Chapter 4

AlAs/GaAs Double Barrier Resonant Tunneling Diodes

4.1 Introduction

The existence of d.c. negative resistance devices has been observed since the late 1950's in many different structures or devices that utilized thin anodic oxides [*Hic62*], degenerately doped p-n junctions (tunnel diodes) [*Esa58*], and heterojunction devices where quantum interference effects are utilized (double barrier resonant tunneling diodes and real space transfer devices) [*TsE73*] [*HeM79*] . The negative differential resistance (NDR) in the I-V characteristics of these devices has been used in many applications involving microwave/millemeter wave oscillators, high speed logic devices and switches.

Of these devices, the double barrier resonant tunneling diode (DBRTD) has gained the most attention in recent years with the improvements in molecular beam epitaxial (MBE) growth. Originally proposed by Tsu and Esaki [*TsE73*], the first experimental observation of resonant tunneling in a double barrier structure occurred in 1974 by Chang, Esaki and Tsu [*ChE74*] . It was only until 1983 that significant interest in the DBRTD occurred when the first high frequency experiments utilizing these structures was performed by Sollner and co-workers [*SoG83*]. Subsequently, a flood of research was performed on the DBRTD ranging from fundamental quantum transport theory, new materials and MBE growth related issues, high frequency oscillator applications, logic and switching applications, and three-terminal resonant tunneling structures [*CaK85*] [*ReF89*] [*SöA88*] [*SeC88*] [*TaM88*] [*SeK92*] . High frequency oscillations of up to 712 GHz have been observed in InAs/AlSb DBRTDs [*BrS91*] . Switching times as low as 2 picoseconds have been observed using electro-optic sampling on AlAs/GaAs DBRTDs [*WhM88*] .

In this chapter, some of the basic principles of the DBRTD will be presented. The current MBE growth and fabrication procedures for the AlAs/GaAs DBRTDs and Quantum Well Injection Transit (QWITT) diodes used in this research will be summarized. In addition, issues concerning barrier asymmetries in the quantum well of our DBRTDs and how they affected the resulting DC-IV characteristics will be addressed in Chapter 5 . Comparisons of the experimental results will be made with device simulations from both a Schrödinger-Drift/diffusion model [*Mil93*] and a Schrödinger-Poisson model [*Gul91*] .

4.2 Background Discussion for the AlAs/GaAs DBRTD

The concept of resonant tunneling of electrons can be generalized in the Kronig-Penney model where the periodic square well potentials result in allowed values of energy and forbidden gaps in energy. In these allowed bands of energy, there can be resonant transmission of electrons. Analogous structures are realized artificially with AlGaAs/GaAs superlattices grown by MBE where the allowed values of energy (minibands) are dependent on the well widths and the barrier heights. In these structures, Tsu and Esaki proposed that the application of an electric field to the structure would result in NDR. A typical AlAs/GaAs double barrier resonant tunneling diode (DBRTD) is formed by sandwiching a GaAs quantum well between two AlAs tunnel barriers as shown in Figure 4.1. Also see Figure 4.8 for a more in-depth cross-sectional layer structure.

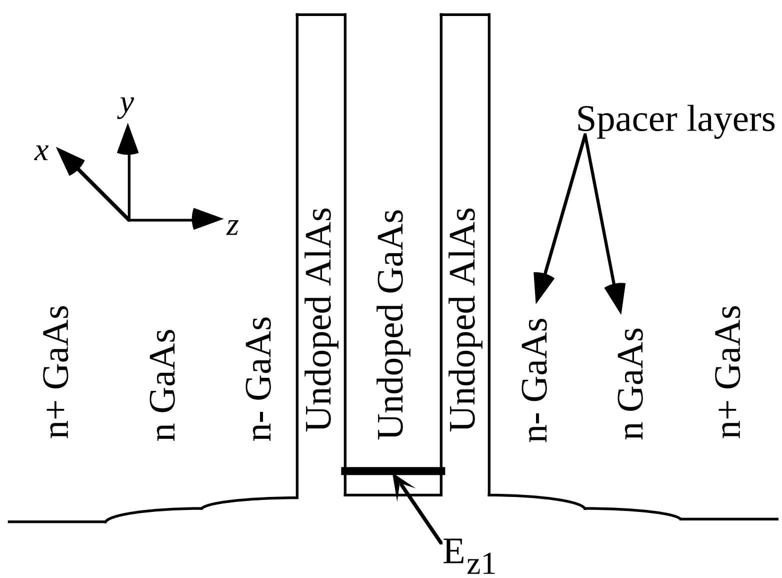

Figure 4.1: Illustration of the Γ-Γ conduction band profile of a typical AlAs/GaAs DBRTD, similar to those grown by MBE in this work. The conduction band offset between Γ_{AlAs} and Γ_{GaAs} is ≈ 1.04 eV.

For the double barrier resonant tunneling diode (DBRTD), the phenomenon of resonant transmission of electrons through the double barriers may not be intuitive at first. In order to describe the nature of resonant tunneling through double barrier structures, we first examine the transmission coefficient, $T_B(E_z)$, as a function of electron energy for a single barrier. Classically the transmission coefficient would be zero. From quantum mechanics, it is well known that even if the electron has an energy less than the height of the potential barrier, there exits some probability that the electron can tunnel through the single barrier if it is thin enough. A plot of the transmission coefficient versus energy is shown in Figure 4.2.

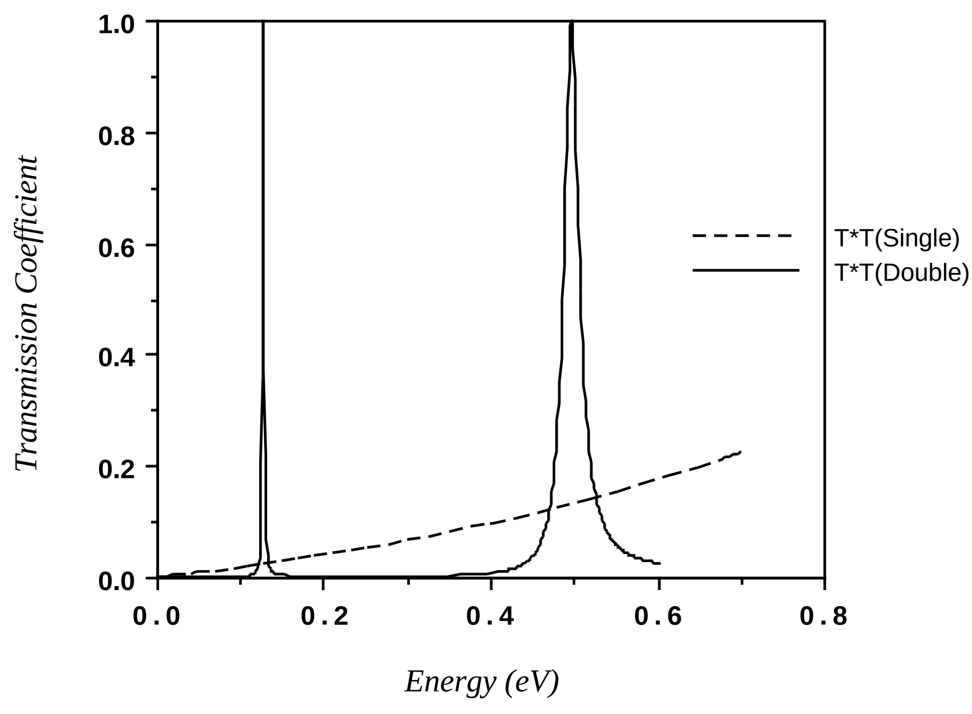

Figure 4.2: Plots of the transmission coefficient versus electron energy for an AlAs/GaAs single barrier structure and a double barrier structure. The barriers are 17Å and the quantum well of the DBRTD is 50Å. Plot courtesy of K.K. Gullapalli.

For a DBRTD with barriers of the same thickness as the single barrier, the total transmission coefficient, $T_{2B}(E_Z)$, for both barriers exhibits an interesting feature. $T_{2B}(E_Z)$ shows resonant peaks in its features that are approximately equal to one at lower energies than $T_B(E_Z)$. The resonant peaks in the transmission coefficient versus electron energy curve imply that when an electron obtains the energy where one of the resonant peaks exists, the probability for the electron to tunnel through the double barriers is approximately one. The way in which the electron obtains this energy in an actual DBRTD is through the application of an electric field across the device. The kinetic energy can be expressed as follows,

$$E = \frac{\hbar^2 k_z^2}{2m^*} + \frac{\hbar^2}{2m^*}\left(k_x^2 + k_y^2\right) = E_z(V) + \frac{\hbar^2 k_t^2}{2m^*} \tag{4.1}$$

where $E_Z(V)$ is the longitudinal energy or energy parallel to current flow, m^* is the effective mass and $E_t = \frac{\hbar^2 k_t}{2m^*}$ is the transverse energy. The transmission coefficient is calculated using the transfer-matrix method [*TsE73*] [*RiA84*] . Subsequently, the current density in the DBRTD can be determined from the Tsu-Esaki model as:

$$J = \frac{em^* kT}{2\pi^2\hbar^3}\int_0^\infty T_{2B}\left(E_z\right)N\left(E_z\right)dE_z \tag{4.2}$$

where $N(E_Z)$ is commonly called the "supply function" [*GuR89*] which is defined as;

$$N\left(E_z\right) = \ln\left(\frac{1+\exp\left[\left(E_f - E_z\right)\Big/ kT\right]}{1+\exp\left[\left(E_f - E_z - eV\right)\Big/ kT\right]}\right) \tag{4.3}$$

Figure 4.3 (a) shows the various steps in the operation of an AlGaAs/GaAs DBRTD. Figure 4.3 (b) defines specific J-V parameters of interest for DBRTDs.

(a)

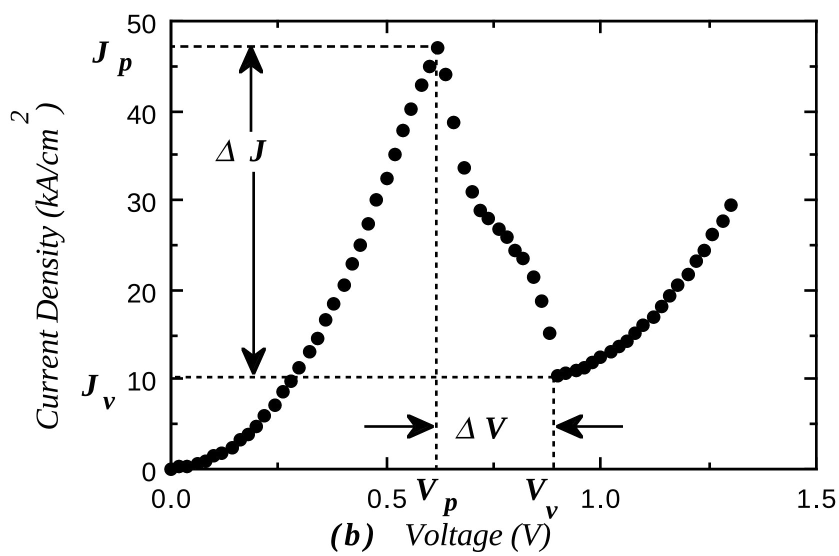

Figure 4.3: (a) Illustration depicting the operation of an AlGaAs/GaAs DBRTD as a voltage is applied to the device. Taken from Sollner, et.al.[*SoG83*]. (b) Measured J-V characteristics of an AlAs/GaAs DBRTD with specific parameters of interest highlighted: Peak voltage (V_p), Valley voltage (V_v), Peak current density (J_p), Valley current density (J_v), $\Delta V = V_v - V_p$, and $\Delta J = J_p - J_v$.

The resonant peaks correspond to the energy levels of the quasi-bound states in the quantum well when at thermal equilibrium or when the device is not under a voltage bias. When a voltage bias is applied to the right side contact of the DBRTD shown in Figure 4.3, the conduction band of the right side contact is pulled down, bringing the first quasi-bound state in the quantum well more in line with the left side contact Fermi level. Note typically these contacts are degenerately doped and therefore the Fermi level is usually in or above the conduction band edge. Thus the supply of electrons in the left side contact begin to tunnel through the double barriers and the current in the I-V characteristic increases. As the voltage is increased further, the electrons in the left hand side contact gain more energy and the supply of electrons on the left hand side contact become lined up with the first quasi-bound state in the well. This is the point at which the transmission probability is a maximum and the supply of electrons is the greatest for tunneling through the double barriers. This point shows up as the peak current, I_p, in the I-V characteristics. The voltage at which the I_p occurs is referred to as the peak voltage, V_p. Ideally, this voltage is given as

$$V_p \geq \frac{2E_1}{e} \tag{4.4}$$

but in most actual devices there are lightly doped spacer layers near the quantum well which result in voltage drops on both the accumulated and depleted side of the well. As the voltage is increased still further, the quasi-bound state is now pulled below the Fermi level of the left hand side contact and the supply of electrons decreases. Thus a drop in the current in the I-V characteristics is seen which is the well known NDR phenomena we mentioned earlier. Finally, as the voltage is still increased even further, the current is seen to rise in the I-V characteristics. This rise is attributed to thermionic emission over the barriers or Fowler-Nordheim type tunneling through the upper corner of the barriers.

4.3 Simulations Used for Analyzing the DBRTD

With the simple Esaki-Tsu model, predictions of the peak current density of DBRTDs are possible, but to date accurate simulations of the I-V characteristics of the DBRTD over its whole operating range have not been achieved. A step closer to those goals was obtained when the space charge effects in DBRTDs were taken into account by using a self-consistent calculation that couples the Schrödinger equation in the quantum well region to the drift-diffusion equation in the rest of the device [*Mil93*] or the Schrödinger-Poisson model [*CaM87*] [*Gul92*]. Both the Schrödinger/Drift-Diffusion models and the Schrödinger-Poisson models have been used by the author as a tool for designing and understanding certain DBRTD device structures that will be discussed later. The main distinction between the Schrödinger/Drift-Diffusion model and the Schrödinger-Poisson model is that the electric field is assumed to be constant across the quantum well in the Schrödinger/Drift-Diffusion model [*Mil93*]. In the more commonly used Schrödinger-Poisson model, the electric field is calculated at every point in the device structure. Accurate values for the peak voltage can be predicted once all the external parasitic and series resistances from the contacts were taken into account in both models. The Schrödinger-Poisson model relies on the converged solutions of V(z), $\Psi(k_Z,z)$, and n(z) in the following equations [*MiT89*] :

$$-\frac{\hbar^2}{2}\frac{d}{dz}\left(\frac{1}{m^*}\frac{d\Psi(k_z,z)}{dz}\right)+V(z)\Psi(k_z,z)=E_z\Psi(k_z,z) \tag{4.5}$$

where m^* is the position dependent effective mass, E_Z is the longitudinal electron energy, $\Psi(k_Z,z)$ is the electron wavefunction, and V(z) is the effective potential energy which can be described as:

$$V(z)=-e\phi_{elec.}(z)+V_h(z)+V_{xc}(z) \tag{4.6}$$

where $\phi_{elec}(z)$ is the electrostatic potential, $V_h(z)$ is the effective potential energy due to the heterojunction offset, and $V_{xc}(z)$ is the local exchange-correlation potential energy [*StS84*] which is included in a few of the models used to simulate DBRTDs [*GaD92*] . The electron concentration, n(z), and the Poisson equation are given as follows:

$$n(z)=2\sum_k f\left(k,E_f(z)\right)\left|\Psi(k_z,z)\right|^2 \tag{4.7}$$

$$\frac{d}{dz}\left(\varepsilon(z)\frac{d\phi_{elec.}}{dz}\right)=-e\left(N_D^+(z)-n(z)\right) \tag{4.8}$$

where $\varepsilon(z)$ is the position dependent dielectric constant, $N_D^+(z)$ is the ionized donor concentration, and n(z) is the free electron concentration. Although the Schrödinger-Poisson model provides good quantitative results for the peak current density, J_p, and the peak voltage, V_p, it does not provide a good estimate of the valley voltage, V_v or the valley current density, J_v, and therefore the peak-to-valley current ratio, PVCR. (See Figure 4.3 (b)) Many mechanisms which contribute to the valley current are not taken into account by the simple Schrödinger-Drift/diffusion model or Schrödinger-Poisson model. Some of these mechanisms include interface roughness scattering [*GuR89*], Γ-X mixing [*MeW87*] , and phonon-assisted tunneling [*ChV89*] .

Furthermore, the above mentioned coherent tunneling models do not take into account the energy loss mechanisms that occur in the depleted spacer layers. Without including these energy loss mechanisms, there is a gross underestimation of the electron concentration in the depleted spacer layers and the magnitude of the space charge resistance is underestimated. A recent "hybrid" model has been used to take into account these energy loss mechanisms in which the coherent tunneling model is used to account for electron injection from the double barrier structure and the Boltzmann transport equation is used to account for the semiclassical dynamics which occur in the depleted spacer layer [*GuM91*] . This hybrid model is useful in properly accounting for the space charge resistance in DBRTDs with long spacer layers. It should also be mentioned that more advanced simulations of DBRTDs have been implemented using the quantum kinetic equations such as the Wigner distribution function or the Lattice Wigner distribution function [*MiN91*] .

4.4 Coherent vs. Sequential Tunneling and Inelastic Tunneling

The Esaki-Tsu model described earlier is based on a coherent tunneling model which is analogous to the Fabry-Perot resonator. This model relies on calculating the total transmission coefficient, $T_{2B}(E_z)$ of the whole structure. In 1985, Luryi proposed an alternative viewpoint that explains the NDR phenomena in DBRTDs which has subsequently been termed "sequential tunneling" or "incoherent tunneling"

[*Lur85*] . From this different perspective, Luryi proposed that NDR could occur without the resonant Fabry-Perot analogy, but rather from considering a step by step procedure in which the electron can first tunnel from the emitting electrode to the quasi-bound state in the well and subsequently tunnel from the quasi-bound state to the collecting electrode. Thus, this two-step process was given the name sequential tunneling [*Lur85*]. In the interim, the electron can face many inelastic scattering events which would cause it to lose its phase coherence [*WeV87*] [*Kho90*] . Viewing the tunneling process of the DBRTD as sequential was quite appealing to many [*WeV87*], [*Kho90*], [*GoT88*] because it allowed the inclusion of inelastic scattering events and other interactions in the well rather than the somewhat idealistic view of total elastic processes in the quantum well.

It was found that although the sequential tunneling model could incorporate scattering events in the well, the calculated peak current densities were equivalent to those calculated in the coherent tunneling model [*Kho90*],[[*JoG90*]. In addition, the PVCRs determined from these models were still overpredicting those measured experimentally. In fact, with the more advanced physically-based quantum transport models, the issue of coherent tunneling or sequential tunneling is a moot point since these more advanced models are based on a many body particle point of view and not a single particle point of view [*Mil91*] .

Nonresonant inelastic tunneling in AlAs/GaAs DBRTDs must be taken into account. Inelastic tunneling is named as such due to tunneling through the lower Γ_{GaAs}-X_{AlAs} barrier due to various inelastic scattering events such as phonon scattering, impurity scattering and interface roughness scattering as shown by Mendez and co-workers who examined the effects of hydrostatic pressure on the DC-IV characteristics of AlAs/GaAs DBRTDs at 77K [*MeC88*] . Through the hydrostatic pressure studies and by using the valley current as a monitor for inelastic tunneling, Mendez found that thicker AlAs barriers exhibit more inelastic tunneling whereas thinner AlAs barriers ($\approx$ 8ML) exhibit reduced inelastic tunneling. Intuitively, one may expect inelastic tunneling since the Γ_{GaAs}-Γ_{AlAs} discontinuity in the conduction band is approximately 1.04 eV and the Γ_{GaAs}-X_{AlAs} discontinuity is approximately 0.19 eV as shown in Figure 4.4. Similar results were found by Kyono and co-workers in which they studied the tunneling processes in AlAs/GaAs single barrier structures as a function of barrier thickness and temperature. They distinguished the temperature dependent inelastic current processes from non-temperature dependent elastic tunneling processes [*KyK89*] . In this study, it was determined that perpendicular transport through a 5 ML AlAs single barrier was primarily Γ_{GaAs}-Γ_{AlAs} elastic tunneling dominated but in an 11 ML AlAs single barrier the current transport began to show inelastic temperature dependent tunneling effects. The question though is whether or not these inelastic tunneling effects exhibit themselves at some point in between 5 ML and 11 ML. These inelastic tunneling effects in AlAs/GaAs DBRTDs exhibit a temperature dependence and are a major cause of the higher valley currents and the reduced PVCRs.

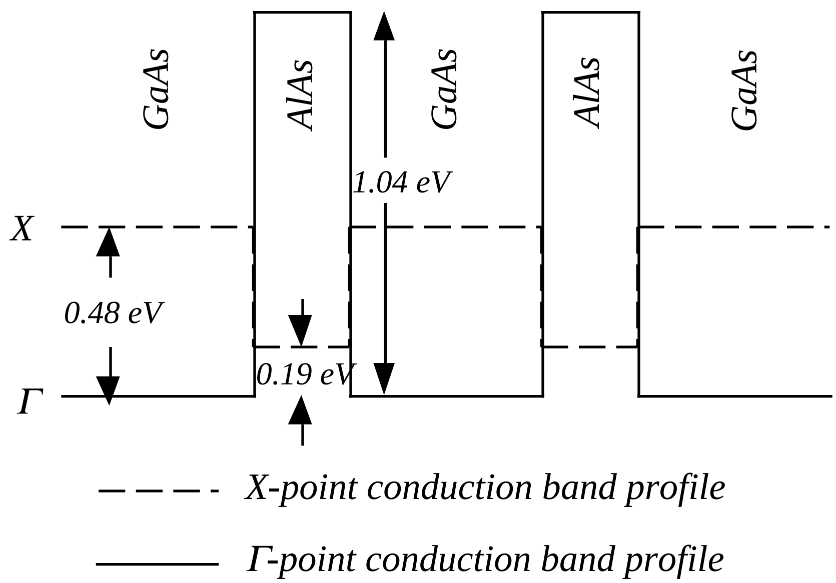

Figure 4.4: Conduction band profiles of both the Γ-point and X-point minimums for an AlAs/GaAs DBRTD. The offset values were taken from Liu [*Liu87*].

4.5 The QWITT diode and the use of depleted spacer layers

The basic distinction between the QWITT diode and the DBRTD is the addition of an undoped or lightly doped GaAs layer downstream or on the collector side of the AlAs/GaAs quantum well, as shown in Figure 4.5. Although the influence of space charge effects on the DC-IV characteristics of DBRTDs has been previously investigated [*CaM87*], the importance of space charge resistance and its impact on the high frequency performance of resonant tunneling structures was made apparent by Kesan, et al. [*KeN87*] .

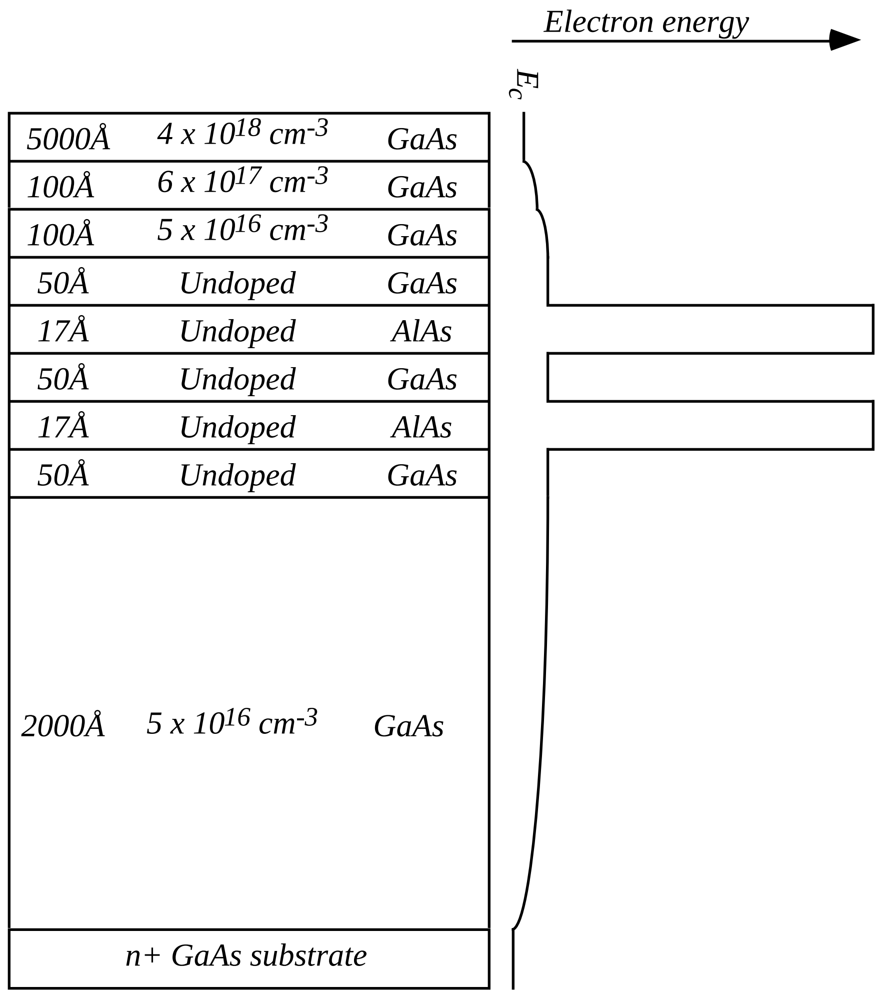

Figure 4.5: Illustration depicting the cross-sectional layer structure and conduction band profile of a typical AlAs/GaAs QWITT. (Not drawn to scale.)

Since a significant part of the voltage is dropped across the depleted spacer layers, it is found that the NDR region is pushed out to higher voltages and also broadened along the voltage axis. As a result, the ΔV can be increased through the addition of a lightly doped spacer layer after the quantum well. From equation 4.12, the r.f. output power of such devices can be improved. The normalized injection conductance, σ, is an important parameter for small signal analysis and is defined as follows [*KeN88*] :

$$\sigma = l\left(\left.\frac{\partial J_{qw}}{\partial V_{qw}}\right|_{V_o}\right) = \left(\frac{\partial J_{qw}}{\partial E_{qw}}\right) \tag{4.9}$$

where l is the length of the quantum well region, J_{QW} is the current density, V_{QW} is the voltage across the quantum well, E_{QW} is the electric field across the quantum well, and V_o is the dc bias voltage. From this injection conductance, the optimum depleted spacer layer width, W, at a given frequency can be determined as well as the specific negative resistance of the device. Unfortunately, extracting the J-E characteristics from the measured DC-IV characteristics can be very difficult because it requires an exact knowledge of parameters such as specific contact resistance of the device being measured, the doping concentration in all the epi-layers, and the thickness of all the epi-layers. Due to realities of MBE growth and fabrication, even at their level of sophistication where layer thicknesses and doping concentrations are now controlled very well, there is always some uncertainties in the parameters mentioned above which may result in deviations in the extraction of parameters such as the injection conductance of the quantum well. In addition, the $\sigma_{inj.}$ is highly dependent on the quantum well structure and the related material system. The two most important parameters for σ are the ΔJ and the ΔV and the quantum well structure should be designed accordingly and not be based just on an optimum peak-to-valley current ratio (PVCR). A comparison of J-V characteristics between an AlAs/GaAs DBRTD and AlAs/GaAs QWITT with the same quantum well structure is shown in Figure 4.6.

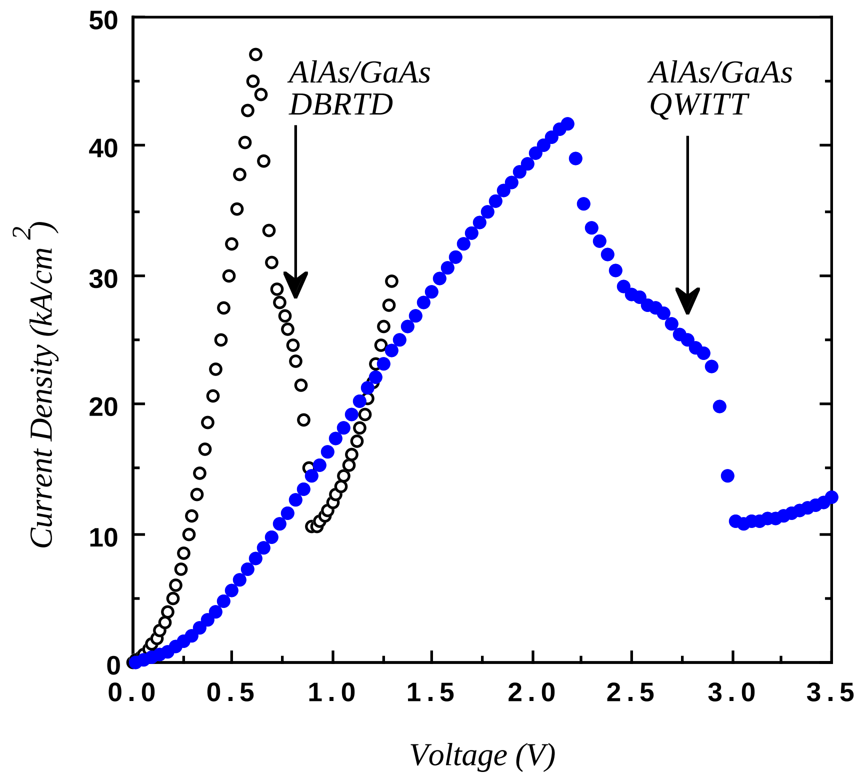

Figure 4.6: Measured J-V characteristics for an AlAs/GaAs DBRTD and an AlAs/GaAs QWITT with the same quantum well structure. The quantum well consists of 17Å AlAs barriers, a 50Å GaAs well and an $\sigma_{inj.} \approx 0.3$ (1/Ω-cm).

Since it is not the purpose of this research to investigate the r.f. aspects of the QWITT and the DBRTD, the reader shall be referred to the appropriate references [*KeN88*]. A few points should be made for completeness about the basic premise and usefulness of the QWITT structure. When the QWITT structure is biased into the NDR regime, the electric fields in the drift region are typically in the 100 KV/cm^2 to 300 KV/cm^2 range and it is assumed that the electrons traverse the drift region at their saturation velocity. In the majority of the calculations used in the analysis of these devices, the saturation velocity, v_s, used is 6x10^6 cm/sec. Once the optimum drift region length has been found, combined with the injection conductance of the quantum well, the r.f. characteristics of the device can be predicted. The two frequency regimes under which the QWITT operation can be described are $\omega > |\sigma|/\varepsilon$ and $\omega < |\sigma|/\varepsilon$. The frequencies that our group has been able to work with have been at $\omega < |\sigma|/\varepsilon$ and therefore the overall specific negative resistance of the QWITT can be expressed as [*KeN88*]:

$$R \cong \frac{W}{\sigma} + \frac{W^2}{2\varepsilon v_s} \tag{4.10}$$

where W is the drift region length, ε is the dielectric constant of the GaAs drift region, and v_S is the saturation velocity. The first term in the above expression represents the space charge resistance associated with the quantum well injector coupled to the drift region using the assumption of a constant saturated velocity and the second term in the above expression represents the space charge resistance of the depleted drift region which is always positive. Note here that it is important to design the drift region length and doping concentration such that the drift region is almost fully depleted when the QWITT is under bias; otherwise any undepleted drift region will contribute a positive series resistance to the overall specific resistance of the device. Typically one must also take into account the specific contact resistance of the contacts if the specific contact resistance is much higher than $5\text{x}10^{-6}$ Ω-cm^2. Therefore, in order to maximize the r.f. characteristics of the QWITT structure, several important structural considerations must be taken into account such as the depleted drift region length, the quantum well structural design for maximum injection conductance, and minimization of contact resistances. The maximum frequency at which an optimized QWITT can oscillate in a circuit is given by [*KeN88*],

$$f_{\max} = \frac{1}{4\pi}\sqrt{\frac{v_s}{\varepsilon\left(R_{cont} + AR_{circuit}\right)}} \tag{4.11}$$

where R_{cont} is the contact resistance, $R_{circuit}$ is the resistance of the external circuit, and A is the device area. As well as being able to predict the maximum frequency at which these devices can operate, it is also useful to approximate the maximum r.f. power that these devices can deliver. From the tunnel diode literature [*KiB61*] [*StN61*] , the maximum r.f. output power is determined by treating the current as a cubic polynomial of the voltage and is expressed as:

$$P_{rf} = \frac{3}{16} \Delta V \times \Delta I \tag{4.12}$$

Experimentally, it has been shown that equation 4.12 seems to underpredict the total output power of some of the QWITT diodes and that the time averaged power, P_{ac}, calculated using a quasi-static method appears to give better results [*ReT90A*] . One reason that the quasi-static power calculation may compare better with experimental data is the fact that the voltage swing is not restricted to lie within the boundaries of the peak and valley voltage and may extend outside these regions. The highest oscillation frequency obtained so far on a DBRTD structure is 712 GHz in an InAs/AlSb DBRTD mounted in rectangular waveguide resonator [*BrS91*]. Here it should be noted that the injection conductance of the DBRTD in the InAs/AlSb is much higher than the AlAs/GaAs system and therefor allows a higher cutoff frequency.

4.6 Temperature Dependence of AlAs/GaAs DBRTDs and QWITTs

The impact of temperature on the DC-IV characteristics of AlAs/GaAs or AlGaAs/GaAs DBRTDs has been examined experimentally with varying degrees of agreement [*HuI87*] [*VaL89*] [*ShX91*] . It is well known that the quantum interference effects observed in DBRTDs become more pronounced at low temperatures as thermionic based current and inelastic tunneling mechanisms are reduced. These mechanisms exhibit their influence primarily in the valley current. Thus, an obvious characteristic in the DC-IV characteristics of a DBRTD as the temperature rises is a corresponding rise in the valley current. Experimentally, other research groups have observed the peak current to increase and decrease as the temperature of the device is increased. These discrepancies may be due to the quantum well structures used in each study [*HuI87*],[*VaL89*],[*ShX91*]. The peak voltages are observed to decrease slightly with increased temperature and the valley voltages are observed to decrease to an even greater extent with increasing temperature. The significant decrease in the valley voltage with temperature corresponds to the increasing thermionic based currents that also occur with temperature. In order to predict the peak current density at low temperature, one can use the following expression from the Tsu-Esaki model which expresses J_p in the limit as T approaches 0 K [*TsE73*]:

$$J = \frac{em^*}{2\pi^2\hbar^3}\int_0^{E_f} T_{2B}(E_z)(E_f - E_z)dE_z \qquad V \geq E_Z \quad (4.13)$$

$$J = \left(\frac{em^*}{2\pi^2\hbar^3}\right)\left[V\int_0^{E_f - V} T_{2B}(E_z)dE_z + \int_{E_f - V}^{E_f}(E_f - E_z)T_{2B}(E_z)dE_z\right] \qquad V < E_Z \quad (4.14)$$

Temperature dependent measurements have also been taken on the AlAs/GaAs DBRTD and AlAs/GaAs QWITT structures fabricated by our group using a two-probe MMR low-temperature probe station that allows one to take DC-IV measurements of devices down to 77K. This low-temperature system utilizes a refrigerator based on the Joule-Thompson effect and therefore the refrigerator itself must be enclosed in vacuum as it is cooled down to 77K in order to avoid condensation. The temperature is monitored with a silicon diode and the sample can be heated with a resistance heater [*MMR84*] . The gas used in the Joule-Thompson refrigerator is nitrogen which is fed into the refrigerator by a capillary at pressures ranging from 1500 psi to 1800 psi. In actual practice, it was found that the system could only be brought down to about 80K and held there for about 30 minutes before the temperature would start rising. Another limitation that made it very difficult to take many measurements on the DBRTDs was the fact that the microscope used for viewing the device under vacuum through a viewport had limited magnification which made it extremely difficult to see the 5μm to 15μm pads of our devices and to probe them repeatably. Nevertheless, a few measurements were taken to examine the change in DC-IV characteristics at low temperatures, but no systematic study could be attempted. In Figure 4.7, a plot of the valley current and the PVCR versus temperature are given for an AlGaAs/GaAs DBRTD (No publications could be found on the systematic variation of the temperature of AlAs/GaAs DBRTDs).

Figure 4.7: Variation of valley current and PVCR versus temperature for AlGaAs/GaAs DBRTDs. Taken from [*VaL89*].

4.7 MBE Growth of AlAs/GaAs DBRTDs and QWITTs

The first use of MBE to grow the layer structures for AlGaAs/GaAs DBRTDs was first performed by Chang, Esaki, and Tsu in 1974 [*ChE74*]. Since that time, improvements in material quality, materials characterization, and epitaxial growth techniques have led to continued improvements in the peak-to-valley current ratio (PVCR), precise control over barrier and well thicknesses, and control over doping profiles. The substrate preparation and system description for MBE growth in the Varian Gen II have already been discussed in Chapter 2. In this section, the specifics related to growth of the AlAs/GaAs DBRTDs will be discussed.

Specific to the MBE growth of DBRTDs is the accurate determination of the AlAs and GaAs growth rates by RHEED. As will be discussed later in Chapter 5, any variation in the barrier thickness of the DBRTD can result in dramatic changes in the current density of the device as well as the $\sigma_{inj.}$. In addition, any variation in the quantum well width will result in different peak voltages. Therefore, in order to be able to repeatably perform MBE growth of DBRTD structures, growth rates were determined both before and after every MBE run. Especially as source material begins to deplete in their respective effusion cells, the growth rates can fluctuate significantly over a couple of hours.

The beam equivalent pressure of the As_2 cracker is set to obtain the proper As/Ga incorporation ratio of $\approx$ 1.5-1.7 for a growth rate of 1 ML/sec.. The native oxide is usually desorbed at raised substrate temperatures of $\approx$ 660°C under an As overpressure. For the device structures grown in this research, the AlAs growth rate was typically kept at 0.25 ML/sec or 0.3 ML/sec and the GaAs growth rate for the quantum well and all spacer layers was kept at 0.4 ML/sec. The GaAs growth rate for all other layers outside of the quantum well and spacer layers was 1.0 ML/sec. Typical growth temperatures were 600°C. CAR rotation is kept at 5 rpm during the whole growth cycle. The quantum well consists of a 50Å nominally undoped GaAs quantum well surrounded by 17Å AlAs barriers. The three-step spacer layers consist of a 50Å nominally undoped GaAs layer, a 100Å GaAs layer n-type doped at 5 x $10^{16}cm^{-3}$, and a 100Å GaAs layer n-type doped at 6 x $10^{17}cm^{-3}$. The growth interrupts at each interface inside or adjacent to the quantum well was 4 seconds. The growth interruption during each change in Si doping setpoints usually required about 10 minutes for the Si effusion cell to stabilize. This structure is referred to as a "baseline" structure since it is the most commonly used structure in our AlAs/GaAs DBRTDs and QWITTs. A typical cross-section of a baseline AlAs/GaAs DBRTD grown in this research is shown in Figure 4.8.

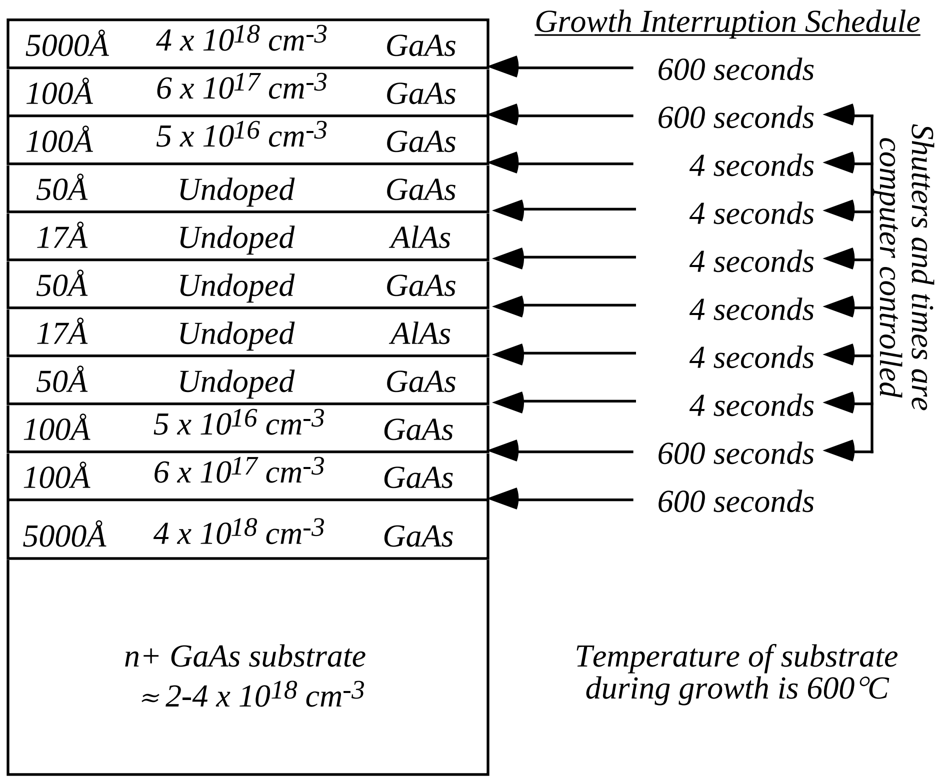

Figure 4.8: Cross-sectional layer structure for a "baseline" AlAs/GaAs DBRTD and the growth interrupts associated with the MBE growth of the device.

4.8 Formation of Ohmic Contacts and Fabrication Issues

The impact and influence of the specific contact resistivity, ρ_c, on the d.c. and r.f. performance of double barrier resonant tunneling diodes (DBRTDs) has been an important area of concern of our group during the initial development of the QWITT and DBRTD structures. The specific contact resistivity can be defined as:

$$\rho_c = \left.\frac{\partial V}{\partial J}\right|_{cont} \left(\Omega - cm^2\right) \tag{4.14}$$

Since the typical device sizes we work with are on the order of 100 μm^2, we need a fairly low specific contact resistivity to reduce the overall series resistance in the device. There are many definitions for an ohmic contact such as: 1) that the I-V curve is linear through the origin 2) that the contacts are not injecting and 3) that the contact resistance is small compared to the device resistance [*PiG83*]. High series resistance effects can degrade the overall r.f. output power by a concurrent drop in the ΔV (V_p - V_v), as seen in equation 4.12. In addition, the maximum oscillator frequency, f_{max}, of the DBRTD is inversely proportional on the contact series resistance in the structures as shown in Equation 4.11. The effects of a high ρ_c can be seen by adding its series resistance to the actual J-V characteristics of a standard AlAs/GaAs QWITT diode. Values of 10Ω, 20Ω, and 30Ω (1×10^{-5} Ω-cm^2, 2×10^{-5} Ω-cm^2 and 3×10^{-5} Ω-cm^2, which can actually occur in a poor process) were added to the I-V characteristics of an AlAs/GaAs QWITT diode, as shown in Figure 4.9.

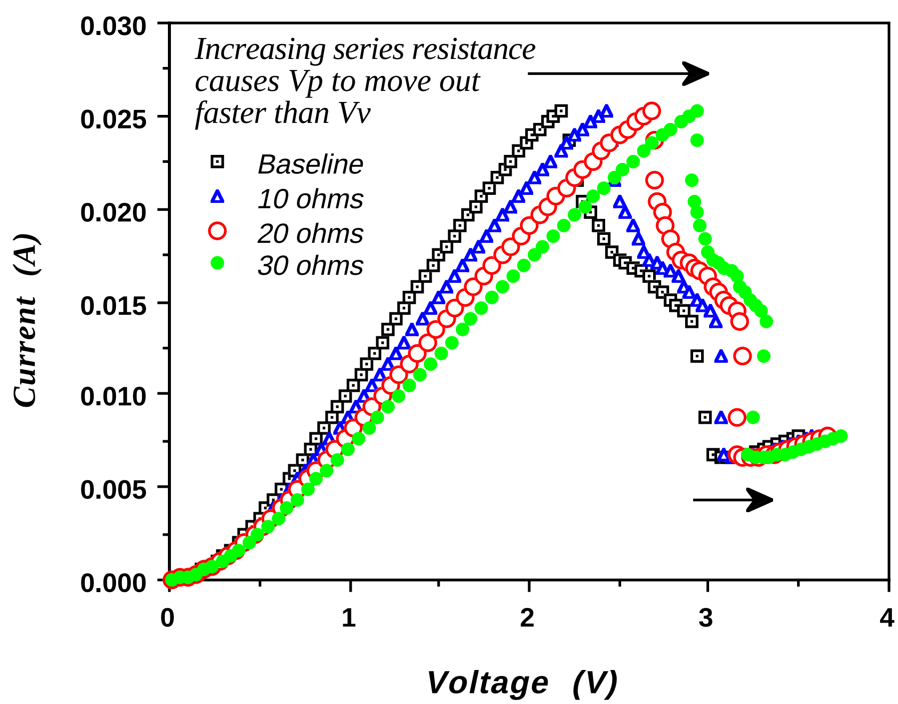

Figure 4.9: Impact of series resistance on the I-V characteristics of an AlAs/GaAs QWITT diode with additional series resistances of 10Ω, 20Ω and 30Ω added to the I-V curves. Note that ΔV decreases with increasing series resistance.

Two significant problems that plagued the r.f. performance of the QWITT and DBRTD oscillators were high specific contact resistivity (Ω-cm^2) and poor contact adhesion to the substrate. Various metallization schemes and surface preparations were used to improve these two parameters. In this section, ohmic contacts to GaAs, specific contact resistivity extraction issues, and fabrication related issues will be discussed.

It is well known that in GaAs the Fermi level is pinned approximately 0.7-0.9 eV below the conduction band as result of interface state formation or the effective workfunction of microscopic clusters of oxide/metal phases [*WoF81*] . Gold, for example, has been observed to have a Schottky barrier height of 0.88 eV on a 100 n-type GaAs surface [*RhW88*] . The source of these interface states have been heavily investigated with many theories as to their origin and effect on ohmic contacts [*WoF81*]. The result of Fermi level pinning in GaAs results in a Schottky barrier when an unalloyed contact is formed on the GaAs. The two most common methods to obtain an ohmic contact are to thin the barrier such that the electrons can tunnel through the barrier or lower the barrier such that the electrons can traverse the metal semiconductor interface unimpeded. Thus, a way must be found to avoid thermionic emission over this Schottky barrier and tunnel through the barrier either by thermionic field emission or field emission. Thermionic field emission occurs when the carriers, in our case electrons for metal-n+ contacts, acquire enough energy such that they can tunnel through the top of the Schottky barrier. Field emission occurs when the electrons can tunnel through the Schottky barrier which usually is very thin through the use of a degenerately doped n+ GaAs layer. Both thermionic emission and thermionic field emission are temperature dependent. In fact, for metals on heavily n-type doped GaAs material, the depletion region does not extend fully into the n+ doped layer, thus allowing electrons to tunnel through the barrier. In our actual contacts, the specific contact resistivity was quite high when the contacts were unannealed. Therefore we had what could be referred to as "poor" ohmic contacts and/or "leaky" Schottky diodes. In order to improve the specific contact resistivity of these "poor" ohmic contacts on heavily n-type doped GaAs layers, work has been performed investigating surface preparation, metallization techniques, alloying recipes, and MBE growth conditions for some of the delta-doped and indium based nonalloyed contacts. In addition, the contact resistance measurements and their limitations will be discussed.

The two most common alloyed ohmic contact metallurgies to n+ GaAs are the AuGe/Ni system and the PdGe system. Germanium is the preferred n-type dopant over tin because it does not diffuse as fast in GaAs. The AuGe/Ni system was originally developed in 1967 by Braslau, Gunn and Staples [*BrG67*] and has been investigated very heavily, even up to the present. The standard Au-Ge eutectic specified is 88% Au and 12% Ge by weight. The use of Ni and its purpose have changed since its initial use in such contacts. Several studies have investigated in-depth the mechanisms that occur when the AuGe/Ni system is alloyed and variations of the metallization scheme to achieve a lower specific contact resistivity [*BrP87*] [*KuB83*] . The number of various metallization schemes or "recipes" that can be cited in the literature is almost endless, thus indicating that there is not a set procedure in making AuGe/Ni contacts [*Oga80*] [*CaP85*] [*MuC86*] . Instead the number of recipes most likely can find their origin based on the type of surface preparation, vacuum/evaporator setup and source material purity that was implemented. Many factors must be taken into account when deciding upon the layers and layer thicknesses to be used in a AuGe/Ni contact scheme such as ohmic contact adhesion, equipment capabilities (two or three evaporation sources), specific contact resistivity, sheet resistance of the final alloyed metallization, whether the contact metallization will be patterned by lift off or by etching, and will the contact metallization be used as an etch mask during mesa isolation.

The original contact metallization used on the AlAs/GaAs DBRTD and QWITT structures consisted of approximately 800Å-1000Å of Au-12% Ge and 200Å of Ni. A typical surface preparation before the evaporation consisted of the 2:1 HCl:DI-H_2O etch for 30 seconds. The metal evaporation was performed in a standard bell jar evaporator, named "Philvac", utilizing a rough pump, Varian coldtrap and diffusion pump. Typical base pressures of 3×10^{-6}-5×10^{-6} Torr were achieved in this system. The alloying of these contacts was performed in a rapid thermal annealer (RTA) at 450 ˚C for 30 seconds in forming gas. This contact metallization resulted in poor contact adhesion in both the metallization lift off process and the metallization etch process that was used in fabricating our QWITT diodes and DBRTDs. In addition, the metallization scheme used above resulted in poor and highly unrepeatable specific contact resistivities in the range of $6x10^{-6}$ Ω-cm^2 - $5x10^{-5}$ Ω-cm^2. Possible causes of the poor quality of these contacts are oxygen contamination of the Ge during thermal evaporation using alumina coated tungsten boats [*MoT88*] , the high base pressure in our bell jar evaporator, or most likely poor formation of a Ni_2GeAs alloy at the metal-semiconductor interface. With this specific AuGe/Ni contact metallization scheme, it was apparent that consistent, repeatable results were not able to be achieved. As stated, many different mechanisms have been proposed during alloying of the AuGe/Ni contacts. Some of the mechanisms that appear to be common among those proposed are [*BrP87*] [*ChL91*],and [*KuB83*]: In the early stages of alloying (300 ˚C) the Ge and Ni diffuse to the contact interface. Ga is found to accumulate where the Au is and As accumulates in the region where the Ni is located. Thus two concurrent reactions have initially occurred: Au + GaAs $\Rightarrow$ AuGa + As and Ni + As $\Rightarrow$ NiAs. The NiAs formation at the interface is critical for the subsequent formation of the Ni_2GeAs phase which provides the ohmic contact. In the later stages of alloying (400 ˚C), the temperature is well above the AuGe eutectic temperature of 356 ˚C and continued movement of Ni and Ge to the NiAs interface occurs. It is the formation of the Ni_2GeAs phase that apparently provides the low specific contact resistivities. Thus the goal is to find the appropriate metallization scheme that produces the Ni_2GeAs phase in large percentages at the interface and reproducibly. The remaining Ni on the top helps prevent the Au from balling up during the latter stages of the alloy cycle.

A new metallization scheme chosen to obtain improved specific contact resistivities repeatably was ≈ 15Å Ni, 800Å-1000Å AuGe and ≈ 150Å Ni. In the metallization lift off process, which is currently used more often since it has two less etch steps than the metallization etch process, a 2:1 HCl:DI-H_2O etch is used as a surface preparation to remove surface oxides. The first layer Ni is used to improve contact adhesion [*MuC86*] to the GaAs as well as aid in the formation of the Ni_2GeAs phase. The top layer Ni, again, aids in the prevention of excessive Au balling during the alloy. A typical SEM micrograph of the Ni/AuGe/Ni contact after alloy is shown in Figure 4.10.

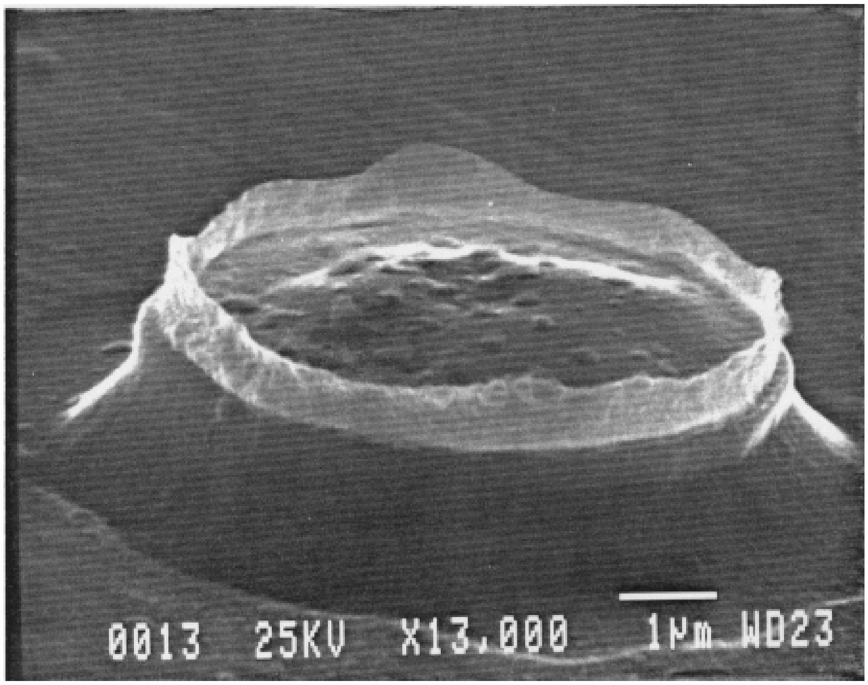

Figure 4.10: SEM micrograph showing an alloyed Ni/AuGe/Ni ohmic contact on top of a mesa isolated AlAs/GaAs DBRTD. This contact was made using a metallization lift off process where AZ1350J-SF photoresist is used to define and lift off the contact pattern. Note the metal flags at the edge of the metallization due to the lift off procedure.

With this metallization we observed lower specific contact resistivities in the range of $7x10^{-7}$ Ω-cm^2 to $4x10^{-6}$ Ω-cm^2.

The specific contact resistivities were determined through the use of Cox-Strack test structures [*CoS67*] as shown in Figure 4.11.

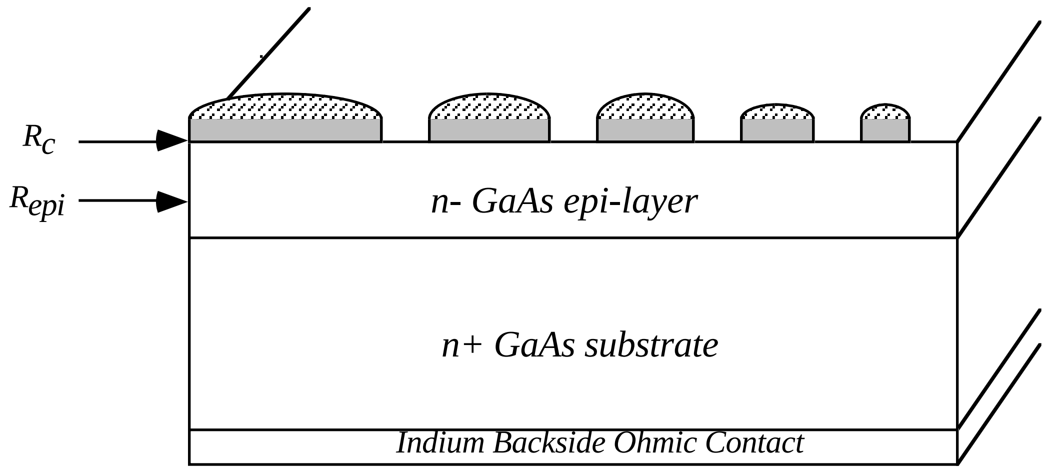

Figure 4.11: Schematic drawing of Cox-Strack structures used to extract the specific contact resistivity of the ohmic contact metallization

In this measurement scheme, specific contact resistivities are extracted with the ohmic contacts on thin epitaxial layers of a certain doping concentration (resistivity). The total measured resistance, R_t, of these structures can be broken up into various terms as follows:

$$R_t = R_c + R_{epi} + R_{ext}\,(\Omega) \tag{4.15}$$

$$R_c = \frac{\rho_c}{\pi\left(\frac{d}{2}\right)^2} \quad (\Omega) \tag{4.16}$$

$$R_{epi} = \frac{\rho_{epi}}{\pi d}\tan^{-1}\left\{\frac{4}{d/t_{epi}}\right\} \quad (\Omega) \qquad (4.17)$$

where R_t is the total measured resistance, R_c is the contact resistance, R_{epi} is the spreading resistance through the epi-layer, R_{ext} is the external resistances in the test setup, ρ_c is the specific contact resistivity, ρ_{epi} is the resistivity of the epi-layer, and t_{epi} is the thickness of the epi-layer. The above resistance terms can be separated out by measuring the total resistance of contacts of different diameters. Although these test structures result in larger errors than other methods, the Cox-Strack structures allow for quick and simple measurements and are based on perpendicular current transport through the epi-film as is also done on our two-terminal DBRTD and QWITT structures. The contact sizes that are measured in this technique must be as small as possible in order to reduce the error that occurs in measuring larger contacts. The typical contact diameters measured are 4μm, 6μm, 8μm, 10μm, 12μm and 14μm. To reduce error, the device area of every test structure is measured with an HMOS optical measurement tool Since the doping concentration is known quite well through C-V and Hall measurements, the resistivity can be determined from tables in any text.

Although the Ni/AuGe/Ni metallization provided a low specific contact resistivity and was repeatable, there were two problems with this metallization. The first problem was related to the fact that the metallization did not act as a good etch mask in the "lift off" process when mesa etching our device structures with any peroxide based etchant such as 8:1:1 H_2SO_4:H_2O:H_2O_2. In fact, most of the etchants used for mesa isolation of our diodes appeared to attack the top layer Ni. The second problem is seen in the fact that the Ni/AuGe/Ni alloyed metallization has a fairly high sheet resistance in the range of 2 Ω/square [*Wil90*] . The sheet resistance of the metallization is not negligible compared to the interfacial sheet resistance or the sheet resistance of the degenerately doped GaAs semiconductor layer underneath the metallization. Thus probing this type of metallization will result in irreproducible results since there is a resistive drop across the pads. Since our tungsten probe tip diameters are on the order of 0.6μm-1.2μm, probing larger pads will result in larger ohmic resistance across the pad. This additional external series resistance on our larger pads which range 20 μm to 50 μm diameter resulted in reduced ΔV's in our DBRTDs and QWITT structures. Therefore, although the Ni/AuGe/Ni metallization provided a low specific contact resistivity it had a high sheet resistance which resulted in series resistance losses in our DBRTDs with larger pads.

In order to circumvent the above mentioned problems, a 1000Å layer of Au was used to cap the standard Ni/AuGe/Ni metallization to reduce its sheet resistance. The concern here was to make sure the Au cap layer did not degrade the specific contact resistivity. Our results showed no degradation in ρ_c when a Au cap layer was used with similar conclusions stated by others [*ChL91*] [*Wil90*]. In addition, the Au cap layer was ideal for an etch mask during mesa etching of our diodes.

Other types of ohmic contact schemes have been reported in the literature with the In-based non-alloyed contacts and the PdGe contacts as the most prominent. These metallizations have been looked at with respect to the processes used in Epitaxial Lift Off, which has been described in Chapters 3 and 6.

With the metallization pattern formed on top of the layer structures for the AlAs/GaAs DBRTDs or QWITTs, they are subsequently mounted on glass cover slips with clear glycophalate wax and mesa etched in 8:1:1 H_2SO_4:H_2O_2:H_2O. The cover slips protect the In backside ohmic contact metallization from the etchant. Etch rates for all these peroxide based etches used are not tabulated in this work. It is strongly suggested that etch calibration samples be made and used in combination with an Alpha-step profilometer before committing a device sample to an etch. After the mesas are etched to their desired depth, the samples are removed from the cover slips and rinsed in acetone, ethanol, and DI-H_2O. The samples are finally alloyed at 450°C for 30 seconds in a rapid thermal annealer (RTA). The devices are tested using a Keithley 230 programmable voltage source and a Keithley 195A digital multimeter controlled by an IBM PC-AT. Once the raw I-V data has been collected, the contact sizes of the devices under test are measured using an HMOS optical measurement tool. Typical sizes range from 5 μm to 45 μm. To extract the parameters of interest for the DBRTDs and QWITT diodes (PVCR, V_p, J_p, V_v, J_v, E_p, E_v, $\sigma_{inj.}$, etc.) would require a significant amount of time. A program written by D.R. Miller has been utilized to extract these parameters from the raw I-V data [*Mil90*] .

4.9 MBE Growth of AlAs/GaAs DBRTDs with As_2 and As_4

The growth of AlGaAs/GaAs heterostructures with either dimeric arsenic, As_2, or tetramic arsenic, As_4, has been studied in-depth. The use of As_2 has gained much interest due to the observations that GaAs growth with As_2 results in lower deep level concentrations, lower recombination velocities at AlGaAs/GaAs interfaces, improved surface morphologies during growth of AlGaAs [*LeS86*] and from a practical point of view, can save up to twice as much of the elemental As source during growth runs. However, there are also certain disadvantages of using As_2. One of the possible problems that may occur when thermally cracking As_4 at high temperatures is the generation of AsO. Furthermore, at these high temperatures (≥ 770°C), contamination and generation of defects may also occur. Typically in the Varian Gen II, the arsenic cracker is set at temperatures of 570°C-670°C (setpoints of 10-12 on the PID controller) such that not all the As_4 is thermally cracked due to observations that higher cracker settings can cause high AsO concentrations which significantly degrade the RHEED intensity oscillations and photoluminescence (PL) intensity [*BlS91*] . In a few occasions, the As_2 cell has been depleted through heavy use and the As_4 had to be used to grow certain layers before bringing the MBE system up to air to re-load materials. In this section, an AlAs/GaAs DBRTD grown with As_4 will be compared to a standard AlAs/GaAs DBRTD grown with partially cracked As_4 at a cracker setting of ≈ 670°C.

The layer structure, MBE growth, fabrication and device testing of the DBRTDs in this section follows those procedures described in section 2.2, section 2.3, section 4.7 and section 4.8. Since the majority of the DBRTDs are grown with the arsenic cracker, the main concern about growing with the As_4 cell is to make sure that the above mentioned problems associated with As_4 do not degrade any parameters of interest in the DC-IV characteristics. In table 4.1, specific DC-IV parameters are given for a standard DBRTD grown with the As_2 cell and the DBRTD grown with the As_4 cell.

DBRTD Structure (Bias)	Peak Voltage, V_p	$\Delta V = V_p - V_v$	Peak Current Density, J_p	PVCR
DBRTD with As_4 (Forward Bias)	0.76 ± 0.04	0.27 ± 0.03	48.8 ± 3.8	4.1 ± 0.1
DBRTD with As_4 (Reverse Bias)	-0.62 ± 0.03	-0.24 ± 0.02	-44.3 ± 3.5	4.5 ± 0.1
DBRTD with As_2 (Forward Bias)	0.71 ± 0.02	0.29 ± 0.01	52.2 ± 2.2	3.8 ± 0.2
DBRTD with As_2 (Reverse Bias)	-0.61 ± 0.02	-0.29 ± 0.01	-48.0 ± 2.1	4.4 ± 0.3

Table 4.1: Table of specific DC-IV parameters for a standard AlAs/GaAs DBRTD grown using the As_2 cell and the AlAs/GaAs DBRTD using the As_4 cell. Note that all the layer thicknesses including those in the quantum well are the same for both devices, as shown in Figure 4.8.

From the data, it is seen that there is no significant difference between the two devices. Although the deleterious effects mentioned above may exist in the samples grown with As_4, the AlAs/GaAs DBRTD devices do not appear to be sensitive to these effects.

4.10 Growth, Fabrication, and Testing of Depletion-Edge Modulation AlAs/GaAs QWITTS

The AlAs/GaAs QWITT structures, as described in section 4.5 and shown in Figure 4.5, utilized lightly doped GaAs spacer layers downstream from the AlAs/GaAs quantum well. In these devices, the depleted spacer layer length,W, was not optimized since the specific injection conductance, $\sigma(V)$, is a function of voltage. One way of optimizing the depleted spacer layer length as a function of voltage is to alter the doping profile downstream from the AlAs/GaAs quantum well. Ideally, one would like to start depleting the spacer layer downstream from the quantum well at the V_p and fully deplete the spacer layer downstream at the V_v. Therefore, one would like to have an optimized W at every voltage bias. To accomplish this task, the spacer layer doping profiles must be altered accordingly. In fact, to succeed in this endeavor requires precise knowledge of the doping concentrations, σ of the quantum well based on a linear fit, layer thicknesses, and repeatable specific contact resistivities. One way to prevent the electric fields from becoming too large before the DBRTD reaches resonance is by inserting a doping spike between the quantum well and the spacer layers [*KeN88*]. Since the electric fields in this region are usually $\geq$ 100 kV/cm, the doping spike can be used to reduce the fields and still maintain a saturation velocity of $\approx 6 \times 10^6$ cm/second. By adjusting the doping spike/depleted spacer layer length and doping accordingly, reduced voltage drops across the device can be obtained resulting in higher dc-to RF conversion efficiencies. Furthermore, if the depleted spacer layer doping is decreased or left nominally undoped, the depleted spacer length can be increased resulting in a larger ΔV and negative resistance. The device structures which utilize this doping spike and a longer depleted spacer layer length are referred to as <u>d</u>epletion-<u>e</u>dge <u>m</u>odulation QWITTs or DEM-QWITTs. Several structures have been designed through simulation involving the "baseline" quantum well structure, a doping spike and 1 μm nominally undoped downstream spacer layer. Simulations have also been used to design a structure where a composite spacer layer/doping spike/spacer layer (1500Å GaAs doped at 1×10^{16}cm^{-3}, 100Å GaAs doped at 5×10^{17}cm^{-3}, and 1 μm GaAs doped at 1.5×10^{16}cm^{-3}) is placed downstream from a quantum well of lower σ since the 1 μm depleted GaAs spacer layer cannot support quantum wells that supply higher current density, ($n > N_d$ must be avoided to prevent large electric fields). The quantum well structure consists of an 8 ML AlAs/ 18 ML GaAs / 6 ML AlAs layer structure grown by MBE, as shown in Figure 4.12. The layer structure, MBE growth, fabrication and device testing of the DEM-QWITT in this section follows those procedures described in section 2.2, section 2.3, section 4.7 and section 4.8.

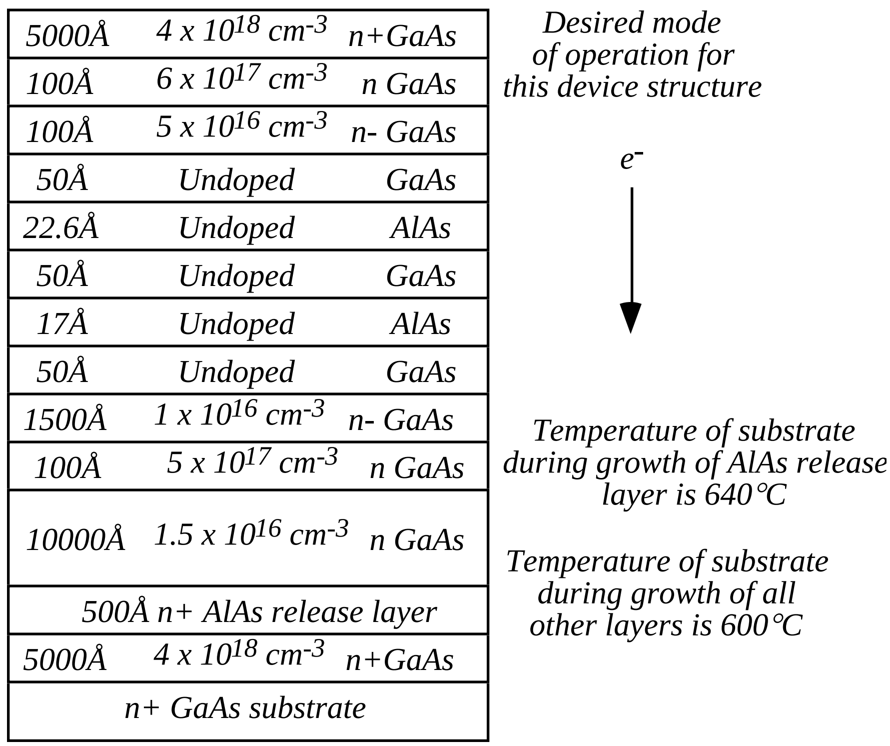

Figure 4.12: Cross-sectional layer schematic for a spacer layer/doping spike/spacer layer combination placed downstream (on the anode side) of an asymmetric AlAs/GaAs/AlAs quantum well. The current density is therefore reduced in such a structure, but the ΔV should be increased. The purpose of the AlAs release layer will be discussed in Chapter 6, section 6.3.

The results that were obtained from this device structure are shown in Figure 4.13 and Table 4.2. These results show the expected lower current densities and larger ΔVs. The unexpected results were the large amount of hysteresis which was observed in the device characteristics and the observation that many of the devices tested would break down when biased much further than the valley voltage. This hysteresis is due to significant series resistance that can occur from any undepleted GaAs in the 1 μm spacer layer.

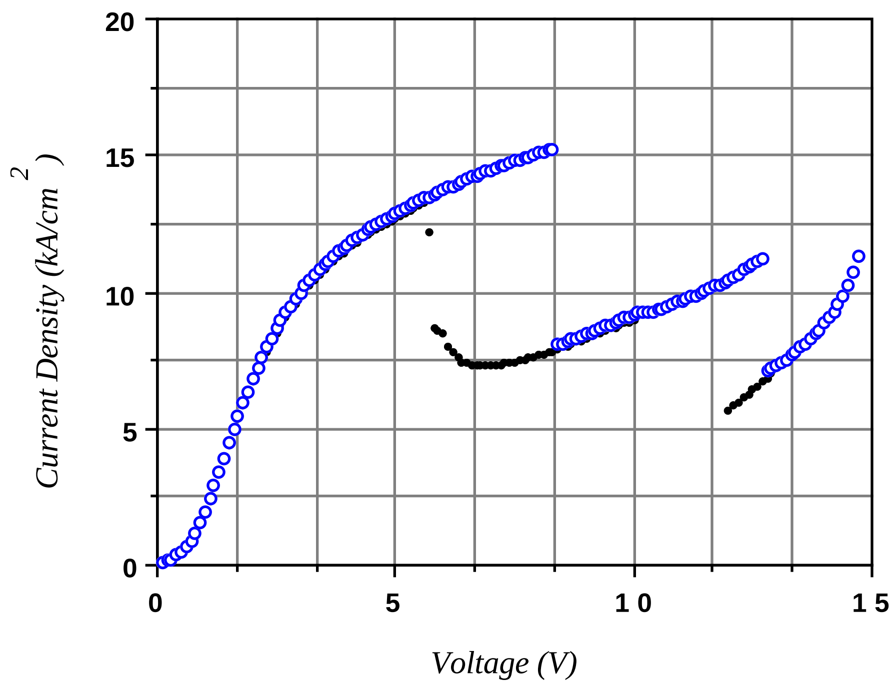

Figure 4.13: J-V characteristics of an AlAs/GaAs DEM-QWITT with a spacer layer/doping spike/spacer layer combination placed downstream (on the anode side) of an asymmetric AlAs/GaAs/AlAs (8ML/18ML/6ML) quantum well. Note that two measurements were taken as a result of the amount of hysteresis observed in the J-V characteristics; one curve with the voltage swept upward and one curve with the voltage swept downward. This hysteresis is a result of significant series resistance from any undepleted GaAs in the 1 μm spacer regions. Also note that the ΔV on such a structure was about 4 V.

Device Structure	Peak Voltage, V_p	$\Delta V=V_p-V_v$	Peak Current Density, Jp	PVCR
AlAs/GaAs DEM-QWITT (Forward Bias)	1.0 ± 0.04	0.2 ± 0.01	21.4 ± 0.8	1.5 ± 0.1
AlAs/GaAs DEM-QWITT (Reverse Bias)	-7.5 ± 0.41	-4.9 ± 0.35	-14.1 ± 1.2	2.6 ± 0.3

Table 4.2: Characteristic J-V data for the AlAs/GaAs DEM-QWITT with an asymmetric quantum well and a spacer layer/doping spike/spacer layer combination placed downstream from the quantum well.

It should be noted that other structures similar to the one above were grown and processed. The devices would breakdown before reaching the NDR regime. Thus, in this set of experiments, the ideal J-V characteristics for an AlAs/GaAs DEM-QWITT were not obtained as a probable result of not achieving the exact doping profiles desired in the doping spike/spacer layer regions.