III–nitrides: Growth, characterization, and properties

Cite as: Journal of Applied Physics **87**, 965 (2000); https://doi.org/10.1063/1.371971
Submitted: 07 April 1999 . Accepted: 05 October 1999 . Published Online: 18 January 2000

S. C. Jain, M. Willander, J. Narayan, and R. Van Overstraeten

APPLIED PHYSICS REVIEWS

III–nitrides: Growth, characterization, and properties

S. C. Jain
IMEC, 3001 Leuven, Belgium

M. Willander
Chalmers University of Technology and University of Gothenburg, Department of Physics, S-41296 Gothenburg, Sweden

J. Narayan
Department of Material Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695.7916

R. Van Overstraeten
IMEC, 3001 Leuven, Belgium

(Received 7 April 1999; accepted for publication 5 October 1999)

During the last few years the developments in the field of III–nitrides have been spectacular. High quality epitaxial layers can now be grown by MOVPE. Recently good quality epilayers have also been grown by MBE. Considerable work has been done on dislocations, strain, and critical thickness of GaN grown on different substrates. Splitting of valence band by crystal field and by spin-orbit interaction has been calculated and measured. The measured values agree with the calculated values. Effects of strain on the splitting of the valence band and on the optical properties have been studied in detail. Values of band offsets at the heterointerface between several pairs of different nitrides have been determined. Extensive work has been done on the optical and electrical properties. Near band-edge spectra have been measured over a wide range of temperatures. Free and bound exciton peaks have been resolved. Valence band structure has been determined using the PL spectra and compared with the theoretically calculated spectra. Strain and its effect on the optical properties of the III–nitride layers have been studied both theoretically and experimentally. Both *n* and *p* conductivity have been achieved. InGaN quantum wells with GaN and AlGaN barriers and cladding layers have been investigated. PL of the quantum wells is affected by confinement effects, band filling, quantum confined Stark effect, and strain. This work has led to the fabrication of advanced optoelectronic and electronic devices. The light-emitting decodes emitting in the blue and green regions of the spectrum have been commercialized. The work leading to these developments is reviewed in this article. The device processing methods and actual devices are not discussed. [S0021-8979(00)01803-X]

TABLE OF CONTENTS

I. INTRODUCTION

A. Introductory remarks

During the last 5 yr evolution and rise of III–nitride semiconductors has been phenomenal. Several conferences proceedings,[1–3] special issues,[4] authored book,[5,6] and edited books[7–9] have been published on these semiconductors and on the devices fabricated using them. Reviews on the III–nitrides have been written by Akasaki and co-workers[10–12] by Nakamura and co-workers[13–17] and more recently by Pearton *et al.*[18,19] The work on field effect transistors (FETs) has been reviewed by Binari,[20] Binari and Dietrich,[21] and by Khan and co-workers.[22–25] Morkoç and co-workers have written comprehensive reviews covering several aspects of the nitride semiconductors.[26–30] In the mid 1990s performance of the nitride devices improved very rapidly. Light emitting diodes (LEDs) in the blue and the green regions of the spectrum were commercialized. More than 6000 h of working life of short wavelength lasers was demonstrated. The estimated life is more than 10 000 h. Recently the FETs have shown excellent microwave high power and high temperature performance. The review by Pearton *et al.*,[18] which appeared recently in this journal, focused on device processing and the role of impurities and devices with extensive tables of the materials properties. In this review we concentrate on the complementary topics of growth, characterization, and mechanical, optical, and electrical properties of III–nitrides.

B. Important properties of III–nitrides

There are many areas where conventional III–V semiconductors cannot be used. Short wavelength light emitters are required for full color display, laser printers, high density information storage, and under water communication. High temperature and high power transistors are needed for automobile engines, future advanced power distribution systems, all electric vehicles, and avionics. Si and conventional III–V semiconductors are not suitable for designing and fabricating optoelectronic devices in the violet and blue region of the spectrum. There band gaps are not sufficiently large. GaAs based electronic devices can not be used at high temperatures. III–nitrides are particularly suitable for applications in these areas. The band gaps of the III–nitrides are large and direct. The band gap values are 1.9 eV for InN, 3.4 eV for GaN, and 6.2 eV for AlN. Because of their wide band gaps and strong bond strength, they can be used for violet, blue, and green light emitting devices and for high temperature transistors. InN and AlN can be alloyed with GaN. This allows tuning of the band gaps and emission wavelengths. Using AlGaN or GaN as the barrier and cladding layers and GaN or InGaN as active layers, quantum wells and superlattices can be fabricated. This is important because the heterostructure technology used in the fabrication of superlattices is essential for fabricating modern devices. InGaN quantum wells are indispensable for light emitting devices because

incorporation of small In in the active GaN layer increases the luminescence efficiency considerably. ZnSe based semiconductors also have band gaps suitable for short-wavelength optoelectronic devices but the bond strength in the II–VI wide band gap semiconductors is low. Landwehr *et al.*[31] have compared the bond energy of several different semiconductors. The values of the bond energy are 2.3 eV/bond in GaN and 1.2 eV/bond in ZnSe. Doping of the III–nitrides with both donors and acceptors has been accomplished. The low field mobilities of electrons and holes in GaN are reasonable although they are not as high as in GaAs.[32,33] The highest reported room temperature electron mobility in GaN epilayer is 900 $cm^2/V\,s$.[21] Monte Carlo simulations predict higher mobilities of electrons.[34–36] A higher mobility of 1700 $cm^2/V\,s$ has been reported in AlGaN/GaN heterostructures.[22] The Monte Carlo simulated peak velocity in GaN is close to 3×10^7 cm/s and the saturation velocity is 1.5×10^7 cm/s.[36] Both values are considerably higher than the values for Si and GaAs. The nitrides have good thermal conductivity.[10] As discussed by Binari and Dietrich[21] and Duboz and Khan[25] III–nitride devices can work not only at much higher temperatures but also in hostile environments. The fact that the nitrides can stand higher temperatures makes device processing easier. However it should be remembered that at higher temperatures vapor pressure of N_2 in contact with GaN is high. At $T>900\,^\circ C$ GaN is prone to preferential loss of N_2. When high temperature processing is necessary, capping layers or high pressure N_2 ambient have to be used to avoid the loss of N_2 (see Refs. 37–39). GaN has high breakdown fields. The breakdown field for Si is 2×10^5 V/cm, for GaAs $4\times10^5\ V/cm$, and for GaN it is 3×10^6 V/cm. High breakdown voltage is necessary for high power devices.[40,41]

The discussion of this section shows why III–nitride semiconductors have emerged as important materials for optoelectronic as well as for high temperature and high power electronic devices.

C. Evolution of GaN materials and devices

Attempts to synthesize GaN were made by Juza and Hahn[42] in 1938. Small needles of GaN could be synthesized by passing ammonia over hot gallium. Grimmeiss *et al.*[43] used the same method and obtained small crystals of GaN in 1959. They measured the PL spectra of the small GaN crystals. Maruska and Tietjen[44] deposited the first large area GaN layers on sapphire in 1969. They used the chemical vapor deposition method. In all the early samples the concentration of electrons due to background doping was very large. Several centers were proposed that could give rise to background *n*-type conductivity. It was shown that a center consisting of a nitrogen vacancy surrounded by four gallium atoms can act as an efficient donor.[40] Residual impurities like oxygen and Si were also considered as possible donors (see Ref. 40, and references given therein). Many groups published papers supporting one model or the other. There is no consensus even today as to which centers give rise to large electron concentration in the as fabricated samples.

The availability of large area samples[44] gave impetus to GaN research. Pankove *et al.*[40,45] were successful in fabricating the first blue GaN LED. It was a metal/insulating GaN:Zn/*n*-GaN (*m-i-n* or MIS) diode. The *n*-GaN layers contained different concentrations of Zn. The wavelength of the emitted light depended on the Zn concentration. The wavelength could be blueshifted by decreasing Zn concentration. Blue and green color could be easily obtained. A Mg-doped electroluminescent diodes emitting in the violet region was also fabricated.[46] Other important research of that period includes[47–59] study of negative electron affinity,[47] surface acoustic wave generation,[48] and solar blind UV detectors.[40] Stimulated emission from small GaN crystals was first observed at 2 K in 1971 by Dingle *et al.*[60] in 1971. The difficulties in *p* doping and making good ohmic contacts persisted. Until the late 1970s the quality of GaN was not very good. The results of optical and electrical measurements were not reproducible.[41] This hampered the progress of research on the III–nitrides. A more detailed account of the early work is given in the reviews of Pankove[40] and Monemar.[41]

The beginning of the growth of good quality epilayers was made by Yoshida *et al.*[49] They were the first to use the two-step method of growth. They showed in 1983 that if an AlN buffer layer is grown between the GaN film and the sapphire substrate, the quality of the layers improves. The two-step method was investigated and perfected by Akasaki, Amano and co-workers in 1988/1989.[50,51] In the first step a thin buffer layer of AlN is grown on the sapphire substrate at a low temperature of $\sim500\,^\circ C$. The GaN layer is grown on the buffer layer. The temperature at which the GaN layer is grown is much higher.

The two step method is described in detail later in this article. Amano *et al.*[50] were also the first to obtain *p*-type conductivity in GaN. As grown Mg-doped GaN layers have a very high resistance. Amano *et al.*[50] were interested in studying cathodoluminescence (CL) of Mg-doped GaN. They observed that electron irradiation has two effects on the behavior of GaN:Mg. First, it increases the efficiency of the luminescence by up to 2 orders of magnitude. Second, the sample becomes conducting and *p* type. Only the top part of the sample in which the electrons penetrate became conducting. This showed that the low energy electron beam irradiation (LEEBI) was responsible for the *p*-type conduction. Akasaki *et al.*[52] also showed that Mg is an acceptor with an energy of activation equal to 160 meV. In 1992 van Vechten *et al.*[53] suggested that Mg:H complexes are formed in the as grown Mg-doped GaN and therefore hydrogen passivates the Mg acceptors. It is known that hydrogen passivates acceptors in Si. Hydrogen is dissociated from Mg and diffuses out during the electron beam irradiation.[40] Later it was found that annealing the Mg-doped GaN layers at $\geqslant750\,^\circ C$ in nitrogen or vacuum also activates the Mg acceptors.[54] First $In_xGa_{1-x}N$ films over sapphire over the whole range $0<x<1$ of composition were grown by Matsuoka *et al.*[55] GaN films were grown at the usual high temperature of $1000\,^\circ C$. For $x>0$, InGaN films were grown at 500–600 °C depending on the In concentration. The V/III ratio used was high, 1.6×10^4 to 2×10^6.

As mentioned earlier the heterostructure technology is most important for fabricating LEDs, laser diodes (LDs), and FETs. The heterostructures need an AlGaN barrier or cladding layers. First AlGaN layered structures were grown by Khan *et al.*[56] in 1990 and by Itoh *et al.* in 1991.[57] First InGaN/GaN layered structures were grown by Nakamura *et al.*[58] in 1993. Electron microscopy studies showed that the concentration of defects in the best GaN epilayers was very large, $\sim 10^{10}\,\mathrm{cm}^{-2}$.[59] Photoconductivity experiments showed a large absorption tail, an abrupt Urbach edge was not observed. This showed that a large density of states in the band gap was produced by the defects.

By mid 1990s so much knowledge and experience in technology had accumulated that the progress in designing and fabricating devices became very rapid. Akasaki and co-workers[10–12,51,61] had already made the first $p-n$ junction GaN LED in 1989. Room temperature stimulated emission and lasing action were demonstrated at about the same time. Nakamura and his colleagues at Nichia Laboratories brought the blue and green LEDs to level so that they could be commercialized[5,18] in the mid 1990s. Nakamura and co-workers published several reviews on the GaN-based LEDs and LDs.[13–17] Khan *et al.*[62] reported the surface mode emission from GaN. Kim *et al.*[63] measured the optical gain. A large number of reports on the fabrication of the optically pumped lasers with improved performance followed.[64–66] Song and Shan[67] have reviewed the subject and have given an extensive list of references. The evolution of thresholds of stimulated and lasing emission has been plotted by these authors in chronological order from 1994 to March 1997.

The injection lasers were first demonstrated in 1995 by Akasaki and collaborators.[12] Nakamura and co-workers reported the laser diode in 1996, which was followed by rapid improvements in the performance of the lasers. Akasaki *et al.*[68] fabricated a laser diode emitting 376 nm also in 1996. This is the shortest wavelength injection laser that has ever been made. A room temperature cw laser emitting at ~4000 nm with more than 6000 h lifetime was demonstrated by Nakamura in 1998.[17]

The GaN-based FETs were also fabricated about the same time, i.e., in 1993.[69] Fabrication of a heterostructure bipolar transistor using *p*-type 6H SiC as the base and *n*-type GaN as the emitter was reported in 1994.[70] Rapid improvements made in the performance of the FETs are discussed in the reviews in Refs. 21, 23, and 25. The FETs have been fabricated on both sapphire and SiC substrates. SiC has higher thermal conductivity and FETs grown on SiC substrates seem to show better performance.

D. Scope and organization of this review

We start with the discussion of growth and characterization of III–nitrides in Sec. II. Interfaces in GaN-based quantum wells and superlattices are also discussed in this section. Because of a large mismatch in the lattice constants and thermal expansion coefficients of the III–nitride epilayers and the substrates, the epilayers are strained. Theoretical and experimental studies of strain and critical thickness are discussed in Sec. III. The large mismatch in the lattice constants and thermal coefficients produces a large number of defects in the III–nitride epilayers. The defects and their effect on lifetime of the minority carriers are discussed in Sec. IV. Band structure of both zincblende (ZB) and wurtzite (WZ) III–nitrides is described in Sec. V. Effects of strain on the band structure and determination of band offsets in GaN-based heterostructures are included in this section. Photoluminescence is important in determining the quality of the epilayers. Epilayers of good optical quality show strong near band-edge PL spectra. The position of excitonic lines depends on the valence band structure. Excitonic lines can be used to determine the splitting of the valence band. The defect related peaks are either absent or become weak. If the layers are not of good quality, PL is dominated by peaks due to deep levels. PL and properties of excitons are described in Sec. VI. Sections VII and VIII are devoted to electrical properties. Doping methods, carrier concentrations, and mobilities are discussed. Results of Monte Carlo simulations of mobility are also described. Electrical and optical properties of alloys of the III–nitrides are given in Sec. IX. Optical properties of quantum wells and superlattices are discussed in Sec. X. A summary of important results is given in Sec. XI.

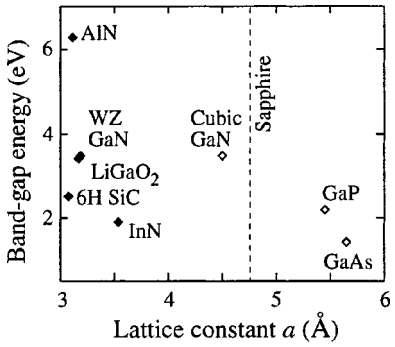

FIG. 1. Band gaps and lattice constants of GaN and related materials.

We have not discussed devices in this article. They have been reviewed by Shur and Khan,[24] by Duboz and Khan,[25] and more recently by Pearton.[18] Interesting articles on the detectors have recently appeared.[71,72]

II. GROWTH

A. Early work

Several attempts were made to synthesize GaN crystals during the period 1930–1960 but good quality crystals of reasonable size could not be grown. Hydrogen passivation and nitrogen vacancies or other unknown defects made it difficult to obtain sufficient *p* doping. Growth of GaN epitaxial layers using hydride VPE (HVPE) was reported in the late 1960s. AlN was grown by the sublimation method. Many attempts to grow thick layers using these methods were made during the 1970s. Although thick bulk-like material could be produced, it was not easy to make electrical measurements, mainly because ohmic contacts could not be made.[30,41]

The lattice constants and band gaps of III–nitrides are shown in Fig. 1. The lattice constants of sapphire and 6H SiC

the most commonly used substrates, are also shown. There is a lattice mismatch as well as thermal mismatch between the III–nitrides and the substrates on which they are grown. Because of this mismatch good epilayers could not be obtained until the late 1980s. The interest in the growth of III–nitride epilayers was revived by the work of Amano, Hiramatsu, and co-workers in late 1980s and early 1990s. These authors showed that MOVPE growth of GaN epilayers on sapphire substrates is facilitated by first depositing a thin AlN buffer layer at low temperature.[10] During the last decade numerous papers have been published on the growth of epitaxial layers of III–nitrides and alloys using both the MOVPE and MBE methods.[1,4,10,73–82] In most cases the epilayers have been grown on sapphire and SiC substrates.

In the early work, VPE was also used to grow the GaN epilayers.[30] Recently Lee *et al.*[83] have grown GaN by VPE. Epitaxial layers of AlN have been grown by pulsed laser deposition on Si substrates.[84] However MOVPE and MBE are by far the most extensively used techniques to grow the III–nitride epilayers.

B. Substrates

1. Sapphire

Sapphire is the most extensively used substrate for growth of the III–nitrides.[85,28] Large area good quality crystals of sapphire are easily available at low cost. They are transparent, stable at high temperature, and the technology of growth of the nitrides on sapphire is quite mature. Crystal orientations of sapphire and GaN (grown on *c*-plane [0001] sapphire) are parallel, but the unit cell of GaN is rotated by 30° about the *c* axis with respect to the sapphire unit cell.[78,86] The $[1\bar{1}00]$ axis of GaN is parallel to the $[1\bar{2}10]$ axis of sapphire. The lattice mismatch between WZ III–nitrides and 6H SiC and ZnO is within 4% (see Fig. 1). The lattice mismatch of GaN with sapphire is ~15%. The quality of GaN grown on *c* plane, *a* plane $[11\bar{2}0]$, and *r* plane $[1\bar{1}02]$ have been studied.[28] In MBE growth, the grain size and inhomogeneous stress of the films grown on the *c* and the *a* plane are similar. However, in the epilayers grown on the *a* plane about 1% of the film was in cubic phase.[28] In MOVPE growth different results were obtained. Growth on the *a* plane sapphire gave better quality films than those grown on the *c* plane. The quality of the films grown on the *r* plane was poor. In general the quality of the films grown directly (i.e., without the buffer layer) on any plane of sapphire is poor. The FWHM of the x-ray rocking curves is large, and strong deep level yellow luminescence is observed. Surface morphology is also poor, but many hillocks are observed. The poor quality is the result of large mismatch in the lattice constants and coefficients of thermal expansion of GaN and sapphire. As discussed in Sec. II D 1, good quality films can be grown by first growing a low temperature buffer layer. Although the importance of the buffer layer was first demonstrated in MOVPE growth, it is also important in MBE growth.[28] The properties of the GaN epilayer are strongly influenced by the thickness and growth temperature of the buffer layer.

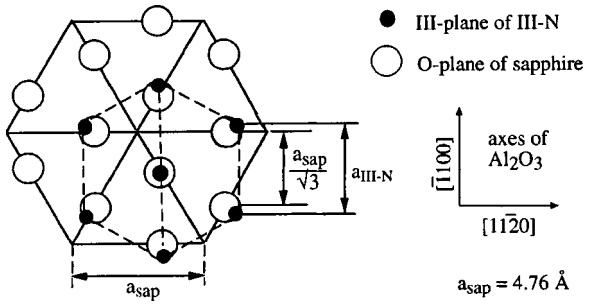

FIG. 2. Schematic representation of in-plane atomic arrangement in the case of (0001) AlN film grown on (0001) sapphire (see Ref. 89).

Experimental critical thickness (defined in Sec. III) of GaN on the *c*-plane sapphire with the AlN buffer is about 3 nm. For films thinner than the critical thickness, the lattice constant of the film has a value equal to that of AlN, i.e., the films are pseudomorphic with the buffer. As the thickness increases, the lattice constant increases due relaxation of strain by misfit dislocations. At a thickness of 100 nm, the relaxation is nearly complete and the lattice constant becomes the same as that of bulk GaN.[28]

Narayan and co-workers have studied the growth and interface structure of the III–nitrides on sapphire and also on ZnO buffer layers grown on sapphire.[87–90] The growth of III–nitrides on sapphire occurs by domain matching where integral multiples of lattice constants or major lattice planes match across the film/substrate interface.[87,88] It is also found that epitaxial growth across the III–nitride/sapphire interface is dictated by Al–O bonding, which results in a 30° in-plane rotation of nitride lattice with respect to the sapphire substrate.[89] The atomic arrangement on the (0001) basal plane involving III–nitride lattice and sapphire is shown schematically in Fig. 2. It delineates the epitaxial relationship between the film (*f*) and substrate (*s*) as follows: $(0001)_f \| (0001)_s$, $[01\bar{1}0]_f \| [\bar{1}2\bar{1}0]_s$, and $[\bar{1}2\bar{1}0]_f \| [\bar{1}100]_s$. The 30° rotation results in a mismatch of about 12% between the substrate $(a_s/\sqrt{3})$ and the film (a_f). The large lattice mismatch is accommodated by domain matching of lattice constants $(7a_s/\sqrt{3} \approx 6a_f)$, where seven $(\bar{1}2\bar{1}0)$ planes of sapphire match six $(01\bar{1}0)$ planes of the film[89,90] as demonstrated for ZnO film on sapphire shown in Fig. 3. It should be mentioned that growth of III–nitrides on ZnO substrates occurs via lattice matching without rotation; however, interfacial reactions at higher temperatures of growth result in the formation of spirals such as $ZnAl_2O_4$ across the AlN/ZnO interface.[90]

In the ideal case the calculated FWHM of the GaN (002) peak is 33.2 arcsec. However due to the large number of dislocations, the measured FWHM of the films grown on sapphire is larger. It is 2–5 arcmin for good quality films and varies up to over 35 arcmin for poor quality films (Ref. 28, Fig. 2.14). The stress in the films depends both on the orientation of the substrate and on the buffer layer.

2. 6H SiC substrate

Extensive work has been done on the growth of III–nitrides on SiC.[85,28] SiC has several advantages over sap-

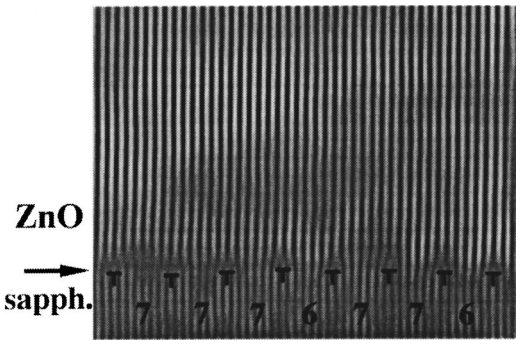

FIG. 3. Fourier-filtered TEM image of domain matching in the growth of ZnO films on sapphire. Misfit dislocations on the interface are indicated. Numbers at the bottom of the picture correspond to the number of planes between the misfit dislocations. Every sixth or seventh $(2\bar{1}\bar{1}2)$ plane of sapphire terminates at the interface (see Ref. 90).

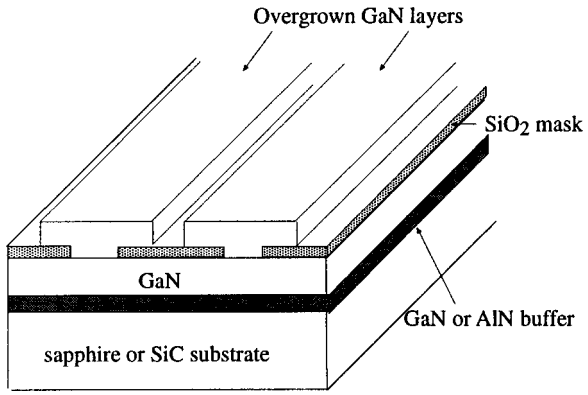

FIG. 4. Design of an ELOG substrate consisting of patterned SiO_2 mask and window stripes. GaN grows in the windows and then laterally on the SiO_2 stripes.

phire. Its lattice mismatch with GaN is only 3.5%, and with AlN, the mismatch is very small (see Fig. 1). SiC has good electrical conductivity. A substrate bias can be applied to reduce the ion energy while using nitrogen ions for the growth of the nitrides.[91] Large good quality SiC substrates are available commercially.[28] For the surface preparation, the SiC wafer is dipped in HF and then treated with a hydrogen plasma. This reduces the oxygen–carbon bonds to very small numbers. Low temperature GaN or AlN buffer layers are also grown on the SiC substrates before depositing the nitrides. The stress in the films grown on SiC is smaller than in the films grown on sapphire.

3. Si substrate

To accomplish integration of III–nitride LEDs or LDs with Si electronics, these devices will have to be grown locally on an already processed Si chip. Therefore considerable work has been done on the growth of III–nitrides on the Si substrate (see Refs. 23, 84 and 92–94 and references given therein). Reviews of this work have been written by Abernathy[85] and Popovici *et al.*[28] A low temperature buffer layer of 30 nm GaN is generally grown on Si before the growth of the main epilayer. GaN grown on (100) Si is predominantly cubic. The epilayer contain a large number of defects such as dislocations, twins, and stacking faults. The layers grown on (111) Si are predominantly WZ with localized inclusions of cubic phase. Better quality WZ GaN has been obtained on (111) Si by ECR MBE on the Si surface containing specially prepared atomically flat terraces.[28] These layers exhibited free exciton PL at low temperature.

4. ELOG substrate

Several authors have grown and studied laterally epitaxially overgrown (LEO) GaN epilayers on patterned substrates, designated as ELOG substrates.[95–100] To fabricate an ELOG substrate, growth of a 0.2 μm GaN layer is performed on a (0001) *c* face of a sapphire wafer. A 0.1 μm thick SiO_2 layer mask is patterned to form 4 μm wide stripe windows separated by 7–8 μm wide SiO_2 stripes (see Fig. 4). The dimensions quoted here are approximate, and different dimensions have been used by different groups. A thick layer of GaN is grown on the mask. The epilayer first grows on sapphire through the windows. As the thickness of the epilayer increases, it grows laterally on the SiO_2 stripes. At about 10 μm thickness the GaN epilayers which grow at the two edges of the SiO_2 stripes coalesce and a continuous flat GaN layer is formed on the patterned substrate. The dislocation density in the GaN surface layer grown on the SiO_2 stripes is small, it is only $2\times10^7\ \mathrm{cm}^{-2}$. The lifetime of the InGaN/GaN/AlGaN-based laser diodes grown on this substrate increased from a few hours to more than 1150 h.[95]

HVPE[98,101] has also been used to grow LEO layers. TEM and SEM have been used to determine the defect structure[98,96,95] and Raman, PL, and cathodoluminescence (CL) techniques have been used to determine the optical properties[97] of the LEO layers grown by different methods. The dislocation concentration in the layers grown in the window area is $\sim2\times10^8\ \mathrm{cm}^{-2}$. In the layers grown laterally on the SiO_2 layer the concentration is drastically reduced. Raman measurements showed that there was a general improvement in the quality of the laterally overgrown layers. Low temperature PL measurements showed features attributed to donor–acceptor recombination. These features were present only in the laterally overgrown layers. They were attributed to Si impurity and an unknown acceptor. There was an additional feature probably due to oxygen impurity.

In a LEO GaN layer, alternate stripes have high and low dislocation density. Since the device is fabricated on the GaN stripes with low density of defects, the size of the device is limited by the width of these stripes. Also, the device fabrication requires careful alignment of the device structure with the underlying mask stripe. Therefore it is desirable to have a contiguous large area layer with a low density of defects. Nam *et al.*[102] were able to produce such a layer by repeating the LEO process a second time. In the second LEO step, the mask stripes are placed on the openings of the first mask, i.e., they cover the GaN stripes with high density of defects. The GaN layers in the second LEO step are therefore seeded on the good GaN stripes. They grow laterally on the second oxide masks. Thus a large area GaN layer with low defect density is obtained. More recently[99] large area GaN films with low density of defects have been fabricated by the so-called pendeoepitaxy (PE). The PE growth process is il-

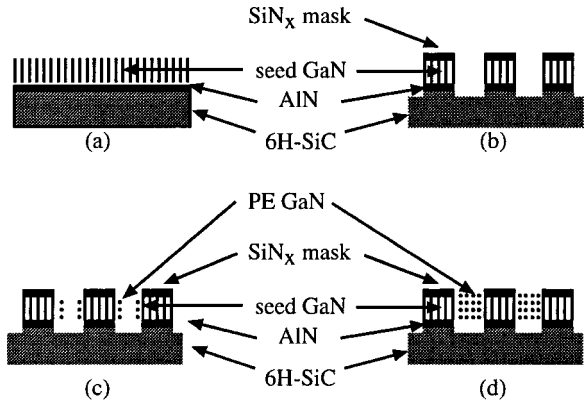

FIG. 5. Schematic diagram of the PE process steps: (a) GaN seed layer, (b) etched GaN columns prior to PE growth, (c) partial growth of PE GaN on the side walls, and (d) coalesced growth of PE GaN.

lustrated in Fig. 5. GaN does not nucleate on the etched surface of SiC. It nucleates only on the side walls of the columns. GaN grows laterally away from the opposite side walls on which it nucleates. Growing surfaces approach each other and eventually meet. A (0001) top surface is created. Now GaN begins to grow vertically on the (0001) surface. When it reaches the top level of the silicon nitride masks, it grows laterally over the mask. Since no nucleation takes place at the substrate or on GaN with large defects, the density of dislocations remains low. A large area of the low defect density GaN is produced.[99]

5. Other substrates

Several other materials have been used as substrates. Some of them are GaAs,[30,103] $MgAl_2O_4$,[28,82] glass,[104] quartz glass,[105] and ZnO.[106] Properties of several substrate materials have been discussed in Refs. 28 and 85. Very high quality GaN films have been grown on ZnO substrates by MBE.[28] The width of the exciton PL line was only 8 meV, comparable to the best results obtained by MBE on sapphire substrates. Deep level yellow PL was absent. However only a few groups have grown the nitrides on the ZnO substrate. The quality of the GaN layer grown on several other substrates has been discussed by Popovici *et al.*[28]

GaN has also been grown homoepitaxially but only on small (mm size) bulk GaN crystals.[86] The bulk GaN crystals were obtained by growth from a nitrogen saturated Ga melt at 1600 °C under a nitrogen pressure of 15 kbar. Large size GaN wafers do not exist at the present time.

C. Pulsed laser deposition (PLD) of III–nitride films

High quality III–nitride films have been successfully grown by pulsed laser ablation using solid nitride targets.[84,107] It is also possible to incorporate dopants into the targets and obtain films with dopant concentrations exceeding thermodynamic solubility limits. Since the average energy of the laser ablated species is considerably higher ($100-200kT$) than the thermal energy kT, it is possible to lower the growth temperature of the films. For example the AlN growth temperature can be reduced from 1000–1100 °C to 700–800 °C for pulsed laser deposition (PLD)[84,89,107] Another advantage of PLD is the convenience with which buffer layers such as ZnO and other III–nitrides can be incorporated during growth in the same chamber without breaking the vacuum.[90] The PLD target holder has 4–6 target slots, therefore, multilayer heterostructure for fabricating LEDs and LDs can be grown with atomic-level control.

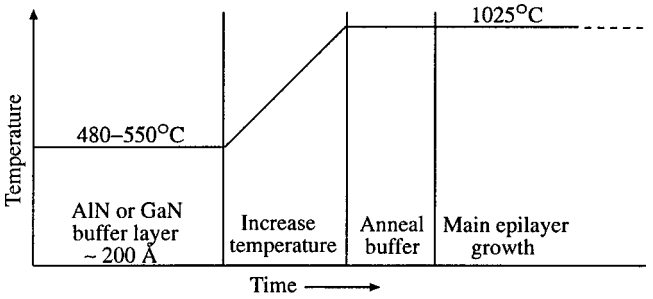

FIG. 6. III–nitride growth by MOVPE (schematic). A nucleation layer (also called buffer layer) of AlN or GaN is grown at a low temperature. The temperature is then raised, the buffer layer is annealed at the growth temperature, and the main epilayer of GaN or its alloy with AlN or InN is grown on the buffer layer.

D. MOVPE

1. The source materials and buffer layers

We first discuss the general features of MOVPE growth of III–nitrides. The source materials generally used are trimethylgallium (TMGa) for Ga, trimethylaluminum (TMAl) for Al, and trimethylindium (TMIn) for In. Ammonia is used as source material for nitrogen.[80] Mg is widely used as a *p*-type dopant and Si as an *n*-type dopant. Biscyclopentadienyl (Cp_2Mg) is used as a source of Mg (Ref. 82) and methyl silane ($MeSiH_3$) as a source of Si. Amano, Hiramatsu, Akasaki and their collaborators showed that good epilayers of III–nitrides on sapphire substrates can be grown by MOVPE if the growth is performed in two steps (see the reviews by Akasaki and Amano[10] and by Hersee *et al.*[78]). In the first step a thin GaN or AlN buffer layer is grown at a low temperature, as shown schematically in Fig. 6. The main epilayer is grown in the second step at a higher temperature. The buffer layer provides the high density of nucleation centers and promotes the lateral growth of the main epilayer. The microstructure of the low temperature buffer layer grown on sapphire has been examined. Atomic force microscopy (AFM) studies show that the structure of the buffer layer is columnar due to islandic growth.[78] Because of the large difference in the lattice constants of the buffer layer and the substrate, the free energy of the system is lowered if the growth is three dimensional (3D). The islands formed in the 3D growth are coherent with the substrate. TEM and AFM studies show that time of ramping and annealing the buffer layer has a large influence on the structure of the buffer layer[10,78] and on the quality of the main epilyaer grown subsequently. The quality of the main epilayer improves by annealing the buffer layer for up to about 20 min. If the buffer layer is annealed for longer times, the quality of the epilayer starts degrading.

The quality of the epilayers grown by the two-step method has been examined by PL, TEM, and x-ray methods.

The quality is indeed very good (see references cited above). The two-step procedure has now become a standard method for the III–nitride growth by MOVPE. The method is also used for the growth of the conventional III–V semiconductor layers directly on Si substrates.[108] In the original work of Akasaki and co-workers discussed in Ref. 10, AlN was used for buffer layers. Nakamura[109] has used GaN layers (instead of AlN layers) as buffers. Subsequent work of Nakamura and co-workers discussed in Secs. VII and VIII suggests that the quality of the GaN epilayers improves if the GaN buffers are used. Nakamura also introduced the so-called two-flow method. In this method, in addition to the main gas flow parallel to the substrate, another flow, designated as the ''subflow'' is used. The subflow transports a mixture of the inactive gases H_2 and N_2 perpendicular to the substrate. The subflow changes the direction of the main flow to bring the reactant gas in contact with the substrate. Nakamura found that subflow suppresses the 3D growth, promotes the 2D growth, and improves significantly the quality of GaN epilayers.

Recently Kachi *et al.*[110] have used InN buffer layers to grow GaN epilayers on (1120) oriented (*A*-face) sapphire substrates using MOVPE. The thickness of the buffer layers was 20–30 nm and they were deposited at a low temperature of ~600 °C by atmospheric MOVPE. GaN epilayers were grown on these buffer layers as well as on the conventional GaN buffer layer. The growth temperature of the main GaN epilayer was 1050 °C in both cases. The dislocation density in the GaN epilayers was measured by cross-sectional TEM. The density was $\sim 6\times 10^8\ \mathrm{cm}^{-2}$ in the epilayers grown on InN buffers. The density in the epilayers grown on GaN buffers was seven times larger. Hall effect measurements were also made. The maximum electron mobility was close to 500 $\mathrm{cm}^2/\mathrm{V\,s}$ in the GaN layer grown on a 30 nm InN buffer. The maximum mobility with the GaN buffer was lower, ~330 $\mathrm{cm}^2/\mathrm{V\,s}$. The melting points of AlN, GaN, and InN are 3000, 1700, and 1100 °C, respectively. The improvements in the quality of the epilayers are probably related to the melting point of the buffer layers. At the substrate temperature of 1050 °C at which the GaN epilayers are grown, the buffer layers should soften and yield to the strain. The softening and plastic yield of the buffer reduce the strain as well as the dislocation density in the GaN epilayers. The softening will be negligible in the case of AlN buffers because the melting point is very high. The softening will be maximum for InN buffers because the growth temperature is close to the melting point of InN.

2. IR oscillations

RHEED cannot be used in MOVPE reactor because RHEED can be used only in high vacuum. A reflected laser beam from the surface of a growing epilayer shows intensity oscillations due to interference effects. Nakamura[79] and more recently Mesrine *et al.* and others (Ref. 111 and references given therein) have used the oscillations to measure the thickness of the film. The layers were grown by MOVPE in the first case and by gas source (GSMBE) in the second case. If the thickness is not uniform the oscillations are not observed. The fluctuations in thickness become larger as the film grows. Therefore the amplitude of the oscillations decreases, and finally the oscillations disappear as the films become thicker. The intensity of the reflected beam is used to determine the film thickness.[79,111]

E. Slow development of MBE

1. Difficulties in MBE growth of III–nitrides

The MOVPE technique for the growth of III–nitrides had become mature by the mid 1990s. Until recently most GaN-based devices were fabricated using MOVPE. The development of MBE as a growth technique for the nitrides has been slow. This has happened because growth temperatures in MBE are low and ammonia, a common source of nitrogen, is stable at low temperatures. Recently, many articles have appeared on the MBE growth of III–nitrides. The main motivation of this work is: (1) to optimize the growth rate and growth temperature, (2) to determine the effect of different sources of nitrogen on the quality and growth rate of the epilayer, and (3) to determine the substrate effects.

Because of the low temperatures, gas-phase reactions do not take place in the MBE. This means that pyrolysis of ammonia containing compounds does not occur efficiently. Until recently the growth rate of GaN by MBE has been $<0.4\ \mu$m/h. Recently growth rates of 1–2 μm/h at 750–800 °C have been achieved.[85,112] These growth rates compare well with those obtained with MOVPE. Very high quality III–nitride epilayers have been grown by GSMBE. The quality and morphology have improved to a level that RHEED oscillations are routinely observed from the MBE epilayers.[112,113] However lower temperatures are required for the growth of buffer layer and also for InGaN alloys. The fraction of ammonia that decomposes at these temperatures is small and therefore growth rates are also small.

2. Nitrogen sources for MBE growth

In spite of the high thermal stability of ammonia, it is a commonly used nitrogen source in the MBE growth of the nitrides. Mesrine *et al.*[111] have examined the efficiency of nitrogen incorporation in the epilayers grown with a NH_3 source in GSMBE. The reflectivity of a laser beam is low from a GaN surface but is high if the surface contains a layer of Ga. First islands of GaN were grown in the 3D mode on a sapphire substrate. The islandic growth suppresses the oscillations in the intensity of the reflected laser beam. An amount of Ga equivalent to 3 monolayers was deposited and ammonia was then delivered to the surface. As the GaN grows, the reflectivity decreases and attains a low level corresponding to the GaN surface. Below 700 °C decomposition of GaN is negligible. Therefore the rate at which reflectivity decreases can be used to determine the rate of formation of GaN, efficiency of cracking, and incorporation of ammonia[111] in the GaN layers. The experiments were performed at different temperatures. It was found that the measured incorporation of ammonia and growth of GaN begins only at 450 °C. The incorporation efficiency is 0.5% at 500 °C, and rises rapidly and becomes 3.8% at 700 °C. This method does not work at higher temperatures because decomposition of GaN and evaporation of Ga become appre-

ciable beyond 700 °C. By measuring the growth rate, the incorporation efficiency was determined at higher temperatures. Above 700 °C, the measured incorporation efficiency increases slowly to 4% at 830 °C. The actual efficiency of cracking at this temperature is probably higher since a part of the nitrogen must have been lost by evaporation and did not contribute to the formation of the epilayer. The work of Mesrine *et al.*[111] is discussed again in Sec. II F. Grandjean and Massies (of the same group as Mesrine) have grown high quality epilayers of GaN and AlGaN,[112] *n*- and *p*-doped GaN,[114,115] and InGaN (Ref. 113) using ammonia as the nitrogen source.

(DMHy)$[(CH_3)_2NNH_2]$ decomposes at lower temperatures (below 600 °C) but introduces carbon impurity in the epilayer. Yoshida[116] has used (DMHy) for the growth of a buffer layer at low temperature and ammonia for the growth of the main epilayer at higher temperature (see Sec. II F).

The difficulty of nitrogen incorporation using ammonia at low temperatures can be overcome by using a reactive species of N_2 obtained by electron cyclotron resonance (ECR) or radio frequency (rf) plasmas. (The methods to create ECR and plasmas are described in detail in Refs. 28, 85, 117, and 118.) rf plasma at a frequency of 13.56 MHz is most commonly used.[85] The plasma is generated by inductively coupling the rf energy into a discharge chamber filled with nitrogen gas to a pressure of $>10^{-6}$. The chamber is water cooled. There are N_2, N_2^+, and atomic nitrogen species in the beams thus generated. The energy of these species is large. High energetic beams are detrimental to the quality of the epilayer. The alternative source which produces reactive nitrogen beams at lower energies is the ECR microwave plasma. In this source the microwave energy at 2.45 GHz is coupled with the resonance frequency of electrons in a static magnetic field. The ECR beam has a higher fraction of reactive nitrogen than the rf plasma beams. The quality of the epilayer seems to degrade if the microwave power is high. The quality of the epilayers can be improved by reducing the ion energy by inserting an aperture in the opening of the ECR source or by keeping the microwave power at low values. Reduction of ion energy also reduces the background impurity concentration in the epilayer. Early use of ECR to grow good quality GaN epilayers on the (0001) sapphire substrate (by the two-step method) was made by Moustakas and Molnar.[117] The GaN buffer layer was grown at 500 °C and the main GaN layer at 800 °C. With low power plasma, the surface was flat and the layers were *n* type. With high power plasma, the layers were insulating and the surface was rough. The *c* plane of the epilayers was parallel to the substrate. Growth rates up to 0.65 μm/h were achieved but detailed studies were made on films grown at a rate in the range 0.20–0.25 μm/h. Si-doped *n*-type and Mg-doped *p*-type layers were also grown. The electrical properties of these films are discussed in Secs. VII and VIII.

Kaufman ion sources which provide flux of N_2^+ ions at lower energies have also been used.[119,91,85] The energy of the N_2^+ ions obtained from these sources is less than 30 eV. The energy can be further reduced by applying a retarding dc bias if the substrate is conducting, e.g., SiC. Decreasing the energy improves the band-edge luminescence and suppresses deep level luminescence.[91] However structural quality as measured by x-ray diffraction does not improve. The structural quality improves on increasing the growth temperature from 600 to 700 °C.

3. Group III sources in MBE

The flux of group III elements is generated by heating the high purity elements in the Knudsen or effusion cells. The vapor pressure of the flux is generally $\sim 10^{-4}$ Torr. The variation and control of the flux pressure is achieved by changing the temperature of the effusion cells. Group V incorporation efficiency depends on the surface coverage by the group III elements. The growth temperature and group V flux are so adjusted that the incorporation of both elements is in a 1:1 ratio. Under these conditions excess group V elements evaporate. However if the temperature is not appropriate and evaporation of excess group V element does not occur, the mobility of group III elements is affected and growth tends to become 3D.

The growth temperatures of III–V nitrides should be low to avoid the loss of group III elements by evaporation from the surface. The sticking coefficient of In is near unity below 540 °C. The sticking coefficients of Ga and Al are practically unity below 710 °C. Above these temperatures the evaporation of group III elements becomes important. The lower limit of the growth temperature is determined by the mobility of the group III atoms on the surface. If the temperature is too low and the mobility is small, the growth becomes 3D and islands are formed.

F. MBE growth of GaN and AlN

1. Growth with ammonia and DMHy

High quality GaN epilayers were already grown by MOVPE in the late 1980s and early 1990s and the MOVPE technique is now quite mature. Recently considerable efforts have been made to grow GaN and AlN epilayers using MBE. Good quality epilayers can now be grown by MBE also. We have already discussed work of Mesrine *et al.*[111] on the growth of GaN epilayers by MBE using ammonia as the nitrogen source. In this work the epilayers were grown at 830 °C with a Ga beam equivalent pressure (BEP) of 2×10^{-6} Torr. Two regimes of growth were observed. In regime 1, the growth rate was 0.5 μm/h at the ammonia flux of $\sim$3 sccm, increased rapidly with the ammonia flux, and reached 1.2 μm/h at an ammonia flux of 12 sccm. In the second regime, for values of ammonia flux greater than 12 sccm, the growth rate became constant, independent of ammonia flux. These two observed regimes of growth can be explained by assuming that in the higher ammonia flux regime, the growth rate is determined by the Ga flux. At the lower flux of ammonia, the growth rate is determined by the available N flux resulting from the cracking of ammonia at the surface. Several other groups[26–28,112–114] have grown GaN, AlGaN, and InGaN epilayers on sapphire substrates by MBE using ammonia as the nitrogen source. In the method used by Morkoç and co-workers[28] ammonia decomposes by pyrolysis when it comes in contact with the surface of the substrate. The method has been designated in Ref. 26 as the

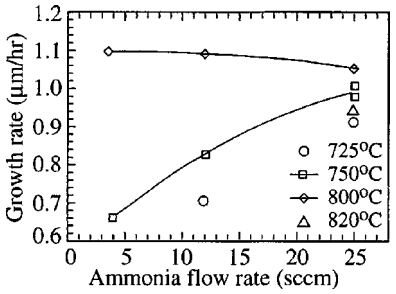

FIG. 7. Variation of RMBE growth rate with ammonia flow rate for four different substrate temperatures (Ref. 26).

reactive MBE or RMBE. The growth temperature is kept in the range 700–820 °C. Conventional effusion cells sources of group III elements are used. After cleaning the substrate, it is nitridated at around 800 °C for about 10 min. This produces a thin AlN layer on the sapphire surface. An AlN or GaN buffer layer is then grown at 500 °C. A high ammonia to group III element flux ratio is used. At a high ratio, the stress generated in the film is low, and deep level luminescence is weak.[28] Typical growth rates of GaN at different ammonia flow rates and at four substrate temperatures are shown in Fig. 7. The highest growth rate reported by the same group is 2.9 μm/h.[28] The high quality of the films was confirmed by strong intrinsic luminescence. High performance devices were fabricated using these epilayers which provide further evidence that the layers were of high quality.

Grandjean and Massies[112] have grown GaN, AlN, and AlGaN alloys on the *c* plane of a sapphire substrate by GSMBE. The method is essentially similar to that used by Morkoç and co-workers discussed above. Ammonia was introduced into the growth chamber via a 50 sccm mass flow controller. Sources of Ga and Al were conventional evaporation cells. The surface of the sapphire was nitridated for 10 min at 850 °C under 20 sccm of ammonia. This resulted in the formation of a thin AlN layer, followed by the deposition of a GaN buffer layer at 500 °C. The growth of the main epilayers was carried out at 830 °C. The V/III ratio was kept in the range 3–5. The growth rate was ~1.2 μm/h. At 500 °C numerous RHEED oscillations were observed from the GaN epilayers. The growth rate at this temperature was low, 1.66 ML/s. At 800 °C only two weak oscillations were observed. The room temperature PL showed a strong band-edge peak at 3.415 eV with FWHM of 32 meV. The intensity of the yellow luminescence due to deep levels was two orders of magnitude smaller than the intensity of the band-edge luminescence band. 8 K emission showed a strong free *A* exciton luminescence peak at 3.478 eV and another peak at 6 meV below the *A* peak which was attributed to the *B* exciton bound to a neutral donor. (The yellow band and the *A, B,* and *C* excitons are discussed in Sec. VI.) The unintentional *n*-type doping was $<4\times10^{16}\,\mathrm{cm}^{-3}$. The Hall mobility at 300 K was 100–150 $\mathrm{cm}^2/\mathrm{V\,s}$. The dislocation density determined from the TEM experiments was $10^9\,\mathrm{cm}^{-2}$. The FWHM of the x-ray rocking curve was 350 arcsec. The RHEED pattern remained streaky like and the amplitude of the laser reflected intensity oscillations remained constants for growth of several microns.

The growth rate of AlN was studied as a function of temperature of the Al evaporation cell.[112] The growth rate versus $1/T$ plot was a straight line with an activation energy of 3.1 eV. This value of the activation energy agrees with the value of activation energy for Al evaporation known from other independent experiments.[112] To grow AlGaN layers, first a few monolayers of GaN were grown at ~500 °C on the *c* plane of sapphire as described earlier. When Al flux was added, RHEED oscillations were still observed but with somewhat reduced intensity. This shows that the surface morphology of the AlGaN epilayers was good but probably not as good as that of the GaN epilayers. The incorporation rate of Al depends on the growth rate. By observing the growth rate before and after the Al flux was switched on, the composition of the AlGaN layer was determined. The compositions of the epilayers was 11% Al in one epilayer and 33% Al in the other. Grandjean and Massies[114] have also grown *n*- and *p*-doped GaN by the same method. The properties of doped epilayers are discussed in Sec. VIII.

Powell *et al.*[115] have succeeded in growing the GaN epilayers using GSMBE directly on sapphire (0001) substrate. The layers were grown at different temperatures in the range 600–850 °C. The substrates were degreased by successive rinses in trichloroethane, acetone, methanol, and isopropyl alcohol followed by a hot (160 °C) H_2SO_4: H_3PO_4 etch. The wafers were then rinsed in de-ionized water and blown dry in dry N_2. The Ga flux was maintained at $7\times10^{14}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and ammonia flux was $1\times10^{17}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The starting base pressure was 10^{-10} Torr but during growth it was 3×10^{-4} Torr. The growth rate was 0.4–0.5 μm/h. Below 770 °C the growth was 3D but it changed to 2D at higher temperatures. The epilayers grown at 780 °C exhibited band-edge PL at room temperature. The FWHM of the $\theta-2\theta$ scan was 1.4 min and that of the x-ray diffraction curve it was 16 min. TEM studies showed that primary defects were threading dislocations (with $a_0/3\langle 2\bar{1}\bar{1}0\rangle$ Burgers vectors) and stacking faults. In the earlier work electron mobilities in GSMBE grown epilayers were small, less than 20 $\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$. Powell *et al.*[115] achieved a value 110 $\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$ of the electron mobility in their samples. The RHEED, XRD, and XTEM observations suggested that the basal plane of GaN epilayers is rotated 30° with respect to that of the sapphire substrate.

To obtain reactive species of nitrogen at low temperatures, Yoshida[116] used both dimethylhydrazine (DMHy) $[(CH_3)_2NNH_2)]$ and ammonia as nitrogen sources. A 50 nm buffer layer was grown on a sapphire substrate at a temperature in the range 710–730 °C. Rate of decomposition of DMHy is considerably larger than that of ammonia in this temperature range. Yoshida[116] has mentioned that he could not grow the buffer layer at this temperature with ammonia as a source of nitrogen. This result is different from that obtained by Grandjean and Massies[112–114] and Morkoç[28] who have grown the buffer layers at temperatures lower than 700 °C. The difference in the results is probably due to the fact that Yoshida did not nitridate the substrate. DMHy decomposes at lower temperatures (below 600 °C) but introduces carbon impurity in the epilayer. Therefore DMHy was used for the growth of the buffer layer. For the growth of the

main GaN epilayer, ammonia was used as the source of nitrogen. For Ga and Si Knudsen effusion cells with solid sources were used. The BEP of Ga was about 6×10^{-7} Torr and that of ammonia was 5×10^{-6} Torr. For Si doping, a Si BEP of 5×10^{-9} Torr was used. GaN epilayers with mirror-like surface could be grown in the temperature range 800–850 °C. Above 850 °C the surface of the epilayer became rough. Below 760 °C Ga droplets were seen on the surface because of low decomposition rate of ammonia. Room temperature PL spectra of GaN showed the band-edge emission at 365 nm and a broad yellow band at 575 nm due to deep levels. (The yellow PL band is discussed in Sec. VI D.) By increasing the V/III flux ratio by a factor 2, the yellow band disappeared. Si-doped layers did not show the broad yellow band.

2. Growth using rf plasma, ECR microwave plasma, and other ion sources

Rizzi *et al.*[120] have grown GaN and AlN on 6H SiC substrates. The lattice mismatch between AlN and SiC is small (in-plane lattice constants are 3.08 Å for 6H–SiC and 3.11 Å for 2H–AlN. The thermal coefficients of the two materials are not so different.[121] A buffer layer of AlN is generally used for the growth of GaN on SiC(0001). Rizzi *et al.*[120] used *n*-type (electron concentration $\sim\times10^{18}\,\mathrm{cm}^{-3}$) 6H–SiC as substrates for the growth of AlN and GaN epilayers by GSMBE. rf plasma source was used for the activated nitrogen supply. After annealing at 750 °C for 15 min in ultrahigh vacuum LEED and Auger spectroscopy, it showed that the surface was (1×1) Si terminated. The AlN buffer layer was grown at 620 °C and the GaN/AlN heterostructure was grown at 760 °C.[120] After each deposition of AlN the surface showed a clear (1×1) LEED pattern. This shows that the layers had the same orientation as the substrate. PL studies showed that the epilayers were of good optical quality.

Korakakis *et al.*[122] have grown Al_xGa_{1-x} on both (0001) sapphire and (0001) Si terminated 6H–SiC using an ASTEX compact ECR microwave plasma source to activate the molecular nitrogen gas. For growth on the sapphire substrate the usual procedure of nitridating the substrate and growing a low temperature GaN or AlGaN buffer layer was followed. The growth temperature of the buffer layer was 550–600 °C and the thickness of the buffer layers was 30 nm. Al_xGa_{1-x} layers were grown at 700–775 °C. In the case of 6H–SiC, the growth was done directly on the substrate without any buffer layer. RHEED and XRD studies showed that the layers were of good quality. The composition of the layers was determined using XRD and energy dispersive spectroscopy. Films with up to 60% Al concentration were grown. Four probe resistivity measurements were made. Hall effect measurements were also made on films doped moderately and heavily. The properties of the alloy layers arc discussed in Sec. IX.

3. Growth using supersonic jet molecular beams

Single phase WZ AlN[94] and GaN and InN (Ref. 123) have been grown on Si (100) substrate using supersonic jet molecular beam gas sources. In this method a seeded jet accelerates precursors to kinetic energies which are one to two orders of magnitude higher than the average kinetic energies in the molecular beams of materials used as sources.[123] The high energies enhance precursor adsorption and increase the growth rates. Good quality III–nitride layers have been grown by this method. The low base pressure and high growth rates make this technique somewhat attractive. However, due to focusing of the flux, the deposition nonuniformity is quite large. Further work is necessary before this technique can be employed to fabricate satisfactory device structures.[123]

4. Other methods of growth

Atomic layer epitaxy is discussed in Refs. 124 and 125. Meng *et al.*[126] have grown AlN films on Si (111) and Si(001) substrates by ultrahigh vacuum reactive sputter deposition. The films were characterized by Raman scattering, ion channeling, and TEM. The films grown on (111) at >600 °C were epitaxial with WZ structure. The growth was islandic, and the islands coalesced as the film thickness increased. The films grown on (001) Si were polycrystalline with highly oriented grains.

G. Growth of InGaN

Recently many articles have appeared on the growth of InGaN alloys using both MOVPE and MBE. A large fraction of these papers are on quantum wells and superlattices (see Sec. X). Yoshida[116] grew InGaN epilayers by MBE on sapphire substrates. Uncracked ammonia was used as a source of nitrogen. In concentration in the layers was 20%. Yoshida[116] found that the growth temperature had a large effect on the quality of the layers. If the growth temperature was more than 800 °C the concentration of In was very low. On the other hand if the temperature was lower than 760, Ga droplets were formed on the surface. As mentioned earlier, ammonia does not decompose well at the lower temperature and the V/III ratio does not remain sufficiently high. The layers grown at temperatures in the neighborhood of 780 °C were of good quality and had a mirror-like surface. Under these conditions epilayers with more than 20% In could not be grown. There was a strong band-edge peak at 420 nm, and a buffer layer peak at 365 nm in the PL. Deep level PL was absent.

Grandjean and Massies[113] have grown $In_xGa_{1-x}N$ layers with high In concentrations (up to $x=0.46$) by GSMBE. The nitrogen precursor was 50 sccm ammonia. A 2 μm thick GaN layer was first deposited on (0001) sapphire at 500 °C after nitridating the substrate at 850 °C.[112] $In_xGa_{1-x}N$ layers were grown at temperatures in the range 500–600 °C. Since at lower temperatures the fraction of ammonia that decomposes is smaller, the growth rate was reduced from 1.2 to 0.2–0.4 μm/h. The effect of growth temperature on the incorporation of In was studied. The indium incorporation rate was smaller at higher growth temperatures. This result is similar to that obtained by Yoshida.[116] Increasing the Ga flux had only a small effect on the In incorporation rate. By in-

creasing the In flux, up to 46% In could be incorporated in the epilayers. RHEED oscillations were used to monitor the change from 2D to 3D in the growth mode.

The InGaN alloys with the entire composition range have been grown by MBE by Blant *et al.*[92] and by Singh *et al.*[127] Blant *et al.*[92] have studied the MBE growth of both the AlGaN and InGaN alloys on clean and oxidized Si surfaces using a rf plasma source of nitrogen. The growth temperature was in the range of 400–750 °C depending on the composition of the alloys. X-ray measurements showed that alloys of (InGa)N of controlled composition can be grown on Si substrate using a plasma enhanced MBE over the entire composition range from InN to GaN. Films grown on oxidized surfaces of Si were polycrystalline or amorphous. The growth on chemically cleaned Si substrates shows the usual columnar structure common to the growth of group III–nitrides.[86] XPS studies indicate that there is little tendency for spinodal decomposition. The composition of the alloys deduced from electron probe microanalysis studies agree well with those from x-ray measurements. Singh *et al.*[127] have obtained somewhat different results. They were also able to grow the InGaN alloys over practically the entire composition range at the growth temperature of GaN (700–800 °C) by MBE. These authors found that if the thickness of the films is $>0.3\ \mu$m, incorporation of more than about 30% indium results in spinodal decomposition. However such phase separation is absent in thin InGaN films (<600 Å) grown as GaN/InGaN/GaN heterostructures. In such configurations up to 81% In could be incorporated.

H. Interfaces

Ponce[86] has given a very good account of the atomic structure at the GaN and sapphire (α-Al_2O_3, trigonal structure) interface. The projection of the basal plane of WZ GaN lies in the horizontal direction parallel to the (0001) substrate surface. Lattice images of GaN and AlGaN show that the interfaces are atomically sharp, i.e., the transition from the oxide to the nitride is abrupt and stable. The mismatch between GaN and sapphire is so large that the calculated critical thickness is less than a monolayer. The strain is therefore quickly relieved by the generation of misfit dislocations. Complete relaxation requires misfit dislocations at an average distance of 2.0 nm. However the dislocations are clustered, leaving a large volume of the epilayer free from defects.

Atomically abrupt and planar interfaces between the GaN layer and SiC can have two configurations: C–Si–C–Si–C–Al–N–Al–N–Al and C–Si–C–Si–N–Al–N–Al. These two configurations cannot be distinguished in the HRTEM. The abrupt interface with these configurations is charged. Charge neutrality and lowest energy configuration require that some mixing should occur. Steps in the basal plane of 6H–SiC cause the creation of misfit dislocations and stacking faults.

Meng *et al.*[126] have studied the GaN/Si (111) interface by ion beam channeling. The thickness of the epilayers was 1000 Å. They found that the strong dechannelling peaks usually found when misfit dislocations are present were absent from the GaN/Si interface. The mismatch between the lattice constants is 23% and the critical thickness for the GaN epilayers is less than 1 ML. The absence of the misfit dislocations is therefore a surprising result. Meng *et al.*[126] suggested that in this case the structure of the interface is quite different. The whole interface accommodates the strain uniformly. In TEM studies the localized strain contrast expected from the dislocations is also not observed in this system.[126]

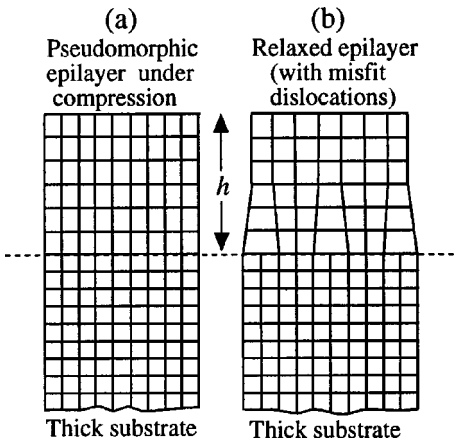

FIG. 8. Structure of an epilayer under biaxial compression: (a) pseudomorphic and (b) relaxed with misfit dislocations.

The interfaces between two III–nitride epilayers are generally smooth and abrupt. The interface mixing and roughening observed in GeSi (Ref. 128) and in III–V multilayers (Ref. 108) are not so much of a problem in the III–nitride multilayers. The $\theta-2\theta$ scans of the InGaN/GaN multilayer structures show that the interfaces of the multilayer are smooth plane and abrupt. Singh *et al.*[129] reported the growth of InGaN/AlGaN MQWs on *c*-plane sapphire by ECR assisted MBE. TEM images showed sharp and abrupt interfaces, thus confirming the high interfacial quality of the MQW structures. Several authors have observed 2D gas of electrons or holes at the modulation doped interface on III–V multilayers (see for example Ref. 130 for the study of 2D electron gas at the GaN/AlGaN interface and several articles in the MRS 97 symposium on III–V nitrides[1]).

In the epilayers grown by the MOVPE method, interface roughness can be controlled to within 5–10 Å. However the residual point defect density remains high at $10^{17}\,\mathrm{cm}^{-2}$. The dislocation density is $\sim 10^{10}\,\mathrm{cm}^{-2}$ in the layers grown on sapphire and $\sim 10^{8}\,\mathrm{cm}^{-2}$ in the layers grown on SiC.[86]

III. STRAIN AND DISLOCATIONS

A. Misfit strain and critical thickness

A pseudomorphic epitaxial layer of a crystal can be grown on a substrate of slightly different lattice constant provided the thickness of the layer is sufficiently small (see, e.g., Ref. 128). The structure of a lattice mismatched layer is shown in Fig. 8. In this figure the lattice constant of the epilayer is bigger than that of the substrate. If the thickness h of the layer is smaller than a certain thickness, known as critical thickness h_c [defined later in Eq. (6)], the misfit between the layer and the substrate is accommodated by a te-

tragonal compression of the layer as shown in Fig. 8(a). The strain in the layer is homogeneous and is known as misfit strain. The in-plane lattice constant of the layer becomes the same as that of the substrate. If the lattice constant of the epilayer is smaller than that of the substrate, the strain is tensile. Such homogeneously strained layers are known as pseudomorphic strained layers. When $h>h_c$, misfit strain begins to relax by the introduction of misfit dislocations as shown in Fig. 8(b). Strain has a profound effect on many properties of the epilayers.[131–167]

Lattice mismatch is measured by the misfit parameter f_m defined as

$$f_m=\frac{a_l-a_{\text{sub}}}{a_{\text{sub}}}, \tag{1}$$

where a_l and a_{sub} are the lattice constants of the epilayer and the substrate. If the epilayer consists of an alloy, such as $In_xGa_{1-x}N$ grown on GaN, its lattice constants depend of the composition x. Assuming that Vegard's law is valid, the lattice constant $a_l(x)$ of the alloy epilayer is given by

$$a_l(x)=a_{\text{GaN}}+(a_{\text{InN}}-a_{\text{GaN}})x. \tag{2}$$

Now the misfit parameter is also a function of the alloy composition x, i.e., $f_m=f_m(x)$. It can be calculated using the values of lattice constants shown in Fig. 1. The in-plane strain components in the x and y directions are given by

$$\epsilon_{xx}=\epsilon_{yy}=-f_m(x). \tag{3}$$

We have formulated the above equations using the convention that the strains ϵ_{xx} and ϵ_{yy} are negative if the layer is under compression and positive if it is under tension.

Expression for h_c has been derived using two different approaches, both known as equilibrium theories of critical thickness. The first approach is based on the principle of energy minimization and was first enunciated by Frank and Van der Merwe in 1949 and the second, known as the force balance theory, was given by Matthews and Blakeslee in the 1960s. Both theories have been discussed extensively in the literature and we have published reviews with extensive lists of references on these theories.[128,131] Due to lack of proper understanding, the two theories have been regarded as different from each other. Many authors have given two different numerical values of h_c for the same layer calculated using the two theories. It has been demonstrated, however, that the two theories are equivalent and give identical numerical values of h_c if same values of dislocation energy are used in the theories (see Ref. 128 and reference given therein). We use here the force balance method to derive the expression for the critical thickness.

A dislocation half loop nucleates at the surface and expands until it touches the interface. Its threading arms then expand and deposit misfit dislocation at the interface. The threading arm of an existing threading dislocation can also move and deposit the misfit dislocation. The force G acting on the threading arm is given by[131]

$$G=2\mu\frac{1+\nu}{1-\nu}h|f_m b_1|-2E_D^{\infty}, \tag{4}$$

where E_D^{∞} is the dislocation tension or its energy per unit length

$$E_D^{\infty}=\frac{b^2\mu}{4\pi(1-\nu)}(1-\nu^2)\ln\frac{\rho_c h}{q}, \tag{5}$$

where b is the Burgers vector, $b_1=b\cos\beta$, q is the core cutoff parameter and is assumed to be equal to b, μ is the shear modulus of elasticity, ν is the Poisson's ratio, and ρ_c is the core energy parameter. β is the angle between the dislocation line and its Burgers vector and has a value 60° for the 60° dislocations. Only 60° dislocation can glide at the interface. If $h<h_c$ the line tension, i.e., the force which opposes the motion of the threading dislocation is larger than the driving force due to misfit strain given by the first term and G is negative. Therefore the misfit dislocations cannot be generated and the layer remains pseudomorphic. As h increases, the driving force increases faster than the opposing force given by the dislocation tension. At $h=h_c$, $G=0$. For $h>h_c$, G is positive and the threading dislocation moves, depositing the misfit dislocation. The thickness h_c is the maximum thickness up to which the layer can remain pseudomorphic. Equating G to 0, we obtain the expression for the critical thickness

$$h_c=\frac{b^2(1-\nu\cos^2\beta)}{8\pi f_m(1+\nu)b_1}\ln\frac{\rho_c h_c}{q}. \tag{6}$$

The theory does not take into account the energy needed for the nucleation and motion of the dislocations. The only material parameter involved in Eq. (6) is the Poisson's ratio ν, which has a value ≈ 0.3 for most semiconductors. Therefore theoretical h_c versus f_m curve is approximately the same for most lattice mismatched epilayers. Numerous measurements of h_c of Ge_xSi_{1-x}, $In_xGa_{1-x}As$ and other strained alloys were made in the mid 1980s.[128,32] In general observed values of h_c are larger than those calculated using Eq. (6). The discrepancy between the theoretical and the experimental values of h_c is believed to be due to the fact that the energy needed for the propagation of dislocations is not available at the growth temperatures. In high quality strained epilayers, the number of existing threading dislocations is very small and nucleation of dislocation loops is also required for the generation of misfit dislocations. If defects, which act as sources for dislocation generation, are absent (i.e., the layers are of high quality), nucleation is homogeneous and needs large energy.[131] This also limits the generation of misfit dislocations.

Critical thickness and strain relaxation in the capped layers and in layers with nonuniform distribution of dislocations have been studied by many authors and have been reviewed in Refs. 128 and 131. The effect of interactions between dislocations on the critical thickness and strain relaxation are also discussed in these references.

For $h>h_c$, the equilibrium theory requires that misfit dislocations be introduced to relieve the misfit strain. Under the ideal equilibrium condition, two perpendicular arrays of misfit dislocations are created. The distribution of dislocations in each array is periodic. In the presence of the dislocations the strain is given by

$$\epsilon_{xx}=\epsilon_{yy}=-\left(f_m(x)-\frac{b_1}{p}\right), \tag{7}$$

where p is the distance between two neighboring dislocations. In practice, the distribution is not periodic and an average value of p is used. The relation between the thickness h and strain ϵ_{xx} is obtained by replacing f_m by the strain ϵ_{xx} in Eq. (6)

$$f_m(x)-\frac{b_1}{p}=\frac{b^2(1-\nu\cos^2\beta)}{8\pi h(1+\nu)b_1}\ln\frac{\rho_c h_c}{q}. \tag{8}$$

We have used the value of ϵ_{xx} given in Eq. (7). The value of p and therefore the number of misfit dislocations in a layer of thickness h can be calculated using Eq. (8). The number of dislocations observed in a thick layer (i.e., a layer with $h>h_c$) is generally smaller than predicted by Eq. (8). The reasons for the discrepancy are the same as those responsible for the discrepancy between the observed and calculated critical thickness and discussed earlier.

In many cases of practical importance the mismatch between the epilayer and the substrate is very large, e.g., GaAs grown on Si and GaN grown on sapphire substrates. In such cases the layers are grown by using the so-called *two-step* method discussed in Sec. II D 1. If the mismatch is $\geqslant$4% and layer thickness is not too small, the strain is almost completely relaxed by generation of misfit dislocations during the growth and/or annealing of the layer at the higher temperatures. The layer can be regarded as ''lattice matched'' to the substrate. On cooling the layers to room temperature, stresses and strains are generated due to the difference in the values of coefficients of thermal expansion α_{th}. The misfit parameter due to thermal strain is given by[108]

$$f_{(m,\text{th})}=(\alpha_{\text{th,layer}}-\alpha_{\text{th,substrate}})(T_g-RT), \tag{9}$$

where T_g is the growth temperature. All results of the previous section can be used for thermal strain if $f_{(m,\text{th})}$ is used instead of $f_m(x)$. Generally the experimental values of thermal strain do not agree with the values calculated using this equation. If the thickness of the layer is not sufficiently large or if the growth temperature is not sufficiently high, relaxation of strain at the growth temperature is not complete and the lattice mismatch strain persists on cooling the layer to lower temperatures. Also additional misfit dislocations may be generated during the cooling process, relieving some of the thermal strain. Both factors will make the observed strain different from the strain calculated using Eq. (9). It is difficult to take these factors into account and calculate the thermal strain more accurately (see however Ref. 108).

B. Experimental values of critical thickness

Critical thickness of GaN on AlN and AlN on GaN has been measured recently.[112] Epilayers of GaN on thick relaxed AlN buffers and of AlN on thick relaxed GaN buffers were grown on (0001) sapphire. The GaN epilayers were grown at 600 °C whereas the AlN layers were grown at 800 °C. During the growth of GaN, only five RHEED oscillations were observed. After 5 ML when the oscillations decay, the RHEED pattern changes from streaky to spotty, indicating the onset of a 3D islandic growth. The change in the lattice parameter due to relaxation of strain begins after the growth of two monolayers. The experimental thickness of GaN on AlN is therefore 2 ML. The behavior of AlN epilayers grown on relaxed GaN is somewhat different. The change in lattice parameter due to strain relaxation was observed after 3 ML, a value similar to that for GaN grown on AlN. However, 15 RHEED oscillations were observed and strain relaxation was rather rapid during this period. Since RHEED observations did not show any signs of the islandic growth, the relaxation of strain probably occurred by the generation of misfit dislocations. The critical thickness in both cases is 2–3 ML. This value is smaller than that determined earlier by other experiments.[112] The value of critical thickness obtained by different workers is different. The critical thickness is 10 Å for AlN/GaN determined in Ref. 132 and 30 Å GaN on AlN determined in Refs. 133 and 28.

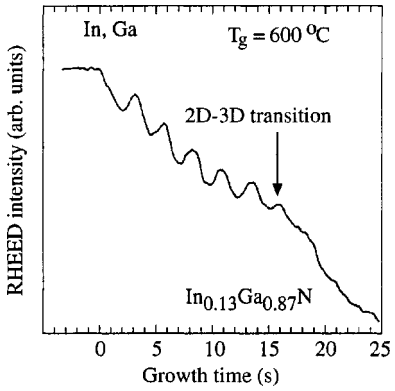

FIG. 9. RHEED specular intensity oscillations recorded during the growth of $In_{0.13}Ga_{0.87}N$ at 600 °C (Ref. 113).

It is well known that the RHEED oscillations disappear as soon as islanding begins and the growth changes from 2D to 3D mode. Grandjean and Massies[113] studied RHEED oscillations in InGaN epilayers for several values of In concentrations. The oscillations observed from an $In_{0.13}Ga_{0.87}N$ epilayer grown on 2 μm thick GaN/sapphire layer[113] are shown in Fig. 9. Grandjean and Massies[113] determined the critical thickness at which the growth of islands of InGaN begins and the oscillations disappear. The critical thickness is 3 ML for $x=0.3$ and 7 ML for $x=0.12$. Below $x=0.12$, islanding was not observed, only surface roughening was seen.

Akasaki and Amano[12] investigated the critical thickness of $In_xGa_{1-x}N$ and $Al_xGa_{1-x}N$ layers. In both cases the layers were grown on relaxed GaN layers. The concentration x of Al and In was in the range $0.05\leqslant x\leqslant 0.2$. The values of critical thickness were in the range 300–700 nm for $Al_xGa_{1-x}N$/GaN and ~400 nm for $In_xGa_{1-x}N$/GaN. There was no significant change of the critical thickness with the composition x of the epilayers.

All experimental values of the critical thickness quoted in this section are considerably larger than the values calculated using the equilibrium theories. This result is similar to that obtained with GeSi, II–VI, and III–V semiconductors.[128,131,134]

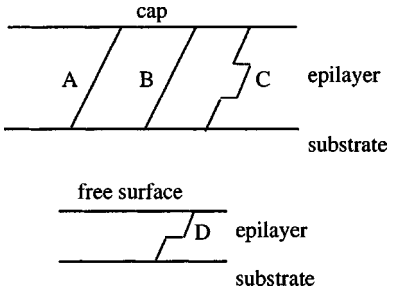

FIG. 10. A double kink C and a single kink D are shown.

C. Propagation of dislocations

1. The kink model

The dislocation velocity v_d can be represented by an equation of the form[128]

$$v_d = v_0(\sigma_{\text{exc}})^m \exp(-E_v/kT), \tag{10}$$

where v_0 is a constant, σ_{exc} is the excess stress, and E_v is the energy of activation for the glide motion of the dislocation. Values of m equal to 1 and 2 in Eq. (10) have been used (Ref. 128, and references given therein). The excess stress can be written as

$$\sigma_{\text{exc}} = \sigma_e - \sigma_t, \tag{11}$$

where σ_e is the misfit stress and σ_t is the self stress associated with the line tension of the dislocation. The values of these stresses are well known.[128] Substituting these values one obtains the expression for the excess stress

$$\sigma_{\text{exc}} = 2S\mu\epsilon\frac{1+\nu}{1-\nu} - \frac{\mu b\cos\alpha(1-\nu\cos^2\beta)}{4\pi h(1-\nu)}\ln\frac{\rho_c h}{b}. \tag{12}$$

Here S is the Schmid factor and has a value $1/\sqrt{6}$. If the strained layer has an ''infinitely thick'' capping layer, the self-energy of the dislocation line (dipole) increases by a factor of ~ 2[131] and therefore σ_t is replaced by $2\sigma_t$ in Eq. (12). σ_{exc} and the velocity of the dislocation are reduced in the capped layers.

The effect of dislocation interactions is not included in the above expression for the excess stress. The actual distribution of the dislocations is not periodic, generally it is irregular. For such a distribution, the second term in Eq. (12) increases and the velocity is reduced. Also, in such cases the stress is not uniform and the velocity becomes a function of the spatial coordinates of the dislocation. Observed velocities of dislocations in different regions of a sample differ by a factor of up to 3.[128]

Dislocations move by the formation of kinks shown in Fig. 10. It is not easy for the whole length of the dislocation line to jump from A to B in one step as it will need energy which the system cannot provide at the temperatures at which experiments are performed. If small portions of dislocation move at a time, by the formation of kinks as shown in Fig. 10, this difficulty is avoided.[135,136,128] A small portion of the line B jumps forward (to the right in Fig. 10) forming a double kink as shown at C. The double kink is formed in a capped or a thick layer. A single kink is formed, as shown at D in Fig. 10, if the layer is uncapped and thin. The two components of the double kink move apart along the dislocation line. In the case of a single kink, the kink moves in a manner such that the length of the dislocation line in the displaced position increases. In this process small bits of the dislocation line move forward until the whole dislocation line has moved by one atomic spacing. If the dislocation line is sufficiently long, several double kinks can be formed at the same time. The process is repeated and the dislocation moves large distances. There is a critical thickness above which the probability of formation of double kinks becomes large. The thickness at which the transition from single kink to double kink nucleation takes place depends on the strain ϵ in the epilayer. It is ~ 1 μm for $\epsilon = 0.2\%$ and ~ 200 Å for $\epsilon = 1\%$.[135,131] For details of the mathematical theory of the formation of kinks and derivation of the expression for the dislocation velocity, Refs. 135 and 128 should be consulted. The expression for the velocity derived from the kink model is of the form given in Eq. (10).

2. Experimental results

Sumino and Yonenaga[137] have reviewed the dislocation velocities in (unstrained) Si, GaAs, and InP. In general the dislocation velocity in II–VI and III–V semiconductors is much larger than in Si. The velocity of α dislocations (in which the extra plane terminates in group III elements) is larger than that of β dislocations (where the extra plane terminates in group V elements). The fact that dislocations do not expand, move, or multiply in GaN epilayers after the epilayers are deposited suggests that the activation energy for propagation of dislocations must be very large in this material. Recent work of Sugiura *et al.*[138,139] shows that this is indeed the case. Sugiura *et al.*[139] found that there is a linear relation between the activation energies for dislocation propagation and the band gaps of the conventional III–V semiconductors. Assuming that the linear relation for undoped α dislocations holds for the III–nitrides also, they calculated the values 2.1, 2.0, and 2.3 eV of E_v for GaN, $In_{0.2}Ga_{0.8}N$, and $Al_{0.2}Ga_{0.8}N$, respectively. Sugiura *et al.*[139] found that the pre-exponential factors in Eq. (10) are not very sensitive to the semiconductor type. A value, 1×10^6 cm/s can be used for all semiconductors including the III–nitrides. At room temperature the mobility of the III–nitrides calculated with these approximations comes out to be smaller by a factor 10^{-10}–10^{-16} than the mobilities in GaAs. It is known that in GaAs the dislocation velocity increases by the nonradiative recombination at the dislocations. The activation energies for the III–nitrides are so large that this effect is likely to be smaller in the nitrides.

D. Stress distribution and dislocations

Stress in the GaN and AlN layers has been measured by the wafer curvature method and by the Raman measurements.[140,126,141,75,142] Yoon *et al.*[141] have studied the shift in the Raman E_2 mode as a function of the thickness of the GaN films grown on (0001) sapphire by GSMBE. The shift in the E_2 Raman mode showed the residual stress changes from tensile to compressive as the epilayer thickness

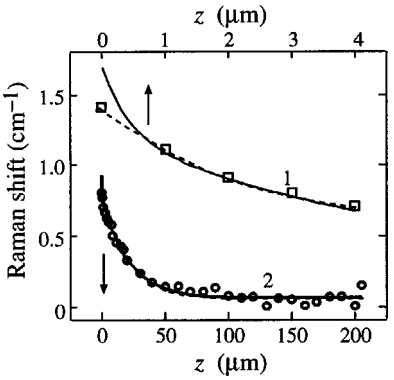

FIG. 11. Square symbols are the measured Raman shift $\Delta E^p_{2\,\mathrm{eff}}(z_0)$ in the 4 μm thick (Ref. 75) and open circles in the 200 μm thick (Ref. 142) GaN epilayers. Solid curves are the corrected values of $\Delta E^p_2(z_0)$.

increases. Hiramatsu *et al.*[143] observed a compressive stress in the thin GaN epilayers and a rapid decrease in the stress of GaN as the thickness of the GaN layer increased to more than 4 μm. However Kozawa *et al.*[140] studied stresses in films of thickness up to 50 μm by the curvature method and observed only a small decrease in the compressive stress as the GaN films became thicker. A complex behavior of stress evolution in AlN films grown on Si substrate was observed by Meng *et al.*[126] As mentioned earlier the layers were grown by the ultrahigh vacuum reactive sputter deposition. The authors also used the wafer curvature method to measure the stress. In the layers grown at room temperature the stress was compressive and small in thin layers. It increased approximately linearly with thickness to values of about 200 N/m for 1500 Å thick films. In the layers grown at 800 °C, the stress was compressive and increased with thickness as before but at about 500 Å thickness, the stress started decreasing. In the layers grown on (111) Si, the stress became practically zero in 1700 Å thick films.

We now discuss the stress distribution with depth z measured by micro-Raman method in Refs. 75 and 142. The shifts $\Delta E^p_{2\mathrm{eff}}(z_0)$ in two GaN epilayers have been observed and is shown in Fig. 11. The laser beam entered through the side of the sample. The geometry used in the experiments and the relative position of the beam are shown in Fig. 12. The stress does not vary in the x and y directions and therefore the observed Raman shift varies only with depth z.[75,142] The effect of the depth of penetration of the beam in the x

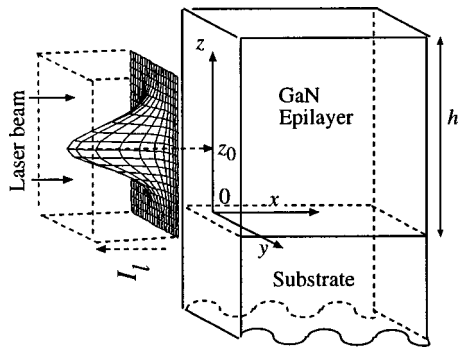

FIG. 12. The sample and the axes of coordinates.

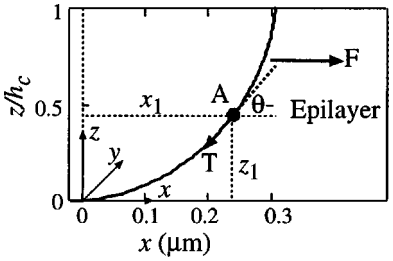

FIG. 13. Shape of the threading dislocations predicted by the equilibrium theory.

direction and that of the beam width in the y direction can be ignored. The intensity of laser light is therefore assumed to vary only in the z direction across the width of the laser beam. In the Raman experiment spectra originate from different points (z) in the volume of the structure sampled by the laser light. Due to the finite beam width, the beam samples not only the point z at which the laser is positioned but the neighboring areas also. Therefore a correction to observed shift is required before it can be used to determine stress at the point z. In the geometry used in Refs. 75 and 142, the E_2 Raman mode is active in GaN. The strain shifts this mode by ΔE_2. The spectrum originating from the point z is given by a Lorenzian line with peak frequency $\Delta E^p_2(z)$ which depends on the stress at z. To account for the effect of the beam width, the spectra from all points z are superposed to obtain the intensity of the spectrum observed in the experiment.[76] Jain *et al.*[146] have applied such a correction. The corrected Raman shifts are also shown in Fig. 11. The stress is maximum at the interface and decreases as distance from the interface increases. Similar stress distribution has been found in ZnTe epilayers containing threading dislocations.[147] Jain *et al.*[146] showed that the stress variation can be attributed to the large number of dislocations. To interpret these observations Jain *et al.*[146] assumed that the concentration of threading dislocations is n cm^{-2} and neglected the interaction between the dislocations. A threading dislocation is shown in Fig. 13. At a point z the force F due to misfit strain tends to pull the dislocation in the direction shown. In equilibrium this force is balanced by the component $T\cos\theta$ of the line tension of the dislocation line. The strain z is relieved by n misfit dislocations (i.e., the x components of the threading dislocations) each of length x. Therefore the strain $\epsilon(z)$ is given by[147]

$$\epsilon(z)=\epsilon(0)-n\frac{b}{2}x(z). \tag{13}$$

The treatment is approximate because it does not properly take into account the image forces. The only adjustable parameter in Eq. (13) is n. The strain values calculated using Eq. (13) do not agree with the observed values for any reasonable concentration of the dislocations. The equilibrium theories of strain relaxation are important, but generally the values of strain and dislocation concentration calculated using these theories do not agree with the experiments.[128] Jain *et al.*[146] therefore gave a phenomenological model to interpret these experiments. The model gives

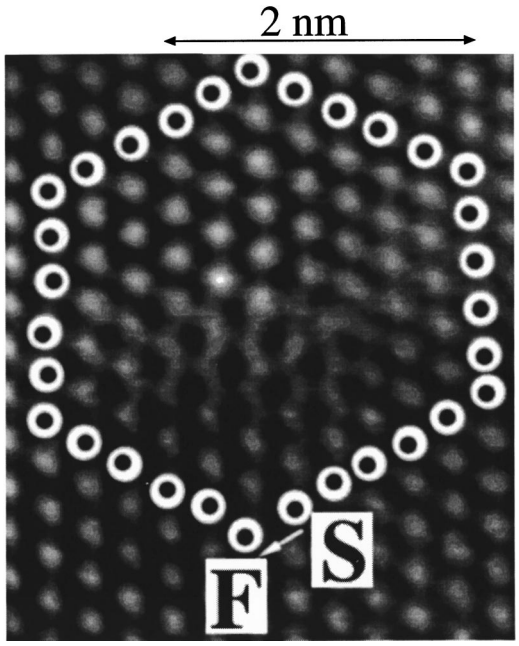

FIG. 14. High resolution transmission electron microscopy micrograph showing edge-on dislocation with Burgers vector $(1/3)[11\bar{2}0]$.

$$Kz + K_0 = \frac{1}{a_1}\left(\ln\frac{\epsilon(z)}{a_1 + a_2\epsilon(z)}\right), \tag{14}$$

where a_1 and a_2 are the adjustable parameters, $K = nb/2$, and K_0 is the boundary value at $z = 0$. The agreement between the experiments and the model is quite good.

IV. HIGH DENSITY OF DEFECTS

A. Structural defects

The point defects, e.g., nitrogen vacancies, oxygen, and other impurities, and donor–acceptor pairs are discussed in sections on the optical and electrical properties of the nitride semiconductors. We discuss the structural defects in this section. The GaN-based III–nitride heterostructures are found to contain characteristic one-dimensional (edge, mixed, and screw dislocations) and two-dimensional (stacking faults and domain boundaries) defects. These defects act as nonradiative recombination centers. Depending upon the growth conditions, it is possible to vary the defect microstructure of the specimens from inversion domain boundaries and stacking faults to edge dislocation with $(1/3)[11\bar{2}0]$ Burgers vector associated with subgrain boundaries.[90,148,149] Figure 14 shows the atomic arrangement around a $b = (1/3)[11\bar{2}0]$ dislocation, where the Burgers circuit is drawn using start–finish right-hand convention in a GaN film.

During growth of the III–nitride thin film heterostructures, polarity plays an important role in the formation of microstructural defects. It was found[150] that GaN grown by MOCVD on low-temperature buffer layer (or without buffer) has N-terminated polarity with inclusions of Ga-terminated domains. An increase of growth temperature of buffer layers results in the formation of unipolar Ga-terminated films. The inversion domains were envisaged to originate from the steps at the substrate surface and expected to be suppressed by the buffer layers. Nanopipes observed in many studies represent yet another type of defect which has been identified as the hollow core of a dislocation with (0001) Burgers vector.[151,152] These nanopipes are envisaged to be domains of N-terminated polarity inside the Ga-terminated matrix having lower than matrix growth rate.[153] The polarity of the films also influences the quality of the surface of the growing layer, which might be especially important from device considerations of heterostructures. It has been proposed that some of the Ga-terminated domains have higher growth rates originating from macroscopic hexagonal pyramids.[153] The nitridation of the sapphire substrates is an important factor for obtaining flat unipolar films. When the nitridation of the samples is optimized, the presence of a discontinuous thin cubic layer was observed at the sapphire/buffer interface.[154]

Lester *et al.*[59] measured the dislocation concentration in high efficiency LEDs fabricated by Nakamura's group. The dislocation density in the active layers of the GaN devices was $\sim 10^{10}\,\mathrm{cm}^{-2}$.[59,86] Until recently the density of the threading dislocations present in practically all the high performance LEDs and laser diodes was $10^8 - \sim 10^{12}\,\mathrm{cm}^{-2}$. The main reason for this high density is the large lattice mismatch between the III–nitrides and the sapphire substrate. In conventional III–V semiconductor devices, a dislocation density of $> \sim 10^3\,\mathrm{cm}^{-2}$ degrades the device performance. A density $> 10^6\,\mathrm{cm}^{-2}$ would be fatal for light emitters fabricated with III–V semiconductors. The dislocations are nonradiative recombination centers and reduce the optical output. The heat generated by recombination at the dislocations induces motion and multiplication of defects and very soon the devices burn out or stop emitting any light. In spite of this large dislocation density, the efficiency of the GaN-based LEDs is higher than AlGaAs and AlInGaP-based LEDs and light emitters. This raises an important question why light emitters fabricated with epitaxial nitride layers work so well for photonic and optoelectronic devices.

Earlier studies suggested that dislocations did not act as recombination centers in GaN-related materials and that the surface and grain boundaries are not sinks for the minority carriers.[59] However recent work shows that dislocations do reduce considerably the light output from the LEDs.[101] They act as pipes through which metals atoms (e.g., Mg) diffuse rapidly. They increase the leakage current.[101] The direct evidence that dislocations are nonradiative recombination centers comes from several other experiments.[150,156] Rosner *et al.*[156] studied the GaN epilayers using TEM, AFM, and wavelength resolved CL imaging techniques. They calculated the density of the light generated carriers near a dislocation line, assuming that the recombination probability at the dislocation is unity. They showed that the hole density p near a dislocation varies as

$$p = G\tau[1 - \exp(-r/L_p)]. \tag{15}$$

Here G is the generation rate, τ is the minority carrier lifetime, and L_p is the diffusion length. Using a computer code, this solution for the minority carrier density was convoluted with the observed pit distribution. The diffusion length L_p was adjusted until the processed map of the pits and the CL image agreed. The value of the diffusion length that gave the

best results was 250 nm. The reduction in the minority carrier density was large. These were probably the first results which demonstrated that dislocations in GaN are not benign and are nonradiative recombination centers. More recently Sugahara *et al.*[150] have confirmed these results, i.e., the dislocations are nonradiative recombination centers. They took plan-view TEM and CL images from exactly the same spot of a sample. They found one to one correspondence in the dark spots in the CL images and the dislocations present seen in the TEM. They found that the efficiency of luminescence is high as long as the diffusion length is smaller than the dislocation spacing. For a dislocation concentration of $10^{10}\,\mathrm{cm}^{-2}$, the average dislocation spacing is 100 nm. The value of the diffusion length determined from CL mapping experiments is about 60 nm (Refs. 157,17, and references given therein). This value is considerably smaller than the value obtained by Rosner *et al.*[156] quoted above.

B. Effect of defects on the minority carrier recombination in GaN

Ponce has given a good account of the distribution of dislocations in the GaN epilayers.[86] The dislocations are clustered in local regions of the epilayer leaving large volumes of the layer defect free. Jain *et al.*[146] have constructed a model to interpret the effect of dislocations on minority carriers in GaN epilayers. The model is based on the concept that the trapping/recombination centers can only cope with a limited number of recombinations per unit time. As the carrier density increases, the loss of carriers by recombination increases in the beginning but saturates at a constant value for large injected carrier density. The ratio of carriers lost by recombination divided by the number of the injected carriers decreases monotonically as the injected carrier density increases. For sufficiently large density of the injected carriers, the fraction lost by the recombination at the dc becomes negligible. The model also allows us to calculate the volume surrounding the dc which is depleted of the carriers and the volume of the quasineutral region, which is practically free from the effect of the dcs. This allows calculation of the characteristics of the device being designed.

Jain *et al.*[146] considered planar dcs uniformly distributed in a GaN epilayer as shown in Fig. 15. The treatment can be easily extended to other geometries.[158] The carriers are injected in GaN either by photogeneration or electrically by making a $p-n$ junction. A uniform generation rate g of the carriers is assumed throughout the crystal to simplify the calculations. The carrier densities at the defect surface, $n(r_c)$ and $p(r_c)$, are expressed in terms of the band bending $\psi_B = \psi(r_c) - \psi(W)$ (ψ is the electrostatic potential). The carrier densities can be written as[158]

$$p(r_c) = n_i \exp q\left(\frac{\psi_B + \psi(W) - \phi_p(r_c)}{kT}\right), \tag{16}$$

$$n(r_c) = n_i \exp q\left(\frac{\phi_n(r_c) - \psi_B - \psi(W)}{kT}\right), \tag{17}$$

where the symbols $\phi_n(r_c)$ and $\phi_p(r_c)$ denote the quasi-Fermi levels of the electrons and holes at the dc. The charge neutrality condition requires that the charges in the dc and in the depletion layer are equal. The charge in the dc is obtained by integrating the density of the trapping states multiplied by their occupancy probability over the energy gap. Known or other reasonable values of capture cross sections are used in the calculations. The charge in the depletion layer is calculated in terms of ψ_B. The band bending ψ_B can be determined by equating the two charges. The $p(r_c)$ and $n(r_c)$ are now calculated using Eqs. (16) and (17). The recombination current J_r at the dc is calculated in terms of known $p(r_c)$ and $n(r_c)$ using the Shockley–Hall–Read equations.

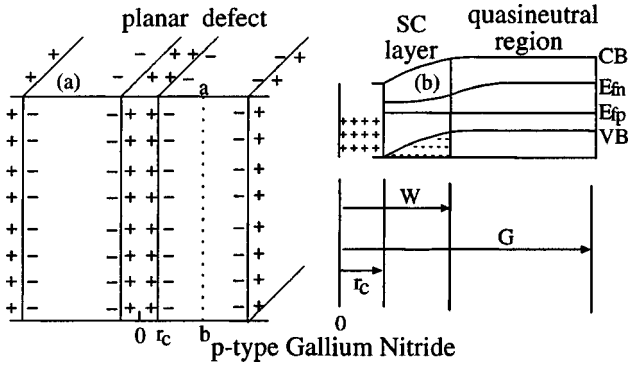

FIG. 15. (a) The + symbols show a planar dc with trapped holes in a p-GaN epilayer. The − symbols show the regions depleted of holes on both sides of the dc. (b) Schematic band bending and carrier concentration around the dc. W is the depletion region and G is the quasineutral region. Quasi-Fermi levels of electrons and holes are also shown.

It is now assumed that the recombination in the space charge (SC) layer is negligible and $J_r = J_n(W)$, where $J_n(W)$ is the current flowing from the quasineutral region into the SC layer. The transport equation in the quasineutral region is solved and the values of $n(W)$ (the minority carrier density at the edge of the SC layer) are obtained.[158] The effective recombination velocity at the edge can be calculated using the following expression:[158]

$$S_{\mathrm{eff}}(W) = \frac{J_n(W)}{q[n(W) - n_0(W)]} = \frac{J_r}{q[n(W) - n_0(W)]}, \tag{18}$$

where n_0 is the thermal carrier density of the minority carriers. Jain *et al.*[146] have made calculations assuming $G \gg W$, a trap density of $1.5\times10^{11}\,\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$ uniformly distributed over the band gap and using the capture cross sections $\sigma_n = 1\times10^{-16}\,\mathrm{cm}^2$ (neutral traps), $\sigma_c = 1\times10^{-14}\,\mathrm{cm}^2$ (Coulombic attractive traps). These calculations show that $S_{\mathrm{eff}}(W)$ changes from 10^6 to $10^3\,\mathrm{cm\,s}^{-1}$ as $n(W)$ changes from 10^6 to $10^{15}\,\mathrm{cm}^{-3}$. It should be noted that the model works only when the recombination velocity at the dcs is small. In other cases $S_{\mathrm{eff}}(W)$ continues to be large for very high densities of the injected carriers.

The model as outlined above explains the high efficiency of light emission only if the dislocations are not efficient recombination centers. However, the model can also be used for the case when the dislocations are efficient nonradiative recombination centers. In this case the capture cross sections are modified to give a high recombination rate at the dislocations. The width of the SC layer can be calculated by the method outlined above. The volume from which carriers are

lost by recombination at the dislocations is the SC layer plus a region equal to the diffusion length extending into the quasineutral region. The model predicts that if the quasineutral region is much larger than the diffusion length, the fractional loss of carriers by recombination at the dislocations will be small. It should be emphasized that this concept is useful only if the radiative recombination in the quasineutral region dominates and determines the diffusion length. The other condition for this concept to work is that the dislocations do not propagate into the quasineutral region and do not multiply and increase in number. We have already seen that the velocity of dislocations in GaN is very small (see Sec. III C). The layer growth is columnar and the point defects are gettered at the grain boundaries. The optical output within a column, which is the quasineutral region, is very high (Ref. 4, p. 18). The modified model as discussed above is consistent with these results.

Local regions of high concentration of In can act as quantum dots and confine the carriers. This can give rise to high efficiency of emission even when the dislocations are present. This topic is discussed in Sec. IX A 4.

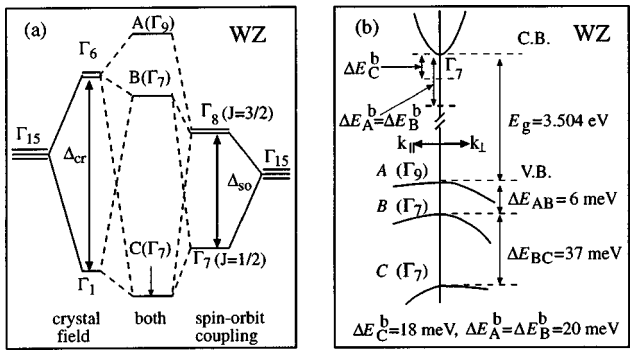

FIG. 16. (a) Effect of crystal field splitting and spin-orbit coupling on the valence band (near the Γ point) of WZ GaN (Ref. 160). At $k=0$, the valence band is split by the combined action of crystal field and spin-orbit coupling into $A(\Gamma_9)$, $B(\Gamma_7)$ and $C(\Gamma_7)$ states. (b) Calculated values of the band energy levels (Ref. 161) near the Γ point. $\Delta E_A^b, \Delta E_B^b, \Delta E_C^b$ are the binding energies of the *A*, *B*, and *C* excitons, respectively.

V. BAND STRUCTURE OF III–NITRIDES

A. Zincblende (ZB) structure

In unstrained zincblende crystals at $\mathbf{k}=0$, the valence band of an unstrained crystal consists of a fourfold $P_{3/2}$ multiplet ($J=3/2$, $m_J=\pm 3/2, \pm 1/2$ in spherical notation) and a twofold $P_{1/2}$ multiplet ($J=1/2$, $m_J=\pm 1/2$). These bands are also described in an alternative notation: $|3/2, \pm 3/2\rangle$, $|3/2, \pm 1/2\rangle$ and $|1/2, \pm 1/2\rangle$ bands, respectively. Holes in the $|3/2, \pm 1/2\rangle$ band have a smaller effective mass than the holes in the $|3/2, \pm 3/2\rangle$ band. The two hole bands are known as the light hole (LH) and the heavy hole (HH) bands, respectively.

Biaxial strain can be resolved into a hydrostatic part and a shear part. The hydrostatic component does not produce splitting of the valence band. Hydrostatic stress changes the distance between the conduction band and the three uppermost valence bands, i.e., the heavy hole, light hole, and spin-orbit split-off (SO) bands. The distance between the conduction band and these bands increases under hydrostatic compression and decreases under hydrostatic tension. The changes in the band gaps can be shown by a shift of the conduction band edge. It is not essential to distribute the shift between the conduction band and the valence band edges. The shear component produces a tetragonal distortion and lowers the symmetry. The degeneracy of the valence band multiplet is removed by the shear component of the biaxial strain. The two hole bands move in opposite directions away from the center of gravity of the two bands. Under compression the $|3/2, \pm 3/2\rangle$ HH band moves up and determines the band gap. Under tensile strain, the $|3/2, \pm 1/2\rangle$ LH band moves up. Now the LH band determines the band gap.

For small biaxial strain ϵ, the eigenvalues of the orbital-strain Hamiltonian can be calculated and the shifts in band gaps can be determined.[159] Since we are interested only in the wurtzite III–nitrides in this article, we do not discuss the splitting of the ZB valence band here. The article of Shahzad *et al.*[159] and our review[108] may be consulted for a detailed discussion of this subject.

B. Wurtzite (WZ) III–nitrides

Band structure of WZ GaN has been investigated by several groups.[161,133,162,67,160] Using the local density approximation as implemented by the all-electron relativistic, full-potential, linearized augmented plane wave (FLAPW) method, several groups[160–162] have performed the first principle band structure calculations for GaN and AlN. Local coordination of the first neighbor atomic positions is the same in the WZ and the ZB structures. The difference in the two structures arises in the positions of the distant neighbors. The layer stacking in the [0001] direction in WZ structure is equivalent to that along the [111] direction in the ZB structure. Based on these considerations, a quasicubic approximation is used for solving the Luttinger–Kohn Hamiltonian using the FLAPW method.[160] A schematic diagram of the splittings is shown in Fig. 16(a) and the calculated values of the band energy levels are shown in Fig. 16(b). The degeneracy of the valence band is lifted due to crystal field and spin-orbit coupling. The valence band splits into three bands designated as *A*, *B*, and *C* bands (see Fig. 16). The symmetry of the *A* band is Γ_9 and that of the *B* and the *C* bands is Γ_7. The calculated band structure near the Γ point is shown in Fig. 16(b). The calculated and observed values of the band structure parameters of WZ III–nitrides are shown in Table I. The main difference between the valence band structure of ZB and WZ crystals is the crystal field splitting in the WZ crystals. Because of the cubic symmetry this splitting is absent in the ZB structures.

The band gaps corresponding to $A(\Gamma_9)$, $B(\Gamma_7)$, and $C(\Gamma_7)$ levels shown in Fig. 16 are given by[162]

$$E(A(\Gamma_9)) = 1/2(\Delta_{SO} + \Delta_{CF}), \tag{19}$$

$$E(B(\Gamma_7)) = +1/2[(\Delta_{SO} + \Delta_{CF})^2 - 8/3(\Delta_{SO}\Delta_{CF})]^{1/2}, \tag{20}$$

and

TABLE I. Band structure parameters Δ_{SO} and Δ_{CF} (both in meV), and average effective masses m_e^*/m_0 and m_h^*/m_0 in WZ AlN, GaN and InN.

Method	ΔSO	Δ_{CF}	m_e^*/m_0	m_h^*/m_0	Ref.
		AlN			
Cal.	20.4	−58.5	—	—	160
Cal.	19	−217	—	—	162
		GaN			
Cal.	12	37.5	0.18	0.60[a]	161
Cal.	15.6	72.9	0.18	0.95–1.1[b]	160
Expt.	11(+5,−2)	20–24	0.6–2	—	160
Cal.	13	42	—	—	162
		InN			
Cal.	1	41	—	—	162

[a]This value of m_h/m_0 is for the *A* band. The values for the *B* and *C* bands are 0.52 and 0.61, respectively.
[b]Average of the hole masses in the *A*, *B*, and *C* bands.

$$E(C(\Gamma_7)) = -1/2[(\Delta_{SO}+\Delta_{CF})^2 - 8/3(\Delta_{SO}\Delta_{CF})]^{1/2}. \tag{21}$$

The parameters Δ_{SO} and Δ_{CF} cannot be directly measured. The measurable quantities shown in Fig. 16(b) are[162]

$$\Delta E_{AB} = E(A(\Gamma_9)) - E(B(\Gamma_7)), \tag{22}$$

and

$$\Delta E_{BC} = E(B(\Gamma_7)) - E(C(\Gamma_7)). \tag{23}$$

The parameters Δ_{SO} and Δ_{CF} are determined by fitting the measured values of ΔE_{AB} and ΔE_{BC} with Eqs. (19), (20), and (21). Numerical calculations show that $\Delta E_{AB} < \Delta E_{SO}$ and $\Delta E_{BC} \sim \Delta E_{CF}$. There has been some confusion in the literature because the difference between ΔE_{AB} and ΔE_{SO} or between ΔE_{BC} and ΔE_{CF} has not been recognized by some authors (e.g., see Fig. 6.1 of Ref. 67). According to Ref. 160, the calculated values of spin-orbit splitting Δ_{SO} are 15.6 meV for WZ GaN and 20 meV for ZB GaN. The corresponding observed values are 11(+5,−2) meV and 17±1 meV, respectively, for the two structures. The agreement between the theoretical and experimental values is reasonable. It is interesting to examine the trends in the values of the spin-orbit parameter Δ_{SO} in going from AlN through GaN to InN. For conventional III–V semiconductors, $\Delta_{SO}(InX) > \Delta_{SO}(GaX) > \Delta_{SO}(AlX)$ for X=P, As, or Sb (Ref. 162, and references given therein). This is because *p* states make a positive contribution and *d* states make a negative contribution to Δ_{SO}. In phosphides, arsenides, and antimonides, there is a weak coupling of the cation *d* orbital at the valence band maximum and Δ_{SO} is determined mainly by the *p* bonding. In the nitrides there is a considerable hybridization of the nitride *p* orbitals with the cation *d* orbitals. This reduces Δ_{SO}, the reduction being more in the heavier cation nitrides. The results of Ref. 162 given in Table I show that for the WZ nitrides the order reverses and $\Delta_{SO}(InX) < \Delta_{SO}(GaX) < \Delta_{SO}(AlX)$. The values of Δ_{SO} calculated in Ref. 162 are smaller in the WZ nitrides than in the ZB nitrides, particularly in the heavier cation InN.

The calculated crystal field splitting Δ_{CF} is 72.9 meV in WZ GaN and −58.9 meV in WZ AlN according to Ref. 160. The corresponding values calculated in Ref. 162 are −217 and 42 meV. It is generally believed that the crystal field splitting for AlN is negative [the order of Γ_6 and Γ_1 in Fig. 16(a) is reversed, for AlN Γ_1 is higher] and for GaN it is positive.[162,160] However there is a considerable discrepancy in the values calculated by different groups. The observed value of Δ_{CF} in WZ GaN is 21–22 meV. The discrepancy between the calculated and the observed values of this parameter is also large. More experimental and theoretical work is necessary to resolve this discrepancy.

The calculated and experimental values of the average effective electron and hole masses are also shown in Table I. The calculated values of hole mass show considerable *k* dependence.[161,160] However, experimentally only average (density of states) values have been determined. For values of effective masses in different *k* directions, articles by Suzuki and Uenoyama[160] and Chen *et al.*[161] may be consulted.

C. Strained WZ GaN

The band structure and electronic properties of semiconductors are considerably modified by strain. To determine the band structure of the strained crystals, the components of the strain tensor are required. We first discuss the strain tensor components that arise in WZ GaN epilayers.

The stress–strain relations for a hexagonal crystal (C_{6v} symmetry) are expressed by a 6×6 matrix. However if the crystal is strained in the (0001) plane and is free in the [0001] direction, there are only three nonvanishing strain components given by[160]

$$\epsilon_{xx} = \frac{a-a_0}{a_0}, \quad \epsilon_{yy} = \frac{b-a_0}{a_0}, \tag{24}$$

$$\epsilon_{zz} = \frac{c-c_0}{c_0} = \frac{C_{13}}{C_{33}}(\epsilon_{xx}+\epsilon_{zz}). \tag{25}$$

Here the subscript ''0'' indicates the lattice constants of the unstrained crystal and *a*, *b*, and *c* are the lattice constants of the strained crystal. C_{ij} are the elastic constants. The values of the elastic constants are given in Ref. 18. If the strain is biaxial, $a=b$. In a pseudomorphic epilayer, *a* and *b* are equal to the lattice constants of the substrate material and we have

$$\epsilon_{xx} = \epsilon_{yy} = \frac{a_{sub}-a_0}{a_0}, \quad \epsilon_{zz} = -\frac{2C_{13}}{C_{33}}\epsilon_{xx}. \tag{26}$$

If the stress is uniaxial, i.e., if only $\sigma_{xx} \neq 0$ and all the other stress components are zero, the nonzero strain components are[160]

$$\begin{pmatrix} \epsilon_{yy} \\ \epsilon_{zz} \end{pmatrix} = \frac{1}{C_{13}^2 - C_{11}C_{33}} \begin{pmatrix} C_{12}C_{33}-C_{13}^2 \\ C_{11}C_{13}-C_{12}C_{13} \end{pmatrix} \epsilon_{xx}. \tag{27}$$

Under the approximation of strain independent isotropic spit-orbit coupling, the effect of strain on the *A*, *B*, and *C* band gaps (shown in Fig. 16) can be described by the following equations:[67]

$$E_A - E_A(0) + D_1\epsilon_{zz} + D_2(\epsilon_{xx}+\epsilon_{yy}) + D_3\epsilon_{zz} + D_4(\epsilon_{xx}+\epsilon_{yy}), \tag{28}$$

$$E_B - E_B(0) + D_1\epsilon_{zz} + D_2(\epsilon_{xx}+\epsilon_{yy}) + \Delta_+[D_3\epsilon_{zz} + D_4(\epsilon_{xx}+\epsilon_{yy})], \tag{29}$$

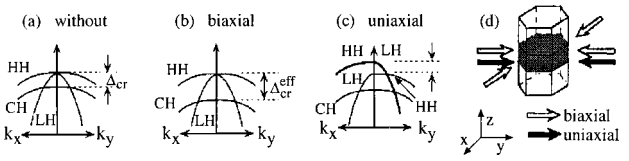

FIG. 17. Schematic band structure near the top of the valence bands of the WZ GaN in the k_x-k_y plane; (a) without strain, (b) with a biaxial strain, and (c) with a uniaxial strain in the *c* plane. The direction of each strain is shown in (d). The figure is taken from Ref. 160.

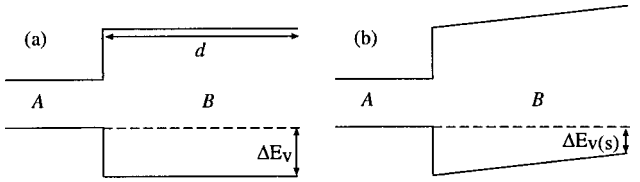

FIG. 18. (a) The heterojunction between epilayer *B* grown on epilayer *A*. The heterojunction is designated as the *B*/*A* heterojunction. ΔE_V is the valence band offset and *d* is the epilayer thickness. (b) Effect of strain induced electric field on the valence band offset. The offset $\Delta E_{V(S)}$ in the strained layer is considerably reduced.

$$E_C - E_C(0) + D_1\epsilon_{zz} + D_2(\epsilon_{xx}+\epsilon_{yy}) + \Delta_-[D_3\epsilon_{zz} + D_4(\epsilon_{xx}+\epsilon_{yy})]. \quad (30)$$

$E_I(0)$ represents the strain free exciton energies, D_I are the deformation potentials, and ϵ_{ij} are the components of the strain tensor. D_1 and D_2 are the deformation potentials for the hydrostatic strain and D_3 and D_4 are the uniaxial deformation potentials which describe the splitting (due to strain) of the three topmost valence band edges for tension or compression along and perpendicular to the (0001) direction. Most numerical values of D_1 vary between -6 and -8 eV and those of D_2 vary between -8 and -12 eV (see Table 6.5 of Ref. 67). Some $+ve$ values of these deformation potentials have also been reported. The uncertainty arises because the strain in the epilayers is induced by both the lattice mismatch and thermal mismatch and is difficult to determine accurately. Similarly both $+ve$ and $-ve$ values of D_3 and D_4 varying between 0.7 and 8 eV have been reported. The coefficients $\Delta_\pm$ are given by[67]

$$\Delta_\pm = \frac{1}{2}\left\{1 \pm \left[1 + 8\left(\frac{\Delta_3}{(\Delta_1-\Delta_2)}\right)^2\right]^{-1/2}\right\}. \quad (31)$$

Here $\Delta_1=\Delta_{CF}$ and $\Delta_2=\Delta_3=\Delta_{SO}/3$. The values of Δ_{SO} and Δ_{CF} are given in Table I.

The band structure of the strained GaN (Ref. 160) is shown in Fig. 17. The uniaxial strain makes the effective hole masses different for k_x and k_y directions. For actual values of hole masses for different values of *k* the article by Suzuki and Uenoyama[160] may be consulted.

D. Band offsets

Knowledge of the band offsets at the heterojunctions is vital for determining the degree of confinement of the carriers in quantum well and quantum dot structures. The confinement plays a vital role in designing the electronic and optoelectronic devices. The band offsets of the III–nitrides have been calculated and measured by several groups.[120,132,133,162] The three nitrides form three heterojunctions: AlN/GaN (*A*/*B* implies that epilayer *A* is grown on top of layer *B*), GaN/InN, and InN/GaN. The band offsets at the *A*/*B* heterojunction are different from the offsets in the *B*/*A* heterojunction.[133] We therefore have six pairs of heterojunctions with six different band offsets. In all six cases the band offsets are of type I, i.e., the smaller gap lies within the wider gap as shown in Fig. 18(a). Valence band offsets can be calculated using first principle methods[162] and can be determined using XPS. The conduction band discontinuity can be determined using the known band gap values (see Fig. 1) of the two nitrides.

The valence band discontinuity ΔE_V is given by[133]

$$\Delta E_V = (E_v - E_{CL})_A - (E_v - E_{CL})_B + (\Delta E_{CL})_{A-B}. \quad (32)$$

The two $(E_v - E_{CL})$ are the conveniently identifiable characteristic core levels of the bulk nitride layers *A* and *B* (see Fig. 18). The quantities are measured by XPS from the valence band maximum on separate *A* and *B* epilayers. They are characteristic of the bulk materials. They do not depend on the structure of the interface. $(\Delta E_{CL})_{A-B}$ is the energy difference between characteristic core levels on the two sides of the actual heterojunction. $(\Delta E_{CL})_{A-B}$ and ΔE_v are determined by the structure of the interface. $(E_v - E_{CL})_A$, $(E_v - E_{CL})_B$, and $(\Delta E_{CL})_{A-B}$, are calculated from the first principles[162] or measured by XPS.[133,132,120] The valence band discontinuity (ΔE_V) is then calculated using Eq. (32). The measured[120,132,133] and the calculated (using the LAPW method)[162] values do not agree well. The values depend on the strain state of the epilayer. The calculated value for AlN grown on GaN is 0.81[162] and the experimental values vary from 0.57 eV[133] for the strained layer to 1.36 eV[132] for the relaxed layer. Values of ΔE_V at the GaN/$Al_{0.2}Ga_{0.8}N$ interface have been discussed in Ref. 160.

It is seen from the results of Martin *et al.*[133] that the value of the band offset for a given *A*–*B* heterojunction is not unique. It depends on whether the pseudomorphic epilayer *A* is grown on the relaxed layer *B* (denoted by *A*/*B*) or of the strained epilayer *B* grown on the relaxed layer *A*. The difference arises because strain gives rise to piezoelectric fields in the nitrides (the piezoelectric effect is discussed in Ref. 18). The field induces band bending shown in Fig. 18. Since escape depth of the photoelectrons is small, the measured values must correspond to a thin surface layer of the epilayer. In the case considered in Fig. 18(b), the valence band offset is reduced by the field, i.e., $\Delta_{V(S)} < \Delta E_V$. If the field direction changes and bands bend in the opposite direction, $\Delta_{V(S)} > \Delta E_V$. Since piezoelectric constants of the layers *A* and *B* are different, the valence band offset in the heterojunction *A*/*B* is different from its value in the *B*/*A* heterojunction. It is also clear from this figure that the measured valence band offset in the strained layers will depend on the epilayer thickness *d*. Indeed Rizzi *et al.*[120] have observed such a dependence on the thickness of the epilayer. They found that the valence band offset in AlN/SiC decreases from 1.7 eV for small AlN thickness to 1.1 eV when the

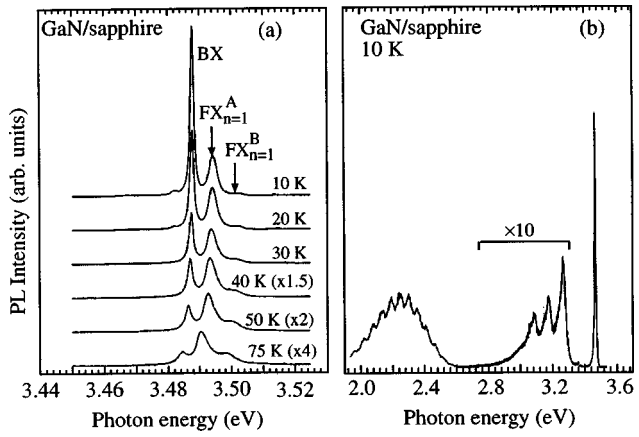

FIG. 19. (a) Near-band-edge exciton luminescence spectra taken from a 7.2 μm GaN epifilm grown on sapphire by MOVPE. (b) Typical luminescence spectra including the structure on the low energy side of the band-edge luminescence from a nominally undoped GaN epifilm (Ref. 67).

thickness is about 6.7 nm. On the other hand the valence band offset in the AlN/GaN heterojunction increases from 0.2 eV for small AlN thickness to about 0.6 eV when the thickness increases to 9 nm. Both Martin *et al.*[133] and Rizzi *et al.*[120] have calculated the effect of strain induced piezoelectric field and found that the observed values are consistent with the calculated values.

VI. OPTICAL PROPERTIES OF III–NITRIDES

A. Near band-edge luminescence

1. Exciton transition energies

Extensive work on the luminescence of III–nitrides (mostly GaN) has been done and several reviews have been published.[67,160–181] Typical near band-edge PL spectra of GaN at different temperature is shown in Fig. 19(a). At low temperatures the spectra are dominated by the *BX* peak due to recombination of excitons bound to impurities. Since the as grown GaN is generally *n* type the *BX* peak is most likely due to the excitons bound to neutral donors. As the temperature rises, the *BX* peak decays faster than the free exciton peaks marked $FX^A_{N=1}$ and $FX^B_{N=1}$ and all the peaks move to lower energies. The two peaks are due to the *A* and the *B* excitons (i.e., due to transitions involving holes in the *A* and *B* valence bands shown in Fig. 16). At 75 K the *A* exciton peak becomes the strongest. The strength of the PL is sensitive to the quality of the crystal. Nonradiative recombination is promoted by multiphonon processes, Auger recombination, recombination at defects, and recombination at surfaces. The excitons may also migrate away from the area from where the detecting system receives the emitted light.

The splitting of the valence band can be determined from the low temperature PL, photoreflectance (PR), calorimetric absorption, and calorimetric reflection measurements.[67,165] Assuming that the binding energy of the *A*, *B*, and *C* excitons are nearly equal and their binding energies with the impurities are known (see below), the difference in the observed exciton peaks due to transitions from the $A(\Gamma_9)$, $B(\Gamma_7)$, and $C(\Gamma_7)$ bands can be used to determine the splitting of the valence band. Transitions from the $A(\Gamma_9)$ and $B(\Gamma_7)$ bands have been observed by several workers. Transitions from $C(\Gamma_7)$ band have been observed in PL by Götz *et al.*[163] and in PR measurements by Shan *et al.*,[164] Gil (unpublished work of Gil quoted in Ref. 165), and others. We have compared the experimental and calculated values of the splittings ΔE_{AB} and ΔE_{BC} of the valence band. We found that the agreement between the calculated and the observed splittings is quite good.[6]

PL peaks have also been measured with ultrahigh intensities of exciting light. At the very high excitation intensities PL lines broaden and new lines on the low energy side are observed. The changes in the PL are attributed to two interacting excitons and/or to electron–hole plasma created by the high excitation intensity.[165]

The *A* and *B* exciton peaks in MBE grown GaN have been observed by Grandjean and Massies.[112] The positions of these peaks are similar to those obtained with the MOVPE grown layers. However extensive measurements of PL at different temperatures have not been made in the MBE grown layers.

2. Binding energies and lifetimes of the excitons

The lowest excited state corresponds to the Wannier–Mott intrinsic exciton. In these excitons the electrons and holes interact with each other through their Coulomb field influenced only by the dielectric constant of the semiconductor. The energy of the crystal is lowered by the formation of the excitons (relative to that of the unrelated free carrier state) by an amount equal to the binding energy of the excitons. The binding energy of the exciton (with zero wave vector) is given by

$$E_n = e^4\mu/2(4\pi\hbar\epsilon)^2 n^2. \tag{33}$$

The absorption and luminescence occurs due to the formation and recombination of excitons and therefore their energies are given by[67]

$$E = E_g - E_b, \tag{34}$$

where E_b is the binding energy of the exciton in the ground state, i.e., the energy necessary to ionize the exciton. It is obtained by using $n=1$ in Eq. (33). The experimental values of E_b for the excitons in GaN vary from 18 to 25 meV. The excitons have a tendency to localize near an impurity. The energy with which the exciton is bound to a donor or an acceptor is denoted by E_{BX}. The energy of an emitted photon becomes

$$\hbar\omega = E_g - E_b - E_{BX}. \tag{35}$$

Chen *et al.*[161] have also calculated the binding energies ΔE^b_A, ΔE^b_B, and ΔE^b_C of the *A*, *B*, and *C* free excitons formed in the $E[A(\Gamma_9)]$, $E[B(\Gamma_7)]$, and $E[C(\Gamma_7)]$ bands, respectively. The calculated values are shown in Fig. 16(b). The calculated binding energies are nearly equal for the three excitons. The reason for this is that the binding energy is mainly determined by the electron mass because hole mass is very large. The calculated values for WZ CdS show similar trends. Chen *et al.*[161] observed the free exciton line at 3.484 eV at 10 K and a line due to the exciton bound to a neutral donor at 3.476 eV. The band gap value at these low tempera-

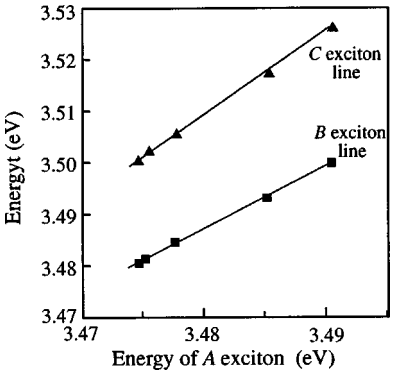

FIG. 20. Trends in the free *B* and *C* exciton energies versus the position of the *A*-exciton for several sample of GaN grown on sapphire. A higher compressive strain in the epilayer leads to higher exciton energies (Ref. 170).

tures is 3.504 eV.[161] The transition due to the exciton bound to an acceptor was at 3.459 eV. These results give binding energy of a free *A* exciton as 20 meV. The binding energy for the exciton bound to a neutral donor comes out to be 8 meV and for the exciton bound to a neutral acceptor, it is 25 meV. The binding energies of free and bound excitons determined by several other authors using different techniques have similar values.[67,160] The calculated value of 20 meV for the *A* exciton agrees well with the experimental value.[161] A more detailed discussion of the binding energies and other properties of excitons in WZ GaN is given in Refs. 155, 160.

Exciton recombination dynamics and lifetimes of excitons have been reviewed by Monemar *et al.*[165] Time resolved experiments on PL decay show that at 10 K the radiative lifetime of the free excitons is ~35 ps and of the bound excitons it is ~55 ps.[67] Similar values are also observed in III–V and II–VI semiconductors.[168,169] Song and Shan[67] have discussed the results of several authors who have determined the binding energy of the excitons using the results of optical measurements. The binding energies of the *A* excitons vary between 20 and 28 meV. Values of binding energies of the *B* and the *C* excitons have been reported only in a few cases, and they vary between 18 and 23 meV.

B. Effect of strain

1. Effect of strain on the exciton transition energies

We saw in section V C that strain strongly modifies the band structure. Both the band gap and the splitting of the valence band are changed. Compressive strain increases the splitting between the *A*, *B*, and *C* bands (see Sec. V C, see also Ref. 165, pp. 96 and 97). Experimental results of several workers confirm that this indeed is the case.[165] The most direct confirmation of the strain effects on the near band-edge PR is provided by Fig. 20. As the strain in the samples increases, the energy of all the three excitons shifts upwards as expected. However the shifts are not equal. The *C* exciton moves by up to 30 meV, whereas the maximum shift in the *A* and *B* is 20 meV. It is expected that the excitons will move downwards in the GaN epilayers grown on 6H SiC. However the widths of the PL and PR spectra in the layers grown on SiC are large and such clear cut results have not been obtained.[165] Other experiments which show the effect of strain on optical properties have been discussed by Song and Shan.[67]

Popovici *et al.*[28] have discussed the stress and strain in GaN layers grown on different planes of sapphire and on different buffer layers. The films grown on AlN buffers and on *a*-plane sapphire are generally under compression. The films grown without the buffer layers are poor in quality and are under tension. The films grown on GaN buffers generally have small compressive stress or no stress. These are general trends, and different results have been obtained in some cases.

2. Effect of RTA on strain and defects

Li *et al.*[171] have studied the effect of rapid thermal annealing (RTA) at 900 °C on the Raman and PL spectra of GSMBE grown GaN layers on (0001) sapphire substrates. In Raman spectra, the high energy E_2 mode was observed at 572 cm^{-1}. The A_1 LO mode at 733 cm^{-1} was not observed probably due to high free carrier concentration.[171] The frequency of the E_2 mode changed from 572 to 568 cm^{-1} after 20 s RTA at 900 °C. The shift in the frequency of the E_2 mode was attributed to the relaxation of residual strain by the annealing process. The A_{1g} and E_{1g} peaks due to the Raman scattering from the sapphire substrate were observed at 417 and 750 cm^{-1}, respectively. Two other peaks at 377 and 644 cm^{-1} were also attributed to the substrate. The bound–exciton PL peak was at 3.38 eV in the as grown layers. After the 20 s RTA, its position changed to 3.40 eV. The authors also attributed this shift to the relaxation of residual strain. No deep level luminescence was observed in the as grown epilayers. After the 20 s anneal, a PL peak at 2.75 eV was observed. The authors found that hydrogen concentration decreased as a result of RTA. In the as grown layers, the defect centers are passivated by the hydrogen and the deep level PL is not observed. During the RTA process hydrogen diffuses out and the defects give rise to the 2.70 eV PL. The defect centers responsible for this peak were not identified.

C. Temperature dependence of the optical transitions

The temperature variation of the band gap and of the excitonic transitions in GaN depends on the method of sample preparation, method of measurements, and values of coefficients α of thermal expansion (Refs. 164, 165, and references given therein). Shan *et al.*[164] used PL and PR techniques to measure the temperature dependence of the exciton energies in the temperature range 15–300 K. 4.2 μm thick GaN epilayers were grown on AlN buffer layers on (0001) sapphire substrate by MOVPE. The temperature dependence of the transition energies of the *A* excitons $[A(\Gamma_9)-\Gamma_7(\mathrm{CB})$, *B* excitons $[B(\Gamma_7)-\Gamma_7(\mathrm{CB})]$, and *C* excitons $C(\Gamma_7)-\Gamma_7(\mathrm{CB})]$ (see Fig. 16) could be described by the Varshni empirical equation[164]

$$E_0(T)=E_0(0)-\alpha T^2/(\beta+T), \tag{36}$$

where $E_0(0)$ is the transition energy at 0 K, and α and β are the temperature coefficients. The directly observed 15 K values of E_0 for the three excitons are 3.485 eV (*A* exciton),

3.493 eV (B exciton), and 3.518 eV (C exciton), respectively. The values of E_0 which fit the above equation are approximately the same. The values of α (in units of eV/K) are 8.32×10^{-4}, 10.9×10^{-4}, and 2.92×10^{-3}. The corresponding values of β are 835.6, 1194.7, and 3698.9 K, respectively. Caution should be exercised in using these results to interpret experiments in general. As we discussed earlier, the transition energies depend on the strain and strain in the epilayer changes with temperature. The effect of strain and temperature have not been separated out in the above experiments or in other experiments reviewed by Monemar *et al.*[165] Monemar *et al.*[165] have pointed out that the temperature dependence of the exciton transition energies determined by different workers varies considerably.

D. Yellow band, DAP band, and persistent photoconductivity

1. Yellow and DAP bands

PL peaks are also observed on the low energy side of the near band-edge spectra as shown in Fig. 19(b). The main features are a yellow band at about 2.5 eV and donor–acceptor pair (DAP) bands.[67,163] As mentioned earlier, the real origin of the yellow band is not known. The vibronic structure of the yellow band shows the strong electron–phonon coupling. This band is associated with deep levels. The deep levels arise due to point defects such as Ga or nitrogen vacancies.[67] The existence of this band indicates the presence of defects which have an adverse effect on the carrier lifetime and on the intensity of the near band-edge spectra. Several theoretical studies have been made and models involving transitions from shallow donors to deep acceptors or from deep donors to shallow acceptors have been proposed.[67] The pressure dependence of the yellow band is the same as that of the GaN band gap. This suggests that the transitions associated with the yellow band are from the shallow donor states to the deep acceptor state.[67] The PL lines in the energy range 2.95–3.27 eV are due to DAP recombination processes.

PL of Mg-doped WZ GaN has also been studied.[67] The PL shows the DAP PL at 3.27 eV at 4.2 K along with its LO phonon replicas.[67] If the Mg concentration is very high, the PL is dominated by a deep level broad band with its peak at 2.95 eV. The blue emission is best obtained if the Mg concentration is between 5×10^{19} and $2\times10^{20}\,\mathrm{cm}^{-3}$.[67] The yellow band is very weak under these conditions. At higher concentrations of Mg only the yellow band is observed. Doping GaN p type with other column II acceptors has not been so successful (see, however, Ref. 28).

2. Persistent photoconductivity

The photoconductivity induced in the GaN epilayers by the incident light persists for a time duration ranging from several minutes to several days.[172] The persistent photoconductivity (PPC) is present in GaN layers grown by MBE as well as by MOVPE. Several centers can be the possible sources of PPC. The suggested centers are defects with bistable character, nonstoichiometry induced potential fluctuations, and inclusions of cubic GaN.[172] It has been observed that if the yellow band is absent in the GaN epilayers, the PPC is also absent. The energies of the photons which give rise to the yellow band and those which induce the PPC are also related. These observations suggested that the yellow band and the PPC originate at the same defect centers.

Recently Reddy *et al.*[172] have investigated in detail the yellow band and the PPC in the GaN epilayers grown by a radio-frequency plasma assisted MBE on sapphire substrates. They measured the PPC with the energy of the incident photons varying in the range 1.38–2.4 eV. At small values of the energy the PPC was small. In fact at 1.38 eV no PPC was observed. As the energy increased, the PPC also increased. The decay time also depended on the energy of the incident photons. For $h\nu=1.77\,\mathrm{eV}$, the decay time was 3 h. For $h\nu=1.96\,\mathrm{eV}$, the decay time was more than 1 d. With white light, the PPC lasted several days. Measurements on cubic GaN grown on GaAs and on cubic SiC were also made. No PPC was observed in the cubic GaN. Reddy *et al.*[172] also found that the PPC is observed only in those samples in which the yellow band is present. The threshold of the yellow band and the PPC were both at $h\nu\sim1.6\,\mathrm{eV}$. The PPC increases monotonically from 1.8 to 2.2 eV which coincides with the rising part of the yellow band. These results strongly support the view that the centers responsible for the PPC are also the centers which give rise to the yellow band. As we mentioned earlier, the centers for the yellow band and background n-type conductivity are still being debated. Reddy *et al.*[172] suggested that neither the Ga interstitial nor the Ga or nitrogen vacancies are the likely centers which give rise to the yellow band or the PPC. A complex consisting of nitrogen antisite and a Ga interstitial, $\mathrm{N_{Ga}}-\mathrm{Ga}_I$ was suggested as the likely defect responsible for both. This defect is similar to the $\mathrm{As_{Ga}}-\mathrm{Ga}_I$ defect which has been found in GaAs. Reddy *et al.*[172] pointed out that simple defects like vacancies should be present in the cubic GaN also. If these defects were responsible for the yellow band or the PPC, both the yellow band and the PPC should have been observed in the cubic GaN.

E. Optical properties of GaN:Er

Several studies of erbium implanted GaN have been made.[173–181] Early work on the optical properties of the III–nitrides Er has been discussed in Refs. 173–175. The transition of the shielded $4f$ electrons of Er produces luminescence at 1540 nm. This wavelength coincides with the minimum in the loss in silica optical fibbers. The luminescence is sharp and its peak wavelength is insensitive to the environments of the host and to the temperature. Major improvements in the link distances, data rates, and reduced needs for signal regeneration have been demonstrated using optical amplifiers based on Er-doped fibers.[173] However the luminescence of Er in most semiconductor hosts quenches with the increase of temperature, particularly in GaAs and other semiconductors with relatively small band gaps. The quenching is smaller in wide band gap III–V nitride semiconductors.

Recently PL of *in situ*-doped GaN:Er grown by hydride vapor phase growth[179] and by and by MOVPE[180] has been

studied. The 11 K emission line of Er in the *in situ* doped GaN:Er samples was at shorter wavelengths by 2.5 nm as compared to its position in the implanted sample.[179] The intensity of the PL in the *in situ*-doped samples was weaker by a factor of 5. The intensity was further reduced on annealing the samples in N_2 atmosphere. The loss in intensity was recovered on subsequent annealing of the samples in H_2 or NH_3 ambient. The 295 K PL of the implanted samples showed a strong deep level luminescence at 500 nm. This emission was absent in the *in situ*-doped GaN:Er samples. Further work is required to understand the complex behavior of the *in situ*-doped GaN:Er and Er-implanted GaN layers. In the MOVPE grown sample[180] codoping with C and O increases the intensity of the 1.54 μm emission band by up to 2 orders of magnitude. C and O also increase the thermal stability of the luminescence. Increase of luminescence by codoping GaN:Er with O has also been observed in the implanted GaN samples.[178] A high concentration of O (5–10 times the concentration of Er) gives the best results. The electrically pumped laser can be fabricated using GaN doped with erbium and oxygen.[40] The hot electron induced luminescence in GaN codoped with Er and O is as efficient at room temperature as it is at 8 K (see Ref. 40 and references given therein).

AlN doped with Er during MOVPE growth and *in situ* plasma hydrogenated at 200–250 °C has been studied.[181] On hydrogenation at 200 °C the luminescence intensity of the 1.54 μm increases fivefold. The increase is not quenched on heating the sample up to 300 °C. Due to its wide band gap the 1.54 μm luminescence is more stable in AlN than in GaN. In GaN:Er the stability of the PL depends on the wavelength of the exciting light. The luminescence intensity decreases by up to 4.8 times in going from 77 K to room temperature if excited by light of wavelength shorter than 998 nm.[177]

VII. ELECTRICAL PROPERTIES: *n*-TYPE III–NITRIDES

A. Background defects responsible for the *n*-type conduction and yellow band

As grown GaN is always *n*-type with carrier concentration $\sim 10^{16}\,\mathrm{cm}^{-3}$. Theoretical calculations regarding the centers responsible for the *n*-type conductivity give conflicting results (see Refs. 28, 182–200, and references given therein). The first principle calculations of Perlin *et al.*[183] show that the nitrogen vacancies act as shallow donors. They give rise to *n*-type conductivity in nominally undoped GaN epilayers. The energy of formation of nitrogen vacancy is lower than that of Ga interstitials. On the other hand, calculations of Neugebauer and Van de Walle[182] show that the energy of formation of nitrogen vacancies is 4 eV and therefore at room temperature the vacancies will not be present in sufficiently high concentration, which is needed to explain the large background concentration of electrons. These calculations also show that the energy of formation of Ga vacancies in *n*-type GaN is low. The electronic energy levels of these defects form deep acceptor states and give rise to the yellow band. There is evidence that the yellow band is suppressed under Ga-rich conditions, which supports this idea. Experiments also show that oxygen atoms act as donors in GaN and suggestions have been made that residual impurities (Si, Ge, O) are responsible for the *n*-type conductivity in nominally undoped GaN.[28,67] These residual impurities may also be responsible for the yellow band. Thus two models exist to explain the *n*-type conductivity and the yellow band observed in GaN epilayers grown without intentional doping: According to one model the nitrogen vacancies are responsible for the yellow band as well as for the *n*-type conductivity. According to the other model residual impurities give rise to the *n*-type conductivity. These impurities and/or Ga vacancies also give rise to the yellow band. Popovici *et al.*[28] have reviewed studies which support one or the other of these two models. In an attempt to resolve this controversy, GaN layers were grown by RMBE with different flow rates of ammonia.[28] The epilayers were characterized by SIMS and Hall measurement. The background carrier concentration decreased as the nitrogen concentration in the layers increased. This result supports the view that nitrogen vacancies are responsible for the *n*-type conduction. These experiments also show that at least in the samples grown by RMBE, Si and O impurities are not responsible for the *n*-type conductivity. The background electron concentration in the epilayers was larger than Si concentration by 1 order of magnitude. Also it was found that in some samples with larger O concentration, the electron concentration was lower. On the other hand it has been observed that the GaN epilayer containing a larger concentration of nitrogen than Ga shows the large background electron concentration.[85] This result argues against the nitrogen vacancies as being the donors. There are other experiments[163] which suggest that in the MOVPE grown GaN the *n*-type conductivity is due to the residual concentration of Si atoms. We will discuss these experiments in Sec. VII B. It is possible that the centers that give rise to yellow band and *n*-type conduction are different in samples grown by different techniques and under different experimental conditions.

B. Hall mobility and concentration of electrons in nominally undoped III–nitrides

1. AlN

In this section we discuss the electrical properties of AlN, GaN, and InN, in that order. As grown AlN always has a large concentration of C, O, and H impurities.[85] However, unlike GaN and InN, carrier concentration and mobility in AlN are very low, and undoped AlN is generally semi-insulating. The carrier concentration in undoped AlN is $<10^{16}\,\mathrm{cm}^{-3}$ and mobility is less than 2 $\mathrm{cm}^2/\mathrm{V\,s}$. However, *p*-type conduction in deliberately C-doped AlN has been observed.[184] Doping of AlN is discussed in Sec. IX B 2.

2. GaN

Nakamura *et al.*[109] made Hall effect measurements on undoped GaN layers. The undoped layers were grown by MOVPE on GaN buffer/sapphire substrates. Before the epilayers were grown, GaN buffer layers of different thicknesses were grown at lower temperatures. The layers were undoped but were *n* type due to unintentional background doping.

Both electron Hall mobility and electron concentration are found to depend on the thickness d_b of the buffer layer. As the thickness increases from 100 Å, the mobility also increases up to a thickness of 200 Å. At larger thicknesses the mobility begins to decrease. The value of the mobility was 600 $cm^2/V\,s$ at room temperature and 1500 $cm^2/V\,s$ at 77 K for a 200 Å thick layer. The electron concentration was minimum at $d_b = 200$ Å and increased monotonically as d_b increased further. The mechanism which gives rise to these effects is not clearly understood. Experimentally it is difficult to control the thickness of the buffer layer exactly at 200 Å.[109,185]

Hall mobility and electron concentration in undoped GaN have been measured by several workers (Refs. 186, 187, and references given therein). Both quantities depend not only on the thickness of the buffer layer but also on the thickness of the epilayers. Look and Molnar[186] investigated the Hall effect in the temperature range 10–400 K in HVPE grown layers on ZnO pretreated sapphire substrate. The background concentration of electrons at room temperature was a high $\sim 10^{17}\,cm^{-3}$. As the temperature decreased, the carrier concentration also decreased, became minimum at $T \sim 100$ K, and then started increasing again. The concentration became constant at $\sim 2\times 10^{17}\,cm^{-3}$ below 30 K. A model was constructed to interpret these results. The model consists of two layers: one thin degenerate *n*-type layer at the GaN, sapphire interface and the other layer with a lower conductivity above it. The degenerate layer dominates at 10 K but also gives a significant contribution at 400 K. Götz *et al.*[187] have investigated the effect of the layer thickness *d* on the Hall mobility and electron concentration in GaN. Two types of HVPE grown layers on sapphire substrates were investigated. In one type, the sapphire substrate was pretreated with ZnO and in the other type, it was pretreated with GaCl. They found that the general behavior was the same in both cases: the mobility increased and the carrier concentration decreased as the thickness *d* of the epilayers increased. A detailed modeling of the Hall effect measurements was done. The model of Look and Molnar[186] mentioned above could be fitted with the results obtained with the sample grown on the GaCl treated substrate. The results obtained with the ZnO pretreated substrate did not agree with this model. The results obtained with this sample could be interpreted with a model in which it is assumed that electron concentration decreases and mobility increases continuously as the thickness of the epilayer increases.

3. InN

Hall measurements[85] show that the as grown InN films are always *n* type with a high concentration ($>10^{20}\,cm^{-3}$) of electrons. The measured mobility and electron concentration in InN films grown at different temperatures are shown in Fig. 21. The carrier concentration is low in the layers grown at higher temperatures. This observation argues against the nitrogen vacancies being responsible for the background *n*-type conductivity. At high temperatures the nitrogen atoms should evaporate and concentration of nitrogen vacancies should increase. If vacancies were the donors, the conductivity of the high temperature grown layers should have been high. Carrier concentration was also measured as a function of TMIn flow rate. In the initial stages the carrier concentration decreased as the flow rate increased. This result supports the view that nitrogen vacancies are not the background doping. Residual oxygen or other impurities are probably responsible for the high background *n*-type doping. Results similar to those given in Fig. 21 have also been reported by Ren[188] for MOMBE grown InN at growth temperatures in the range 400–530 °C.

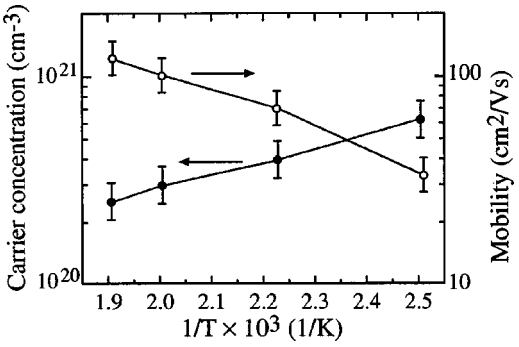

FIG. 21. Carrier concentration and mobility vs growth temperature for InN films deposited by GSMBE (Ref. 85).

C. Doping with Si and Ge during growth

Both Ge and Si substitute Ga sites in GaN and form shallow donors. Nakamura *et al.*[189] fabricated Ge- and Si-doped GaN layers using MOVPE. The layers were deposited on GaN buffer layers grown on sapphire substrates. The carrier concentration as a function of the flow rates of germane and silane is shown in Fig. 22. Figure 22 shows that incorporation efficiency of Si in GaN is very high. In Si-doped layers a high electron concentration of $2\times 10^{19}\,cm^{-3}$ was obtained. The quality of the surface was good up to the highest Si doping. The surface of the layers was like a mirror. No cracks or etch pits were observed. Crystal quality degraded slightly. The FWHM of the (0002) x-ray diffraction peak increased from 4 to 6 arcmin when the silane flow rate was changed from 0.1 nmole/min to 10 nmole/min.[189,190] The surfaces of the Ge-doped layers degraded for large electron concentration, for $n > 1\times 10^{19}\,cm^{-3}$. Etch pits were observed

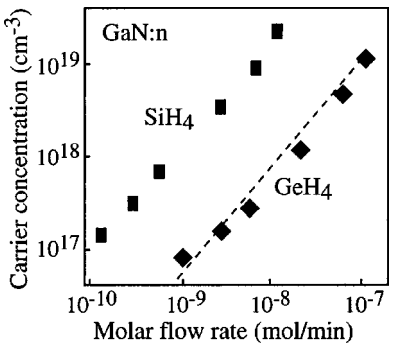

FIG. 22. Incorporation of Si and Ge in GaN in MOVPE growth of GaN using silane and germane as the precursors (Ref. 189).

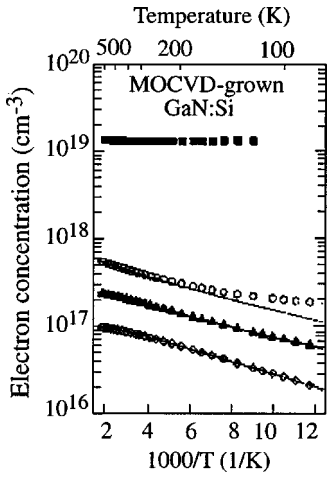

FIG. 23. Electron concentration vs reciprocal temperature for GaN doped with four different concentrations of Si. The symbols indicate the experimental results. The sample number for a given symbol is the same as given in Fig. 24. The solid lines are the least square fit of Eq. (38) (Ref. 163).

at the surfaces. The intensity of the yellow band increased with doping concentration in both Si- and Ge-doped GaN layers.

Activation energy of Si in GaN determined by Tanaka *et al.*[191] using Hall effect measurements is 30 meV. More recently, Götz *et al.*[163] have studied the Si-doped MOVPE grown GaN layers and determined the activation energies of the Si donors. One nominally undoped sample (sample 1) and four samples doped with different concentrations of Si (samples 2–5) were studied. Hall measurements were made at different temperatures in the range 70–500 K. The carrier concentration was determined by using the measured Hall constant R_H and the equation

$$n = r_H \frac{1}{R_H}, \tag{37}$$

where r_H is the Hall scattering factor. Since r_H for GaN is not known, its value was assumed to be unity. The symbols in Fig. 23 show the values of electron concentration determined in this manner in samples 2–4. To obtain the activation energies, the experimental data were fitted with the following equation:

$$n + N_{\text{comp}} = \sum_{i=1}^{M} \frac{N_{Di}}{1 + (g_1 n / N_C)\exp(\Delta E_{Di}/kT)}, \tag{38}$$

where M is the number of donors having different activation energies ΔE_{Di}, N_{Di} is the concentration of the ith donors, and N_{comp} is the concentration of background compensating impurities. The degeneracy factor g_I of the electronic states in the band gap was taken to be $=2$. The effective electron mass of $0.2m_0$ (m_0 is the free electron mass) and an effective conduction band density of states $N_C = 4.3 \times 10^{14} T^{1.5}\,\text{cm}^{-3}\,\text{K}^{-1.5}$ were used in the calculations. ΔE_{Di}, N_{Di}, and N_{comp} were used as adjustable parameters. The fit could be obtained only if two independent donor levels were used. Solid lines in Fig. 23 show the fitted values of n. The fit is good for samples 2 and 3. It is also good for sample 4 at high temperatures but there is a discrepancy between the experimental data and the fitted curve at low temperatures. Götz *et al.*[163] attributed the discrepancy to the defect band conduction in this sample. The carrier concentration in sample 5 is above the degeneracy limit and therefore it does not depend on the temperature. The values of ΔE_{Di} obtained by the best fit varied in the range 12–17 meV for donors N_{D1} and 32–37 meV for donors N_{D2}. The concentration N_{D1} was greater than N_{D2} by a factor of about 3 in samples 2 and 3 and a factor of about 10 in sample 4. Since the concentration of the donors with higher energy of activation is small, the electrical properties of GaN are determined mainly by the donors with the activation energy $E_A = 12-17$ meV. The activation energy determined from PL measurements is 22 meV (see Sec. VI D 1). We conclude that the values of the activation energy for Si determined by electrical methods are in the range 12–17 meV and determined by the optical methods, 22 meV. The ionization energies derived from the PL measurements do not suffer from these uncertainties. The energy of activation derived from the PL measurements was 22 meV. The values of the activation energy derived by the electrical measurements are somewhat lower than those derived from the PL measurements. In electrical measurements the ionization occurs by thermal activation. It is well known that optical ionization energies are higher than the thermal ionization energies in nonmetallic crystals.[192]

The thermal activation energies derived by earlier workers were higher, ~27 meV (see Ref. 163). The high values were obtained by earlier workers mainly because a simple exponential law for the ionization of donors [instead of Eq. (38)] was used to interpret the data.[163]

D. Electron mobility

1. Measurements

We have already discussed the electron mobility in undoped MOVPE grown GaN epilayers in Sec. VII B. Mobility in MBE grown undoped GaN layers has been reported recently by several workers.[119,112] Nakamura *et al.* [189] and others (see Refs. 117, 21, 163, and references given therein) have measured the room temperature mobility as a function of carrier concentration in Si- and Ge-doped GaN layers. In the experiments of Götz *et al.*[163] discussed earlier, Hall mobilities of electrons were also determined in Si-doped GaN layers. Hall mobilities of electrons in samples 2–5 are shown in Fig. 24. The high temperature mobilities are well approximated by the $\sim T^{-1.5}$ dependence on temperature. The reasons for the decrease of mobility with temperature in the low temperature range are not clear. Similarly the electron concentration dependence of the mobility is also not clearly understood. As expected, in samples 2, 4, and 5 the mobility decreases with increasing electron concentration. Sample 3 has a higher electron concentration than sample 2, but the mobility in sample 3 is not smaller. In fact the mobility in sample 3 is the highest among all the samples. It is 501 $\text{cm}^2/\text{V s}$ at 300 K and 764 $\text{cm}^2/\text{V s}$ at 160 K. Moustakas *et al.*[117] observed that when the electron concentration in their samples decreases to less than $10^{16}\,\text{cm}^{-3}$, the mobility starts to decrease Moustakas *et al.*[117] suggested that if the

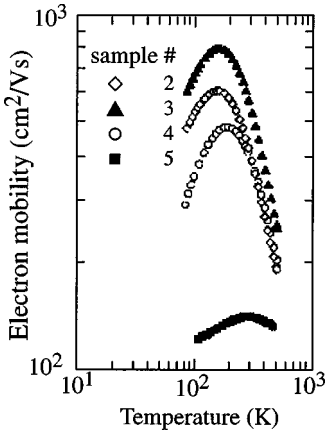

FIG. 24. Hall mobility of electrons vs temperature for GaN doped with four different concentrations of Si. The symbols indicate the experimental results. The carrier concentrations for the same samples are given in Fig. 23 (Ref. 163).

concentration of background deep level defects is more than the electron concentration, the material is fully compensated and the conduction is dominated by the hopping transport.

Nakamura *et al.*[189] measured the mobilities for electron concentration in the range $\sim 10^{16} - > 10^{19}\,\mathrm{cm}^{-3}$ (see Fig. 25). For carrier concentration $< 6\times 10^{16}\,\mathrm{cm}^{-3}$, undoped layers were used. The mobility values were practically the same, irrespective of whether Si or Ge dopants were used. As shown in Fig. 25, the room temperature mobility measured by Nakamura *et al.*[189] is 600 $\mathrm{cm}^2/\mathrm{V\,s}$ for an electron concentration of $\sim 3\times 10^{16}\,\mathrm{cm}^{-3}$, and decreases slowly with the increase in carrier concentration, reaching a value of about 100 $\mathrm{cm}^2/\mathrm{V\,s}$ at a carrier concentration of $3\times 10^{18}\,\mathrm{cm}^{-3}$. For an electron concentration of $1\times 10^{18}\,\mathrm{cm}^{-3}$, the mobility is 250 $\mathrm{cm}^2/\mathrm{V\,s}$.[189] As mentioned earlier, a high room temperature mobility of 501 $\mathrm{cm}^2/\mathrm{V\,s}$ at an electron concentration of $2-3\times 10^{17}\,\mathrm{cm}^{-3}$ was achieved by Götz *et al.*[163] This value is

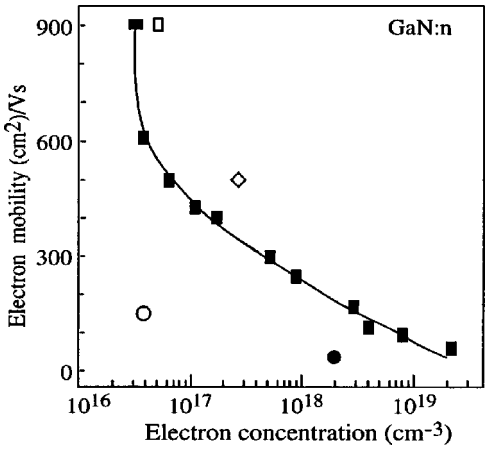

FIG. 25. The room temperature Hall mobility of electrons as a function of carrier concentration. The filled vertical rectangles are the data of Ref. 189 for MOVPE grown films. The filled circle (Ref. 119) and open circles (Ref. 112) show the values for the MBE grown undoped films. The diamond shows the mobility measured by Götz *et al.* (Ref. 163). The highest mobility, 900 $\mathrm{cm}^2/\mathrm{V\,s}$, has been reported by two groups, by Binari and Dietrich (Ref. 21) shown by the open rectangle and by Nakamura *et al.* (Ref. 185) shown by the horizontal filled rectangle.

also shown by the diamond symbol in Fig. 25. The value is larger than the value measured by Nakamura *et al.*[189] in this concentration range. The highest reported room temperature value of electron mobility is 900 $\mathrm{cm}^2/\mathrm{V\,s}$. This value has been achieved by two groups. In a later publication, Nakamura *et al.*[185] measured this value (shown by horizontal filled rectangle) at an electron concentration $=\sim 3\times 10^{16}\,\mathrm{cm}^{-3}$. Binari and Dietrich[21] were able to obtain this value at a higher electron concentration, $=5\times 10^{16}\,\mathrm{cm}^{-3}$. The value of Binari and Dietrich [21] is shown by the open rectangle in the figure. Moustakas and Molnar [117] have shown that a log–log plot of the mobility versus electron concentration is a good straight line. This line extrapolated to lower values of electron concentrations predicts a mobility of 10^4 $\mathrm{cm}^2/\mathrm{V\,s}$ at an electron concentration of $10^{14}\,\mathrm{cm}^{-3}$, which is comparable to the mobility values in GaAs. The highest published room temperature value of electron mobility in GaN is 1700 $\mathrm{cm}^2/\mathrm{V\,s}$ (not shown in the figure) in an AlGaN heterostructure.[22]

The mobility in the MBE grown layers is generally smaller. Typical values in the nominally undoped MBE grown films measured by Powell *et al.*[119] is shown by the filled circle and by Grandjean and Massies [112] by the open circle in Fig. 25.

Nakamura *et al.*[185] have also reported the electron mobility as a function of temperature. Starting from the value of 900 $\mathrm{cm}^2/\mathrm{V\,s}$ at room temperature, it increases to $\sim$3000 $\mathrm{cm}^2/\mathrm{V\,s}$ at 70 K and then decreases at lower temperatures. Above 70 K it varied as T^{-1}. We have already discussed the temperature dependence of electron mobility observed by Götz *et al.*[163] earlier in this section.

2. Monte Carlo simulations of electron mobility

An analytical multivalley conduction band model based on empirical pseudopotential band structure calculations is available for WZ GaN (Ref. 34, and references given therein). The calculations show that the lowest conduction band minimum is located at the Γ_1 point and the conduction band satellite valleys occur at the U point located 2/3 of the way between the L and M symmetry points. Higher conduction band valleys are located at the Γ_3 point, at the M point, and at the K point. The following expression gives the dispersion relations for the valleys of the two lowest conduction bands of GaN:

$$\epsilon(1+\alpha_X\epsilon+\beta_X\epsilon^2)=\frac{\hbar^2k^2}{2m_X^*},\quad \epsilon=E-E_X\,. \tag{39}$$

Here E denotes the electron energy and E_X is the minimum energy of a conduction band valley labeled X. The effective masses m_X^* and the nonparabolicity factors α_X and β_X are obtained by matching Eq. (39) with the relevant conduction band. However for the Γ valley, the experimental value of the effective mass was used. The value of β_X for the lowest Γ valley is 0.08 eV^{-2} and zero for other valleys. The values of other parameters are given in Ref. 34. Using expression (39) for the band structure, Albrecht *et al.*[34] calculated the electron mobilities μ_e by Monte Carlo simulations. The calculations were made for different concentrations of the ion-

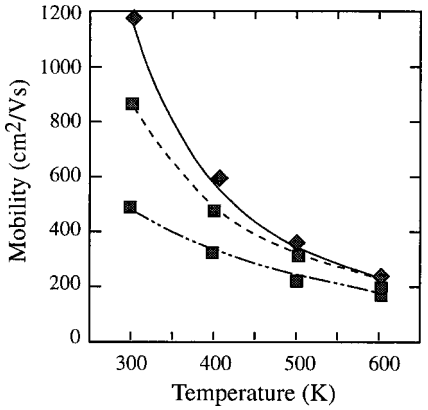

FIG. 26. Values of electron mobility in uncompensated GaN calculated using Eq. (40) are plotted as a function of temperature. Solid curve is for electron concentration $n=10^{16}\,\mathrm{cm}^{-3}$, dashed curve is for $n=10^{17}\,\mathrm{cm}^{-3}$, and dashed-dotted curve is for $n=10^{18}\,\mathrm{cm}^{-3}$. Symbols are the mobilities calculated by the Monte Carlo method (Ref. 34).

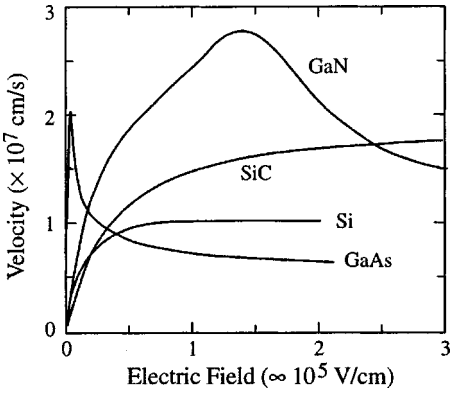

FIG. 27. Electron drift velocity as a function of electric field for Si, GaAs, 6H–SiC, and GaN (Ref. 21).

ized impurities and at different temperatures. The calculated values of the mobility are shown by symbols in Fig. 26.

For many applications it is helpful to have an analytical expression describing the dependence of the mobility on temperature and ionized impurity concentration. Albrecht *et al.*[34] have shown that the following parameterized analytical expression describes μ_e calculated by Monte Carlo simulations satisfactorily:

$$\frac{1}{\mu_e}=a\left(\frac{N_I}{10^{17}\,\mathrm{cm}^{-3}}\right)\ln(1+\beta_{\mathrm{CW}}^2)\left(\frac{T}{300\,\mathrm{K}}\right)^{-1.5}+b\left(\frac{T}{300\,\mathrm{K}}\right)^{1.5}+\frac{c}{\exp(\Theta/T)-1}, \tag{40}$$

where

$$\Theta=\frac{\hbar\omega_{\mathrm{LO}}}{k_B}=1065\,\mathrm{K},$$

$$\beta_{\mathrm{CW}}^2=3.00\left(\frac{T}{300\,\mathrm{K}}\right)^2\left(\frac{N_I}{10^{17}\,\mathrm{cm}^{-3}}\right)^{-2/3},$$

$$N_I=(1+k_c)N_D, \tag{41}$$

$$a=2.61\times10^{-4}\,\mathrm{V\,s\,cm}^{-2},$$

$$b=2.90\times10^{-4}\,\mathrm{V\,s\,cm}^{-2},$$

and

$$c=1.70\times10^{-2}\,\mathrm{V\,s\,cm}^{-2}.$$

Here N_D is the ionized donor concentration in cm^{-3}, and $k_c=N_A/N_D$ is the compensation ratio. The mobility calculated using Eq. (40) is also shown in Fig. 26. This expression gives values of mobility which agree with the results of Monte Carlo simulations (with a maximum error of 6%) in the following ranges of the parameters; temperature: $300\,\mathrm{K}<T<600\,\mathrm{K}$, ionized donor concentration: $10^{16}\,\mathrm{cm}^{-3}<N_D<10^{18}\,\mathrm{cm}^{-3}$, and compensation ratio: $0<k_c<1$. The values predicted by this equation and by Monte Carlo simulations are higher than the experimental values discussed earlier in Sec. VII C (see also Fig. 25). Theoretical values of the mobilities calculated by other workers also have similar high values.[35] It is hoped that as the technology of growth improves, the experimental mobility values will approach the calculated values.

Monte Carlo simulations of electron velocity versus electric field have been made.[36,21] The results of these simulations are shown in Fig. 27. Results for GaAs, Si, and SiC are also shown for comparison. The peak velocity in GaN is close to 3×10^7 cm/s and the saturation velocity is 1.5×10^7 cm/s. Both values are considerably higher than the values for Si and GaAs.

Consistent with the Monte Carlo predictions, recent measurement show that mobility above 1000 $\mathrm{cm}^2/\mathrm{V\,s}$ may be achieved for WZ GaN. The values in cubic GaN are expected to be 40% higher.[23] Transient Monte Carlo calculations show that even though conduction band effective mass in GaN is $0.22m_0$ as compared to $0.067m_0$ in GaAs, electrons in GaN are expected to show much larger overshoot effects, although fields required for the effect will be considerably larger.[23]

VIII. ELECTRICAL PROPERTIES: *p*-TYPE III–NITRIDES

A. Hole concentration and mobility

1. Activation of Mg and hole concentration

Doping GaN with acceptors to obtain a high concentration of holes has been a difficult problem. Attempts to dope GaN with Mg acceptors were not so successful in 1970s and early 1980s because Mg was always compensated by the background *n*-type defects or was passivated with hydrogen atoms. It was discovered in the late 1980s by Amano *et al.* (Ref. 51, and references given therein) that Mg-doped GaN can be converted into *p*-type GaN by LEEBI. They performed experiments on samples grown by MOCVD on AlN buffer layers and sapphire substrates. The layers were doped with 10^{20} Mg atoms/cm^3. The LEEBI was performed with 10 kV electrons with a beam current of 60 μA. After the LEEBI the layer resistivity decreased to 35 Ω cm. Hall measurements were made and it was found that this resistivity corresponds to a hole concentration of $2\times10^{16}\,\mathrm{cm}^{-3}$ and a mobility of $\mu=8\,\mathrm{cm}^2/\mathrm{V\,s}$. These results give an activation ratio of 2×10^{-4} of the Mg atoms, a rather low value. A higher activation ratio was obtained by Nakamura *et al.*[193] with

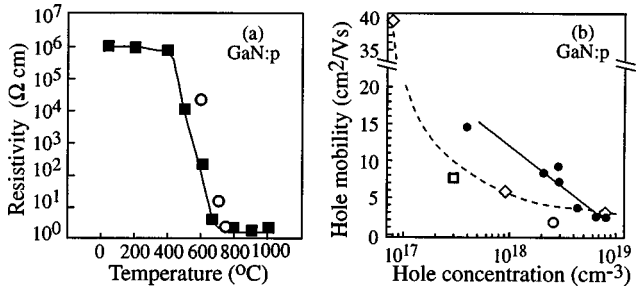

FIG. 28. (a) Effect of thermal annealing on the resistivity of GaN:Mg as a function of annealing temperature. Filled squares are from Ref. 194 and open circles are from Ref. 195. (b) Hole mobility in Mg-doped GaN as a function of hole concentration. Filled circles are from Ref. 193 for LEEBI treated GaN, open circle is from Ref. 196, and open square is from Ref. 114 for a MBE grown layer. Open diamonds show the results for ZB cubic MBE grown GaN (Ref. 197).

GaN layers deposited on GaN buffer layers. The as grown layers were *p* type with a hole concentration $=2\times10^{15}/\mathrm{cm}^3$, the hole mobility $=9\ \mathrm{cm}^2/\mathrm{V\,s}$, and the resistivity $=320\ \Omega\,\mathrm{cm}$. After LEEBI treatment, the hole concentration became $3\times10^{18}/\mathrm{cm}^3$, and the resistivity became $0.2\ \Omega\,\mathrm{cm}$. The mobility remained unaffected. In subsequent publications Nakamura *et al.*[54,194] showed that the effect of LEEBI is in fact to heat the layers and probably to dissociate hydrogen atoms which form complexes with Mg. They found that thermal annealing in vacuum or in nitrogen atmosphere also activates the Mg acceptors without the LEEBI treatment. Thermal annealing of the MOVPE grown Mg-doped layers increases the conductivity by up to 6 orders of magnitude as shown in Fig. 28(a). A hole concentration $>10^{18}\,\mathrm{cm}^{-3}$ is easily obtained at room temperature by the thermal treatment. The effect of thermal treatment is reversible. By heating the activated layers in hydrogen atmosphere, the Mg atoms are deactivated. More recently Götz *et al.*[195] have made extensive Hall measurements on Mg-doped GaN epilayers grown by MOCVD on sapphire substrates. They also activated the Mg impurity by thermal annealing and confirmed that on thermal annealing the resistivity decreases by more than 6 orders of magnitude. Open circles in Fig. 28(a) show some of the data points measured by Götz *et al.*[195]

Recent experiments by Miyachi *et al.* and others (Ref. 198, and references given therein) suggest that the presence of minority carrier electrons helps the activation process. Miyachi *et al.*[198] fabricated the GaN $p-n$ diode using Si and Mg dopants. The layers were grown by MOVPE. As deposited Mg-doped layers were insulating. Different diodes were annealed under different biasing conditions. Under forward bias of 8 V, the resistance of the Mg-doped layer begins to decrease at 300 °C, and saturates at a value five orders of magnitude smaller at 400 °C. These annealing temperatures are lower than those shown in Fig. 28(a). Since the device had metal coatings, hydrogen did not diffuse out. It was confirmed by SIMS experiments that the hydrogen profile did not change after the anneal. Mg atoms are deactivated by annealing the diodes at 350 °C. If Mg atoms are activated at higher temperatures and without the presence of electrons, as was the case in the experiments[194,195] discussed above, hydrogen ambient and higher temperatures are necessary for deactivation. IR measurements and theoretical calculations show that in Mg-doped GaN layers grown by MOVPE in hydrogen rich environments, Mg–H complexes are formed.[198] A model to explain the result of activation and deactivation of the GaN diodes in the presence of electrons can now be constructed.[198] The complexes dissociate at the low temperature of about 300 °C. H^+ ions capture electrons and become H^0 atoms. This prevents the back reaction, i.e., the formation of the Mg–H complex. However, since hydrogen stays in the layer, annealing the diode in the absence of the bias results in ionization of the hydrogen atoms, formation of the complexes at lower temperatures, and deactivation of Mg atoms. Obviously external hydrogen ambient is not necessary in this case. The behavior of the diode annealed under open circuit configuration was different. Practically no activation of Mg was observed even at 600 °C. This result is also different from the earlier result of Nakamura *et al.*[194] who observed activation of Mg at 500 °C. As mentioned earlier, because of the metal coating on both sides of the sample, hydrogen does not diffuse out. This suppresses the activation of Mg in these experiments.[198]

2. Energy needed to activate Mg in GaN

The earliest determination of the activation energy of Mg was made by Akasaki *et al.*[52] They measured the temperature dependence of the PL intensity due to donor–acceptor recombination in MOVPE grown layers. The layers were Mg doped and LEEBI treated. They obtained the ionization energy of Mg in the range 155–165 meV. Subsequently extensive measurements of conductivity and Hall effect in Mg-doped GaN were made by Moustakas and Molnar.[117] Moustakas and Molnar[117] deposited GaN films on (0001) sapphire substrates using ECR assisted MBE. These authors doped the GaN films with Mg during growth and obtained Mg concentration in the range of $10^{17}-10^{19}\,\mathrm{cm}^{-3}$. LEEBI or thermal annealing to activate the Mg atoms were not required. Values of hole concentrations were in the range $10^{13}-10^{18}\,\mathrm{cm}^{-3}$. The log(conductivity) versus $1/T$ plot consisted of two straight lines: one at high temperatures (>200 K) with large slope and the other at lower temperatures with smaller slope. The slope of the high temperature plot was analyzed to extract the energy of ionization of the Mg atoms. The interpretation of the slope is complicated because the slope depends on whether the concentration of compensated acceptors (which is equal to the background donor concentration N_D) is smaller or larger than the concentration p of free holes. If it is assumed that $N_D \gg p$, the slope yielded an activation energy of 150 meV. On the other hand, if $N_D \ll p$, the ionization energy came out to be 300 meV. Results of Akasaki *et al.*[52] mentioned above and subsequent work of Götz *et al.*[195] discussed below showed that the ionization energy is close to 150 meV.

Götz *et al.*[195] determined the hole concentration by Hall measurement and plotted the concentration as a function of $1/T$. The plot was a straight line for the sample annealed at 600 °C. They analyzed the slope of the 600 °C annealed samples using a model that assumes single acceptor species, and partial compensation by unidentified donors specified by their bulk concentration. This yielded a donor concentration

of $3.9\times10^{18}\,cm^{-3}$ and an activation energy of 182 meV. The plots for the samples annealed at 700 and 775 °C were straight lines only above 200 K. At lower temperatures the plots were curved because of the onset of impurity band conduction. The straight line portions of these plots were analyzed assuming a fixed background donor concentration of $3.9\times10^{18}\,cm^{-3}$. The activation energies obtained from these plots are 170 and 176 meV.

Activation energy of Mg in GaN determined by Tanaka *et al.*[191] by Hall effect measurements is 150 meV. Johnson *et al.*[200] have determined the ionization energy of Mg acceptors in GaN using PL measurements. Mg-doped GaN epilayers were grown using rf-assisted MBE. They observed a PL peak due to the transition from conduction band to acceptors at 3.2730 eV at 60 K. The 60 K band gap of their sample was 3.494 eV. This gave an optical ionization energy of Mg acceptors as 221 meV, and the contribution due to thermal energy was estimated to be 2.6 meV, leading to a value of 224 meV for the ionization energy. The value determined from the 2 K PL measurements by Götz *et al.*[163] is 209 meV. Both values determined by the PL method are considerably larger than the energy of ~160 meV determined from the temperature variation of conductivity and Hall effect measurements made at high temperatures. The difference in the values could arise due to the large difference in the temperatures at which the electrical and the optical measurements were made. The temperature dependence of the ionization energy is not known. The difference in the thermal and the optical ionization energies are discussed in Ref. 192 and references given therein.

In view of the difficulties with Mg as a *p*-type dopant, other dopants are being tried to obtain high *p*-type conductivity.[85,28] We have already discussed the results of Be-doped GaN layers earlier. Carbon, cadmium, and calcium are the other potential acceptors. *p*-type GaN (hole concentration $=3\times10^{17}\,cm^{-3}$) and AIN (hole concentration $=3\times10^{18}\,cm^{-3}$) have been obtained with C doping.[85] C is a donor in InN. *p*-type conductivity has been observed in Ca implanted GaN.[85] Zn forms a single acceptor if it substitutes Ga and it forms a triple acceptor with deeper energy levels if it substitutes a nitrogen. Thus Zn gives rise to four levels above the valence band at 0.57, 0.88, 1.2, and 1.72 eV.[40] A more detailed discussion of the donor and acceptor levels in GaN is given in Ref. 18.

3. Hole mobility

Nakamura *et al.*[193] have measured Hall mobility of holes as a function of carrier concentration in LEEBI treated Mg-doped GaN epilayers grown on GaN buffers. The results of these measurements are shown in Fig. 28(b). The value shown by the open circle is for a recent MOVPE grown layer.[196] Hole mobilities have also been measured in ZB cubic GaN.[197,199] Lin *et al.*[197] deposited ZB GaN epilayers on vicinal (100) GaAs substrates by ECR plasma-enhanced MBE. Mg was also injected into the ECR source using N_2 carrier gas. The growth temperature was low, 580 °C. Hole concentrations between 8×10^{16} and $8\times10^{18}\,cm^{-3}$ were obtained. The mobility at a few typical concentrations used by Lin *et al.*[197] are shown by open diamonds in Fig. 28. The highest mobility is 39 $cm^2/V\,s$. This high value is obtained because a layer with very low hole concentrations could be fabricated. At higher hole concentrations, the mobilities are lower or comparable to the mobilities obtained by Nakamura *et al.*[193] and shown by filled circles.

Brandt *et al.*[199] have codoped ZB GaN with Be and oxygen. The layers were grown on the (113) plane of GaAs by MBE. In a sample containing $10^{20}\,cm^{-3}$ Be atoms, the free hole concentration was $5\times10^{18}\,cm^{-3}$, conductivity was 50 $(\Omega\,cm)^{-1}$, and mobility was 70 $cm^2/V\,s$. In another sample, with $>5\times10^{20}\,cm^{-3}$ Be atoms and $1\times10^{18}\,cm^{-3}$ holes, a mobility of 150 $cm^2/V\,s$ was observed. These high values of hole mobilities are attributed to the formation of Be–O pairs which do not scatter the holes efficiently. The high conductivity is the direct result of the high hole mobility. For a Be concentration $>5\times10^{20}\,cm^{-3}$, the conductivity is not thermally activated, i.e., it is metallic in nature. Brandt *et al.*[199] have mentioned that it should be possible to obtain similar results in WZ GaN and also with Mg impurity if oxygen is incorporated simultaneously to form Mg–O pairs. Such high values of mobilities have not been reported in WZ GaN so far.

As discussed earlier good *p*-type Mg-doped epilayers have also been fabricated by MBE.[114,91] Using the method described in Ref. 112 and discussed earlier in Sec. II F, Grandjean and Massies[114] fabricated Si-doped (*n*-type) and Mg-doped (*p*-type) GaN epilayers. In *p*-type epilayers the Mg concentration was $\sim9\times10^{18}\,cm^{-3}$ and the hole concentration was $3\times10^{17}\,cm^{-3}$. The resistivity was 5.6 Ω cm and the hole mobility was 8 $cm^2/V\,s$. We have already mentioned earlier that it is not necessary to use LEEBI or high temperature annealing to activate the Mg acceptors in the MBE grown GaN epilayers.[117,114,91,28] The open square in Fig. 28(b) shows the value of mobility in the MBE grown layers.[114]

IX. PROPERTIES OF ALLOYS

A. In containing alloys

1. InAlN

Growth of III–nitride alloys was discussed in Secs. II F and II G. The discussion in this section is focused on the electrical and optical properties of the alloys.[201–215]

Carrier concentration in InAlN epilayers grown in hydrogen ambient have been measured.[85] If the In mole fraction is less than 30%, the epilayers show AlN like behavior. Both carrier concentration and mobility are very small. As the In concentration increases, both the carrier concentration and the mobility increase approximately linearly with In concentration up to 100% In.

2. InGaN

The band gap of $In_xGa_{1-x}N$ alloys depends on the alloy composition and also on temperature. At low temperatures (~10 K) the measured band gap is given by the following expression:[202]

$$E_g(x)=3.5-2.63x+1.02x^2,\ \text{eV}. \tag{42}$$

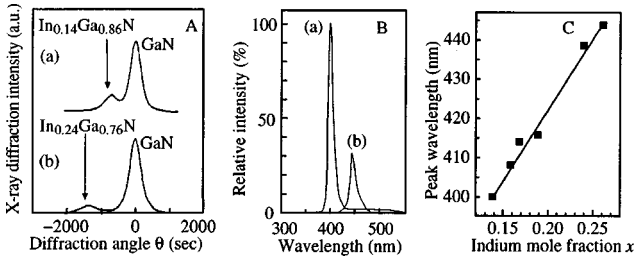

FIG. 29. (A) The x-ray rocking curve for (0002) diffraction from the InGaN films. (B) Room temperature PL spectra of the $In_xGa_{1-x}N$ films. The growth temperature was 830 °C for film (a) and 780 °C for film (b). The value of x is 0.14 for sample (a) and 0.24 for sample (b). (C) The peak wavelength of the PL spectra of $In_xGa_{1-x}N$ films as a function of In mole fraction x. The figures are taken from Ref. 203.

The band gap decreases with temperature and at 295 K it is smaller by about 0.06 eV for all values of x. Since the band gap of the InGaN alloy is smaller than that of GaN, the alloy is very suitable for making quantum well active layers in the GaN/$In_xGa_{1-x}N$/GaN double heterostructures. Nakamura and Mukai[203] have grown high quality $In_xGa_{1-x}N$ films by the two-flow MOVPE method at a growth temperature in the range 780–800 °C. TMG, TMI, and NH_3 were used as the source precursors. First a sapphire substrate was heated to 1050 °C in H_2 ambient and then the temperature was lowered and a GaN buffer was grown at 510 °C. The thickness of the InGaN films grown at ~899 °C on these buffers was approximately 2 μm. The x-ray rocking curves (XRCs) of two samples (a) and (b) are shown in Fig. 29(A). Two peaks, one due to the GaN substrate layer and the other due to the $In_xGa_{1-x}N$ layer are seen in the XRC curve for each sample. The indium mole fraction in the $In_xGa_{1-x}N$ samples was calculated from the observed separation of $In_xGa_{1-x}N$ and GaN peaks, assuming that a Vegard's type of law is valid for the relation between separation and the indium concentration. The values of x are 0.14 for sample (a) and 0.24 for sample (b). The FWHM of the curves is 6 min for the GaN film and 8 min for the $In_xGa_{1-x}N$ sample (a) and 9 min for sample (b). The small values of the FWHM show that the films were of high quality. The room temperature PL spectra for both $In_xGa_{1-x}N$ samples are shown in Fig. 29(B). The FWHM of the two PL peaks are 70 and 110 meV, respectively. The narrow widths show that the PL is due to the band-edge emission. The deep level emission at 550 nm was extremely weak. PL results support the x-ray diffraction results that the films were of very high quality. The PL peak wavelength for several different concentrations of indium are plotted in Fig. 29(C). It is seen that the emission wavelength can be tuned from 400 nm to about 445 nm by changing the alloy composition.

In a later publication,[204] Nakamura *et al.* grew Si-doped InGaN films using the same method. SiH_4 was used as a source of nitrogen. XRC curves showed that for a small concentration of Si the quality of the InGaN layers did not degrade. FWHM of the XRC curves was only 6.4–6.5 min and deep level luminescence was not observed in the PL curves. The near band-edge PL peak was narrow. The intensity of the PL peak increased by a factor of 36 on doping the films with Si, Nakamura *et al.*[204] suggested that Si-doped layers may be good candidates for use in LEDs because of this high efficiency of luminescence.

Electrical properties of the undoped InGaN films have also been reported for $\sim 0.1 < x < 1$.[85] For $x > 0.2$ the electron concentration remains high, characteristic of InN films. The carrier concentration decreases abruptly as the In concentration decreases to less than 0.2. The mobility increases monotonically and approximately linearly with In concentration in the whole range of composition x investigated.

3. Bowing parameter of InGaN

Matsuoka *et al.*[55] studied the bowing parameter of InGaN layers. Vegard's law was assumed to hold and concentration of indium was determined using the lattice constants determined by x-ray techniques. Absorption coefficients α of the InGaN films were measured as function of $h\nu$. α vs $h\nu$ plots had a long tail on the low energy side. In the high photon energy range α^2 vs $h\nu$ plots were straight lines confirming that the alloy layers had direct band gaps. The band gap values were derived from these plots for values of x up to $x=0.42$. The value for $x=0.42$ was close to the value of the band gap of 1.89 eV as we know it today. This indicates a very large value of the bowing parameter.

Recently McCluskey *et al.*[205] have studied experimentally and theoretically the bowing parameter of $In_xGa_{1-x}N$ alloys with low values of x. The 0.25 μm thick layers were grown by MOCVD on a 1 μm thick GaN layer on sapphire substrate. The layers were pseudomorphic with GaN and were strained. Absorbance for several values of x in the range $0 \leqslant x \leqslant 0.12$ was measured. Values of band gap E_g determined from the absorbance versus photon energy plots are given by the following equation:

$$E_{g,\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}}(\text{stained}) = 3.42 - 3.93x \text{ eV}, \quad (0 \leqslant x \leqslant 0.12). \tag{43}$$

The change in E_g due to strain ϵ is given by[205]

$$\Delta E_g = -9.3\epsilon \text{ eV}. \tag{44}$$

The values of strain for given values of x were calculated assuming that Vegard's law is valid. Substituting these values into Eq. (44) the strain induced shift becomes

$$\Delta E_g = 1.02x \text{ eV}. \tag{45}$$

The band gap of the relaxed layer can now be calculated by subtracting the strain induced shifts from the observed band gaps given by Eq. (43)[205]

$$E_{g,\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}}(\text{relaxed}) = 3.42 - 4.95x \text{ eV}. \tag{46}$$

The bowing parameter for the relaxed alloy layer was calculated using the equation

$$E_{g,\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}}(\text{relaxed}) = xE_{g,\mathrm{In\,N}} + (1-x)E_{g,\mathrm{GaN}} - bx(1-x). \tag{47}$$

The value of the bowing parameter was large, 3.8 eV for $x = 0.1$. First principle calculations were made to determine the band gaps of the alloys with ZB structure and it was assumed that the values for the WZ structure should be simi-

lar. The calculated values of the bowing parameter were 4.8, 3.5, and 3.0 eV for $x=0.0625$, 0.125, and 0.25, respectively. Several previous workers determined the values of bowing parameter of $In_xGa_{1-x}N$ layers (Ref. 205, and references given therein). The values obtained by previous workers were lower, only 1 eV at $x\sim0.1$. Recently Akasaki and Amano[12] have determined the band gaps of the strained $In_xGa_{1-x}N$ and $Al_xGa_{1-x}N$ layers. The layers were coherently grown on GaN and were strained. The band gap was determined from the PL measurements. Values of x in the range 0.05–0.2 were used. The average value of the bowing parameter was 3.2 for $0<x\leqslant0.12$.

As expected the bowing parameter is different for the strained and the relaxed layers. The PL and absorption techniques give different values of the band gap and of the bowing parameter. The measured values of the bowing parameter also depend on the technique of measurement of the band gap. Uncertainties arise because In concentration is not known accurately and is nonuniform (see Sec. IX D).

4. Nonuniform distribution of indium

Experiments show that In concentration in InGaN alloys is not uniformly distributed. Chichibu *et al.*[157] measured the temperature dependence (in the range 10–300 K) of the luminescence spectra of GaN and undoped and Si-doped $In_xGa_{1-x}N$ 3D alloys (layers which were sufficiently thick so that there were no 2D effects). Two values of x: 0.03 and 0.06 were used. The temperature dependence of the emission spectra in InGaN was much weaker compared to the temperature dependence in GaN. This suggests that the emission in the InGaN samples is from the localized excitons. The authors provided other evidence which shows that the excitons in InGaN are localized. The origin of the localized states are the In-rich regions due to fluctuations in the alloy composition. More direct evidence of the fluctuation of composition is provided by the experiments of Narukawa *et al.*[206] In these experiments structure analysis of $In_{0.20}Ga_{0.80}N(3\,nm)/In_{0.05}Ga_{0.95}N(6\,nm)$ was performed by using TEM, and energy dispersive x-ray microanalysis. A contrast in the light and shaded regions of the well was observed and was attributed to the differences in the In composition. It was suggested that the In-rich regions act as quantum dots and give rise to efficient radiative recombination.

Recently Behr *et al.*[207] have studied the Raman scattering from GaN and $In_{0.11}Ga_{0.89}N$ epilayers (grown by MOVPE on the c plane of sapphire substrates) at different temperatures in the range 300–900 K. In this temperature range the change in band gap of GaN with temperature is given by

$$E_g=3.56\,eV-T\times5.6\times10^{-4}\,(eV/K). \tag{48}$$

In the backscattering geometry used in the experiments and with 2.34 eV excitation, A_1 (LO) mode at 570 cm^{-1} and high energy E_2 mode at 735 cm^{-1} were observed from GaN layers. The frequency of the $In_{0.11}Ga_{0.89}N$ modes was only slightly smaller and the modes were very weak because the thickness of the $In_{0.11}Ga_{0.89}N$ layers was relatively small. These modes were not separately resolved. The E_{g1} phonon mode of the sapphire substrate was observed at 751 cm^{-1}. For excitation with 3.05 eV light the A_1 (LO) InGaN mode became stronger, but there was no change in the GaN A_1 (LO) mode. The InGaN band gap is smaller and close to the exciting energy of 3.05 eV, which gives rise to the resonant Raman scattering. At 570 K the band gap of InGaN becomes smaller and coincides with the exciting energy. Now the resonant effect and the InGaN A_1 (LO) mode become very strong. At a higher temperature of 870 K, the band gap of InGaN becomes smaller than the energy of the exciting photons and the intensity of the InGaN A_1 (LO) mode is reduced. The frequency difference $\Delta\nu=A_1(LO)-E_2$ was plotted as a function of temperature T for both GaN and InGaN epilayers. The A_1 (LO) mode of InGaN is smaller than the A_1 (LO) mode of GaN by about a wave number. Frequency of the A_1 (LO) modes in both epilayers decreases linearly with temperature at a rate of $3.1\times10^{-2}\,cm^{-1}/K$ up to $\sim500\,°C$ and $\Delta\nu$ remains small and constant. As the temperature rises further, the frequency of the A_1 (LO) mode in GaN continues decreasing at the same rate but in InGaN, the rate becomes very small, and the frequency becomes nearly independent of the temperature. The rate remains small until the temperature reaches 900 K, the highest temperature at which the measurements were made. This absence of expected decreases of the frequency of the A_1 (LO) modes in InGaN with rising temperature was attributed to the inhomogeneous distribution of In. As the temperature approaches 500 K, the band gap of the InGaN layer with large In concentration comes in to resonance with the exciting energy. At further rise of temperature, the band gap of the regions with smaller In concentration and larger band gaps comes in to resonance with the exciting energy. The temperature induced decrease of the frequency is compensated due to this effect over a large temperature range of 400 K. This shows that the inhomogeneity in the distribution of In is large.

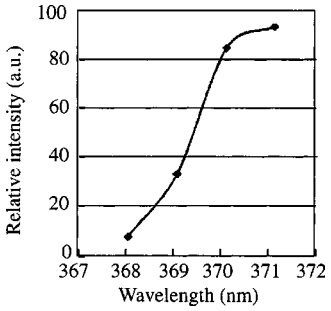

FIG. 30. The relative intensity of the UV LEDs as a function of wavelength with different In mole fractions in the active layer (Ref. 17).

Nakamura[17] has pointed out that high efficiency light emitters cannot be fabricated with GaN active layers. It is necessary to use InGaN layers in the active region to obtain high quality LEDs and laser diodes.[17] The effect of In concentration on the efficiency of the laser diodes is demonstrated in Fig. 30. The output power of the LED containing a small amount of indium in the active layer was about ten times greater than that of the diode containing no In. The In concentration was so small that the wavelength of the emitted light changed from 368 to only 371 nm. Using TEM and

energy-dispersive x-ray microanalysis, Narukawa *et al.*[206] found that In concentration is indeed nonuniform. Narukawa *et al.*[206] have modeled the regions of local high concentration of In as quantum dots. Excitons in the quantum dots give rise to radiative recombination with high efficiency.

There are experimental studies which do not support the idea that regions of high local In concentrations act as quantum dots and excitons localized in the dots give efficient luminescence.[208,209,201] Crowell *et al.*[208] fabricated a 3 nm $In_{0.22}Ga_{0.88}N$ QW on 1.8 μm GaN on a sapphire substrate and with a 1.8 μm GaN cap layer. They studied the morphology of the InGaN QW by AFM. Small pits ~20 nm in size and with a concentration of $10/\mu m^2$ were observed at the surface where threading dislocations terminated. Some larger hexagonal pits, ~400 nm in diameter and extending up to 200 nm below the surface, were also observed. The pits have been observed by several other workers and have been discussed by Nakamura.[17] The pits were also seen in the near-field transmission measurements using 3.16 laser light. Keller *et al.*[210] observed spiral-like growth on nanoscale islands at the surface of GaN where dislocations with threading components terminated[210] using AFM. The islands are formed if the rate of flow of trimethylgallium is low. Crowell *et al.*[208] obtained important information about the inhomogeneity by PL measurements by near-field optical spectroscopy. Room temperature (RT) PL consisted of two bands: one related to deep levels and extending from 2.0 to 2.5 eV, and the other at 2.95 eV attributed to band to band recombination in the InGaN QW. The deep level low energy peak saturates at higher density of excitation and is quenched at low temperatures. Near-field scanning microscopy with a 100 nm diam fiber was used to study the spatial distribution of the PL. The distance between the tip and the surface was kept constant at a value less than the wavelength of light used. At RT the intensity of the deep level band increased by a factor of about 10 whenever the tip was over a pit. The other regions were essentially dark. The 2.95 eV emission was observed from the entire area. It was slightly enhanced at the edges of the pits. In the case of II–VI and III–V quantum dots, very sharp and narrow PL lines are obtained. Crowell *et al.*[208] measured the entire PL spectrum at each position of the tip as the tip was moved across the sample. Experiments were performed at two temperatures: 50 and 295 K. The PL half width remained at 200 meV everywhere even when the carriers were excited only in the 100 nm region of the sample. No sharp PL lines characteristic of quantum dots were observed. At RT a variation of PL intensity which could be attributed to the efficient emission of light from the quantum dots formed by a local high concentration of In was not observed. The authors concluded that they did not find any evidence of PL from strongly localized carriers at least at RT.

B. AlGaN alloys

1. Optical absorption, band gap, and bowing parameter

Optical properties of the $Al_xGa_{1-x}N$ alloy were measured in late 1970s and 1980s by several groups (Ref. 211, and references given therein). The early results were conflicting. The band gap of the $Al_xGa_{1-x}N$ alloy as a function of Al molar fraction x was reported by different authors to bend upward, downward, and also remain linear. The first reliable optical measurements of the alloy epilayer measurements were made by Koide *et al.*[211] and values of the band gap were derived. The layers were grown by MOVPE on (0001) sapphire substrates with a 50 nm thick AlN buffer layer. The buffer layer was grown at 800 °C and the alloy layer at 1000 °C. The absorption coefficients α were measured in the photon energy range 3.2–4.4 eV and for $0<x<0.4$. In direct band gap semiconductors the relation between the absorption coefficient α near the band edge and the band gap is given by

$$(\alpha d)^2 = \text{constant}(h\nu - E_g), \tag{49}$$

where d is the thickness of the layers. The plots of $(\alpha d)^2$ versus the photon energy $h\nu$ were straight lines near the absorption edge as expected from Eq. (49). The band gap was determined by extrapolation of the linear portions to $\alpha^2=0$. The band gap determined in this manner could be represented by the following equation:

$$E_{g,Al_xGa_{1-x}N} = xE_{g,AlN} + (1-x)E_{g,GaN} - bx(1-x). \tag{50}$$

A value of 1.0 ± 0.3 was found for the bowing parameter b.

Recently Angerer *et al.*[212] have repeated these experiments with the plasma-induced MBE grown $Al_xGa_{1-x}N$ epilayers. They covered the whole composition range of the alloy from $x=0$ to $x=1$. They also determined the lattice constants and showed that Vegard's law is valid for this alloy to a good approximation. The absorption coefficient α as a function of photon energy $h\nu$ is shown in Fig. 31(a). The band gap values were calculated using these plots and Eq. (49). In alloys with high Al concentration the assumption of a parabolic density of states near the band edge does not remain valid and Eq. (49) does not give good results.[212] Absorption due to impurity levels and free excitons also contributes to the absorption coefficient. Therefore, the band gap values were also determined from the photon energy at which the absorption coefficient has a value 7.4

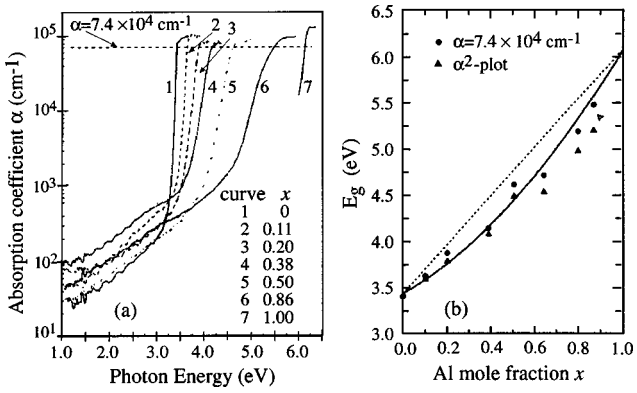

FIG. 31. (a) The room temperature absorption coefficient of $Al_xGa_{1-x}N$ as a function of photon energy. The band edge was defined at $\alpha=7.4\times10^4$ cm^{-1}. (b) Band gap of the epitaxial $Al_xGa_{1-x}N$ layers as a function of Al molar fraction x. (Solid circles) at $\alpha=7.4\times10^4$ cm^{-1}, (triangles) Eq. (49) (Ref. 212).

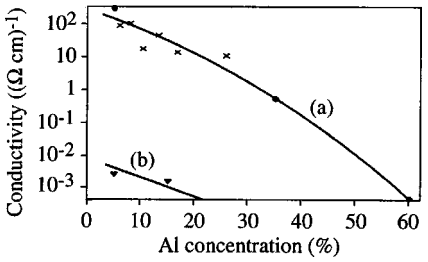

FIG. 32. Effect of Al fraction x in $Al_xGa_{1-x}N$ layers on the conductivity. (a) Grown on sapphire (crosses) and on 6H–SiC (filled circles) and doped with Si (flux 2×10^{10} atoms $cm^{-2}\,s^{-1}$). (b) Undoped GaN on 6H–SiC (triangles) (Ref. 122).

$\times10^4\,cm^{-1}$. At this value the alloys even with high Al mole fraction have a clear direct band-edge absorption.[212] The band gap as a function of Al mole fraction x is shown in Fig. 31(b). A value of 1.3 eV of the bowing parameter was determined. Akasaki and Amano[12] have determined the band gaps of the strained $Al_xGa_{1-x}N$ layers using the PL method. The bowing parameter was 0.25 for $Al_xGa_{1-x}N$ layers for values of x in the range 0–0.25.

2. Doping of the AlGaN alloys

The cladding regions of a III–nitride based laser diodes consist of $Al_xGa_{1-x}N$ layers. It is therefore important to be able to dope $Al_xGa_{1-x}N$ layers both n and p type. Stampfl and van de Walle[213] have discussed experimental work on doping these alloys and given a list of references which describe the earlier work. As Al mole fraction x increases from 0.4 to 0.6, the concentration of electrons and mobility decrease by up to 5 orders of magnitude and the epilayers become semi-insulating. Doping experiments have been performed on $Al_xGa_{1-x}N$ alloys with x varying from 0 to 0.75. For $x>0.4$, the addition of Si was not effective in decreasing the resistance of the $Al_xGa_{1-x}N$ layers.

Recently Korakakis *et al.*[122] have made extensive measurements of electrical conductivity and Hall mobility in AlGaN layers grown by ECR microwave plasma. Values of the measured conductivity of Al_xGa_{1-x} vs Al concentration are shown in Fig. 32. Figure 32 shows that the conductivity does not depend on the substrate used. The conductivity decreases rapidly with the increase of Al fraction x. Suppression of conductivity with the increase of Al in the Al_xGa_{1-x} layers has been observed in undoped and Ge-doped layers also (See Ref. 122, and references given therein). Hall effect measurements showed that both the carrier concentration and mobility of the carriers decrease as the Al fraction increases. The suppression of the mobility is due to alloy scattering. The resistivity of a GaN and two Al_xGa_{1-x} samples (for $x=0.08$ and 0.18, respectively) was also measured as a function of temperature. Below 100 K, the conduction was thermally activated. Above this temperature the conductivity became practically independent of temperature. The authors derived the approximate values of the activation energy of Si dopant atoms from these measurements. The activation energy was 17 meV for the GaN samples and 54 meV for the AlGaN samples. The activation energy for the donors increases (i.e., the donor states become deeper) as the Al fraction increases.

The difficulty in doping the $Al_xGa_{1-x}N$ layers for large values of x has been examined theoretically.[213] It was assumed that the GaN layers contained residual oxygen and Si impurities. When the Al concentration is more than 40% oxygen atoms change from being shallow donors to deep *DX* centers. The same effect arises when hydrostatic pressure is applied to GaN. At a hydrostatic pressure of about 20 GPa, the carriers in oxygen-doped n-type GaN freeze out and the layers become insulating when the hydrostatic pressure increases to 20 GPa. In this case also the band gap becomes wider and oxygen atoms are converted into deep *DX* centers. This interpretation is not valid for Si-doped GaN. Si donors do not form deep *DX* states. If Si is the residual impurity, triply charged Al vacancies, V_{Al}^{3-} are formed for $x>0.4$ in the highly n-doped $Al_xGa_{1-x}N$ layers. These defects compensate the n-type doping and reduce the conductivity when dopants are Si atoms.[213]

Attempts to dope p-type $Al_xGa_{1-x}N$ layers with Mg have not been very successful. Activation energy of Mg in $Al_{0.075}Ga_{0.925}N$ is about 180 meV, higher than the value 150 meV in GaN.[191] It has not been possible to dope the p-type layers with Mg if $x>0.13$ (see references in Ref. 213). However Akasaki and Amano have mentioned in a recent publication[12] that a hole concentration of $\sim2\times10^{17}\,cm^{-3}$ can be achieved in MOCVD $Al_{0.2}Ga_{0.8}N$ layers. Yun *et al.*[184] have reported p-type conduction in C-doped AlN, where AlN contains 11% carbon had a hole concentration of $3.8\times10^{19}\,cm^{-3}$. The concentration of holes is much smaller than the concentration of C. Yun *et al.*[184] suggested that C occupies both Ga sites and nitrogen sites. C is an acceptor at the nitrogen sites and a donor at the Ga sites. C at the Ga sites compensates the acceptor C at the nitrogen sites. The resistivity of undoped AlN was $\sim10^8\,\Omega$ cm. On doping with C, the resistivity decreased by 8 orders of magnitude of 1 Ω cm.

Some acceptors in $Al_xGa_{1-x}N$ layers from *AX* deep level centers, analogous to the *DX* centers formed by the donors.[213] However, Mg acceptors do not form the deep level centers in $Al_xGa_{1-x}N$ layers. In $Al_xGa_{1-x}N$ layers formation energy of the triply charged nitrogen vacancies V_N^{3+} is low. These defects compensate the Mg acceptors and reduce the hole concentration. Both experiments and calculations show that ionization energy of Mg increases as x in the $Al_xGa_{1-x}N$ layers increases.[213] This reduces the efficiency of p doping. Calculations also show that the formation energy of interstitial Al_I^{3+} in WZ AlGaN is large. This center is not the likely cause of reducing the efficiency of the p doping.[213]

X. QUANTUM WELLS AND SUPERLATTICES

Recently several articles have appeared on the growth of InN by MOVPE as well as by MBE. Akasaki and Amino[10] have grown the GaInN/AlN multilayered structure by MOVPE. Six pairs of InGaN and GaN layers sandwiched between AlGaN layers were grown. Akasaki and Amino[10] examined the quality of InGaN/GaN multilayers using x-ray

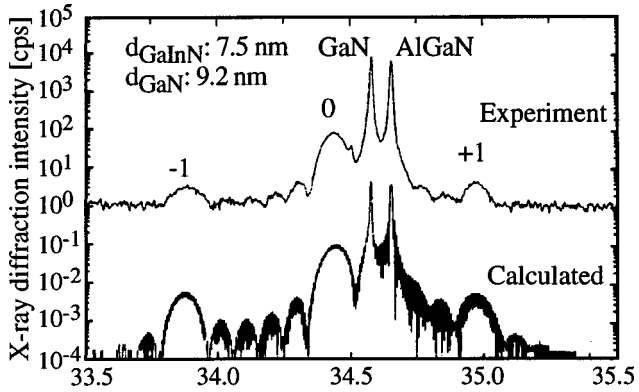

FIG. 33. High resolution x-ray diffraction ($2\theta/\theta$ scan profile) from a $(GaInN/GaN)_6$ MQW/GaN/buffer/sapphire structure (Ref. 10).

and PL techniques. A typical $2\theta/\theta$ mode x-ray diffraction profile from a GaInN/GaN multilayered structure is shown in Fig. 33. The calculated results using a kinematical theory are also shown. The agreement between the theory and the experiment is very good. There are four fingers between the zeroth and −first order peaks of the multilayer structure in both the experimental and the theoretical curves. This shows that the interfaces are smooth, and the composition and thickness are well controlled. At RT and PL intensity of InGaN layers in the multiquantum wells grown on the low temperature buffer is 2.5 orders of magnitude higher than that in the InGaN layers grown without the low temperature buffer. At 77 K, the intensity is three orders of magnitude higher.[10]

Good quality InGaN/GaN multiquantum wells have been grown by atmospheric MOVPE on the *c*-plane (0001) sapphire substrate by Minsky *et al.*[81] The substrate was annealed at 1050 °C. A 19 nm GaN nucleation layer was deposited at 600 °C. A 1.8 μm GaN layer, doped with $1\times10^{18}\,cm^{-3}$ Si atoms, was grown at 1080 °C on top of this layer. The quantum well structure was now grown. The barriers consisted of 7 nm GaN layers doped with $1\times10^{18}\,cm^{-3}$ Si atoms. The quantum wells were 2 nm layers of $In_{0.2}Ga_{0.8}N$ with background *n*-type doping of $3-5\times10^{17}\,cm^{-3}$. The structure was capped with a 100 nm $Al_{0.06}Ga_{0.94}N$ which also had an *n*-type background doping of $3-5\times10^{17}\,cm^{-3}$. Three structures were grown; with one, five, and ten quantum wells, respectively. Time resolved PL measurements were made. The decay of PL was rapid in the single quantum well. The decay rate decreased with the number of quantum wells, being slowest in the ten quantum well structure. The FWHM of the PL from the ten quantum well structure was 98 meV at room temperature and 58 meV at 4 K. This shows that the quality of the epilayers was good and interfaces were abrupt. The FWHM on the single quantum well PL was larger. Minsky *et al.*[81] suggested that the gettering of defects becomes more effective as the number of interfaces increases, and therefore the quality of the ten quantum well structure is better.

GaN quantum wells with AlGaN barriers have been fabricated and studied by Khan *et al.*[56] and by Itoh *et al.*[57] We first discuss the experiments of Khan *et al.*[56] These authors fabricated GaN and AlGaN single epilayers as well as GaN quantum wells (with AlGaN barriers). The layers were grown using low pressure (LP)MOVPE on AlN buffers, and the buffers were deposited on sapphire substrates. The source materials were triethylaluminum, triethylgallium, and high purity ammonia. The growth rate was 0.1–0.5 μm/h and the growth temperature was 850 °C. Single layers of 0.2 μm thick GaN and 0.2 μm thick $Al_{0.14}Ga_{0.86}N$ were grown. The 30 K PL spectra of the layers showed peaks at 360 nm (3.444 eV) for the GaN layer and 335 nm (3.70 eV) for the $Al_{0.14}Ga_{0.86}N$ layer. Three quantum well structures were fabricated.[56] The barrier layers of all the quantum wells consisted of 0.2 μm $Al_{0.14}Ga_{0.86}N$ layers. The thicknesses of the GaN well layers were 10, 15, and 30 nm, respectively. The PL spectra of the first two quantum wells showed two peaks: one from the $Al_{0.14}Ga_{0.86}N$ barrier layers and the other from the GaN quantum well layers. The peak from the barrier layer was at the same position, 3.70 eV, as from the single $Al_{0.14}Ga_{0.86}N$. The strain in these layers must have been negligible. The GaN PL peaks shifted to higher energies by 38.7 meV for the 30 nm well and by 43.6 meV for the 15 nm well. In the 10 nm well the PL peak was not resolved as it merged with the low energy tail of the $Al_{0.14}Ga_{0.86}N$ PL peak. To interpret the observed PL shift in the quantum wells to higher energies, change in the band gap due to confinement energy was calculated. A conduction band offset of 0.6 eV, effective electron mass of $0.2m_0$, and effective heavy hole mass of $0.6m_0$ were used in the calculations. The calculations gave a shift of 2.3 meV for the 30 nm well and 8.50 meV for the 15 nm well. These values of the confinement energy are smaller than the observed values by 36 meV for the 30 nm sample and 35 meV for the 15 nm sample. The fact that the observed PL shifts are larger than the calculated confinement energies indicates that there are other mechanisms that give rise to the widening of the band gap of the GaN well layers. There are two other factors which can modify the band gap. In the heavily doped semiconductors band tails are formed due to fluctuation of dopant concentration. The effective band gap changes due to the band tails and also due to carrier interactions and carrier–dopant interactions.[214] The other effect that can contribute to the widening of the band gap is the strain present in the GaN quantum well layers due to lattice mismatch and thermal mismatch with the buffer layer and substrate. These effects were not calculated.

In the GaN quantum wells studied by Itoh *et al.*,[57] $Al_{0.1}Ga_{0.9}N$ barrier layers were used. The samples were grown by MOVPE on (0001) sapphire substrates with AlN buffer layers. A 6 μm thick GaN layer was first grown over the AlN buffer layer. Single layers and multilayer structures were grown on the GaN layer. If the thickness of the $Al_{0.1}Ga_{0.9}N$ was more than 0.3 μm, the layers showed cracks. The $Al_{0.1}Ga_{0.9}N$ layers with smaller thickness did not show cracks and had good surface morphology.

The multilayer structures (superlattices) consisted of 23–240 periods, each period consisting of a GaN and an $Al_{0.1}Ga_{0.9}N$ layer. The thicknesses of the two layers were equal in all the samples. The thickness of the periods varied in the range 4.5–60 nm. The total thickness of the samples was in the range 1.0–1.5 μm. The XRC curves of the 20

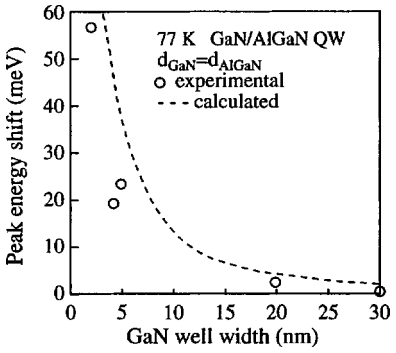

FIG. 34. Relationship between the GaN well width and the shift in the PL peak position. Symbols show the experimental points and the solid curve shows the values of the calculated confinement energy (Ref. 57).

nm/20 nm superlattice (23 periods) were measured and also simulated using dynamical x-ray diffraction theory. The GaN layers were assumed to be unstrained and the $Al_{0.1}Ga_{0.9}N$ layers were assumed to be strained due to lattice mismatch with the 6 μm thick GaN buffer. The XRC curves showed a prominent line from the GaN underlayer. In addition, the curves showed -2, 0, $+1$, and $+2$ satellite peaks arising from the superlattice. The -1 satellite peak could not be resolved as it was too close to the strong GaN peak. The FWHM of the zero order peak and the GaN peak were nearly the same and were small, 2–3 min. This indicates the high quality of the superlattices. There was good agreement between the calculated and the observed curves consistent with the high quality of the superlattice. The zero order peak position represents the average of the vertical lattice constants. Average Al concentration was calculated from the average lattice constant determined from the position of this peak. The position remained constant for different values of the periodicity. This shows that the Al concentration in the barrier layers was quite reproducible. The thickness of the layers was calculated by fitting the calculated curves with the experimental curves for each superlattice. The PL spectra of the GaN wells in the superlattices was measured. The PL peak shifted to higher energies as the thickness of the well decreased. The shifts of the PL peaks as a function of the well thickness are shown by open circles in Fig. 34. The experimental shifts were compared with the confinement energy calculated using the Kronig–Peny model. The values $0.22m_0$ and $0.8m_0$ for the effective electron and hole masses were used. The values used for the conduction and valence band discontinuities were 168 and 12 meV, respectively. The calculated values of the shifts are shown by the solid curve in Fig. 34. The authors stated that the experimental and the calculated values of the shifts are in good agreement. However, the discrepancy between the values of the shifts for well thicknesses <10 nm is very large. The effect of heavy doping and the strain were not taken into account in the calculations. There are uncertainties in the values of effective masses and band offsets used in the calculations. The effect of these uncertainties is not known at the present time.

DenBaars *et al.*[215] have studied InGaN quantum wells and have fabricated LEDs. For InGaN/GaN heterostructures they have obtained direct band-edge transitions with FWHM as narrow as 7.9 nm (59 meV) for 50 Å thick $In_{0.16}G_{0.84}N$ quantum wells at 300 K, which were among the best reported values.[215] The quantum wells PL shifted towards high energies with decreasing well thickness. The shift agreed with the calculated quantum effects.

XI. SUMMARY AND CONCLUDING REMARKS

Considerable research on GaN was done in the 1960s and early 1970s. Early investigations were made on compressed powder tablets. The main difficulty was in fabricating good quality samples. Later epilayers were grown but the quality was poor. The as grown samples had a large concentration of background donors. A large concentration of point defects and extended defects were also present. Doping with acceptors to obtain *p*-type conductivity was not successful. Ohmic contacts with low resistance could not be fabricated. In spite of these difficulties several useful results were obtained. Notable among them was the MIS short-wavelength LED emitting in the blue and green region of the spectrum.

Even after the MBE and MOVPE techniques became available, difficulties in the growth of the III–nitride epilayers persisted. The major obstacle was the lack of a suitable substrate. The large GaN wafer is not available even today. The difficulty of doping with acceptors was not solved until the late 1980s. Both these difficulties were overcome by the pioneering work of Akasaki and his group in 1988/1989. Akasaki and collaborators showed that good quality GaN epilayers can be grown on sapphire substrates by a ''two-step'' method. In the original work of Akasaki and collaborators a thin buffer layer of AlN was first grown on the sapphire substrate. The thickness of the buffer layer was around 50 nm and it was grown at a low temperature of $\sim$500 °C. Later Nakamura showed that the GaN buffer layer gives equally good or better results. The growth temperature is then raised and the main GaN layer is grown at a temperature $\sim$1050 °C on the buffer layer. Both Akasaki and Nakamura use MOVPE for the growth of the III–nitrides by the two-step method. Most other groups have also used MOVPE for fabricating and investigating these semiconductors. The development of MBE for growth of the nitrides has been slow because suitable sources of nitrogen were not available. Now rf plasma or ECR sources of nitrogen are used for the MBE growth of III–nitrides. Akasaki and collaborators were also the first to obtain good *p*-type conductivity in GaN in 1988/1989. In the as grown Mg-doped GaN *p* conduction is not observed. Akasaki and collaborators discovered that on irradiating the GaN:Mg layers by a low energy electron beam good *p*-type conductivity is obtained. Soon afterwards Nakamura showed that Mg can also be activated by high temperature annealing without LEEBI.

In the layers grown by the two-step method the density of the threading dislocations is very high, $\sim 10^{10}\,cm^{-2}$. In spite of these large density of defects, blue and green LEDs were perfected by Nakamura and commercialized in the mid 1990s. This raised an important question. GaAs based LEDs and lasers do not work at these large concentrations of dislocations. Why do the GaN LEDs work in spite of the large number of these defects? One possibility discussed by many

groups was that the dislocations in the nitrides are benign, i.e., they do not produce nonradiative recombination. Later work showed that this may not be true. TEM and CL experiments showed that dark spots in CL occur exactly at the same spots where there is a cluster of dislocations. Alternative explanations have been proposed. The dislocations are highly clustered in localized areas and large regions of the nitrides remain free of dislocations. The diffusion length of the injected carriers is not large and most of them decay by radiative recombination before they reach the dislocations. It has also been suggested that the distribution of In in the InGaN active layer is not uniform. The regions of local high In concentrations act as quantum dots and produce efficient emission of light. The carriers in the dots do not ''see'' the dislocations. The dislocations do not degrade the device performance significantly.

Although the LEDs worked well and had a long lifetime, the laser diodes did not work so well. It was identified that the heat generated during the operation of a laser diode is so large that the dislocations do degrade the diode. Therefore it became urgent to reduce the concentration of dislocations and to improve further the quality of the layers. This was achieved by using the LEO layers. The dislocation density was reduced to less than $10^7\,\mathrm{cm}^{-2}$ in the LEO layers. The recently developed new technology of growth, known as PE promises to give dislocation free layers.

Deep level defects continue to be a problem. A yellow band related to deep levels is observed at about 2.5 eV. The centers responsible for this band have not been identified. Photoconductivity in GaN persists long after the exciting light is turned off. It can persist from a few hours to several days depending on the wavelength of the exiting light and temperature of the sample. The mechanism responsible for the persistent photoconductivity is not known. It has not been possible to grow GaN layers free from background donors. The lowest concentration of electrons in good quality undoped GaN layers is generally $\sim 10^{16}\,\mathrm{cm}^{-3}$.

Considerable work has been done on the critical thickness and stress in the GaN layers. The experimental values of the critical thickness vary from a few monolayers to 30 Å. The values of stress measured by different groups do not agree. The difficulty arises because the stress is caused by both the lattice and thermal mismatch. If the layers are thick, stress relaxes at the growth temperature by generation of misfit dislocations. The stress is produced again due to thermal mismatch during the cooldown period. The actual stress is strongly dependent on the experimental conditions. The stress distribution is not uniform. Different techniques sample different volumes of the layer. Therefore the measured values of the stress depend on the techniques used.

The valance band structure of WZ GaN is different from the valence band of the ZB crystals. The degeneracy of the valence band is completely lifted by the combined action of crystal field and SO interaction. The three valence bands obtained are designated as *A*, *B*, and *C* bands. The *A* band is the topmost band, i.e., it is closest to the conduction band. The spacing between the *A* and *B* band is 6 meV and between the *B* and *C* bands it is 37 meV. The excitons involving the three bands are known as *A*, *B*, and *C* excitons. At low temperature all three excitons have been resolved in PL. The effective mass of electrons in GaN is $0.18m_0$ and that of holes lies in the range $0.6m_0-1.1m_0$. The experimental values of valence band offsets at a heterojunction between two different III–nitrides have a large scatter. They vary between 0.6 and 1.7 eV. Piezoelectric coefficients of the nitride are larger than those of conventional III–V and II–VI semiconductors. Because of the piezoelectric effect the measured values of the band offsets depend upon the thickness of the layers. The band structure of AlGaN and InGaN alloys has been investigated by several groups. The bowing parameter is large in both cases. The actual values of the band bowing parameters measured by different groups vary widely.

At low temperatures near band-edge spectra dominate in good quality GaN epilayers. Both free and bound excitons have been resolved. The yellow PL band mentioned earlier is almost always observed. DAP bands are also observed. A phonon replica of the low energy band is observed. The energy of the *A*, *B*, and *C* excitons has been determined using the PL peak positions. The experimental values agree well with the calculated values quoted above. PL of doped GaN has also been studied. Extensive studies of PL in GaN quantum wells have been made. The PL peak moves to high energies as the well width decreases. A part of this shift is due to the quantum confinement effects. Exact calculation of the observed shift is difficult because the shift depends on band filling, strain, and the quantum confined Stark effect. The quantum confined Stark effect arises due to the combined effect of the strain and large piezoelectric coefficients in GaN.

Extensive work has been done on the electrical properties. Si and Ge are shallow donors in GaN. Large concentrations of Si can be built into GaN without degrading the surface or the crystal quality. The ionization energy of Si in GaN is less than 20 meV. Mg is the most commonly used acceptor in the three nitrides. Immediately after doping with Cd, the layers do not show the *p*-type conductivity. Either LEEBI or high temperature annealing are required to activate Mg. The ionization energy of Mg in GaN is large, about 150 meV. Therefore at room temperature only a mall fraction of Mg is thermally activated. The large concentration of unionized Mg acceptors scatter holes and reduce the hole mobility to small values. The electron mobility in GaN varies from 100 $\mathrm{cm}^2/\mathrm{V\,s}$ when electron concentration is $\sim 10^{20}\,\mathrm{cm}^{-3}$ to 900 $\mathrm{cm}^2/\mathrm{V\,s}$ when electron concentration is $\sim 10^{17}\,\mathrm{cm}^{-3}$. These experimental values of mobility are smaller than the values calculated by Monte Carlo simulations for uncompensated GaN epilayers. Monte Carlo simulations also show that the peak and saturation velocities in GaN are considerably larger than in Si or GaAs. The optical phonon energy is high, 90 meV. The scattering of charge carriers by optical phonons is not so efficient in GaN. The hole mobility is considerably smaller. The hole mobility in GaN is 3 $\mathrm{cm}^2/\mathrm{V\,s}$ when hole concentration is $\sim 10^{18}\,\mathrm{cm}^{-3}$. Doping of AlGaN is more difficult. Doping efficiency of donors and acceptors decreases drastically as the concentration of Al and band gap of AlGaN increase. The ionization energy of Si in AlGaN increase from 20 to 54 meV for Al concentration of about 20%. No *p*-type conductivity in

AlGaN has been achieved for more than 13% Al. Fortunately only a small concentration of Al is required in cladding layers in device structures.

Good ohmic contacts on *n*-type GaN can now be made but contacts on *p*-type GaN continue to be a problem. The contacts are high resistance leaky Schottky barriers. In spite of these difficulties, spectacular device performance has been obtained during the last few years. Blue and green LEDs have been commercialized and can be bought off-the-shelf. Laser diodes are likely to be commercialized soon. High frequency and high temperature FETs have been fabricated. The FETs have shown very promising results.

There are several areas where further research is required. MOCVD has been used most extensively to grow nitride epilayers for both research and fabrication of devices. Improvements in MBE growth technique are required. It is yet to be determined whether other methods of growth, e.g., laser deposition will become important. So far sapphire substrates have been used most extensively. Recent work shows that at least for transistors, SiC substrates may prove better. Further work is necessary to understand the nature of centers which give rise to the yellow band, the persistent photoconductivity, and the background donors. Efforts should be made to grow large GaN crystals. If GaN substrates become available, the quality of the epilayers can be improved. Point defects and dislocation density will decrease and there will be an overall improvement in the device performance. There is an urgent need to find acceptors with a lower activation energy to improve *p* doping and hole mobility. If heavier *p* doping becomes possible, it will help in solving the problem of making ohmic contacts to the *p*-type nitrides. If these problems are solved, many more devices will be commercialized. Scientists working in this area are looking forward to new exciting developments in the next few years.

[1] F. A. Ponce, *III–V Nitrides*, edited by T. D. Moustakas, I. Akasaki, and B. A. Monemar, Mater. Res. Soc. Symp. Proc. **449** (1997).
[2] The Second International Conference on Nitride Semiconductors, Tokushima, Japan, October 27–31, 1997.
[3] Mater. Res. Soc. Symp. Proc. **482**, 875 (1998).
[4] Several interesting papers on III–nitride materials and devices are contained in the February 1997 special issue of MRS Bull. **22**, (1997).
[5] S. Nakamura and G. Fasol, *The Blue Laser Diodes* (Springer, Heidelberg, 1997).
[6] S. C. Jain, M. Willander, and R. Van Overstraeten, *Compound Semiconductor Strained Layers and Devices* (Kluwer Academic, Dordrecht, to be published).
[7] *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997).
[8] *Group III Nitride Semiconductor Compounds*, edited by B. Gil (Clarendon, Oxford 1998).
[9] *GaN*, edited by J. I. Pankove and T. D. Moustakas (Academic, New York, 1998), Vol. 1.
[10] I. Akasaki and H. Amano, Tech. Dig. Int. Electron Devices Meet. **96**, 231 (1996).
[11] I. Akasaki, H. Amano, and I. Suemune, Inst. Phys. Conf. Ser. **142**, 7 (1996).
[12] I. Akasaki and H. Amano, in *GaN*, edited by J. I. Pankove and T. D. Moustakas (Academic, New York, 1998), Vol. 1, pp 459–72.
[13] S. Nakamura, MRS Bull. **22**, 29 (1997).
[14] S. Nakamura, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 471–507.
[15] S. Nakamura, in ''Applications of LEDs and LDs'' in *GaN*, edited by J. Pankove and T. D. Moustakas (Academic, New York, 1998), Vol. 1.
[16] S. Nakamura, Sel. Top. Quantum Electron. **3**, 712 (1997).
[17] S. Nakamura, Science **281**, 956 (1998).
[18] S. J. Pearton, J. C. Zolper, R. J. Shul, and F. Ren, J. Appl. Phys. **86**, 1 (1999).
[19] S. J. Pearton, F. Ren, J. C. Zolper, and R. J. Shul, Mater. Res. Soc. Symp. Proc. **482**, 961 (1998).
[20] S. C. Binari, ''GaN FETs for microwave and high temperature applications,'' Proceedings Topical Workshop on III–V Nitrides, 1995.
[21] S. C. Binari and H. C. Dietrich, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 509–534.
[22] M. A. Khan, Q. Chen, J. W. Yang, and C. J. Sun, Inst. Phys. Conf. Ser. **142**, 985 (1995).
[23] M. Asif Khan, Q. Chen, J. Yang, M. Z. Anwar, M. Blasingame, and M. S. Shur, Tech. Dig. Int. Electron Devices Meet. **96**, 27 (1996).
[24] M. S. Shur and M. A. Khan, MRS Bull. **22**, 44 (1997).
[25] J.-Y. Duboz and M. A. Khan, in Ref. 8, pp. 343–90.
[26] W. Kim, Ö. Aktas, A. E. Botchkarev, A. Salvador, S. N. Mohammad, and H. Morkoç, J. Appl. Phys. **79**, 7657 (1996).
[27] S. N. Mohammad, W. Kim, A. Salvador, and H. Morkoç, MRS Bull. **22**, 22 (1997).
[28] G. Popovici, H. Morkoç, and S. Noor Mohammad, in Ref. 8, pp. 19–69.
[29] H. Morkoç, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, J. Appl. Phys. **76**, 1363 (1994).
[30] S. Strite and H. Morkoç, J. Vac. Sci. Technol. B **10**, 1237 (1992).
[31] G. Landwehr, A. Waag, F. Fischer, H.-J. Lugauer, and K. Schüll, Physica E **3**, 158 (1998).
[32] H. Morkoç, B. Sverdlov, and G.-B. Gao, Proc. IEEE **81**, 493 (1993).
[33] L. Pfeiffer, K. W. West, H. L. Stormer, and K. W. Baldwin, Appl. Phys. Lett. **55**, 1888 (1989).
[34] J. D. Albrecht, R. P. Wang, P. P. Ruden, M. Farahmand and K. F. Brennan, J. Appl. Phys. **83**, 4777 (1996).
[35] T. L. Tansley, E. M. Goldys, M. Goldewski, B. Zhou, and H. Y. Zou, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 233–94.
[36] B. Gelmont, K. Kim, and M. Shur, J. Appl. Phys. **74**, 1818 (1993).
[37] J. C. Tan, J. S. Williams, J. Zou, D. J. H. Cockayne, S. J. Pearton, and J. C. Zoppler, Appl. Phys. Lett. **72**, 1190 (1998).
[38] J. C. Zolper, R. M. Biefeld, S. B. Van Deusen, W. R. Wampler, D. J. Reiger, S. J. Pearton, J. S. Williams, H. H. Tan, R. F. Karlicek, Jr., and R. A. Stall, J. Electron. Mater. **27**, 179 (1998).
[39] X. A. Cao, C. R. Abernathy, R. K. Singh, S. J. Pearton, M. Fu, V. Sarvepalli, J. A. Sekhar, J. C. Zoppler, D. J. Rigger, J. Ham, T. J. Drummond, R. J. Shul, and R. G. Wilson, Appl. Phys. Lett. **73**, 229 (1998).
[40] J. I. Pankove, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 1–9.
[41] B. Monemar, The Second International Conference on Nitride Semiconductors, Tokushima, Japan, October, 27–31 1997, pp. 6–8.
[42] R. Juza and H. Hahn, Z. Anorg. Allg. Chem. **234**, 282 (1938), **244**, 133 (1940).
[43] H. Grimmeiss and Z. H- Koelmans, Nature (London) **14a**, 264 (1959).
[44] H. P. Maruska and J. J. Tietjen, Appl. Phys. Lett. **15**, 327 (1969).
[45] J. I. Pankove, E. A. Miller, and J. E. Berkeyheiser, J. Lumin. **5**, 84 (1972); J. I. Pankove, E. A. Miller, and J. E. Berkeyheiser, RCA Rev. **32**, 383 (1971); J. I. Pankove, M. T. Duffy, E. A. Miller, and J. E. Berkeyheiser, J. Lumin. **8**, 89 (1972).
[46] H. P. Maruska, W. C. Rhines, and D. A. Stevenson, Mater. Res. Bull. **7**, 777 (1972).
[47] J. I. Pankove and H. E. P. Schade, Appl. Phys. Lett. **25**, 53 (1974).
[48] M. T. Duffy, C. C. Wang, G. D. O'Clock, S. H. McFarlane, and P. Z. Zanzucchi, J. Electron. Mater. **2**, 359 (1973).
[49] S. Yoshida, S. Misawa, and S. Gonda, Appl. Phys. Lett. **42**, 427 (1983).
[50] H. Amano, I. Akasaki, T. Kozawa, K. Hiramatsu, N. Sawak, K. Ikeda, and Y. Ishi, J. Lumin. **40–41**, 121 (1988)
[51] H. Amano, N. Sawaki, I. Akasaki, and Y. Toyoda, Appl. Phys. Lett. **48**, 353 (1986); H. Amano, I. Akasaki, T. Kozawa, K. Hiramatsu, N. Sawaki, K. Ikeda, and Y. Ishii, J. Lumin. **40&41**, 121 (1988); H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **28**, L2112 (1989).
[52] I. Akasaki, H. Amano, M. Kito, and K. Hiramatsu, J. Lumin. **48&49**, 666 (1991).
[53] J. A. Vechten, J. D. Zook, and R. D. Horning, Jpn. J. Appl. Phys., Part 1 **31**, 3662 (1992).
[54] S. Nakamura, N. Iwasa, M. Senoh, and T. Mukai, Jpn. J. Appl. Phys., Part 1 **31**, 1258 (1992).

[55] T. Matsuoka, H. Tanaka, T. Sasaki, and A. Katsui, Inst. Phys. Conf. Ser. **106**, 141 (1990).
[56] M. A. Khan, R. A. Skogman, and J. M. Van Hove, Appl. Phys. Lett. **56**, 1257 (1990).
[57] K. Itoh, T. Kawamoto, H. Amano, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys., Part 1 **30**, 1924 (1991).
[58] S. Nakamura, T. Mukai, and M. Senoh, Jpn. J. Appl. Phys., Part 2 **32**, L8 (1993).
[59] S. S. Lester, F. A. Ponce, M. G. Craford, and D. A. Steigerwald, Appl. Phys. Lett. **66**, 1249 (1995).
[60] R. Dingle, K. L. Shaklee, R. F. Leheny, and R. B. Zetterstrom, Appl. Phys. Lett. **19**, 5 (1971).
[61] H. Amano, T. Asahi, and I. Akasaki, Jpn. J. Appl. Phys., Part 2 **29**, L 205 (1990).
[62] M. A. Khan, D. T. Olsen, J. M. Van Hove, and J. N. Kuznia, Appl. Phys. Lett. **58**, 1515 (1990).
[63] S. T. Kim, H. Amano, I. Akasaki, and N. Koide, Appl. Phys. Lett. **64**, 1535 (1994).
[64] I. Akasaki, H. Amano, S. Sota, S. Saki, T. Tanaka, and M. Koike, Jpn. J. Appl. Phys., Part 2 **34**, L1517 (1995).
[65] K. Kim, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **50**, 1502 (1994).
[66] T. J. Schmidt, X. H. Yang, W. Shan, J. J. Song, A. Salvador, W. Kim, O. Aktas, A. Botchkarev, and H. Morkoç, Appl. Phys. Lett. **68**, 1820 (1996).
[67] J. J. Song and W. Shan, in Ref. 8, pp. 182–241.
[68] I. Akasaki, S. Sota, T. Tanaka, M. Koike, and H. Amano, Electron. Lett. **32**, 1105 (1996).
[69] M. A. Khan, J. N. Kuznia, A. R. Bhattarai, and D. T. Olsen, Appl. Phys. Lett. **62**, 1786 (1993).
[70] J. I. Pankove, S. S. Chang, H. C. Lee, R. Molnar, T. D. Moustakas, and B. Van Zeghbroeck, Tech. Dig. Int. Electron Devices Meet. **94**, 389 (1994).
[71] J. M. Van Hove, R. Hickman, J. J. Klaassen, P. P. Chow, and P. P. Ruden, Appl. Phys. Lett. **70**, 2282 (1997).
[72] Q. Chen *et al.*, Appl. Phys. Lett. **70**, 2277 (1997).
[73] M. B. Nardelli, K. Rapcewicz, E. L. Briggs, C. Bungaro, and J. Bernholc, Mater. Res. Soc. Symp. Proc. **449**, 893 (1997).
[74] S. Limpijumnong, W. R. L. Lambrecht, B. Segall, and K. Kim, Mater. Res. Soc. Symp. Proc. **449**, 905 (1997).
[75] J. W. Ager III, T. Suski, S. Ruvimov, J. Krueger, G. Conti, E. R. Weber, M. D. Bremser, R. Davis, and C. P. Kuo, Mater. Res. Soc. Symp. Proc. **449**, 775 (1997).
[76] S. C. Jain, K. Pinardi, M. Willander, A. Atkinson, H. E. Maes, and R. Van Overstraeten, Semicond. Sci. Technol. **12**, 1507 (1997).
[77] J. A. Majewski, M. Städele, and P. Vogl, Mater. Res. Soc. Symp. Proc. **449**, 917 (1997).
[78] S. D. Hersee, J. C. Ramer, and K. J. Malloy, MRS Bull. July, Vol. 22, 45 (1997).
[79] S. Nakamura, Jpn. J. Appl. Phys., Part 1 **30**, 1620 (1991).
[80] D. Walker, X. Zhang, A. Saxler, P. Kung, and M. Razeghi, Appl. Phys. Lett. **70**, 949 (1997).
[81] M. S. Minsky, S. B. Fleischer, A. C. Abare, J. E. Bowers, and E. L. Hu, Appl. Phys. Lett. **72**, 1066 (1998).
[82] C. J. Sun, J. W. Yang, B. W. Lim, Q. Chen, M. Zubair Anwar, and M. Asif Khan, Appl. Phys. Lett. **70**, 1444 (1997).
[83] H. Lee and J. S. Harris, J. Cryst. Growth **169**, 689 (1996).
[84] R. D. Vispute, J. Narayan, H. Wu, and K. Jagannadham, J. Appl. Phys. **77**, 4724 (1995).
[85] C. R. Abernathy, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 11–51.
[86] F. A. Ponce, MRS Bull. **22**, 51 (1997).
[87] J. Narayan, P. Tiwari, X. Chen, R. Chowdhury, and T. Zheleva, Appl. Phys. Lett. **61**, 1290 (1990).
[88] J. Narayan, U. S. Patent No. 5406123 (11 April 1995).
[89] K. Dovidenko, S. Oktyabrsky, J. Narayan, and M. Razeghi, J. Appl. Phys. **79**, 2439 (1996).
[90] J. Narayan, K. Dovidenko, A. Sharma, and S. Oktyabrsky, J. Appl. Phys. **84**, 2597 (1998).
[91] T. C. Fu, N. Newman, E. Jones, J. S. Chan, X. Liu, M. D. Rubin, N. W. Cheung, and E. R. Weber, J. Electron. Mater. **24**, 249 (1995).
[92] A. V. Blant, T. S. Cheng, C. T. Foxon, J. C. Bussey, S. V. Novikov, and V. V. Tret'yakov, *Studies of Group III-Nitride Growth on Silicon, in III–V Nitrides*, Symposium 449 (Mater. Res. Soc. Symp. Proc. Pittsburgh, PA, 1997), pp. 465–9.
[93] S. C. Binari, J. M. Redwing, and W. Kruppa, Electron. Lett. **33**, 242 (1997).
[94] K. A. Brown, S. A. Ustin, L. Lauhon, and W. Ho, J. Appl. Phys. **79**, 7667 (1996).
[95] S. Nakamura *et al.*, Appl. Phys. Lett. **72**, 211 (1998).
[96] T. S. Zheleva, O.-H. Nam, M. D. Bremser, and R. F. Davis, Appl. Phys. Lett. **71**, 2472 (1997).
[97] J. A. Freitas, Jr., O.-H. Nam, and R. F. Davis, Appl. Phys. Lett. **72**, 2990 (1998).
[98] A. Sakai, H. Sunakawa, and A. Usui, Appl. Phys. Lett. **71**, 2259 (1997).
[99] T. S. Zheleva, S. A. Smith, D. B. Thomson, K. J. Linthicum, and R. F. Davis (unpublished).
[100] J. Narayan (unpublished work).
[101] C. Sasaoka, H. Sunakawa, A. Kimura, M. Nido, A. Usui, and A. Sakai, J. Cryst. Growth **189/190**, 61 (1998).
[102] O. H. Nam, T. S. Zheleva, D. B. Thomson, and R. F. Davis, Mater. Res. Soc. Symp. Proc. **482**, 301 (1998).
[103] A. Nakadaira and H. Tanaka, Appl. Phys. Lett. **70**, 2720 (1997).
[104] T. Beierlein, S. Strite, A. Dommann, and D. D. J. Smith, MRS Internet J. Nitride Semicond. Res. **2**, 29 (1997), Universal Resource Locator http://nsr.mij.mrs.org/2/29.
[105] K. Iwata, H. Asahi, K. Asahi, R. Kuroiwa, and S. Gonda, Jpn. J. Appl. Phys., Part 2 **36**, L 661 (1997).
[106] R. F. Davis, M. J. Paisley, Z. Sitar, D. J. Kester, K. S. Ailey, K. Linthicum, L. B. Rowland, S. Tanaka, and R. S. Kern, J. Cryst. Growth **178**, 87 (1997).
[107] R. D. Vispute, H. Wu, and J. Narayan, Appl. Phys. Lett. **67**, 1549 (1995).
[108] S. C. Jain, M. Willander, and H. Maes, Semicond. Sci. Technol. **11**, 641 (1996); Erratum **11**, 975 (1996).
[109] S. Nakamura, Jpn. J. Appl. Phys., Part 2 **30**, L 1705 (1991).
[110] T. Kachi, K. Tomita, K. Itoh, and H. Tadano, Appl. Phys. Lett. **72**, 704 (1998).
[111] M. Mesrine, N. Grandjean, and J. Massies, Appl. Phys. Lett. **72**, 350 (1998).
[112] N. Grandjean and J. Massies, Appl. Phys. Lett. **71**, 1816 (1997).
[113] N. Grandjean and J. Massies, Appl. Phys. Lett. **72**, 1078 (1998).
[114] N. Grandjean and J. Massies, Appl. Phys. Lett. **72**, 82 (1998).
[115] R. C. Powell, N.-E. Lee, and J. E. Greene, Appl. Phys. Lett. **60**, 2505 (1992).
[116] S. Yoshida, J. Appl. Phys. **81**, 7966 (1997).
[117] T. D. Moustakas and R. J. Molnar, Mater. Res. Soc. Symp. Proc. **281**, 253 (1993).
[118] J. C. Zolper and R. J. Shul, MRS Bull. **22**, 36 (1997).
[119] R. C. Powell, N.-E. Lee, Y.-W. Kim, and J. E. Greene, J. Appl. Phys. **73**, 189 (1993).
[120] A. Rizzi, *6H-SiC(0001)/AlN/GaN Epitaxial Heterojunctions and Their Valence Band Offsets*, Nato ASI series, Heterostructure Epitaxy and Devices, edited by Peter Korols (NATO, to be published).
[121] A. Rizzi, D. Freundt, D. Holz, and H. Lüth, The Second International Conference on Nitride Semiconductors, Tokushima, Japan, October 27–31, 1997, paper Th1-1, pp. 416–7.
[122] D. Korakakis, H. M. Ng, K. F. Ludwig, Jr., and T. D. Moustakas, Mater. Res. Soc. Symp. Proc. **449**, 233 (1997).
[123] B. A. Ferguson and C. B. Mullins, J. Cryst. Growth **178**, 134 (1997).
[124] J. G. Nelson, J. Vac. Sci. Technol. A **5**, 2140 (1987).
[125] F. Briones, I. González, and A. Ruiz, Appl. Phys. A: Solids Surf. **49**, 729 (1989).
[126] W. J. Meng, J. A. Sell, T. A. Parry, L. E. Rehn, and P. M. Baldo, J. Appl. Phys. **75**, 3446 (1994).
[127] R. Singh and T. D. Moustakas, *Growth of InGaN films by MBE at the Growth Temperature of GaN, Gallium Nitride and Related Materials*, First International Symposium (Mater. Res. Soc. Symp. Proc. Pittsburgh, PA, 1996), pp. 163–8.
[128] S. C. Jain, *Germanium-Silicon Strained Layers and Heterostructures*, Advances in Electronics and Electron Physics series, Supplement 24 (Academic, Boston, 1994).
[129] R. Singh, W. D. Herzog, D. Doppalapudi, M. S. Unlu, B. B. Gldberg, and T. D. Moustakas, Mater. Res. Soc. Symp. Proc. **449**, 185 (1997).
[130] W. Knap *et al.*, Appl. Phys. Lett. **70**, 2123 (1997).
[131] S. C. Jain, A. H. Harker, and R. A. Cowley, Philos. Mag. A **75**, 1461 (1997).
[132] J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **68**, 2879 (1996).
[133] G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, Appl. Phys. Lett. **68**, 2541 (1996).

[134] K. Pinardi, U. Jain, S. C. Jain, H. E. Maes, R. Van Overstraeten, and M. Willander, J. Appl. Phys. **83**, 4724 (1998).
[135] J. P. Hirth and J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968).
[136] F. Louchet, D. Cochet Muchy, and Y. Brechet, Philos. Mag. A **57**, 327 (1988); F. Louchet, Inst. Phys. Conf. Ser. **60**, 35 (1981).
[137] K. Sumino and I. Yonenaga, Solid State Phenom. **19&20**, 295 (1991).
[138] L. Sugiura, J. Nishio, M. Onomura, and S. Nunoue, Mater. Res. Soc. Symp. Proc. **482**, 1157 (1998).
[139] L. Sugiura, J. Appl. Phys. **81**, 1633 (1998).
[140] T. Kozawa, T. Kachi, H. Kano, H. Nagasw, N. Koide, and K. Manabe, J. Appl. Phys. **77**, 4389 (1995).
[141] S. F. Yoon, X. B. Li, and M. Y. Kong, J. Cryst. Growth **180**, 27 (1997).
[142] H. Siegle, A. Hoffmann, L. Eckey, C. Thomsen, T. Detchprohm, K. Hiramatsu, T. Davis, and J. W. Steeds, Mater. Res. Soc. Symp. Proc. **449**, 677 (1997).
[143] K. Hiramatsu and I. Akasaki, Jpn. J. Appl. Phys., Part 1 **32**, 1528 (1993).
[144] S. C. Jain, H. Maes, K. Pinardi, and I. De Wolf, J. Appl. Phys. **79**, 8145 (1996).
[145] S. C. Jain, H. E. Maes, and R. Van Overstraeten, Curr. Opin. Solid State Mater. Sci. **2**, 722 (1997).
[146] S. C. Jain, K. Pinardi, H. E. Maes, R. Van Overstraeten, M. Willander, and A. Atkinson, Mater. Res. Soc. Symp. Proc. **482**, 875 (1998).
[147] N. Lovergine, L. Liaci, J. D. Ganiere, G. Leo, A. V. Drigo, F. Romanato, A. M. Mancini, and L. Vasanelli, J. Appl. Phys. **78**, 229 (1995).
[148] K. Dovidenko, S. Oktyabrsky, and J. Narayan, J. Appl. Phys. **82**, 4296 (1997).
[149] K. Dovidenko, S. Oktyabrsky, and J. Narayan, J. Mater. Res. (in press).
[150] T. Sugahara *et al.*, Jpn. J. Appl. Phys., Part 2 **37**, L398 (1998).
[151] S. Ruvimov *et al.*, Mater. Res. Soc. Symp. Proc. **468**, 287 (1997).
[152] P. Vennegues, B. Beamont, M. Vaille, and P. Gibbart, J. Cryst. Growth **173**, 249 (1997).
[153] J. L. Rouviere *et al.*, J. Nitride Semicond. Res. **1**, 33 (1996), Universal Resource Locator http://nsr.mij.mrs.org/1/33.
[154] J. L. Rouviere *et al.*, Mater. Sci. Eng., B **43**, 161 (1997).
[155] B. K. Meyer, A. Hoffmann, and P. Thurian, in Ref. 8, pp. 242–306.
[156] S. J. Rosner, E. C. Carr, M. J. Ludowise, G. Girolami, and H. Erikson, Appl. Phys. Lett. **70**, 420 (1997).
[157] S. Chichibu, T. Azuhata, T. Sota, and S. Nakamura, Appl. Phys. Lett. **70**, 2822 (1997).
[158] P. De Pauw, R. Mertens, R. Van Overstraeten, and S. C. Jain, Solid-State Electron. **25**, 67 (1984).
[159] K. Shahzad, D. J. Olego, and C. G. Van de Walle, Phys. Rev. B **38**, 1417 (1988).
[160] M. Suzuki and T. Uenoyama, in Ref. 8, pp. 307–342.
[161] G. D. Chen, M. Smith, J. Y. Lin, H. X. Jiang, S.-H. Wei, M. A. Khan, and C. J. Sun, Appl. Phys. Lett. **68**, 2784 (1996).
[162] S.-H. Wei and A. Zunger, Appl. Phys. Lett. **69**, 2710 (1996).
[163] W. Götz, N. M. Johnson, C. Chen, H. Liu, C. Kuo, and W. Imler, Appl. Phys. Lett. **68**, 3144–6 (1996).
[164] W. Shan, T. J. Schmidt, X. H. Yang, S. J. Hwang, J. J. Song, and B. Goldenberg, Appl. Phys. Lett. **66**, 985 (1995).
[165] B. Monemar, J. P. Bergman, and I. A. Buyanova, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 85–139.
[166] V. A. Savestenko and A. U. Shelag, Phys. Status Solidi A **48**, K135 (1978).
[167] A. Polian, M. Grimsditch, and I. Grzegony, J. Appl. Phys. **79**, 3343 (1996).
[168] J. Feldmann, G. Peter, E. O. Göbel, P. Dawson, K. Moore, C. Foxon, and R. J. Elliott, Phys. Rev. Lett. **59**, 2337 (1987).
[169] J. Martinez-Pastor, A. Vinattieri, L. Carraresi, M. Colocci, Ph. Roussignol, and G. Weimann, Phys. Rev. B **47**, 10456 (1993).
[170] B. Gil, O. Briot, and R. L. Autombard, Phys. Rev. B **52**, R17028 (1995).
[171] X. B. Li, D. Z. Sun, J. P. Zhang, and M. Y. Kong, Appl. Phys. Lett. **72**, 936 (1998).
[172] C. V. Reddy, K. Balakrishnan, H. Okumura, and S. Yoshida, Appl. Phys. Lett. **73**, 244 (1998).
[173] J. M. Zavada and Z. Dahua, *Luminescence of Erbium Atoms in III–V Compound Semiconductors*, Proceedings of the twenty-first state-of-the-art Program on Compound Semiconductors (SOTAPOCS XXI) (Proc.-Electrochem. Soc. Pennington, NJ, 1995), pp. 250–69.
[174] S. J. Pearton, C. R. Abernathy, J. D. MacKenzie, R. N. Schwartz, R. G. Wilson, J. M. Zavada, and R. J. Shul, *Er-Doping of GaN and Related Alloys*, Rare Earth Doped Semiconductors II Symp. (Mater. Res. Soc. Symp. Proc. Pittsburgh, PA, 1996), pp. 47–56.
[175] J. T. Torvik, R. J. Feuerstein, C. H. Qui, M. W. Leksono, J. I. Pankove, and F. Namavar, *Annealing Study of Erbium and Oxygen Implanted Gallium Nitride*, Rare Earth Doped Semiconductors II Symp. (Mater. Res. Soc. Symp. Proc. Pittsburgh, PA, 1996), pp. 199–204.
[176] S. Kim, S. J. Rhee, D. A. Turnbull, E. E. Reuter, X. Li, J. J. Coleman, and S. G. Bishop, Appl. Phys. Lett. **71**, 231 (1997).
[177] J. T. Torvik, R. J. Feuerstein, C. H. Qui, J. I. Pankove, and F. Namavar, J. Appl. Phys. **82**, 1824 (1997).
[178] J. T. Torvik, C. H. Qiu, R. J. Feuerstein, J. I. Pankove, and F. Namavar, J. Appl. Phys. **81**, 6343 (1997).
[179] D. M. Hansen, A. R. Zhang, N. R. Perkins, S. Safvi, L. Zhang, K. L. Bray, and T. F. Kuech, Appl. Phys. Lett. **72**, 1244 (1998).
[180] J. D. MacKenzie, C. R. Abernathy, S. J. Pearton, U. Hommerich, J. T. Seo, R. G. Wilson, and J. M. Zavada, Appl. Phys. Lett. **72**, 2710 (1998).
[181] S. J. Pearton, C. R. Abernathy, J. D. MacKenzie, U. Hommerich, J. M. Zavada, R. G. Wilson, and R. N. Schwartz, J. Vac. Sci. Technol. A **16**, 1627 (1998).
[182] J. Neugebauer and C. G. Van de Walle, Mater. Res. Soc. Symp. Proc. **339**, 687 (1994).
[183] P. Perlin *et al.*, Phys. Rev. Lett. **75**, 296 (1995).
[184] L. X. Yun, L. Riba-Salamanca, M. G. Spencer, K. Wongchigul, P. Zhou, X. Tang, V. Talyansky, and T. Venkatesan, Mater. Res. Soc. Symp. Proc. **449**, 555 (1997).
[185] S. Nakamura, T. Mukai, and M. Senoh, J. Appl. Phys. **71**, 5543 (1992).
[186] D. C. Look and R. J. Molnar, Appl. Phys. Lett. **70**, 3377 (1997).
[187] W. Götz, L. T. Romano, J. Walker, N. M. Johnson, and R. J. Molnar, Appl. Phys. Lett. **72**, 1214 (1998).
[188] F. Ren, in *GaN and Related Materials*, edited by S. J. Pearton (Gordon and Breach, New York, 1997), pp. 433–69.
[189] S. Nakamura, T. Mukai, and M. Senoh, Jpn. J. Appl. Phys., Part 1 **31**, 2883 (1992).
[190] O. Briot, in Ref. 8, pp. 70–122.
[191] T. Tanaka, A. Watanabe, A. Amano, H. Kobayashi, I. Akasaki, S. Yamazaki, and M. Koike, Appl. Phys. Lett. **65**, 2024 (1994).
[192] A. E. Hughes and S. C. Jain, Adv. Phys. **28**, 717 (1979).
[193] S. Nakamura, M. Senoh, and T. Mukai, Jpn. J. Appl. Phys., Part 2 **30**, L 1708 (1991).
[194] S. Nakamura, T. Mukai, M. Senoh, and N. Iwasa, Jpn. J. Appl. Phys., Part 2 **31**, L 139 (1992).
[195] W. Götz, N. M. Johnson, J. Walker, D. P. Bour, and R. A. Street, Appl. Phys. Lett. **68**, 667–9 (1996).
[196] N. Akutsu, H. Tokunaga, I. Waki, A. Yamaguchi, and K. Matsumoto, Mater. Res. Soc. Symp. Proc. **482**, 113 (1998).
[197] M. E. Lin, G. Xue, G. L. Zhou, J. E. Greene, and H. Morkoç, Appl. Phys. Lett. **63**, 932 (1993).
[198] M. Miyachi, T. Tanaka, Y. Kimura, and H. Ota, Appl. Phys. Lett. **72**, 1101 (1998).
[199] O. Brandt, H. Yang, H. Kostial, and K. H. Ploog, Appl. Phys. Lett. **69**, 2707 (1996).
[200] M. A. L. Johnson, Z. Yu, C. Boney, W. C. Hughes, J. W. Cook, Jr., J. F. Schetzina, H. Zhao, B. J. Skromme, and J. A. Edmond, Mater. Res. Soc. Symp. Proc. **449**, 215 (1997).
[201] C. J. Sun *et al.* Appl. Phys. Lett. **71**, 2978 (1997).
[202] W. Shan, B. D. Little, J. J. Song, Z. C. Feng, M. Schurman, and R. A. Stall, Appl. Phys. Lett. **69**, 3315 (1996).
[203] S. Nakamura and T. Mukai, Jpn. J. Appl. Phys., Part 2 **31**, L 457 (1992).
[204] S. Nakamura, T. Mukai, and M. Senoh, Jpn. J. Appl. Phys., Part 2 **32**, L 16 (1993).
[205] M. D. McCluskey, C. G. Van de Walle, C. P. Master, L. T. Romano, and N. M. Johnson, Appl. Phys. Lett. **72**, 2725 (1998).
[206] Y. Narukawa, Y. Kawakami, M. Funato, S. Fujita, S. Fujita, and S. Nakamura, Appl. Phys. Lett. **70**, 981 (1997).
[207] D. Behr, J. Wagner, A. Ramakrishnan, H. Obloh, and K.-H. Bachem, Appl. Phys. Lett. **73**, 241 (1998).
[208] P. A. Crowell, D. K. Young, S. Keller, E. L. Hu, and D. D. Awschalom, Appl. Phys. Lett. **72**, 927 (1998).
[209] C.-K. Sun, T.-L. Chiu, S. Keller, G. Wang, M. S. Minsky, S. P. DenBaars, and J. E. Bowers, Appl. Phys. Lett. **71**, 425 (1997).
[210] S. Keller, U. K. Mishra, S. P. DenBaars, and W. Seifert, Jpn. J. Appl. Phys., Part 2 **37**, L 431 (1998).

[211] Y. Koide, H. Itoh, M. R. H. Khan, K. Hiramatsu, N. Sawaki, and I. Akasaki, J. Appl. Phys. **61**, 4540 (1987).
[212] H. Angerer *et al.*, Appl. Phys. Lett. **71**, 1504 (1997).
[213] C. Stampfl and C. G. Van de Walle, Appl. Phys. Lett. **72**, 459 (1995).
[214] S. C. Jain and D. J. Roulston, Solid-State Electron. **34**, 453 (1991); S. C. Jain, R. P. Mertens and R. J. Van Overstraeten, *Advances in Electronics and Electron Physics*, edited by P. W. Hawkes (Academic, New York, 1991), Vol. 82, pp. 197–275.
[215] S. P. DenBaars, U. K. Mishra, Y.-F. Wu, D. Kapolnek, P. Kozodoy, M. Mack, J. S. Speck, and S. Keller, Mater. Res. Soc. Symp. Proc. **423**, 23 (1996).