Intersubband Transition-Based Processes and Devices in AlN/GaN-Based Heterostructures

This review covers the physics, epitaxial growth, fabrication, and characterization of optoelectronic devices for use in video players and other consumer electronics as well as in commercial systems.

By Daniel Hofstetter, Esther Baumann, Fabrizio Raphael Giorgetta, Ricardo Théron, Hong Wu, William J. Schaff, Member IEEE, Jahan Dawlaty, Paul A. George, Lester F. Eastman, Fellow IEEE, Farhan Rana, Prem K. Kandaswamy, Fabien Guillot, and Eva Monroy

ABSTRACT | We report on the physics, epitaxial growth, fabrication, and characterization of optoelectronic devices based on intersubband transitions in the AlN/GaN material system. While in 1999, only results of optical absorption experiments could be shown, photodetectors and modulators with operation frequencies beyond 10 GHz as well as optically pumped light emitters have been demonstrated recently. This is the reason for a comprehensive report on the most important properties of such devices. Beside some basic theoretical considerations, we will concentrate on the fabrication and characterization of modulators, switches, photodetectors, and light emitters. At the end of this paper, an outlook to future trends and developments in this emerging field will be given.

KEYWORDS | III-nitrides; intersubband transitions; optical rectification; quantum cascade detectors; quantum wells

I. INTRODUCTION

Intersubband transitions have already quite a long history; their first description dates back to 1982 by Ando et al. [1]. These three authors observed a novel kind of optical transition between bound electronic levels which were present in the two-dimensional carrier gas at a semiconductor heterojunction. Because of the small energetic separation between adjacent levels, the observed absorption effects were visible only under excitation with far-infrared radiation. Therefore, the practical use of this effect was not immediately evident. After first proposals of using such intersubband transitions in infrared photodetectors by Coon et al. in 1984 [2], West and Eglash demonstrated in 1985 intersubband absorption in a very thin semiconductor layer sandwiched between two higher bandgap semiconductors. This “quantum well” (QW) configuration exhibited basically the same behavior as the heterojunction three years before: it absorbed infrared light [3], but this time at shorter midinfrared wavelengths. The big advantage of this new result was its possibility of being tailored: by simply growing thicker QWs, one could reduce the
energy of the absorbed photons. Two years after this important milestone, the first quantum-well infrared photodetector (QWIP) was realized by Levine et al. [4]. This was a very crucial step towards practical devices such as focal plane arrays for military and surveillance applications [5], [6]. Another seven years later, in 1994, the first laser emission, from the quantum cascade (QC) laser, was demonstrated by a research group headed by Capasso at Bell Laboratories [7]. This was the beginning of an impressive development that culminated in a room-temperature midinfrared continuously driven semiconductor laser in 2002 [8], [9]. Nowadays, QWIPs and QC lasers are commercialized and are present in many important applications. An especially interesting aspect of intersubband physics has been addressed only in 1999: it was the demonstration of intersubband transitions approaching the technologically important wavelength range around 1.55 µm. Such transitions are possible in carefully selected material systems offering a sufficiently high conduction band discontinuity. One possibility is the use of AlN/GaN, which exhibits a band offset of nearly 2 eV; this is clearly sufficient for the exploitation of near-infrared intersubband transitions [10]–[17]. However, one has to take into account certain particularities of this material system such as its hexagonal crystal structure and its pyro- and piezoelectric behavior. These properties lead to a relatively complicated band structure in which the simulation of electronic levels is substantially more difficult than for instance in GaAs/AlAs [18]–[20]. Nevertheless, recent experiments with AlN/GaN-based superlattice structures have revealed very interesting physical properties, which have eventually resulted in high-speed optoelectronic devices for the 1.3/1.55 µm wavelength region. That this statement is actually more than just a hypothesis has been confirmed already by several experiments which demonstrated intersubband scattering times on the order of hundreds of femtoseconds [21], [22].

This paper is organized as follows. In Section II, we will present some theoretical aspects of intersubband physics and introduce Fermi’s golden rule. In Section III, some important facts of crystal growth will be described. Section IV deals with devices such as photodetectors, light emitters, and saturable absorbers. In Section V, we will discuss future trends and possible new developments in this field.

II. PHYSICS/THEORY

A. Physics of Intersubband Transitions

The physics of intersubband transitions can be understood on the base of Fermi’s golden rule. In our case of a two-level system (with energies $E_i$ and $E_f$) interacting with electromagnetic radiation, it describes the probability $W_{if}$ of an electron passing from the lower (energy $E_i$) to the upper level (energy $E_f$). In very general terms, Fermi’s golden rule writes as

$$W_{if} = \frac{2\pi}{\hbar} \left| \langle \psi_i | H' | \psi_f \rangle \right|^2 \cdot \delta(E_f - E_i - \hbar \omega) \quad (1)$$

with $\hbar \omega$ being the photon energy, $H'$ the interaction Hamiltonian, and $\psi_i$ and $\psi_f$ the wave functions of the initial and the final electronic states, respectively. If we use the dipole approximation, then the interaction Hamiltonian can be expressed as $H' = (q^2 E_0^2/(4m^*) \omega^2) \vec{e} \cdot \vec{p}$ with $q$ being the elementary charge, $E_0$ the amplitude of the electric field, $m^*$ the effective mass, $\omega$ the angular frequency of the incoming radiation, $\vec{p}$ the momentum operator, and $\vec{e}$ the unit vector of the light’s oscillation direction. If this term is inserted into (1), then we get

$$W_{if} = \frac{2\pi}{\hbar} \left| \frac{q^2 E_0^2}{4m^* \omega^2} \langle \psi_i | \vec{e} \cdot \vec{p} | \psi_f \rangle \right|^2 \cdot \delta(E_f - E_i - \hbar \omega). \quad (2)$$

In the next step, the absorption probability given by (2) is normalized towards the number of incoming photons and multiplied with a geometrical factor to take into account the angle between the incoming radiation and the surface normal. Integration over all allowed directions has to be done to get rid of the scalar product in the matrix element, while the energy-conserving delta function is replaced by a Lorentzian lineshape [23]. All these operations result in a compact formula for the peak absorbance $\alpha_{QW}$

$$\alpha \cdot \tau_{QW} = \frac{2\pi^2 q^2}{\varepsilon_0 n \lambda} \cdot n_i \cdot n_{QW} \cdot (z_{if})^2 \cdot \frac{\sin^2 \theta}{\cos \theta} \cdot \frac{1}{\Delta E_{if}} \quad (3)$$

where $\theta$ is the incidence angle measured towards the surface normal, $z_{if}$ the dipole matrix element, $n_i$ the sheet carrier density, $n_{QW}$ the number of QWs, $n$ the refractive index of the material under investigation, $\lambda$ the wavelength corresponding to the energy difference between states $i$ and $f$, $\varepsilon_0$ the vacuum permittivity, and $\Delta E_{if}$ the full width at half-maximum of the optical transition. Careful analysis of the formula shown above reveals that interaction between light and an intersubband transition is possible only for transverse magnetic (TM) polarized light. This interesting feature is known under the term polarization selection rule. Unfortunately, it prevents us from using the most convenient geometry for all of our experiments, namely, vertical incidence. The latter would be possible only if one would add a surface grating for a redirection of the light within the sample. For any particular sample geometry, for instance, a multipass zigzag waveguide, (3) contains the number of passes and
the number of active region QWs. According to these considerations, a typical 1.5-nm-thick GaN QW sandwiched between AlN barriers and with a silicon doping level on the order of $1 \times 10^{20}$ cm$^{-3}$ will produce roughly 0.5% absorption per pass. This is valid for an incidence angle of 45° into the sapphire substrate. Using, for instance, 40 active region periods, it is therefore possible to fabricate structures which absorb nearly 50% of the light per double-pass. An example of such absorbing structures is given in Fig. 1; it compares the absorption curves for five double-passes through structures with different QW thickness. Due to the decreasing size of the dipole matrix element in thinner QWs, the absorption goes down if the well thickness is reduced. From the broadening of the absorption curves at higher transition energy, an interfacial roughness of about one monolayer can be estimated. This number will be confirmed below by investigations using a transmission electron microscopy (TEM).

B. Saturation of Absorption

Another interesting aspect of intersubband transitions is the effect of absorption saturation. In optical pump and probe measurements on AlN/GaN superlattice structures, several authors have reported very short carrier relaxation times. For nitride structures absorbing at 1.55 μm, the most typical absorption recovery times are on the order of 170–370 fs [21], [24], [25]. Based on this value, one can compute at which intensity saturable absorption should occur. The basic relation describing this behavior is given by

$$\alpha / \alpha_0 = \frac{1}{1 + (I_{in}/I_{sat})}$$

(4)

without saturation, and $I_{in}$ and $I_{sat}$ the incoming and the saturation intensities, respectively. Since pump probe experiments using different input intensities reveal the saturation intensity of this two-level system, these measurements allow us an indirect determination of the lifetime of the system [26]. This lifetime is in fact linked to the saturation intensity via

$$\tau_{life} = \frac{\Delta E_{12}}{2I_{sat}} \cdot \frac{\cos \theta \cdot n_{QW} \cdot n_i}{\alpha_0} \cdot \frac{1}{I_{QW}}$$

(5)

In (5), $\Delta E_{12}$ is the photon energy, while the other quantities are defined as in the previous formulae. It turns out that the saturation intensity for nitride-based QW structures is quite high, on the order of 1 GW/cm$^2$. Although certain other applications might be able to produce such a high intensity, the main application field will most likely lie in ultrafast solid-state lasers with saturable absorbers. First experiments have indeed revealed a great potential for such switching devices. In the nitride material system, it seems to be crucial to adapt the device geometry to achieve saturation intensities. By making waveguides with a very small cross-section, it is possible to fabricate saturable absorbers which can be operated with moderate power levels. This technique has been successfully implemented by the lizuka group [27], and it has been demonstrated by the Moustakas group at Boston University that one can still refine it in order to further reduce the required saturation power [28].

C. Photodetectors

Intersubband photodetectors based on AlN/GaN superlattices show an interesting mode of functioning that could be clarified only several years after their first experimental demonstration. For a correct understanding of their mechanism, it is crucial to know the exact size and orientation of the internal polarization fields. According to the theory presented in Ambacher et al. [29], these internal pyro- and piezoelectric polarizations can be treated like

$$p_{AlInGaN}^{AlI} (x) = -0.090x - 0.034(1 - x) + 0.021x(1 - x)$$

(6)

$$p_{AlIN/GaN}^{AlI} (x) = 0.026(1 - x) + 0.0248x(1 - x)$$

(7)

$$p_{AlIN/GaN}^{piezo} (x) = -0.0525x + 0.0282x(1 - x)$$

(8)

where $x$ is the Al concentration of the AlGaN compound and $p_{AlInGaN}^{AlI}(x)$ is the pyroelectric polarization for AlGaN. Equation (6) describes the spontaneous polarization of AlGaN compounds, (7) stands for the piezoelectric polarization $p_{AlIN/GaN}^{piezo}(x)$ of an AlGaN layer grown fully strained on AlN, and (8) describes the piezoelectric polarization
Fig. 2. Schematic representation of a GaN-based QW with a slightly “diagonal” transition due to the potential asymmetry induced by the piezo- and pyroelectric field. The [0001] growth direction is marked by an arrow [36, Fig. 16].

For AlGaN/GaN(x) for an AlGaN layer on top of GaN. According to the work of Bernardini and Fiorentini [30], the internal fields \( F_{\text{int}} \) in the superlattice structure are related to the respective polarizations and layer thicknesses by

\[
F_{\text{GaAs}} = \frac{t_{\text{AlN}}}{e_{\text{AlN}}} \left( \frac{p_{\text{GaAs}}}{p_{\text{piezo}}} - \frac{p_{\text{AlN}}}{p_{\text{piezo}}} \right)
\]

\[
F_{\text{AlN}} = \frac{t_{\text{GaAs}}}{e_{\text{GaAs}}} \left( \frac{p_{\text{AlN}}}{p_{\text{piezo}}} + \frac{p_{\text{AlN}}}{p_{\text{piezo}}} + \frac{p_{\text{GaAs}}}{p_{\text{piezo}}} \right)
\]

where \( e_{\text{AlN}} \) and \( e_{\text{GaAs}} \) are the dielectric constants for AlN and GaN, respectively, while \( t_{\text{AlN}} \) and \( t_{\text{GaAs}} \) are the corresponding layer thicknesses. For AlN/GaN hetero-interfaces with all layers fully strained on the underlying AlN substrate, one obtains thus polarization fields on the order of 5–8 MV/cm. The resulting asymmetry of the QW has important consequences for the functioning of intersubband photodetectors. As shown in the schematic picture of Fig. 2, the ground state wavefunction is slightly displaced towards the left (i.e., in the [0001] growth direction) while the first excited state is displaced towards the right (i.e., in the [000-1] direction). Therefore, excitation of an electron into the excited state is always accompanied by a displacement of the electron along the growth axis. However, since the positively charged doping atom does not change its place, a small dipole moment occurs for each electron/dopant atom pair. If, due to strong absorption, a certain fraction of electrons will undergo such a transition, the resulting total light-induced polarization can become appreciable. It will eventually be large enough to produce an electrical voltage between a dark and an illuminated area of the device [31]. The photovoltage generated by this effect can be described by

\[
V_{\text{rect}} = \frac{n_{QW}t_{QW}}{e_{\text{GaAs}}n_{\text{GaAs}}} \cdot \tau_{12} \cdot \delta_{12} \cdot n_{3D} \cdot |(n_{12})|^2 \cdot \frac{e^2}{2I \hbar \cos \theta}
\]

where \( t_{QW} = 1.5 \text{ nm} \) is the thickness of the QW, \( n_{QW} = 80 \) (according to one double-pass through the active region) is the number of QWs, \( \tau_{12} = 370 \text{ fs} \) is the lifetime of the upper quantized state, \( \delta_{12} = 3.0 \text{ Å} \) is the lateral displacement of the electrons, \( n_{3D} = 10^{20} \text{ cm}^{-3} \) is the volume carrier density, \( n_{12} = 3.5 \text{ Å} \) is the dipole matrix element of the optical transition, \( \theta = 37^\circ \) is the angle of incidence within the GaN material (diffracted by the sapphire substrate), \( E^2 = (2 I)/(c \varepsilon_0) = 7.5 \times 10^{10} \text{ V}^2/\text{m}^2 \) is the square of the incoming electric field amplitude (10 mW focused to an area of 10 \( \times 10 \) \( \mu \text{m}^2 \)), \( 2I = 80 \text{ mW} \) is the full width at half-maximum of the transition, \( n_{\text{GaAs}} = 2.2 \) is the refractive index of the QW material, and \( \varepsilon_{\text{GaAs}} = 10.2 \) the static dielectric constant of GaN. Using the parameters stated above results in a value of 80 \( \mu \text{V}/\text{W} \), which needs to be compared to the experimental value of 130 \( \mu \text{V}/\text{W} \) at 150 K. It has to be stressed, however, that in particular, the lifetimes and the lateral displacements suffer from large uncertainties. This is so because of an eventual transfer of carriers into additional impurity levels being resonant with the upper detector state. The disadvantage of such an optical rectification type of operation is obviously that the lateral displacements are relatively small, and the lifetimes of the excited electrons are short; both effects limit the size of the photovoltaic signal to relatively small values. However, since again the lifetime is short, the system recovers very quickly so that high-frequency operation is easily possible. In the discussion about the experimental results, we will see that such detectors have already been demonstrated for signal frequencies in excess of 10 GHz. Recent experiments using a quantum dot-based active region have revealed a marked performance improvement, most likely due to lateral coupling of the excited quantum dot state with the ground state of the wetting layer.

D. QC Lasers and Light Emitters

As a final point, quantum cascade lasers remain to be demonstrated in this material system. Although there has been quite a considerable effort in this direction, results are somewhat discouraging so far. Since, in addition, no truly convincing evidence of resonant tunneling has been reported at this point, an electrically injected QC light emitter remains an extremely challenging target. Nevertheless, there has been substantial progress in terms of optically pumped light emission. The most advanced experiments have been conducted at the University of Paris-Sud in the group of Julien [32].
One of the most serious problems for a QC laser in this material system is clearly the short lifetime of an electron in the upper quantum state. Even if this value could be made as “long” as, say, approximately 350 fs using tricks like diagonal transitions, the lower state lifetime should still be considerably shorter in order to establish a sufficiently large population inversion. Considering the good laser performance achieved with the so-called bound-to-continuum design in InGaAs/InAlAs-based QC lasers, such a design is also proposed for nitride-based light emitters. It is shown in Fig. 3. The injector is a regular superlattice whose electronic states split into a miniband under application of a small bias voltage. Owing to the equal layer thickness of GaN wells and AlN barriers, strain will be fully compensated in the injector. The gain region itself is designed as two strongly coupled QWs. The lower lasing state is aligned with the uppermost downstream injector states, while the upper lasing state is a hybrid state between the two QWs of the gain region. It is energetically aligned with the lowest upstream injector states and has a good overlap with the injector wavefunctions. Thanks to the continuum-like lower lasing state, the electron lifetime could in theory be reduced to 150 fs, which is short enough to achieve an inversion. The proposed laser structure is designed to emit at a photon energy of roughly 500 meV (2.5 µm). Since each period would show a voltage drop of at least 600 mV, a maximum of 15–20 active region periods can be grown. Otherwise, the total voltage drop on the structure would exceed 12 V, which can be considered as maximum value to be tolerated. The waveguide of such a device can be formed using the sapphire substrate as a lower cladding layer. The upper cladding is then fabricated with an evaporated SiO₂ layer; altogether this would result in an appreciable overlap factor for the active region. Based on the different lifetimes and the expected population inversion, an estimation of the maximal current density $J_{\text{max}}$ is possible

$$\Delta n = n_3 - n_2 = \frac{J_{\text{max}}}{q} \frac{1}{\tau_3} \left(1 - \frac{\tau_3}{\tau_{32}}\right)$$

where $n_2$ and $n_3$ are the sheet carrier densities in the lower and upper laser levels 2 and 3, respectively, $\tau_2$ and $\tau_{32}$ are the lifetimes of the lower and the upper laser level (considering only the 3 → 2 transition), and $\tau_{31}^{-1} = \tau_{32}^{-1} + \tau_{31}^{-1}$ is the total lifetime of level 3 considering carrier loss into level 1 as well. Using the lifetimes cited earlier and a population inversion of $\Delta n = 3 \times 10^{10}$ cm⁻², we get $J_{\text{max}} = 24$ kA/cm². In addition, it is important to know the approximate threshold current for a given structure. This quantity can be estimated if parameters like spectral linewidth $2\gamma = 0.1$ eV, oscillator strength of the transition $f = 0.6$, effective mass $m^* = 0.2 m_e$, and combined waveguide/mirror losses $\alpha_m + \alpha_{WG} = 1$ cm⁻¹ are known. We then get

$$J_{\text{th}} = \frac{\varepsilon_0}{4\pi q} \frac{1}{\tau_3} \left(1 - \frac{\tau_3}{\tau_{32}}\right) \frac{2\lambda n_{\text{ref}} L_p \gamma (\alpha_m + \alpha_{WG})}{\Gamma_{\text{opt}} \varepsilon_0^2 \gamma^3}.$$  

When using realistic values of $\lambda = 2.5$ µm, $\tau_3 = 150$ fs, $\Gamma_{\text{opt}} = 40\%$, $\Gamma_{\text{opt}} = 40\%$, $n_{\text{ref}} = 2.5$, and $L_p = 15$ nm, this computation yields minimal values on the order of $J_{\text{th}} = 10$ kA/cm². In contrast to passive devices such as photodetectors or optical switches, which rely on intersubband absorption effects only, such lasers would unfortunately not be ultrafast. This might be somewhat counterintuitive at first glance, but as described in a paper by Bouadma et al. [33], a short laser cavity with a very short photon lifetime offers a very efficient way to push the relaxation oscillation frequency of a semiconductor laser to its highest values. Since GaN QC lasers will obviously be working in a low-loss/low-gain regime, they will most likely have relatively high thresholds and long cavities. Therefore, they are not very good candidates for high-speed modulation. For this reason, only moderate maximal modulation frequencies in the low gigahertz range can be expected.

III. FABRICATION

AlN/GaN QW structures can be grown either by molecular beam epitaxy (MBE) or by metal-organic vapor phase epitaxy (MOVPE). Most of the results to be presented here are based on material grown by plasma-assisted MBE (PAMBE). The low growth temperature of this technique results in sharp AlN/GaN interfaces, which is a critical
requirement to achieve decent device operation at 1.55 μm. Here we give only some general information concerning the growth procedure; details can be found in the work of Monroy [34], [35]. All PAMBE samples are grown on AlN-on-sapphire templates fabricated by MOVPE. The AlN layer has a thickness of roughly 1 μm and the orientation of the substrate is C-face. During the growth of the PAMBE layers, different methods have been investigated in order to guarantee metal-rich growth conditions, which in turn result in smooth surfaces. More particularly, In used as a surfactant, Ga excess, and growth interruptions along with Ga and Al excess for GaN and AlN layers, respectively, have been explored [35]. From a morphological point of view, the three different methods gave comparable results; however, the crystalline quality of the Ga excess samples was highest, so that most layers of this paper were grown using this method. A typical characterization of sample morphology and interface quality using atomic force microscopy (AFM) and TEM is presented in Fig. 4(a) and (b), respectively. The AFM analysis showed a root-mean-square surface roughness of about 0.6 nm in an area of 2.5 × 2.5 μm². From the analysis of the TEM picture, we concluded that the AlN/GaN interfaces are abrupt at the monolayer scale [36].

If, on the other hand, MOVPE is used for growth of such structures, one has to deal with somewhat larger interface roughness [37], [38]. However, for certain device applications involving thick layer stacks, it is obviously an advantage to have a higher growth rate like is typically achieved using MOVPE. Different authors have also reported structures that contain layers from both growth techniques; in these samples, all thick cladding layers are typically grown using MOVPE, while the thin QWs and barriers in the active region are grown by MBE. This is an ideal combination of the high growth rate and good crystal quality of MOVPE and the high interface control of MBE. A good example of such a device is the lattice-matched InAlN/GaN waveguide reported by Lupu et al. [39].

IV. DEVICES

In terms of devices, we will first briefly talk about saturable absorbers, optical switches, and modulators. Since considerable work has been devoted to photodetectors, it will constitute the main part of this paragraph. Lastly, some results of luminescence devices will be presented and discussed.

A. Saturable Absorbers/Modulators

As we have seen in the theory paragraph, the intensities required to achieve saturable absorption in GaN are quite considerable. Nevertheless, some successful demonstrations have been published. The most advanced structures are those reported by Iizuka et al. from Toshiba Corporation in Japan [40], [41]. As shown in the schematic cross-section of Fig. 5, the device was grown on a 70-nm-thick high-temperature AlN (HT-AlN) buffer followed by a 500 nm AlN/GaN multiple intermediate layer (MIL). The active zone consisted of a 480-nm-thick GaN lower waveguide layer, 10 AlN/GaN multi-QWs (Si, 5 × 10¹⁹ cm⁻³), and a 960-nm-thick GaN upper waveguide layer. The QW thickness was 1.5 nm while the barriers measured 2 nm. Most recent devices were mounted in a butterfly package and fiber-pigtailed for further use, as presented in the photograph of Fig. 6. Fabrication was based on lateral waveguide etching with an electron cyclotron resonance reactive ion beam etching

Fig. 4. (a) Atomic force microscopy surface scan and (b) transmission electron microscopy image of a superlattice sample grown by PAMBE [36, Fig. 4].
(ECR-RIBE) system. The switch consisted of a 400 µm long and 1.5 µm wide waveguide with tapered ends on both sides. Antireflection coatings helped to minimize the in- and out-coupling losses to very small values. The total insertion loss of the component was on the order of 6.6–8.6 dB for TE mode.

Fig. 7 shows a series of experiments proving the correct functioning of the device. Using the signal and idler beams of an optical parametric oscillator (OPO) as signal and control pulse generators, respectively, switching times of below 1 ps could be achieved. Thanks to the waveguide scheme, the switching energy was reduced to a value on the order of 100 pJ. Extinction ratios of the saturable absorption in excess of 11.5 dB were seen. Since the recovery times were on the order of <1 ps, a series of pulses separated by only 1 ps in time could be demultiplexed using these structures.

As far as modulators are concerned, several prototype devices were demonstrated first by the authors at the University of Neuchatel [42], but later mainly from the Julien team at the University of Paris-Sud. While our device was based on an electron transfer between the two-dimensional carrier gas underneath a 5 QW active region (see Fig. 8), their first tests were performed using coupled double QWs with a very thin tunneling barrier in between. With our prototype modulator, we faced the problem of an insufficient signal on–off ratio. There was a strong absorption tail from the two-dimensional electron gas’ reaching into the wavelength region of interest. Since this absorption could not be suppressed, the on–off ratio did not become higher than a couple of decibels. This is shown in Fig. 9, which presents the absorption spectra for three different voltages on the top contact. On the other hand, the main problem of the double QW structures was to adjust the doping levels and the layer thicknesses in a suitable way, which would allow moving the carriers from one well into its neighbor [43]. The published work showed a modulator with a quite convincing performance, however. The device consisted of 20 periods of a 3-nm-thick reservoir GaN QW coupled via an ultrathin AlN barrier of 1 nm to an active GaN (Si, 5 x 10^19 cm^-3) QW of 1 nm thickness. This active region was sandwiched between two Al_{0.6}Ga_{0.4}N (Si, 5 x 10^19 cm^-3) contact layers. The samples were processed as square mesa with 700 µm side length having 500 x 500 µm^2 large windows in the center. This process allowed testing under Brewster geometry. As shown in Fig. 10, electromodulated absorption in two different wavelength windows around 1.2–1.67

![Fig. 5](image5.png)

**Fig. 5. Schematic representation of a GaN-based switch as used by Iizuka. The structure was grown on an HT-AlN buffer layer, followed by a MIL while the mesa was etched with an ECR-RIBE system [41, Fig. 1].**

![Fig. 6](image6.png)

**Fig. 6. Photograph of a GaN-based switch as used by Iizuka. Fibers for optical input and output are coming from the left/right, while the electrical connections are shown on the bottom.**

![Fig. 7](image7.png)

**Fig. 7. Experimental results of a GaN-based all-optical switch as used by Iizuka. Whenever signal and control pulse coincided in time, a bleached absorption (increased transmittance) was observed. The shortest pulse interval demonstrated was 0.67 ps [41, Fig. 6].**
and 2.1–2.4 μm was achieved. The maximum modulation depth was 44% and the RC-limited cutoff frequency reached a value of 11.5 MHz.

B. Photodetectors

In some sense, AlN/GaN-based infrared photodetectors play a privileged role because they constituted the first working optoelectronic device using intersubband transitions in this material system [44]. It took, however, quite a long time until the functioning of these components was well understood. Here, we review the current status and give an outlook of future developments.

The generic structure for such photodetectors is based on an MBE grown, 40-period AlN/GaN superlattice with layer thicknesses of 1.5 nm for the Si-doped (0.5–1.0 × 10²⁰ cm⁻³) GaN QWs, and between 1.5 and 15 nm for the AlN barriers (undoped). This active region is typically grown on a 500-nm-thick AlN buffer layer and covered with a 100 nm AlN cap. Growth is performed on top of MOVPE-grown templates consisting of a 1-μm-thick AlN layer. Characterization of such structures is done in two stages requiring almost identical sample preparation: Ti/Au (10/400 nm) based contacts are being evaporated on the sample surface; the patterning is stripe-shaped for low-frequency electrical tests and square-shaped for high-frequency measurements. Absorption experiments do usually not require any electrical contacts. The stripes are 800 μm wide and 3 mm long, while the squares measure typically 100 μm across. In order to allow an efficient coupling of the incoming radiation into the sample, two

Fig. 8. Electrooptical modulator prototype with an applied bias of –2 V (a) at the top contact and (b) without bias. Under the action of the applied bias voltage, the carriers will be transferred from or to the two-dimensional electron gas [42, Fig. 3].

Fig. 9. Experimental results of the modulator presented schematically in Fig. 8. An on-off ratio on the order of 3 dB was seen [42, Fig. 4].

Fig. 10. Experimental results of the modulator presented by the Julien group. An on-off ratio on the order of 11.5 dB was seen [43, Fig. 3].
parallel 45° inclined facet mirrors are being polished on both sides. To form a zigzag waveguide, the back of the sample needs to be polished as well. A schematic picture of such a sample along with the direction of the incoming light is shown in Fig. 11. For simple absorption measurements, it is sufficient to mount this waveguide sample on a holder. A comparison between normalized absorption data for both TE and TM polarization allows the correct extraction of absorption curves. The absorbance is then given by

$$\alpha(\nu)L = \ln \left( \frac{I_{TE}(\nu)}{I_{TM}(\nu)} \right)$$

(14)

where $I_{TM}$ and $I_{TE}$ are transmitted intensities for both TM and TE polarized light and $L$ the effective interaction length. Experiments involving photodetection were done by illumination of the device at only one of the two electrical contacts, while leaving the other one in the dark. The exact reason for using this technique can be explained by the nonlinear optical rectification mechanism used in these devices: the lateral conductivity between the two contacts is quite low. According to (11), illumination underneath one contact and at the correct wavelength induces an optical polarization, whereas the dark contact does not produce any effect. A measurement between dark and illuminated contacts therefore results in a measurable photovoltage. Typical spectral characteristics of the first such detector are shown in Fig. 12 [44]. It worked at a peak wavelength of 1.85 μm, but due to its relatively broad sensitivity curve, it could nevertheless detect down to the technologically important wavelength of 1.55 μm. This feature was tested with a superluminescent diode. The maximum operating temperature was 170 K. The high-frequency behavior of this first device was determined by several inherently slow processes and the large size of the device; they limited its operation to values below 100 kHz. In the following years, considerable optimization work was done, mainly in terms of crystal growth but also in terms of process and characterization [45], [46]. Later generations of such photodetectors delivered responsivity curves as shown in Fig. 13. In this experiment, three different samples with 3 QW thicknesses were grown; this resulted in a series of photodetectors being sensitive between 2.33 and 1.49 μm [47].

As mentioned in the Introduction, it was only in 2007 when the functioning of these devices could be explained in a satisfactory way. This was also the moment when more serious high-frequency tests could be run [48]. For these experiments, a more sophisticated mounting directly on a

![Fig. 11. Schematic cross-sections through a typical intersubband absorption sample using nitride semiconductors [36, Fig. 5].](image)

![Fig. 12. Energetic peak positions of absorption (red line 77 K, green line 300 K) and photovoltage (blue dashed line, 10 K) for the first GaN-based intersubband photodetector demonstrated in 2003 [44, Fig. 2].](image)

![Fig. 13. Photovoltage responsivity spectra of a series of three samples having different QW thicknesses (17, 22, and 38 Å) showing the quantum confinement shift. The sample with the thickest QW shows two higher order transitions [47, Fig. 2 (bottom)].](image)
subminiature Version A connector socket with short bond wires was used. Otherwise, parasitic effects due to the highly inductive bonds severely hampered the high-frequency characteristics of the detector. For characterization of a high-speed device, we carefully separated optical signal generation and detection. On the generation side, we used a commercially available continuous wave operated telecommunication laser at 1.55 μm. Its output was directly delivered to an optical modulator with a maximum modulation frequency going up to 10 GHz. The signal then entered a 30-m-long optical single mode fiber and was directed towards the AlN/GaN-based detector. The latter was held at room temperature and mounted on an x-y-z-stage for alignment purposes. The result of these measurements is shown in Fig. 14, which presents electrical detector signal as a function of optical input at frequencies between 100 kHz and 10 GHz. The rollover was seen at 43 MHz, while a signal was detected up to modulation frequencies of nearly 3 GHz. It is obvious that parasitic effects limited the performance of this device; as a matter of fact, Julien et al. recently reported a nitride-based quantum cascade detector with appropriate high-frequency stripe-line mounting (see below) [49]. This device worked well beyond 10 GHz, which is very encouraging for future developments.

In a different line of research, we tested such a detector also for the 1 → 3 intersubband transition, which can be exploited via exactly the same detection mechanism [50]. The detection wavelength is then half of 1.55 μm, namely, 775 nm. For this wavelength, there exist very powerful pulsed solid-state light sources such as Ti:sapphire or Nd:YAG lasers. By using such a laser device, but still the same detector as for the 3 GHz experiment, we were able to go to considerably higher frequencies as well. In the signal versus frequency characteristic, a pronounced change in slope from -20 to -40 dB/dec can be seen. It happens at exactly the same frequency as in the 1.55 μm experiment, suggesting an extrinsic mechanism as the limiting factor. Similar to a high-frequency experiment involving a midinfrared quantum cascade detector (QCD), where a second-order low-pass filter characteristic was observed, we suspect here that the parasitic inductance of the bond wire and the parasitic capacitance of the detector contacts produce this frequency behavior [51]. This is shown in Fig. 15, where the measured frequency response is reported for different experimental configurations. In the first test, where the 1 → 2 transition was excited using an Yb-doped mode-locked fiber laser, the signal was amplified by an SHF806P broadband amplifier. This resulted in a high-frequency cutoff of 2.2 GHz. Thanks to the better signal-to-noise ratio achieved with an additional NexTec NBL 00558 amplifier, the maximum measurement frequency could be pushed further, up to 5 GHz. When finally using a 780 nm pulsed high-power Nd:YAG laser by exploiting the 1 → 3 optical detection transition and the same amplification scheme as in the previous experiment, the highest detection frequency was 13.3 GHz.

C. Quantum Cascade Detectors

Recent experiments using a QCD based on a periodic GaN/AlGaN structure revealed even higher detection frequencies [49]. The device structure is shown in Fig. 16, while the frequency response of the detector is presented in Fig. 17. This detector contains 40 active region periods, each comprising a doped (Si, 1 x 10^19 cm^-3) GaN-based main well and an AlN/Al0.25Ga0.75N-based (6 x 1 nm AlN/1 nm Al0.25Ga0.75N) carrier extraction region. The latter is coupled to the precedent and the following stage’s main well. Besides the possibility of actively designing a real transport mechanism for the carriers, another important
Fig. 16. Schematic band structure of a near-infrared QCD. The transport direction is given by the slope of the potential ramp [49, Fig. 1].

Fig. 17. High-frequency behavior of a QC detector with a size of $17 \times 17 \mu m^2$ (red curve) and of $25 \times 25 \mu m^2$ (blue curve). The dotted lines are simulations taking into account the wire inductance and the device capacitance. The insets show, on top, the electrical equivalent circuit and, at the bottom, a microscope picture of the finished device [49, Fig. 2].

advantage of this device is its small size of $17 \times 17 \mu m^2$ and its correct high-frequency mounting using coplanar stripe line technology. As Fig. 17 shows, a $-3$ dB frequency of 11.4 GHz was seen, along with a cutoff beyond 30 GHz. A responsivity on the order of $2 - 2.5$ mA/W was reported for this device. Altogether, these are impressive results that clearly reveal the future potential of GaN-based photodetectors for telecommunication wavelengths. Whether the demonstration of this device already paves the way towards a more advanced intersubband system such as a QC laser remains an open question. For a QC detector, the designed potential asymmetry next to the active QW along with some sort of impurity-assisted electron scattering will suffice to allow a more or less correct function. For a QC laser, however, an extremely well-defined resonant tunneling process would be required. Although resonant tunneling cannot be excluded in the reported QCDs, the presently existing level broadening will render injection of electrons into an upper laser state very difficult and inefficient.

D. Light Emitters

A considerable amount of theoretical work has been published on designs for QC light emitters and lasers. The majority of these designs use the classical approach of the electron-longitudinal optical (LO-) phonon scattering to depopulate the lower laser level [52], [53]. Although this was also the first experimentally explored approach of the University of Paris-Sud group, no successful experimental demonstration of this concept has been published so far. An alternative to the LO-phonon resonance design is the bound-to-continuum design presented in Section II-D. Yet another approach was used for the first functional intersubband light emitter based on a 250-period regular AlN/GaN-superlattice (2.1 nm GaN/3 nm AlN), which was tested at the University of Paris-Sud as well by Nevou et al. [32]. No doping was intentionally added to the structure; the electron concentration is therefore intrinsic by residual doping [54]. The estimated doping level through this intrinsic doping was $4 \times 10^{12}$ cm$^{-2}$. In the device, the $1 \rightarrow 3$ transition was exploited for optical pumping with a 980 nm high-power semiconductor pump laser and the $3 \rightarrow 2$ transition for the emission of radiation at 2.1 $\mu m$. In order to minimize the mirror losses, two parallel, 90° oriented facet mirrors were polished onto the sample to couple out the intersubband light; pumping took place via two parallel, polished 45° facets (placed orthogonally to the 90° facets) and a multiple zigzag path through the sample. The power density of the pump laser was on the order of $10^4$ Wm$^{-2}$.

In order to measure the output radiation, lock-in amplification at 4 kHz was necessary. As Fig. 18 shows, the output spectrum is centered around 2.1 $\mu m$ and fully TM polarized. A comparison with the corresponding intersubband absorption shows very good agreement. The measured luminescence efficiency is on the order of 10 pW/W. A slightly better performance could be achieved in a more recent sample having an active region with a 200-period stack of quantum dots [55]. While pumping with an YVO$_4$ solid-state laser took place at a wavelength of 1.34 $\mu m$, the emission from the quantum dots was exactly one LO-phonon energy away, namely, at 1.48 $\mu m$. This is shown in Fig. 19. By fitting a Lorentzian lineshape to the sharply dropping high energy shoulder of the emission spectrum, it was possible to achieve an excellent matching with the measurement. The full width at half-maximum was as small as 4 meV, roughly a factor of ten smaller than in standard QW samples. Very recently, the group of Moustakas et al.
demonstrated such an optically pumped GaN-based QW light emitter under pulsed excitation with an OPO. Thanks to the much higher peak pump power, an increased output power on the order of hundreds of nanowatts could be observed [56].

V. OUTLOOK

After these encouraging results in different crucial fields of optoelectronics, we believe that GaN is the material of choice for several emerging applications. Ultra-high-speed devices such as photodetectors or optical switches might soon play an important role in modern telecommunication systems. Since GaN and AlN are chemically inert and mechanically very robust materials, applications under harsh environmental conditions might become especially appealing for optoelectronic components fabricated from AlN/GaN. It is clear, too, that the development toward higher frequencies will continue until the intrinsic physical limits of the material system are reached. At this moment, there is however no sign that this will happen in the near future. Based on the short intersubband lifetime and the absence of an internal parasitic capacitance, cutoff frequencies on the order of several hundred gigahertz are expected. In this same context, detectors with larger asymmetries—be it with asymmetric potentials and physical transport of electrons or using optical rectification employing additional electronic levels—will be promising candidates for improved responsivity and higher speed. Along the same lines, highly sensitive photodetectors with quantum dot active regions will be further developed. Preliminary experiments have shown already a responsivity improvement of a factor of 60.

An application with considerable potential for future development is the use of GaN saturable absorbers as mode-locking elements in solid-state laser cavities. Since GaN is chemically robust, it will be able to withstand the required high pump powers. The search for the ideal device geometry and the suitable place within the laser cavity are some of the important points to be resolved here. Another interesting potential direction for novel devices might be the use of nonpolar III-nitrides. The absence of internal polarization fields would lift one of the big uncertainties and would clearly constitute an advantage for intersubband device design. As far as intersubband light emitters are concerned, the general prognosis might be somewhat less enthusiastic than for detectors and switches. At this point, neither electroluminescence nor optically pumped lasing has been demonstrated. Nevertheless, optically pumped luminescence has been observed. Further progress will certainly be made using device configurations involving distributed feedback, more sophisticated pump geometries, and quantum-mechanical optimizations such as the Raman laser. The most important point for future QC laser development, however, is the use of high-quality dislocation-free GaN bulk crystals. The unambiguous demonstration of a resonant tunneling diode on such material will be a highly important milestone that would give us confidence that the necessary material quality for laser development is now reached. Without this demonstration, transport measurements in
any III-nitride device always suffer from a lack of clarity and meaningfulness. Like 17 years ago with the violet-blue laser diodes based on GaN, it is expected that a particularly annoying bottleneck present in the “real world” will be able to trigger a dedicated effort in the development of nitride intersubband devices for practical applications. In 1992, this bottleneck was the short lifetime of blue lasers based on ZnSe. GaN was the only material capable of outperforming ZnSe-based devices, and it soon did.

REFERENCES


Hong Wu, photograph and biography not available at the time of publication.

William J. Schaff, (Member, IEEE) photograph and biography not available at the time of publication.

Jahan Dawlaty, photograph and biography not available at the time of publication.

Paul A. George, photograph and biography not available at the time of publication.

Lester F. Eastman, (Fellow, IEEE) photograph and biography not available at the time of publication.

Farhan Rana, photograph and biography not available at the time of publication.

Prem K. Kandaswamy, photograph and biography not available at the time of publication.

Fabien Guillot, photograph and biography not available at the time of publication.

Eva Monroy, photograph and biography not available at the time of publication.