 and -26 nm using $\alpha=18$ mrad with the lens of $C_s=0.1$ mm (b). The CPU time for the simulation of images in (a) is almost the same as the simulation of the images shown in Figure 5. Since the semiangle is large, it takes a long time for the simulation in (b). The CPU time for a BS image is about 2 h and that for a TDS image is about 4 h. Additional calculation of an image at different thickness or focus takes about 11 s. Then, all these calculations were done in about 6 h [(2 h +11 s $\times$ 5) + (4 h + 11 s $\times$ 5)].

$I(R)H^\dagger P(R)O(R)$, under an approximation of nondispersive Bloch waves. Then, an experimental HAADF-STEM image that is deformed by the incident probe can be deduced to the object function by the deconvolution, $O(R) = \mathcal{F}^{-1}\{\mathcal{F}[I(R)]/\mathcal{F}[P(R)]\}$, where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent Fourier and inverse Fourier transformation, respectively (Figure 8). We have developed a deconvolution processing technique, where the defocus value...
assigning the effective probe function \( P(R) \) is determined using the maximum entropy method [15]. Images having unresolved Si dumbbells or the artificial bright spots are retrieved to the projected atomic structure (Figure 9). Thus, the deconvolution processing is completely different from any conventional image processing method such as noise filtering through Fourier transform, smoothing or averaging of images, which cannot eliminate the influence of the probe, for instance, the artificial spots.

HAADF-STEM images recorded with field-emission electron microscopes were often deformed by instrumental and environmental instability. The deconvolution processing was applied to retrieve a systematically distorted image by referring a HRTEM image [18]. Figure 10 shows a deconvolution processing which successfully derived the \( Z \)-contrast image of a \( \text{Zn}_2\text{Sb} \)-monolayer inversion twin-boundary from a distorted HAADF-STEM image of \( \text{Sb}_2\text{O}_3 \)-doped \( \text{ZnO} \).

### 2.4 Quantitative Composition Analysis by High-resolution HAADF-STEM

Sections 2.2 and 2.3 imply that we have two approaches to the structural and chemical analysis by HAADF-STEM; the image simulation and the retrieval image processing. These two approaches have been discussed, illustrating images of Sb-rich inversion boundaries in \( \text{ZnO} \) ceramics [20]. It was shown that the artifacts may mislead the structural analysis unless one finds them from image simulation or eliminates them by an appropriate image processing. Paradoxically speaking, the simulation of HAADF-STEM images provides the structural and compositional information of every atomic column along the incident beam. In fact, we have reported HAADF-STEM concentration analyses of As atoms along every [110] column in As-doped Si wafers (Figure 11) [8], and Bi atoms along every [001] column in the boundary layer in \( \text{SrTiO}_3 \) ceramic condensers (Figure 12) [9,10], with the aid of the image simulation.

### 3. ELECTRON MICROSCOPY INVESTIGATIONS OF GAN-BASED LIGHT EMITTING DIODES AND LASER DIODES

The lifetime of LDs with multiple InGaN/GaN quantum wells has attained to more than 15,000 h [38, 47]. This was achieved thanks to the epitaxial lateral overgrowth (ELOG) method for a \( \text{GaN:Si} \) layer on the sapphire substrate, which reduces dislocation density in the \( \text{GaN:Si} \) layer [48,49], and to the development of strained-layer superlattices (SLSs) of \( \text{AlGaN/GaN} \) for cladding [38]. The nanostructures of MQW InGaN/GaN active layer and SLS cladding layers greatly influence the final laser properties. However, there have been very few structural investigations of the InGaN/GaN layers and AlGaN/GaN layers. Figure 13 illustrates one of the LD devices we used for electron microscopy observations [50]. Figure 14 reproduces a HAADF-STEM image of a prototype ultra-violet LED wafer [35]. Al\(_{0.005}\)In\(_{0.02}\)Ga\(_{0.975}\)N QWs (bright) and Al\(_{0.12}\)In\(_{0.005}\)Ga\(_{0.875}\)N:Si barriers (dark) are definitely resolved in the MQW active layer, and \( n \)-type and \( p \)-type cladding layers of Al\(_{0.15}\)Ga\(_{0.85}\)N/GaN SLSs can be seen. It may be noted that the results shown in section 3 would not be obtained by any conventional microanalysis technique such as HRTEM alone, but were successfully achieved by using HAADF-STEM.
Formalism of the deconvolution process

Under non-dispersion approximation, image intensity $I(R)$ can be described by convolution:

$$I(R) = \mathcal{F}^{-1}\{\mathcal{F}(I(R))\} = \mathcal{F}^{-1}\{\mathcal{F}(O(R))\} = \mathcal{F}^{-1}\{\mathcal{F}(\Delta f)\}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent Fourier and inverse Fourier transformation, respectively.

$O(R)$: Object function or Projected atom structure
$\Delta f$: Fourier transformation of eq 1 is $\mathcal{F}[f(R)] = \mathcal{F}[\mathcal{F}^{-1}\{\mathcal{F}[O(R)]\}]$

so that object function is obtained by

$O(R) = \mathcal{F}^{-1}\{\mathcal{F}[O(R)]\} = \mathcal{F}^{-1}\{\mathcal{F}[\mathcal{F}^{-1}\{\mathcal{F}[\Delta f]\}]\}$

In this processing, $\Delta f$ of effective probe function $\mathcal{F}^{-1}\{\mathcal{F}[\Delta f]\}$ is determined by maximum entropy method.

Figure 8. Formalism of deconvolution process. $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent Fourier and inverse Fourier transformation, respectively.

Figure 9. Deconvolution processing of HAADF-STEM images. HAADF-STEM image can be regarded as a convolution of the object function (projected atom structure) and the effective probe function, which is schematically shown for two typical probes. The effective probe function is not the real one. Observed images in (a) were retrieved to Z-contrast images in (c) by the deconvolution. The effective probe functions used for the deconvolution were obtained from $\Delta f$ corresponding to the maximum entropy values shown in (b). (The images were taken with JEM 2010F-TEM/STEM ($Cs=1.0$ mm), at $V=200$ keV, $\alpha=6$ mrad, and $D=50\sim110$ mrad.)
Figure 10. Deconvolution processing of a [010] HAADF-STEM image of an Sb-rich basal plane inversion boundary in Sb₂O₃-doped ZnO ceramics. Left: ZnO in the matrix. Right: Inversion boundary. HAADF-STEM images (a) were additionally processed using the Savitzky-Golay smoothing. The process includes correction of the diffractograms of HAADF-STEM images (c) by referring to the diffractograms of HRTEM images (d), as a pre-treatment of the deconvolution. The corrected diffractograms (e) were used as [I(R)] for the deconvolution. The reconstructed images (f) were obtained by inverse Fourier transformation of (e), and are the representatives of the experimental images, noise-filtered but including the effect of the probe. The most probable defocus for the images was determined to be Δf = -36 nm from the entropy calculation for the matrix image. Retrieved HAADF-STEM images (g) were obtained by deconvolution using a probe at Δf = -36 nm, and exhibit Z-contrast, indicating Zn columns (bright spots) and Sb columns (brighter spots). FT: Fourier transformation. IFT: inverse Fourier transformation. (The images were taken with JEM-2010F TEM/STEM (Cs=0.48 mm), at V=200 keV, α=10 mrad, and D=100~220 mrad.)

Figure 11. Two-dimensional distribution of As atoms doped in a Si crystal. (a) Experimental (noise-filtered) [110] HAADF-STEM image of a pure region in the Si wafer. Si dumbbells are not completely resolved under this imaging condition. (b) Experimental (noise-filtered) [110] HAADF-STEM image of a region implanted with 1.9 at. % As. Insets are evaluated number of As atoms substituted for Si atoms on the 65 atom sites along each [110] atomic column in the crystal 25 nm thick. (c) The most fitted simulated intensity profiles with experimental profiles, along rows 1-8 in (a and b). Experimental intensity profiles are not shown in this figure (See ref. 8). Si dumbbells in (a) are symmetric (see fine curves), which means the atom sites are occupied only by Si. Si dumbbells in (b) are asymmetric due to the substitution of As atoms (see thick curves). The thickness and the numbers of As atoms in (b) were determined from the simulation of the intensity profiles. (The images were taken with JEM 2010F-TEM/STEM (Cs=1.0 mm), at V=200 keV, Δf = -50 nm, α=6 mrad, and D=50~110 mrad.)
Figure 12. Atomic-scale quantitative elemental analysis of boundary-layer semiconducting SrTiO$_3$ ceramic condenser. (a) TEM image of a grain boundary having Bi diffusion layers of insulators. (b) Atomic resolved (001) HAADF-STEM image near the grain edge. (c) Enlarged image of a part of (b). (d,e) Bi concentration in Sr and Ti(O) columns near the edge, and observed and calculated HAADF image intensity of Sr, O, and Ti(O) columns. The intensity profiles are along X-Y on the (100) Sr-O-Sr layer and X’-Y’ on the (100) O-Ti(O)-O layer in (c). The observed profiles were averaged along (010) layer over the area in (c). The calculated profiles from a preliminary simulation which took no account of the lattice distortion are also shown together with the Bi concentration used in this preliminary simulation (in bracket). (The image was taken under the same condition as that shown in Figure 11 with JEM 2010F-TEM/STEM (Cs=1.0 mm)).

Figure 13. (a) Schematic structure of a violet LD with In$_{0.25}$Ga$_{0.75}$N (2.5 nm)/GaN (8nm) MQWs and Al$_{0.14}$Ga$_{0.86}$N/GaN SLS cladding layers, prepared by production-scale metalorganic vapor phase epitaxy. (b) Room temperature emission spectra of the 10 $\mu$m × 10 $\mu$m ridge-waveguide InGaN LD measured under pulsed operation. At the injection pulsed current above a threshold current of 0.87 A, a strong stimulated emission appears at around 415.5 nm with a FWHM of 0.2 nm, as well as two small emission peaks between 416 and 417 nm. (Ref. 50).
3.1 V-defects in InGaN MQWs and Light Emission Mechanism of GaN-based Devices

Even low magnification HAADF-STEM images gave certain information about the structure and composition of the ultrathin InGaN layer, as seen in Figure 14. The images shown hereafter are from prototypes of the LED or LD devices where some parts of the heterostructures were not deposited. Figure 15a shows a low magnified image of capped MQW In\textsubscript{0.25}Ga\textsubscript{0.75}N (2.5 nm) and GaN (13.9 nm) layers deposited on underlying GaN:Si layer. Bright bands are In\textsuperscript{31}Ga\textsuperscript{49}N\textsubscript{7} layers and dark ones are Ga\textsuperscript{31}N\textsubscript{7} layers, as predicted from Figure 3c. This was confirmed by energy dispersive X-ray spectroscopy (EDX) [26]. High-resolution HAADF-STEM image, processed by the deconvolution method [15], clearly defined the interface between InGaN and GaN layers at the atomic scale [26], as seen in Figure 16a.

Threading dislocations (TDs), which were formed by the big lattice misfit between the GaN and sapphire, are seen with bright contrast due to large distortion around the defects (Figure 15a). HAADF-STEM also revealed V-defects or inverted hexagonal pyramid defects (IHPs). The names come from the empty V-pit in hexahedron cone sharp, with six sidewalls on \{101\} planes, which nucleates at a TD extending along the [0001]-axis growth direction, in the InGaN layers [51-53]. We demonstrated that the V-defects grow in the form of a thin six-walled structure with InGaN/GaN \{101\} layers (Figures 15b and c) [28]. This structure had been predicted by Wu et al. [52], Sharma et al. [54] had reported the InGaN MQWs ending abruptly at the V-defect interfaces without these thin InGaN/GaN layers. Therefore, we were the first to establish the structure of V-defects having thin sidewalls of InGaN/GaN QWs, which are different from the main \(c\)-plane QWs. Then, we succeeded in distinguishing the MQW In\textsubscript{0.25}Ga\textsubscript{0.75}N (2.5 nm) and GaN (13.9 nm) layers in a secondary electron image and a backscattered electron image in an HR-FESEM, and confirmed the thin six-walled structure of the V-defect [40]. We also found that the V-defects start even at In-rich dots, as seen at B in Figure 15b [28]. On the basis of these observations, we have proposed a formation mechanism of the V-defects, taking into account the growth kinetics of the GaN crystal and the masking effect of In atoms on the analogy of ELOG [33, 36]. The In-rich area at B in Figure 15b may be regarded as an example of the masks.

A green LD having thick QWs and barrier layers of In\textsubscript{0.18}Ga\textsubscript{0.82}N (4 nm)/GaN (40 nm) has been observed [36]. Figures 15d and e are a CTEM image and a HAADF-STEM image of a V-defect in the QWs of the green LD, respectively. The HAADF-STEM image provided undoubted evidence of the thin InGaN/GaN \{10\overline{1}\} MQW walled structure of the V-defect. A bright spot in Figure 15d or strong contrast spot in Figure 15e that appears at the apex of the V-defect indicates an In-rich region. Indium atoms were trapped and segregated in the strained field (Cottrell atmosphere) around the core of the incorporated TD and then formed a small mask for the growth of V-defect. The curved corners of the thin MQW walls with the main MQWs seen in Figures 15d and e were explained on our formation mechanism of V-defects [36].

A HAADF-STEM image (Figure 16a), processed by two-dimensional smoothing and deconvolution, provides precise positions of the atomic columns (Figure 16b) and clear atomic number dependent contrast, thereby allowing us to map both the strain field (Figure 16c) and In atom distributions (Figure 16d) [27]. The lattice expanded largely along the \(c\) axis as seen in Figure 16c but hardly in the \(\epsilon\)
The maps, thus, directly led to the observation that there is a local fluctuation of In atoms in the InGaN layers, and that the In-rich regions, considered to be quantum dots (QDs), correspond with lattice expansion along the c-axis. This was the first direct evidence of local phase separation or QDs.

The QDs are observed as diffraction contrast caused by the strain field (Figure 15d). Several reports, including the paper by Y. Narukawa et al. [55] who first proposed the QDs for explanation of the high light emission efficiency, had assumed indirectly the In concentration from the lattice strain only using the Vegard law approximation. They explained the high emission efficiency of InGaN-based LED to be mostly due to the large localization of excitons because the pathways of nonradiative recombination are hindered once the excitons are captured in a small volume. However, the mechanism of the high emission of InGaN-based LEDs and LDs has not been fully established, at the present becoming more and more confused by some reports doubting the In-rich quantum dots and their effect [56, 57]. In any case, the side-walled InGaN/GaN layers might strongly correlate with unfavorable weak long-wavelength emissions like the 416-nm and 417-nm emissions seen in Figure 13b [33, 36, 52]. The detection and control of these V-defects are, therefore, very important for fabrication of these LEDs and LDs.
3.2 Observations of Strained-layer AlGaN/GaN Superlattice Cladding Layers

The claddings of AlGaN/GaN strained-layer superlattices (SLSs) have been used for the following two reasons. Thick AlGaN cladding layers need for optical confinement, but cracks and dislocations induced by the lattice mismatch between GaN and AlGaN make the formation of the thick cladding layers impossible. It is the first reason to depress the formation of these defects by strain relief of the AlGaN/GaN SLSs [38].
Figure 16. (a) Experimental deconvolution-processed high-resolution HAADF-STEM image of an InGaN QW in the In$_{0.25}$Ga$_{0.75}$N (2.5 nm)/GaN (13.9 nm) layer. The defocus of the probe-forming lens was evaluated to be $\Delta f = -44$ nm in deconvolution process. (b) Projected atomic column positions derived from (a). The distortion hardly occurred along the [1100] direction. (c) Strain field maps along the [0001] direction. Arrowheads show the interfaces between GaN and InGaN layers. (d) Two-dimensional In atom distribution. There is local fluctuation of In atoms in the InGaN layers, and the In-rich regions enclosed, considered as quantum dots, cause large expansion along the [0001] direction. (The image was taken under the same condition as that shown in Figure 15 in Tecnai F30 ($Cs=1.2$ mm)).

There are very few reports on the structural analysis of AlGaN/GaN SLSs. It was confirmed by image simulation that HRTEM images exhibit no difference in contrast between Al$_{0.14}$Ga$_{0.86}$N and GaN in the cladding (Figure 17a) [31,34]. The TEM diffraction contrast due to small strain field appears along their interfaces but TEM images never provide the real widths of AlGaN or GaN layers. We have observed the SLSs (Figure 17b) which are 200 coupled layers of $n$-Al$_{0.14}$Ga$_{0.86}$N (3 nm) and $n$-GaN (3 nm) on the $n$-GaN:Si grown directly on a (0001) sapphire substrate, and distinguished the Al$_{13}$Ga$_{31}$N$_7$ and Ga$_{31}$N$_7$ layers as dark and bright fringes, respectively, in a [1010] HAADF-STEM image (Figures 17c) [34].

From a HAADF-STEM image shown in Figure 17e, the averaged thicknesses of the AlGaN and GaN layers were determined, with the aid of image processing, to be $2.24 \pm 0.09$ nm and $2.34 \pm 0.15$ nm, respectively, which corresponds to nine atom layers in the c direction [34]. The AlGaN lattice shrunk for $\sim 4\%$ along the c direction while no change was measured in the basal plane, as compared with the GaN lattice. This is a direct illustration of the strained lattice in the superstructure cladding layer. Namely, a good lattice matching in the basal plane between the AlGaN and the GaN suppresses the generation of misfit dislocations on the Al$_{0.14}$Ga$_{0.86}$N/GaN interfaces in the SLS as well as on the interface between the Al$_{0.14}$Ga$_{0.86}$N and GaN:Si layers.
Figure 17. N-type strain Al$_{0.14}$Ga$_{0.86}$N (3 nm)/$\mu$-GaN (3 nm) superlattice cladding layer that was grown on the $\mu$-GaN:Si deposited on a (0001) sapphire substrate. (a) HRTEM images of AlGaN and GaN layers. The AlGaN and GaN layers cannot be distinguished in the HRTEM images. (b) CTEM image of the whole specimen. (c) HAADF-STEM image of the whole SLS cladding layer. 200 pairs of AlGaN (dark bands) and GaN layers (bright bands) are clearly resolved. The AlGaN layers are numbered. (d) BF-STEM image of the same area. Dislocations are visible in strong diffraction contrast. (e) HAADF-STEM image, distinguishing between the AlGaN (dark stripes) and GaN layers (bright stripes). (f) HAADF-STEM image of the area I in (b). Dislocations are seen, exhibiting bright contrast due to lattice distortion. Note a dislocation turning its course from running along the $c$ axis to lying in the basal plane. It apparently cannot reach the MQW. As a result the number of defects in MQW is reduced. Semi-quantitative energy dispersive X-ray spectroscopy (EDX) nanoanalysis was performed along A, B, C, and D. (Observations were performed in JEM-2010F TEM/STEM ($C_s=0.48$ mm). $V=200$ keV. $\Delta f =$ the optimum condition. For HAADF-STEM, $\alpha=10$ mrad and $D=100\sim220$ mrad.)
Semi-quantitative EDX nanoanalysis for the areas B and D in Figure 17f confirmed the chemical composition of the Al$_{0.14}$Ga$_{0.86}$N [34]. Dislocations run either parallel or perpendicular to the $c$ axis. They appear as bright contours in HAADF-STEM images as a result of random static atom displacement around the dislocations, while they appear as dark contours due to diffraction contrast in BF-STEM images (Figure 17d). EDX nanoanalysis also revealed that the area A along a TD and the area C along a dislocation in the basal plane were Al-rich, suggesting the precipitation of Al along the dislocations. Figure 17f shows TDs reverted from running along the $c$ direction to running in the basal plane. The converted dislocations would not reach the upper MQW layer. This evidently shows an important role of the SLS in suppressing the defect propagation, which results in the overall reduction of structural defects in the active MQW layer.

The same SLS sample was examined in the FE-SEM [39]. The fringes due to SLSs were observed in the secondary electron images taken at voltages of 10~25 keV, but they could not be detected in the backscattered electron images. This is obvious because the mass difference between Al$_{0.14}$Ga$_{0.86}$N and GaN is so small that the image contrast due to the back scattered electrons or Rutherford scattered electrons is undetectable. The secondary emission yield is greatly depending on the surface state of materials so that we might resolve the Al$_{0.14}$Ga$_{0.86}$N and GaN layers in the secondary electron images. We also observed the $p$-Al$_{0.14}$Ga$_{0.86}$N(3.0 nm) / $p$-GaN(3.0 nm) structure at the same imaging conditions. Any fringe was detected neither in the secondary electron image nor in the backscattered one. Difference in the secondary emission yield between $p$-Al$_{0.14}$Ga$_{0.86}$N and $p$-GaN seems to be very small. Then, we tried a HAADF-STEM observation of an SLS cladding including thirty four pairs of $p$-Al$_{0.14}$Ga$_{0.86}$N and $p$-GaN:Mg layers [35]. The Al$_{0.14}$Ga$_{0.86}$N layers and the GaN layers were distinguished as dark and bright bands, respectively. The average thickness of the pair was roughly estimated to be $\sim$12 nm, which corresponds to 48 atom layers in the $c$ direction.

### 3.3 Ultra-high Density InGaN Quantum Dots

The performance of LDs has been expected to greatly improve if the density of QDs can be made to far exceed that of dislocations. Therefore, structures of the active layer with QDs, instead of QWs, have been pursued. One approach for fabricating the 3D InGaN QDs is the deposition of silicon anti-surfactant or SiN$_x$ nano-masks, which altered the morphology of the InGaN films from that of step flow to that of a three-dimensional island. We have observed a device of ultra-high density InGaN QDs [37] that was prepared by Tu et al. [60]. It exhibited strong PL emission at room temperature. Figure 18a shows its structure together with growth condition. Masks with a rough SiN$_x$ layer were formed on the GaN:Si underlying layer, at flow rates of 5 slm for NH$_3$ and 50 sccm for diluted Si$_2$H$_6$. The deposition of a pair of the SiN$_x$ and In$_x$Ga$_{1-x}$N was repeated using a spacer of the GaN:Si barrier layer deposited at 1000°C, and finally the last In$_x$Ga$_{1-x}$N layer was capped with an undoped GaN layer. All the layer thicknesses shown in Figure 18a are nominal values. Figure 18c shows a low magnified HAADF-STEM image of the specimen, in which the areas for EDX analysis are indicated. Figure 18d shows EDX spectrum in the area P, which comprises emissions from In, Ga, N, and Si. Figures 18e and f (or Figures 18g and h) are the line profiles of HAADF-STEM image intensity and In-L$_\beta$ emission intensity along
A-B (or C-D), respectively. The bright area along C-D in Figure 18c might be a layer containing InGaN QDs (QDs layer). From Figure 18f the width of this layer is estimated to be ~15 nm. Figure 18h indicates that very roughly estimated width of the QDs is 10 nm on the C-D.

Figure 18b shows a HAADF-STEM
image near the top of the specimen. The dark line is identified, due to the small thermal diffuse scattering cross-section of Si$_{14}$ (and N), as the projection of a layer composed of SiN$_x$ mask islands. The height of the SiN$_x$ masks is roughly estimated to be less than 2 nm from the thickness of the dark line. The In$_{49}$Ga$_{31}$N$_7$ QDs are observed as bright triangles, which look peaks in a mountain chain, in the QDs layer between two-dotted lines. Figure 18i shows an HRTEM image near the specimen edge, where a part of the capping layer was removed during the ion-milling. Quite high QDs (indicated by black arrowheads) are seen on the SiN$_x$ mask layer (indicated by an white arrowhead). The dark contrast for the QDs may be caused by diffraction effect of the high structure amplitude of In and large lattice strain. The rectangle frame in Figure 18i encloses an area containing the SiN$_x$ layer and a QD and the square frame encloses an area in the GaN barrier layer. In Figure 18j the former area is inset on the latter area. The SiN$_x$ mask surely has a heavily strained lattice. The images do not allow us to determine the crystal structure and the composition x of the SiN$_x$. The lattice in the QD was coherently connected with the GaN lattice in the capping layer although it was also heavily strained.

The QDs can be regarded as nanoisland crystals with the \{10\overline{1}\} side walls and a height of several nanometers or more. This crystal habit can be understood in taking account of ELOG or the crystal growth through masks at lower reactor temperature [33,36]. Thus, we can propose the structure for this three-dimensional InGaN quantum dots as the inset in Figure 18a.

In conclusion, functional devices, such as quantum well lasers, tunneling devices, random access memories, and high electron mobility transistors, usually comprise nanoscale multilayer heterostructures that exploit quantum confinement effects. Since the in situ monitoring in most processing equipment is still absent, atomic-scale analysis is required for understanding the real structure in the final product, including the thickness and the composition of each layer. Only then can improvement of the functional properties and control of the product’s quality be carried out. HAAD-STEM has become one of the most powerful analytical tools for characterization of these nanostructures because it enables us to make atomically-resolved quantitative composition analysis. The other part of my lectures can be acquired through our review papers [61-63].

ACKNOWLEDGEMENTS

Thanks are owing to collaborators, particularly coauthors of our papers cited in references, and also deeply to Dr. Torranin Chairungsri, Department of Industrial Chemistry, Chiang Mai University, for motivating me to write this paper.

REFERENCES


