DEVICES BASED ON QUANTUM TUNNELING

SHRI VIDHATRI M M
International Institute Of Information and Technology, Hyderabad

November 22, 2020

Abstract

This paper presents review of the principles of operation, properties and implications of tunnel diode, MIM diode, and some other devices based on tunneling. These electronic devices work based on quantum tunneling that gives many advantages over traditional electronics. The paper touches upon the basics of Quantum tunneling and its properties and tries to provide an overall and succinct picture of how two different disciplines of science i.e. Quantum mechanics and electronics come together to bring out devices such as tunnel diodes and MIM diodes having the best of both fields.

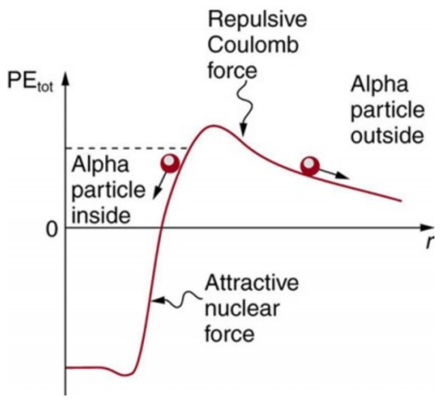
1 Introduction

MOS and CMOS are some of the classic semiconductor devices used in Integrated circuits. The last few decades has seen extensive geometry shrink of these devices combined with an increase in the circuit density. But this shrink was limited to 20-30 nm.

Improved photo-lithography technology could take this shrinkage further, only to be stonewalled by quantum mechanical leakage currents. A milestone in this field was achieved when devices were built that were based on the electron wave behaviour that involved a much smaller geometry. Quantum tunneling concept was integrated into diodes and transistor to develop tunnel diodes, MIM diodes, resonant tunneling diodes, Quantum dot transistors and many more.

2 Literature review

The 20th century has seen many remarkable discoveries that seem to have an impact in various fields. One of those discoveries is quantum tunneling. It was predicted and described theoretically by Oskar Klein long before it was proved. When we go deeper, Quantum tunneling was identified through the study of the concept of radioactivity [1].



Friedrich Hund [a], Leonid Mandelstam and Mikhail Leontovich discovered the phenomenon of tunneling independently, Friedrich in his research of deep well potential and the latter two, during their study of Schrodinger's equations. The idea was then used to understand alpha decay process and radioactivity deeply. It was further extended in various researches by scientists in many parts of the globe. Max Born was one of the first scientists to realise that tunneling could be used to understand many other phenomenon besides and beyond nuclear processes. A few years later, in 1973, Leo Esaki along with several others was awarded the noble prize[b] for the discovery of Tunneling effects in semiconductor materials which will be discussed further in this paper.

3 Theory

In quantum tunneling, an electron has a finite probability of penetrating beyond the classical reflection point by passing through a potential barrier(tunneling through the barrier) without surmounting it.

The discovery of quantum mechanical tunneling effects in electronic devices has created many challenging problems in the theory of semiconductors. Despite the problems, they had advantages too, ranging from amplification to generation of signals of high frequency and fast acting storage as well as logic devices[1].

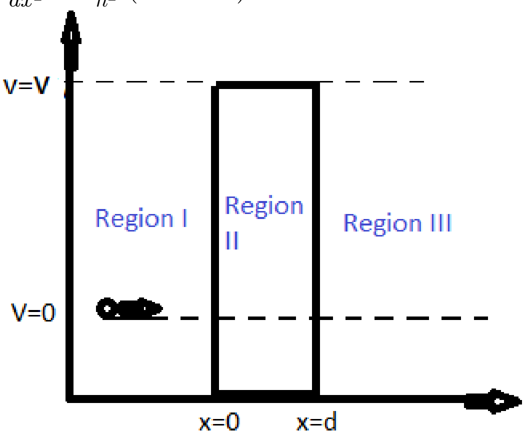
A two-terminal negative resistance device was a rarity up until the point of this discovery. Thus it could be said that the so called problem of quantum mechanical tunneling of electrons created a new field of engineering[11].

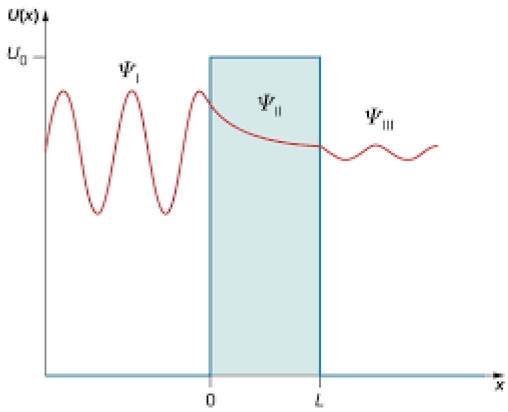
We have established that in tunneling effect, an electron crosses a barrier with potential energy more than that of the electron.

We know that,

$\frac{d^2\Psi}{dx^2} + \frac{8\pi m}{h^2}(E - V)\,\Psi = 0$

$\frac{d^2\Psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\,\Psi = 0$





$$V(x) = \begin{cases} 0 & x < 0 \\ V & 0 \le x \le d \\ 0 & x > d \end{cases}$$

Outside the barrier, V=0

In region I & III,

$\frac{d^2\Psi}{dx^2} + \frac{2m}{h^2}(E)\,\Psi = 0$

Inside the barrier , region II,

$\frac{d^2\Psi}{dx^2} - \frac{2m}{h^2}(V - E)\,\Psi = 0$

We can use the WKB approximation to get the solutions of Ψ in these regions[13].

In region I,

$\Psi_I = A_1 cos\alpha x + A_2 sin\alpha x$

In region III,

$\Psi_{III} = B_1 cos\alpha x + B_2 sin\alpha x$

We can say that

$(\ \alpha^2 + \frac{2m}{\hbar}E)\ (\ A_1 cos\alpha x + A_2 sin\alpha x)\ \ = 0$

so, $\ \alpha^2 = \frac{2mE}{\hbar}$

In region II,

$\Psi_{II} = C_2 e^{\beta x} + C_1 e^{-\beta x}$

since $\int \Psi\Psi^* dv < \infty$

thus $C_2 = 0$ (so that the barrier width $< \infty$

$\Psi = C_1 e^{-\beta x}$

$\beta^2 = \frac{2m}{\hbar^2}(\ V - E)$

Using boundary conditions, since Ψ is continuous,

$A_1 = C_1$

Since $\frac{d\Psi}{dx}$ is continuous,

$\alpha A_2 = -\beta C_1$

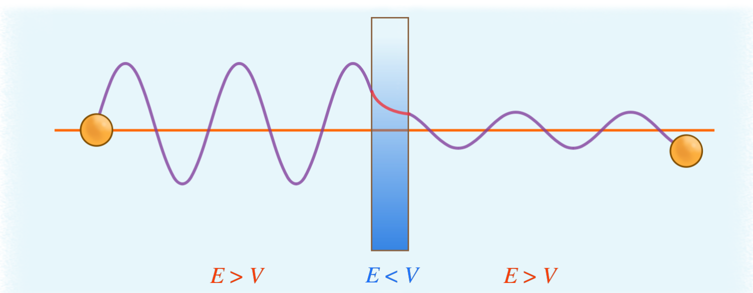
at $\ x = d,\ \ \Psi_{II} = \Psi_{III}$

thus $\ \ C_1 e^{-\beta d} = B_1 cos\alpha d + B_2 sin\alpha d$

or $-\beta C_1 e^{-\beta d} = \alpha(\ -B_1 sin\alpha d + B_2 cos\alpha d)$

$$\begin{aligned}\Psi\ &=\ C_1(\cos\alpha x - \tfrac{\beta}{\alpha} sin\alpha x)\\ &= C_1\ e^{-\beta x}\\ &= C_1\ e^{-\beta d}\ \{(\ cos\alpha d + \tfrac{\beta}{\alpha} sin\alpha d)\ cos\alpha x + (\sin\alpha d - \tfrac{\beta}{\alpha} cos\alpha d)\sin\alpha x\}\end{aligned}$$

An obvious observation from the above equation is that, $|\Psi|_{III} \neq 0$



The equation is simplified further

$\Psi_{III} = C_1\, e^{-\beta d}\, \{cos\alpha(\, x - d) + sin\alpha(\, x - d)\, \}$

Or in other words, $|\Psi|_{III} \infty C_1\, e^{-\beta d}\, |\Psi|_I$

Tunneling probability is defined as the ratio between square of absolute value of transmitted electron wave amplitude and square of absolute value of incident electron wave amplitude[1].

In other words, it is the the ratio between square of $|\Psi_{III}(\, x\,)|$ and square of $|\Psi_I(\, x\,)|$

Let the tunneling probability be P_t

$P_t = \frac{|\Psi_{III}(\, x\,)|^2}{|\Psi_I(\, x\,)|^2}$

Or in other words,

$P_t \infty\, e^{-a\beta d}$

We notice here that the tunneling probability decreases exponentially with the barrier width and barrier energy.

The tunneling probability can be approximated after a series of steps into a simple equation for a high energy thick barrier that barely can facilitate tunneling.

$P_t \approx 16\, \frac{E}{V}(\, 1 - \frac{E}{V})\, e^{-2\beta d}$ [13].

4 Tunnel diode

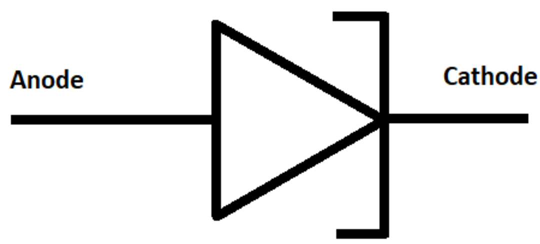
Tunnel diode is a p-n junction diode with heavy doping on both sides and exhibits negative resistance. This level of doping reveals the effect of electron tunneling in semiconductors and has many differences as opposed to regular p-n junction diodes.

Tunnel diode is also referred to as Esaki diode, named after the scientist Leo Esaki who first studied this type of device.

Tunnel diode has a broken band gap due to heavy doping. The doping levels in an Esaki diode are very high, there is 1 dopant for every 10^3 atom compared to 1 dopant in 10^8 for a classic p-n diode. or in other words, the doping is 1 million times higher!![7].

This level of doping causes the two sides of depletion layer to be nearly aligned.

Circuit symbol for tunnel diode:



In classic semiconductor devices, the current is contributed by the minority charge carriers i.e. holes in n-type semiconductor and electrons in p-type. In tunnel diode, current is in fact contributed by majority charge carriers.[9]

This difference in tunnel diodes provides three useful advantages:

- The current carriers are not limited by the low value of diffusion velocity in a tunnel diode.
- These type of diodes can work in high frequency or high speed applications. This is because there is no limitation on time due to low carrier velocity.
- This type of current is not subject to recombination process, making it inherently more efficient.

Note: Diffusion is the main driving force for current in a semiconductor device. It is determined by the concentration gradient between the p and n regions. The diffusion velocity is very less i.e.its a slow process and the main limiting factor when the device is used in high frequencies.

Note: Minority charge carriers recombine by mutual annihilation of positive holes and negative electrons and this recombination rate is high in the surface. Recombination causes loss of energy to the system.

A seemingly innocent act of heavy doping of the diode gives unexpectedly many advantages to the tunnel diode. These advantages are often the reason this diode is preferred to the regular semiconductor in special conditions such as high frequency, inconsistent temperature etc.

- This type of diode is resistant to nuclear radiation and can be used in the core of reactors without worrying about the accuracy of the device changing.
- These heavily doped tunnel diodes are insensitive to temperature.
- A notable advantage of the heavy doping is that, the composition mimics an impure semiconductor material. This has enabled scientists to explore the properties of materials whose purification technology has not yet been discovered.
  Also, the material used in the tunnel diode need not be heavily purified, simultaneously decreasing the production cost as well as increasing the ease of operation[12].

There can always arise a question that, even though the classical diodes and transistors work with electrons and holes, why doesn't tunneling come into play there?
The answer to this question lies in the understanding of doping and tunneling.

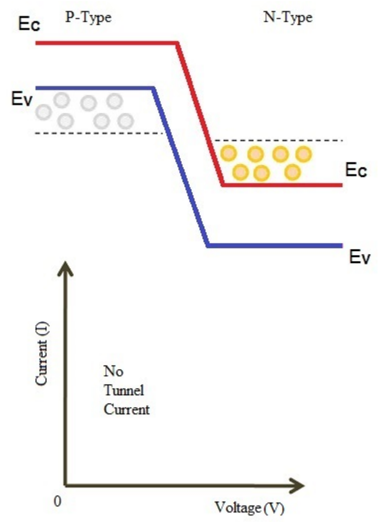
It is an established fact that heavy doping decreases the band gap between conduction and valency band. The depletion layer between p-side and n-side acts as the barrier.

A GaAs p-n junction diode has a depletion layer width of approx. 870nm, whereas the heavy doping decreases the depletion layer width in tunnel diodes to almost 5-10nm.

Tunneling can ideally take place only when the barrier width is 1-3nm.
Thus in the tunnel diode, as the voltage increases, the depletion layer width reduces until this barrier width is achieved, at this point we notice the negative resistance trend. Further explanation about the negative resistance trend is given below.

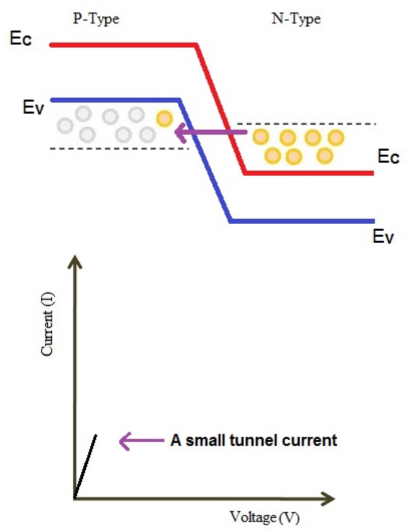
Lets us plot the I-V graph for a tunnel diode starting from an unbiased condition to a high voltage and observe how the current changes with increasing bias voltage.

1. Unbiased region:



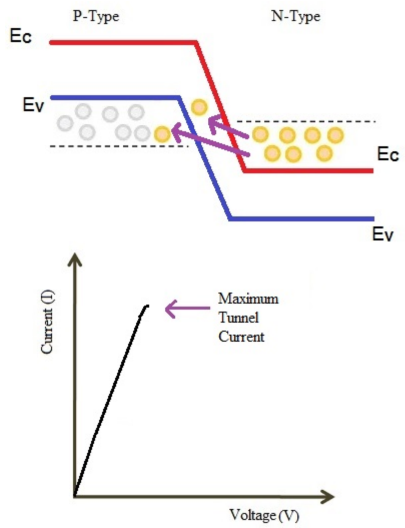
Here, there is no overlapping between the conduction band of p-side and valency band of n-side. And the lack of any bias voltage means that the depletion layer is thick and does not facilitate tunneling. Thus there is no current.

2. Small current region:



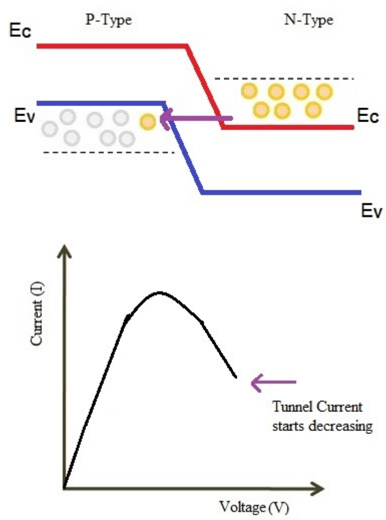
When a small voltage is applied across the p-n junction of a tunnel diode, the conduction band of p-side and valency band of n-type begins to overlap. Though the voltage here is not sufficient enough to surmount the depletion layer barrier, a small number of electrons tunnel through the barrier from the valance band. This produces a small amount of linear tunnel current as shown in the graph.

3. Maximum current region:



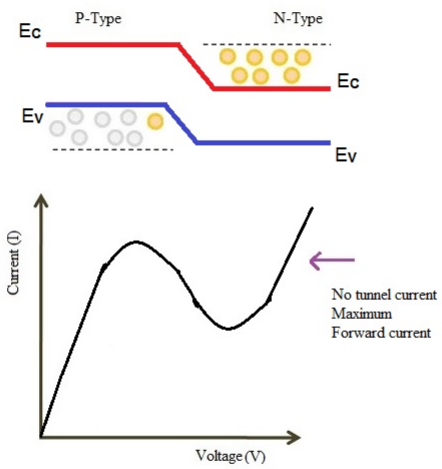
Here, There occurs a maximum overlap of conduction band of p-type and valance band of n-type semiconductor of tunnel diode. This maximum overlap results in a peak in the tunnel current, depicted in the graph as peak current. Note that this peak takes place when the bias voltage caused the depletion layer to be thin enough to promote the maximum tunneling.

4. Negative differential resistance region:



Beyond the peak current region, as the bias voltage is increased, the overlaps between the bands reduce, almost like uncrossing of the conduction and valency bands of p-type and n-type respectively, this results in A decrease of current with increasing bias voltage, resulting in a part of the graph to follow a negative resistance trend as shown in the graph above.

5. Diffusion current region:



The negative resistance trend hits a valley when the conduction and valance bands completely uncross. At this point, the diffusion current due to the bias voltage is large enough for the electrons to move across the barrier in a classical way similar to a typical p-n diode at this region in forward bias.

6. Reverse bias region:

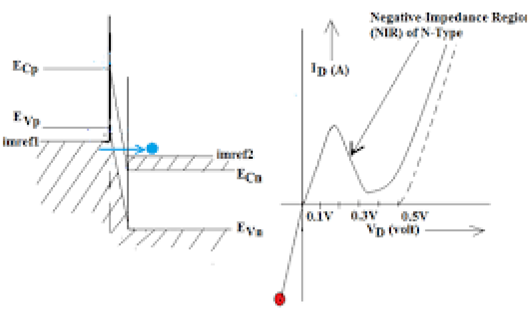
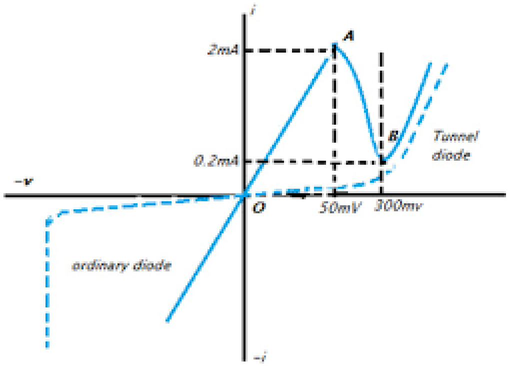


Figure 3.22.G.Under Reverse Bias at small reverse voltage Zener Breakdown occurs and large current in reverse direction starts flowing giving it a backward diode characteristics in backward direction and tunnel diode characteristics in forward direction.

In reverse bias, the electrons from valance band of p-type tunnel into the conduction band of n-type. This tunneling happens monotonically where a large amount of current flows with a little reverse bias voltage. The tunnel diode behaves like a zener diode with almost zero breakdown voltage.[11]

Comparison between I-V graphs of tunnel diode and p-n junction diode



The next step would be to put the values of peak current, maximum negative resistance, peak voltage, valley current, and valley voltage into formulas.

$T_{tunnel} = \frac{V}{R_0} e^{-(\frac{V}{V_0})^m}$

where V is bias voltage, V_0 ranges from 0.1 to 0.5V , m ranges from 1 to 3 in value and R_0 is T.D resistance.

Peak Voltage: the voltage where the current is maximum.

This implies, at $\frac{dT_{tunnel}}{dV} = 0$, V = V_{peak}

$\implies \frac{1}{R_0} e^{-(\frac{V}{V_0})^m} - m(\frac{V}{V_0})^m e^{-(\frac{V}{V_0})^m} = 0$

$\implies 1 - m(\frac{V}{V_0})^m = 0$

or

$V_{peak} = ((\frac{1}{m})^{\frac{1}{m}}) V_0$

And

$I_{peak} = ((\frac{1}{m})^{\frac{1}{m}}) V_0 e^{\frac{-1}{m}}$

The negative resistance of a tunnel diode can be defined as

$R = \frac{1}{\frac{dI}{dV}}$

$\implies R = R_0/(1-(m(\frac{V}{V_0})^m) e^{-(\frac{V}{V_0})^m})$

When R is maximum, $\frac{dI}{dV} = 0$

in other words,

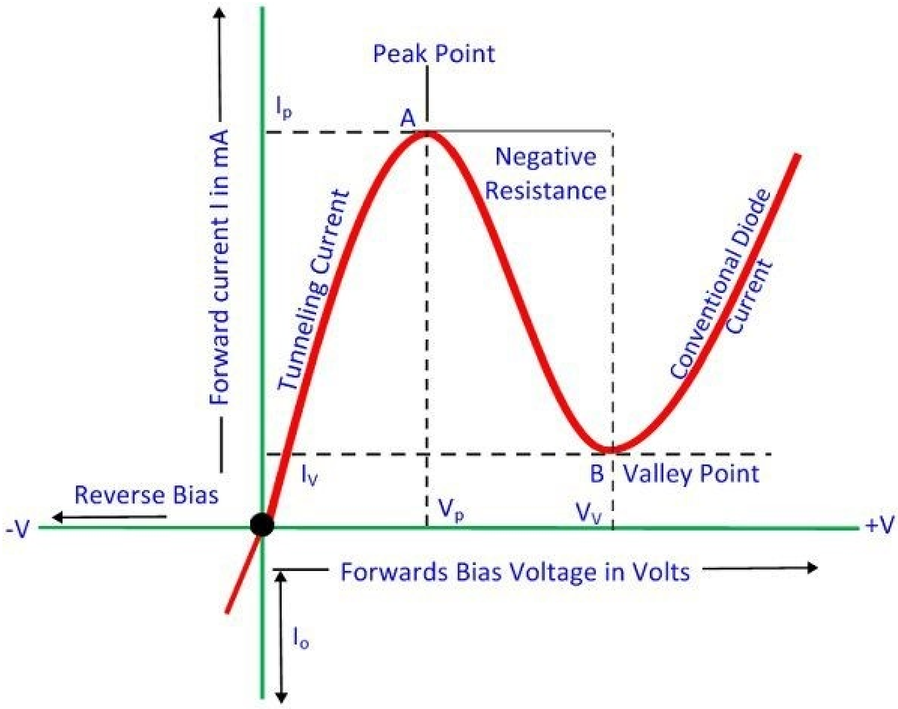
$(1-(m(\frac{V}{V_0})^m) e^{-(\frac{V}{V_0})^m}) / R_0 = 0$

At the point where $V = V_0(1+\frac{1}{m})^{(\frac{1}{m})}$

So,

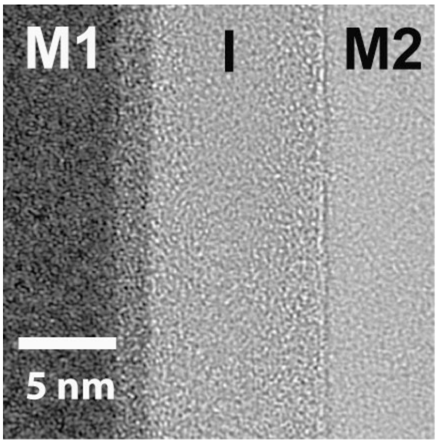
$R_{max} = -(R_0(e^{(\frac{1+m}{m})}))/m$[12]

The below graph gives a better picture of what these values exactly represent.



5 MIM diode

The first diode looked like this:



Metal-insulator-metal diodes are semiconductor devices with non-linear properties. The thin film contact area of this diode is in micro meters, and the insulator layer is in single digit nanometer range.

In the conduction of MIM diode, tunneling current plays a dominant part. This diode is used largely for its ability to work as high frequency devices.

MIM diodes have large current densities and high values of asymmetry in their IV characteristics. These two distinguishable characteristics of a MIM diode open up many different applications for this device.

In this device, two metals(usually different metals) are present in the ends and an insulator, generally a metal oxide is sandwiched between these layers. The asymmetry in the construction of this diode is the basis for its asymmetry in I-V graph behavior.[3]
Some examples of MIM diodes are $Ni/NiO/Cr,\ Al/AlO_x/Gr,\ Al/Al_2O_3/Ag$

Thickness of the insulator plays a crucial role in the behavior of a MIM diode. This is because the tunneling efficiency drops exponentially with increasing width of insulator. The insulator thickness should be less than 10nm for accurate working of the device.

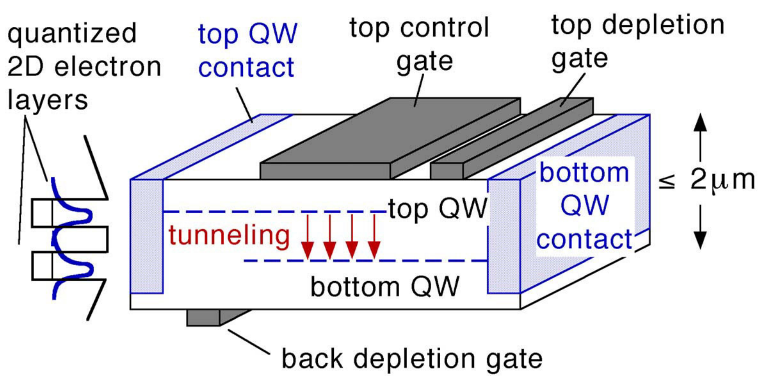
The insulator layers of an MIM diode can be increased, and these modified diodes have been studied to have amplified properties as compared to the regular MIM diode. Even 4 insulator layer diodes have been synthesised. The asymmetry values of these modified devices are exponentially high when compared to the regular device. [2]
Some examples of M-nI-M devices are $Cr/TiO_2/Al_2O_3/Ti,\ Pt/TiO_2/TiO_1.4/Ti$

6 Other devices

1. Double-electron-layer-tunneling transistor

   Also called as DELTT, this is a planar device of gate controlled 2D-2D layers of electrons acting as an insulator between a double quantum well that is conductive. This type of transistor was initially developed to make up for the lack of three terminals in its 3D-2D counterpart. DELTTs can be fabricated in very large numbers with an easy ease.

   The resonant tunneling factor in a DELTT is noticeably sharp, owing to the 2d electron layer. Here, the gates at surface modulate and control the tunneling of electrons by varying density of the layers.. Tunneling invariably produces prominent negative differential resistance characteristics in a DELTT. The peak and value ratios for current and voltage are around 20:1 at 1.5 Kelvin which are pretty good numbers. The tunneling is precisely controlled by similar momentum and energy that is in-plane of the electron, Only a particular resonance will facilitate appropriate tunneling through the insulating layer.

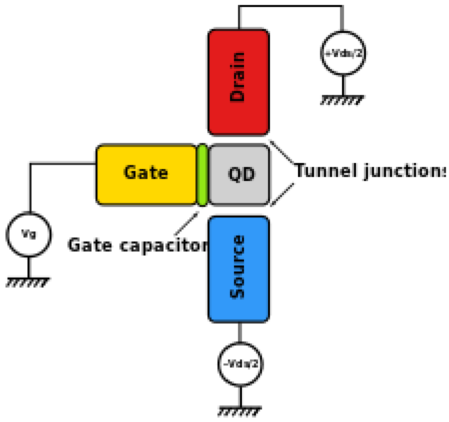
   We can look at the DELTT having two terminals with electrical contacts. Among these terminals, the one at the source has contact to the top quantum well, and the one at drain has contact to bottom quantum well

   The gates in this device are very closely placed, the whole device is not more than 2 microns in width. This is to prevent as much tunneling leakage current as possible. These leakage of currents take place in between the control gate present and the depletion gate.
   Some of the advantages of a DELTT are sharp resonance in tunneling due to the precise

restriction in momentum and energy and the density control by the gates present on surface.[10]

2. Single-electron quantum-dot transistor

   A QD-SET device has a small quantum dot with source and drain leads attached to it. A quantum dot is a 3D nano-material that has localised electrons. It is often referred to as artificial atom.



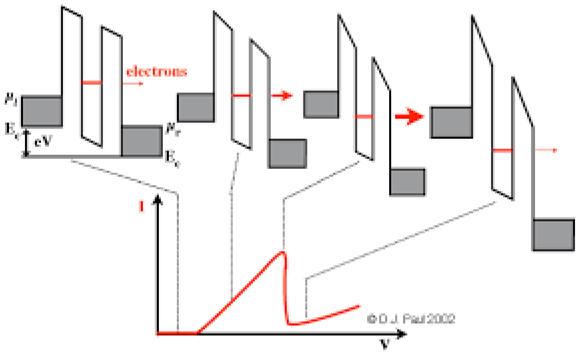
   The QD-SET used controlled electron tunneling that causes an amplification of current. A SET is very similar to a field emission transistor except that the channel in a FET is replaced by a Quantum Dot.[8]

   Quantum dot is separated from the source and drain by a small high band gap semiconductor that acts as the tunnel barrier.

   A gate is connected to Quantum dot by a capacitor which controls the amount of current by tunneling. The gate is connected to the Quantum dot by a capacitor.[5]

3. Resonant tunnel diode

   It is a double barrier quantum well diode, where we can notice two heavily doped conductive semiconductors with a small gap, and the well has is a high band gap semiconductor.



As we can see in the graph above, the RTD has a negative resistance region similar to a tunnel diode

In an RTD, the tunneling across the quantum well barrier happens when there is a resonance in both of the conducting semiconductors. This resonance occurs only for some specific values of bias voltage.[4]

This resonance condition is what makes an RTD an effective ultra high switch.

7 Conclusion

From everything stated above, we can arrive at a number of conclusions. All those conclusions are listed below.

- The tunneling probability decreases exponentially with the barrier width and barrier energy.

- for a high energy barrier with thick width, the tunneling probability P_t is $P_t \approx 16\,\frac{E}{V}(1 - \frac{E}{V})\,e^{-2\beta d}$

- Tunneling effects are seen in semiconductors with doping concentrations as high as 1 in 10^3 . The high dopiant concentration decreases depletion layer's width to less than 10nm. The narrow width is the reason why tunneling effects are prominent.

- The devices with tunneling effects have some common points that are observed in all of the devices mentioned above:

- They are used in high frequency and high speed signal applications
- They involve elements of the higher atomic number.
- The I-V graphs are different from the I-V graphs of conventional semiconductor devices.
- They generally follow a negative resistance trend
- The efficiency of tunneling in these drops exponentially with the increasing width of layers.

References

(citations)

1. Razavy, Mohsen (2003). Quantum Theory of Tunneling. World Scientific. pp. 4, 462. ISBN 978-9812564887.
2. Yavuz M(2013),Higher Performance Metal-Insulator-Metal Diodes using Multiple Insulator Layers, Austin J Nanomed Nanotechnol
3. Choi, Kwangsik & Yesilkoy, Filiz & Ryu, Geunmin & Cho, Si & Goldsman, Neil & Mario, Dagenais & Peckerar, Martin. (2011). A Focused Asymmetric Metal–Insulator–Metal Tunneling Diode: Fabrication, DC Characteristics and RF Rectification Analysis. Electron Devices, IEEE Transactions on. 58. 3519 - 3528. 10.1109/TED.2011.2162414.
4. Sun, Jian & Haddad, George & Mazumder, P. & Schulman, Joel. (1998). Resonant tunneling diodes: Models and properties. Proceedings of the IEEE. 86. 641 - 660. 10.1109/5.663541.
5. Karmakar, Siddhartha & Jain, Faquir. (2015). Circuit Model of Different Quantum Dot Based Field Effect Transistors. Silicon. 7. 10.1007/s12633-014-9252-9.
6. Sophya Garashchuk, Bing Gu, James Mazzuca, "Calculation of the Quantum-Mechanical Tunneling in Bound Potentials", Journal of Theoretical Chemistry, vol. 2014, Article ID 240491, 11 pages, 2014. https://doi.org/10.1155/2014/240491
7. RICHARD H. BUBE, 10 - Junctions, Editor(s): RICHARD H. BUBE, Electrons in Solids (Third Edition), Academic Press, 1992, Pages 214-241, ISBN 9780121385538,
8. Om, Kumar Manjit, Kaur. (2010). Single Electron Transistor: Applications Problems. International Journal of VLSI Design Communication Systems. 1. 10.5121/vlsic.2010.1403. (set)
9. Nikhil M. Kriplani, Stephen Bowyer, Jennifer Huckaby, Michael B. Steer, "Modelling of an Esaki Tunnel Diode in a Circuit Simulator", Active and Passive Electronic Components, vol. 2011, Article ID 830182, 8 pages, 2011. https://doi.org/10.1155/2011/830182
10. Morath, Christian Lilly, M. Simmons, J. Reno, J. Baca, W. Moon, J.. (2002). Temperature Dependence of Three-Layer Double Electron Layer Tunneling Transistor.

11. J.O.Scanlan, Analysis and Sysnthesis of Tunnel Diode Circuits, (1996) (book), Library of congress catalog card number 66-24360
12. MA Lee, B.Easter, H A Bell, Tunnel diodes-Tunnel Diodes (Modern Electrical Studies), 1967
13. Övgün, Ali. (2015). Quantum Tunneling and Hartman Effect. 10.13140/RG.2.1.2979.3449.

(websites)

- Early History of quantum tunneling
- Press release for Noble prize for Leo Esaki
- MIM diode
- MIM diode
- MIM diode
- MIM diode
- Tunnel diode
- MIM diode
- DELTT
- RTD
- RTD
- Image
- Tunnel diode
- Tunnel diode
- Tunnel diode
- Image
- Image
- ImageT
- image SET