Experiment E1: Quantum Semiconductor Devices

Duc Le

October 29, 2003

Abstract

Quantum effects are becoming more important in electronics as devices get smaller. This paper will discuss the tunnelling effects in semiconductors, and what factors affect them. We present experimental data that supports the predictions of quantum mechanics. Finally, we will discuss resonant tunnelling diodes that uses a quantum well to create a negative differential resistance, which can be used in oscillators and amplifiers.

1 Introduction

As semiconductor devices become ever smaller, they become more subjected to quantum mechanical effects, such as tunnelling. The nearly free electron model tells us that in semiconductors there is a small bandgap which electrons can cross due to thermal excitations. These electrons can then conduct because the conduction band into which they had crossed has many unfilled states. However, in different semiconductors, the bandgap, and hence the conduction band, is at different energies. If we place a semiconductor with a small band gap next to one with a large bandgap, then electrons moving from the first to the second would encounter a potential barrier. It is possible to produce such a barrier using III-V semiconductors such as GaAs, AlAs or AlGaAs. This is because these semiconductors have very similar lattice constants, so that no dislocations will be introduced when there is an abrupt change from one semiconductor to another. Furthermore, GaAs and AlAs have just such a bandgap difference as we want, with an energy difference in their conduction bands of $\Delta E_c = 1.1eV$.

Thus we can produce simple one dimensional barriers and wells by layering these semiconductors. This is done by Molecular Beam Epotaxy (MBE), with each layer deposited one atomic layer at a time. These layers can produce simple barriers, as shown in Figure 1 (a). We can also produce quantum wells by combining two barriers, as in Figure 1 (b). In this paper, we shall investigate quantum mechanical tunnelling across these barriers, and their potential application in microelectronics.

2 Quantum Tunnelling

2.1 Theory

Tunnelling occurs when electrons encounter a potential barrier with height greater than their energy. Classically the electrons may not pass through the barrier onto the other side. However, the wave-particle nature of electrons predicted by quantum mechanics requires that the wavefunction representing the electrons be continuous at the barrier interface. This means that the wavefunction dies away exponentially within the barrier region, rather abruptly go to zero at the interface. As the square of the wavefunction represents the probability of finding the electrons at a certain position, there is the probability that the electron can be found within the barrier. If the barrier is narrow enough, then the wavefunction will not have decayed to zero at the opposite interface, and hence the electron may be found on the other side of the barrier. The electron will thus have tunnelled through the barrier. Solving the Schrödinger equation for the quantum barrier, with the

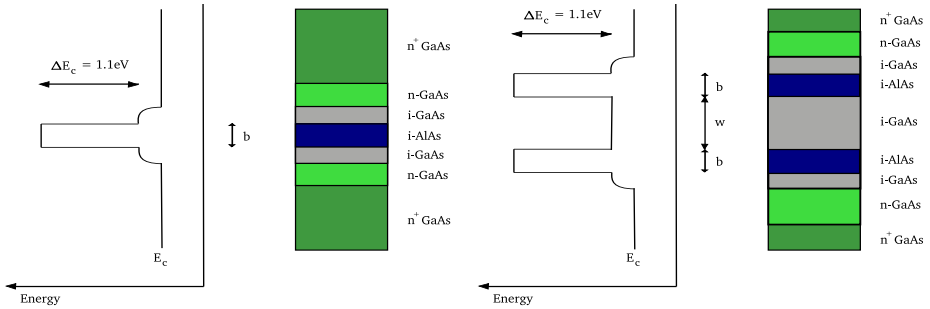

Figure 1: *Quantum barriers and wells produced by layering III-V semiconductors with different conduction band energies. The width of the single barrier is either 2nm or 3nm depending on devices. The width of the wells is either 6nm or 9nm. The doped layers surrounding wells allow electrons to flow to the wells. These layers are then connected by gold wire to soldered contacts with copper wire to our measuring equipment.*

appropriate boundary conditions we get the transmission coefficient as:

$$T = \left[1 + \frac{V_0^2}{4E(V_0 - E)} \sinh^2(kb)\right]^{-1} \tag{1}$$

where $k = \sqrt{\frac{2m^*(V_0-E)}{\hbar^2}}$, V_0 is the barrier height, b is the barrier width, E is the energy of the electron, and m^* is the effective mass of the electron. In the limit $kb >> 1$, $\sinh kb \to \frac{1}{2}e^{kb}$ so Equation 1 becomes:

$$T = \frac{16(V_0 - E)}{V_0^2} E e^{-2kb} \tag{2}$$

since $e^{2kb} >> 1$. In the low bias limit, $V_0 >> E$, so we can simplify Equation 2 further to $T \approx \frac{16E}{V_0} e^{-2kb}$.

Now the current density is given by $J = \int T(E) g(E) f(E) dE$ where $g(E)$ is the density of state and $f(E)$ is the Fermi function. The density of state is proportional to $E^{\frac{1}{2}}$. The Fermi function can be taken as unity from zero energy up to the maximum Fermi energy, E_f, at which all the states are filled, and zero beyond E_f. This is only strictly true at absolute zero, and at finite temperatures the step at E_f is somewhat blurred so that there are some states above the Fermi energy, and correspondingly fewer below it due to thermal excitation. However, we will

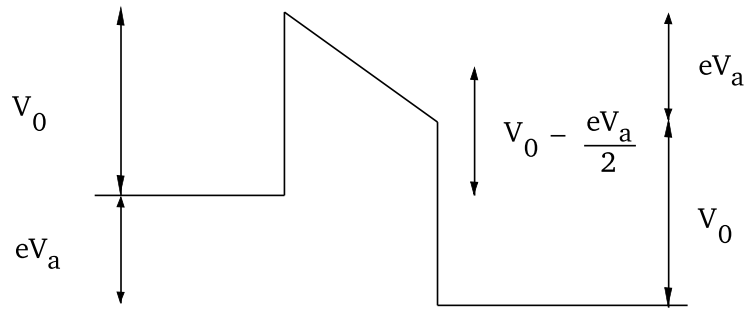

Figure 2: *With a voltage across it, the barrier behaves as shown, and now has an effective height of* $V_0 - \frac{eV_a}{2}$

take the situation at $0K$ as a first approximation. Substituting all of this into the integral we get:

$$J = const \times e^{-2kb} \int_0^{E_f} E^{\frac{3}{2}} dE = const \times e^{-2kb} \tag{3}$$

because E_f is a constant of the semiconductor. Now when we apply a voltage across the barrier, we effectively raise the potential of one side of the barrier with respect to the other, and so lower the height (energy) of the barrier by $\frac{eV_a}{2}$, where V_a is the applied voltage. This is summarised in Figure 2. Hence the energy of electrons effectively becomes $E + \frac{eV_a}{2}$, and the expression for k in Equation 1 becomes $\sqrt{\frac{2m^*}{\hbar^2}(V_0 - E - \frac{eV_a}{2})}$. Taking the natural logarithm of both side of Equation 3, and substituting in the equation for k, and taking $V_0 - E \approx V_0$, we obtain:

$$\ln J = \ln(const) - 2\sqrt{\frac{2m^*}{\hbar^2}\left(V_0 - \frac{eV_a}{2}\right)}b \tag{4}$$

Since the current, I, is proportional to the current density, J, we can plot $\ln I$ vs $\sqrt{V_0 - \frac{eV_a}{2}}$ to obtain a straight line with a gradient of $2\sqrt{\frac{2m^*}{\hbar^2}}b$.

2.2 Experiments

We used two types of barriers, one of which had a nominal width of 2nm and the other 3nm. The barrier itself is a layer of intrinsic AlAs semiconductor, sandwich between two layers of intrinsic GaAs. Outside these layers are layers of GaAs doped with silicon, to provide a conduction path to the quantum barrier. These doped layers are connected by gold wire to soldered contacts with the copper wire which connect the devices to voltmeters. The devices also had variable areas of semiconductor layers. These came in two sets - 10×10, 14×14 and 20×20 μm, and 30×30, 40×40 and 50×50 μm.

A sine wave signal varying from -2V to +2V at 10kHz was passed through the barrier. The voltmeters measured the voltage at the signal generator, V_a, and across a known resistance in series with the barrier, from which the current was calculated. This resistor was choosen so that the devices would not burn out, but would give high enough resolution in the results. It was found that 100Ω and $1k\Omega$ resistors worked well.

Figure 3 show the results for the wider, 3nm, barrier. The graphs do not show exact straight lines, but we only expect such behavior in the low bias limit, that is in the bottom right corner of the graphs. It is then from this region that we obtained the slope of the graphs, and hence an estimate of the barrier width, as shown in Figure 4.

We made similar calculations for the narrow barrier, of nominal width 2nm, but a lack of working devices meant we only obtained an estimate of 4 ± 1nm for the $14 \times 14\mu$m wafer. These result approach close to the nominal width, but there tends to be an overestimate with the larger sized wafers, and an under estimate with the smaller area ones. This may be due to the difference in the resistance of different area wafers. We can see from Figure 5, that the slope of the I-V curve increases with increasing area. Since the resistance is inversely proportional to the slope, the larger the area of the device, the smaller resistance it has. This resistance will alter our value of the applied voltage, V_a in Equation 4, and hence our width estimate. With the smaller area devices, the resistance will be higher, so we will measure a higher voltage for a given current. So, as V_a is larger, $\sqrt{V_0 - \frac{eV_a}{2}}$, our y-coordinates, become smaller, and we get a smaller, shallower slope. Since the width, b, is proportional to the slope, we get a smaller width for smaller sized devices.

In addition, we must also consider the doped layers on either side of the barrier, since they will have their own resistances, and the contact resistance with the connections to the voltmeters. When we measure V_a, we are effectively measuring the voltage across all of these components, rather than just across the barrier, as we had assumed in the derivation above. Also, we have to consider the effect of the two GaAs layer sandwiching the AlAs barrier. Although the con-

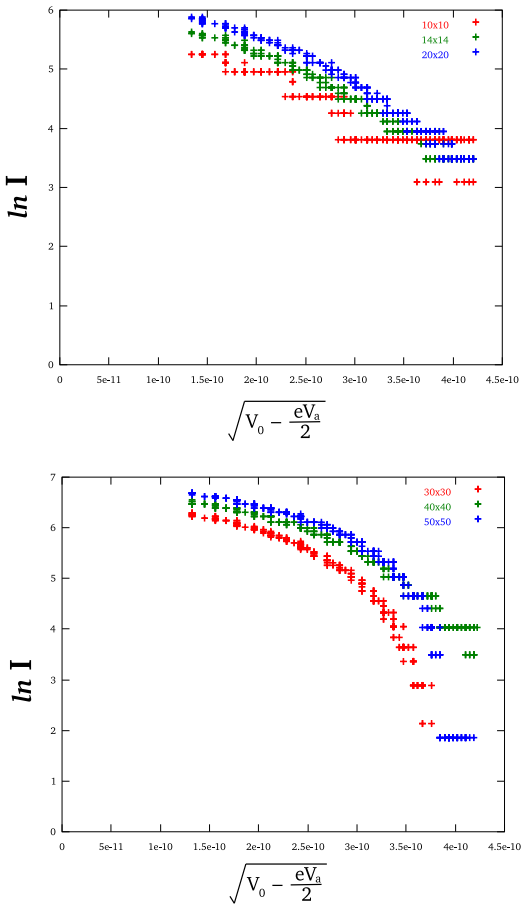

Figure 3: *The gradient of lower right part of this slope gives an estimate of the barrier width. This is the 3nm barrier, and the key shows the areas of the semiconductor layers, in microns. The horizontal areas arise due the lack of precission in the voltmeter in measuring very small currents.*

	10×10	14×14	20×20
Least Slope	0.83nm	1.11nm	1.52nm
Greatest Slope	1.42nm	1.33nm	2.46nm
Average	1.1 ± 0.3nm	1.2 ± 0.1nm	2.0 ± 0.5nm
	30×30	40×40	50×50
Least Slope	5.63nm	1.82nm	5.3nm
Greatest Slope	7.91nm	3.02nm	11.5nm
Average	6 ± 1nm	2.4 ± 0.6nm	8 ± 3nm

Figure 4: *Estimated widths of the wide barrier. The error estimates are obtained from comparing the greatest and least slopes.*

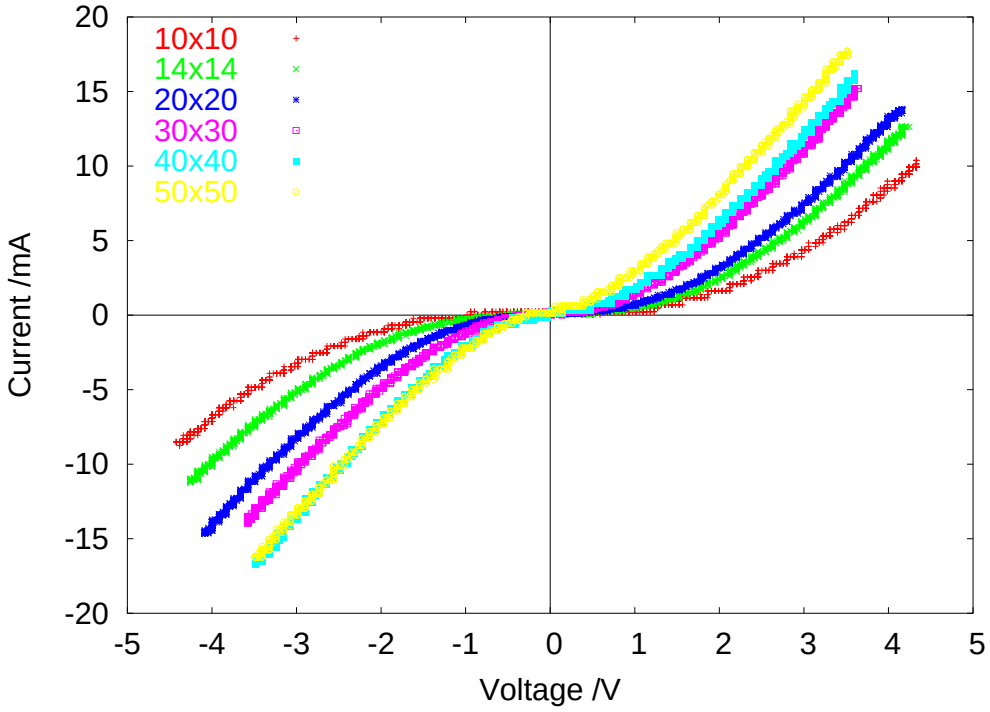

Figure 5: *The I-V curves of the variously sized area wafers of the wide single barrier*

duction band energy for GaAs is lower than AlAs, it is still higher than that of the doped layers surrounding, so that electrons flowing through may still perceive the doped/intrinsic GaAs contact as a potential barrier. This may cause us to over estimate the true width of the barrier. With this many factors contributing to the errors in the estimates of the barrier width, we nonetheless obtain values that are within an order of magnitude of the nominal width. This demonstrates that Equation 4 is indeed correct, within limits, and hence our concepts of the physics of the situation is applicable.

Finally, we consider the effects of temperature on the current-voltage behaviour of the single barrier. At high temperatures, such as room temperature, we expect there to be more electron scattering, and hence higher resistance due to thermal noise. Conversely we should see lower resistance at lower temperature. So we immersed our devices in a bath of liquid nitrogen at 77K, and took measurements as per the methods above. We found that the I-V altered as illustrated in Figure 6, with the 77K curve being shallower than the room temperature one. This means that it has a higher resistance, since resistance is inversely proportional to slope, rather than lower as expected.

We know that in order to tunnel through, the electron needs to have a high enough initial energy so that it will have a finite wavefunction at the other end of the barrier, after exponentially decaying within the barrier. So the more electrons we have at higher energy, the more will tunnel through and the higher our current, and lower resistance. So it would seem that at 77K the effects of fewer high energy electrons outweigh the effects of fewer phonon scattering in which the electrons would loose energy, giving us a higher resistance at lower temperatures.

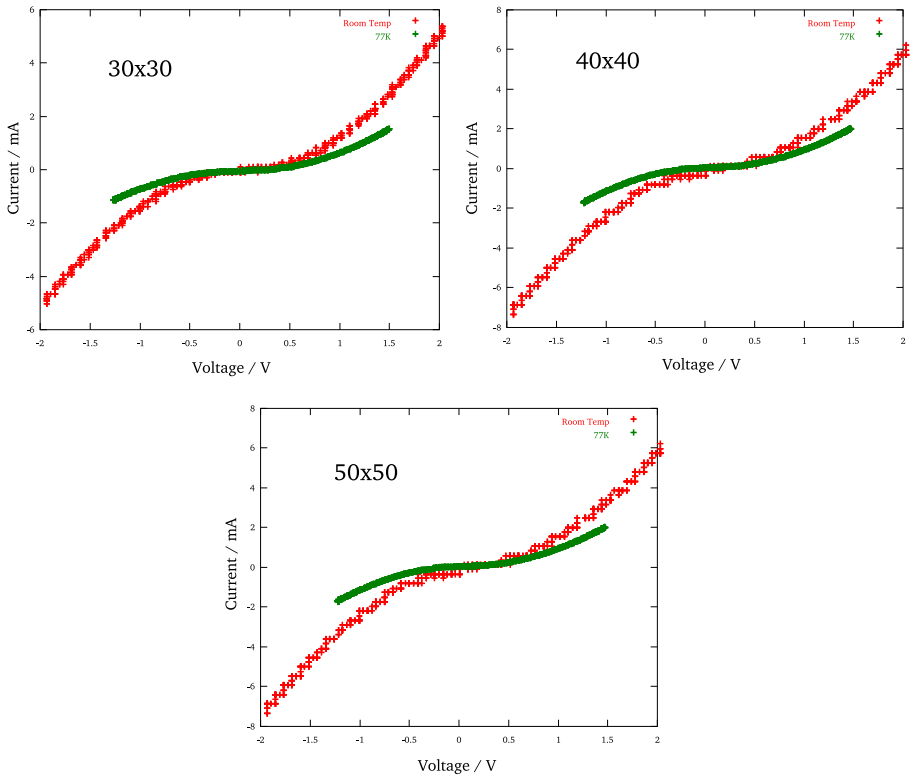

Figure 6: *A graph showing the effects of temperature on the I-V characteristics of a single barrier. The red curves represent roomt temperature, and the green ones 77K.*

3 Resonant Tunnelling Diodes

3.1 Theory

We now turn to the quantum well configuration, as shown in Figure 1. This configuration is used in a class of electronic devices called Resonant Tunnelling Diodes. We can derive the probability that an electron will tunnel through the well by applying its wavefunction to the well boundary conditions. It turns out that this probability, or transmission coefficient T, depends on the transimission, T_L for the left and T_R for the right barrier, and reflection coefficients, R_L, and R_R, of the electron at each barrier, and the phase, ϕ of the wavefunction.

$$T = \frac{T_L T_R}{(1-\sqrt{R_L R_R})^2 + 4\sqrt{R_L R_R} sin^2(\frac{\phi}{2})}$$

We can assume that the transimission and reflection coefficients is slowly varying with respect to energy, relative to the phase, and hence are approximately constant. Therefore, the transimission will resonate when $sin^2(\frac{\phi}{2})$ is zero, or when $\phi = 2n\pi$, giving a peak transmission of:

$$T_{res} = \frac{T_L T_R}{(1-\sqrt{R_L R_R})^2} \approx \frac{4T_L T_R}{(T_L + T_R)^2}$$

Further, if the barriers around the well are identical, then $T_L = T_R$, so T_{res} will be approximately unity. If we plot these resonant peaks on a transmission-energy graph, we find that they coincide with the energy levels inside the well. Hence the resonating case represents an electron which, once it tunnels into the well has one of the permitted energies, and so will certainly be transmitted. Electrons with different energy can only be transmitted if they loose (or gain) enough energy to enter one the allowed energy levels. They can do this through phonon or defect scattering. Phonon scattering is important at high temperatures (including room temperature) due to higher lattice vibrations, but at low temperatures, this is largely damped, so the imperfections in the lattice become relatively more important. As there are relatively fewer defects than phonons at the temperatures where we see phonons, so in addition to the resonant peaks, we should find that as the temperature (energy) increases, there is greater transmission, since there is more phonons for the electrons to scatter off and loose energy, as shown in Figure 7

We can vary the energy of the electrons by varying the applied voltage across the well, increasing the electron energy from E to $E + \frac{eV_a}{2}$ as mentioned in Section 2.1, where V_a is the applied voltage. As the transmission coefficient is proportional to the current, we should see a curve similar to Figure 7 in an I-V graph, with the peak voltage, V_p corresponding to the energy levels of the well. That

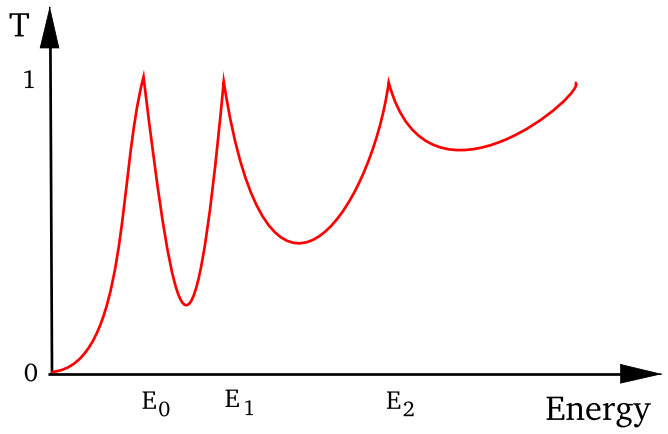

Figure 7: *The transmission coefficient as a function of energy. The resonant peaks correspond to allowed well energy levels. The general increase in T with energy (temperature) is due to greater number of phonons for electrons to scatter off, to enter the permitted energy levels.*

is for the first peak, and energy level, $E_0 \approx \frac{eV_p}{2}$. Since E_0 is small compared to the well height, V_0, we can approximate the well as one of infinite height, so the energy levels are given by:

$$E_n = \frac{(n+1)^2 \pi^2 \hbar^2}{2m_0 w^2} \tag{5}$$

where w is the well width.

3.2 Experiments

Using the same experimental setup as in Section 2, we measured the I-V characteristics of three different wells. The first has a wide well, $w = 9$nm, with narrow barriers at either end, $b = 2$nm. This gave a value of 0.17 ± 0.01eV for the smaller area device, and 0.31 ± 0.2eV for the larger area device, compared with a theoretical value of 0.13eV for E_0. The difference between the theory and experimental estimates may be due to how V_p was measured. We effectively measure the voltage across the well, its doped sandwhich layers and its contacts, rather than simply across the well itself. Since each of the outer layers add their own resistance - especially the contacts - we measure a voltage that is higher than that across the well only, and hence overestimate the energy level.

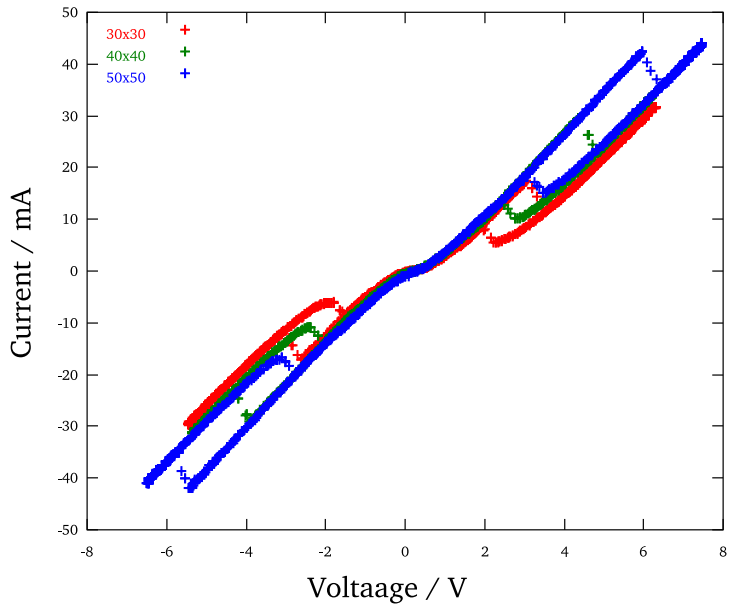

Figure 8: *Hysteresis in resonant tunnelling diodes. The hysteresis is more noticable they larger the area of the device.*

The narrower second well, with $w = 6$nm, and $b = 2$nm gave results of 2.1 ± 0.4eV, compared with a theoretical value of 0.3eV. This is much higher than even contact resistance can account for. The well also exhibited some hysteresis, as shown in Figure 8. The wide well also showed some hysteresis with the larger area devices (40×40 and 50×50 microns), but it was much smaller. The hysteresis loop seems to get larger with increasing size of device.

Since our estimate of the energy level, E_0, in the well depends on the peak voltage, V_p, which also increases as the hysteresis increases, then our estimates appear to increase with increase area of device. However, the energy level is quantised in the z-direction only, and is not affected by the x- and y-directions, and hence the area. Therefore, this must be an artifact of the experiment.

Hysteresis occurs because we are not passing a direct current through the diode, but rather a sine wave signal. This means the direction of current flow, and bias alters each half cycle. If the frequency of the signal is higher than the time for an electron to pass through the barrier, it becomes stuck in the barrier, and then moves in the opposite direction as the bias changes sign. However, we

can only have a finite number of electrons in the well in each energy level due to Pauli exclusion. This means the electrons stuck in the well prevents other electrons (from the reverse current) from entering the well, and hence reducing the current on the return leg of the hysteresis curve.

3.3 Current-Voltage Characteristics

We now turn to the other I-V characteristics of the resonant tunnelling diode. Figure 9 shows the I-V graph for a well with $w = 6$nm and $b = 3nm$. We see an initial peak, before V_p, that in the figure appears as a step due to the poor resolution of the voltmeter. This is due to the intrinsic GaAs layers sandwiching the AlAs barriers. These layers also have a conduction band energy higher than that of the surrounding doped layers, but lower than the AlAs barrier. Hence they appear as barriers to the lower energy electrons from the doped layers, just as the AlAs is a barrier to electrons from the GaAs layers. When we apply a bias across the device, we pull down one side of this smaller barrier, and so cause a small quantum well, with one side being the AlAs potential barrier, and the other barrier being the bias. This well will have energy levels and hence resonant peaks in the current as well, which is what we see.

The next noticable feature is a step after V_p. This is due to phonon scattering as electrons from a higher state looses energy and fall to the ground state, and then flow out of the well, and so maintain a current as the bias increase, even though the chemical potential is now below the energy level, so no electrons can tunnel through.

However, the most important feature of the graph is the Peak to Valley ratio after the negative differential resistance, which would determine the use that the diode can be put to in electronics. We found that the peak to valley ratio increased at lower temperatures. This is because there will be less excitation at lower temperatures, and so fewer electrons flowing after the chemical potential passes below the ground state of the well. Hence the current after V_p will be lower.

References

[Paul] D.J. Paul *Semiconducting Quantum Devices: Resonant Tunnelling of Electrons in Semiconductors* Experimental Brief, Cavendish Laboratory (2003).

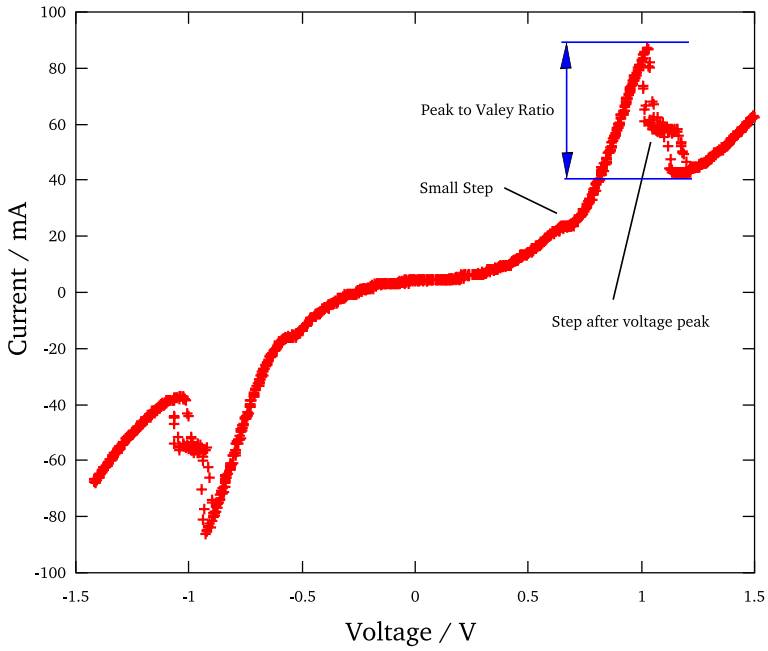

Figure 9: *A current-voltage plot for a resonant tunnelling diode. The device used has a well width of 6nm, the surround barriers are 3nm wide, and the area of the layers is* 20×20 *micron.*