Spatial Inhomogeneity of Luminescence in III-Nitride Compounds

Gintautas TAMULAITIS∗

Semiconductor Physics Department and Institute of Applied Research, Vilnius University, Sauletekio 9 – III, LT-10222 Vilnius, Lithuania

corresponding to the band gap of AlN (see [1–3] for review).

*Corresponding author. Tel.: +370-5-2366071, fax.: +370-5-2366070. E-mail address: gintautas.tamulaitis@ff.vu.lt (G.Tamulaitis)

ISSN 1392–1320 MATERIALS SCIENCE (MEDŽIAGOTYRA). Vol. 17, No. 4 2011

1. INTRODUCTION

The high-brightness light-emitting diodes (LEDs) with worldwide market of up to $11 billion in 2011 became an important substitute for conventional lamps in many applications. A substantial breakthrough in LED applications came after the development of blue and white LEDs with active layers based on InGaN and GaN heterostructures. Introduction of aluminum into ternary AlGaN, InAlN or quaternary AlInGaN compounds enables growth of epilayers with wide band gaps of up to 6.2 eV corresponding to the band gap of AlN (see [1–3] for review), while increasing content of indium in InGaN shifts the band gap of this compound down to 0.77 eV, the band gap of binary InN (see [4] for review).

Though the lattice constant of the ternary compounds A_xB_{1-x}N, where A and B stands for In, Ga, and Al, linearly depends on the content x (Vegard’s law), the composition dependence of the band gap is nonlinear and can be approximated by empirical expression

\[ E_g^{ABN}(x) = (1-x)E_g^{BN} + xE_g^{AN} - bx(1-x), \]  

where \( E_g \) denotes the band gap of the corresponding compound. At room temperature, 0.77 eV, 3.42 eV, and 6.13 eV are the band gaps of InN, GaN, and AlN, respectively. The bowing parameter b depends not only on compound composition but, possibly, also on strain, which strongly depends on substrate material, buffer layer used, growth conditions, layer thickness, etc. Thus, selection of the b value is ambiguous. For example, analysis of 20 references on b for AlGaN by Lee et al. [5], led to conclusion that b values can be grouped into four groups with b in the vicinity of quite different values of −0.8 eV, +0.53 eV, +1.3 eV, and +2.6 eV. Bowing parameters 2.4, 1, and 3.0 eV can be considered as typical for InGaN, AlGaN, and AlInN, respectively [1].

Whatever the value of b, the band gap of III-nitrides covers the entire range from infrared to deep UV, and LEDs emitting in this range can, in principle, be produced. However, growing of the compounds with any desirable composition often encounters substantial difficulties due to phase separation, structural quality of the epilayers, impurities and extended defects, etc. Moreover, even in the III-nitride epilayers and heterostructures without phase separation, a spatially inhomogeneous emission distribution is observed. The most obvious origin of the inhomogeneity is the band gap modulation due to fluctuations in compound composition. For example, locally higher indium content in InGaN crystal results in a narrower band gap in this location. Other mechanisms can also be responsible for potential fluctuations and are discussed shortly in this review.

The most straightforward approach for study of the spatial emission inhomogeneity is application of photoluminescence (PL) or cathodoluminescence (CL) spectroscopy techniques enabling spatial resolution. These techniques will be introduced before discussion of the results obtained in InGaN and AlGaN epilayers and heterostructures.

2. TECHNIQUES FOR SPATIALLY-RESOLVED LUMINESCENCE SPECTROSCOPY

The simplest and least expensive technique for achieving spatial resolution in luminescence spectroscopy
is micro-photoluminescence (µ-PL). The excitation light is focused as tightly as possible, and the luminescence is collected from the excited spot. The spatial resolution in this approach is, first of all, limited by aberrations of the optical systems used. The aberrations can be diminished or compensated by using appropriate design and materials for the optical systems. Nevertheless, the minimal spot size is limited by self-diffraction even for ideal optical systems. According to the Rayleigh criterion, the minimal spot size for a planar wavefront \[ d_R = 0.61\lambda / NA, \] (2) where \( NA = n \sin \theta \) is the numerical aperture, \( n \) is the refraction index, \( \theta \) is the half-angle between the marginal converging rays of the light cone. For modern microscope objectives, \( NA \) values of up to \( \sim 1 \) can be achieved. Application of immersion liquids increases \( NA \) by \( n \): 1.33 for water, 1.56 for oil. Thus, the self-diffraction limit of the spatial resolution obtainable with high-quality objectives equals approximately half the wavelength of the light used. A typical spatial resolution in real µ-PL systems equals \( \sim 1 \) µm.

A higher spatial resolution can be achieved in scanning confocal microscopes (SCMs) [6–8]. The configuration of the confocal microscope is depicted in Fig. 1. The same objective is used both for focusing the excitation light and for collecting the luminescence. The image plane of the collection system coincides with a screen containing a tiny aperture directly on the optical axis of the system. Thus, only the light coming out from a small spot on the sample surface passes the aperture and is detected directly by a detector or is analyzed by a spectrometer. The lateral resolution can be approximated as \[ d_{xy} \approx 0.4\lambda / NA. \] (3)

A typical value for the lateral spatial resolution of conventional SCMs in the green region is \( \sim 200 \) nm. The surface of the sample under study can be scanned by moving the piezoelectric stage with the sample.

**Fig. 1.** Configuration of confocal microscope

Another advantage of SCMs is their capability of on-axis spatial resolution. As shown in Fig. 1, only the light emitted from the sample surface is focused into the aperture, while the light coming from the deeper parts of the sample forms a large, blurred-out image, and only a small part of it passes the aperture. However, the focal plane can be shifted deeper into the sample by moving the objective. Thus, a spatial resolution along the z-axis can be achieved. The resolution depends on the effective wavelength \( \lambda \), which is related to the wavelengths of excitation \( \lambda_{exc} \) and luminescence \( \lambda_{lum} \) as \( 1 / \lambda = 1 / \lambda_{exc} + 1 / \lambda_{lum} \), and can be approximated as

\[ d_z \approx 1.4\lambda n / NA^2. \] (4)

For \( \lambda = 0.5 \) µm and typical objective \( (NA = 0.9) \) in air \((n = 1)\), the \( d_z \) value equals 0.86 µm.

The axial spatial resolution can be increased by the factor of up to five by using a group of techniques called 4-Pi confocal. Interference of two excitation light beams is utilized in the 4-Pi confocal microscope.

Self-diffraction limits the spatial resolution of optical systems only in the far field, i.e., when the distance from the light source considerably exceeds the wavelength. The limit can be substantially decreased in the near field when applying the technique that is called either Scanning Near-Field Optical Microscopy (SNOM) or Near-Field Scanning Optical Microscopy (NSOM). An aperture with the diameter smaller than the wavelength blocks the light in the far field but the light extends outside such an aperture in the near field, approximately to the distance equal to the aperture diameter \( d << \lambda \). As illustrated in Fig. 2, making the aperture on the apex of a narrow tip enables scanning the sample surface with a high spatial resolution considerably exceeding the wavelength. Tapered metal-coated fused-silica single-mode optical fibers or hollow silicon cone tips are currently used in SNOM.

**Fig. 2.** Field distribution near the tip of a scanning near-field optical microscope (SNOM)

Spatial resolution in the devices containing apertures is basically determined by the aperture diameter. However, the throughput of the probe decreases as the sixth power of the aperture diameter. Thus, a trade-off between the spatial resolution and the intensity of luminescence available for detection has to be considered in planning experiments. Resolution reaching \( \lambda / 40 \) is demonstrated [9] but majority of the SNOM systems currently in use exploit apertures enabling resolution of \( \sim 100 \) nm (\( \sim \lambda / 5 \)). A good spatial resolution can be achieved in apertureless configuration devices, where a sharp metal tip is used as a probe. Due to electron oscillations in the probe tip excited by illumination in the far field, the light field directly below the tip is enhanced by many orders of magnitude, and the sample is effectively excited only there. Unfortunately, the background luminescence detected from the area outside the tip might be substantial and impede the detection of the desirable signal from beneath the tip.
The surface is usually scanned by moving the sample mounted on a piezoelectric sample stage. The tip in SNOM devices is maintained at a fixed distance above the sample surface by using the same feedback systems as those used in atomic force microscopy (AFM), which is a more mature measurement technique.

SNOM devices can be configured in different modes. In illumination (I) mode, the sample is excited through the probe in the near field, and the luminescence is detected in the far field, thus the emission is detected from the entire excited area that can exceed the excitation spot due to diffusion or other spreading mechanisms. In collection (C) mode, the sample surface is homogeneously excited in the far field, while the spatial distribution of the luminescence is probed by collecting the light through the probe in the near field. In illumination-collection (I-C) mode, both excitation and detection are accomplished through the probe.

Short pulse laser excitation and luminescence detection in time-resolved mode by exploiting streak camera are also used in SNOM devices [10]. A dual-probe SNOM system independently controlling the distances between probe and sample and between two probes provides additional information, as demonstrated in real time study of carrier diffusion in InGaN [11].

High spatial resolution and capability of simultaneous measurement of PL distribution and surface morphology can be pointed out as the main advantages of SNOM technique. However, SNOM did not become a routine technique is spectroscopy due to several difficulties. In addition to the small signal to be detected, since the excitation volume or/and the probed volume are small, comparability of the experimental results is a considerable problem in SNOM applications due to deviations in aperture diameters of different tips, roughness of the aperture edges or the metallic coating of the tip, and possible wearing of the tip during experiment.

According to equation (1), the excitation spot size can be diminished by using excitation with a shorter wavelength. In cathodoluminescence (CL), using excitation by electrons with energy $E$ corresponding to the wavelength $\lambda = hc/E$, which is substantially shorter than that of the light, enables a considerable gain in the spatial resolution. The spatial resolution in CL experiments can reach 1 nm and is usually restricted by instrumental limitations rather than by physical limits. Electrons in CL experiments are usually provided by electron gun and have energies of 0.1 keV–30 keV. The electron beam creates backscattered electrons, secondary electrons, Auger electrons, X-rays and photons in visible, UV, or IR ranges. Thus, interpretation of the CL results is sometimes more complicated but informative. The CL experiments are usually carried out in either a scanning electron microscope (SEM) or a scanning transmission electron microscope (TEM). Matching CL results with structural characterization using SEM or TEM is a considerable advantage of CL technique. CL spectroscopy can also be accomplished in time-resolved mode. Fast deflecting of the electron beam away from the spot under study using deflection plates or an optically driven electron gun [12] are used to accomplish the time resolution. The latter technique ensures a better time resolution up to subpicosecond domain.

3. ORIGIN OF INHOMOGENEOUS SPATIAL DISTRIBUTION OF LIGHT EMISSION IN III-NITRIDE EPILAYERS AND HETEROSTRUCTURES

III-nitride epilayers exhibit inhomogeneous spatial distribution of luminescence properties. There are several origins for this inhomogeneity. First of all, most of III-nitride epilayers are grown on foreign substrates. Sapphire currently dominates as substrate material but SiC, Si, and other monocrystals are also used, though at a considerably lower extend [13]. Bulk GaN [14,15] and AlN [16] substrates are being developed but their size is still small and their price is still high. Recently, hydride vapor phase epitaxy (HVPE) is proved to be prospective for growing thick, bulk-like GaN [17,18] and AlN [19] layers that can be removed from the sapphire substrate by laser lift-off and used as free-standing substrates for homoepitaxy of these materials.

Most of III-nitride layers are grown by metal-organic chemical vapor deposition (MOCVD) using heteroepitaxy on sapphire. Low-temperature-grown buffer layers are usually deposited to accommodate the lattice mismatch and to facilitate growing of the subsequent layers of better structural quality. The growth of the buffer layer starts by formation of nucleation islands that gradually coalesce. Temperature and V/III precursor ratio serve as the main parameters influencing formation of the buffer layer. For example, a lower V/III ratio enhances 3D growth mode and surface mobility, though increases the roughness of the layer [20]. The larger are the islands, the smaller are the areas of coalescence, where formation of structural defects, especially dislocations, has a higher probability. The dislocation density in the coalescence regions of the buffer layer is higher, and most of these dislocations extend to subsequent layers deposited on the buffer. Very high dislocation density is one of the major problems in growing III-nitride epilayers and heterostructures. Various more sophisticated buffer layers, including epitaxial lateral overgrowth (ELOG) [21] and ELOG variations using patterned substrates, have been developed. However, spatial inhomogeneity in the structure of the buffer layer might often be the origin of luminescence inhomogeneity in the layers deposited above the buffer.

Another origin of the PL spatial inhomogeneity in ternary or quaternary III-nitrides is composition fluctuations in access of local statistical fluctuations expected in a random alloy. Since the variations in composition results in local variations in the band gap, they can cause carrier (exciton) localization in small-scale potential minima and their accumulation in extended regions, where the average potential is lower. Large indium content in InGaN might also lead to phase separation of the compound. Thermodynamic calculations showed that there is a miscibility gap in the InGaN system and that for a typical growth temperature of 600 °C the alloy decomposes into In-rich and In-poor regions at In content exceeding just 6% [22]. Strain shifts the
decomposition limit to considerably higher indium contents and the phase separation also depends on the growth conditions. Nevertheless, difficulties in avoiding the phase separation in high indium content InGaN strongly inhibits fabrication of InGaN-based LEDs emitting in green, yellow and red regions. No phase separation was theoretically predicted for AlGaN compound [23]. However, this compound also exhibits carrier localization effects.

The inhomogeneous distribution of indium atoms should cause an inhomogeneous strain. III-nitride semiconductors are piezoelectric. Thus, strain occurring at a heterointerface causes an electric field. Electric field also occurs due to different polarity of the adjacent materials. The built-in field due to both piezoelectric and spontaneous polarization in quantum wells results in the quantum-confined Stark effect: the field transforms the quantum well from rectangular to triangular shape, the emission band shifts to long-wavelength side, and the probability of radiative recombination decreases, since the overlap of electron and hole wavefunctions decreases due to spatial separation of the recombining electron and hole in the triangular quantum well. The built-in electric field can be calculated using material parameters [24]. However, interpretation of certain experimental results in AlGaN heterostructures requires smaller field values than those calculated using material parameters available to date [25, 26]. Anyway, spatial inhomogeneity in strain might result in inhomogeneity in PL properties. At high density of nonequilibrium carriers, the built-in field might be screened. The screening effect can be taken into account theoretically by simultaneously solving the stationary Schrödinger equation and the Poisson equation.

The localized regions of strain contrast [27, 28] and lattice fringe displacement [29, 30] observed in high resolution transmission electron microscopy (HR-TEM) have been attributed to the inhomogeneous composition in InGaN compound. However, the further TEM study showed that the strain inhomogeneities might be introduced in the material due to electron beam damage [31].

The effective band gap of quantum wells might have substantial spatial variations due to well width fluctuations. For narrow QWs with widths of the order of several nanometers (typical for LED and laser diode structures), even fluctuations of the well width by one monolayer result in variation in the effective band gap by tens of meV. Thus, the areas with broader QWs serve for carrier localization or accumulation.

Inhomogeneous strain at the interfaces might also cause potential fluctuations leading to inhomogeneous PL distribution.

According to theoretical calculations, carriers might be localized even in homogeneous InGaN due to localization of the hole wave functions around In along randomly formed In-N-In chains [32, 33]. There are no corresponding calculations in hexagonal InGaN, which is the usual modification of InGaN, but the recent results on positron annihilation are in consistence with the model of localization on such In-N-In chains in hexagonal InGaN [34].

4. SPATIAL DISTRIBUTION OF LIGHT EMISSION IN InGaN EPILAYERS AND HETEROSTRUCTURES

Though InGaN is currently the material of choice for active media in blue and white LEDs and blue laser diodes, peculiarities of carrier (exciton) localization in this compound are still not fully understood. It is generally accepted that nonequilibrium carriers in ternary InGaN compounds are localized in potential fluctuations occurring due to inhomogeneous distribution of indium in epilayers or/and due to variation in the well width in multiple quantum well (MQW) structures. The spatial scale of these fluctuations is still a matter of discussions and might range in different samples from micrometers [35] down to interatomic distances. In-plane localization on the length scale of approximately 2 nm was observed in InGaN/GaN QWs by analyzing the intensity ratio between the zero-phonon low-temperature PL bands and their phonon replicas to estimate the Huang-Rhys factor. The dependence of the Huang-Rhys factor on electron-hole separation was utilized. The separation depends on the piezoelectric field due to strain at the QW interfaces, while the piezoelectric effect depends on indium content in the GaN/InGaN/GaN heterostructure [36]. Sharp lines emerging in low-temperature PL spectra of InGaN mesas, when the mesa size is diminished down to submicrometer range, are also considered as an evidence of emission by localized excitons [37]. On the other hand, no evidence on deviations of indium distribution from that expected in a random alloy was found using three-dimensional atom probe analysis [38].

Probable, different InGaN epilayers and QWs, especially those with different indium content, have the potential profile fluctuations with considerable variations in their depth and spatial extent.

The carrier localization usually causes an abnormal temperature dependence of PL band peak position and width. In semiconductors without localization, the temperature dependence of the PL band peak follows the continuous band gap shrinkage phenomenologically described by the Varshni formula [39] or a Bose-Einstein-like formula [40]. In semiconductors with considerable carrier localization, the temperature dependence of the PL band peak position has an S-shape. The PL band initially redshifts due to the thermally enhanced ability of localized excitons to reach the lowest available energy sites via hopping [41–43]. The further increase in temperature facilitates thermalization of the exciton system, causes a high-energy shift of the exciton distribution and, consequently, a blue shift of the PL band [41]. Finally, the temperature-induced band gap shrinkage starts to dominate and shifts the emission band to longer wavelengths again. Simultaneously, the band width experiences a W-shaped temperature dependence.

Based on comparison of the experimentally observed temperature dependences of PL band characteristics with results obtained using Monte Carlo simulation of exciton hopping, a double-scale potential fluctuation model has been suggested: excitons in InGaN accumulate in areas with higher indium content and experience random
potential fluctuations at a smaller spatial scale when moving within these areas [44].

Bright spots of the order of hundreds of nanometers in diameter on the background with lower emission intensity have been observed more than a decade ago in CL images of InGaN single quantum wells (see Fig. 3) and were interpreted by existance of quantum discs due to compositional undulations in the quantum well [45].

Fig. 3. Monochromated CL images of In$_{0.2}$Ga$_{0.8}$N single quantum well at wavelengths of 400 nm (a) and 420 nm (b). Reproduced from [45] with permission

As discussed above, early results on transmission electron microscopy (TEM) have been interpreted as a direct evidence of the inhomogeneous In distribution in InGaN epilayers. However, the further TEM study demonstrated that the In-rich “clusters” observed might be created by the electron beam used in TEM characterization [31, 46]. Nevertheless, cathodoluminescence [45, 47–49], μ-PL [50–52], confocal spectroscopy [53–55], and scanning near field optical microscopy (SNOM) [56–59] reveal inhomogeneous photoluminescence intensity distribution in InGaN epilayers and QWs.

Well width fluctuations are pointed out as the main origin of exciton localization in InGaN QWs in many publications [60, 36, 61]. Correlation between well width variations and emission intensity has been demonstrated by comparison of atomic force microscopy and CL images [62].

Usually, the PL bands in the bright areas of InGaN epilayers and MQWs are redshifted in respect of their positions in the dark areas. This is in consistency with the model assuming that the carriers tend to accumulate in the areas with a lower potential. Thus, the photons emitted in these regions have, in average, lower energies than those emitted in the dark areas. Nevertheless, an opposite correlation between PL intensity and band position is also observed in certain InGaN structures [63, 59]. The correlations of three types were observed in violet, blue, and green-emitting SQWs (see Fig. 4). Comparison of the results obtained by using SNOM in I-C and I modes led to conclusion that the different correlations are caused by indium-content-dependent carrier localization conditions that influence capability of the carriers to reach nonradiative recombination centers [59].

Fig. 4. Relationship between peak intensity and wavelength of PL probed by SNOM in the I–C mode in In$_{x}$Ga$_{1-x}$N SQWs emitting in violet (a), blue (b), and green (c) regions. Reproduced from [59] with permission

The red shift of PL band in dark areas in respect to its position in the bright areas observed in InGaN-based LED structures were shown to occur as a result of predominant nonradiative recombination of less localized carriers (excitons), which are responsible for short-wavelength side of PL band (see Fig. 5). This effect is more pronounced in areas with a higher defect density. Thus, it was shown that the inhomogeneous defect density might be responsible for inhomogeneous PL distribution [63].

Fig. 5. PL intensity mapping (2 μm × 2 μm) image of InGaN-based LED structure (a) and spatially-integrated PL spectrum in bright areas (solid curve) and spectral distribution of ratio of PL intensity spatially integrated in bright and dark areas (dashed curve) (b). After [63]

Inhomogeneous carrier distribution in the well plane might also be influenced by threading dislocations. Hexagonal V-shaped pits formed around the threading dislocations under appropriate conditions exhibit narrow sidewall quantum wells with an effective band gap significantly larger than that of the regular quantum well [64]. SNOM spectroscopy in combination with TEM study has been successfully used to reveal “antilocalizing” potential barriers around such dislocations and to clearly elucidate an important mechanism preventing nonradiative carrier recombination at dislocations in InGaN MQWs [64] (see Fig. 6). The quantum wells on the side walls of the V-shaped pit in the vicinity of a threading dislocation crossing the InGaN MQW structure are narrower than those in the areas free of dislocations. Accordingly, high and low energy bands appear in the PL spectra recorded by using SNOM (see Fig. 6, c). Formation of the region with a higher potential due to the narrower quantum well prevents the carriers from reaching the dislocation and recombining nonradiatively there.

347
5. SPATIAL DISTRIBUTION OF LIGHT EMISSION IN AlGaN STRUCTURES

The S-shaped temperature dependence of PL band position that, as discussed above, is an indication of carrier localization was observed in AlGaN epilayers together with other evidences showing that alloy clustering is quite possible in this compound [65–68]. A double-scaled potential profile has been elucidated by comparing experimental results with Monte Carlo simulation of exciton hopping and recombination [68]. It was also shown that the density of nonradiative recombination centers is a key issue in achieving a better LED performance, since a considerable difference in carrier lifetimes was observed in Al_{0.26}Ga_{0.74}N epilayers with very similar potential profiles.

Spatially nonuniform intensity distribution of redshifted emission was observed in CL images at room-temperature in AlGaN samples grown by plasma-assisted molecular-beam epitaxy with 20%–50% of aluminum [69]. The emission inhomogeneities were attributed to compositional inhomogeneity. The results obtained using time-resolved CL spectroscopy of Al_{x}Ga_{1-x}N epilayers with 0.05 < x < 0.25 were consistent with the potential fluctuation model. At low temperatures, when the excitons are strongly localized, the exciton lifetime increases monotonically with aluminum content, while at elevated temperatures the excitons are delocalized, the decay becomes faster and predominantly nonradiative, regardless of the aluminum content.

There are only a couple of papers reporting on SNOM study of AlGaN [70, 71]. Using SNOM in I-C mode, a dual localization pattern in AlGaN epilayers with the AlN molar fraction varying from 0.30 to 0.50 was elucidated [71]. The potential fluctuations on 100 nm spatial scale were evaluated from the width of the photoluminescence spectra. The fluctuation depth varied between 0 meV and 51 meV and increased with increased Al content. These potential variations have been assigned to small-scale compositional fluctuations due to stress variations, dislocations, and existance of Al-rich grains. The larger area potential variations of 25 meV–40 meV in depth were especially pronounced in the low Al-content epilayers and have been attributed to Ga-rich regions close to grain boundaries or atomic layer steps. Maps of localization parameter \( \sigma_L \) and sketches of potential profiles in AlGaN epilayers containing different molar fraction of aluminum are presented in Fig. 7.

6. CONCLUSIONS

Spatial inhomogeneity of radiative and nonradiative recombination of nonequilibrium carriers in III-nitride epilayers and heterostructures reflect complicated composition fluctuations, inhomogeneous point defect distribution, influence of dislocations, strain variations. Better understanding of these properties and processes is desirable for purposeful development of light emitting devices, sensors and transistors based on this prospective semiconductor family. Techniques for luminescence study with spatial resolution down to subwavelength range are powerful tools to study the spatially inhomogeneous emission properties. Combination of these techniques with conventional luminescence spectroscopy and structural analysis provided a considerable contribution in study of III-nitrides and have good prospects for further applications.
Acknowledgments

The article was prepared under support of the European Social Fund Agency implementing measure VP1-3.1-SMM-05-K of the Human Resources Development Operational Programme of Lithuania 2007-2013 3rd priority “Strengthening of capacities of researchers and scientists” (project No. VP1-3.1-SMM-05-K-01-003)

REFERENCES


