A Perspective of GaAs$_{1-x}$N$_x$ and GaP$_x$N$_{1-x}$ as Heavily Doped Semiconductors

A. Mascarenhas$^1$), Yong Zhang, and M. J. Seong

National Renewable Energy Laboratory, 1617 Cole Blvd. Golden, CO 80401, USA

(Received August 31, 2001; accepted September 1, 2001)

Subject classification: 71.55.Eq; 72.20.Jv; 78.30.Fs; 78.55.Cr; S7.14

The behavior of GaAs$_{1-x}$N$_x$ and GaP$_x$N$_{1-x}$ is contrasted with that of Al$_x$Ga$_{1-x}$As and In$_x$Ga$_{1-x}$As with respect to irregular and regular alloy behavior. It is proposed that GaAs$_{1-x}$N$_x$ and GaP$_x$N$_{1-x}$ behave as heavily nitrogen doped semiconductors rather than dilute nitride alloys and that their abnormal or irregular alloy behavior is associated with impurity band formation that manifests itself in the giant bowing and poor transport properties characteristic of these materials.

Introduction

Semiconductor heterostructures constitute the building blocks of several electronic and photonic devices such as solar cells, diode lasers, light emitting diodes, and heterojunction bipolar transistors. The heterostructure components and their stacking geometry are often utilized for tailoring features for improving device performance. Although the choice of semiconductors for synthesizing a heterostructure for a particular device application is guided by the specific electrical or optical design requirements, this choice is also guided by the ability to grow the required stack of semiconductor layers epitaxially. Thus the semiconductor lattice size in addition to its electronic bandstructure is an important criterion for heterostructure design. The requirement of a specified bandstructure feature and lattice size is most often fulfilled by the use of substitutional alloys, thus circumventing the discreet and finite repertoire of available elemental and compound semiconductors. They have thus been the subject of extensive studies as regards their alloy thermodynamics and electronic properties. Recently, the phenomenon of giant band gap “bowing” that has been observed in several III–V dilute nitride alloys offers the exciting promise of increasing the flexibility in choice of semiconductor band gaps available with specified lattice constants. However, the poor electrical transport properties that these materials exhibit, seriously limit their usefulness. In this paper, these limitations will be discussed from the perspective of impurity band phenomena.

Background

The lattice size of a substitutional semiconductor alloy A$_{1-x}$B$_x$C has experimentally been found to be well approximated by the concentration weighted average of that of its constituents AC and BC, which is usually referred to as Vegard’s rule or law [1]. Also, the band gap is found to be close to but generally lower than the concentration weighted average of that of the constituent compounds and this deviation well described in terms of a bowing coefficient which is usually quite small (about 1 eV) [2]. Thus the band gap and lattice size of the alloy covers the range of values between that for the two end point constituent compounds provided

$^1$) Corresponding author; Phone: +01 303 384 6608; Fax: +01 303 384 6481; e-mail: amascar@nrel.gov
that they are not immiscible in this range. But in spite of the variety of band gaps and lattice sizes that are made possible by alloying semiconductors, there have been many situations encountered over the past few decades, where although it is possible to obtain an alloy having the desired lattice constant, the alloy does not have the required band gap. Such frustrations impose severe penalties in device design. An example of this is the technologically important field of vertical cavity surface emitting lasers (VCSELs) used for fiber optic communications [3]. The difficulty with Bragg mirror stacks of quaternary alloys grown on InP substrates lies in their poor thermal conductivity. Although high reflectivity AlAs/GaAs Bragg mirror stacks can readily be grown on GaAs substrates, it is not possible to find an alloy for the active region of the VCSEL, with a band gap near 0.8 or 0.95 eV and which is lattice matched to GaAs. To grow such VCSELs, it has thus been necessary to resort to complicated techniques which often involve “lift-off” and “fusion bonding” steps in the laser fabrication process [3]. Another alloy constraint example is the quadruple junction solar cell where the optimal heterostructure for realizing a 40% efficient solar cell would be comprised of light absorbing semiconductor active regions grown epitaxially with the sequence of band gaps 0.67 eV/1.05 eV/1.42 eV/1.9 eV on a Ge substrate [4]. GaInP$_2$, GaAs and Ge are ideally suited lattice matched semiconductors for the top, next to top, and bottom cells, but there existed no material lattice matched to GaAs with a band gap of 1.05 eV until quite recently [4].

Giant Bandgap “Bowling” Phenomena About a decade ago, Weyers et al. [5] succeeded in incorporating almost 1% N into GaAs using OMVPE (organo metallic vapor phase epitaxy) techniques and observed a surprising lowering of the band gap by about 140 meV which is contrary to what is expected by Vegard’s rule. This rekindled interest in a subject that is over three decades old, namely, the N isoelectronic trap. Research from the 60’s to the 80’s had established that N behaves as an isoelectronic trap in GaP and GaAs [6–8]. Thermodynamic calculations [9] predicted that N was insoluble in both these materials and in the early work on bulk crystals and LPE material it was only possible to achieve N doping levels less than $10^{19}$ which allowed for the identification of N induced trap levels in these semiconductors [6–8] as shown in Fig. 1. The ability to incorporate over an order of magnitude more N (>1%) into these semiconductors is attributed to the use of non-equilibrium growth techniques such as OMVPE and MBE (Molecular Beam Epitaxy). The experiments of Weyers et al. [10] were soon repeated by several groups around the world, confirming that the band gap of GaAs was lowered by about 180 meV with just 1% of N incorporation into GaAs, and by 400meV with just 3% N. GaAs$_{1-x}$N$_x$ began to be viewed as an alloy and the anomalous large lowering of its band gap to be described as a giant band gap “bowing” (the bowing parameter varies with x but is <20 eV) [11]. These results were followed by a flurry of research activity after the suggestion by Kondow et al. [12] about the possibility of realizing 1.3 μm lasers using GaAs$_{1-x}$N$_x$ active regions grown on GaAs substrates. At around the same time it became evident that GaAs$_{1-x}$N$_x$ could also be used to provide the 1.05 eV band gap material lattice matched to GaAs for realizing the 40% efficient quadruple junction solar cell and several research groups raced towards this goal [4]. It thus appeared that the anomalous and giant “bowing” of the band gap in GaAs$_{1-x}$N$_x$ offered a way out of the semiconductor alloy constraints that limited the design of some technological important devices.
Soon it was evident, however, that the giant band gap “bowing” in GaAs$_{1-x}$N$_x$ came with a caveat. The electron mobilities in this material were severely degraded and several efforts to improve this using rapid thermal annealing resulted only in marginal improvements, thus dampening the initial excitement of using this material for both solar cells and lasers [13]. In the remainder of this paper, the exciting research during the past few years on understanding the physical mechanisms underlying the phenomenon of the anomalous giant band gap “bowing” will be discussed from the viewpoint of determining whether the limitations of the GaAs$_{1-x}$N$_x$ system can be overcome.

**Heavy Doping and Impurity Bands** Using photoluminescence studies at low temperatures and high pressures on GaAs samples doped with 0.01% N, Wolford et al. [7] and Liu et al. [8] had determined that the isolated N impurity forms a resonant level 150–180 meV above the conduction band edge of GaAs. Almost a decade later, in an effort to understand the origin of the giant “bowing” phenomenon, Perkins et al. [14] used electro-modulated reflectivity studies to investigate the giant “bowing” phenomenon in heavily N doped GaAs, and directly observed a level (denoted as $E_+$) above the conduction band edge whose existence was inferred by Shan et al. [15] using photoluminescence studies under high pressure. Employing a simple two level repulsion model, Shan et al. assumed that the giant band gap “bowing” in GaAs$_{1-x}$N$_x$ arose from level repulsion between the isolated N resonant level and the GaAs conduction band edge. But soon thereafter, Zhang et al. [16] demonstrated a very unusual behavior for the conduction band effective mass in GaAs:N. Since the carrier lifetimes in GaAs:N are too short to be able to measure the effective mass with conventional techniques that involve cyclotron resonance, this was determined by fabricating GaAs/GaAs:N quantum wells with several well widths for each value of N doping, and inferring the effective mass from the confinement induced shifts in the QW ground state transition energies. The results obtained revealed that the conduction band effective mass exhibited a
very abrupt increase by almost an order of magnitude on going from 0.01 to 0.1% N and then gradually decreased towards the normal GaAs value as the N doping approached a value of about 3% as shown in Fig. 2. This surprising behavior was a clue to the presence of an impurity band in GaAs:N induced by the heavy N doping. Subsequent low temperature photoluminescence and selective excitation of photoluminescence studies by Zhang et al. [17] showed the evolution of the discreet N trap levels in GaAs into a N induced impurity band. Recently, peculiarities of the N resonant levels in GaAs:N have been dramatically demonstrated in resonant Raman scattering studies by Cheong et al. [18]

The physics of the N isoelectronic trap in GaP and GaAs$_{1-x}$N$_x$ was extensively studied in the 1960’s and 70’s because of the potential for use of these materials for light emitting diodes (LED’s), however, because of the N solubility problem analogous to that for GaAs:N, the doping levels studied were limited to <0.1%. This work has recently been reviewed by Zhang and Ge [19]. In the early 1990’s it became possible to incorporate much larger amounts of N into GaP using non-equilibrium growth techniques such as MBE, and Baillargeon et al. [20] and Liu et al. [21] revealed the existence of a large red shift of the absorption edge or a band gap reduction in GaP doped with about 1% N, suggesting the formation of an N induced impurity band in this material as well. Almost a decade later, research by Xin et al. [22] showed that the absorption edge in GaP:N for N concentrations exceeding 1% appeared to have the energy dependence characteristic of a direct gap semiconductor ($\sqrt{E-E_g}$ dependence), renewed the interest in this material for LED technologies. The evolution of the conduction band edge in GaP:N from the merger of bands formed from the N trap levels in this material is quite apparent in Fig. 3.

As mentioned earlier, the caveat with both the nitrogen doped GaAs and GaP was the poor carrier mobilities. Both these are a natural consequence of the peculiar nature of the conduction band minimum in GaAs:N and GaP:N. The conduction band edge in these nitrogen doped materials evolves out of the formation of an impurity band. As the conduction band effective mass studies in GaAs have indicated, the increase in curvature (see Fig. 4) of the impurity band with increased N doping causes the effective
mass to decrease from its abnormally high value in the flat impurity band that is characteristic of low N doping concentrations, and that is reminiscent of hopping transport between impurity levels that overlap only weakly. The impurity bands which evolve from the NN trap levels, broaden with heavy nitrogen doping and merge with the bottom of the conduction band, thus giving rise to the phenomenon of giant band gap “bowing”. A question that thus emerges in this unusual situation, is whether one should view GaAs$_{1-x}$N$_x$ as an abnormal alloy or rather as a heavily doped semiconductor GaAs:N [16, 17, 23].

Irregular Alloys When two semiconductors AC and BC are mixed, this typically results in the formation of a disordered alloy A$_x$B$_{1-x}$C if the physical properties of A and B do not differ greatly from each other, and in that case the properties of the alloy change smoothly from those of BC to those of AC as $x$ is changed from 0 to 1. Such is the case for Al$_x$Ga$_{1-x}$As and In$_x$Ga$_{1-x}$As for example. The semiconductors AC and BC must of course be miscible for a range of $x$ which is mostly true when the properties of A and B do not differ greatly. This contrasts
with the n or p-type charge doping of semiconductors, where the solubility of the donor or acceptor in the host is limited (as for e.g. GaAs:Si and GaAs:Zn). In some situations the differences between alloys and doped semiconductors is not so clear. Heavy n-type doping of the order of $10^{19}$ (or 0.1% dopant concentration) causes a nearly 200 meV band gap reduction in GaAs which results from impurity band formation [24], whereas, a 0.1% N doping in GaAs results in a band gap reduction of less than 20 meV [10]. However, one does not speak of GaAs:Si as an alloy and so the question arises as to why is GaAs:N being referred to as a GaAs$_{1-x}$N$_x$ alloy? The reason for this is that N is an isoelectronic impurity in III–V alloys and thus does not result in charge doping. There are two types of isoelectronic impurities: those that do not give rise to bound states (such as GaAs:Al or GaAs:In) and those that do give rise to bound states in the host. If the isoelectronic impurity generates bound states located in the band gap (either through isolated centers or pairs) such as is the case for GaAs:N and GaP:N, then with heavy doping the impurity levels associated with these bound states evolve into impurity bands that broaden and merge with the conduction band edge, and this merger gets manifested as a giant band gap “bowing”. The formation of an impurity band in heavily n or p-type doped semiconductors has been well studied together with its associated Mott metal-insulator transition [25]. Although the formation of impurity bands in heavily isoelectronically doped alloys like GaAs:N and GaP:N is relatively new, the phenomenon was actually observed in the CdS$_{1-x}$Te$_x$ system for concentrations $x > 10^{-4}$ almost thirty five years ago [26]. More insight into the common physical relationships between alloys, heavily charge doped semiconductors and heavily isoelectronically doped semiconductors is provided by a scaling rule recently observed by Zhang et al. [23]. As shown in Fig. 5, the band gap reduction in a doped semiconductor

![Graph](image)

Fig. 5. Comparison of band gap reduction as a function of impurity concentration for three typical systems. Data for a) GaAs:Si (n-type doping) are from Ref. [24], b) GaAs:In (regular alloy) are from Laurenti et al. [32], and c) GaAs:N are from Ref. [23]
is observed to follow the scaling rule
\[ \delta E_g(x) = \beta x^a, \]
where \( x \) is the mole fraction of the dopant. For alloys like In\(_x\)Ga\(_{1-x}\)As, the scaling exponent \( a \) is very close to unity. For heavily doped p-type semiconductors \( a \approx 1/3 \). The physics underlying the \( x^{1/3} \) scaling rule is simply that the bandwidth of the impurity band or the band gap reduction is proportional to the electron–electron interaction, and this interaction is proportional to the average impurity separation. For isoelectronically doped GaAs:N the scaling exponent \( a \approx 2/3 \). The significance of the scaling exponent \( a \) being close to 2/3 lies in that it confirms that the band gap reduction in heavily doped GaAs:N is primarily due to the formation of an impurity band associated with nitrogen pair bound states. This is the primary reason for the irregular or abnormal behavior of these alloys as will be discussed below.

**Physics of Isoelectronic Traps** Because of the difference in valence between the dopant atom and the host atom that it replaces, a non iso-electronic donor (acceptor) atom donates an electron (hole) to the conduction (valence) band of the host crystal. The Coulomb potential of the resulting ionized donor (acceptor) atom varies with distance as \( r^{-1} \) and generates a shallow donor (acceptor) bound state. In contrast, for isoelectronic traps that are generated by isoelectronic impurities such as N in GaAs or GaP, it is the difference in electronegativity, size, and pseudopotential between the isoelectronic impurity and the host atom it replaces that generates the trap state \[27\]. Such traps are characterized by a potential that varies with distance much faster than \( r^{-1} \) \[28\]. The potential well created by the isoelectronic trap is therefore much steeper than that created by the non-isoelectronic donor (acceptor) and because of this an electron (hole) trapped around the isoelectronic impurity atom is localized much more tightly around it. This spatial localization of electrons (holes) by isoelectronic traps smears out the electronic eigenstates in \( k \)-space causing them to be delocalized in the Brillouin zone. This proves advantageous in allowing radiative transitions from these states to the band edges, thus enabling light emission from indirect gap semiconductors like GaP. As discussed above, heavy N doping in GaP and GaAs leads to impurity band formation, red-shifts in the photoluminescence, and the giant band gap “bowing” phenomenon. However, the spatial localization around the isoelectronic traps that generate the impurity bands is precisely what disadvantageously affects the carrier mobility. This is the caveat with heavily isoelectronically doped semiconductors wherein the properties of the resulting alloy are irregular. Evidently, the very success in incorporating large amounts of insoluble isoelectronic dopants using non-equilibrium growth techniques is what leads to “irregular alloy” behavior.

**Resonance Raman Studies** A powerful technique for probing the structure of the impurity bands is resonant Raman scattering, dealing with the interaction between lattice vibrations and the intermediate electronic states involved. If the intermediate electronic state is a Bloch state with a well-defined momentum \( \mathbf{k} \), only zone-center (\( \Gamma \)) phonons with momentum \( \mathbf{q} = 0 \) are involved in the light scattering process, in the dipole approximation, and the phonon line shape is not expected to change near the resonance. However, if the intermediate electronic state originates from a strongly localized deep-impurity level so that its wavefunction is delocalized in \( k \)-space, phonons with \( \mathbf{q} \neq 0 \)
become allowed near resonance and, as a result, the phonon line should show broadening near resonance. The changes in the Raman spectrum near resonance thus provide strong clues to the nature of the intermediate electronic state. The resonance Raman profile for the full width at half maximum (FWHM) of the LO(Γ) phonon and TO(Γ)/LO(Γ) intensity ratio for \( x = 0.62\% \) sample exhibits distinct double maxima, labelled \( e_\text{W} \) and \( E_\text{W} \), as illustrated in Fig. 6. The resonance profile of the LO(Γ) phonon intensity as a function of excitation energy also shows two maxima, labeled \( e_\text{i} \) and \( E_\text{i} \), located near \( E_0 + \Delta_0 \) and \( E_+ \), respectively, as displayed in the inset of Fig. 6. It is important to note that the line width of the LO(Γ) phonon in pure GaAs was observed to be almost constant for the excitation energy from 1.55 to 2.0 eV, which encompassed both \( E_0 + \Delta_0 \) of pure GaAs and \( E_+ \) of GaAs\(_{1-x}\)N\(_x\) for \( x < 3\% \) at \( T = 300 \) K [18]. The similar anomalous double resonance maxima were more distinctly observed in the resonance profile of the line width of the 2LO(Γ) phonon for the same \( x = 0.62\% \) sample as shown in Fig. 7.

The distinct double resonance of the LO(Γ) phonon line width near \( E_0 + \Delta_0 \) and \( E_+ \) for the \( x = 0.62\% \) sample strongly suggest that the electronic states for \( E_0 + \Delta_0 \) transition have quite similar characteristics as those involved in \( E_+ \) transition. It is interesting to note that nitrogen incorporation hardly affects the valence band structure of GaAs whereas it changes the GaAs conduction bands drastically since all the nitrogen induced states (N\(_X\) and NN pairs) are located near the conduction band edge but are very far from the valence bands of GaAs. Since the valence bands are hardly perturbed upon dilute incorporation of N into GaAs, the anomalous RRS behavior near \( E_0 + \Delta_0 \) of GaAs\(_{1-x}\)N\(_x\) originates from the nitrogen-induced change in the nature of the conduc-
tion band edge at $\Gamma$. Therefore, the similarity in the line width resonance near the $E_0 + \Delta_0$ and $E_+$ transitions is a strong indication that the conduction band minimum around $\Gamma$ does contain significant contributions from various $k$-components other than $\Gamma$, just as for $E_+$. It is well known that NN pair states as well as the isolated nitrogen resonant state (NX) are strongly localized due to the short-range potential produced by the isoelectronic nitrogen atoms [7, 8]. The $E_+$ state, whose origin can be traced back to the isolated nitrogen resonant state (NX), contains in its wavefunction significant amount of $k$-components from other than the Brillouin-zone center ($\Gamma$) [29]. The significant non-$\Gamma$ components in the electronic states involved in the $E_0 + \Delta_0$ transition are expected to be introduced by NN-pair and cluster states as they form impurity bands that merge with GaAs conduction band minimum at $\Gamma$ [23, 30, 31].

**Conclusion** A description of the electronic properties that result from substitution of isovalent impurities which do not generate bound states in a host semiconductor such as GaAs or GaP can be obtained quite satisfactorily within the framework of conventional semiconductor alloy models. In contrast to this, the electronic properties that result from the substitution of non-isovalent dopant impurities which generate shallow bound states in a host are quite well described by conventional models for shallow dopants. In the case of substitutional N impurities in GaAs or GaP, N although isovalent generates bound states in the host, posing the question: should GaAs$_{1-x}$N$_x$ and GaP$_{1-x}$N$_x$ be described as conventional alloys or as heavily doped semiconductors? The abnormal or irregular behavior of these materials together with the results of a scaling rule are used to argue that they are better described as heavy isoelectronically doped semiconductors rather than as alloys.

**Acknowledgements** This research was supported by the U.S. Department of Energy under contract No. DE-AC36-83CH10093 and by the NREL DDRD program No. 0659.0004.

**References**


Y. Zhang, A. Mascarenhas, Y. Zhang, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz, Phys. Rev. Lett. 82, 3312 (1999);


