UNIVERSITY of CALIFORNIA
SANTA CRUZ

RESONANT TUNNELING AND APPLICATIONS

A thesis submitted in partial satisfaction of the
requirements for the degree of

BACHELOR OF SCIENCE

in

PHYSICS

by

Adrian Shestakov

June 2018

The thesis of Adrian Shestakov is approved by:

Professor Michael Dine
Advisor

Professor David Smith
Senior Theses Coordinator

Professor Robert Johnson
Chair, Department of Physics

Copyright © by

Adrian Shestakov

2018

Abstract

Resonant Tunneling and Applications

by

Adrian Shestakov

Resonant tunneling is a quantum mechanical phenomenon which occurs in a quantum tunneling system with two or more barriers separated by a quantum well. Resonant Tunneling Diodes (RTDs) are a type of tunnel diodes which employ such a system. They exhibit nonlinear electrical properties which can be utilized in many different ways. RTDs can operate at terahertz frequencies and can be used to supply gain to an electrical signal. This paper seeks to examine the quantum mechanics of resonant tunneling in a thorough and accessible way, to connect the quantum mechanics to the electrical properties of RTDs, and to describe some of their applications.

Contents

List of Figures v

List of Tables vi

Acknowledgements vii

1 Introduction 1

1.1 Resonant Tunneling . . . 2
1.2 Wave Packets . . . 5
1.3 Resonant Tunneling Diodes . . . 6
1.4 History and Alternative Theories . . . 9
1.4.1 McMillan Anderson Model . . . 10
1.4.2 Gadzuk Model . . . 11

2 Quantum Mechanics of Resonant Tunneling 12

2.1 Incident Wave Packet . . . 15
2.2 Dwell Time . . . 16

3 RTD operation 22

3.1 Switching . . . 26
3.2 RTD Static Memory Operation . . . 29

4 Conclusion 34

Bibliography 35

List of Figures

1.1 Single Barrier Quantum Tunneling . . . 2
1.2 Quasi-bound States Summary Plot . . . 3
1.3 Regional Wave Function Forms . . . 4
1.4 Motion of a Gaussian Wave Packet . . . 6
1.5 RTD Biases . . . 8

2.1 Double Barrier Tunneling Probability . . . 14
2.2 Schematic Probability Peaks . . . 15
2.3 Simulated Wave Packet Evolution . . . 18
2.4 Dwell Time v. Barrier Thickness . . . 21

3.1 Single Peak I-V Curves . . . 24
3.2 RTD Small-Signal Model . . . 25
3.3 RTD Switching Circuit and Large-Signal Equivalent . . . 27
3.4 Piecewise Linear I-V Approximation . . . 29
3.5 RTD Memory . . . 31
3.6 RTD and 6T-SRAM Memory Cells . . . 32
3.7 Cascade RTD Memory Cell . . . 33

List of Tables

2.1 Dwell Times 20

3.1 RTD Structure 23
3.2 RTD Current Values 26

Acknowledgements

I owe a deep debt of gratitude to my father, to my dear friend Natalie Moore, and to my advisor Professor Michael Dine, all of whom made the completion of this paper possible.

1

Introduction

Quantum tunneling is a familiar phenomenon from the undergraduate physics curriculum. When a particle is incident on a barrier with higher energy than that of the particle, the particle has a nonzero probability of tunneling through the barrier and appearing on the other side. This happens despite the particle not having enough energy to cross the barrier according to classical physics. This is explained mathematically using the particle's wave function, Ψ. In the classically forbidden region the particle's wave function decreases exponentially as a function of distance. It does not however, equal zero, so for barrier of finite thickness the amplitude of the wave function is exponentially decreased but still nonzero on the other side of the barrier, as shown in Figure 1.1. Devices that make use of tunneling typically have barriers on the order of a few nanometers thick.

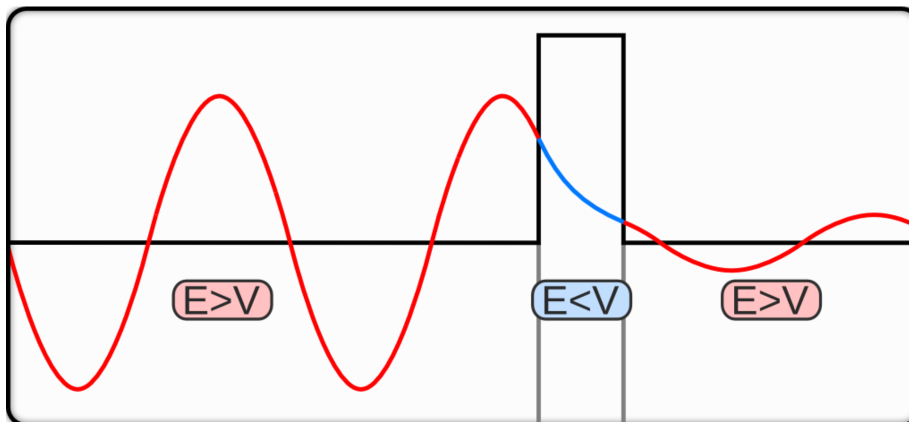


Figure 1.1: In the classically allowed (E>V) regions a particle's wave function is of the form e^{ikx} where $k = \frac{p}{\hbar}$ is the wavenumber of the particle. In the classically forbidden region (E<V) the wavefunction is of the form $e^{-\kappa x}$ [1].

Quantum Tunneling is the mechanism behind nuclear fusion in stars and radioactive decay. Many instruments, scanning tunneling microscopes and tunneling diodes for example, make use of tunneling. The application of interest for this paper is tunneling diodes with two potential barriers. These diodes, known as resonant tunneling diodes (RTDs), exhibit interesting quantum behavior and possess useful nonlinear electrical properties to be described in later sections.

1.1 Resonant Tunneling

The particular phenomenon of interest for this paper is resonant tunneling. Resonant tunneling occurs in systems with two or more potential barriers separated by quantum wells. In the quantum well there exist quasi-bound states which are analogous to bound states in a finite or in an infinite well. An example of this is pictured in Figure 1.2

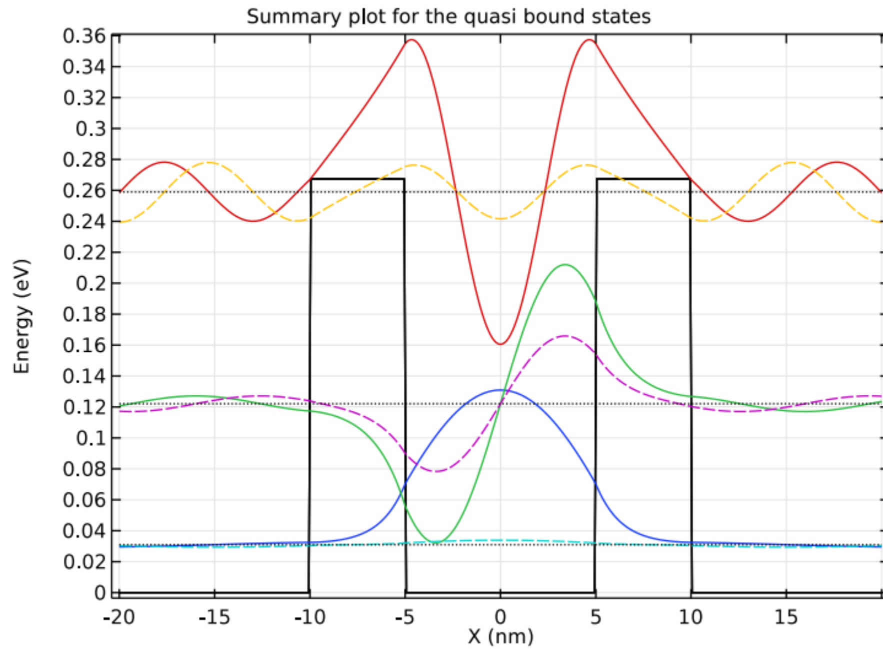


Figure 1.2: The first, second, and third quasi-bound states for a resonant tunneling system with barrier height 0.27 eV, barrier width 5 nm, and well width 10 nm. The oscillatory nature and exponential decay are visible in the classically allowed and forbidden regions respectively [2].

The quasi-bound states differ from their true bound analogues because the quasi-bound states decay over time. This decay is due to the chance that the particle will tunnel through the confining barriers. This is seen graphically in the small amplitude oscillations of the wave functions in Figure 1.2, particularly the second and third quasi-bound states.

The tunneling probability takes an interesting and illuminating form in the case where a particle is incident on the barriers from one side. The simplest example of a double barrier system is depicted in Figure 1.3

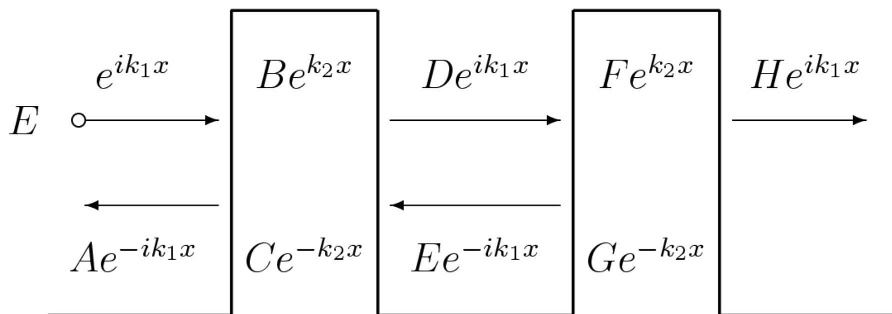


Figure 1.3: A particle incident on double barriers will have its wavefunction split into transmitted and reflected waves. Letters A through H are the amplitudes of the wavefunctions in each region. The direction of the wave is given by the sign of the exponent; a right traveling wave has positive exponent, e^{+ikx}, and a left traveling wave has a negative exponent, e^{-ikx}. This does not apply in the tunneling region however; in this region both the $e^{\kappa x}$ and $e^{-\kappa x}$ terms describe tunneling in both directions [3].

We obtain the transmission coefficient by equating the wave functions and their first derivatives at each boundary and solving the resulting system of linear equations for H. The transmission coefficient is simply $|H|^2$. In general this coefficient takes a very small value. For certain energies however the coefficient can be well approximated by the Breit-Wigner form

$$\frac{(\Gamma/2)^2}{(E-E_R)^2+(\Gamma/2)^2}$$

characterizing a resonance distribution. In the transmission coefficient, E is the energy of the particle and E_R takes discrete values equal to the energies of the quasi-bound states of the well. The exponential suppression due to the barriers is captured in the factor Γ. We see then that for incident energies equal to the resonance energies of the well the tunneling probability is equal to unity. It is interesting to note that in the case of a single barrier, the tunneling probability is always less than one, but a double barrier can actually have a higher

tunneling probability and in some cases theoretically be equal to one. When the energy of the incident wave is not close to the resonance energy the wave function is exponentially suppressed twice, once by each barrier, in the same manner as the suppression due to a single barrier. Therefore the double barrier exhibits very low tunneling probability except for certain energies. This property is used in RTDs and gives them their useful electrical properties.

1.2 Wave Packets

Before discussing RTDs we must revisit the concept of wave packets. A wave packet is comprised of a superposition of waves with distinct energies and wave numbers. We use wave packets to represent particles traveling through space to ensure the wave function is normalizable. The general form for a wave packet is

$$\Psi(x,t) = \int_{-\infty}^{\infty} \varphi(k) A(k) e^{i(kx-\omega(k)t)} dk$$

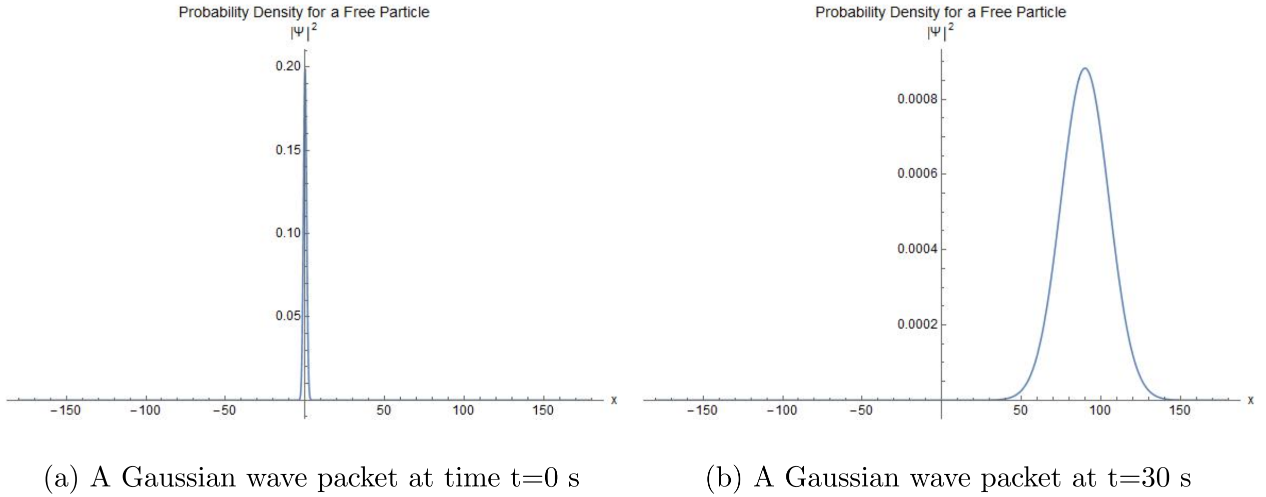
where $\omega(k) = \hbar k^2/2m = E(k)/\hbar$ is the frequency for a given wave number, x is a spatial coordinate and $\varphi(k)$ is some distribution peaked at some value k_0 and that goes to 0 as $k \to \pm\infty$. For a particle we may set $A(k) = 1 \ \forall \ k$ without loss of generality.

We most commonly model particles using a Gaussian wave packet. This wave packet can be described in many forms. It is particularly illuminating to describe such a wave packet as the Fourier transform of its wave number distribution:

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(k) e^{ikx-i\omega(k)t} dk.$$

where wave number distribution is Gaussian and centered around a specific wave number k_0:

$$\varphi(k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(k-k_0)^2/2\sigma_k^2}.$$



(a) A Gaussian wave packet at time t=0 s

(b) A Gaussian wave packet at t=30 s

Figure 1.4: The motion of a wave packet for a free particle with average wavenumber $k_0 = 3$ m^{-1}, wave number uncertainty $\sigma_k = 0.5\ m^{-1}$. In this simulation the particle's mass and $\hbar$ are set equal to one.

A Gaussian distribution is normalized and symmetric around its mean value, in this case k_0, for all t.

Composing the wave function, $\Psi(x, t)$, in this way ensures that the spatial probability distribution, $\Psi^*\Psi = |\Psi|^2$, is Gaussian as well. The wave packet travels through space with velocity $v_g = \frac{\partial \omega}{\partial k} = \hbar k_0/m$. It is then centered at $x_0 = \hbar k_0 t/m$ assuming it is centered at $x = 0$ initially. The motion of a wave packet for a free particle is plotted in Figures 1.4a and 1.4b. We model the tunneling electrons in RTDs with Gaussian wave packets.

1.3 Resonant Tunneling Diodes

RTDs are useful circuit components in optoelectronics and other applications. Charged particles tunnel through the barriers in the diode. The flow of charged particles is called the tunneling current. It is simply a macroscopic manifestation of the tunneling probability;

the higher or lower the tunneling probability, the higher or lower the tunneling current. Quantum tunneling is a very fast process, though exactly how fast is still debated. Some studies claim that quantum tunneling can be instantaneous [4] while others report that tunneling through barriers a few Bohr radii ($1a_0 = 5.291772 \times 10^{-11}$ m) thick takes on the order of 10^{-16} seconds [5]. The speed of quantum tunneling enables RTDs to operate at very high frequencies. Brown et al [6] observed RTDs operating at room temperature at 420 GHz using 1.1 nm thick AlAs barriers and a 4.5 nm GaAs quantum well and at 712 GHz in a different experiment [7] using InAs/AlSb double barrier diodes. The latter experiment produced a power density of 90 W cm^{-2} at 360 Ghz, approximately 50 times greater than the power density generated by the GaAs/AlAs diodes at approximately the same frequency. Recently RTDs have been observed to operate at frequencies in the THz range [8, 9]. The potential functions for typical RTDs resemble those in Figure 1.5.

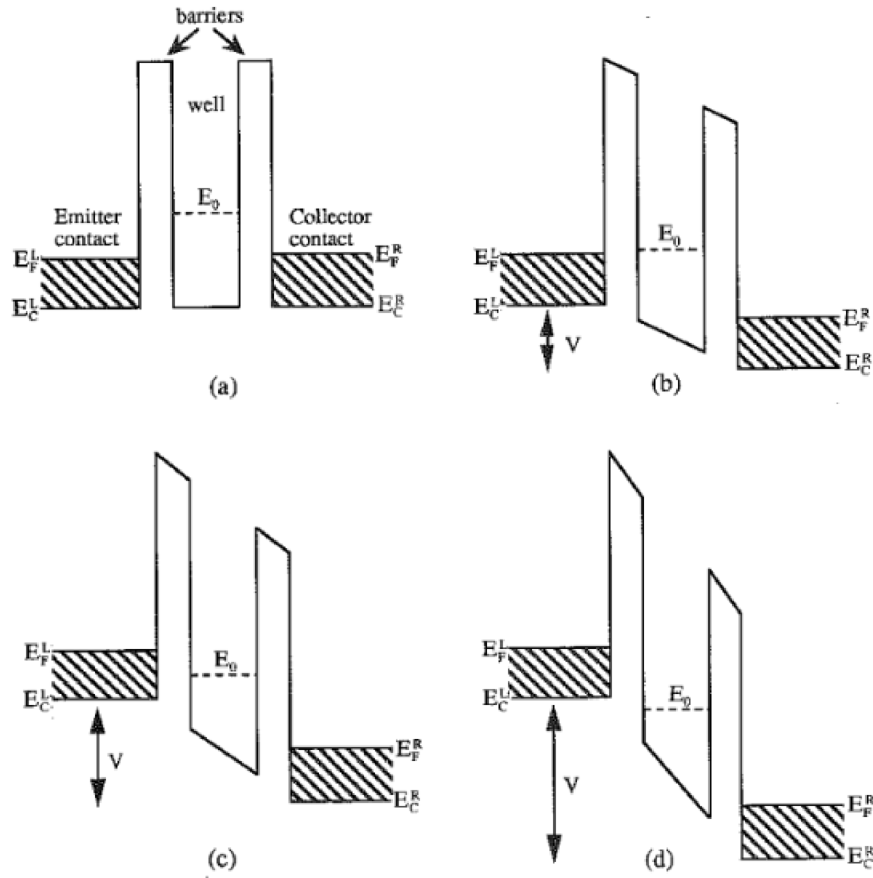


Figure 1.5: A potential function for an RTD at different biases: (a) zero bias, (b) threshold bias, (c) resonance, and (d) off-resonance. The barriers correspond to insulating material. The areas outside the barriers correspond to metals. The shaded areas outside of the barriers represent the energies available to electrons in the conduction bands in the metals. E_F^L and E_F^R are the fermi energies in the emitter and collecter respectively. E_C^L and E_C^R are the energies of the conduction band edges. The resonance energy for the well is labeled by E_0. The resonance energy for this RTD is contained in the range of energies available to electrons incident on the barriers from the left, the so called "emitter region," so this RTD would exhibit a high tunneling current.

Additionally resonant tunneling diodes can exhibit negative differential resistance or con-

ductance

$$R_{\text{diff}} = \frac{dV}{dI} \qquad G_{\text{diff}} = \frac{1}{R_{\text{diff}}}.$$

This can be seen by examining the transmission coefficient

$$T = \frac{(\Gamma/2)^2}{(E - E_R)^2 + (\Gamma/2)^2}.$$

If the tunneling particles have energy close to E_R the tunneling current will be maximized. If the voltage applied to an RTD is increased, the energies of the tunneling particles are increased and the $(E - E_R)^2$ term will increase, causing an overall decline in the tunneling current $\frac{dV}{dI} = R_{\text{diff}} < 0$, hence negative differential conductance. This can also be illustrated using Figure 1.5. Increasing the applied voltage to the diode raises the energies of the conduction bands without affecting the resonance energy. If the voltage is increased to a point where the bottom of the conduction band on the incident side is above the resonance energy, the tunneling current will decrease dramatically.

Negative differential conductance (NDC) makes RTDs useful for amplifying electrical signals. A positive resistor consumes power as current flows through it but, NDC can be used to produce power. The power does not appear from nothingness however; a D.C. bias has to be applied to the diode for it to exhibit NDC. The D.C. bias power is converted into A.C. power by the diode [10]. RTDs operate similarly to traditional transistors in this regard but use only two terminals as opposed to three.

1.4 History and Alternative Theories

Resonant tunneling was first observed experimentally by W.J. Tomasch in 1965 [11]. Since its first observation resonant tunneling has become an extensively documented phe-

nomena. Tomasch observed what he referred to as "geometrical resonance" as periodic oscillations in $\frac{dV}{dI}$ as a function of V in thick ($\sim$5 μm) superconducting film diodes. Originally seen only in superconducting film diodes, resonant tunneling was soon observed in non-superconducting metal diodes [12] under the influence of an applied magnetic field. Gadzuk [13] showed that resonance existed in metal-insulator-metal junctions as well. As Tomasch continued to observe the effect in lead and other materials, his results were examined from a theoretical perspective [14, 15].

1.4.1 McMillan Anderson Model

McMillan and Anderson explain Tomasch's observations as being the result of a quasiparticle interference effect on or near the surface of superconducting film. Particles in a solid behave in complicated ways, but can sometimes be modeled as particles with a different mass in free space, thus simplifying calculations. This is the essence of a quasiparticle. Using quasiparticles McMillan and Anderson derive an expression

$$\omega_n = \left(\Delta^2 + \left(\frac{\pi n \hbar v_F}{d} \right)^2 \right)^{1/2}$$

in agreement with Tomasch's experimental results [11, 15]. In this expression ω_n is the energy of the tunneling quasiparticle, Δ is the difference in energy gaps of the materials in the diode, v_F is the fermi velocity, and d is the thickness of the film. The energy gap is simply the difference between valence band and conduction band of a material.

1.4.2 Gadzuk Model

Taking a time-dependant perturbational approach, Gadzuk formulated a result for the tunneling coefficient

$$T = \langle f| V_F |m\rangle + \frac{\langle f| V_F |a\rangle \langle a| V_F |m\rangle}{E - E_F - i\Gamma}$$

where $|f\rangle$, $|m\rangle$, and $|a\rangle$ represent the wave functions of the tunneling particle in the free, metalic, and atomic part of the potential, V_F is the potential of the applied electric field, and E_F is the Fermi energy. This form closely resembles the Breit-Wigner form

$$\frac{\Gamma/2}{(E - E_R) + i\Gamma/2}$$

characterizing a resonance distribution.

Resonant tunneling, its applications, and its limitations remain of interest. This paper examines the phenomenon through an accessible quantum mechanical lens.

2

Quantum Mechanics of Resonant Tunneling

We begin by obtaining the transmission coefficient for a particle incident on a double barrier. In all regions the wave function can be taken to have the form

$$\Psi(x,t) = \int_{-\infty}^{\infty} \varphi(k)\psi_k(x)e^{-iEt/\hbar}$$

with $E = \frac{\hbar^2 k^2}{2m}$ and $\psi_k(x)$ is a solution to the Schrödinger Equation. Far left of the first barrier we have

$$\Psi_1 = \int_{-\infty}^{\infty} \varphi(k)\left(e^{ikx} + B(k)e^{-ikx}\right)e^{-iEt/\hbar}dk.$$

Far to the right we have

$$\Psi_5 = \int_{-\infty}^{\infty} \varphi(k)A(k)e^{ikx}e^{-iEt/\hbar}dk.$$

The two wave functions are connected using the WKB method [16]. In the classically allowed region the wavenumber given by the WKB approximation is $k(x) = \frac{1}{\hbar}\sqrt{2m(E - V(x))}$. In the classically forbidden region $\kappa(x) = \frac{1}{\hbar}\sqrt{2m(V(x) - E)}$ gives the rate of exponential growth or decay of the wavefunction. For a generalized system with barriers from $x = -b$

to $x = -a$ and $x = a$ to $x = b$ with a quantum well from $x = -a$ to $x = a$ and $V = 0$ elsewhere, the general relation is [17]

$$\begin{pmatrix} 1 \\ B \end{pmatrix} =$$

$$\frac{1}{2}\begin{pmatrix} e^{-2i\rho}\left(\left(4\theta^2 - \frac{1}{4\theta^2}\right)\cos L - 2i\sin L\right) & i\left(4\theta^2 - \frac{1}{4\theta^2}\right)\cos L \\ -i\left(4\theta^2 - \frac{1}{4\theta^2}\right)\cos L & e^{2i\rho}\left(\left(4\theta^2 - \frac{1}{4\theta^2}\right)\cos L + 2i\sin L\right) \end{pmatrix}\begin{pmatrix} A \\ 0 \end{pmatrix}$$

with

$$L = \int_{-a}^{a} k(x)dx, \quad \theta = \exp\left(\int_a^b \kappa(x)dx\right), \quad \text{and} \quad \rho = \int_b^a k(x)dx - kb.$$

This gives

$$A = \sqrt{T}e^{-i\phi} = \frac{e^{2i\rho}}{\frac{1}{2}\left(4\theta^2 - \frac{1}{4\theta^2}\right)\cos L - i\sin L}.$$

The transmission coefficient is

$$T = |A|^2 = \frac{4}{\left(4\theta^2 - \frac{1}{4\theta^2}\right)^2\cos^2 L + 4\sin^2 L}$$

. The exponential suppression from the barriers is captured in the factor θ which is later used in the definition of Γ. The behavior of this coefficient is pictured in Figure 2.1.

For most energies the tunneling probability is very low, $\sim \theta^{-4}$. However, when the energy is equal to the energy of a bound state in an infinite well of equal width and depth, $L = (2n+1)\pi/2$. For these energies we have $\cos^2 L = 0$ and $\sin^2 L = 1$, giving $T = 1$. In this vicinity

$$\cos L \approx \mp \left(\frac{\partial L}{\partial E}\right)\Bigg|_{E-E_R}(E - E_R) \qquad \text{and} \qquad \sin L \approx \pm 1$$

This gives

$$\sqrt{T}e^{-i\phi} = \frac{\Gamma/2}{E - E_R + i\Gamma/2}e^{2i\rho}$$

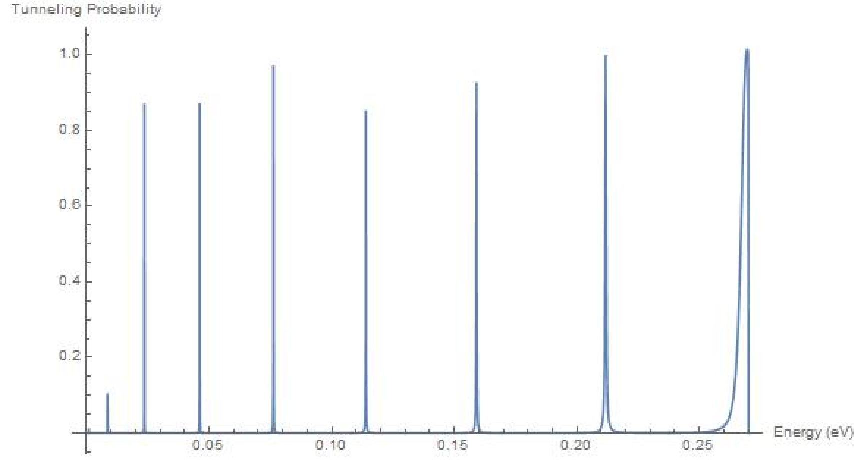


Figure 2.1: Tunneling probability as a function of energy for a double square barrier system with mass $m = m_e = 0.510998$ MeV/c^2, barrier width 5 nm, well width 10 nm, and barrier height 0.27 eV. The tunneling probability peaks sharply at resonance energies. Each peak reaches 1 but this plot is limited by machine precision.

or

$$T = \frac{(\Gamma/2)^2}{(E - E_R)^2 + (\Gamma/2)^2}$$

with $\Gamma = 1/\left(\theta^2 \left.\frac{\partial L}{\partial E}\right|_{E=E_R}\right)$.

Γ is inversely proportional to θ^2, the integral of the wave number in the classically forbidden region. If barrier width or height is increased, Γ decreases exponentially. For most systems Γ will be much smaller than the energy. Γ also defines the Full Width at Half Maximum (FWHM) of the tunneling probability peaks as a function of energy. Similarly we denote the FWHM of the peak as a function of k with Γ_k, as shown in Figure 2.2.

We have total transmission at resonant energies due to destructive interference in the reflected waves; the wave reflected off the first barrier is π radians out of phase and equal in amplitude with the wave reflected off the second barrier. They therefore cancel, and there is no reflected wave. This is possible only in theory of course, as a physical particle must have

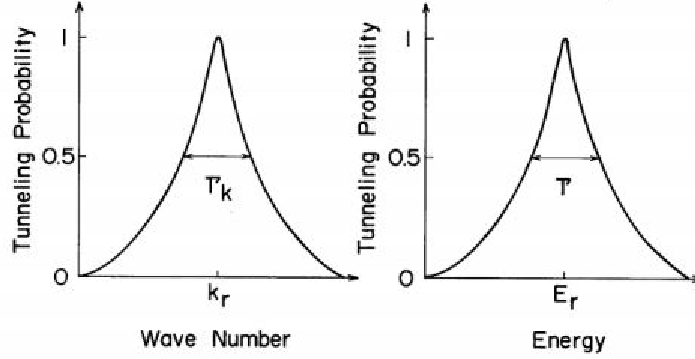


Figure 2.2: Schematic representations of Probability Peaks as functions of wave number and energy [18].

energy uncertainty.

2.1 Incident Wave Packet

Consider a particle described by a Gaussian wave packet incident on a double barrier potential, as in Figure 2.3a. The incident wave function is given by

$$\Psi_I(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(k-k_0)^2/2\sigma_k^2} e^{ikx - i\frac{\hbar k^2}{2m}t} dk.$$

Let us presume that the mean value of the wave number distribution is such that $E_R = \frac{\hbar^2 k_0^2}{2m}$. Let us also presume that the width of the wavenumber distribution, σ_k is small enough that the energy uncertainty of the wave packet $\sigma_E = \frac{\hbar^2 k_0 \sigma_k}{m}$ includes only one resonance.

Upon hitting the first barrier, the particle tunnels through and oscillates in the well. This is visible in the amplitudes of Ψ and $|\Psi|^2$ in Figures 2.3b and 2.3c. The amplitude of $|\Psi|^2$ in the well then decays exponentially with a characteristic time to be discussed in Section 2.2. This decay is visible in Figures 2.3c, 2.3d, and 2.3e.

The transmitted portion of the wave function can be given by

$$\Psi_T(x,t) = \sqrt{T}e^{i\delta_0}\Psi_I(x,t)$$

with $\delta_0 = -k(2d + D)$ where d is the barrier thickness and D is the well width.

We may then use the approximation for $\sqrt{T}e^{i\delta_0}$ in the vicinity of resonances:

$$\Psi_T(x,t) = \frac{\Gamma/2}{E - E_R + i\Gamma/2}e^{i\delta_0}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\varphi(k)e^{ikx - \frac{i}{\hbar}Et}dk.$$

If the wave number uncertainty σ_k is greater than the FWHM, Γ_k, of the tunneling probability peak, the incoming wave packet contains nonresonant components. As a consequence a significant portion of Ψ_I will be reflected. The transmitted wave packet tunnels out from the resonant state with width on the order of Γ_k. Therefore, the width of Ψ_T in x-space, $\sigma_x \cong \frac{1}{\Gamma_k}$ will be large in comparison to that of the incoming packet [18].

2.2 Dwell Time

Far to the right of the barrier the wave function can be expressed as

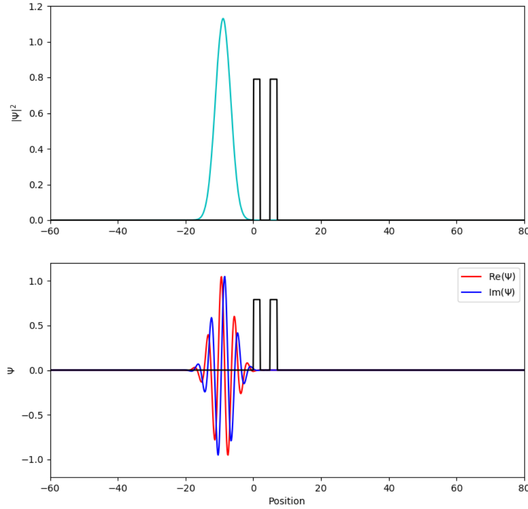
$$\Psi(x,t) = \int_{-\infty}^{\infty}\varphi(k)e^{i(kx+\delta(k)t-\frac{i}{\hbar}Et)}dk$$

where $\varphi(k)$ is peaked at k_0. We may then use the stationary phase approximation and evaluate at k_0

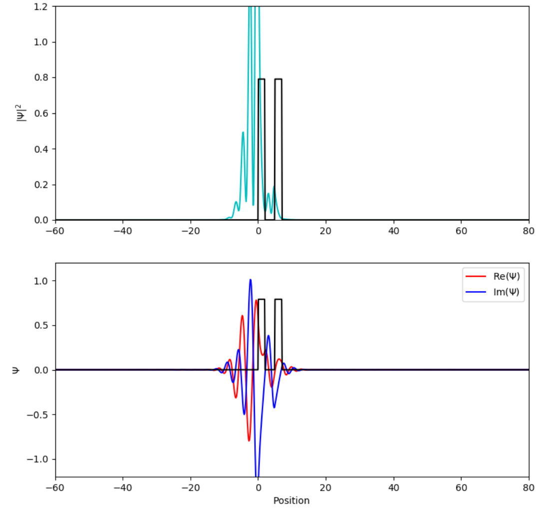
$$\frac{d}{dk}\left(kx + \delta(k)t - \frac{\hbar k^2}{2m}t\right)\Bigg|_{k=k_0} = 0.$$

This gives the position for the center of the wave packet

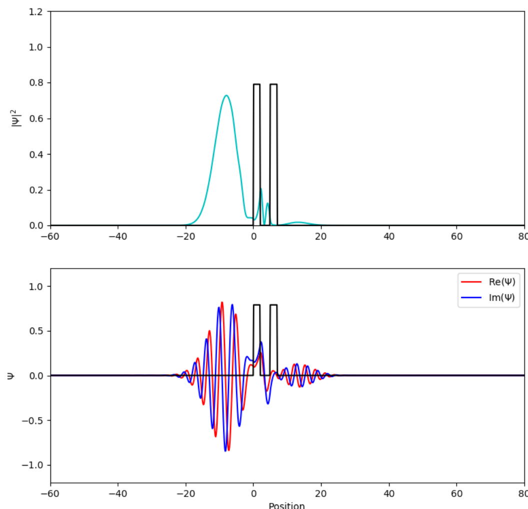
$$x = \frac{\hbar k_0}{m}\left(t - \frac{m}{\hbar k_0}\left.\delta'(k)\right|_{k=k_0} t\right).$$



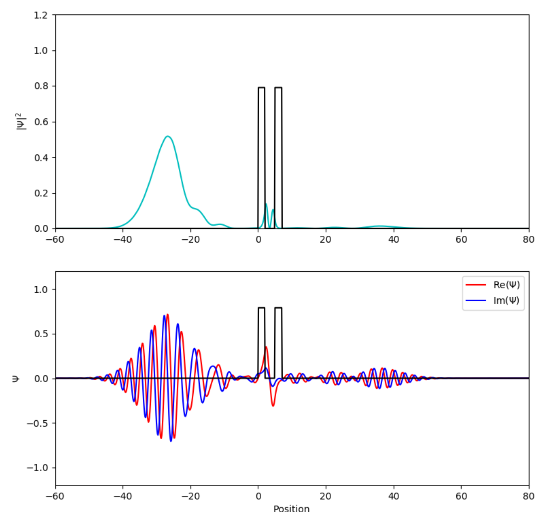
(a) A Gaussian wave packet incident from the left on a double barrier structure



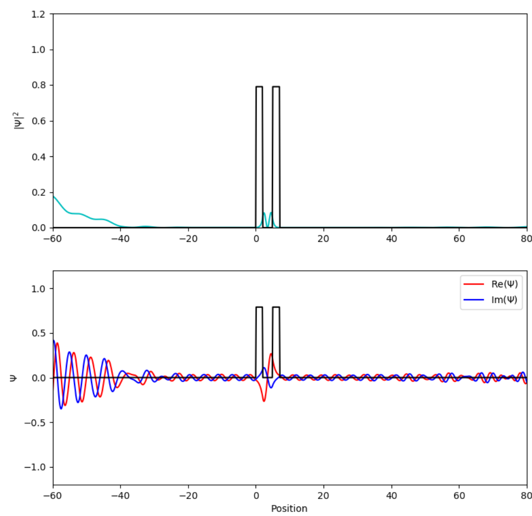
(b) The wave packet hits the barrier. The fluctuations in $|\Psi|^2$ are due to interference from the incident and reflected waves. We see the amplitude of $|\Psi|^2$ increasing as the wave packet reaches the first barrier.



(c) The reflected wave packet is approximately Gaussian in its shape. Right of the barriers we see a wave packet tunneling out. This leading wave front corresponds to the particle tunneling through both barriers without being reflected.



(d) $|\Psi|^2$ decays exponentially in the well as the particle oscillates between the two barriers, tunneling out probabilistically.



(e) Continued evolution of Ψ and $|\Psi|^2$.

Figure 2.3: Evolution of a wave packet incident on a double barrier structure. The evolution equations for this simulation were obtained by applying the finite differences method to the Schrödinger equation for an arbitrary potential.

We see that the wave packet incurs a time shift of $\frac{m}{\hbar k_0}\delta'(k_0)$. The phase shift, δ, is given by the arctan of the transmission coefficient in the complex plane

$$\delta = \delta_0 - \tan^{-1}\left(\frac{\Gamma/2}{E - E_R}\right).$$

Taking the derivative and multiplying by $\hbar$ gives the dwell time, τ_d, or simply τ:

$$\tau_d = \hbar \left.\frac{d\delta(E)}{dE}\right|_{E=E_R} = \frac{m}{\hbar k_0} \left.\frac{d\delta(k)}{dk}\right|_{k=k_0}.$$

which gives

$$\tau_d = \hbar \frac{\Gamma/2}{(E - E_R)^2 + (\Gamma/2)^2}$$

or

$$\tau_d = \frac{2\hbar}{\Gamma}$$

at resonance.

Harada et al. ran a simulation similar to that pictured in Figure 2.3 and calculated the dwell time for various well and barrier dimensions. These results are presented in Table 2.1 and graphed in Figure 2.4.

V (mV)	L (nm)	d (nm)	τ (ps)
400	2	2	0.022
		4	0.37
		6	6.1
	5	2	0.23
		4	8.9
		6	360
	10	2	2.1
		4	110
		6	5900

Table 2.1: Dwell times obtained from wave packet simulation for various well widths (L), barrier thicknesses (d), and a 400 mV potential [18].

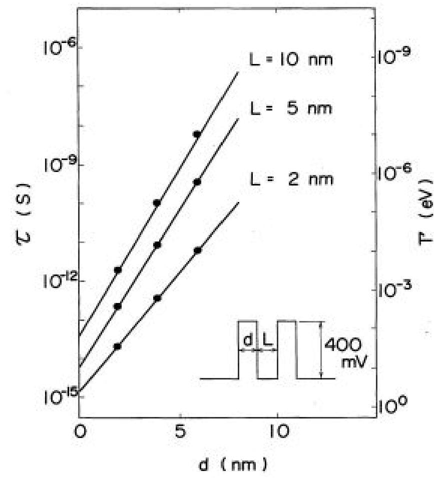


Figure 2.4: Dependence of dwell time on barrier thickness. The solid circles indicate the points obtained from the simulation of the wavepacket. The lines show the dwell time calculated from Γ [18].

3

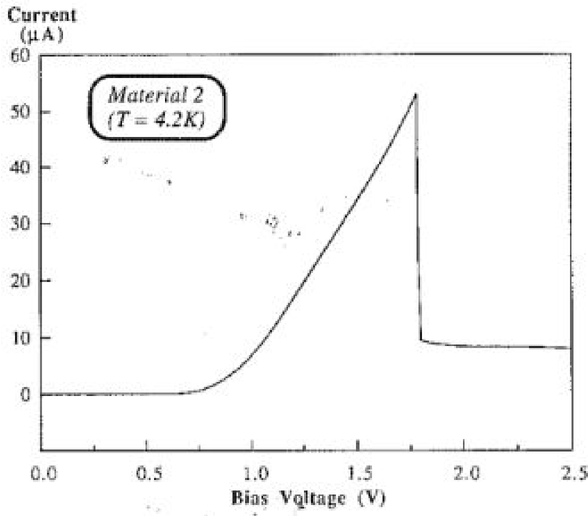
RTD operation

We turn our attention now to RTDs and their electrical properties. Typical RTDs consist of the following: a heavily doped emitter region, an undoped quantum well sandwiched between two undoped barriers, and a heavily doped collector region. The bias applied to the diode, labeled V in Figure 1.5, functions to change the alignment of the emitter and quasi-bound states. In Figure 1.5 tunneling current begins to flow in state (b), is maximized in state (c) and drops off after state (d). We see exactly this behavior in Figure 3.1. The structural parameters of the diodes for this figure are listed in Table 3.1.

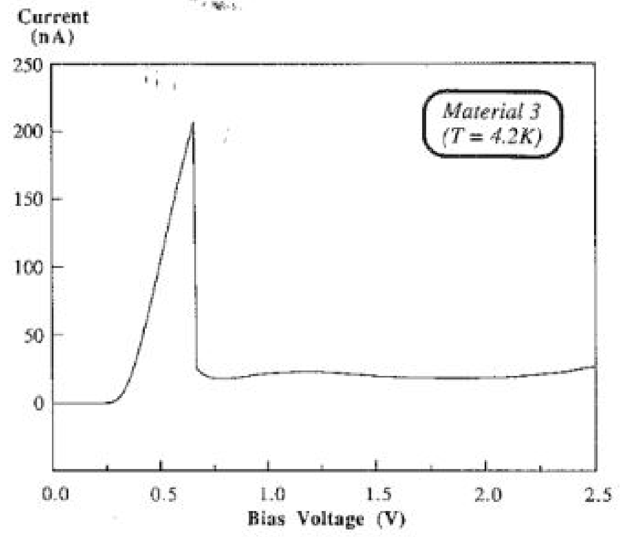
Typical tunnel diodes must be doped heavily enough that the semiconductor degenerates in order to exhibit NDC. When this occurs the impurity atoms create an impurity band in the semiconductor which increases the capacitance. RTDs however offer flexibility in this regard; the structural parameters of the barriers and surrounding doping profile can be selected to give the RTD low capacitance while maintaining NDC. This makes RTDs favorable for generating high frequency signals. Their low capacitance and short dwell times ensure that there is little lag between input and output signals.

	Emitter	
n-GaAs	(Si: $5 \times 10^{18}/\mathrm{cm}^3$)	150 nm
n-GaAs	(Si: $2 \times 10^{17}/\mathrm{cm}^3$)	150 nm
n-GaAs	(Si: $8 \times 10^{16}/\mathrm{cm}^3$)	200 nm
n-GaAs	(Si: $3 \times 10^{16}/\mathrm{cm}^3$)	250 nm
n-GaAs	(Si: $1 \times 10^{16}/\mathrm{cm}^3$	450 nm
	undoped-GaAs	5 nm
	Double barrier structure	
	undoped-GaAs	5 nm
n-GaAs	(Si: $1 \times 10^{16}/\mathrm{cm}^3$	450 nm
n-GaAs	(Si: $3 \times 10^{16}/\mathrm{cm}^3$)	250 nm
n-GaAs	(Si: $8 \times 10^{16}/\mathrm{cm}^3$)	200 nm
n-GaAs	(Si: $2 \times 10^{17}/\mathrm{cm}^3$)	150 nm
n-GaAs	(Si: $5 \times 10^{18}/\mathrm{cm}^3$)	500 nm
	Collector	

Table 3.1: General Emitter and Collector structures. Materials 2 and 3 in Figure 3.1 follow this structure with an AlAs/GaAs/AlAs double-barrier structure with dimensions (4.2 nm/5.9 nm/4.2 nm) and (5.0 nm/7.0 nm/5.0 nm) respectively [19].



(a) I-V curve for Material 2



(b) I-V curve for Material 3

Figure 3.1: I-V curves for a 14 μm diode for Materials 2 and 3 at 4.2 K [20].

To make use of their NDC a DC bias must be applied to the diode. The amplitude of the DC signal can be made large in comparison to that of the AC signal. We may then use a small-signal model[1] for an RTD, pictured in Figure 3.2. In this figure R_s denotes the series resistance due to ohmic contacts[2], the resistivity of the emitter and collector regions, and spreading resistance[3]; C_D denotes the parallel capacitance due to charging and discharging at the wiring and semiconductor depletion regions; $-G$ denotes the conductance of the diode and is made negative to express the NDC; L_{QW} is an inductance element which corresponds to the delay of the diode current with respect to voltage. It takes time to accumulate charges in the quantum well to change the diode current, thus causing the delay. Generally L_{QW} is chosen such that $L_{\mathrm{QW}}G = \tau_d$.

The impedance of the diode can be expressed as a function of frequency in terms of its

[1] A model approximating a circuit element as a smaller circuit of linear components that can be used when the AC signal is small compared to the DC signal.

[2] An electrical junction with a linear I-V curve.

[3] Resistance due to distribution of dopant atoms

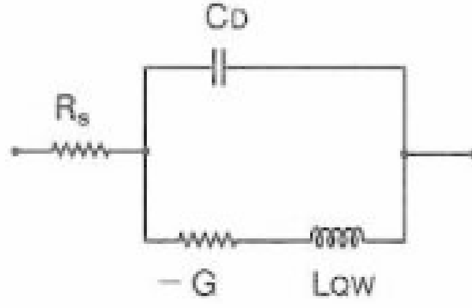


Figure 3.2: Small-signal model for an RTD [21].

real and imaginary parts:

$$Z(f) = R(f) + iX(f).$$

The power consumed by the diode is given by $R(f) \cdot I^2$. If $R(f)$ is negative the diode supplies electrical power to the circuit at frequency f by converting power from the D.C. bias required to operate the RTD in the NDC regime. The output signal is thereby amplified. The maximum oscillation frequency is defined as the frequency where $R(f_{\mathrm{max}}) = 0$ and is given by [22]

$$f_{\mathrm{max}} = \frac{1}{2\pi}\left(\frac{1}{2L_{\mathrm{QW}}^2 C_D}\right)^{\frac{1}{2}}$$
$$\times \left\{ 2L_{\mathrm{QW}} - \frac{C_D}{G^2} + \left[\left(\frac{C_D}{G^2} - 2L_{\mathrm{QW}}\right)^2 - \frac{4L_{\mathrm{QW}}(1 + R_s G)}{R_s G}\right]^{\frac{1}{2}} \right\}^{\frac{1}{2}}.$$

For diodes with negligibly small L_{QW} this expression simplifies to

$$f_{\mathrm{max}} = \frac{1}{2\pi C_D}\left(\frac{-G}{R_s} - G^2\right)^{\frac{1}{2}}.$$

The high frequency performance can be improved by increasing G or decreasing R_s or C_D. To achieve large G the peak current density J_p must be large and the valley current density J_v must be small. The peak current is the local maximal current flow as a function

Material	J_p (kA cm^{-2})	J_p/J_v	$J_p - J_v$ (kA cm^{-2})
InGaAs/AlAs	450	4	345
InAs/AlSb	370	3.5	260

Table 3.2: Peak current density, peak-to-valley current ratio, and available current density for different RTDs [23, 24].

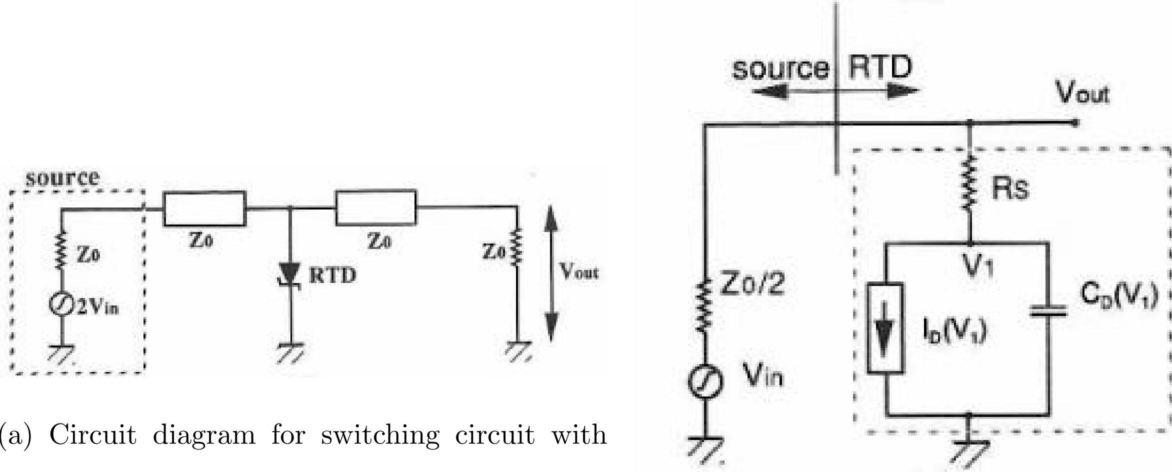
of voltage. The valley current is the local minimal current flow. RTDs with high J_p typically have a peak-to-valley current ratio $\frac{J_p}{J_v}$ of ~ 10. High J_p corresponds to a high transmission probability and therefore a shorter dwell time. Table 3.2 lists current values for two RTDs. The two listed RTDs have an averaged NDC of over 650 kS cm^{-2}. If we assume they have a specific C_D of 100 nF cm^{-2} which corresponds to a depletion region of slightly over 100 nm, the time constant C_D/G is 0.15 ps. This corresponds to an expected oscillation frequency of 1 THz [19].

3.1 Switching

RTD's high operation speeds make them useful for high-frequency switching circuits. The short dwell times of tunneling particles allow high-frequency A.C. signals to be treated as being quasi-static, provided that the period of the A.C. signal is significantly longer than the dwell time. An illustration of an RTD switching circuit and its large-signal[4] equivalent is pictured in Figure 3.3.

The large-signal model differs from the small signal model in two ways: first, the differential conductance of the diode, $-G$ in Figure 3.2, is replaced by the current function $I_D(V_1)$

[4] Approximate model for large AC compared to DC current.



(a) Circuit diagram for switching circuit with RTD switching element

(b) Equivalent circuit. The box given by the dashed line indicates the equivalent circuit for the RTD.

Figure 3.3: RTD switching circuit and its large-signal equivalent [19]

of the voltage across the diode; second, the inductance L_{QW} is not included. The differential conductance must be replaced by the current function in the large-signal model because the differential conductance changes during switching. This current function is assumed to be piecewise-linear.

The circuit equation for Figure 3.3 according to Kirchoff's current law is

$$C_D(V_1)\frac{dV_1}{dt} = \frac{V_{\text{in}} - V_1}{R_L + R_s} - I_D(V_1) \tag{3.1}$$

where $R_L = \frac{Z_0}{2}$. By rewriting $\frac{V_{\text{in}} - V_1}{R_L + R_s}$ as $I_L(V_1)$ equation 3.1 can be simplified as

$$C_D(V_1)\frac{dV_1}{dt} = I_L(V_1) - I_D(V_1). \tag{3.2}$$

This has steady state solutions

$$I_L(V_1) = I_D(V_1).$$

Two solutions to this equation are plotted in Figure 3.4. There are two solutions for V_{in_1} and one solution for V_{in_2}. Suppose the diode is initially in state $V_{\text{in}} = V_{\text{in}_1}$ and $V_1 = V_{\text{p}}$. If the input voltage is then increased to V_{in_2} the diode will switch from V_{p} to V_{f}.

The transient behavior of the diode is governed by Equation 3.2. We may compute the 10-90% rise time[5] for the switch by rewriting Equation 3.2 as

$$dt = dV_1 \frac{C_D(V_1)}{I_L(V_1) - I_D(V_1)}$$

and integrating from $V_p + 0.1\Delta V$ to $V_f - 0.1\Delta V$ where $\Delta V = V_f - V_p$. This gives

$$T_{\text{rise}} = \int_{V_p+0.1\Delta V}^{V_f-0.1\Delta V} \frac{C_D(V)}{I_L(V) - I_D(V)} dV.$$

If we assume C_D is independent of V and that the difference between V_{in_1} and V_{in_2} is small, the rise time is given by [25]

$$T_{\text{rise}} = |R_n| C_D \left[\frac{x}{x-1} \ln \left(\frac{10(x+y)}{x(x-y)} \right) + \frac{xy}{x+y} \ln \left(\frac{10y(x+y)}{x(x-y)} \right) \right]. \tag{3.3}$$

In this equation $x \equiv \frac{R_L+R_s}{|R_n|}$ and $y \equiv \frac{R_D}{|R_n|}$, where $\frac{1}{R_n}$ is the slope of I_D in the NDC region (just right of V_p in Figure 3.4) and $\frac{1}{R_D}$ is the slope of I_D at V_f.

From Equation 3.3 we can see that the rise time is dominated by R_n and C_D. There is also an optimal x value for a given y value. As x decreases towards 1, $I_L - I_D$, the current charging the diode capacitance, shrinks. This causes an increase in T_{rise}. If x becomes large due to an increase in load impedance, T_{rise} increases due to an increase in the input voltage swing necessary for switching. We see then that to achieve optimal performance R_s, R_n, and C_D must be reduced, as one might expect. RTD switches have been used in triggering circuits with frequencies up to 110 GHz [26, 27]. Switching times as low as 2.2 ps have been observed [28].

[5]The time for the signal to change from 0.1 to 0.9 of the output step.

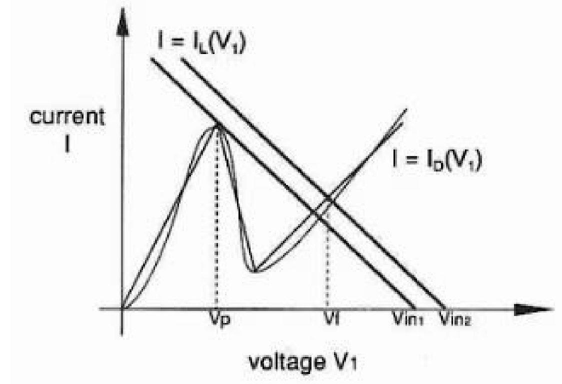


Figure 3.4: $I_D(V_1)$ for an RTD and load lines $I_L(V_{\text{in}_i})$ [19]. The straight lines represent the piecewise linear approximation I_D.

3.2 RTD Static Memory Operation

The last application for discussion is RTDs as static memory elements. RTDs can be used to construct inverter circuits[6] that exhibit hysteresis. These circuits can be used as static memory elements and use fewer components than typical static memory circuits [19]. Their operation is similar to the switching described in Section 3.1.

A simple memory circuit using an RTD is pictured in Figure 3.5a. Figure 3.5b shows the I-V model for the diode and three different load lines. We see that the circuit can have between 1 and 3 steady-state solutions. The stability of each solution is governed by

$$V_{\text{out}} = C_1 e^{-C_D^{-1}\left(R_L^{-1}+R_i^{-1}\right)t} + \frac{R_L R_i}{R_L + R_i}\left(\frac{V_{\text{in}}}{R_L} - I_i\right)$$

where C_1 is some constant not to be confused with capacitance, R_i is the differential resistance at intersection $i \in (1,\ 2,\ 3)$, and I_i is the current axis intersect of each linear piece of I_D. For $R_L^{-1} + R_i^{-1} > 0$ the sign of the exponential is negative and the exponential tends to zero for large t, therefore the solution is stable. Solutions 1 and 3 in Figure 3.5b are stable;

[6] A circuit that converts direct current to alternating current

Solution 2 is not. Switching between these two stable points allows the circuit to function as a memory device.

The operation of the device is simple. If we imagine V_{in} increasing from zero, the diode voltage increases along the first part of I_D until it reaches V_p. After reaching V_p it abruptly switches to state 3 and continues to increase with V_{in}. The process is analogous in reverse: the diode voltage decreases along the third part of I_D, through state 3, reaches V_v and abruptly switches to state 1 as V_{in} continues to decrease. This behavior is plotted in Figure 3.5c. The width in terms of V_{in} between the two switching points determines the noise margin[7]. The state of the device is read by simply sensing the output voltage.

Due to their simplicity, these devices are well suited to use in conjunction with other active components[8] such as FETs[9] to create ICs[10]. An example RTD memory cell is pictured next to a conventional memory cell in Figures 3.6a and 3.6b respectively. We see that the RTD cell requires far fewer components than a typical cell: a resistor, an RTD, and an FET compared to six transistors and two resistors. Opening gates Q_1 and Q_2 connects the output node of the RTD memory element to the bit line. The element is read by sensing the output voltage and written to by changing the node voltage through the bit line.

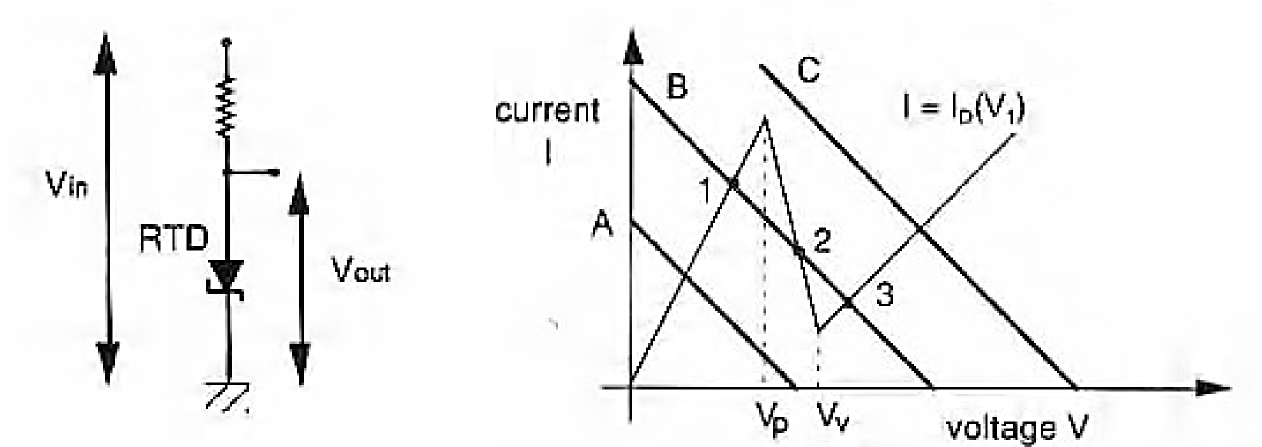
For an RTD to function as a memory element we must have two stable points for memory operation, as in Figure 3.5b. The valley current for the diode must be sufficiently small or the second stable state will be lost. To maintain a low valley current the load resistance R_L and therefore the input voltage must be large. This causes significant power loss during operation. The valley current is due to scattering in the quantum well [19]. Electrons with

[7]The amount by which the signal exceeds the minimum amount for proper operation.

[8]Components that require power, usually from a DC source, for their operation.

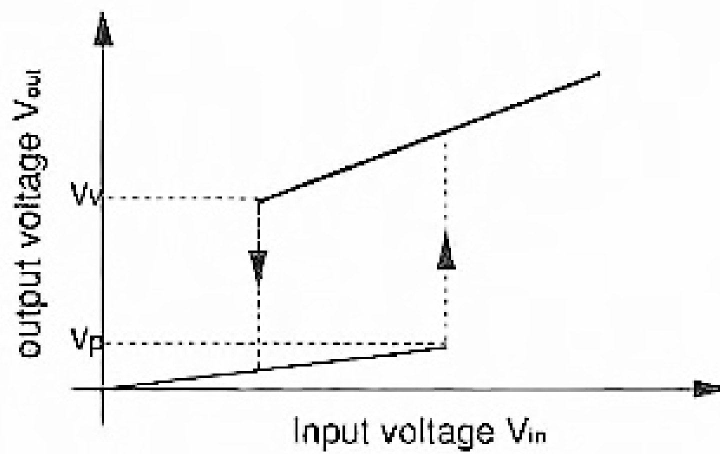
[9]Field Effect Transistors: transistors which use electric fields to control the output.

[10]Integrated circuits, commonly known as microchips: a set of circuits on a small flat piece (a "chip") of semiconductor material, usually silicon.



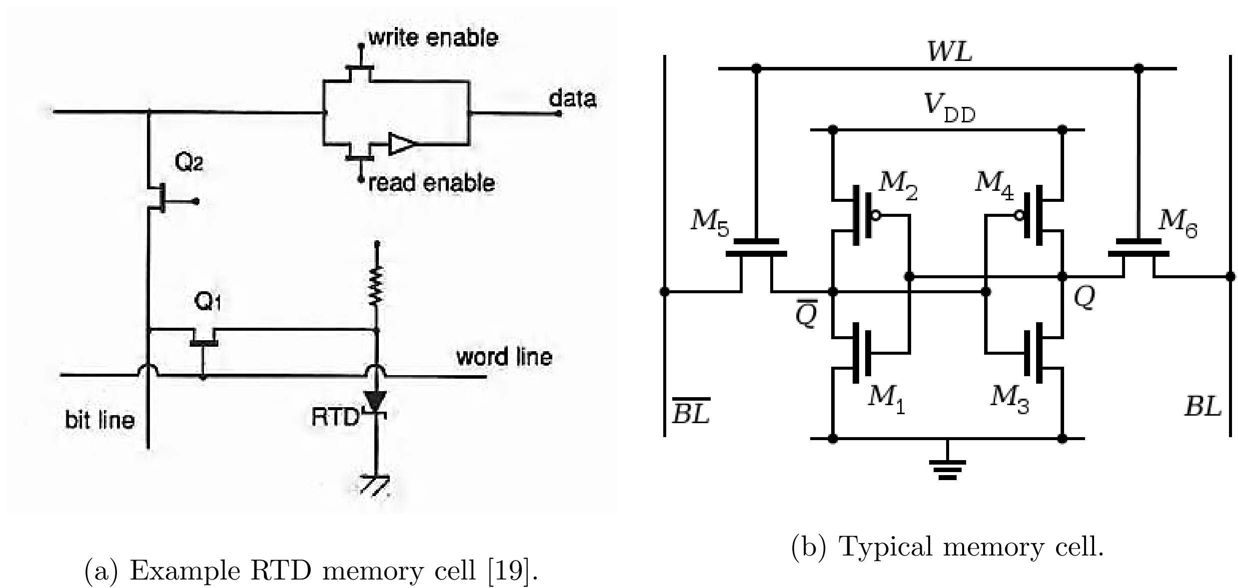
(a) A simple RTD memory circuit built from an RTD and a resistor.

(b) Piecewise-linear model for RTD and three load lines. V_p and V_v denote the voltages for peak and valley currents respectively.



(c) V_{out} vs V_{in} plot for the circuit in Figure 3.5a

Figure 3.5: RTD memory circuit, I_D model, and output-input plot [19].



(a) Example RTD memory cell [19].

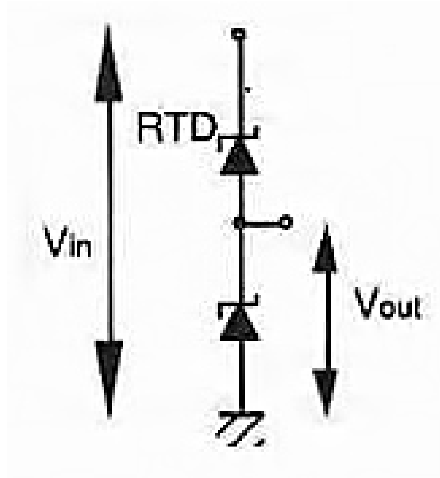
(b) Typical memory cell.

Figure 3.6: Circuit diagrams for RTD memory cell and 6T-SRAM cell.

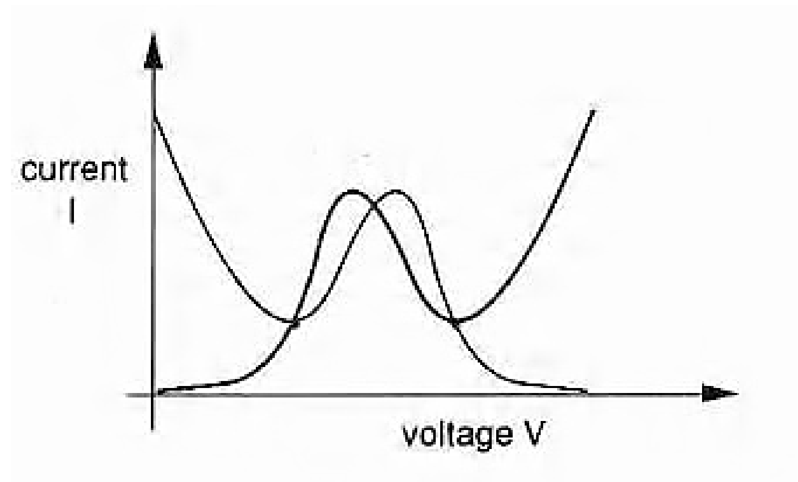
initial energy greater than the resonance energy may scatter off of impurities in the diode. Some of their longitudinal momentum may then be converted into transverse momentum, allowing them to inhabit resonant states in the well. A more efficient way to control the valley current has yet to be found and RTD memory cells still consume too much power to be practical.

Replacing the load resistance with a nonlinear circuit element offers a different way of reducing power consumption in RTD memory cells. A number of devices have been used in this pursuit including D-FETs and Bipolar Junction Transistors. Both D-FETs and BJTs reduce power consumption in RTDs, but increase the space required for each cell. A more promising approach is using a second RTD as the load resistor in a cascade [29, 30] as in Figure 3.7a. The operation of this circuit is essentially the same as that for the linear load resistance, despite the more complicated load line. The two stable operation points occur in the valley current for the two RTDs, as seen in Figures 3.7b. This reduces power consumption while maintaining a compact cell. Cascade RTD memory cells have been manufactured to

be as small as $10 \times 13\,\mu$m [30], only slightly larger than the smallest reported 6-transistor memory cell (0.1 μm^2) [31].



(a) Example double RTD memory cell. [19].



(b) I-V curve for a two RTD memory cell

Figure 3.7: Cascade RTD memory cell and its I-V Curve

4

Conclusion

This paper aims to clarify the connection between the quantum mechanics of resonant tunneling and the properties exhibited by RTDs. The transmission coefficient for a tunneling particle was derived and images from a simulation were presented. The lifetime of the quasi-stable state was also derived. Some properties of the transmitted wave packet, such as its width in x and k space were presented.

Additionally this paper serves to outline a few fundamental electrical applications of RTDs. RTD applications are still a field of active research and some topics have been excluded from this paper including multiple-state memory and logic [32, 33], optical control [34, 35] and near THz wireless transmission frequencies [36]. As RTD technology continues to improve they may become more ubiquitous circuit components used for a wide range of applications.

Bibliography

[1] Felix Kling - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=11921

[2] https://www.comsol.com/blogs/learning-quantum-mechanics-concepts-with-double-barrier-structures/

[3] https://en.wikipedia.org/wiki/Resonant-tunneling˙diode Public Domain

[4] Torlina, Lisa et al (2015), Interpreting attoclock measurements of tunnelling times, Nature Physics volume 11, pages 503–508

[5] Camus, Nicholas et al (2017), Expiremental Evidence for Quantum Tunneling Time, Physical Review Letters, Volume 119, Issue 2,

[6] E.R. Brown et al (1989), Oscillations up to 420 GHz in GaAs/AlAs resonant tunneling diodes, Applied Physics Letters, vol. 55, p. 1777-1779.

[7] E.R. Brown et al (1989), Oscillations up to 712 GHz in InAs/AlSb resonant tunneling diodes, Applied Physics Letters, vol 58, p. 2291-2293.

[8] Feiginov et al (2011), Resonant-tunnelling-diode oscillators operating at frequencies above 1.1 THz, Applied Physics Letters, Volume 99, Issue 23, id. 233506

[9] Kanaya et al (2012), Fundamental Oscillation up to 1.31 THz in Resonant Tunneling Diodes with Thin Well and Barriers, Applied Physics Express, Volume 5, Issue 12, article id. 124101

[10] Lesurf, Jim (2006). "Negative Resistance Oscillators". The Scots Guide to Electronics. School of Physics and Astronomy, Univ. of St. Andrews. https://www.st-andrews.ac.uk/ www˙pa/Scots˙Guide/RadCom/part5/page2.html

[11] W.J. Tomasch (1965), Geometrical Resonance in the Tunneling Characteristics of Superconducting Pb, Physical Review Letters, 15, 672-675.

[12] E. E. H. Shin (1968), Geometrical resonance in E H tunneling in normal metal diode, Physics Letters A, Volume 26, Issue 7, p. 303-304.

[13] J. W. Gadzuk (1970) Resonance Tunneling Through Impurity States in Metal-Insulator-Metal Junctions, Journal of Applied Physics, Volume 41, Issue 1, p.286-291

[14] J.W. Gadzuk (1970), Resonance-Tunneling Spectroscopy of Atoms Adsorbed on Metal Surfaces: Theory, Physical Review B, vol. 1, Issue 5, pp. 2110-2129

[15] McMillan, W.L.; Anderson P.W. (1966), Theory of Geometrical Resonances in the Tunneling Characteristics of Thick Films of Superconductors, Physical Review Letters, 16, 85-87

[16] Bohm, David (1951), Quantum Theory, Prentice Hall

[17] Merzbacher, Eugene (1961), Quantum Mechanics, John Wiley and Sons

[18] Harada, Naoki; Kuroda, Shigeru (1986), Lifetime of resonant state in a resonant tunneling system, Japanese Journal of Applied Physics, Part 2, vol. 25, p. L871-L873.

[19] Mizuto, Hiroshi; Tanoue, Tomonori (1995), The Physics and Applications of Resonant Tunneling Diodes, Cambridge University Press

[20] C.J. Goodings (1993), Variable-area resonant tunneling diodes using implanted gates, PhD thesis, Cambridge University.

[21] E.R. Brown et al (1989), Effect of quasibound-state lifetime on the oscillation power of resonant tunneling diodes, Applied Physics Letters, vol. 54, p. 934-936.

[22] T.C.L.G Sollner, E.R. Brown, C.D. Parker and W.D. Goodhue, High frequency applications of resonant-tunneling diodes, *Electronic Properties of Multilayers and Low-dimensional Semiconductor Structures*, edited by J.M Chamberlain *et al.* (Plenum Press, New York, 1990), pp. 283-96.

[23] Broekaert, Tom P. E.; Fonstad, Clifton G (1990), $In_{0.53}Ga_{0.47}As/AlAs$ resonant tunneling diodes with peak current densities in excess of 450 kA/cm^2, Journal of Applied Physics, vol. 68, p. 4310-4312.

[24] Soderstrom, J. R.; Brown, E. R.; Parker, C. D.; Mahoney, L. J.; Yao, J. Y. (1991), Growth and characterization of high current density, high-speed InAs/AlSb resonant tunneling diodes, Applied Physics Letters, vol. 58, p. 275-277.

[25] Diamond, S. K.; Ozbay, E.; Rodwell, M. J. W.; Bloom, D. M.; Pao, Y. C. (1989), Resonant tunneling diodes for switching applications, Applied Physics Letters, vol. 54, p. 153-155.

[26] Özbay, E.; Bloom, David M. (1991), 110-GHz monolithic resonant-tunneling-diode trigger circuit, IEEE Electron Device Letters, vol. 12, p. 480-482.

[27] Özbay, E.; Bloom, D.M.; Diamond, S.K. (1990), Pulse forming and triggering using resonant tunneling diode structures, *Electronics Letters*, Vol. 26.

[28] Shimizu, N.; Nagatsuma, T.; Shinagawa, M.; Waho, T. (1995), Picosecond-switching time of $In_{0.53}/Ga_{0.47}$/As/AlAs resonant-tunneling diodes measured by electro-optic sampling technique, IEEE Electron Device Letters, vol. 16, issue 6, pp. 262-264

[29] Mori, T.; Mutoh, S.; Tamura, H.; Yokoyama, N. (1993), An SRAM cell using a double emitter RHET for gigabit-plus memory applications, Japanese Journal of Applied Physics, Vol. 33, pp. 790-793.

[30] Watanabe, Y.; Nakasha, Y.; Imanishi, K.; Takikawa, M. (1992), Monolithic integration of InGaAs/InAlAs resonant tunneling diodes and HEMT for single transistor cell SRAM application, *Tech Dig.*, IEDM 92, pp. 475-478.

[31] Haran, B.S. *et al.* (2008), 22 nm technology compatible fully functional 0.1 μm^2 6T-SRAM cell, IEEE International Electron Devices Meeting, San Francisco, CA, pp. 1-4.

[32] Soderstrom, Jan; Andersson, Thorwald G. (1988), A multiple-state memory cell based on the resonant tunneling diode, IEEE Electron Device Letters, vol. 9, p. 200-202.

[33] Waho, Takao; Chen, Kevin J.; Yamamoto, Masafumi (1997), A Novel Functional Logic Circuit Using Resonant-Tunneling Devices for Multiple-Valued Logic Applications, Japanese Journal of Applied Physics, vol. 36, issue Part 1, No. 3B, pp. 1818-1821

[34] Lee, Kiwon; Park, Jaehong; Lee, Jooseok; Yang, Kyounghoon (2015), Optically controlled low-power on-off mode resonant tunneling oscillator with a heterojunction phototransistor switch, Optics Letters, vol. 40, issue 6, p. 1022.

[35] Romeira, Bruno; Figueiredo, José M. L.; Ironside, Charles N.; Kelly, Anthony E.; Slight, Thomas J. (2010), Optical Control of a Resonant Tunneling Diode Microwave-Photonic Oscillator, IEEE Photonics Technology Letters, vol. 22, issue 21, pp. 1610-1612.

[36] Oshima, Naoto; Hashimoto, Kazuhide; Suzuki, Safumi; Asada, Masahiro (2017), Terahertz Wireless Data Transmission With Frequency and Polarization Division Multiplexing Using Resonant-Tunneling-Diode Oscillators, IEEE Transactions on Terahertz Science and Technology, vol. 7, issue 5, pp. 593-598