Chapter 1

INTRODUCTION

Semiconductors rank as among the technologically most useful materials known to man. For a long time the most dominant semiconductor in the electronics industry has been silicon. Silicon has many useful properties, the most important of them being the existence of a very stable oxide [1]. Hence silicon is the semiconductor used in the metal-oxide-semiconductor (MOS) integrated circuit technology, the most important VLSI technology today. However, there are some useful characteristics that silicon does not possess, which has caused us to turn to compound semiconductors.

Binary compound semiconductors, such as GaAs and InP, have many useful properties [2] including a large, direct bandgap, high carrier velocity, and high electron mobility that find applications in microwave and optoelectronic devices. However, the fact that GaAs or InP does not have a stable native oxide [3] precludes its wide acceptance as a major semiconductor in VLSI applications requiring MOS transistors. Ternary compound semiconductors such as AlGaAs, InGaAs, or GaAsP have added a new dimension by permitting us to engineer the energy bandgap over a range of values [4, 5]. When a layer of a ternary compound semiconductor, such as AlGaAs, is placed next to a binary semiconductor, such as GaAs, to form a heterostructure, the energy band discontinuity at the interface gives rise to interesting transport phenomena that have been known for a long time. The original William Shockley transistor patent in 1951 mentioned a "wide-gap emitter device" [6], and in

1957 Herbert Kroemer proposed his theory of the wide-gap emitter [7]. This search for high-performance, high-frequency devices based on new heterostructure transport phenomena has resulted in more stringent structure requirements [8] at small dimensions. Although many thin-film technologies, such as vapor phase epitaxy (VPE), liquid phase epitaxy (LPE), chemical vapor deposition (CVD), sputtering, and vacuum evaporation have been investigated [9] to fabricate these structures, it was not until the advent of molecular beam epitaxy (MBE) [10] and metal-organic chemical vapor deposition (MOCVD) [11, 12] that devices using transport properties in heterostructures became a reality. Figure 1.1 indicates the pattern of current research in the field of heterostructures and reduced dimensionality systems--physics, materials, and devices.

MBE is an epitaxial growth process involving the reaction of one or more thermal beams of atoms or molecules with a crystalline substrate under ultrahigh vacuum conditions. It permits precise control of the beam fluxes and deposition conditions [8], resulting in abrupt changes in chemical composition and doping profiles over a few atomic layers. Conventional microwave devices such as the hyperabrupt varactor, IMPATT (impact ionization avalanche transit time) diode, mixer diode, field-effect transistor, metal-insulator-semiconductor (MIS) devices, bipolar transistors, Gunn diodes, and charge-coupled devices (CCD) fabricated by MBE have either equaled or excelled the performance obtained from these devices made by other fabrication technologies.

Novel heterostructure devices unique to MBE and MOCVD, such as the heterostructure bipolar transistor (HBT), high electron mobility transistor (HEMT), superlattice avalanche photodiode (APD), and quantum-well devices have been the

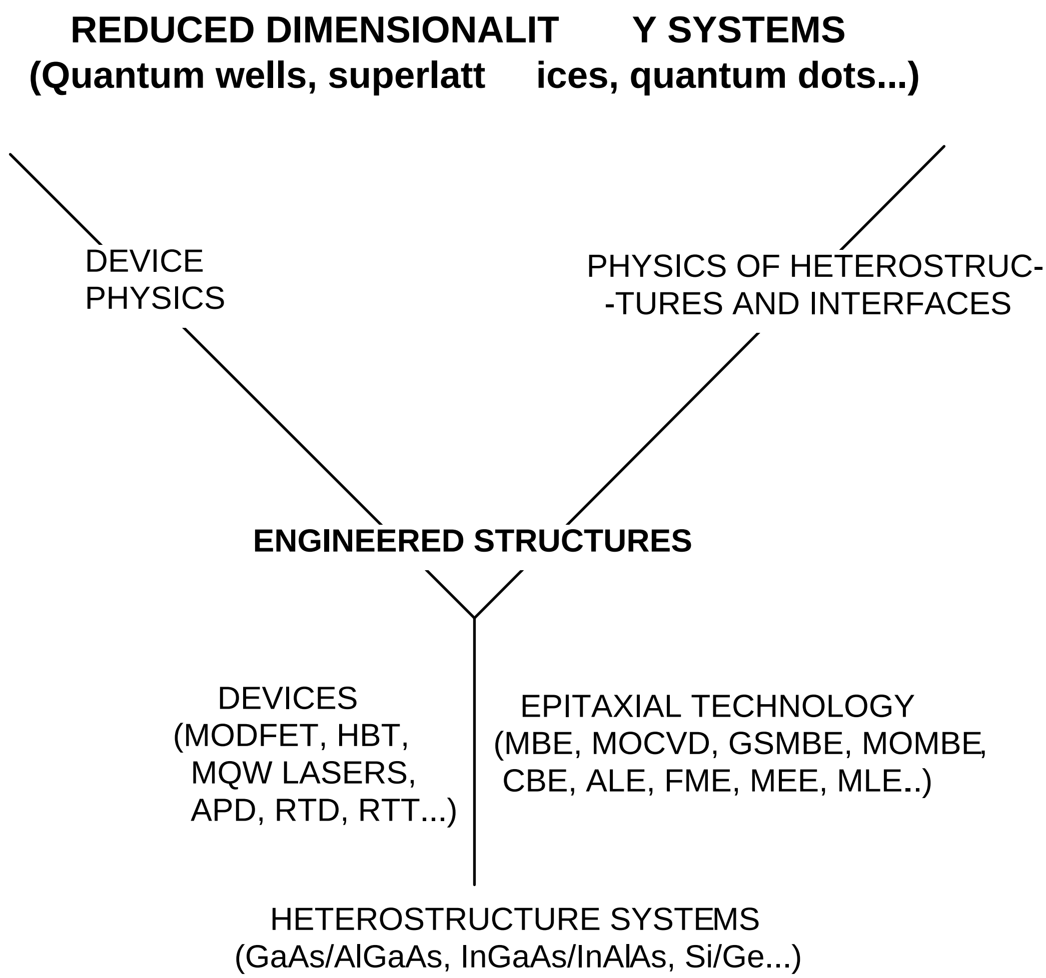

Fig. 1.1: Pattern of current research in semiconductor heterostructures. Adapted from Ref. 45.

subject of numerous studies recently. For microwave and millimeter wave applications, HEMTs show promise as low noise devices [13] and HBTs find their niche in high power applications [14].

Quantum well devices being investigated include quantum well lasers [15], resonant tunneling/double barrier diodes (RTD/DBD) [16, 17], and quantum well transistors [18, 19]. As digital devices beyond the post-shrink era are being proposed and fabricated, functional integration achieved through the use of quantum well transistors looks promising. Resonant tunneling diodes and quantum well transistors are also finding acceptance as microwave and millimeter wave devices [20, 21].

1.1 Brief Review of Microwave/Millimeter Wave Oscillators

The ability to generate energy at high frequencies (1-1000 GHz) is desirable for a number of military, industry, and commercial applications. Diodes, and in particular the IMPATT (impact ionization avalanche transit time) diode and the Gunn diode, have been the main sources of power at frequencies between 1-220 GHz [21-23]. The IMPATT diode is particularly useful for high power applications from 1-60 GHz, but at higher frequencies the noise inherent in the avalanche multiplication process is a major limitation. The highest frequency of oscillation obtained from a IMPATT diode is 220 GHz [23], with cut-off frequencies being limited by avalanche induced dispersion [24] and carrier energy relaxation effects [25]. The Gunn diode is particularly useful for low noise applications at 94 GHz, and the highest frequencies of oscillation are around 200 GHz for InP Gunn diodes. Other microwave diode structures investigated are the BARITT (barrier injection and transit time) diode, the TUNNETT (tunnel injection transit time) diode, and the TRAPATT (trapped plasma avalanche triggered transit) diode, but these devices have shown limited promise. Carrier injection over a potential barrier in the BARITT diode results in an unfavorable injection angle of $\pi/2$ [26], and its high frequency performance is limited to about 60 GHz by low drift velocities and high diffusion coefficients around the injection point. In contrast tunneling, such as that used in the TUNNETT diode is a low noise process that appears capable of providing reasonable high frequency performance. However, poor injection phase delay in the TUNNETT diode affects its efficiency of operation [27], and pure large-signal TUNNETT mode operation in homojunction devices may be virtually unattainable, due to difficulties in obtaining band-to-band tunneling while avoiding impact ionization. Hence, it is clear that the need for a stable, low noise millimeter wave source at frequencies greater than 100 GHz still persists. A local oscillator at these frequencies would be desirable for a number of applications such as telecommunications, space exploration, atmospheric propagation studies, plasma diagnostics, and various military needs. Resonant tunneling diodes (RTDs), which use the phenomenon of tunneling through a bound state in a quantum well to produce a dc negative resistance, are a promising way of realizing a millimeter wave oscillator at frequencies between 100-1000 GHz.

1.2 A Resonant Tunneling Diode Oscillator

The phenomenon of resonant transmission of electrons through potential barriers can be understood using the Kronig-Penney model of a one-dimensional crystal [28]. According to this model, a one-dimensional crystal exhibits allowed bands of perfect transmission for electron waves separated by forbidden bands of attenuation [29]. If a thin region of a small bandgap semiconductor is sandwiched between two similar regions of large bandgap material, distinct energy levels are created in the small bandgap semiconductor. Interesting transport properties such as resonant tunneling through these energy states, giving rise to negative differential conductivity, can be observed and explained on the basis of this model [29, 30].

A typical quantum well heterostructure (QWH) consisting of one quantum well of GaAs between two $Al_xGa_{1-x}As$ barriers is schematically shown in Fig. 1.2(a). From this figure we can see that when particular values of dc bias are applied across this structure, the injected carriers have certain energies corresponding to the quasistationary energy states in the potential well. For such a bias condition, resonant tunneling can occur giving rise to a large current. Equivalently, the electron wave function reflected at the first barrier can be considered to be cancelled by the wave which leaks from the well in the same direction, causing resonant tunneling. As the dc bias is increased the bands pass each other, causing a decrease in the tunneling current. This initial increase in current is followed by a subsequent decrease, as the dc bias is increased further. The result is a negative resistance

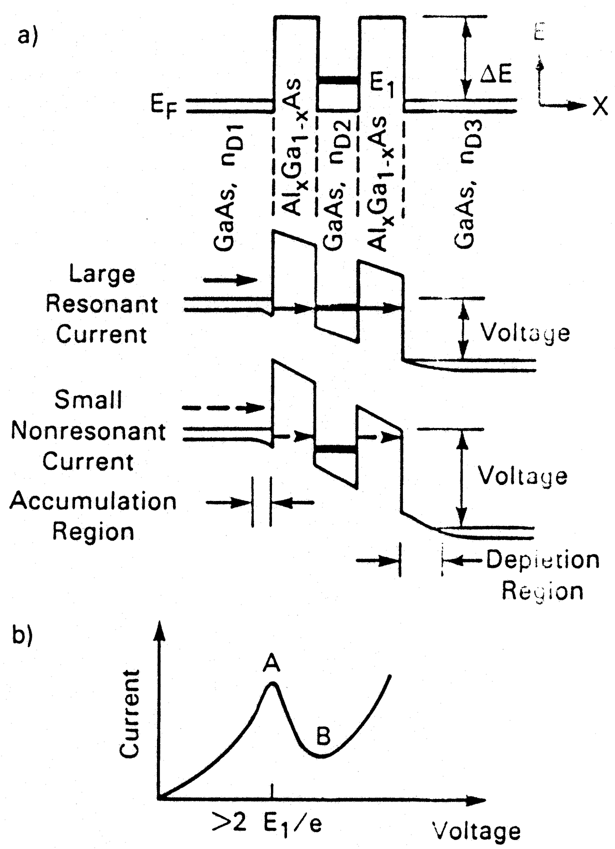

Fig. 1.2 (a): Electron energy as a function of position in the double barrier resonant tunneling structure. The energy level E_1 occurs above the bottom of the bulk conduction band because of confinement in the x direction.
(b) Diode current as a function of incident electron energy. From Ref. 78.

region, as shown in Fig. 1.2(b). We can see from the figure that for an energy state E_1 in the well, based on this simple picture the negative resistance occurs at approximately a value of applied voltage $>2E_1/q$ because the voltage drop across each of the two barriers is at least half the applied voltage. Thus, for this case the energy bands in the barrier shift only half the amount of applied voltage causing tunneling at energy E_1. A solution to the wave equation for these square potentials shows that there are peaks in the transmission coefficient corresponding to the energy levels in the quantum well. This solution is very similar to the analysis of wave interference in a Fabry-Perot interferometer where similar resonance peaks in the transmission coefficient are observed [31]. Hence tunneling in these double barrier quantum wells has been referred to as 'resonant tunneling'.

The negative resistance characteristics obtained from a resonant tunneling diode can be used to build a microwave oscillator by embedding the device in a circuit appropriately designed to match the impedance of the diode at the desired frequency of operation.

1.3 The Quantum Well Injection Transit Time (QWITT) Diode

1.3.1 Introduction

There are several reasons to expect that for any given state of semiconductor technology the highest frequency solid state sources will be negative resistance transit time diodes. One reason is that such devices use an effect (the transit time delay) which limits the performance of other devices. This transit time delay, in combination with any delay due to the current injection mechanism, causes the induced particle currents to lag behind the ac voltage in the device. If the total delay is appropriate, a dynamic negative resistance results. The principles of operation of existing transit time negative resistance devices, and the various limitations arising from the nature of the injection mechanism used, have been reviewed in the literature [32]. Resonant tunneling diode oscillators are intrinsically low power [33-36], due to low dc bias voltages and area and impedance limitations associated with the capacitance of very narrow junctions.

We propose here a new transit time device which uses resonant tunneling through a single quantum well to inject carriers into the drift region of the device [37, 38]. Compared to other transit time devices this device uses an injection mechanism which may provide more favorable injection angles, has superior high frequency characteristics, and is relatively low noise. Depending on the voltage drop across the quantum well, the device may inject single current pulses peaked at $\pi/2$ or $3\pi/2$, or multiple pulses peaked at $\pi/2 \pm \delta$ or $3\pi/2 \pm \delta$. Compared to double barrier diodes (DBDs), the presence of a drift region increases both the specific negative resistance and impedance of the device so higher rf output power can be expected. Since this is a quantum well injection transit time effect device, we call it a QWITT diode.

1.3.2 Basic Device Description

The physical structure of this device, consistent with a typical molecular beam epitaxy growth sequence, is shown in Fig. 1.3. The device structure consists of a single GaAs quantum well between two $Al_xGa_{1-x}As$ barriers, together with a drift

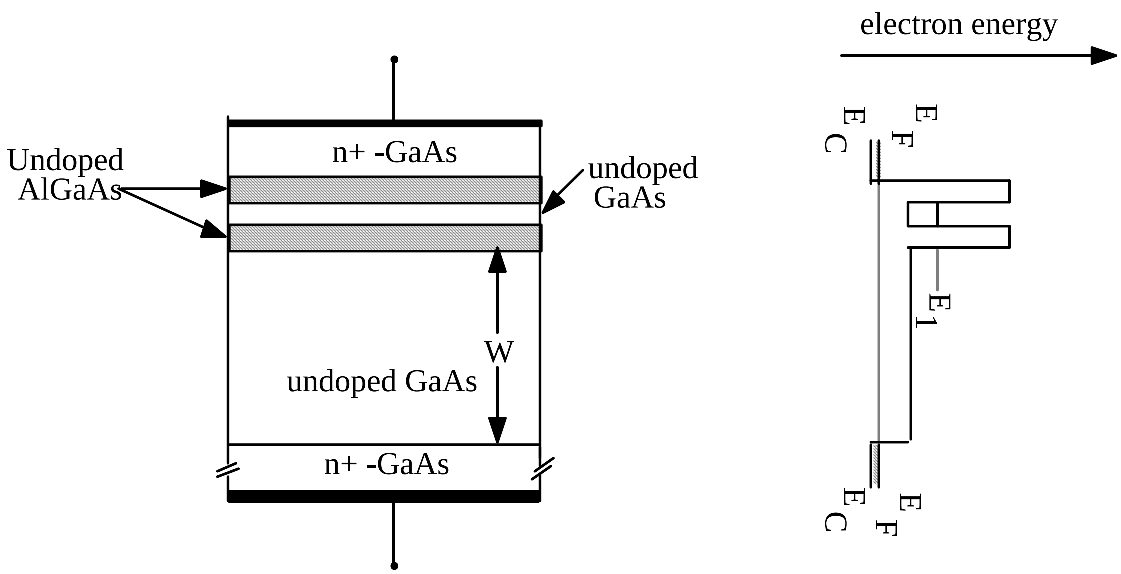

Fig. 1.3: Physical structure of the QWITT diode and energy band diagram of the device when no bias is applied. The length of the transit time region, W, would be much greater than the quantum well thickness for typical millimeter wave frequencies.

region of undoped GaAs. This structure is then placed between two n^+-GaAs regions to form contacts. The energy band diagram of the device, when no dc bias is applied, is shown in Fig. 1.3. For an isolated quantum well, when the voltage across the well equals $2E_1/q$, a large peak in the current through the device occurs. In the QWITT diode, an idealized picture of operation of this device can be explained using Figs. 1.4 and 1.5. In order to achieve maximum resonant tunneling of electrons through the well at the correct time, for any desired injection phase angle ϕ, the dc bias must be adjusted so that the following expression is satisfied (Fig. 1.4b):

$$2E_1/q = V_{RF}\sin\phi + V_i \qquad (1.1)$$

where V_{RF} is the amplitude of the ac voltage across the quantum well, V_i is the effective dc voltage across the well, and E_1 is the energy level in the well. Note that these are the voltage drops across the injection region only. Assuming a constant electric field throughout the injection and drift regions, the ratio of the voltage drops across the two regions is simply the ratio of their thicknesses. A typical quantum well structure is about 0.015 µm thick, and typical drift region thicknesses are 0.2 to 5.0 µm, depending on the desired frequency of operation. The voltage across the quantum well will thus be much less than the total voltage across the entire device. If an appropriate bias is applied so that the Fermi level in the n^+-GaAs is below the energy level E_1 by an amount V_{RF} (Figs. 1.4a & 1.5a), and an ac signal $V_{RF}\sin\omega t$ is superimposed on this bias, then resonant tunneling of electrons through the well will peak at an injection angle of $\pi/2$ in the ac cycle (Figs. 1.4b & 1.5b). The electrons will then traverse through the drift region at their saturation velocity, resulting in a dynamic negative resistance. Although in this mode of operation the

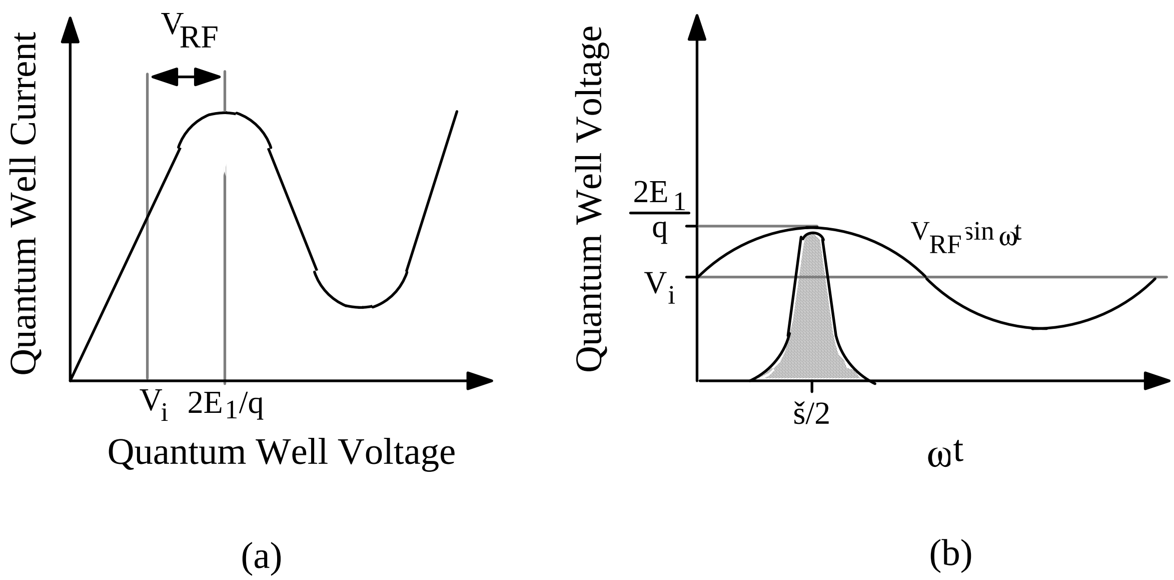

Fig. 1.4: a) The bias point for the QWITT diode is adjusted so the bias voltage across the quantum well injection region, V_i, is less than the resonant bias by an amount equal to the rf voltage across the well, V_{RF}; b) current injection through the well will then peak at $\omega t = \pi/2$ when the total instantaneous voltage across the well is equal to $2E_1/2$. Note that the I-V curve shown is for the quantum well injection region only, and does not represent the I-V curve for the complete QWITT diode.

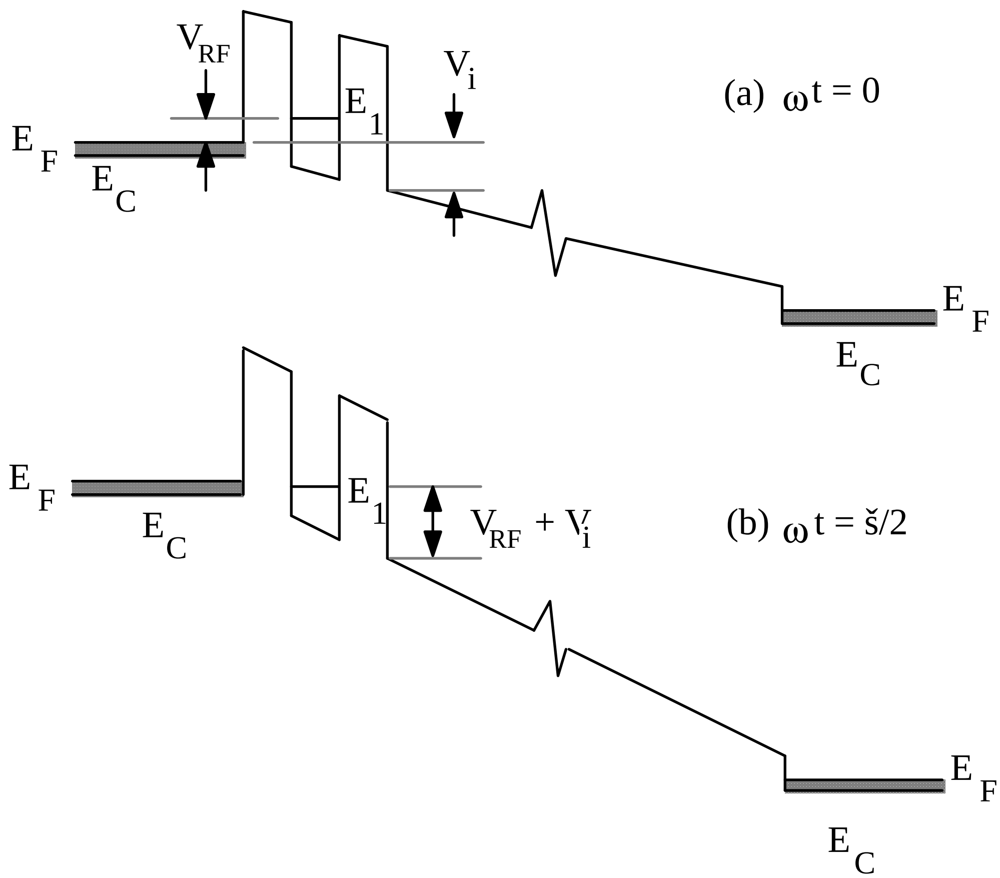

Fig. 1.5: (a) Schematic energy band diagram at ωt=0 when the dc voltage drop across the quantum well V_i is below resonance by an amount equal to the amplitude of the rf voltage V_{RF}; (b) At π/2 in the rf cycle the well will now be at resonance, causing current to be injected into the transit time region of the device. In actual operation the device would not be at equilibrium, and the electric field in the drift region would be changed by the presence of the drifting current pulse.

QWITT diode exhibits an injection angle of $\pi/2$, like that in the BARITT and idealized TUNNETT diodes, it would not suffer from the other limitations associated with these devices.

The QWITT diode can also be biased to move the Fermi level in the n^+-GaAs above the energy state in the GaAs well so that current injection peaks around $3\pi/2$ in the negative half of the ac voltage cycle. Compared to injection near $\pi/2$, this mode may have lower output power (due to a shorter drift region sustaining a lower dc voltage, and requiring a reduced area due to impedance constraints) but it should have higher efficiency [32]. Note that no other transit time device exhibits an injection angle of greater than π. For an injection angle of $3\pi/2$ the maximum negative resistance would occur for a transit angle of about $\pi/2$. As the transit angle decreases, the length of the drift region is reduced and one can envisage this device looking more like a pure resonant tunneling oscillator [33-35]. It has been reported that in order to achieve oscillations near 40 GHz with a resonant tunneling quantum well oscillator, 300 nm thick, $2\text{x}10^{17}\text{cm}^{-3}$ doped GaAs buffer layers between the AlGaAs barriers and the n^+-GaAs contacts were necessary [39]. Since these layers would be expected to contribute a small transit time delay, it is possible that this device was actually operating near a $3\pi/2$ injection mode, rather than in a pure quantum well oscillator mode. However, it is not clear to what extent the advantages of buffered resonant tunneling diode oscillators arise from (i) the improved impedance matching of a thicker structure, (ii) higher dc and rf voltages achievable in a thicker structure and (iii) transport through a depletion region on the anode side causing transit time effects. As we shall see in the later chapters, this could be resolved experimentally by comparing the performance of structures using one-sided buffers, placed only before, or after, the quantum well. The QWITT structure should permit maximum performance optimization by considering all the factors mentioned above.

In existing transit time devices the injection mechanisms used produce broad current pulses [22]. Hence, it should be emphasized that the QWITT concept does not depend on an extremely sharp resonant tunneling current peak. Small-signal analysis has been performed, and large-signal analysis is being performed, based on experimental J-V curves [35, 40-42] of quantum well diodes. These results demonstrate the validity of the QWITT concept using injection through existing quantum well structures.

1.3.3 Device Design Considerations

The design issues specific to the QWITT diode center on the new features of (i) the choice of optimum drift/depletion region length, (ii) quantum well design for injected pulse optimization, (iii) the choice of injection phase angle, as discussed above, and (iv) multiple pulse injection during an rf cycle. In order to obtain the optimum depletion region length self consistent knowledge of the carrier concentration and velocity in the presence of time varying fields will be required. The choice of quantum well dimensions for the QWITT diode presents several non-trivial problems. The well must produce suitable injected carrier pulses, at dc bias levels consistent with required drift region fields. There will be trade-offs between the number of carriers in an injected pulse and the sharpness of the pulse, as determined by the quantum well and barrier thicknesses [40, 43]. Another design issue unique to the QWITT diode is that for any combination of dc and rf voltage there are two points in the rf cycle at which the injection peaks, except for degenerate injection around $\pi/2$ or $3\pi/2$. In general, then, there will be two injected carrier pulses per rf cycle. Idealized drift region optimization assuming equal amplitude pulses injected symmetrically at $\pi/2 \pm \delta$ or $3\pi/2 \pm \delta$ is straightforward. In practice, however, the field perturbation due to the space charge of the first pulse will mean that the pulses will have different shapes, and will not be injected symmetrically about a terminal voltage extremum. This field perturbation due to the first pulse may enhance or quench, and will certainly phase-shift, the second pulse.

Several other factors must also be considered in the design of the device. Operation at submillimeter wave frequencies will involve very thin drift regions. In this case, it may be necessary to use a heterojunction anode to avoid Debye-tail smearing of the drift region boundary. Device performance will be significantly affected by transient and quantum transport at this interface [44]. Additional consideration may have to be given to the theory of tunneling in the presence of time-dependent fields.

In summary, compared to other transit time devices and resonant tunneling diode oscillators, the QWITT diode should have several advantages. Quantum mechanical tunneling in the QWITT diode is a low noise injection mechanism with superior high frequency characteristics. The presence of the depletion/drift region should enhance the specific negative resistance available from the device compared to a resonant tunneling diode. Depending on the bias level, the device may permit injection of carriers into the drift region at more favorable phase angles than other transit time devices, so higher efficiencies can be expected. The characteristics of the injection region, and therefore the injected current pulse, can be adjusted by changing the physical dimensions of the quantum well structure. Thus, both the length of the drift region and the shape of the current pulse can be optimized to obtain the best power-frequency performance from the device. The QWITT diode should extend the normal frequency limit associated with transit time devices, while providing higher output power than resonant tunneling diodes.

1.4 Thesis Overview

The goal of this research effort was to analyze, design, fabricate, and test a QWITT diode oscillator.

1.4.1 Contributions of This Work

In order to accomplish the goal of realizing a QWITT diode oscillator, at first small signal and large signal models for the QWITT diode that relate physical device parameters to the dc terminal characteristics and predicted rf performance, were developed. These models were then used to provide optimum device dimensions to maximize rf performance at a desired frequency of operation. The large signal device model was also used to design a planar QWITT diode oscillator.

Initially GaAs/AlAs resonant tunneling diode structures were grown by MBE (molecular beam epitaxy) and devices with good room temperature negative resistance characteristics were fabricated. A number of QWITT devices were then fabricated and planar and waveguide oscillator circuits were designed to test these devices at frequencies between 1-35 GHz. The dc and rf characteristics of these devices were compared to predicted device behavior based on the device models developed. Improved QWITT structures designed to obtain higher dc-to-rf conversion efficiency were fabricated and tested.

The rf performance of QWITT devices was shown to be significantly better than resonant tunneling diodes. At microwave frequencies (8-12 GHz) the highest output power density obtained to date from a quantum well oscillator was achieved during the course of this work. The first implementation of a quantum well oscillator in a planar circuit was also reported. Self-oscillating mixers in waveguide and planar circuits, using the QWITT diode as the LO (local oscillator), have also been fabricated. Conversion gain at X-band (8-12 GHz) was observed in both waveguide and planar circuits over a wide range of intermediate frequencies (10-1000 MHz). This is the first report of conversion gain obtained from a self-oscillating mixer using a quantum well device.

1.4.2 Thesis Organization

Chapter 1 has discussed the phenomenon of resonant tunneling, and introduced the concept of a QWITT diode oscillator. A brief overview of heterostructure devices was also presented.

Chapter 2 presents the small signal and large signal behavior of the QWITT device. In particular, device design criteria to optimize rf performance are presented.

Chapter 3 deals with MBE of $GaAs/Al_xGa_{1-x}As$ heterostructures. Device fabrication procedures employed in fabricating QWITT diodes are described.

Chapter 4 discusses the dc characteristics of a number of resonant tunneling diodes (RTDs) and QWITT devices fabricated in the course of this study. Terminal characteristics are related to the physical structure of the devices studied.

Chapter 5 presents the rf performance of the QWITT diode oscillator at microwave and millimeter wave frequencies in waveguide and planar circuits. Results on a self-oscillating QWITT diode mixer are also presented.

Chapter 6 looks at promising research directions in three major areas--(a) alternative heterostructure material systems, (b) improved device structures and transport models, and (c) improvements in RF circuit design and testing.