1 Introduction

Spontaneous polarization exists along the (0001) (metal-polar) and (0001) (N-polar) directions of wurtzite III–V nitride semiconductor crystals. Strain-induced piezoelectric polarization can either add to, or subtract from the spontaneous polarization in heterostructures based on the composition and polarity. At a sharp heterojunction, the discontinuity in polarization leads to the appearance of an immobile polarization sheet charge (unit: C/cm²) of density:

$$\sigma_x = (P_1 - P_2) \cdot \hat{n},$$

(1)

where $P_1$, $P_2$ are the total (spontaneous + piezoelectric) polarization vectors across the junction, and $\hat{n}$ is the unit normal to the heterointerface plane. If the heterojunction is compositionally graded (say along the growth direction $z$) instead, Eq. (1) is modified to its 3D counterpart:

$$\rho_\tau(z) = -\nabla \cdot P(r) = -\frac{\partial P_z(z)}{\partial z},$$

(2)

where $\rho_\tau(z)$ is the immobile 3D bulk charge density (unit: C/cm³). The built-in real electric fields originating from the polarization charges, when combined with the quasi-electric fields at sharp and graded heterojunctions lead to a rich range of phenomena that impact every conceivable wurtzite nitride heterostructure device. This work reviews the existing body of knowledge on the device consequences of the high polarization fields, and outlines several future possibilities that exploit as-yet unrealized applications of polarization engineering. The materials challenges in realizing the proposed novel applications are mentioned. Though structures grown along semi-polar directions are not discussed explicitly, similar arguments are expected to apply.

2 Polarization in Al$_x$In$_y$Ga$_{1-x-y}$N alloys

Figure 1a shows the contours of constant polarization sheet charge at Al$_x$In$_y$Ga$_{1-x-y}$N/GaN heterojunctions grown coherently strained on (0001) oriented relaxed GaN substrates. Bowing parameters of $b = -1.0$, $-2.0$, $-4.0$ eV for AlGaN, InGaN, and InAlN ternaries are used, and quaternary material parameters are taken from Ref. [1]. We note that these are representative values, and in particular, parameters for InAlN are currently under investigation. The polarization coefficients used are taken from Ref. [1]. Recent investigations of the growth and applications of InAlN and InAlGaN layers lattice matched to GaN (falling on the red line) have yielded very encouraging results for both electronic and optical devices, greatly expanding the range of material choices available for III–V nitride devices. The rapidly developing growth technology – by both MOCVD and MBE makes the realization of quaternaries and InAlN ternaries with a range of compositions a tantalizing prospect.
of new applications. These materials offer a potential method of solving existing problems encountered in typical Al(Ga)N/GaN and In(Ga)N/GaN – containing heterostructures that form most devices today. To aid crystal growers, Al(Ga)N/GaN and In(Ga)N/GaN – containing heterostructures are yet to be investigated for E-Mode HEMTs, but will require the source/drain and access regions to be populated by carriers similar to what was done recently for such devices grown along the non-polar direction [5].

In optical devices, the role of polarization has been investigated in the active regions where the built-in electric field lowers the oscillator strength by reducing the electron–hole wavefunction overlap [6], the quantum-confined Stark effect (QCSE). Under high carrier injection conditions, the polarization field is partially screened, but it causes spectral shift of emission with bias, and may prove problematic for ultralow threshold light emitters. Though light emitting diodes and lasers over a large spectral window have been demonstrated with heterostructures grown along the polar directions, non- and semi-polar orientations are being intensively investigated recently [7].

4 Sharp heterojunctions

4.1 Tunable quasi band-offsets Consider an ultra-thin layer (thickness = t) of polarization-mismatched nitride layer inserted between thick GaN layers as shown in Fig. 2a. The dipole layer could be of wider, narrower, or the same bandgap as the cladding layers with proper choice of composition and polarity; the cases are indicated schematically in Fig. 2a. The polarization discontinuities of Eq. (1) result in a charge dipole at the heterojunction with sheet charges \( \pm \sigma_x \). The electric field in the thin layer due to the polarization charges is \( F_x = \sigma_x/\varepsilon_0 \), and the corresponding energy drop is \( \Delta E_x = qF_xt = q\sigma_x t/\varepsilon_0 \), where \( \varepsilon_0 \) is the dielectric constant of the semiconductor. For example, if \( t \approx 1 \) nm, and \( F_x \approx 1 \) MV/cm, then \( \Delta E_x \approx 0.1 \) eV, which is substantial. This is achievable with, for example (but not limited to), a thin layer of AlInN that is bandgap-matched to GaN, which we assume does not introduce a chemical band offset (this assumption will be relaxed later). The value of \( \Delta E_x \) can be much larger given the large polarization fields in nitrides.

Note that an ultrathin polarization dipole layer can be effectively transparent to electron transport in the vertical
direction, and therefore the dipole layer effectively mimics a band offset. The magnitude of the band offset is equal in both the conduction and valence bands ($\Delta E_{c\tau} = \Delta E_{v\tau} = \Delta E_{\pi}$) and is therefore of the staggered type if the polarization-induced band offset $\Delta E_{\pi} < E_g$ is less than the bandgap ($E_g$) of the thicker cladding layers (Fig. 2b). Since such a band offset changes as $t$ and $\sigma_\tau$ are varied, it is tunable; indeed, it may even be possible to achieve an effectively broken gap offset if $\Delta E_{\pi} > E_g$ as shown in Fig. 2c. The origin of this polarization-dipole induced band offset is a real electric field and therefore it acts equally on electrons and holes, giving the symmetry in the CB and VB offsets. This is in stark contrast to a quasi-electric field at heterojunctions, which is born out of changes in the chemical nature of constituent atoms and thereby can act very differently on electrons and holes leading to $\Delta E_c \neq \Delta E_v$ in general, and resulting in all three possible (straddling, staggered, and broken) band offsets [8].

Such artificial band-offsets have been investigated in GaN/AlN/GaN junctions [9, 10] and in GaN/InGaN/GaN junctions [11]. The ultrathin layer for obtaining a polarization induced band offset may be of a smaller bandgap than the cladding layers for high transparency to charge transport in the growth direction, in which case it may require the expulsion of bound states in the thin asymmetric quantum wells to prevent charge trapping. The ultrathin dipole layer could also be of a larger bandgap which will reduce the current density by requiring tunneling/thermionic emission. By the proper choice of compositions, such polarization induced band offsets can be introduced in conjunction with real band offsets to increase, reduce, or cancel $\Delta E_c$ (or $\Delta E_v$) selectively. The added degree of freedom provided by the large polarization in the III–V nitrides can prove to be a powerful device design tool.

### 4.2 Coupled electron–hole gases

Consider an AlN layer of thickness $t$ embedded deep inside an undoped GaN bulk layer shown schematically in Fig. 3a. A critical thickness $t_{cr} = E_g/F_{\pi}$ is easily identified – if the dipole layer is thicker, the potential drop exceeds the bandgap of the cladding layers ($\Delta E_{\pi} > E_g$), resulting in an effectively broken-gap offset, whereby the valence band electron states are aligned with the empty conduction band states on the other side. In a nominally undoped structure, assisted by the large polarization field, there is a finite probability for valence band electrons on the left to tunnel into the empty conduction band states on the right, leaving behind holes as indicated in Fig. 3a. A rough estimate of the “time constant” involved in this process can be obtained. The effective velocity of carriers in the valence band is $v \sim h\sigma_{\pi}/m_0$, where $\sigma_{\pi} \sim \pi l_0$ is a reciprocal lattice vector, $l_0$ is a lattice constant, and $m_0$ is the free electron mass. An electron confined to a region of length $l_0$ collides the forbidden gap barrier at a rate $v \sim \hbar l_0$, and the probability of tunneling is $T \approx \exp(-2S)$, where $S = \int \kappa(z) \, dz$ is the action, with $\kappa^2(z) = 2m[V(z) - E]/\hbar^2$. The time for “escape” is estimated as $\tau \sim 1/\gamma T$, which for relevant GaN parameters evaluates to the order of seconds.

After such a process of tunneling relaxation has occurred, the potential difference far from the junction region is $\Delta \phi_\pi = (\sigma_\pi - qn_s)/\epsilon_s$ as shown in the energy band diagram in Fig. 3b. Here $n_s$ is the sheet charge density (units: cm$^{-2}$) of a mobile electron gas, equal to the mobile hole charge density at the other interface. The mobile charge dipole forms in response to the potential drop, in an attempt to screen the polarization dipole $\sigma_\pi$. We note that charge neutrality requires the net charge to be 0, but the mobile charge dipole need not be equal to the polarization dipole. This novel mechanism of formation of mobile carriers may also be pictured as “modulation-doping.” However, for example, for the 2DEG, the source of electrons are not donor impurities in a wider bandgap region, but the valence band on the other side of the heterojunction barrier. This is a novel mechanism for creating mobile electrons and holes without impurity doping, and understanding it helps us exploit polarization in powerful ways.

When the 2DEG and 2DHG form as indicated in Fig. 3b, the Coulombic coupling between them can be tuned by changing $t$. A major challenge in observing such coupled electron–hole gas is forming separate electrical contacts to...
them. Such closely coupled 2DEG and 2DHG assemblies are desirable for fundamental studies of Coulomb drag, and with suitable choice of the materials (and their bandstructures) can enable tunneling field-effect transistors (TFETs) of both of the “classical” type [12], and perhaps also of the phase-coherent many-particle type [13].

4.3 Resonant- and Zener-tunneling structures

Tunneling structures are best understood in polar III–V nitride heterostructures by building upon the nominally undoped structure shown in Fig. 3a and b and discussed in the earlier section. When the structure is uniformly doped (say n-type), the energy band diagram is as shown in Fig. 3c. The heterojunction side that has a positive polarization sheet charge (+\(\sigma_{n}\)) sees an accumulation layer of electrons, whereas the other side, the role played by the mobile holes in the undoped structure is taken up by the immobile ionized donors. If two such barriers are arranged to enable resonant tunneling, polarization essentially increases the tunneling distance by the depletion region thickness and leads to a lowering of the net tunneling current drive over a distance by the depletion region thickness and leads to a polarization-free structure. This problem can be remedied by growing along non-polar orientations, or even when the growth is along the polar orientation, by choosing polarization-matched barriers (e.g., by using Al\(_{0.7}\)In\(_{0.3}\)N barriers as indicated in Fig. 3a, b, and d, though the barrier height may be restricted).

If instead of uniform doping, the polarization dipole layer is surrounded by a p–n junction, the extra potential drop due to the dipole can either assist, or oppose the natural built-in potential of the p–n junction. In the case where it opposes the built-in potential, it causes an extension of the depletion regions and is akin to a built-in reverse bias, and can be used for tuning the turn-on voltage and/or the reverse-bias breakdown voltage. The opposite case is when the polarization dipole assists the dipole formed by ionized dopants by virtue of being in the same direction. Is such a case, the polarization dipole shrinks the depletion region thickness and is similar to a built-in forward bias. If the polarization charges are greater than necessary for dropping the entire built-in potential of the control p–n junction (\(|\sigma_{n}| > qN_{A}N_{D} = qN_{A}N_{Vp}\)), then accumulation regions form at the two heterojunctions, and the effective band offset between the two sides is now similar to a broken-gap heterojunction. Such situations are shown in the energy band diagrams of Fig. 3e and f. This arrangement enables Zener tunneling of electrons from the valence band of the p-side to the empty conduction band states on the n-side under reverse bias voltage, whereas no current flows under small forward bias since the valence band states are aligned with the bandgap of the n-side. Thus the structure acts as a “backward diode.” Such polarization-enabled Zener tunnel junctions have been recently experimentally demonstrated using GaN/Al\(_{0.55}\)In\(_{0.45}\)N/Al\(_{0.55}\)In\(_{0.45}\)N heterostructures [16], and carefully modeled [17]. They hold much promise for future applications in wide-bandgap semiconductor devices.

Such interband tunneling diodes in traditional narrow-gap semiconductors are achievable primarily by heavy impurity doping with associated problems of discrete dopant effects and bandgap shrinkage. In nitride heterostructures the role of dopants is taken up by polarization charges that are confined to atomic planes, and can be far higher than what can be achieved by impurity doping. Furthermore, the dipole layer may be chosen to be a narrower bandgap than the surrounding layers as in Fig. 3f, and thus the choice of polarization and composition offers a clear path for boosting the current drive. Such polarization-induced Zener tunnel junctions can prove to be useful in zero-bias rectifiers, for electrically connecting multijunction solar cells and multicolor light emitters [15], and also possibly in TFETs in the future.

The large built-in potential drop can enable tunneling-enhanced ohmic contacts. Low-resistance ohmic contacts to p-type wide-bandgaps has proven difficult due to the unavailability of very high work-function (\(q\phi_{M}\)) metals. The resulting misalignment of the Fermi level of the metal with the valence band results in high barriers (\(\phi_{M} = q\phi_{D} + E_{F} - q\phi_{M}\)). This barrier can be lowered by an amount \(\Delta E_{n} = qF_{n}t\) by introducing a polarization dipole layer. The high electric field, and a narrower bandgap dipole layer can enable low-resistance contacts that are also reasonably optically transparent, as proposed in Ref. [18].

4.4 Polarization-balanced electronic devices

The precise equality of the positive and negative components in a polarization dipole at symmetric nitride heterojunctions has led to the concept of “Natural Super-Junctions” (NSJs) for high-voltage applications. Such devices when made using p–n junctions (e.g., in silicon) by impurity doping place stringent requirements on the doping densities of donors and acceptors so that the vertical field is entirely supported by ionized dopants. This requirement is naturally met in III–V nitride heterojunctions, and in addition the high density (and presumably high mobility) polarization-induced 2DEGs and 2DHGs can contribute to low on resistances [19]. Back-barriers for strong electron confinement is required for highly scaled GaN HEMTs. Typical double heterostructure designs employ AlGaN back barriers for metal-polar devices, which causes a degree of reverse polarization and depletion of the 2DEG densities, causing higher sheet resistances. One possibility for reducing this loss of charge is by using an AlInN layer that is polarization-matched to GaN.

4.5 Polarization-balanced optical devices

Though growth along non-polar or semi-polar directions helps reduce QCSE based reduction of oscillator strengths by removing the built-in polarization fields, it is possible to do so by growing along the polar directions by choosing polarization-matched compositions [20]. We discuss one possibility as an example, along with a few device electrostatics arguments. Consider the case of a bandgap-matched GaN-Al\(_{0.65}\)In\(_{0.35}\)N heterostructure grown along the (0001) direction. Polarization sheet charges ±\(\sigma_{n}\) form at the
two heterojunctions without a band offset as indicated schematically in Fig. 4a and b. If the entire structure is uniformly doped with donor density $N_d$, the polarization charges cause depletion and accumulation of mobile electrons as indicated. The peak electric fields at the heterojunction is $F_{\text{max}} = \sigma_d/2\varepsilon_0 = F_a/2$ where $\varepsilon_0$ is the dielectric constant. The potential barrier height at the depleted junction is $\phi_d = \sigma_d^2/8\varepsilon_0 qN_d$, and the potential “well” depth in the accumulation side is $\phi_a \approx F_{\text{max}} L_d$, where $L_d = \sqrt{\varepsilon_0 kT/q^2N_d}$ is the Debye length. Accumulation regions do not impede the motion of free carriers; thus, if a p–n junction structure as shown in Fig. 4c is grown, accumulation regions form at both the heterojunctions. If one chooses InGaN composition on the same polarization contour as the AlInN region for quantum wells, the active region is freed of polarization fields. The material choices are easily seen in Fig. 1a or b: move from GaN $\rightarrow$ InAlN along the bandgap-matched line, then move to the InGaN composition staying along the constant polarization contour (dashed line). The materials challenges in doing so are indeed high, and it may be more prudent to free active regions from polarization by simply growing along non-polar orientations. But the fundamental physics does not rule out the possibility of achieving polarization-free active regions even if the underlying structure is inherently polar. In cases where the injection efficiency of holes (and/or electrons) are low due to deep dopant levels, similar polarization-matching approaches could prove highly attractive when combined with the added advantage of higher injection efficiency in structures grown along the polar orientation. The possibility of polarization-enhanced p-type doping can be integrated into the device design along polar directions in graded bandgap heterostructures, which is discussed next.

5 Graded heterojunctions When one or both polar heterojunctions are compositionally graded, a number of additional utilities of polarization emerge. Consider the canonical structure of the last section as shown schematically in Fig. 5a, and assume we keep one junction abrupt and compositionally grade the other as shown in Fig. 5b. In that case, the fixed bulk polarization charge density in the graded region is given by Eq. (2): $\rho_p = -\nabla \cdot \mathbf{P}(r)$. In response to the macroscopic electric field created by the fixed polarization charge, neutralizing mobile charges on the graded side also spread from a 2D form to a 3D bulk form to screen/cancel the electric field. We note here that the neutralizing charge may be localized in the presence of a high density of defects and traps, but can always be made mobile by compensation doping of the traps and defect states.

5.1 Optoelectronic devices: doping and contacts Highly conductive p-type layers have been difficult to achieve in GaN based heterostructures owing to the high activation energy of Mg acceptors ($E_A \sim 180$ meV), leading to associated problems with hole injection efficiency in optical devices. In addition, both Mg acceptor levels as well as Si donor levels ($E_D$) become deeper with increasing bandgap of Al(Ga)N layers that are useful for UV optoelectronics [21, 22]. When the activation energies are much larger than the thermal energy ($E_A, E_D >> kT$), the low thermal ionization efficiency results in high resistivity layers. The polarization-induced charge in short-period superlattices offers a method to field-ionize these deep dopants, and extract mobile carriers. However, in such structures there still exist band offsets that impede carrier motion in the direction of most interest. Interestingly, just as a buried polarization dipole results in a mobile electron–hole gas without impurity doping, a defect/trap-free compositionally graded layer should necessarily result in mobile carriers to screen the macroscopic electric field. The role of dopants ($N_A$) in compositionally graded layers is reduced to compensating the trap states ($N_T$) that may localize the mobile charges, and maintaining the Fermi level close to the valence band (for p-type) or the conduction band (for n-type). Thus the mobile carrier concentration $p$ is determined not by the impurity doping, but by the polarization charge density by the charge neutrality condition:

$$\rho_p = -\nabla \cdot \mathbf{P}(r).$$
\[ \rho_n + q N^A_n \approx q(p + N^A_p) \], whereby the hole density is \( p \approx \rho_n/q + (N^A_n - N^A_p) \). If complete compensation of the traps are achieved, \( p \approx \rho_d/q \). This form of polarization-induced doping was demonstrated for conductive n-type layers on both metal [23] and N-face heterostructures [24], and very recently for p-type layers, and has been incorporated into prototype UV LED structures [25]. Figure 5c shows an embodiment of a polarization-doped p–n junction with \( n \approx \rho_n/q + (N^A_n - N^A_p) \).

A number of added features of such polarization-induced doping may be identified. The graded bandgap layer can be of a larger bandgap and thus optically transparent for the active layer wavelengths. It can enable natural optical mode confinement in the correct orientation, similar to graded-index separate confinement heterostructure (GRINSCH) lasers. If the graded layer is p-type (a similar argument holds for n-type), \( E_g - E_v \) is nearly constant, and the entire difference in the bandgap appears in the conduction band \( E_c(z_1) - E_c(z_2) = \Delta E_c \), leading to a quasi-electric field for electrons. With the proper magnitude and orientation, this barrier can act as an electron blocking layer as shown in Fig. 5c. In abrupt heterostructures with straddling band offsets, a conduction band offset \( \Delta E_c \) based electron blocking layer necessarily comes with a valence band offset \( \Delta E_v = E_v(z_1) - E_v(z_2) \), which results in hole blocking to a certain degree. The only way to obtain a barrier for electrons and not for holes is to use a p-type doped graded-bandgap layer, where the entire bandgap difference appears in the conduction band, whereas the valence band is left without any barriers impeding hole flow. This feature is well known in heterostructure bipolar transistors (HBTs), where such doping can prove specially attractive.

5.2 Heterostructure bipolar transistors

The high resistance of p-type layers has proven to be the bottleneck in achieving high-performance HBTs using GaN. Significant improvements in GaN HBTs have been made recently by incorporating narrower bandgap InGaN base layers. The polarization-induced p-type doped graded gap layers described in the last section can help accelerate the progress in GaN HBTs [26]. Such p-type layers allow for wide-bandgap base and collector layers to take full advantage of the high breakdown of wide bandgaps, and simultaneously offer large built-in quasi-electric fields in the base to reduce the base transit time and improve the speed of such devices.

6 Summary

In summary, the recent expansion of GaN material research into quaternaries opens up new opportunities for polarization-based heterostructure device design. A sampling was discussed here, and is expected to be extended in the future.

Acknowledgements

The authors acknowledge financial assistance from DARPA, ONR, NSF, and AFOSR.

References

[1] F. Bernardini, in: Nitride Semiconductor Devices: Principles and Simulation, edited by J. Piprek (Wiley-VCH, Berlin, 2007), p. 49.
[2] H. Morkoc and J. Leach, in: Polarization Effects in Semiconductors, edited by C. Wood and D. Jena (Springer, Berlin, 2008), p. 373.
[3] M. Gonschorek, J.-F. Carlin, E. Feltin, M. A. Py, and N. Grandjean, Appl. Phys. Lett. 89, 062106 (2006).
[4] T. Lim, R. Aidam, L. Kirste, P. Waltereit, R. Quay, S. Muller, and O. Ambacher, Appl. Phys. Lett. 96, 252108 (2010).
[5] T. Fujiwara, S. Keller, J. S. Speck, S. P. DenBaars, and U. K. Mishra, Appl. Phys. Express 3, 101002 (2010).
[6] R. Butte and N. Grandjean, in: Polarization Effects in Semiconductors, edited by C. Wood and D. Jena (Springer, Berlin, 2008), p. 467.
[7] J. S. Speck and S. F. Chichibu, MRS Bull. 34, 304 (2009).
[8] H. Kroemer, Rev. Mod. Phys. 73, 783 (2001).
[9] S. Keller, S. Heikman, L. Shen, I. P. Smorchkova, S. P. DenBaars, and U. K. Mishra, Appl. Phys. Lett. 80, 4387 (2002).
[10] Y. Cao, T. Zimmermann, H. Xing, and D. Jena, Appl. Phys. Lett. 96, 042102 (2010).
[11] T. Palacios, A. Chakraborty, S. Heikman, S. Keller, S. P. DenBaars, and U. K. Mishra, IEEE Electron. Dev. Lett. 27, 13 (2006).
[12] Q. Zhang, W. Zhao, and A. Seabaugh, IEEE Electron. Dev. Lett. 27, 297 (2006).
[13] S. K. Banerjee, L. F. Register, E. Tutuc, D. Reddy, and A. H. MacDonald, IEEE Electron. Dev. Lett. 30, 2 (2009).
[14] J. Simon, Z. Zhang, K. Goodman, H. Xing, T. Kosel, P. Fay, and D. Jena, Phys. Rev. Lett. 103, 026801 (2009).
[15] M. Grundmann and U. K. Mishra, Phys. Status Solidi C 4, 2830 (2007).
[16] S. Krishnamoorthy, S. N. Nath, F. Akyol, P. S. Park, M. Esposto, and S. Rajan, Appl. Phys. Lett. 97, 035302 (2010).
[17] M. Schubert, Phys. Rev. B 81, 035303 (2010).
[18] M. Schubert, Appl. Phys. Lett. 96, 031102 (2010).
[19] H. Ishida, D. Shibata, H. Matsuou, M. Yanagihara, Y. Uemoto, T. Ueda, T. Tanaka, and D. Ueda, IEEE IEDM Tech. Dig. 1–4 (2008).
[20] M. Schubert, J. Xu, J. K. Kim, E. F. Schubert, M. H. Kim, S. Yoon, S. M. Lee, C. Sone, T. Sakong, and Y. Park, Appl. Phys. Lett. 93, 041102 (2008).
[21] Y. Taniyasu, M. Kasu, and T. Makimoto, Nature 414, 325 (2001).
[22] A. Khan, K. Balakrishnan, and T. Katona, Nat. Photonics 2, 77 (2008).
[23] D. Jena, S. Heikman, D. Green, D. Buttari, R. Coffie, H. Xing, S. Keller, S. DenBaars, J. S. Speck, U. K. Mishra, and I. P. Smorchkova, Appl. Phys. Lett. 81, 4395 (2002).
[24] S. Dasgupta, Nidi, D. F. Brown, F. Wu, S. Keller, J. S. Speck, and U. K. Mishra, Appl. Phys. Lett. 96, 143504 (2010).
[25] J. Simon, V. Protasenko, C. Lian, H. Xing, and D. Jena, Science 327, 5691 (2010).
[26] P. M. Asbeck, E. T. Yu, S. S. Lau, W. Sun, X. Dang, and C. Shi, Solid State Electron. 44, 211 (2000).