كلية الدراسات العليا

Sudan University of Science and Technology

College of Graduates Studies

# Confinement Effects on Electronic Properties of Nano Scale Semiconductors

# آثار التحجيم على الخصائص الإلكترونية لأشباه الموصلات النانوية

A Thesis Submitted for Fulfillment Requirement of Degree of
Doctor of Philosophy in Physics

**Submitted by:**

**Mohamed Yassin Mustafa**

**Supervisor:**

**Prof. Dr. Mubarak Dirar Abd-All**

# Acknowledgement

It is my pleasure to express my gratitude to Sudan university of science and technology, graduate college, and faculty of science department of physics for moral support during doing this work. Special thanks to my supervisor prof. Mubarak Dirar Abd-Alla for encouragement and fruitful suggestions. Thanks extends to Dr. Asim Fadol for encouragement and invaluable help. Special thanks to Mr. Tageldeen Mohamed Saeed for typing this manuscript

# Abstract

The behavior of macroscopic electronic components can be easily understood using classical electric and electronic relations. Unfortunately such classical relations can no longer capable of describing nano electronic components. This is due to the fact that nano particles can not be described by classical laws, but one needs quantum laws in this case. This needs new quantum electronic model that can describe nano electronic components behavior. Using quantum uncertainty principle and statistical distribution current, one can find conductivity, conductance and resistance in terms of Plank constant, density of energy states and diffusion constant. These quantities are shown to be quantized. These relations enables writing ohms law as a sum of classical part sensitive to the conductor length beside a quantum part sensitive to chemical energy and diffusion current at diode interface. The formal definition of resistance and conductance in terms of current and volt has been used to find the quantum resistance. The energy density of states manifests it self through the statistical distribution function, while the plank constant was incorporated through the chemical potential. The statistical nature of these parameters manifests itself through the diffusion constant which is related to the current density. The same arguments were used to obtain quantum capacitance in terms of the energy density of states, for non-equilibrium systems.

The Boltzman transport equation has been written in terms of a chemical potential variation instead of the statistical distribution function, which has been smeared out using simple mathematical tricks. The Boltzman equation was thus used to obtain conductance and resistance for quantum non equilibrium systems.

# مستخلص

يمكن وصف سلوك النبائط الالكترونية المرئية باستخدام العلاقات التقليدية في الكهرباء والإلكترونيات بسهولة. ولسوء الحظ لا تستطيع هذه العلاقات التقليدية وصف النبائط الإلكترونية النانوية وهذا يعزى لحقيقة أن الجسيمات النانوية، لا يمكن وصفها بالقوانين التقليدية بل تحتاج للقوانين الكمية في هذه الحالة. وهذا يحتاج لنموذج إلكتروني كمي جديد يمكن أن يصف سلوك النبائط الإلكترونية النانوية باستخدام مبدأ عدم التيقن الكمي وتيار التوزيع الإحصائي. يمكن إيجاد الموصلية والتوصيلية والمقاومة بدلالة ثابت بلانك وكثافة مستويات الطاقة وثابت الانتشار. وتم توضيح أن هذه المقادير مكممة. وهذه العلاقات مكنت من كتابة قانون أوم كمجموع تقليدي يعتمد على طول الموصل بجانب جزء كمي يعتمد على الجهد الكيميائي وتيار الانتشار عند السطح الفاصل للوصلة الثنائية. تم استخدام الصيغة النموذجية للمقاومة والتوصيلية بدلالة التيار والجهد لإيجاد المقاومة الكمية حيث ظهرت كثافة مستويات الطاقة من خلال دالة التوزيع الإحصائي في حين ضمن ثابت بلانك عبر الجهد الكيميائي وقد ظهرت الطبيعة الإحصائية لهذه المعاملات عبر ثابت الانتشار الذي له علاقة بكثافة التيار. واستخدمت نفس الفرضيات لإيجاد السعة الكمية بدلالة كثافة مستويات الطاقة للمنظومات الغير متزنة. تمت كتابة معادلة الانتقال لبولتزمان بدلالة تغيرات الجهد الكيميائي بدلاً من دالة التوزيع الإحصائي الذي أزيل باستخدام بعض الحيل الرياضية البسيطة. وعليه تم استخدام معادلة بولتزمان هذه لإيجاد التوصيلية والمقاومة للمنظومات الكمية غير المتزنة.

# Table of Contents ☐

# Chapter One

# Introduction

## 1.1 Research motivation

This research is motivated by the recognition of the future of nanotechnology. Nanotechnology evolved very fast over the last forty years. Electronic computers have grown more powerful as their basic subunit, the transistor, has shrunk. However, the laws of physics changed. This change in physics laws put the scientists in a good opportunity of time not met before, to give them the capability to manufacture objects down to the size of few hundred atoms.

General speaking, if the manipulation of material under one micrometer to be put into that perspective, that 0.001 of one micrometer (about 800,000 times smaller than one strand of hair), when applied, many things change dramatically with this scale, specially the laws of physics. This research focuses on the innovation of nanoscale devices uniquely enabled by emerging electronics for a wide range of applications, such as computing, sensing, energy conversion, and communication. The research philosophy is to address major challenges in these domains by pioneering the fundamentals of new concept prototypes to enable smarter life, by employing distinctive “atoms to applications” through examination of traverse low-dimensional material physics approach, to learn more about predictive ab-initio (*from scratch*) modeling electrical transport theory.

## 1.2 Importance of the research:

This research is very important because the transistor is the key invention of modern society which have paved the way for innovative technologies. Due to their economic importance and pressure to keep production costs low, while simultaneously increasing efficiency, transistors have been scaled to deka-nanometer regime during the last decades. This excessive down scaling of dimensions has now reached several physical limits, whereby, transistor building and manufacturing fabrication have been challenged numerous times. Although some of these challenges have been overcome by the usage of new materials and clever building of device geometries. A lot of additional efforts must be put into the research work to innovative nanotechnology in order to open up for new opportunities. In this work recent development in transistor technologies has been summarized to give short overview of new future technologies people are going to win, such as TMOSFET.

## 1.3 Overview of the research problem:

One of the problems facing this research is the confusion made by the scientists about how they can define the nanotechnology. Most of them focus on the study and control of the phenomena and material at length scale (100nm). Since the basic of conventional procedure for scaling down from top to bottom approach is rapidly approaching inherent physical limits of miniaturization, the nanoelectronics circuitry requires the realization of three dimensions integration based on nanodevices. To achieve this goal, significant progress has been made, and miniaturized electronic circuits are integrated on semi-conductor chips where the basic elements is the transistor, whose size is taken under (100nm). The problem with this is how to make these devices self-consistent and extremely comfortable, feasible, and adaptable according to the needs of people.

Continued success in device scaling is necessary for further development of the semiconductor industry in the years to come. A group of leading companies publishes their projections for the next decade in the most recent International Technology Roadmap for Semiconductors (ITRS-99) *[International Technology Roadmap for Semiconductors, Semiconductor Industry Association, CA2]*. The roadmap projects a device gate-length down to ~30 nm around 2014 *[ International Technology Roadmap for Semiconductors, Semiconductor Industry Association, CA2*]. This forecast promises another ten years of brightness. Scaling beyond 30 nm, however, can be much more difficult and different. Now the research work is quite close to the fundamental limits of semiconductor physics. How much further down it can go? It is hard to answer. Nevertheless, without doubt, the research work is facing numerous challenges, both practically and theoretically, it requires new theory and approaches to help in understanding device physics, so as to build devices at the sub-30nm scale. Efforts have been put forth in recent years, but much more is needed. For these purposes, this research project has been proposed, the results of which may help to understand device physics to build devices at nanoscale.

## 1.4 Research objectives:

Nanoelectronics aims at improving the capabilities of electronic devices and display while shrinking them to a very small size, reducing their weights and power consumption. Hence the objective of this research is to develop research efforts to increase the usefulness and reliability of these nanoelectronics' devices inventory by continued realization of Gordon Moore law using new methods and materials to build electronic devices with feature size on the nanoscale.

## 1.5 Historical background

In the year 1984 researchers were involved in the research of how they can get to smaller devices **[1,2,3,4,5,6]**. They looked at the active region (the channel) in the following model consisting of a resistor represented by a channel and two big contacts to bring it down in few (microns) in size **[7]**,

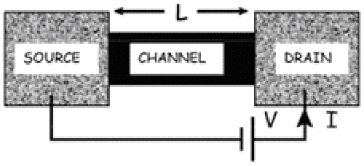


Figure 1.1 ;Three region device representing three dimensional device

What has happened, they brought it in a submicron to about (100nm), and now researchers are thinking to bring it to sub (10nm). Nobody now knows how to bring it to less than that in the near future. The device has been driven in the submicron scale due to the need of the continuous miniaturization to pack more and more of them into small area of a chip **[8,9].**

The model is basically a resistor, when a voltage is applied across it, current flows. A third terminal to control the current flow is needed to make it a transistor.

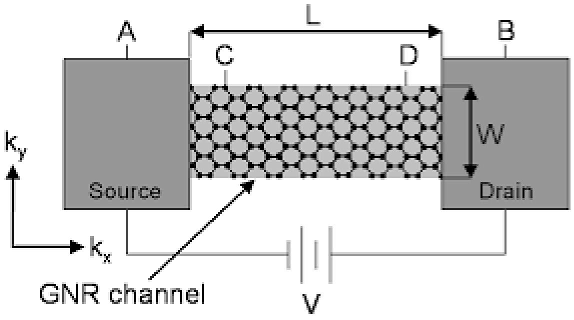


Figure 1.2 ;Three region device representing two dimensional homogenous material

During the last three decades the dimensions of a transistor channel has gone down by a factor of one hundred, although the transistor is somewhat greater than the channel, because of the two big contacts, roughly the dimensions is going down by one hundred, that means ($10^4$) more transistors can be packed in there. The largescale integration probability may reach hundred thousand ($10^5$) to overcome the ($10^4$) by a factor of ten more. But what has happened today the number of the accommodated transistors reach about a billion transistors in an area of not more than (3x3) $cm^2$.

Now researchers are debating what will happen after this since they are getting so close to the atomic dimension and this research will not be dealing with this, because up to now researchers don't know the answer and they don't know the type of new devices they are going to win in the near future.

This research discussion will be confined to the basic understanding of the current flow and how it could be controlled by studying how it evolved and how it has been affected by what is known today in comparison of what is not known before.

Back to 1984 what is known, the resistance is given by;

$$R = \frac{\rho L}{A} = \frac{V}{I} \qquad (1.1)$$

Where, L is the conductor length.

A is cross-sectional area of the conductor

ρ is the material property (resistivity), the reciprocal of which is the conductivity σ;

$$\sigma = \frac{1}{\rho} \qquad (1.2)$$

There are a lot of theories which deals with the dimensions of the resistivity, and the conductivity.

By 1989 ,five years after 1984 a completely new thing appeared at that time, when there is really a short conductor, the electrons instead of going through the conductor in a diffusive random manner (random walk) it passes through as bullet without hitting anything in between the two contacts, known later as (ballistic transport) between the source and the drain, in this case the resistance is quantized by $(\mathrm{h}/q^2)$ the value of which is 25kΩ and the resistance of the ballistic conductor will be given by **[10]** ;

$$R_B = (h/q^2).\frac{1}{M} \qquad (1.3)$$

Where: $h$ is the Plank's constant

q is the charge of the electron

M is the number of modes

The number of modes (M) is going to be explained later inside this dissertation, but for the time being it can be looked at as a numerical factor in equation (1.3) above.

Because of this ballistic phenomena what has happened in 1989,there were two groups *(from Nano hub-u)* who reported on this experimentally and before that there was a lot of debate arguing what is really the resistance of a very short conductor or what exactly would it be, because if they use the resistance of 1984 , given by equation (1.1) above , it must go to zero when the length of the resistor (L) goes to zero, but it doesn't, it goes to some other constant value .

During the period (1984 _ 1989) the answer came in a clear cut to show the importance of the experimental work.

Since the resistances $(R = \rho\mathrm{L}/\mathrm{A})$ and $R_B = \left(\frac{h}{q^2}\right)\left(\frac{1}{\mathrm{M}}\right)$ are totally look different, but they could be connected by defining something, that is called the mean free path (mfp) **[11,12,13]** in the following form;

$$R = R_B(1 + \frac{L}{mfp}) = (\frac{h^2}{q})\left(\frac{1}{M}\right)\left(1 + \frac{L}{mfp}\right) \qquad (1.4)$$

Equation (1.4) above describes the resistance of a very short conductor and the very long one as well. When the length (L) is zero the ballistic form is given, and when (L) is longer compared to the mfp, by dropping the one, Ohm's law is given in which (R) is directly proportional to the length (L), so a nice clear picture appeared from (the bottom up) approach. (*National Science Foundation, the Intel Foundation, and Purdue University who laid the foundation for the Computational Nanotechnology's "Electronics from the Bottom Up" initiative.)*

## 1.6 Change in paradigm (Theory)

During the last three decades (20 to 25) years ago, current flow developed in a different prospective compared to what is used to be in the elementary physics (fresh man's physics). Simply the picture carried in physician mind starts at the diffusion end, but now the nature of the current flow in small devices is the ballistic transport **[15]**.

In nanoelectronics the most important device is the Nano transistor (see figure1.3), which in its simplest form consists of a channel through which electrons can flow between the two contacts (the source and the drain), the flow of electrons give rise to current flow in the external circuit according to Ohm's law.

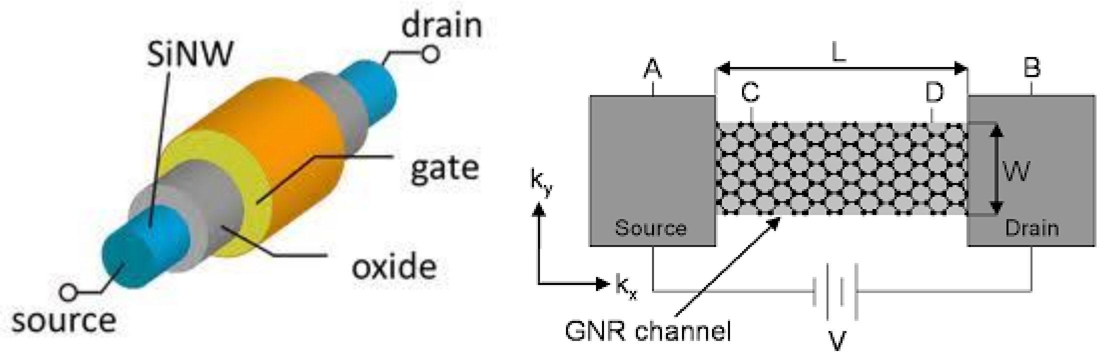


Figure 1.3 ; Simplest form of nanotransistor construction

The word Nano stands for a nanometer in length scale, equals to ($10^{-9}$ meter). The word micron is ($10^{-6}$ meter) also called a micrometer, therefore a nanometer is one thousandth of a micron. The distance between atoms is a little less than one nanometer.

Historically as stated above during the last five decades the transistor channel continuously reduced to minimize the transistor size so as, to get more transistors to fit in a chip at a given area, and that is why computers of today are more powerful than those of the 1985. In a chip ($3x3cm^2$), if the channel is reduced to (10 μm) long, the whole transistor will be ten times that much, because of the extra two big contacts, then the size of the transistor will be (100 μm) or (0.1mm), that means each millimeter can accommodate ten transistors, three centimeters accommodate (300) transistors, in total there will be ($10^5$) transistors in an area of 9 $cm^2$, that is to say (300) transistors each way. Today each dimension is reduced by a factor of (100) to fit (300000) transistors each way in the same area to go up to ($10^9$) transistors, approximately one billion transistors. This very small device that has come about has changed the paradigm in which the electron goes from source to drain straight through in a ballistic transport, which represents the change in paradigm in contrast to the diffusion transport in the devices of 1985.

Ohm's law sometimes is used to be denoted by ($R = \frac{V}{I} = \frac{\rho L}{W}$) the first part ($R = \frac{V}{I}$) does not change, but due to the shrinking of the device the new Ohm's law can be represented by the second part ($\frac{\rho L}{W}$), after being modified. If this work has been taken in the order of nanometer, Ohm's law will take the form;

$$R = \left(\frac{V}{I}\right) = \frac{\rho}{W}(L + something)$$

This (*something)* is the mean free path as stated above in equation (1.4). The mean free path can be defined by the distance traveled by the electron before it turned around, by this equation when the channel length (L) turns to zero the resistance does not go to zero, but it goes to some other constant value. To connect the resistance to this (*something)*, in that case there will not be a mean free path, because when the device goes to a short distance, (ρ) doesn't mean much by itself and so the mean free path (mfp), though their product means that (*something*), and that is a constant to denote the resistance of something that is really short.

### 1.7 (Bottom -up) Approach

*(National Science Foundation, the Intel Foundation, and Purdue University who laid the foundation for the Computational Nanotechnology's "Electronics from the Bottom Up" initiative.)*

Understanding of current flow usually based on much bigger resistances evolved from($R = \frac{\rho L}{A}$). Getting to small devices the usage of this relation is needed with the trial to understand what has been observed in a very short conductor. This approach seems to be rather difficult, and it looks very puzzling, because after all if the length went to zero the resistance should go to zero, and there will be no reason to go to one term resistance ($\frac{h}{q^2}$) divided by (M) whose significance is the number of modes, this is a very important concept which will be discussed later, it emerges from the field of nanoelectronics and microscopic physics**[18,19,20].**

At hand now if the bottom-up approach is taken to start from ($R_B = \frac{h}{q^2} . \frac{1}{M}$), which applies to a very short conductor to be compared to the diffusive conductor, where($R = \frac{\rho L}{A}$), usually those two relations look very different in forms, but actually they are the same when connected together, whereas for the resistance of a long conductor is the ballistic resistance ($R_B$) multiplied by a factor $\left(1 + \frac{L}{\lambda}\right)$, where (L) is the length of the conductor, ($\lambda$) is the mean free path (mfp). This what is known as the new prospective, in general we could write the resistance as;

$$R = R_B\left(1 + \frac{L}{\lambda}\right) = (h/q^2)\left(\frac{1}{M}\right)\left(1 + \frac{L}{\lambda}\right) \qquad (1.5)$$

As the length (L) approaches zero, ($R = R_B$), when (L) is very long the (one) can be omitted and (R) will be equal to $\left(\frac{R_B L}{\lambda}\right)$ getting back to Ohm's law where (R) is directly proportional to the length of the conductor (L). Therefore, the bottom up approach gives a different view of how current flows, but also there is something relatively simpler to understand, to clarifies many concepts, which's going to be discussed in this thesis.

## 1.8 Thesis layout.

The thesis consists of seven chapters. Chapter (1) is the introduction, while chapter (2) deals with Nano electronics. Chapter (3) is concerned with the new modified ohms law. Chapter (4) deals with energy band model while the Nano scale transistor is given in chapter (5). Chapter (6) is devoted for potential change inside channel, whereas chapter (7) is concerned with conclusion.

# Chapter Two

# Nano electronics and it's new prospective

## 2.1 Scientific descriptions

Before ballistic conductors were studied back in 1980's, what is confusing was how can such conductors have any resistance if electron went from source to drain ballistically? What is known whenever current flows, there will be heat loss generated given by ($I^2$R). For heat to be generated electrons need to interact with the solid. Then if the electrons go right through the conductor without any interaction there will be no heat generated, though how can there be a resistance of any kind? As an answer to this question, a very clear idea of an Elastic Resistor was described by an elastic process, where there is no exchange of energy in in the elastic resistor itself, all the heating takes place at the ends of that elastic resistor **[18];**

Figure 2.1: An elastic resistance process

There was a good experimental evidence to proof that small conductors behave this way. In carbon nanotube the amount of current passing through is enough to burn it up, if the heating happened to be inside the nanotube, but it doesn't burn, because heating appears at the contacts, those big things, which are able to get rid of that heat.

The separation of the elastic part from the heat generated part makes this type of resistance conceptually very much simpler than what people used to, because this elastic process is very different from what happens at the ends. The elastic part can be described physically by mechanics and the end parts can be described by thermodynamic (flow of heat);

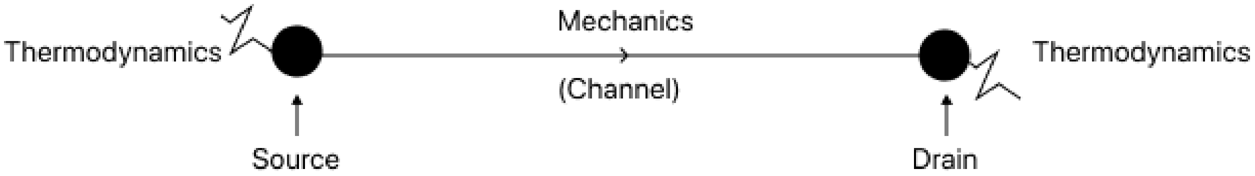


Figure 2.2: Separation of the elastic resistance from the heat generated contacts

Mechanics and thermodynamic are two different branches of science, separated from each other by a very long span of time. Newton's laws came first, then several centuries later heat engines arrived, and heat has considered to be a form of energy which can be converted into work. During the eighteenth and nineteenth centuries both mechanics and thermodynamics were put together, and one of the famous scientists who played a very important role in that area was Boltzmann, who put together Newton's laws and added to them the thermodynamics processes driven by entropy, to create his famous Boltzmann Transport Equation (BTE) as follows; see appendix (I)

Newton's laws +Entropy processes=BTE

What Boltzmann did here is very subtle, and it causes a lot of debate at his time, and even today still people arguing about this. In general, BTE has been accepted as a standard starting point to all semiclassical transport theory (flow of all kinds of things). Boltzmann equation is used as a statistical mechanics in the sense that it consists of mechanics and entropic processes added to it, that is to say (nonequilibrium statistical mechanics), it represents the starting point to all semiclassical transport theory of what has been done in this research (dissertation), is to obtain a lot of results like the thermoelectric coefficients, conductivity, diffusion coefficient ..etc. Because this (BTE) takes some time to get used to and get familiar with, all the above results can be obtained in a relatively elementary way without getting into this (BTE), what made that possible was the introduction of the idea of the elastic resistor, because the mechanics and the thermodynamics are well separated in an elastic resistor, whereas what makes the transport so complicated using (BTE) is the mesh up of the mechanics and thermodynamics like an intertwin. *In this proposed new prospective, mechanics and thermodynamics are nicely separated from each other* **[21,22],** and that was really made everything so simple to handle in an elementary way, because thermodynamics processes continually restore contacts into equilibrium, so whenever an electron comes in from the contact can be described by its equilibrium distribution function (Fermi-function), and just with this statement all the details of what is really going on inside the contacts can be bypassed, but this only applies to relatively small things. Out of the small scale lies the big things, to extend these ideas to big devices, conceptually these results can be applied approximately by thinking of a big device as a punch of a very little conductors in series. This picture provides a nice approximation, and it also agrees with rigorous theory for low bias to give all the standard results given by rigorous theory (what would be got from BTE) **[23,24].**

The quantum version of the subject, where the electron will be thought of as a wave, will be introduced instead of Newton's laws, where the electron will be thought of as a particle, therefore Newton's laws get replaced by Schrodinger equation, then still entropy driven term should be added. The resulting equation is known as; Quantum Boltzmann Transport Equation (QBTE), but normally called None Equilibrium Green's Function (NEGF) method *which involve work of many people,* that why it is preferred to take an impersonal name rather than just Boltzmann Quantum Transport Equation;

NEGF =Schrodinger Equation + Entropy driven

This equation seems to be rather difficult stuff, because it consists of a none equilibrium statistical mechanics, but what makes it possible to be applied is the idea of the elastic resistor. Schrodinger equation looks as shown here below;

$$H\psi = E\psi \qquad (2.1)$$

Where; $H$ is a differential operator which can be converted into a matrix version.

$E$ is the eigen values of energy.

$\psi$ is the wave function.

Schrodinger equation as it stands in this form, is not enough to analyze devices, it needs to include the entropy driven term (the contacts and sources), the extra term to be added will look like this,$(\Sigma\Psi + S)$, to represent the effect of all contacts where all the entropic processes take place, therefore;

$$E\psi = H\psi + (\Sigma\psi + S) \qquad (2.2)$$

Then what about NEGF since equation (2.2) above doesn't look more than something totally different modified Schrodinger equation, where matrix $(H)$ describes the electronic highway (the available energy levels for electrons to flow), matrices $(\Sigma)$ represent the flow of electrons in and out of the contacts, interactions in the channel can be added as a matrix $(\Sigma_0)$, all these processes develop NEGF method. It is quite clear that this modified Schrodinger equation cannot be used directly or at least should be carefully treated, because of the assumption that if two electrons are coming in from the two contacts, $(S_1)$ from the source with a wave function$(\psi_1)$, $(S_2)$ from the drain with a wave function $(\psi_2)$, if both of them present at the same time there will be no capability to add them together algebraically, because they are complex numbers, and their addition is a vector sum , therefore their squares should be added **[25];**

$$(S_1 + S_2) = \psi_1(\psi_1)^* + \psi_2(\psi_2)^* \qquad (2.3)$$

Where the star indicates the conjugate of the wave functions.

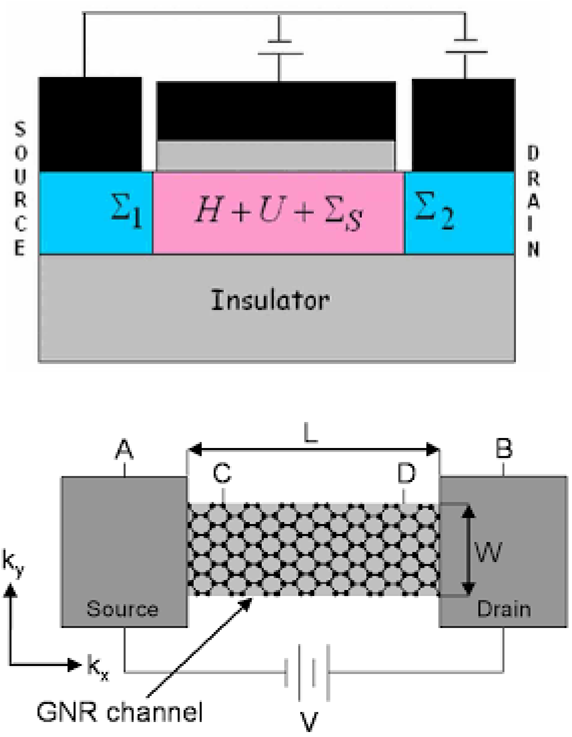


Figure 2.3: Effect of all contacts where entropic processes take place

By analogy to light sources in this case, there will be two types of sources, incoherent thermal source (the ordinary thermal light bulb), and the laser source. In the first one powers of the light bulbs can be added, but their electric fields cannot be added. In the laser source electric fields can be added. Electrons in normal solids have the equivalent of thermal sources and not the equivalent laser sources, although the superconductors are more like a laser sources, but this not a part of this research, those who are interested may revise (*Fireman's Lectures, volume iii, the last chapter).*

Instead of working directly with the modified Schrodinger equation using the wave functions ($\psi' S$), when there are multiples of sources, NEGF equation is used to avoid adding every single wave function from each separate source.

NEGF directly relates ($\psi\psi *$) to be written like this **[26];**

$$(\psi\psi *) \dashrightarrow G^n = G^R \Sigma^{in} G^A \quad (2.4)$$

Where $G^n$ is the none equilibrium Green's function

This is the basic equation for NEGF method or the None Equilibrium Quantum Statistical Mechanics originated back in the 1960's with *the work of many people*, although the notion is slightly different (*see1965, papers by; Keldish Chile*), and the difference requires fairly advanced formalism, because it involves many body perturbation theory, which need more than a complete year to master. Fortunately by the application of the new prospective (elastic resistor principle), all that can be avoided without getting into many body perturbation theory to go forward to one electron Schrodinger equation to show that how it can be used to understand all types of experiments of today, involving one conductor quantization, quantum Hall effect and spin transport experiments. That is why intentionally this part is left for future work.

## 2.2 The Basic Conceptual Framework

The continuous miniaturization processes of making small devices, has changed the nature of current flow from the diffusive random walk based on Ohm's law given by equation (1.1) above, to the ballistic form in which the resistance is quantized by the value ($h/q^2$) with units of (KΩ), it's exact value is (25KΩ) at room temperature, which could be multiplied by $\left(\frac{1}{M}\right)$, as given by equation (1.3) above, where M is the number of modes. The number of modes is an important concept going to be explained later as this research goes on. Equations (1.1) and (1.3) look very different in their forms, but they are connected by the generalized form of the resistance given by equation (1.4) above, after rearrangement it takes the following form stated here below;

$$R = R_B \left(1 + \frac{L}{\lambda}\right) \quad (1.4)$$

When the length $(L)$ is much smaller than (0.01λ) the second term $\left(\frac{L}{\lambda}\right)$ can be dropped to have the ballistic resistance $(R_B)$, on the other hand when $(L)$ is hundred times longer than (λ), the (one) can be dropped to have the diffusive resistance which is directly proportional to the length where Ohm's law connects.

Expanding the generalized form of the resistance given by equation (1.4) mentioned above, it consists of two terms $(R_B)$ and $\left(\frac{R_B L}{\lambda}\right)$, that it is one term is directly proportional to the length of the channel $(L)$ and the other one is independent of the length $(L)$, then physically where these resistances are located?

Any one might think the term which is directly proportional to the length (L) is associated with the length of the channel, and the other term $(R_B)$ is independent of the length is associated with the contacts (the source and the drain), or the two interfaces, the overall $(R_B)$ is a series combination of the two interfaces, $(\frac{R_B}{2}+\frac{R_B}{2})$.

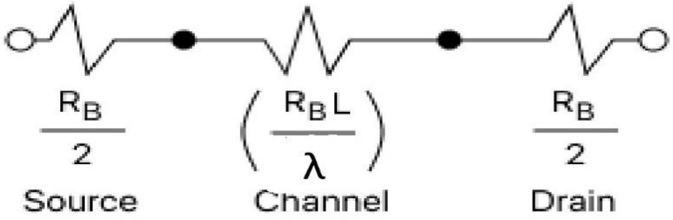


Figure 2.4: Generalization form of resistance given by equation (1.4)

This description is not quite convincing which is going to be justified by an experimental evidence, as the discussion is going along with this dissertation.

Thinking about current flow driven by an applied voltage, usually a resistance is defined given by equation (1.1), the inverse of which is known as the conductance $G$,

$$G = \frac{\sigma A}{L} \qquad (2.5)$$

Where: A is the cross-sectional area of the conductor
$L$ is the length of the conductor

σ is conductivity

The conductivity (σ) is the reciprocal of the resistivity (ρ) given by the Drude formula **[24,];**

$$\rho = \frac{m}{q^2 \tau n} \qquad (2.6)$$

Where; $m$ is the effective mass

$n$ the number of electrons (*the electron density per unit length, area, or volume).*

$\tau$ is the mean free time (for how long an electron travel before it get to hit something)

$\rho$ is the electron charge

So, there would be an expression for the conductivity (σ); the used material parameter, inside which electrons travel by a diffusive process or random walk, given by;

$$\sigma = \frac{q^2 \tau n}{m} \qquad (2.7)$$

This what is called the old prospective which is going to refer to occasionally to connect with the new prospective, which starts from what happens in small devices using the bottom-up approach. Since the resistance in short devices is quantized by ($R_B = \frac{h}{q^2 M}$), see equation (1.3), therefore, the quantum conductance can be written as;

$$G_B = \left(\frac{q^2 M}{h}\right) \qquad (2.8)$$

Where; $q$ is the electron charge

$h$ is Boltzmann's constant

$M$ is the number of modes.

The fundamental constant ($q^2/h$) has the dimensions of conductance (mho) exactly $\left(\frac{1}{25}\right)$Ʊ. This result got established more than twenty five years ago, in mesoscopic physics, viewed as a separate result for small devices, whereas for big devices still people do the traditional way, here people argued that for big devices they could actually start from equation (2.8), and show that the conductance G can be written as the ballistic conductance $G_B$ multiplied by the factor $\left(\frac{\lambda}{L+\lambda}\right)$;

$$\mathrm{G} = \frac{G_B\lambda}{(\,L+\lambda)} \qquad (2.9)$$

Where; L is the length of the device

λ is the mean free path

The mean free path can be defined as the distance an electron travels in the mean free time $(\tau)$ before it hits something. Therefore $(L)$ can be omitted if the device is very short compared to the mean free path $(\lambda)$, and the conductance (G) will be equal to the ballistic conductance $(\mathrm{G}_B)$, in the other hand for a long device $(L)$ is very large compared to $(\lambda)$, so $(\lambda)$ can be ignored in the denominator to have a conductance, $\mathrm{G} = (\frac{G_B\lambda}{L})$, which goes inversely as length, just as expected from Ohm's law for long devices.

Let us now compare equation (2.5) with equation (2.9); for long devices

$$G = \frac{\sigma A}{L} = \mathrm{G_B}\lambda/(\mathrm{L}+\lambda)$$

$$\sigma A = G_B\lambda. \qquad (2.10)$$

This result arrived to, is the new prospective, starting from the bottom end (10nm to 1um) up to the diffusive end, by defining something on the ballistic conductance $(G_B)$ depending on the number of modes $(M)$ which is the material property.

From equation (2.10) above for any real short section of any material experimentally $\left(G_B = \frac{q^2M}{h}\right)$, can be measured. To make it longer, it would have the usual behavior described by the conductivity;

$$\sigma = \frac{G_B\lambda}{A} \qquad (2.11)$$

How to obtain this new prospective shown by equation (2.11) is going to be explained and justified in detail when convenient to understand the new prospective of the current flow in the next chapter.

To see the difference between the new prospective and the old prospective represented by the Drude formula given by equation (2.7).

The Drude formula can be described by the standard prospective of the current flow learned from freshman's physics, due to the electrons driven by an electric field. From Newton's laws the force is given by mass times acceleration;

$$\vec{F} = m\vec{a} = m\left(\frac{dV_d}{dt}\right) = qE$$

Where; $V_d$ is the drift velocity

$q$ is the charge of the electron

$E$ is the electric field

$\vec{F}$ is the force that an electron feels

But this true only in vacuum, in solids there will be a frictional force $\left(m\frac{V_d}{\tau}\right)$, momentum per time, or the compensation for viscous mechanism (a kind of viscous drag) which can be described by a mean free time, therefore;

$$qE - m\left(\frac{V_d}{\tau}\right) = m\left(\frac{\mathrm{dV_d}}{\mathrm{dt}}\right)$$

At steady state nothing is changing with time, therefore, $m\left(\frac{\mathrm{dV_d}}{\mathrm{dt}}\right)$ is equal to zero, then;

$$qE = \frac{m\mathrm{V_d}}{\tau}$$

And the drift velocity is given by;

$$V_d = \left(\frac{q\tau}{m}\right)E$$

The factor $\left(\frac{q\tau}{m}\right)$ is known as the mobility which is the material property always people ask about it for any material. The current depends on the electron density $(n)$ times the velocity;

$$I = Aqn(V_d) = Aqn\left(\frac{q\tau}{m}\right)E = A(q^2n\tau/m)\,E$$

$$J = \frac{I}{A} = \left(\frac{q^2n\tau}{m}\right)\ E = \sigma E \qquad (2.12)$$

The current density $(J)$ depends on the electric field **[27,28]**, which determine the voltage, through the factor $\left(\frac{q^2n\tau}{m}\right)$, which is called the conductivity $(\sigma)$ **[29].** Electrons driven by an electric field causes a very interesting problem, because all the electrons in the solid must feel the electric field, and start to move, but this statement is completely wrong, and no one claims that it is true. All scientists say that only the free electrons in the valance band close to the Fermi-level are the only ones able to move. Electrons deep down in a solid have a lot of states, and up to now only the (*tip of Iceberg)* just shown, because there are many other states well down there in the solid containing most of the electrons, but no one claims that they play any role in current flow. This very well-known and very well established, the argument is the number of electrons used as shown in by equation (2.12), is the number of free electrons $(n)$ only, because all the bands other than the valence band are filled completely, filled bands don’t conduct. But in the new prospective of the current flow, the answer to the question; why filled bands don’t conduct, is very pretty explained in a very clear way, as will be shown in the next paragraph.

To understand this new prospective of current flow the energy levels inside the channel need to be known. For hydrogen atom for example there will be discrete leveled represented by $S, 2S, \ldots etc$, for molecules energy levels get closer, and for solids energy levels get very dense represented by lots of states typically they form a band (an energy range), with certain states followed by an energy gap with no states in it, again there will be an energy band full of states, and so on, energy levels might look like this, (see the figure below);

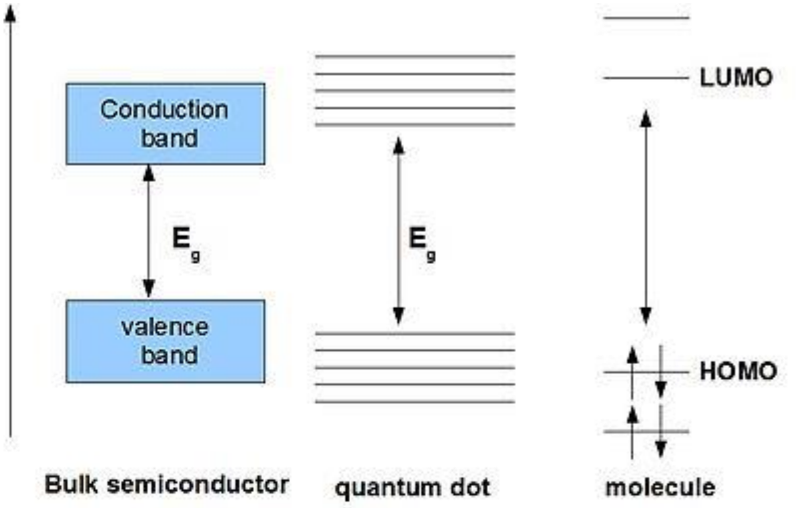


Figure 2.5: Energy levels configuration

For contacts there are many states continuously distributed, the details of which doesn’t matter, because the flow of current takes place in a very big region, while the channel is in the form of a narrow bottle neck to control the flow of current. The contacts can be represented by a lot of States filled up to a certain level called the Fermi-level or the electrochemical potential denoted by $(\mu)$ **[30,31,32,33]**. These two expressions shall be used almost interchangeably to tell up to what point all the levels are filled at low temperature.

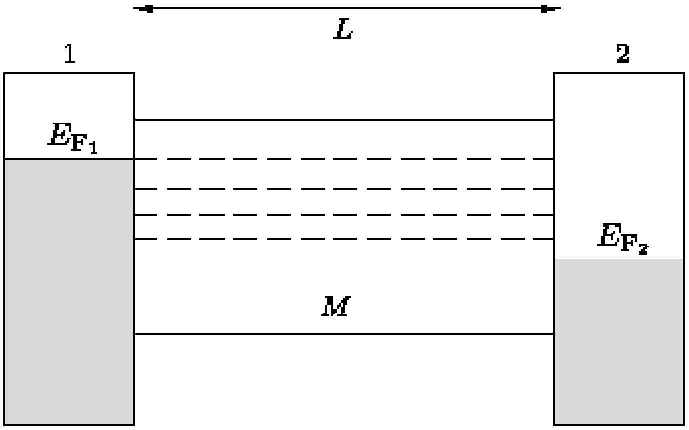


Figure 2.6: Contact states with Energy levels inside the channel

Electrons if they are left free, they tend to occupy the lowest energy levels, but all of them are not able to do that, because of Pauli exclusion principle. When a positive voltage is applied to the drain side, it lowers all the energy levels by an amount $(qV)$

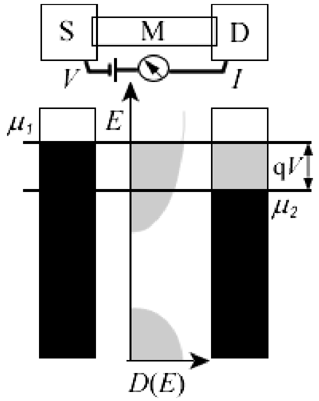


Figure 2.7: Energy diagram

This what is called the energy diagram. Both contacts in this energy diagram try to bring the channel into equilibrium with themselves, and that is why current flows, because one contact tries to fill the channel, and the other contact tries to empty it. Using this viewpoint, an expression for the conductance ($G$) can be obtained from the following relations;

Number of electrons in the channel = (electrons/second) * (time spent)

$$n = \left(\frac{I}{q}\right) t$$

Density of states D(E) = number of levels per unit energy

$$\frac{1}{2} D(E) = (\frac{n}{qV})$$

Then, an expression for the conductance G is given by,

$$G = \frac{q^2 D(E)}{2t} \qquad (2.13)$$

Where; $D(E)$ is the average density of states

$t$ is the transfer time

$q$ is the electron charge

By assuming $D(E)$ is not constant, equation (2.13) can be generalized, and from which current can be estimated. For instance, this will be ignored to be discussed later in another contest, in the next chapter. Equation (2.13) leads naturally to ballistic conductance; $G_B = \left(\frac{q^2 M}{h}\right)$ ; given by equation (13) above, it is possible to use this viewpoint for diffusive transports well to get an expression for the conductivity (σ) as;

$$\sigma = \left(\frac{q^2 D(E) Đ}{AL}\right) \qquad (2.14)$$

Where; $D(E)$ is the density of states close to the Fermi-level energy

Đ is the diffusion coefficient

$A$ is the cross-sectional area

$L$ is the length of the channel

Note that equation (2,14) above does not look like the Drude formula, although this equation is not a new expression, it is well known a long time ago, and referred to as Kupo formula. Everyone agrees that, this is really a much more general expression for the conductivity ($\sigma$), which can be obtained from an advanced formulism like Boltzmann equation, see appendix (I). What is new about this equation the way by which it is found using the concept of an elastic resistor (Landauer resistor) - *after Ralf Landauer- the first of pioneered this way of thinking a long before it become experimentally relevant,* to come so easily and naturally from equation (2.13) with no need to rigorous theory of Boltzmann's formula.

# Chapter Three

# The New modified Ohm's law

## 3.1 Key concepts

In this chapter the resistance or the conductance in the nanoscale devices is described. A couple of concepts need to get across, namely;

(1) The density of states dented by $D(E)$.
(2) The Fermi- Function denoted by $f(E)$.

### 3.1.1 The density of states (DOS)

The density of states denoted by, DOS or $D(E)$ **[34,35,36,37]**, is based upon some experiments, the most common one is by photoemission means discovered hundred years ago by hitting a solid with an ultraviolet light (photons) to knock out an electron to the vacuum out of the solid with sufficient energy, otherwise it stays inside the solid **[38]**. By considering the vacuum as zero reference, that means all energy levels are negative quantities to keep the electrons inside the solid, or they would jump out of the solid (see figure 3.1 below);

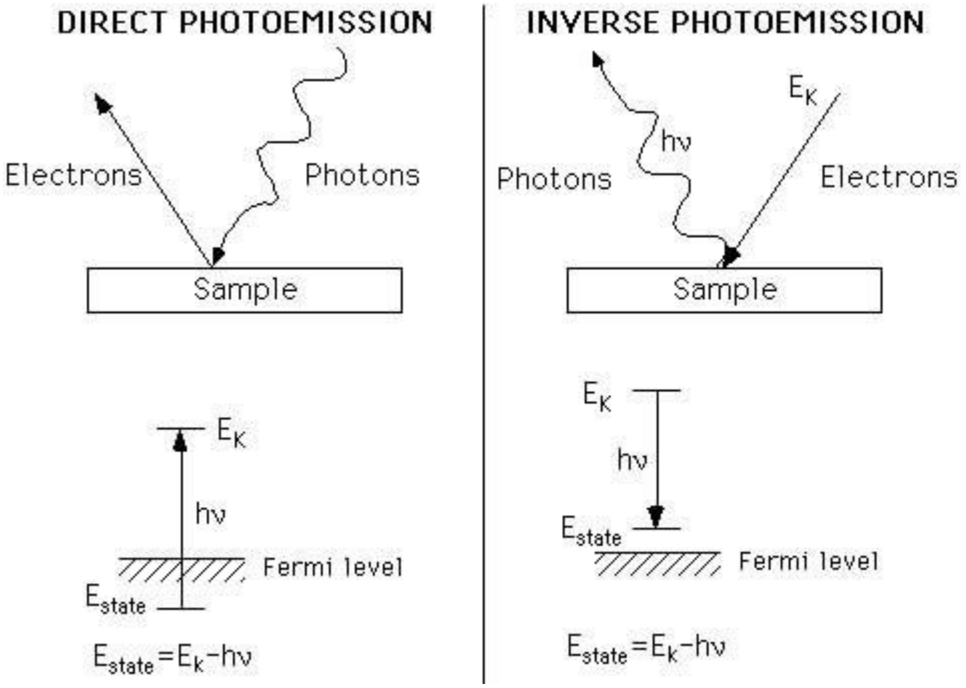


Figure 3.1: Photoemission processes

As far as current flow is concerned, what really matters is the top part, the valence electrons, where the highest energies are located, but still there are a lot more electrons down which need more energy to knock them out. Therefore, inside the solid there are some states filled, and some others are empty, and that is why an electrochemical potential or Fermi-level line ($\mu$) need to be drawn, above which all the states are empty, and below it all the states are completely filled. Empty states cannot be measured with spectroscopy **[38],** because there are no electrons there to be knocked out, hence no current can be calculated. $D(E)$ can be plotted schematically in a more convenient manner by drawing the energy in the vertical axis, and horizontally the density of states $D(E)$, which tells how many states per unit energy are in there. This not like what usually people used to, by drawing the independent variable ($E$) horizontally (see Figure 3.2 below);

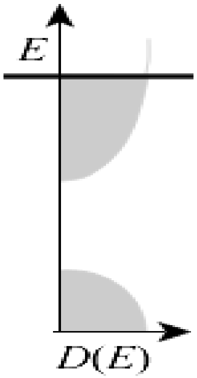


Figure 3.2: Schematic plot of density of states against energy

Energy is measured in International Standard units ($SI$) given in (Joules) or (Colombo volts), but for small nanoscale devices it is much convenient to measure the energy in electron volt ($eV$), which is very smaller than the (Joule) unit;

$$E = 1\ eV = 1.6X10^{-19} joules$$
$$E = 1\ eV = 1.6X10^{-19}\ coulomb\ volt\ (qV)$$

Electrons fill up a lot of states obeying the Pauli- exclusion principle, so the number of filled states *equal* to the number of electrons, because each electron occupy a separate level or state, *this statement may cause a little confusion*, because people always say; isn't like every state can hold two electrons (*an up spin and down spin*), the right way is every state *hold only one electron*, because there two separate states one for the up spin and the other for down spin, so as if states come in pairs, but each state can hold one electron, that is the exclusion principle, therefore there is a level $(\mu)$, which separates all the filled states from the empty ones. The density of states can be defined by the number of levels per unit energy;

$$D(E) = \frac{n}{qV} \qquad (3.1)$$

Where; $D(E)$ is density of states
$q$ is the electron charge
$V$ is the applied voltage
$n$ is the number of levels (the number of electrons).

### 3.1.2 The Fermi-Function $f(E)$

The mathematical form of the Fermi-Function is given by;

$$f(E) = \left(\frac{1}{1 + e^{\frac{E-\mu}{KT}}}\right) \qquad (3.2)$$

Where; $E$ is the energy
$K$ is the Boltzmann’s constant
$T$ is the absolute temperature
$\mu$ is the electrochemical potential

This relation tells the probability of occupation of the energy levels. The concept of the electrochemical potential should be introduced to separate all the filled states from the empty ones, since some of the available levels are filled, others are empty. The function $f(E)$ can be plotted as a function of $(E)$ using the MATLAB or something equivalent. When $(E)$ is very small, just below the Fermi-level $(\mu)$, $\left(\frac{E-\mu}{KT}\right)$ will be negative large number, so the Fermi-Function $f(E)$ is equal to one, the reverse if $(E)$ is above the Fermi-level $(\mu)$, $\left(\frac{E-\mu}{KT}\right)$, will be a large positive number, then $f(E)$ turns to zero, and the Fermi-Function is expected to behave as shown in Figure 3.3, as follows ;

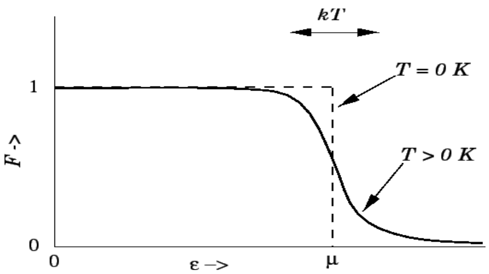


Figure 3.3: The Fermi-Dirac distribution for different temperatures

At zero temperature, the Fermi-Function takes an abrupt change, to resemble a rectangular shape, but at none zero temperature the function is somewhat diffused boundary, in an energy range in the order of $(KT)$, it spreads out with $(KT)$, as shown in the figure above. This the shape of the function normally found in most of the books, but it more convenient to be flipped to the left-hand side and rotated ninety degrees to take the following shape (see Figure 3.4);

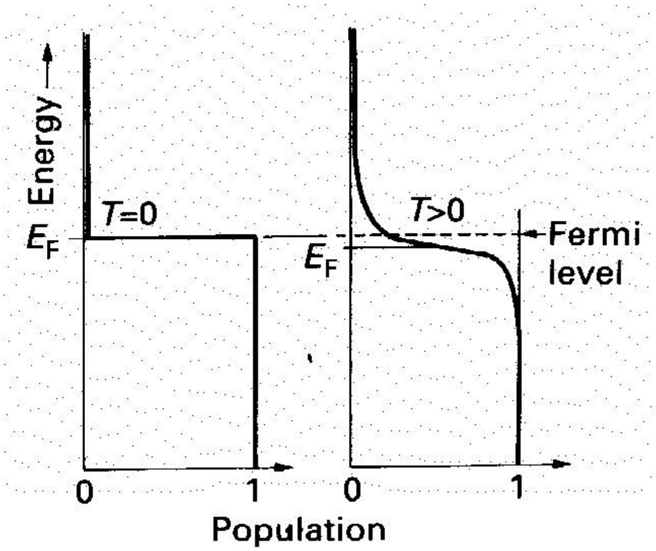


Figure 3.4: Convenient plot of Fermi function population

## 3.2 The Transistor Model

To introduce the Nano-transistor basic model, which is going to be referred to from time to time. The transistor is nothing more than a resistor in the sense that, when a voltage is applied across, current flows using the available energy channels, in small energy window, between the two Fermi-levels of the contacts ($\mu_1$ & $\mu_2$) **[39,40].** Based on this the conductance, $G = (\frac{q^2D}{2t})$ can just be obtained, from which both the ballistic and the diffusive regimes are connected by making use of the idea of time the electron takes from source to drain. From the picture developed in the last section, namely; there are the density of states $D(E)$ in the channel, and two big reservoirs (the contacts), full of electrons to a certain level (the Fermi-levels of the source and the drain). When a voltage is applied between the source and the drain, they separate by an amount of $(qV)$. As a result of current flow through the channel from one terminal to the other giving rise of the current in the external circuit. The model can be described in its simplest form by a conductor called the channel, through which the electrons can flow between the two contacts the source and the drain. Since the transistor has three terminals, the third one (the gate) is to control the resistance of the channel to change it into different orders of magnitudes. The third terminal has not been shown in the following diagram for simplicity, and it is going to be added later when convenient when putting a voltage in that gate.;

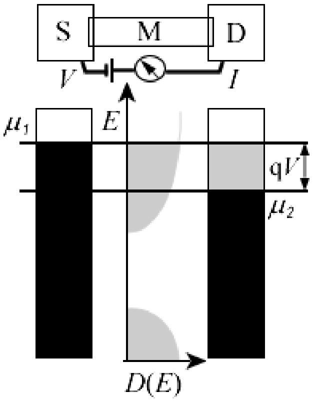


Figure 3.5: Small energy window separation due to an applied voltage between the source and the drain

### 3.3 The Current Flow

There are two important concepts needed for the current flow discussions;

(1) If the whole of the function $f(E)$ is at equilibrium then a common Fermi-level in both contacts, and no current flow.
(2) If a positive voltage is applied across the two terminals (source and drain), the two Fermi-levels $\mu_1$ and $\mu_2$ separates by an amount of energy $(qV)$, hence the current flows.

### 3.3.1 Equilibrium Case

To understand the nature of the current flow in small devices consider the equilibrium case by short circuiting the two contacts, under these conditions no current flows. Some photoemission experiments can be done on this system to draw the energy levels diagram. Since the channel has certain energy levels available to it, in the form of bands and energy gaps, it is more convenient to consider the density of states $D(E)$ which tells how many energy levels per unit energy are there. In the contacts a continuous density of states up to the level $(\mu_0)$ is drawn to which all the states are filled. The level $(\mu_0)$ separates the filled states from the empty ones at zero temperature (0k).

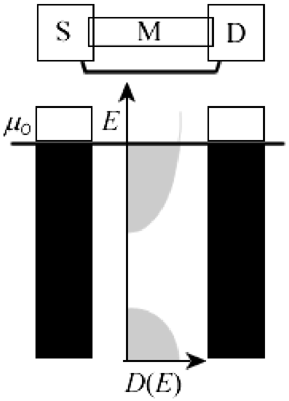


Figure 3.6: Electron distribution at equilibrium state given by Fermi function

At room temperature (300k) some electrons below $(\mu_0)$ will have enough thermal energy (KT) to jump up and occupy the higher states above $(\mu_0)$, the distribution of the electrons at equilibrium will be given by the Fermi-function given by equation (3.2) above, this equation can be normalized to $\frac{(E-\mu)}{KT}$ in the vertical axis, that means energy referred to as $(\mu)$ gives the zero axis at $(\mu_0)$ , the plot of which as shown here below **[41];**

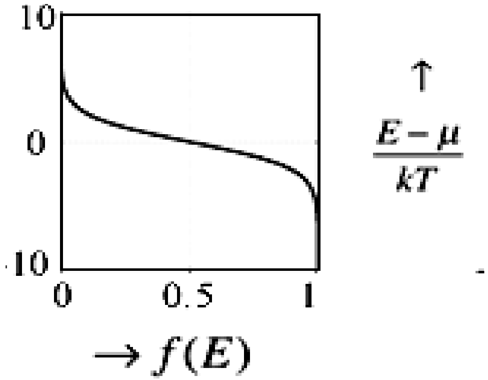


Figure 3.7: Energy function normalization to $\frac{(E-\mu)}{KT}$

For $(E-\mu)$ greater than zero, $\frac{(E-\mu)}{KT}$ is a large positive number, so $f(E)$ equals to zero; for $(E-\mu)$ less than zero, $\frac{(E-\mu)}{KT}$ is a large negative number, so $f(E)$ quals to one, and for $(E=\mu)$, $f(E)$ equals 0.5. In other words, the function $f(E)$ given by equation (3.2) tells what fraction of states is filled. For $(E)$ far below $(\mu_0)$, 100% of states are filled, for $(E)$ far above $(\mu_0)$, 100% of the states are empty, and for $(E=\mu)$ or $(E)$ is around $(\mu_0)$, 50% of states are filled. The thermal energy $(KT)$ is roughly $(25\,meV)$, and it has got the dimensions of energy. In MKS system of units, energy is given by (coulomb volts), but it is quite common to describe the electronic energy in (electron volts) rather than joules (Colombo volt);

$$25\,meV = 25X10^{-3}\ eV = 25\ X10^{-3}X1.6X10^{-19}\ cV\ or\ joules$$

### 3.3.2 Out of Equilibrium Case

In this case there is a current flow through nanodevice, in order to obtain an expression for that current, the nature of the current flow through the device should be discussed. From the current expression the conductance $(G)$ or the resistance $(R)$ can be deduced.

From the picture developed in the last section (3.3.1), namely the density of states $D(E)$ in the channel and the Fermi-level in the two big contacts full of electrons to certain Fermi-level, both contacts have the same Fermi-level $(\mu_0)$, because they are at equilibrium. When a bias is applied across the device, the positive side, the drain, all energies including the electrochemical potential sink down by an amount $(qV)$ with respect to the negative side, the source, and the two contacts are held at two different electrochemical potentials $(\mu_1 and\, \mu_2)$ respectively. This is the basic fundamental to be known before discussing why electrons flow in the first place.

The reason for current flows very simple; the states in the channel are below the Fermi-level of the source contact, therefore the source wants to fill up the channel and bring it in equilibrium with itself all the way up to $(\mu_1)$. If the channel happened to be empty, the source will push an electron to fill the empty space in the channel, on the other hand the drain contact sees an electron above its Fermi-level $(\mu_2)$, it tends to pull it up to empty the channel to bring it in equilibrium with itself, therefore one contact wants to fill up the channel and the other contact want to empty it. No contact wants to give up, and that will go for ever, another electron comes in then goes out. Both contacts have totally different agenda, the source fill up and the drain takes out, then the electron leaves the drain to the battery, and a new electron comes into the source, so the process goes for ever.

The electrons in the lower bands, energy levels below the band gap don't conduct, because all the energy levels therein are below the Fermi-levels of both contacts so that the source contact wants to fill up and so the drain contact as well, that means both contacts have the same agenda, therefore all the states down there stay filled, and more happen the electrons keep pumping into the channel from both sides. As far as these states are concerned, both contacts want to keep them filled, so they just stay filled and no current flows. This the point that causes a lot of argument and confusion, because it is so hard to understand when thinking of electric field driving electrons, as in the freshman's physics. Because in that case electrons down therein should start moving because every electron should feel the electric field and should be moving essentially, but this is not true and widely accepted is that the current flows only in the small energy window around the Fermi-level. States deep down teleplay no role in current conduction, and this what causes the whole confusion in case of electric field, the reasons given in that case is by stating :((*Filled bands don't conduct*.)) without giving any explanation for that, but as far as the ballistic current is concerned, this is not true by making a pun that; *what really drive the electron is the difference in agenda between the two contacts*, which is quite different from what stated in case of electrified drive.

Another important reasoning for how well a material conduct. Copper for example conducts very well with high conductivity, while glass does not, although both has the same number of electrons. Copper conducts well because it's Fermi-level passes through the upper band where conduction takes place, and it has some free electrons for conduction, while in the glass Fermi-level passes through the region where $D(E)$ is almost zero below

the band gap, and no free electrons are available for conduction (see figure 3.8 below);

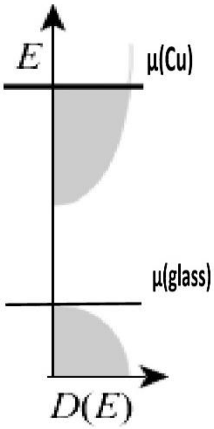


Figure 3.8: How well different materials with the same number of electron differ in conduction

To turn all the above information into a quantitative expression, assume that the density of states almost constant over the entire range of energy window $(qV)$, which is opened by the applied voltage (see Figure 3.9 below);

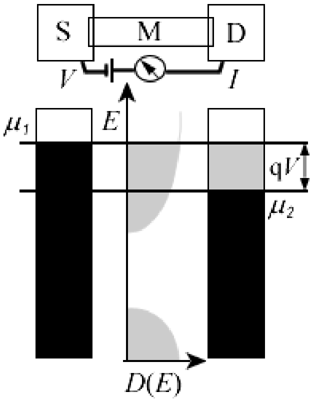


Figure 3.9: Constant density of states over entire (qV) window opened by an applied voltage

By equating the number of electrons in the channel $\left(\frac{DqV}{2}\right)$ to the number of electrons flowing per second multiplied by the time spent inside the channel to get immediately an expression of the conductance $(G)$ given by equation (2.13) above. Why $D(E)$ is divided by two, the argument goes like this, some electrons are flowing from source to drain, and some electrons are flowing from drain to source and all of them are sharing the same channeling both directions.

## 3.4. Current Estimation

In the last paragraph the steady state number of new electrons coming in continuously and leave, are estimated, that number of electrons will be the flow of electrons multiplied by the time, each electron spreads inside the channel, by accepting that the conductance can obtained right away in condition that the density of states is assumed to be constant over the energy window. To be done in a better way by assuming that the density of states is constant in small energy range $(dE)$, in which there is a current and there is a time $t(E)$ spent inside the channel depending on energy. At certain energies the electrons might have fast velocities and zip right through, other energies might take longer time, therefore the number of states can be estimated either by the average density of states within small energy range $(dE)$ given by $(\frac{D(E)dE}{2})$, or by $\left(\frac{I*t(E)}{q}\right)$, by equating these and rearrange current can be written as;

$$I = \frac{qD(E)dE}{2t(E)} \qquad (3.3)$$

This is the current which should be obtained if the states at the source contact $(\mu_1)$ is completely full, and zero states at the drain contact $(\mu_2)$, but the states at both ends may be partially full or partially empty, so in general a difference factor $\Delta f_F = \left(f_1(E) - f_2(E)\right)$, should be introduced

$$I = \frac{qD(E)dE\Delta f_F}{2t(E)}$$

If $\left(f_1(E) = f_2(E)\right)$ implies $\Delta f_F$ equals to zero, putting the system into equilibrium and no current flows. The current is driven only by difference in the Fermi-functions $\Delta f_F$, because one contact wants to fill up the channel and the other contact wants to empty it tending to bring the system into equilibrium, but neither of contacts will give up, so their trial will go alternatively forever, hence current continues to flow. This is the current in single energy level without changing its energy, so the different energy levels are independent to each other, because the elastic resistor idea states that the current is carried by different channels independently. To add different amounts of currents in different energy levels, (highways), can be done by integration

$$I = \int_{-\infty}^{\infty} \frac{qD(E)\Delta f_F dE}{2t(E)}$$

Multiplying by $\left(\frac{q}{q}\right)$

$$I = 1/q \int_{-\infty}^{\infty} \left(\frac{q^2 D(E)}{2t(E)}\right) \Delta f_F dE$$

Where; $\left(\frac{q2D(E)}{2t(E)}\right) = G(E)$

$$I = 1/q \int_{-\infty}^{\infty} G(E) \Delta f_F dE \qquad (3.4)$$

Where; $G(\ E)dE$ is the ballistic conductance function

This is the general form of the current equation, and this is what had been done earlier in a simple version given by equation (2.13) above, where we assumed the conductance is independent of energy in the energy window.

### 3.5 General Conductance Equation

The conductance factor $G(E)$ in equation (3.4) above, given by $\left(\frac{q2D(E)}{2t(E)}\right)$, tells how easily electrons can follow through the channel. This factor depends on the density of states $D(E)$, which tells how many lines in the highway, and $G(E)$ also depends on the time $t(E)$ to tell for how long an electron takes from source to drain. To obtain the general expression for the conductance $G(E)$, the current equation can be divided by the voltage. For low voltages a little approximation can be done to get the required general conductance expression G, which is equal to the conductance function $G(E)$ averaged on energy according to the following relation;

$$G = \frac{I}{V} = \int_{-\infty}^{\infty} G(E) \left(-\frac{\partial f_0}{dE}\right) dE \quad (3.5)$$

Note what enters this relation is the derivative of the Fermi-function with respect to energy E, and this is not obvious at all, but everything will be explained in the following paragraphs.

The derivative of the Fermi-function looks like what is shown in the following figure (3.10);

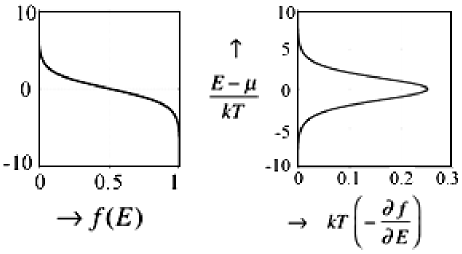


Figure 3.10: Fermi function derivative curve

Right around the zero axis the derivative is the highest where most of the change take place.

$4KT\left(-\frac{\partial f}{dE}\right)$ Is plotted, and the peak value of this derivative is $\left(\frac{1}{4KT}\right)$, away from the electrochemical potential it goes to zero, roughly $(4KT)$, therefore the area under the curve is unity. To get from the current equation (3.4), to the *inverse Ohm's law,* equation (3.5), or the conductance expression is by using *Taylor expansion*. At constant temperature $(T)$, the Fermi-function given by equation (3.2) depends on both (E) and (μ), therefore $\Delta f_F$ in the current equation (3.4) can be written;

$$\Delta f_F = \left( f_1( E) - f_2( E)\right) = (f(E_1\ , \mu_1) - f(\ E_2, \mu_2)$$

For small voltages much less than $(KT)$, the difference between Fermi-levels $(\mu_1 - \mu_2) << K$, i.e. they are very close together, then;

$$\frac{\mathrm{f}_1(\mathrm{E}) - \mathrm{f}_2(\mathrm{E})}{(\mu_1 - \mu_2)} = \left(\frac{\partial \mathrm{f}_0(\mathrm{E})}{\partial \mu}\right);\ \mu = \mu_0$$

$$f_1(E) - f_2(E) = \left(\frac{\partial \mathrm{f}_0(\mathrm{E})}{\partial \mu}\right)(\mu_1 - \mu_2)$$

Where $f_0(E)$ is the average between $f_1(E)$ and $f_2(E)$. Now let;

$$f_0(E) = \frac{1}{1 + e^{\frac{E-\mu_0}{KT}}} = \frac{1}{1 + e^X}$$

Where; $X = \left(\frac{E-\mu_0}{KT}\right)$.

Using the chain rule;

$$\frac{\partial f_0(E)}{\partial \mu_0} = \frac{\partial f_0(E)}{\partial x} * \frac{\partial x}{\partial \mu_0} = \frac{-1}{KT}\left[\frac{\partial f_0(E)}{\partial x}\right]$$

And,

$$\frac{\partial f_0(E)}{\partial E} = \frac{\partial f_0(E)}{\partial x} * \frac{\partial x}{\partial E} = \frac{1}{KT}\left[\frac{\partial f_0(E)}{\partial x}\right]$$

So $\left[\frac{\partial f_0(E)}{\partial E}\right]$ and $\left[\frac{\partial f_0(E)}{\partial \mu}\right]$ are the negative of each other, and any one of them can be put instead of the other, therefore;

$$f_1(E) - f_2(E) = \left(\frac{-\partial \mathrm{f}_0(\mathrm{E})}{\partial \mathrm{E}}\right)(\mu_1 - \mu_2)$$

The difference between the two electrochemical potentials $(\mu_1 - \mu_2) = qV$ , then

$$f_1(E) - f_2(E) = \left(\frac{-\partial \mathrm{f}_0(\mathrm{E})}{\partial \mathrm{E}}\right)qV$$

The general current equation (23), can be rewritten as;

$$I = 1/q \int_{-\infty}^{\infty} G(E)\big(f_1(E) - f_2(E)\big)dE$$

$$I = 1/q \int_{-\infty}^{\infty} G(E)\left(\frac{-\partial \mathrm{f}_0(\mathrm{E})}{\partial \mathrm{E}}\right)(\mu_1 - \mu_2)dE$$

$$I = 1/q \int_{-\infty}^{\infty} G(E)\left[\frac{-\partial f_0(E)}{\partial E}\right]qVdE$$

Since V is a constant value and independent of the energy can be pulled out

$$I = V/q \int_{-\infty}^{\infty} G(E) \left[\frac{-\partial f_0(E)}{\partial E}\right] q dE$$

$$G = I/V = \int_{-\infty}^{\infty} G(E) \left[\frac{-\partial f_0(E)}{\partial E}\right] dE$$

This is the same conductance equation given by equation (3.5) above. The conductance-function $G(E)$ is multiplied by $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ , but $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ has its peak value right around the Fermi-level (E=μ) and then dies out. In fact the conductance-function $G(E)$ is very strongly peaked as the temperature $(T)$ approaches zero, because $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ is infinitely tall and rather thin, exactly representing an impulse or a $(\delta)$ function, therefore $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ can be recognized as a (δ-function) as the temperature tends to be zero, and the average of which is equal to unity;

$$\int_{-\infty}^{\infty} \left(-\frac{\partial f_0(E)}{\partial E}\right) dE = 1$$

When the conductance function $G(E)$ is multiplied by $(\delta)$ it will be taken to a particular energy at zero temperature, but the general value of the conductance $(G)$ in the other side of its equation is at nonzero temperature, so the conductance function $G(E)$ should be averaged over the entire range of energy by integration, to get the inverse of Ohm's law; given by equation (3.5)

### 3.6 Ballistic Conductance

In the section (section 5) above, the small device is considered. The current equation (3,4) is written down first, then linearized to get an expression for the general conductance (G) equation (3.5), by averaging the conductance function $G(E)$ at any given energy. Recall equation (3,5) brought here below;

$$G = I/V = \int_{-\infty}^{\infty} G(E) \left[\frac{-\partial f_0(E)}{\partial E}\right] dE$$

If $G(E)$ is taken out of the relation the integral, $\int_{-\infty}^{\infty}(\frac{-\partial f_0(E)}{\partial E})dE$ , will average to one in either one of the following two cases;

(1) *At low bias, low temperature, and the Fermi- levels are so close together,* $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ is infinitely tall and thin to be recognized as a (δ) function to take the value of $G(E)$ to the Fermi- level, in this case G can be looked at by the value of the conductance function $G(E)$, where; $G(\ E) = \frac{q^2 D}{2t}$ , so it is more convenient to work with $G(E)$ when looking for the conductance without writing the integral explicitly every time when using this relation.

(2) *If the temperature is relatively high* the conductance $G(E)$ is varying strongly with energy, therefore an average will be involved by doing the integral, where;

$$G(\ E) = \frac{q^2 D(E)}{2t(\ E)}$$

And the area under the curve $\left[\frac{-\partial f_0(E)}{\partial E}\right]$ is equal to unity.

Philosophically two different things are discussed, a value about certain energy, and another value about the integral number of electrons. These two types of values will be connected in another context later when the relationship between the Drude formula, and the general relation of conductivity, given by equation (2.14), are discussed to show under what conditions they are equal.

For large devices usually the density of states $(D)$ is a constant value for particular material measured per unit energy per unit volume, where as in $G(E)$ equation (2.13) the density of states $(D)$ has been defined as density of states per unit energy only and not per unit volume, so the density of states will be directly proportional to the volume $(AL)$, plugging in the density of states of any material in $G(E)$ equation (2.13), where the time $(t)$ in that equation depends on the length of the channel $(L)$ and the average electron velocity $(\tilde{V})$, therefore the ballistic conductance $(G_B)$ depends on the dimensions of the channels follows:

$$G_B = \frac{q^2\, D}{2t} = \frac{q^2 D\tilde{V}}{2L} \sim \frac{AL}{L} \qquad (3.6)$$

Then $(G_B \sim A\ )$, this relation was observed 1969, often called Shaven resistor **[42],** which shows the ballistic conductance $G_B$ would be linearly proportional to the area $(A)$ in three dimensions, or the width $(W)$ in two dimensions.

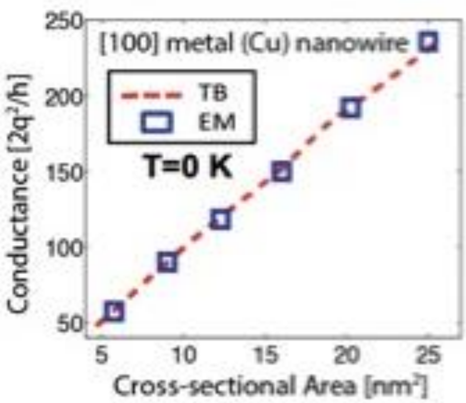


Figure 3.11: Sharven resistance

In semi classical view the electron is looked to as a particle, then the ballistic conductance $G_B$ would be $(D\tilde{V})$ to get this linear increase.

Twenty years later 1988 some experiments are carried on small ballistic conductors with a relatively small areas, instead of going linearly with area, started to go in steps to show a quantized version of conductance due to the wave nature of the electron.

To understand this version, take one dimensional wire then confine it to a very small area to show that $(2t/D = h)$, then plugging this number into $G(E)$ to get the quantum conductance $\left(G_B(E) = \frac{q^2 D}{2t}\right) = \left(\frac{q^2}{h}\right)$ ;

$$G_B = \left(\frac{q^2}{h}\right) \qquad (3.7)$$

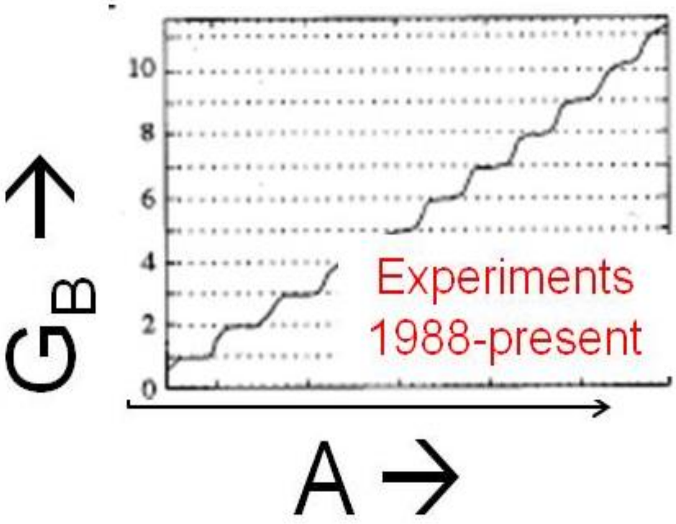


Figure 3.12: Quantum conductance

The value of $\left(\frac{q^2}{h}\right)$ is exactly the inverse of (25KΩ), this can be explained by saying, the particle (energy levels) are not sharp, but they experience broadening out somewhat, and the fundamental unit of that broadening is given by the uncertainty principle. If an electron stays in one level for a length of time $\Delta t$, then;

$$\Delta E.\Delta t > h \qquad (3.8)$$

Therefore, the minimum broadening is $(\Delta E = \frac{h}{\Delta t})$, if the level spread out over ΔE, the density of states is given by the inversion ΔE;

$$D = \frac{1}{\Delta E} = \frac{\Delta \mathrm{t}}{\mathrm{h}} \Rrightarrow \mathrm{h} = \frac{\Delta \mathrm{t}}{\mathrm{D}} \qquad (3.9)$$

Figure 3.13: Energy level spread out; minimum broadening is $\Delta x$

But in general, $\frac{\Delta t}{D} \sim h$ , for this proportionality there could not be a numerical factor of this argument directly, but needs to go deeper to see why this has a fundamental limit there, with this the ballistic conductance can be rationalized as $G_B = \left(\frac{q^2}{h}\right)$ for one dimensional conductor, the real conductor can be viewed as a lot of one dimensional conductors in parallel, which could be called the conducting channels. If there are $(M)$ of them, the conductance would be, $G_B = \left(\frac{q^2 M}{h}\right)$. This what experiments of 1988 show. There is a fundamental quantum unit $\left(\frac{q^2}{h}\right)$ times an integer $(M)$ which is the number of modes. This number of modes tells how many wavelengths of the electron fit into the cross-section, and that is why ballistic conductance is given by $(M)$ times $\left(\frac{q^2}{h}\right)$ to fit this, and that is the reason which make people to think of the actual conductor as a lot of one-dimensional channels in parallel.

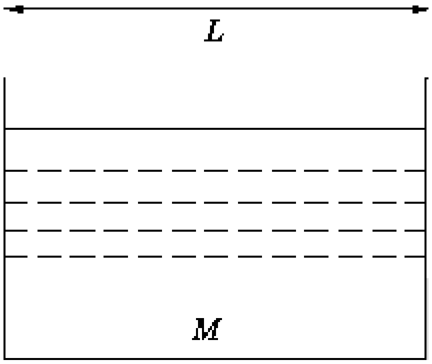


Figure 3.14: Channel as a collection of (M) one dimensional channels

In the figure above the conductance should not be normalized to $\left(\frac{q^2}{h}\right)$, but to $2\left(\frac{q^2}{h}\right)$ , because energy levels in normal materials come into degenerate pairs, and experimentally that $(M)$ is an even integer usually 2,4,6,8,..etc, these even integers has something to do with spin (up & down spin), normally in nonmagnetic materials spin degenerate even in one dimensional conductor can be considered. To have two spins exactly like two things in parallel and the conductance as big as $2\left(\frac{q^2}{h}M\right)$ for $(M)$ channels as shown a lot in all kinds of systems.

## 3.7 Diffusive conductance

When an electron flows in a diffusive manner through any conductor, it doesn't go in in straight line like a bullet as in ballistic transport, but it takes a random manner (random walk).

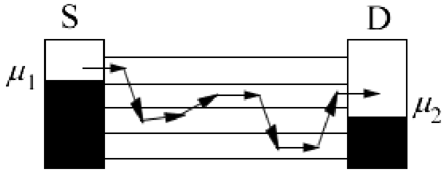


Figure 3.15: Diffusive Transport of an electron

The time it takes to cross the channel is directly proportional to the length of the channel squared and can be written as;

$$t = \frac{L^2}{2Đ} \qquad (3.10)$$

Where; $L$ is the length of the channel

Đ is diffusive coefficient, and ($Đ = V^2$ τ),

where τ is the mean free time $(mft)$ taken by the electron before it changes its path, V is the velocity of the electron to travel a distance $(V\tau)$ called the mean free path $(mfp)$.

Since the time $(t)$ goes directly with $(L^2)$, this where Ohm's law is expected to be seen, and the conductance will not be directly proportional to the area $(A)$ but directly proportional to $\left(\frac{A}{L}\right)$ as an inverse to Ohm's law to go inversely with $(L)$.

The time will be a lot longer and directly proportional to $(L^2)$ rather than $(L)$ compared to the ballistic transport. To combine the time given by equation (3.10) with the general conductance expression;

$$G = \frac{q^2 D}{2t} = \frac{q^2 DĐ}{2L^2} \qquad (3.11)$$

Multiply by $\left(\frac{A}{A}\right)$ and, rearrange and equate with equation (2.5)

$$G = \left(\frac{q^2 D Đ}{A\,L}\right)\left(\frac{A}{L}\right) = \sigma\,\left(\frac{A}{L}\right)$$

The conductivity σ as given by equation (2.14) above, this conductivity with a little discussion could also be written in the form of modes per unit area multiplied by the mean free path $(\lambda)$ as shown by equation (2.11);

$$A = \left(\frac{q^2}{h}\right)\left(\frac{M}{A}\right)\lambda = \left(\frac{G_B \lambda}{A}\right)$$

The form usually seen in literature is as in equation (2.14), namely,

$$\sigma = \left(\frac{q^2 D(E) Đ}{AL}\right)$$

This form is not a new result, it is very well known, but most of the people don’t remember it as they remember the Drude formula, because the usual derivation of it requires Boltzmann’s equation which needs a lot of time to master, and it needs a lot more back ground of knowledge. What is new is the way how to get so easily by using a completely different new prospective. The conductivity equation (2.14) depends on the density of states $(D)$ and the diffusion coefficient (Đ), this equation is really more correct, because it generally more applicable compared to the Drude formula which sometimes makes trouble in many contexts, to explain this recall the Drude formula, equation (2.7), namely;

$$\sigma = \frac{q^2 \tau n}{m}$$

In this equation $(n)$ is the number of free electrons, τ is the mean free time and there is a mass $(m)$ which causes some problems, because it is not the mass of the electron in free space, but it the mass of electron in a solid, and usually its energy -momentum relation is parabolic, and in this case is very well defined mass. But in Graphene for example the energy is directly proportional to momentum $(E = \mathrm{V}_0 P)$, it takes a linear form and looks like a relativistic thing, and that is why it is not clear what to do for mass exactly, but it possible to account for that so easily and at the same time so difficult to confirm or explain. This issue with the problem of the number of free electrons $(n)$ is going to be discussed when Graphene material is studied later.

### 3.8 Ballistic and Diffusive time Justification

Before proceeding to connect the ballistic transport to the diffusive one, it is very important to justify the time used for either of them **[43,44,45]**. One way to think about this justification by the usage of electronic version expression introduced already in chapter (2) to get equation (2.13), as the number of electrons $(Q)$ stored in the channel is related to the flux $(I)$ of electron per second multiplied by the time $(t)$ each electron spends inside the channel;

$$Q = I * t.$$

On that basis the number of electrons stored in the channel for ballistic transports $Q = \left(\frac{D}{2}\right)$, the same for diffusive transport, as shown in the following figures of electron distributions in both cases,

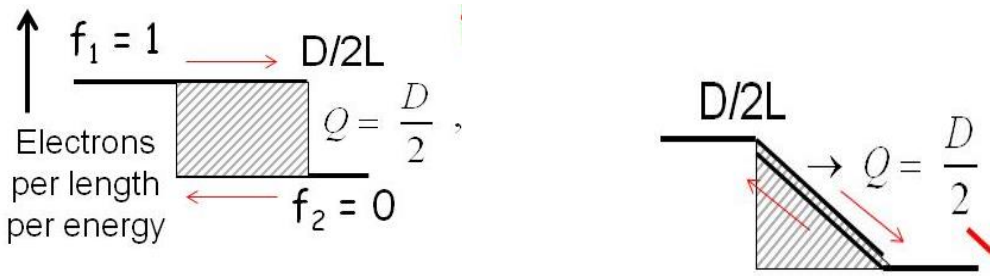


Figure 3.16: Electron distribution for ballistic and diffusive transports

For the ballistic transport all the states carrying the current flow from source to drain occupy half the states in the channel $\left(\frac{D}{2}\right)$, and all the states are fully occupied while the other half in the reverse from drain to source are completely empty.

For the diffusive transport, half the states from source to drain and from the drain to source are equally filled, but the occupation goes down gradually as the electrons flow from source to drain, because the electrons entered the channel from the source are a lot of $\left(\frac{D}{2}\right)$ and some of them hit something to turn around and go back to the source, and very few remain the drain are about $\left(\frac{D}{4}\right)$. The return path starts to build up gradually till the source to drain path is equally the same as the return path from drain to source, $\left(\frac{D}{4}\right)$ both ways.

For the ballistic transport the number of electrons entering the channel from the source are $\left(\frac{D}{2L}\right)$ given in the density of states per unit length and divided by two because half the states are available for them to flow from source to drain, to get the fatal number of electrons stored in the channel is by multiplying by $(L)$, therefore ;

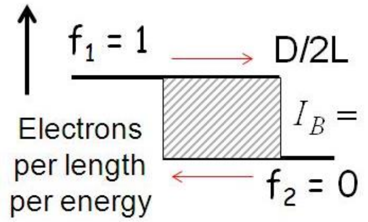


Figure 3.17: Stored number of electrons in the channel for ballistic transport

$$\mathrm{Q} = \left(\frac{D}{2L}\right) * L = \left(\frac{D}{2}\right) \qquad (3.12)$$

The time for the ballistic electron to cross the channel is given by;

$$\mathrm{t_B} = \frac{L}{\tilde{\mathrm{V}}} \qquad (3.13)$$

Where; $(\tilde{\mathrm{V}})$ is the average velocity of the electron

$(L)$ is the channel length.

Therefore, current can be calculated as;

$$\mathrm{I_B} = \frac{Q}{t_B} = \frac{\left(\frac{D}{2}\right)}{\frac{L}{\tilde{\mathrm{V}}}} = \frac{D\tilde{\mathrm{V}}}{2L} \qquad (3.14)$$

For the diffusive transport it is not so easy to write the time, the process will be reversed by trying to write the current first using the standard diffusive equation to deduce the time from it.

The current in diffusive is given by the diffusive equation;

$$I = -\text{Đ}\left(\frac{dn}{dz}\right) \qquad (3.15)$$

Where, (Đ) is the diffusive coefficient

$\left(\frac{dn}{dz}\right)$ is the derivative of the electron density.

The minus sign shows that the going direction in which the function increases most rapidly, it decreases the most rapidly.

The total number of electron density per unit length (i.e. the number of levels) at the source end equals $\left(\frac{D}{L}\right)$, where $\left(\frac{D}{2L}\right)$ going from source to drain, and $\left(\frac{D}{2L}\right)$ in the reverse from drain to source. (see figure 3.18, below)

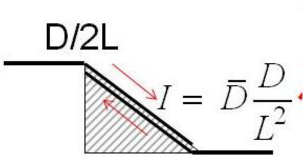


Figure 3.18: Total number of diffusive electrons density per unit length in two directions

The slope of the straight line is given by,

$$\left(\frac{dn}{dz}\right) = \left(\frac{D}{L^2}\right)$$

Then,

$$I = \left(-\frac{\text{Đ}D}{L^2}\right)$$

Initially $\left(\frac{D}{2L}\right)$ electrons are waiting to get in the highway from source to drain while there are no electrons in the reverse way back from drain to source.

Therefore, the total number of electrons stored in the channel will be;

$$Q = \left(\frac{D}{2}\right)$$

Combine $I\ with\ Q$; to get the time,

$$t = \left(\frac{Q}{I}\right) = \left(\frac{L^2}{2\text{Đ}}\right) \qquad (3.16)$$

### 3.9 Connection of the ballistic and diffusive transports

To connect the ballistic regime to the diffusive regime, the general conductance expression $G = (q^2 D/2t)$ is used, and by making use of times given in equation (3.13) for the ballistic transport, and equation (3.16) for the effusive transport. The main difference between the two regimes simply in the ballistic limit, time goes up as $(L)$ whereas in the diffusive limit, time goes up as $(L)$. To get the general answer for both limits and in between time can be written as;

$$t = \left(\frac{L}{\tilde{\mathrm{V}}}\right) + \left(\frac{L^2}{2D}\right) \qquad (3.17)$$

By this equation all the physics can be captured, but it not quite obvious to add them, one way to justify this, the equation looks as if it is written as a polynomial expression in the form of $(L_1, L^2)$. In principle it could have been written $L^3$ ... etc. But according to the ballistic theory, when the length of the device is too short, the time goes with length $(L)$, and the first term with the lower power dominates. When the length of the device is slightly longer, the time goes with $(L^2)$ according to diffusive theory, hence the second term with the higher power still dominates.

For very long device it is very well known that the answer stops at the diffusive limit and no need for higher powers $L^3, L^4$ ...etc. If this is accepted, the time reads as:

$$t = \left(\frac{L}{\tilde{\mathrm{V}}}\right) + \left(\frac{L^2}{2D}\right) = \left(\frac{L}{\tilde{\mathrm{V}}}\right) + \left(\frac{L^2}{2\text{Đ}}\right) * \left(\frac{\tilde{\mathrm{V}}}{\tilde{\mathrm{V}}}\right) = \left(\frac{L}{\tilde{\mathrm{V}}}\right)\left(1 + \frac{L\tilde{\mathrm{V}}}{2\text{Đ}}\right)$$

Let $\left(\frac{L\tilde{V}}{2\text{Đ}}\right) = 1/\lambda$, where $\lambda = \left(\frac{2\text{Đ}}{\tilde{V}}\right)$, which can be defined as the ratio of the diffusive coefficient **[46]** to the average velocity, and at the same time note that $(\lambda)$ is the mean free path, since;

$$\left(\frac{2\text{Đ}}{\tilde{V}}\right) = \frac{2\tilde{V}^2\tau}{\tilde{V}} = 2\tilde{V}\tau$$

Rearrange

$$t = \left(\frac{L}{\tilde{V}}\right) + \left(\frac{L^2}{2D}\right) = \left(\frac{L}{\tilde{V}}\right)\left(1 + \frac{L\tilde{V}}{2\text{Đ}}\right)$$

$$t = \left(\frac{L}{\tilde{V}}\right) + \left(\frac{L^2}{2D}\right) = t_B\left(1 + \frac{L}{\lambda}\right) \qquad (3.18)$$

This (time equation) describes three situations;

(1) All ballistic case
(2) All diffusive case
(3) Semi-ballistic (intermediate between the two)

To generalize the three cases, define a general fraction (T) for the electrons that succeeded to pass from source to drain. This $(T)$ is sometimes known as (the transmission coefficient). This fraction equal one for ballistic transport, 0.5 for diffusive transport, and can take any other value for semi-ballistic case, see the figure 3.19, below;

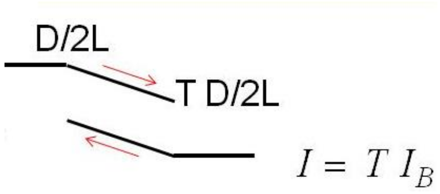


Figure 3.19: Transmission coefficient T

Note that the number of electrons stored inside the channel is still the same for ballistic, diffusive, and probably for semi-ballistic, as $Q = \left(\frac{D}{2}\right)$ in the three cases, because if the number of electrons passing from source to drain are added to the number of electrons in the return path would get exactly the same number in all of the three cases. But the current differs from one case to the other, because the times in all three cases are different. The current has the highest value according to the shortest time (ballistic), and relatively low current for longer times and in between, then what ever the current may be, the ballistic current is,

$$I = I_B T \Rrightarrow \mathrm{Q} = (\mathrm{I_B})\mathrm{t}$$

Where, $T$ is the transmission coefficient

$I_B$ is the ballistic current

t is the average time an electron spends inside the channel

Therefore

$$t = \left(\frac{Q}{I_B}\right).\left(\frac{1}{T}\right) = \left(\frac{\mathrm{t_B}}{\mathrm{T}}\right) \qquad (3.19)$$

Equate with equation (3.18)

$$\left(\frac{t_B}{T}\right) = t_B\left(1 + \frac{L}{\lambda}\right)$$

Then,

$$T = \frac{\lambda}{(L + \lambda)} \qquad (3.20)$$

The transmission coefficient $T$ can be used almost as a definition of the mean free path, in the sense that the length of the device is equal to the mean free path if 50% of electrons get through the device from source to drain, because if the length of the device is one free path ($\lambda$), from equation (3.20), $T$ is equal to 0.5, that means 50% of the electrons get through the device from one end to the other. If the length of the device (L) is very short compared to the mean free path , $L$ can be dropped in equation (3.20), and $T$ will be equal to (one), that means 100% of the electrons get through the device.

The conventional definition of the mean free path is not exactly given by equation (3.20), it has been chosen as a *definition of the mean free path measured* *with the length $L$,* because the expression (3.20) comes out nicely to comply with the previous discussions. It is possible to choose a different definition in which the length of the device equals the mean free path for $(T = 30\%)$ or (60%) of the electrons to path through the device from one end to the other, the last case might have got,

$$\mathrm{T} = \frac{m\lambda}{L + m\lambda}$$

$$\mathrm{T} = \frac{m\lambda}{\lambda + m\lambda} = 0.6 \Rrightarrow T = \frac{1.5\lambda}{\lambda(1 + 1.5)}$$

How to define the mean free path is somewhat can be different from one definition to the other, but the exact definition will come later after the generalization of ohm's law. For instance, the fraction $T$ (the transmission coefficient) can be written as in equation (3.20).

To combine $(T)$ with $(\mathrm{t})$ given by equation (3.19),

$$t = \frac{t_B}{T} = t_B(1 + \frac{L}{\lambda})$$

This is the same expression given earlier by equation (1.4), chapter (1), where the time in general can be written as in equation (3.18), namely found here below,

$$t = \left(\frac{L}{\tilde{\mathrm{V}}}\right) + \left(\frac{L^2}{2D}\right) = t_B\left(1 + \frac{L}{\lambda}\right)$$

When $L$ is small compared to the mean free path, then equation (3.18) gives the ballistic time $(t_B)$, when it is large, gives the diffusive time. If the ballistic time $(\mathrm{t_B} = \frac{L}{\tilde{\mathrm{V}}})$ is applied, directly to equation (3.13), namely found here below;

$$G = \frac{q^2 D(E)}{2t}$$

The ballistic conductance is given as,

$$G_B = \frac{q^2 D\tilde{V}}{2L}$$

At the same time if the diffusive time $t = t_B(1 + \frac{L}{\lambda})$ is applied on equation (3.13), the conductance equation, it gives;

$$G = \frac{G_B}{(1 + \frac{L}{\lambda})} = \frac{G_B \lambda}{(L + \lambda)}$$

That means for a little bigger device the conductance will be less by the factor $(1 + \frac{L}{\lambda})$ . The conductivity $(\sigma)$ and the ballistic conductance $(G_B)$ can be related in the following form;

$$G = \frac{G_B \lambda}{(L + \lambda)} = \frac{\sigma A}{(L + \lambda)}$$

Where, the conductivity can be given by $\sigma = (\frac{G_B \lambda}{A})$, when the length gets really short, so neither the resistance goes to zero nor the conductivity goes to infinity, but they tend to go to some constant value. For a long conductor $(\lambda)$ can be dropped, hence $G = (\frac{\sigma A}{L})$ just like Ohm's law learned in the freshman's physics.

The usual understanding of conductivity starts from long conductors while in the proposed new prospective any material when made very short its property is characterized by;

$$\sigma A = G_B \lambda$$

Which can be used for long conductors as well, because the ballistic conductance when multiplied by the mean free path $(\lambda)$ gives the general conductance multiplied by area.

### 3.10 The Generalization of Ohm's Law

Ohm's law states that the resistance is changing linearly with length. For a very short conductor Ohm's law has got an additional term to take the form;

$$R = \frac{\rho}{x}(L + mfp)$$

Where, $(mfp)$ is the mean free path.

$x = (1, w, A)$ for one-dimensional conductor, two-dimensional conductor or, three-dimensional conductor respectively

$\rho$ is the resistivity

This form of resistance seems to be very useful when a conductor goes to the ballistic limit with length almost equal to zero, the resistance does not go to zero, but goes to another constant value. This form of resistance is slightly misleading, because $(mfp)$ does not mean much by itself so the resistivity $(\rho)$, but their product has got some meaning, what does it means is basically;

$$G_B = \frac{q^2 D\tilde{V}}{2L} = \frac{x}{\rho . mfp}$$

Where, $x = (1, w, A)$, for 1D, 2D or3D respectively. When writing $(\rho . mfp)$ as a function of $G_B$ in its quantum form without any extra factors is slightly misleading, because it gives an impression that it is independent of any mean free time $(\tau)$ or scattering at all. To understand this, recall the resistance equation;

$$R = \frac{2t}{q^2 D} = \frac{1}{G}$$

Where $(t)$ is the time taken by the electron to get from source to drain through the channel, this time can be generalized to cover the ballistic, the diffusive, and in between transports, by adding the diffusive component which is directly proportional to $(L^2)$ to the ballistic component which is directly proportional to $(L)$. Since $D$ is directly proportional to $(L)$, then ballistic component is not representing Ohm's law while the second diffusive component is exactly an Ohm's law.

$$\mathrm{t} = \frac{L}{\tilde{V}} + \frac{L^2}{2Đ}$$

It is not immediately clear to add them but this has been justified before by saying the time is a function of the length $(L)$, and it is written in the form of

Taylor series; comes first ballistic linear term followed by the diffusive quadratic term, and no further terms will be added to it, this also has been justified by the diffusive transport theory. Substituting $(t)$ in the resistance equation to get;

$$R = \frac{2t}{q^2 D} = \frac{2}{q^2 D}\left(\frac{L}{\tilde{V}} + \frac{L^2}{2\text{Đ}}\right)$$

$$R = \frac{2t}{q^2 D} = \frac{2}{q^2 (\frac{D}{L})}\left(\frac{1}{\tilde{V}} + \frac{\mathrm{L}}{2\text{Đ}}\right)$$

Where, $\left(\frac{D}{L}\right)$ is the density of states per unit length for one dimensional conductor, because D is directly proportional to $(L)$ when dividing $(D)$ by $(L)$ makes $(D)$ independent of the length. To generalize the resistance for different dimensions of conductors in one shot, multiply by $\left(\frac{X}{x}\right)$ , where $x = (1, w, A)$ for $(1D, 2D, 3D)$ respectively.

$$R = \frac{2t}{q^2 D} = \frac{2}{xq^2\left(\frac{D}{xL}\right)}\left(\frac{1}{\tilde{V}} + \frac{\mathrm{L}}{2\text{Đ}}\right)$$

The ratio $\left(\frac{D}{xL}\right)$ is the material property, this quantity would not change, because if we change the dimensions of the material the density of states $(D)$ will also change with the same proportion, so the ratio $\left(\frac{D}{xL}\right)$ remains constant. Therefore $(D)$ is given per unit energy per unit (length, area, or volume) when $(x = 1)$ for one dimension, or $(x = w)$ for two dimensions, or $(x = A)$ for three dimensions.

Pull 2Đ out of the bracket in the resistance equation above

$$R = \frac{2t}{q^2 D} = \frac{2}{2xq^2\left(\frac{D}{xL}\right)\text{Đ}}\left(\frac{2\text{Đ}}{\tilde{V}} + \mathrm{L}\right)$$

$$\sigma = \left(\frac{q^2 D\text{Đ}}{xL}\right)$$

$$R = \frac{2t}{q^2 D} = \frac{1}{x\sigma}\left(\frac{2\text{Đ}}{\tilde{\text{V}}} + \text{L}\right)$$

$$R = \frac{2t}{q^2 D} = \frac{\rho}{x}\left(\frac{2\text{Đ}}{\tilde{\text{V}}} + \text{L}\right)$$

The ratio of the diffusion coefficient (Đ) to the average velocity ($\tilde{\text{V}}$) looks like the mean free path since ($\text{Đ} = \tilde{\text{V}}^2\tau$).

In last **section 3.8** under the title of time justification, the parameter ($\lambda$) is defined as $\left(\frac{2\text{Đ}}{\tilde{\text{V}}}\right)$ , starting from the conductance expression ($G = \frac{q^2 D}{2t}$), then this expression had been applied to the ballistic limit to give ($G_B = \frac{q^2 D\tilde{\text{V}}}{2L}$ ), then to the diffusive limit to give ($G = \frac{q^2 D\text{Đ}}{L^2}$), when connecting them together, the fundamental parameters ($G_B$) and ($\sigma$) are written as,

$$\left(G_B = \frac{q^2 D\tilde{\text{V}}}{2L}\right)$$

And,

$$\left(\sigma = \frac{q^2 D\text{Đ}}{Lx}\right)$$

hence;

$$\sigma x = G_B\lambda$$

The bar on the diffusive coefficient (Đ) and the bar over the velocity ($\tilde{\text{V}}$) indicate that an average is involved therein, (see appendix II). The factor (2) appeared in the relation $\left(\lambda = \frac{2\text{Đ}}{\tilde{\text{V}}}\right)$ is due to the backscattering. Assuming that the process is isotropic, scattering take place either in the positive direction or the negative direction in one-dimensional conductor, but our real interest on those which have been completely turned a round (backscatter) into the right direction from source to drain, and the time getting turned a rounds ($2\tau$), where one ($\tau$) is before hitting anything and the second one after hitting any object to turn around in the opposite direction, therefore ($\lambda = 2\tau\tilde{\text{V}}$) by considering the longer time taken by the mean free path for backscattering. The exact scattering factor is equals (2) for one dimensional conductor. In two-dimensional conductor the scattering take place in a form of a circle in all directions, averaging any single velocity over an energy range of few $(KT)$ around the electrochemical potential, turned around in the right direction, the factor will come out to be $\left(\frac{2}{\pi}\right)$. For three dimensional conductors scattering tale place in rather complicated mathematical form, and it will be treated as a polar coordinate with two angles to project the velocity vector at an angle $(\theta)$ with respect to the (Z-direction) on the (XY) plane to see the azimuth angle $(\Phi)$ which defines a particular direction in the three dimensions, the factor in this case will be (1/2). To collect all these factors together the average value of the velocity can be written as;

$$\tilde{V} = V\left(1, \frac{2}{\pi}, \frac{1}{2}\right)$$

For (1D, 2D, 3D) respectively, where V is the maximum velocity.

Similarly, for diffusive transport the integration for $(\tau V^2 Cos^2\theta)$ will be taken over a solid angle to give;

$$Đ = V^2\left(1, \frac{1}{2}, \frac{1}{3}\right)$$

For 1D, 2D, 3D respectively, where V is the maximum velocity, the final result seems to be like the reciprocal of the number of dimensions. Therefore;

$$\lambda = \frac{\tau V^2\left(1, \frac{1}{2}, \frac{1}{3}\right)}{V\left(1, \frac{2}{\pi}, \frac{1}{2}\right)}$$

$$\lambda = V\tau\left(2, \frac{\pi}{2}, \frac{4}{3}\right) \qquad (3.21)$$

This definition of $(\lambda)$, the mean free path $(V\tau)$ multiplied by certain factors depending on the dimensions of the conductor. These factors are always greater than one, equals (2) in one dimensional conductor, (1.6) in two-dimensional conductor, and (1.3) in three-dimensional conductor, due to the back scattering in each case. $(\tau)$ is the mean free time for backscattering, where only small fractions of all scattering processes are considered, or only those backscattering which are able to turn around.

This definition of $(\lambda)$ can be used to connect the fundamental parameter $(G_B)$ to the conductivity $(\sigma)$ by some factor F;

$$\sigma = G_B.F \ \Rrightarrow \mathrm{F} = \left(\frac{\sigma}{\mathrm{G_B}}\right)$$

$$F = (q^2 D\text{Đ}/xL)/(\, q^2 D\tilde{\mathrm{V}}/2L)$$

$$\mathrm{F} = \frac{\frac{q^2 D V^2 \tau}{xL}\left(1, \frac{1}{2}, \frac{1}{3}\right)}{q^2 (\frac{DV}{2L}) \left(1, \frac{2}{\pi}, \frac{1}{2}\right)}$$

$$\mathrm{F} = \frac{V\tau}{x}\left(2, \frac{\pi}{2}, \frac{4}{3}\right) = \frac{\lambda}{\mathrm{x}} \qquad (3.22)$$

$\left(\frac{DV}{2L}\right)$ has got a name called (the density of modes) which is equal to $\left(\frac{M}{h}\right)$, where $(M)$ is the number of modes, $(h)$ is the Planck's constant. In equation (3.22) above the mean free path $\lambda = V\tau\,(2, \frac{\pi}{2}, \frac{4}{3})$ can be written as in the resistance equation;

$$\mathrm{R} = \frac{\rho}{x}(L + mfp)$$

Where, $x = (1, w, A)$ for (1D, 2D, or3D) respectively.

It is slightly written in a misleading way, because it gives an impression that it is independent of the mean free time $(\tau)$ or any scattering at all.

Then,

$$\mathrm{R} = \frac{\rho}{x}(L + mfp) \ \Rrightarrow G = \frac{x}{\rho(L + mfp)}$$

$$\mathrm{G_B} = \frac{x}{\rho(0 + mfp)} = \frac{x}{\rho.mfp} = \frac{q^2 M}{h}$$

$$G_B\,.mfp = G_B\,.\lambda = \sigma x$$

But when the factor $(F)$, from the relation $(G_B.F = \sigma)$ is used;

$$\frac{\mathrm{G_B}.\lambda}{x} = \sigma \;\Rrightarrow\; \sigma x = G_B \lambda$$

Where $(\lambda)$ contains all the scattering processes factor for different dimensions plus the mean free path $(\tau)$.

## 3.11 Wrapping up summary

This chapter (3) started by the current equation of an elastic or Landauer resister as a function of the difference between the two Fermi-functions,

$$I = 1/q \int_{-\infty}^{\infty} G(E) \Delta f_F dE \qquad (3.23)$$

Using the low bias approximation to get an expression for the conductance,

$$G = \frac{I}{V} = \int_{-\infty}^{\infty} G(E) \left( -\frac{\partial f_0}{dE} \right) dE \qquad (3.24)$$

Where the conductance $(G)$ looks like the average of the conduction function $G(E)$. The conductance function $G(E)$ depends on the density of states and the time taken by an electron to go through the channel,

$$G(\,E) = \frac{q^2 D(E)}{2t(\,E)}$$

Starting from this relation of $G(E)$ a general expression for the conductance can be obtained, for both the ballistic and the diffusive limit;

$$G = \frac{G_B \lambda}{(L + \lambda)} = \frac{\sigma \mathrm{A}}{(L + \lambda)}$$

When the length of the channel $(L)$ is very short compared to the mean free path $(\lambda)$ the length $(L)$ can be dropped to have the ballistic conductance, in contrast if $(L)$ is large compared to $(\lambda)$, so the mean free path $(\lambda)$ can be dropped to get back as Ohm’s law requires for the channel to go with the length $(L)$. By this way different expressions for conductivity are given, using this new prospective,

$$\left(G = \frac{G_B \lambda}{(L+\lambda)} = \frac{\sigma A}{(L+\lambda)}\right)$$

The conductivity is given by,

$$\sigma = \frac{G_B \lambda}{(A)}$$

Where the conductivity equals the ballistic conductance per unit area times the mean free path. Whereas in the old prospective, the Drude formula conductivity is given by,

$$\sigma = \frac{q^2 \tau n}{m}$$

where, $(n)$ is the free electron density

$(m)$ is the effective mass

$(\tau)$ is the mean free time

One of the main objectives is to connect the two conductivity relations by an introduction of specific model of the energy levels, therefore, everything should be discussed in the next chapter (4) in terms of the density of states to get the conductance, the current, etc. To compare the two-conductivity equations is by the usage of the following correspondence,

$$\sigma = \left(\frac{G_B}{A}\right) \lambda$$

$$\sigma = q^2 \left(\frac{n}{m}\right) \tau$$

$$\lambda \leftrightarrow \tau \quad and \quad \left(\frac{G_B}{A}\right) \leftrightarrow q^2 \left(\frac{n}{m}\right)$$

# Chapter Four
# Energy Band Model

## 4.1 Introduction

The density of states plays a central role in determining how well something conducts, basically the material conducts if there is a large density of states around an equilibrium chemical potential μ. The density of states has not been discussed much to see from where it originates and how it can be modeled **[47]**. In big conductors it is more convenient to talk in terms of the density of states which tells how much states are there at a given range of energy, that is to say, the number of states per unit energy, and it could be measured by photoemission experiments.

Generally the density of states can be modeled starting from Schrodinger equation to see the wave nature of electrons, but in this chapter the modeling will be done in a relatively simple model based on energy-momentum relation $E(P)$; taking into account the electron in solids behaves almost as if it is in vacuum , but with energy-momentum relation $E(P)$, $E(K)$, or with different mass.

$E(P)$ curve and the density of states curve look similar in shape, but they are totally different in the sense that, momentum can take any direction, thus the $E(P)$ curve can be in both the positive and negative sides of the curve, while the density of states curve should take the positive side only. (see figures (4.1: left, right) below,

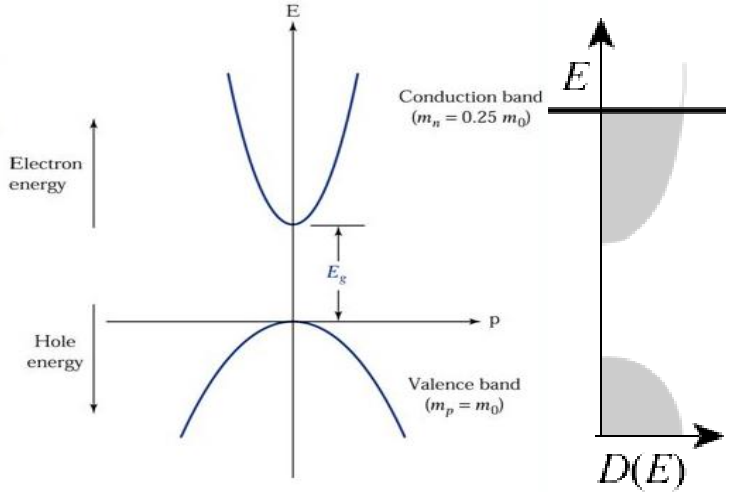


Figure 4.1: Left/Schematic energy momentum diagram; Right/Density of states diagram

Starting from $E(P)$ relation, by $E(P)$ translation, density of states $D(E)$, and the number of modes $M(E)$ can be got. Also in this chapter how the Drude formula relates the conductivity to the electron density will be discussed, whereas the ballistic conductance $G_B = q^2\left(\frac{DV}{2L}\right) = q^2\frac{M}{h}$ plays a central role in the discussion depending on either of $D(E)$, or $M(E)$ at a given range of energy , that means over that energy, averages at high temperatures and low temperatures' are needed.

Philosophically as if two different things are dealt with, one of them is the value about certain energy and the other is the integrated number of electrons.

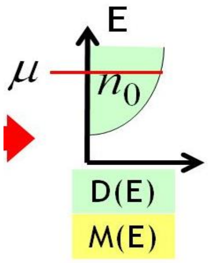
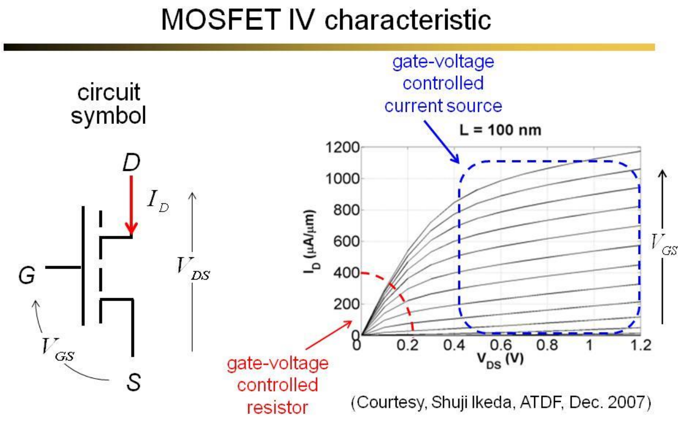


Figure 4.2: Electron density, value of integrated number of electrons

To connect these two types by discussing the electron density and explaining why the new prospective usually agrees with what is got from the Drude formula, but sometimes they don't agree.

Finally, one of the most important devices, the Nano transistor is going to be discussed. The main goal of this thesis is the ability to apply Moor's law for continuous miniaturization of the electronics devices. Up to now an over simplified model of a transistor has been shown, because the third terminal (the gate) has been left out, which is going to be added in this chapter, and it is completely isolated from the channel, in that sense ideally no current flows through that terminal, but in practice here is a negligible leakage current. The open circuited gate voltage $(V_g)$ moves the energy levels in the channel up and down according to whether the gate voltage $(V_g)$ is negative or positive respectively. Since the conductivity $(\sigma)$ depends on where is the chemical potential $(\mu)$ is, the conductivity can be changed by raising the energy levels to stop the current flow if the chemical potential $(\mu)$ gets into the gap. There are two ways to think about this; either the chemical potential is kept fixed and moving the energy levels, or vice versa. The essence of the transistor is the ability to control how well the channel conducts controlling the resistivity $(\rho)$ by several orders of magnitudes. The characteristics of the transistor looks something like this (see figure below);



(Courtesy, Shuji Ikeda, ATDF, Dec. 2007)

Figure 4.3: MOSFET IV characteristics

For more positive gate voltage $(V_g)$, the gate conducts better to get more current. Till now low bias is considered, where the conductance around the origin $(\frac{dI}{dV})$, is taken. This controlled low bias conductance introduces a very important concept of quantum capacitance. At high bias the shape of the transistor characteristics will change tending to saturate as the voltage bias is increased. For saturation to be perfect, it seems to be a very difficult task to do as the device gets smaller.

## 4.2 Energy-Momentum Relation

Energy of an electron in vacuum is the sum of the kinetic energy plus the potential energy, written as;

$$E = \frac{P^2}{2m_0} + U \qquad (4.1)$$

Where, $(P)$ is the momentum

$(m_0)$ is the free electron mass (the rest mass)

$(U)$ The potential energy of the electron or (the well)

Plotting the energy-momentum relation, the bottom of the parabolic band at $(P = 0)$ equals the potential energy $(U)$, The parabolic dispersion looks as follows,

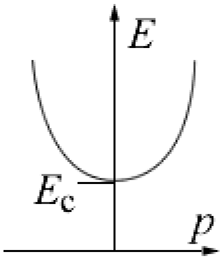


Figure 4.4: Parabolic dispersion of energy momentum relation

An equivalent shape can be shown relating the velocity using the kinetic energy equation;

$$E = \frac{m_0 \ V^2}{2} \qquad (4.2)$$

These two relations; $E = \frac{m_0 \ V^2}{2}$ and $= \frac{P^2}{2m_0}$ , are equivalent only in this special case of parabolic energy-momentum relation **[48,49,50,51,52]**, but in general we have many other types of energy-momentum relations, where $(P)$ is not always equals $(mV)$ in the sense that it is not always true to multiply the velocity by the mass to get the momentum. The general way, the momentum must come first and from which the velocity can be deduced in a certain way.

In solids, atoms are arranged in a very complicated periodic manner, and there should be an energy- momentum relation where the potential energy $(U)$ is not constant but changing periodically all over. If there is a unit cell energy-momentum relational be drawn according to the laws of solid states physics, which can be justified by quantum mechanics models, which is going to be avoided for the time being till the follow up is taken, when getting to quantum models, left for the future work.

By accepting that the electrons in a solid behave almost as if they are in vacuum, but with different mass, called the effective mass, gives a chance not to go deep in quantum mechanics in this chapter. The energy – momentum relations are different even in complicated solids such as silicon and ($GaAs$), because the electron behaves somewhat with different mass (the effective mass), and it won't look anything as simple as a parabola, but it may take any shape, typically parabolic bands as shown here bellow in some solids;

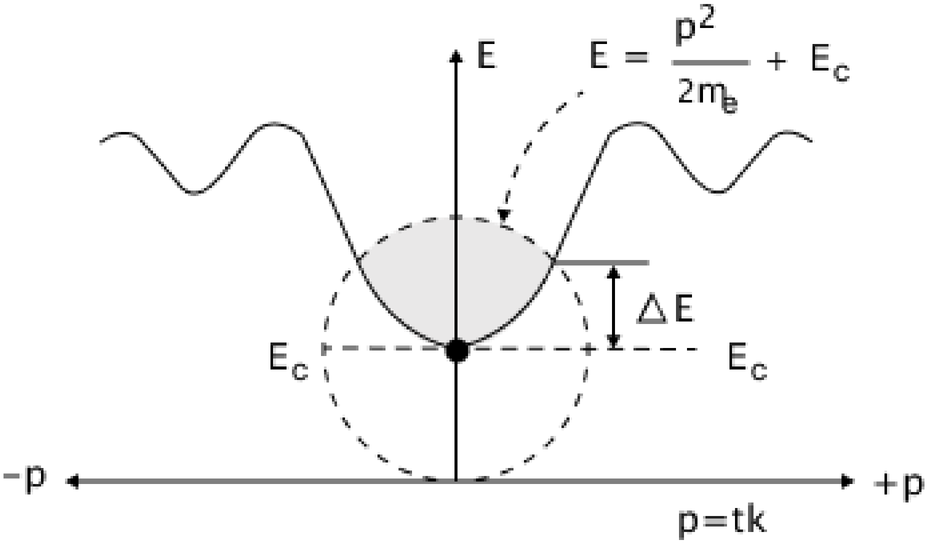


Figure 4.5: Parabolic bands of energy momentum relation

The bottom of the band is ($E_c$). Since the electron in the solid behaves in the energy range of interest ΔE like in vacuum, with a different effective mass ($m_e$), if attention is restricted in small region shown by a circle in the figure above, it looks like a parabola with different mass other than the free electron mass ($m_0$) which is going to be different in solids, but still the energy can be written as;

$$E = \frac{P^2}{2m_e} + E_c$$

Where, $m_e$ is the effective mass

This is just a true approximation over small range of energy (ΔE). If the electrochemical potential passes through ($E_c$) the conduction processes depend on the energy levels around the electrochemical potential ($\mu$), energies far away from the chemical potential don't matter, what matters only those energies inside the circle, possibly not even the upper parts of the curve. To describe the entire curve with the same relation $E(P)$ as far as the ability to fit the parabola well. As far as the bands are parabolic momentum equals $(P = mV)$, so the laws of mechanics apply for all the entire shape of the energy- momentum relation.

For a parabolic band,

$$E = \frac{P^2}{2m_e} + U \;\; \Rrightarrow \;\left[\frac{\partial E}{\partial P} = \frac{P}{m_e} + 0 = V = \left(\frac{\partial Z}{\partial t}\right)\right] \qquad (4.3)$$

Second Newton’s law states that the rate of change of momentum with respect to time is the gradient of energy, which is equal to the force,

$$\mathbf{P = m_e V} \Rrightarrow \left[\frac{\partial P}{\partial t} = m_e \left(\frac{\partial V}{\partial t}\right) = m\vec{a} = \vec{F} = Force \Rrightarrow \frac{\partial P}{\partial t} = \left(-\frac{\partial E}{\partial Z}\right)\right] \qquad (4.4)$$

The minus sign in equation (45) above shows that, the going direction in which the function increases most rapidly, it decreases the most rapidly.

In non- parabolic band energy -momentum relation is not given by $E = \frac{P^2}{2m_e} + U$ , instead it can be more complicated than that, and there are many popular materials such as graphene can take a linear relationship with (P),

$$E = V_0\, P \qquad (4.5)$$

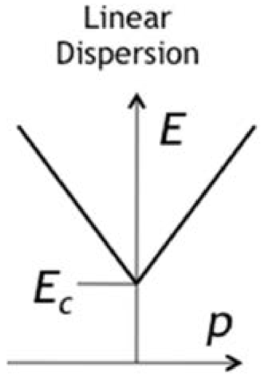


Figure 4.6: Linear dispersion of energy momentum relation

Sometimes the most suitable relation is the relativistic energy-momentum of the form;

$$E^2 = (mc^2)^2 + C^2 P^2 \qquad (4.6)\ see\ appindix\ (III)$$

When $(P = 0)$, $E = mC^2$ , the famous Einstein equation. Einstein special theory of relativity describes how space and time should be modified when systems and particles are moving close to the speed of light $(C)$, this the generalization of classical ideas of kinematics which describes the space and time, but how about dynamics? Likely to talk about forces, momentum, and energy to come up with some relativistic generalization of these ideas (*Einstein was able to generalize them).*

Using the non-relativistic, the momentum is equal to the mass multiplied by the velocity $(P = m_0 V)$, to understand how things work, the conservation of momentum is very useful. Conservation of momentum does not mean that the momentum is the same for all observers while the velocity is a relative concept in different reference frames, hence an object will have different velocities, so the momentum, and that is not what the conservation of momentum says.

The interaction between any two particles having no external forces acting on them, they collide and recoil, conservation of momentum says the total momentum of the system is the same before and at any time after.

If the momentum is conserved in one reference frame it better to be conserved in the other reference frame, and if $(P = m_0 V)$ is used, by looking at some collision where the momentum is conserved, Lorenz equation should be used when shifting to the other reference frame, to tell the new position and the new velocities of particles, then by addition of the of the equations, the velocities of all particles are worked out before and after in the new co-ordinate system, to discover that the momentum is not conserved when viewed in different reference frames. To fix it up in a very simple way by;

$$P = \gamma m_0 V \qquad (4.7)$$

That means a relativistic generalization by making sure that the definition agrees with the old definition in the limit that velocities are small compared to the velocity of light if $(\gamma)$ is close to one. Whenever the velocity is close to zero, sure enough that ($P = m_0 V$), that is, the equation reduced to the non-relativistic form. When something is moving fast the formula is ($P = \gamma m_0 V$), thats is, the correct momentum will be significantly larger. But in solids these relativistic speeds do not exist, in spite of that energy-momentum relations could easily possesses the form (see equation 4.6), shown here below;

$$E^2 = (mc^2)^2 + C^2 P^2$$

The equation is not in terms of actual constants, C is not the velocity of light in this case, but much smaller. The equation takes that shape from the mathematical point of view. There could have been any shape of energy-momentum relation to describe the dynamics of an electron in non-parabolic band by writing laws of mechanics in the following forms;

$$\left[\left(\frac{\partial E}{\partial P}\right) = \frac{\partial Z}{\partial t} = velocity\right]$$

$$\left[\frac{\partial P}{\partial t} = \left(-\frac{\partial E}{\partial Z}\right) = force\right]$$

Force is equal to the gradient of ($E$) with distance or the rate of change of momentum with time. The relation between momentum and velocity can be seen by the derivative of ($E^2$) with respect to momentum;

$$E^2 = (mc^2)^2 + C^2 P^2$$

$$2E\frac{\partial E}{\partial P} = 0 + (2PC^2)$$

$$E(V) = PC^2 \Rrightarrow V = \frac{PC^2}{E} \quad (4.8)$$

$$m \uparrow = \frac{E \uparrow}{C^2} \qquad (5.9)$$

Consider ($C$) to be constant, therefore mass goes with ($E$), it increases and decreases with the increase or decrease of ($E$). When ($E$) increases mass gets bigger. Since the mass is directly proportional to ($E$) the mass is not constant in non-parabolic energy-momentum relations, it is relativistic and change with energy. The mass then is an energy dependent, and amounts to;

$$m = \frac{m_0}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} \qquad (4.10)$$

Where, $(m_0)$ is the rest mass. This is the relativistic enhanced masses seen in the freshman's physics. The mass could be an energy dependent, so if the electrochemical potential happened to be higher up in the band the usage of different mass should be used.

## 4.3 Counting States

The basic idea of counting states is not clear how to count the density of states, for any energy- momentum relation, an electron could have any momentum corresponding to more than one energy.

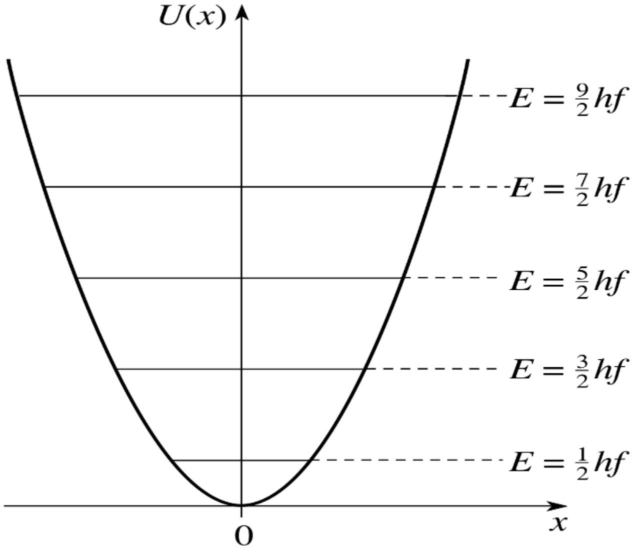


Figure 4.7: Energy momentum relation

It is like continuum, the argument concerned in a solid with a certain finite size, all the electrons momenta are not allowed which lead to discrete values of momentum corresponding to an allowed energy level in the upper band;

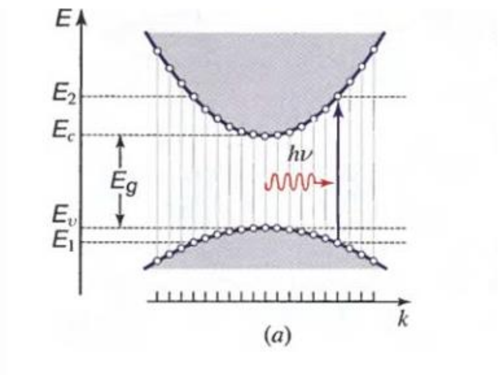


Figure 4.8: Allowed energy levels corresponding to discrete values of momentum

These discrete set of momenta can be translated into density of states $D(E)$, or into number of modes $M(E)$.

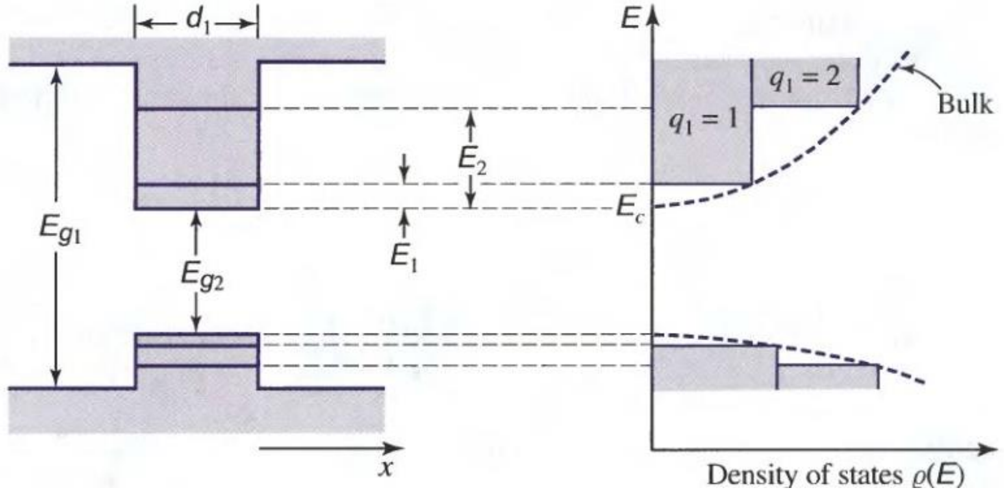


Figure 4.9: Translation of momenta into density of states

For certain maximum value of momentum $(P)$ how many allowed states therein whose momentum is less than P? A function $N(P)$ is defined to tell the number of states within the solid of one-dimensional conductor (a line conductor), two-dimensional conductor (a circle with radius P), or three-dimensional conductor (a sphere with radius P).

Since the energy-momentum relation is known the momentum $(P)$ can be eliminated to get the number of energy states $N(E)$ for an energy less than some value $(E)$ to deduce the density of states $D(E)$, given by;

$$D(E) = \frac{dN(E)}{dE} \qquad (4.11)$$

If the energy $(E)$ is increased by $(dE)$, extra states are added which is equal to the density of states in that energy range $(dE)$, to discretize the momentum, the wave nature of electrons is brought in, that means, basically to introduce the quantum mechanics. Not to go deep into quantum mechanics, it is preferred to be done in an elementary way based on DeBrogle wavelength $(\lambda_D)$ related by;

$$\lambda_D = \frac{h}{P} \qquad (4.12)$$

Where, $\lambda_D$ is DeBrogle wavelength

$h$ is the Planck's constant

$P$ is the momentum

The essence of this counting scheme for one dimensional conductor of length $(L)$, the length must be an integer times the wavelength, often called the Period Boundary Condition $(PBC)$. The only allowed momenta are given by;

$$L = n\left(\frac{h}{P}\right) \qquad (4.13)$$

Therefore, the electron cannot have any momentum other than some certain values $(P_n)$, the condition imposed on that, an electron with certain wavelength to fit into a box of length $(L)$. Turning equation (4.12) around gives the spacing between allowed momenta as a discrete multiple of $\left(\frac{h}{L}\right)$ for a one-dimensional conductor;

$$P_n = n\left(\frac{h}{L}\right)$$

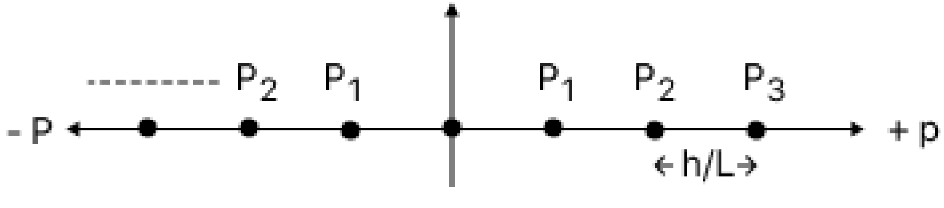


Figure 4.10: Allowed momenta as discrete multiple of (h/L)

Therefore, so easily $N(E)$ can be found as;

$$N(P) = \frac{2P}{\frac{h}{L}} = 2L\left(\frac{P}{h}\right) \qquad (4.14)$$

For big values of $(L)$, the allowed values of $(P_n)$, are very dense and close together, and further apart for short lengths, that is why a bigger box is chosen to get more states, since the states depend on the size of the solid.

Considering the wave property of an electron, the wave possesses a wave number property $(K)$, where $(K = \frac{2\pi}{\lambda_{\mathrm{D}}})$, then;

$$(P = \hbar K = h/\lambda_D)$$

This equation gives the wave-particle duality, that is the momentum $(P)$, the particle, and the wavenumber $(K)$, the wave. Both are related by the reduced Planck's constant $(\hbar = h/2\pi)$, therefore, $(K)$K is just an equivalent to $(P)$;

$$P_n = n\left(\frac{h}{L}\right) = \hbar K_n \Rrightarrow K_n = \frac{P_n}{\hbar} = n\left(\frac{h}{L}\right)\left(\frac{1}{\hbar}\right) = n(\frac{2\pi}{\mathrm{L}})$$

All the states are going from $(-P)$ to $(+P)$ and that is why the range is taken as $(2P)$ and spaced by $\left(\frac{h}{L}\right)$.

The total number of states $N(P)$ in two-dimensional conductor with $(P)$ happened to be less than any momentum value $(P)$ is directly proportional to $(P^2)$;

$$N(P) = \frac{\pi P^2}{\left(\frac{h}{L}\right) \cdot (\frac{h}{W})} = \pi W L (P/h)^2 \qquad (4.15)$$

Where, $(L)$ is the length in one direction, and $(W)$ is the length in the other direction, so these states are spaced by $\left(\frac{h}{L}\right)$ on $(P_x)$, and $\left(\frac{h}{W}\right)$ on the $\left(P_y\right)$ direction to occupy an area of $\left(\frac{h^2}{LW}\right)$ for each state.

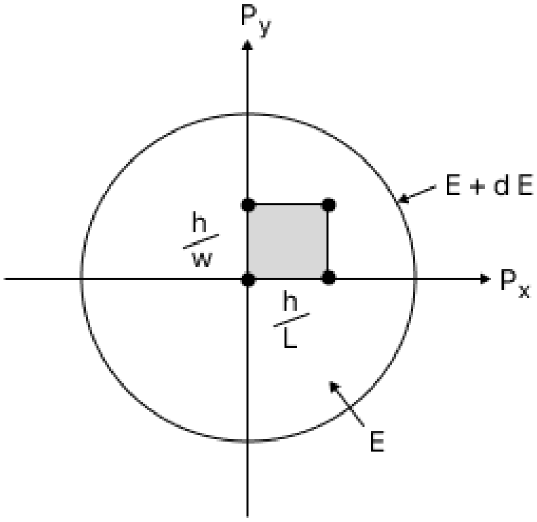


Figure 4.11: An area of each state in 2D conductor

For three-dimensional conductor instead of an area a volume of sphere is considered with a radius $(P)$ happen to be less than the momentum value $(P)$

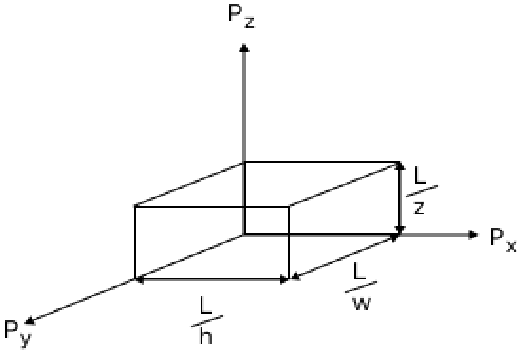


Figure 4.12: Volume of each state in 3D conductor

The total number of states is directly proportional to $(P^3)$;

$$N(P) = \frac{\frac{4}{3}\pi P^3}{\left(\frac{h}{L}\right)\cdot\left(\frac{h}{W}\right)\cdot\left(\frac{h}{Z}\right)} = \frac{\frac{4}{3}\pi}{AL}(P/h)^3 = \frac{4\pi AL}{3}(P/h)^3 \qquad (4.16)$$

To collect all the three relations together to write it in a single expression for $N(E)$ in one dimension, two dimensions, or three dimensions conductors respectively,

$$N(P) = \left(\frac{P}{h}\right)^{d} \left(2L, \pi WL, \frac{4}{3}\pi AL\right) \qquad (4.17)$$

Therefore, $N(E)$ is directly proportional to $\left(P^{d}\right)$, where $(d)$ is the number of dimensions we are in, all the other bracketed factors represent the constants which hoes in front of $\left(P^{d}\right)$ of the general relation;

$$N(P) \sim P^{d} \Rrightarrow N(P) = AP^{d} \qquad (4.18)$$

Exactly what constant goes in front of $\left(P^{d}\right)$, even what that constant is not known the overall picture can be seen by combing $E(P)$ with $N(P)$, then by elimination of $(P)$ the number of energy states $N(E)$ can be got for a value less than the energy value $(E)$, and by using equation (52) the density of states $D(E)$ can be calculated.

## 4.4 The Density of States $D(E)$

In the last section, section (4.3) above the counting of states rule $N(P)$ is obtained in equation (4.17) by using the Periodic Boundary Conditions $(PBC)$ to tell how many states are allowed to have a momentum less than a momentum value $(P)$.

The rule $N(P)$ comes simply by fitting only the corresponding momentum into a box using the DeBrogle wavelength making no use of any energy-momentum relation. This is the general method of counting the number of states based on $(PBC)$ as widely used in solid state physics.

To get $N(E)$, the number of states allowed to have an energy less than an energy value $(E)$, by coupling $N(E)$ with the energy momentum relation $E(P)$, which in general could have many complicated forms of relations, and by eliminating $(P)$, $N(E)$ is given , the derivative of $N(E)$ with respect to energy $(E)$ gives the density of states $D(E)$.

If an isotropic energy-momentum is assumed where energy is the same for any momentum in all directions **[53,54,55]**. In the figure below the shaded region corresponds to a certain value $(E)$. Increasing the momentum, a little bit, the corresponding $(E)$ will increase by an amount $(dE)$, so slightly bigger circle results,

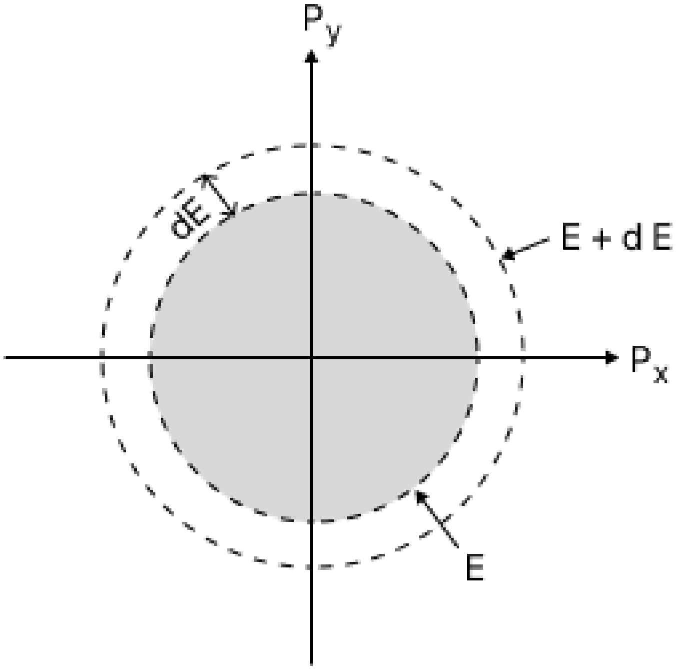


Figure 4.13: An isotropic energy momentum

The number of states available are increased representing the density of states at that energy.

From equation (4.18), having the function, $N(P) = AP^{ȡ}$ where the unimportant constant $(A)$ varies with $(ȡ)$, in general $E(P)$ can be described by;

$$E(P) = E_c + \beta P^{\alpha} \qquad (4.19)$$

Where, $E_c$ is the bottom of band, and $E_c$ varies with $(P)$ in a certain way.

$\beta$ is an unimportant constant multiplied by $P^{\alpha}$, also varies with $P$ in a certain way depending on $\alpha$.

Parabolic bands are often dealt with where $(\alpha = 2), (\beta = \frac{1}{2}m)$ then,

$$E(P) = E_c + \frac{P^2}{2m},$$

But it is better to keep the equation in its general form for the purpose of other $E(P)$ relations.

From equation (4.19);

$$P^{\alpha} = \frac{(E(P) - E_c)}{\beta} \Rrightarrow P = \left(\frac{E(P) - E_c}{\beta}\right)^{\frac{1}{\alpha}}$$

Eliminate $(P)$ by substituting it in equation (4.18), to get $N(E)$;

$$N(E) = \frac{A}{\beta}(E(P) - E_c)^{\frac{\text{ȡ}}{\alpha}}$$

$$D(E) = \frac{dN(E)}{dE} = \frac{A\text{ȡ}}{\beta\alpha}(E(P) - E_c)^{(\frac{\text{ȡ}}{\alpha}-1)}$$

$$D(E) \sim (E(P) - E_c)^{\left(\frac{\text{ȡ}}{\alpha_1}-1\right)} \qquad (4.20)$$

This the general expression for the density of states, $D(E)$ depends on the number of dimensions (ȡ) and the factor $(\alpha)$ which describe how the $E(P)$ relation varies with momentum.

To apply these results for a specific case, the most common one is the parabolic band, where (α=2);

(1) *For one-dimensional conductor*, ȡ =1; then $D(E) \sim r(E(P) - E_c)^{\frac{1}{2}}$ ,which can be represented by the following figure;

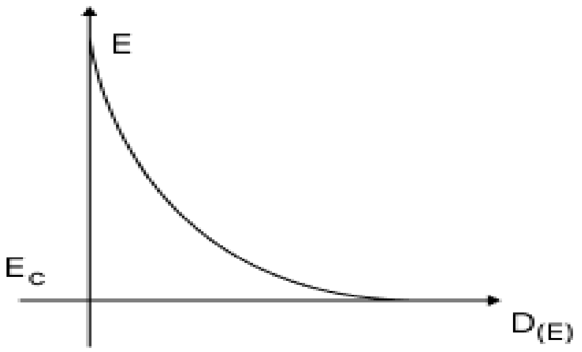


Figure 4.14: Density of states of 1D conductor for parabolic relation

That figure has a singularity right near the band edge ($E_c$) with energy. These things are observed experimentally

(2) *For a two-dimensional conductor*, α = ȡ =2, then, $D(E) \sim (E(P) - E_c)^0 = constant\ K = \frac{A\text{ȡ}}{\alpha\beta}$ , which can be represented by the following figure;

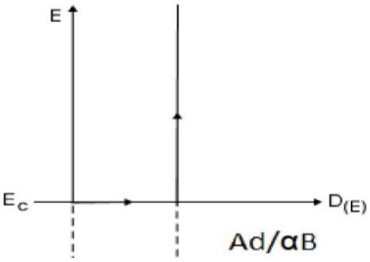


Figure 4.15: Density of states of 2D conductor for parabolic relation

(3) *For three-dimensional conductor*, ȡ=3, α=2, then $D(E) \sim (E(P) - E_c)^{1/2}$, which can be represented by a parabolic band for (3D) bulk semi-conductor shown by the following figure;

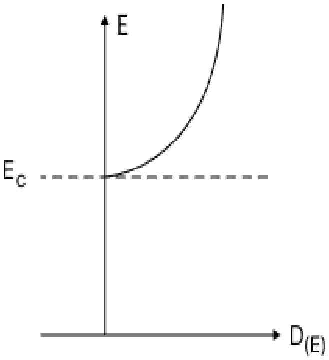


Figure 4.16: Density of states of 3D conductor for parabolic relation

Also, there are very popular materials such as graphene, in graphene $E(P)$ can take a linear relationship with $(P)$;

$$E(p) = V_0 P^{\alpha}$$

Where, $(\alpha = 1)$ for graphene, (see appendix (IV));

To combine $E(P)$ with the rule of counting states $N(P)$ to obtain $N(E)$ by eliminating the momentum $(P)$;

$$N(E) = \left(\frac{E(P)}{V_0}\right)^{\frac{\text{ȡ}}{\alpha}}$$

$$D(E) = \frac{dN(E)}{dE} = \frac{A\text{ȡ}}{\alpha}\left(\frac{E(P)}{V_0}\right)^{\left(\frac{\text{ȡ}}{\alpha}-1\right)} \qquad (4.21)$$

As in $E(P)$ relation instead of parabolic it is linear in this case, and all the above results will change;

(1) For (1D), α = ȡ =1, then $D(E) \backsim E = constant = \left(\frac{A\text{ȡ}}{\alpha V_0}\right)$, which can be represented by the following figure;

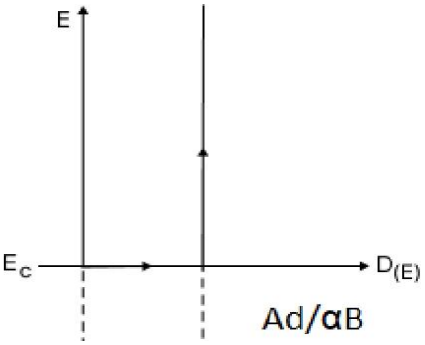


Figure 4.17: Density of states of 1D conductor for linear relation

(2) For (2D), α=1, ȡ=2, then $D(E) \backsim E$ which can be represented by a linear relationship by the following figure

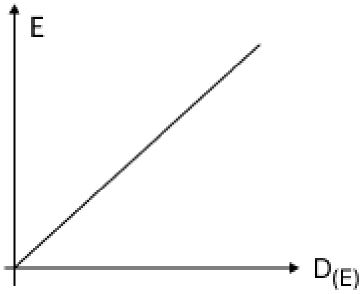


Figure 4.18: Density of states of 2D conductor for linear relation

(3) For (3D), α=1, ḑ=3, then $D(E) \backsim E(P)^2$ which can be represented by the following figure;

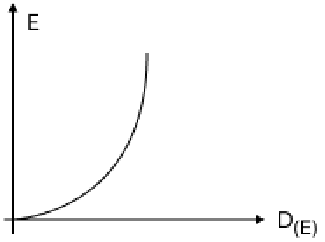


Figure 4.19: Density of states of 3D conductor for linear relation

The energy-momentum relation $E(P)$ can be more complicated than that, it could have been in a relativistic form, where energy is the square root of some constant plus some $(P^2)$,

$$E(P) = (Eg^2 + {V_0}^2 P^2)^{\frac{1}{2}}$$

In this case formula given by equation (61) cannot be used, therefore, $E(P)$ should have been taken to find $N(P)$, then to combine it with $N(P)$ to get $D(E)$ by the derivative of $N(E)$ with respect to $E$;

$$P = \left(\frac{E(P)^2 - E_g^2}{V_0^2}\right)^{\frac{1}{2}}$$

Eliminate $(P)$;

$$N(E) = \ A\left(\frac{E(P)^2 - E_g^2}{V_0^2}\right)^{2ḑ}$$

### 4.5 Number of Modes

Here a very useful relation independent of specific energy-momentum $E(P)$ is introduced to have a general relation valid for all the $E(P)$ relations and could be used to obtain an expression for the number of modes $M(E)$**[56,57,58].** Back to the rule of counting states $N(P)$ given by equation (4.17), the rule does not include any energy-momentum relation $E(P)$. It is just based on the idea of discretized momentum using the number of DeBrogle wavelength to fit in a box. Differentiate $N(P)$ with respect to $(P)$ to get a new important relation according to the following procedure,

$$N(P) = \left(\frac{P}{h}\right)^{\text{ȡ}} \left(2L, \pi WL, \frac{4}{3}\pi AL\right) \qquad (4.17)$$

$$\frac{dN(P)}{dp} = \frac{\text{ȡ}P^{\text{ȡ}-1}}{h^{\text{ȡ}}} \left(2L, \pi WL, \frac{4}{3}\pi AL\right)$$

$$\frac{dN(p)}{dP} Can\ be\ written\ as \frac{dN(E)}{dE} . \frac{dE}{dP}$$

$$\frac{dN(E)}{dE} . \frac{dE}{dP} = \frac{\text{ȡ}P^{\text{ȡ}-1}}{h^{\text{ȡ}}} \left(2L, \pi WL, \frac{4}{3}\pi AL\right)$$

Multiply both sides by $(P)$

$$P\frac{dN(E)}{dE} . \frac{dE}{dP} = p\frac{\text{ȡ}P^{\text{ȡ}-1}}{h^{\text{ȡ}}} \left(2L, \pi WL, \frac{4}{3}\pi AL\right)$$

$$P\frac{dN(E)}{dE} . \frac{dE}{dP} = \text{ȡ}\left(\frac{P}{h}\right)^{\text{ȡ}} \left(2L, \pi WL, \frac{4}{3}\pi AL\right)$$

$$P.D(E).V = \text{ȡ}.N(P) \qquad (4.22)$$

This basic relation should hold in general, because it is totally independent of any $E(P)$ relation, but there is $(\frac{dE}{dP})$ representing the velocity $(V)$, without any assumption of any particular form of energy-momentum relation $E(P)$.

The number of modes $M(E)$ can be related to the ballistic conductance $G_B$ given by $G_B = \frac{q^2 D\tilde{\mathrm{V}}}{2L}$ where the average velocity involves some numerical factors depending on the number of dimensions

$$G_B = \frac{q^2 D\tilde{\mathrm{V}}}{2L}\left(1, \frac{2}{\pi}, \frac{1}{2}\right)$$

The quantity $\left(\frac{D\tilde{\mathrm{V}}}{2L}\right)$ known as density of modes, which is equal to the number of modes $M(E)$ divided by the Planck's constant, and it has special significance, the material property, or (density of modes)

$$\frac{D\tilde{\mathrm{V}}}{2L} = \frac{M(E)}{h}$$

Therefore,

$$M(E) = \frac{\mathrm{h}D\tilde{\mathrm{V}}}{2L}\left(1, \frac{2}{\pi}, \frac{1}{2}\right) \qquad (4.23)$$

The number of modes can be found by multiplying the density of states by the velocity, but it can be done another way by using the general relation given by equation (63); where,

$$D(E).\tilde{\mathrm{V}} = N(P).\frac{\mathrm{d}\wp}{\mathrm{P}}$$

By substituting $(D\tilde{\mathrm{V}})$ in equation (64),

$$M(E) = \frac{\mathrm{h}}{2L}\left(\frac{N\mathrm{d}\wp}{P}\right)\left(1, \frac{2}{\pi}, \frac{1}{2}\right) = ((\frac{h}{P})\frac{\mathrm{Nd}\wp}{2L}\left(1, \frac{2}{\pi}, \frac{1}{2}\right)$$

$$N(P)=\left(\frac{P}{h}\right)^{\text{ȡ}}\left(2L,\pi WL,\frac{4}{3}\pi AL\right)$$

Both of $N(P)$ and $M(E)$ have an $(h/P)$ factor, so $M(E)$ can be written as;

$$M(E)=\frac{\text{ȡ}}{2L}\left(\frac{P}{h}\right)^{\text{ȡ}-1}\left(2L,\pi WL,\frac{4}{3}\pi AL\right)\left(1,\frac{2}{\pi},\frac{1}{2}\right)$$

Then,

$$M(E)=\frac{\text{ȡ}}{2L}\left(\frac{P}{h}\right)^{\text{ȡ}-1}\left(2L,2WL,\frac{2\pi}{3}AL\right)$$

Substitute $\left(\frac{\text{ȡ}}{2L}\right)$ inside the bracket,

$$M(E)=\left(\frac{P}{h}\right)^{\text{ȡ}-1}(1,2W,\pi A) \qquad (4.24)$$

Compare $N(P)$ with $M(E)$, the one-dimensional result $(2L)$ in $N(P)$ is shifted to the two-dimensional result $(2W)$ in $M(E)$, the same with two-dimensional result $(\pi WL)$ in $N(P)$ is shifted to the three-dimensional result $(\pi A)$ in $M(E)$, while the number of modes in one-dimensional conductor is unity in $M(E)$.

These results can give a totally different physical interpretation by writing these expressions of $M(E)$ and $N(P)$ in a different way as follows,

$$N(P)=\left(\frac{P}{h}\right)^{\text{ȡ}}\left(2L,\pi WL,\frac{4}{3}\pi AL\right)=\left(2L\left(\frac{P}{h}\right),\pi WL\left(\frac{P}{h}\right)^{2},\frac{4}{3}\pi AL\left(\frac{P}{h}\right)^{3}\right)$$

$$N(P)=\left(\frac{L}{\left(\frac{h}{2P}\right)},\frac{\pi WL}{4\left(\frac{h}{2P}\right)^{2}},\frac{\pi AL}{6\left(\frac{h}{2P}\right)^{3}}\right)$$

$$N(P) = \left( \frac{L}{\left(\frac{\lambda}{2}\right)}, (\frac{\pi}{4}) \frac{WL}{\left(\frac{\lambda}{2}\right)^2}, (\frac{\pi}{6}) \frac{AL}{\left(\frac{\lambda}{2}\right)^3} \right) \qquad (4.25)$$

Similarly, with $M(E)$,

$$M(E) = \left( 1, \frac{W}{\left(\frac{\lambda}{2}\right)}, (\frac{\pi}{4}) \frac{A}{\left(\frac{\lambda}{2}\right)^2} \right) \qquad (4.26)$$

Therefore, the number of modes physically equals one for (1D) conductor, how many half wavelengths of DeBrogle fit into the width in (2D) conductor, and how many DeBrogle half wave lengths fit in either direction of the cross-sectional area for the the(3D) conductor.

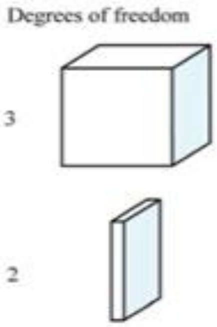


Figure 4.20: Number of modes in 2D and 3D conductors (Degree of freedom)

Similarly, the number of momentum states physically equals to how many DeBrogle half wave lengths fit into the length of (1D) conductor, how many DeBrogle half wave lengths fit in either direction of the cross- sectional area for (2D) conductor, and how many DeBrogle half wavelengths fit in each direction of the volume in (3D) conductor.

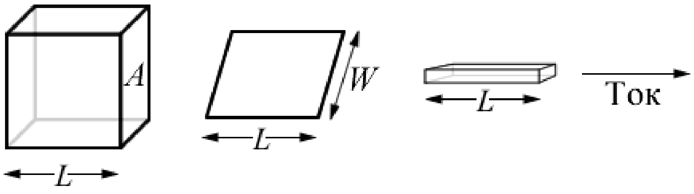


Figure 4.21: Number of momentum states in 1D ,2D, and 3D conductors

The word mode originates from the same ideas of electromagnetic wave guides, when the determination of the transverse modes for electromagnetic wave guides are required.

An important consequence of equation (4.26), if multiplied by a constant value $(q^2/h)$ gives the ballistic conductance $(G_B)$. Since the number of modes $M(E)$ is directly proportional to the width of the cross-sectional area, the ballistic conductance changes linearly with the width $(W)$ or the cross- sectional area $(A)$ of the conductor.This had been observed back in 1970's and sometimes called Sharven resistance.

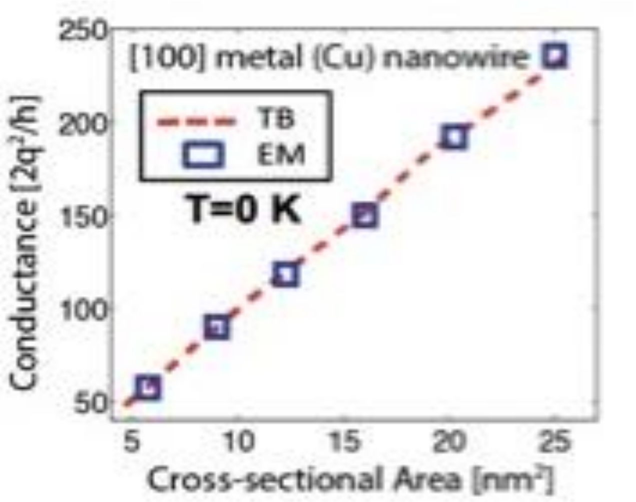


Figure 4.22: Ballistic metal conductance

In 1990's a very important development take place, that in small cross-sections, where the number of modes is very small in the order of ten to twenty thousand it was observed that the ballistic conductance take a quantized form rather than to go linearly with $(W)$ or $(A)$ **[59,60,61,62,63,64],** but it goes in steps;

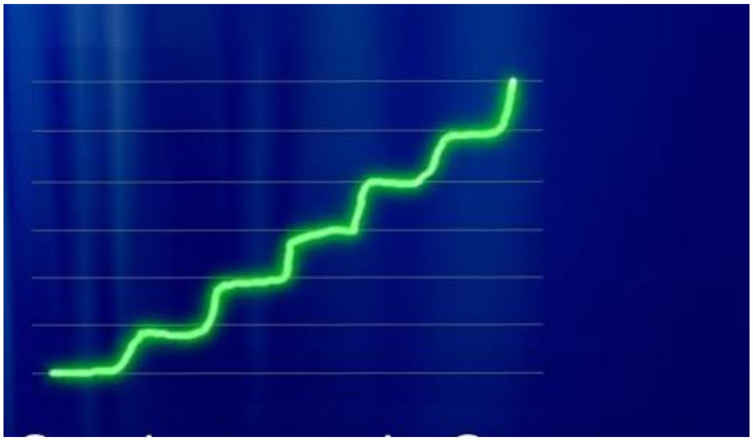


Figure 4.23: Quantized conductance in quantum point contacts

To understand this phenomenon, $M(E)$ is not exactly equal to the quantity $\left(1, \frac{W}{\left(\frac{\lambda}{2}\right)}, (\frac{\pi}{4}) \frac{A}{\left(\frac{\lambda}{2}\right)^2}\right)$, but if $\left(\frac{W}{\left(\frac{\lambda}{2}\right)}\right)$ happen to be a rational number it is actually a whole positive number, because the number of modes is given by how many states are available in that cross-section. Therefore, the correct expression is the integer part of $\left(\frac{W}{\left(\frac{\lambda}{2}\right)}\right)$, when $M(E)$ is small and $(G_B)$ is quantized, $M(E)$ is the integer value of the quantity inside the bracket,

$$M(E) = integer\left(1, \frac{W}{\left(\frac{\lambda}{2}\right)}, \left(\frac{\pi}{4}\right) \frac{A}{\left(\frac{\lambda}{2}\right)^2}\right)$$

But normally nobody cares about that when $M(E)$ is a large number in the order of thousand or more, that will make no difference if it is (1000 or 1001). Because the steps are very close together with these big numbers, so they need to be averaged with $\left[-\frac{\partial f}{\partial E}\right]$ factor over an energy $(KT)$. The steps smoothed out at high temperatures, but with low temperatures specially with conductors of small cross-sections this phenomenon is very well established. This was first seen in semi-conductors like $(G_B As)$ material even in small hydrogen molecules

## 4.6 Electron Density (n)

The ballistic conductance and the ballistic conductivity involve the density of states $D(E)$ and the number of modes $M(E)$, $G_B = \frac{q^2 D\tilde{V}}{2L} = \frac{q^2}{h} M$, do not quite involve the electron density $(n)$.

A band with an electrochemical potential $(\mu)$, shown here below,

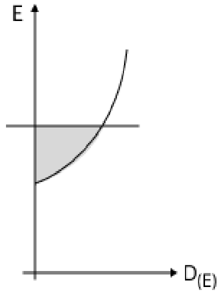


Figure 4.24: Band with electrochemical potential

The density of states at zero temperature can be determined by the number of electrons using the counting of states rule $N(P)$ to tell how many states are there below the electrochemical potential $(\mu)$, whose momentum is less than a momentum value $(P)$. The electron density is given by the number of electrons per unit length for (1D), per unit area for (2D), or per unit volume for (3D),

$$N(P) = \left(\frac{P}{h}\right)^{d} \left(2L, \pi WL, \frac{4}{3}\pi AL\right)$$

Then, the electron density $n(P)$ is given by,

$$n(P) = \left(\frac{P}{h}\right)^{d} \left(2, \pi, \frac{4}{3}\pi\right) \qquad (4.27)$$

To combine $n(P)$ with $E(P)$, after converting the momentum into $(E)$, the electron density as a function of energy $(E)$ will be given at zero temperature for all the states below $(\mu = E)$ per unit volume.

At reasonable temperature when $(T)$ is not equal $(0K)$ the states below $(\mu)$ are not fully occupied, and their occupation follow the Fermi-function which is changing in the order of $(KT)$ between one far below the Fermi-level, and zero far above it (see figure below),

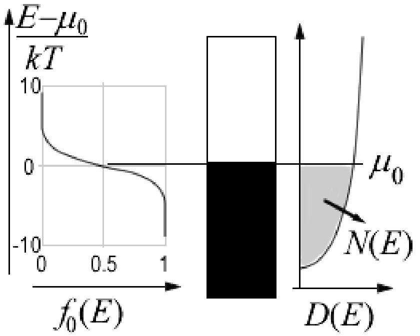


Figure 4.25: States occupation below electrochemical potential

To calculate the number of electrons $(N_0)$ using the following relation,

$$N_0 = \int_{-\infty}^{\infty} D(E).f_0(E)dE. \qquad (4.28)$$

Where, $D(E)$ is the density of states (the number of states per unit energy).

$f_0(E)$ is the Fermi-function to tell whether the state is occupied or not.

$$D(E) = \frac{dN(E)}{dE}$$

Then,

$$N_0 = \int_{-\infty}^{\infty} \frac{dN(E)}{dE} . f_0(E) dE. \qquad (4.29)$$

Integrate by parts, not that the derivative $(dE)$, instead of being at the first function $N(E)$ will be transferred over to the second function $f_0(E)$ with a minus sign associated with it.

$$N_{.0} = [(N(E) f_0(E)]^{\infty} + \int_{-\infty}^{\infty} N(E) \left[-\frac{\partial f_0}{\partial E}\right] dE$$

For the surface term $(Nf_0)$ at the limit of $(+\infty)$, $(f_0 = 0)$ but $N(E)$ is not equal to zero, their product equals zero, at $(-\infty)$ limit $(f_0)$ is not equal to zero but $N(E)$ is, because $N(E)$ starts at the bottom of the band where $N(E)$ equals zero and from there it starts to grow, again their product equals zero, therefore, $(Nf_0)$ is equal to zero.

$$N_0 = \int_{-\infty}^{\infty} N(E) \left[-\frac{\partial f}{\partial E}\right] dE \qquad (4.30)$$

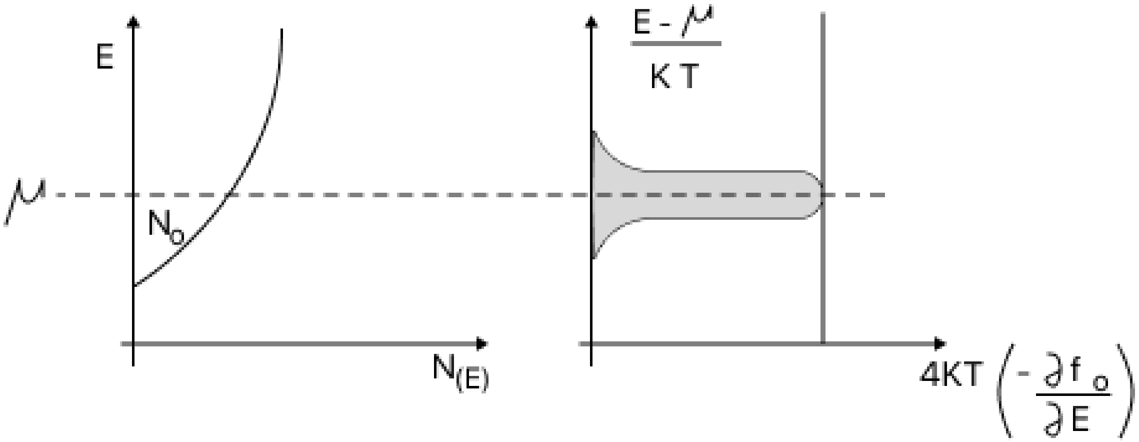


Figure 4.26: calculation of the number of electrons ($N_0$)

Equation (4.29) and (4.30) are mathematically the same, both give the same result. Note that at a very low temperature $\left[\frac{\partial f_0}{\partial E}\right]$ acts like a delta function $(\delta)$ right at $(E = \mu)$, and the value of $(n)$ will be given at the energy $(\mu)$. At non-zero temperature $N(E)$ should be averaged over a certain range of energy to be determined by $\left[-\frac{\partial f_0}{\partial E}\right]$.

Back to equation (4.22), where the electron density, $n(P) = \left(\frac{P}{h}\right)^{d}\left(2, \pi, \frac{4}{3}\pi\right)$, from which the number of electron density $(n_0)$ with a given electrochemical potential $(\mu)$ can be got from the following relation,

$$n_0 = \int_{-\infty}^{\infty} n(\, E) \left[-\frac{\partial f_0}{\partial E}\right] dE \qquad (4.31)$$

Where, $n(E)$ tells the electron density as a function of energy when all the states below $(\mu)$ are all filled up to $(E - \mu)$, therefore $n(E)$ is the total number of states per unit (length, area, or volume) for (1D, 2D, 3D) respectively for an energy value less than $(E)$ at zero temperature. In this special case $f_0(E) = 1,\ (-\frac{\partial f}{\partial E}) = \delta$.

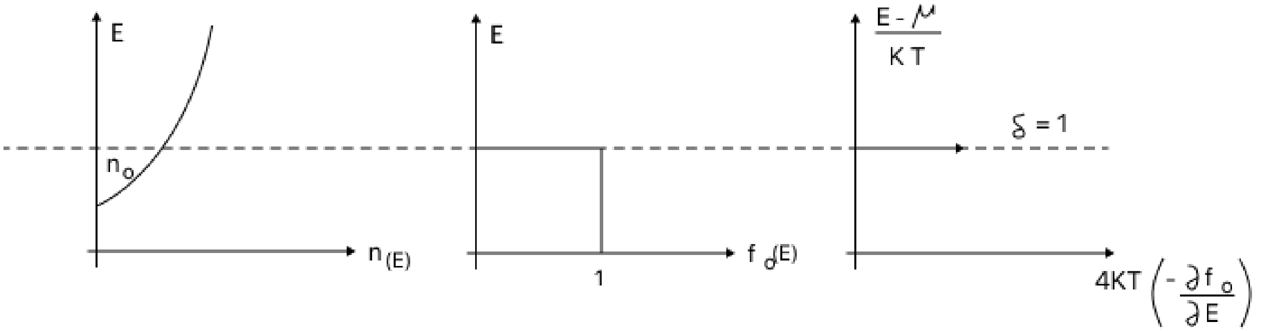


Figure 4.27: Special case when $f_0(E) = 1,\ (-\frac{\partial f}{\partial E}) = \delta$.

At non-zero temperature equation (4.31), can be generalized by averaging over certain energy range and the number of electron density $(N_0)$ is simply given by the value $N(E)$ as stated in equation (4.30), which has got the status of similarity with the conductance equation (3.5) chapter (3), brought here below,

$$G = \frac{I}{V} = \int_{-\infty}^{\infty} G(E)\left(-\frac{\partial f_0}{dE}\right) dE \qquad (3.5)$$

Therefore, the quantity $n(E)$ in equation (4.31) above, should be understood as the electron density at zero temperature.

The energy momentum relation in the valence band is running downwards, the top of which is denoted by $(E_v)$. When it is translated into a density of states or the number of modes, it takes some shape confined to the positive side of the curve, and if the electrochemical potential happen to pass somewhere through this band, in that case $N(E)$ gives the number of holes above $(\mu)$, or the number of empty states, the density of which $(N_0)$ is given by,

$$N_0 = \int_{-\infty}^{\infty} D(E)\big(1 - f_0(\,E)\big) dE \qquad (4.32)$$

Where, $N_0$is the number or density of empty states.

$f_0(\,E)$ tells whether the states are occupied or not.

$\big(1 - f_0(\,E)\big)$ tells whether the states are empty or not

$$D(E) = -\frac{dN(E)}{dE}$$

The minus sign shows the direction in which the function $D(E)$ decreases rapidly while $(E)$ increases rapidly or vice versa, then;

$$N_0 = \int_{-\infty}^{\infty} -\frac{dN(E)}{dE}\big(1 - f_0(\,E)\big)\, dE$$

Where, $N(E)$ represents all the states whose energy is higher than $(E = \mu)$, at the same time $N(P)$ tells how many states are available for a momentum less than a certain value $(P)$, to corresponds to the number of the empty states above certain energy value $(E)$.

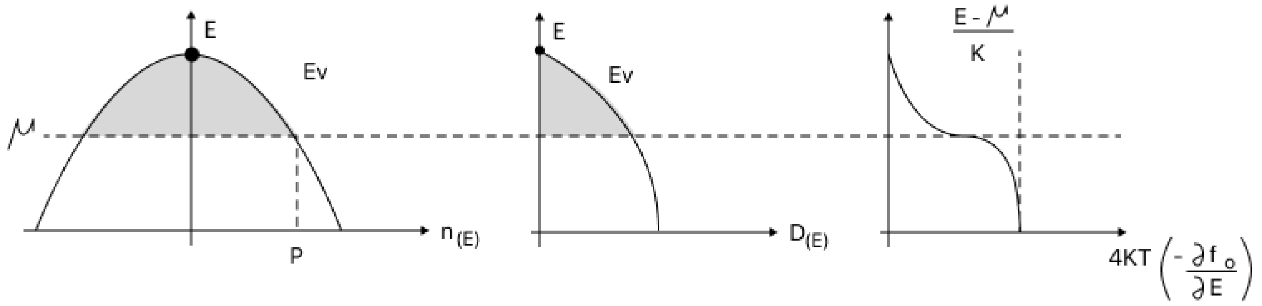


Figure 4.28: Number of empty states above certain energy value (E)

Integrate by parts,

$$N_0 = [-N(\,E)(1 - f_0(\,E)]^{\infty} + \left[\int_0^{\infty} N(\,E)\left[-\frac{\partial f_0}{\partial E}\right] dE\right]$$

The surface term $[-N(\,E)(1 - f_0(\,E)]^{\infty}$ vanishes at both ends and can be ignored, note that when $f_0(E)$ is used instead of $1 - f_0(E)$, the surface term $(Nf_0)$ would not be zero at one end of the limits, then

$$N_0 = \left[\int_0^{\infty} N(\,E)\left[-\frac{\partial f_0}{\partial E}\right] dE\right] \qquad (4.33)$$

It is quite important to note then when dealing with the Drude formula $(\sigma = \frac{q^2 \tau n}{m})$ the value of $(n)$ should be interpreted as the number of empty states in the valence band above $(\mu)$.

## 4.7 The Conductivity (σ) Against the Electron Density (n)

The introduction of electron density in the last section (4.6), gives the capability to connect up the old prospective, the Dude formula, ( $\frac{q^2 \tau n}{m}$), to the new prospective $\left(\sigma = \frac{G_B \lambda}{x}\right)$, where the ballistic conductance $G_B$ in the new prospective plays the same role of electron density $(n)$ in the old prospective. Also, the mean free path $(\lambda)$ in the new prospective corresponds to the mean free time in the old prospective. $(G_B/x)$ gives either of $(G_B)$ per unit (length, area, or volume) according to the dimension $(x)$, we are in.

To connect these perspectives, an expression for the electron density given by equation (4.27) is needed for the sake of comparison;

$$n(P) = \left(\frac{P}{h}\right)^{\text{ȡ}} \left(2, \pi, \frac{4}{3}\pi\right) \qquad (4.27)$$

Now bring in equation (4.31), shown here below;

$$n_0 = \int_{-\infty}^{\infty} n(\,E) \left[-\frac{\partial f_0}{\partial E}\right] dE$$

So, these are the electron density per unit length, area, or volume.

To see if the two conductivity equations of the old and new prospective are equivalent or not, this can be done by the correspondence between the quantities $(\tau\ to\ \lambda)$ and $(G_B/x)\ to\ (n/m)$ of the new prospective to the old one.

The Drude formula is fundamentally correct and works well except in some cases in non-parabolic bands, where some careful treatment is needed for either of mass or the electron density $(n)$.

In the new prospective the ballistic conductance $(G_B/x)$ depends only on the density of states $D(E)$ and the velocity $(V)$, or the number of modes $M(E)$, which tells how many half wave length fits into the cross-section for (3D) or how many wave length fits in the width for (2D), and equal to one in (1D).

Since; $G_B = \frac{q^2 DV}{2L} = \frac{q^2 M}{h}$ ,then using equation (4.24), to find $\frac{G_B}{x}$,

$$M(E) = \left(\frac{P}{h}\right)^{\text{ȡ}-1} (1, 2W, \pi A) \qquad (4.24)$$

$$\frac{G_B}{x} = (q^2/h) \left(\frac{P}{h}\right)^{\text{ȡ}-1} (1, 2, \pi) \qquad (4.34)$$

Where, $(x)$ gives the three dimensions in one shot, then using equation (4.27) to find $\frac{q^2 n}{m}$,

$$n(P) = \left(\frac{P}{h}\right)^{\text{ȡ}} \left(2, \pi, \frac{4}{3}\pi\right) \qquad (4.27)$$

$$\frac{q^2 n}{m} = \frac{q^2}{m}\left(\frac{P}{h}\right)^{\text{ȡ}} \left(2, \pi, \frac{4}{3}\pi\right) \qquad (4.35)$$

Compare equation (4.34) by equation (4.35) above; to compare;

$$\frac{\left(\frac{G_B}{x}\right)}{\frac{q^2 n}{m}} = \frac{m}{P}\left(\frac{1}{2}, \frac{2}{\pi}, \frac{3}{4}\right) \qquad (4.36)$$

Then compare the other ratio $(\lambda/\tau)$;

$$\frac{\lambda}{\tau} = \mathrm{V}\left(2, \pi/2, \frac{4}{3}\right) \qquad (4.37)$$

Multiply equation (4.36) by equation (4.37), to compare the corresponding parameters of the old prospective to the new prospective at the left-hand side of the equation where the result is found in the right-hand side of the equation; as $\frac{mV}{P}(1{,}1{,}1) = 1$, that means they are equivalent, but this is only true in the parabolic bands where the energy goes quadratically with the momentum. It is not necessarily true in general, to how to define mass physically should be taken so carefully, it could be defined as;

$$\frac{dV}{dP} = \frac{1}{m} = \frac{d}{dp}\left(\frac{dE}{dP}\right) = \frac{d^2 E}{dP^2}$$

This definition of mass is found in different context of solid-state physics. This definition is not correct in Dude formula, and should not be used at all, the correct expression is ( $\frac{V}{P} = \frac{1}{m}$ ), and not ( $\frac{dV}{dP} = \frac{1}{m}$ ), these two relations are the same for parabolic bands, but in general they must not be considered the same, because looking back to see how the Drude formula has been derived, there are two quantities required,

(1) The current density ($J = qnv$), which depends on the velocity.

(2) The momentum ($\frac{dP}{dt} + \frac{P}{\tau} = qF$), which comes from Newton's laws, where ($\frac{dP}{dt} = F$), and ($\frac{P}{\tau}$ ) is the frictional force in the solidity viscous drag.

Following the same procedure adopted in section (2) of chapter (2) the Drude formula for conductivity ($\sigma$) is derived;

$$\sigma = \frac{q^2 n\tau}{m}$$

While the derivation of the Drude formula, the mass is introduced as a ratio of $(P/V)$, and $(dP/dV)$ or anything else, so the mass in the Drude formula should be viewed as $(P/V)$.

For non-parabolic bands, the mass depends on what energy -momentum relation $E(P)$ to deal with, because sometimes the mass keeps changing with energy, and $(V/P)$ ratio will not be constant, in such case when using Drude formula care should have to be taken. The formula $(P/V)$ correlates very well with Drude formula, and the ratio between $\left(\frac{G_B}{x}\right)$ $to$ $\left(\frac{q^2 n}{m}\right)$ equals one. When using $\left(\frac{1}{m} = \frac{dV}{dP}\right)$, the density of states $D(E)$ and the number of modes $M(E)$ should be known, all these values are very well defined, but taking care is required when dealing with the number of free electrons and the effective mass in the Drude formula and must know what they mean. If they don't agree, the answer will be found in the new prospective got from Boltzmann's equation, which is the corner stone of the semi-classical transport, and that why people tend to use the Drude formula, because it is easy to learn and derive, not like the new prospective version which require the complicated Boltzmann's equation, but now by the new (bottom-up) approach the new prospective expression ($\sigma = \frac{G_B \lambda}{x}$ ) comes out in a relatively simple way.

## 4.8 Quantum Capacitance

The Nano transistor is the most important electronic device, it is like a resistor in the since that when a voltage is applied between the source and the drain, current flows. For this two terminals device, the density of states of the channel with a chemical potential ($\mu$) to tell the level up to which the states are filled as shown here below;

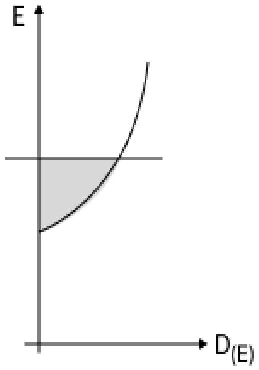


Figure 4.29: Level up to which the states are filled

The valence band (the lower band will be ignored and concerned only with the conduction band, the upper band). The transistor normally has three terminals, the third one is to control the channel, and this gives the essence of the transistor, for this purpose the third terminal will be supplied by a voltage $(V_g)$. Schematically this terminal is called the gate terminal, and it is insulated from the channel by any insulating material. Ideally no current flows through this terminal, but the insulator should be very thin, and that is why there will be an undesirable leakage current. The gate voltage $(V_g)$ changes the potential inside the channel.

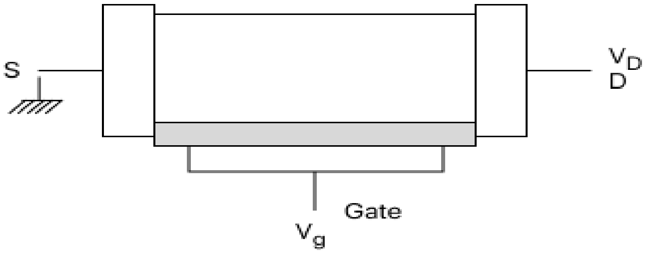


Figure 4.30: Schematic diagram of three terminal nano transistor

The voltage $(V_g)$ moves the energy states in the channel up or down to change the resistance of the channel, hence the energy inside the channel by an amount;

$$U = \beta(-V_g q) + U_0(N - N_0) \quad (4.38)$$

When $(N = N_0)$ the channel is neutral that means all the electrons $(N_0)$ has a compensating positive charge in them (or protons), then;

$$U = \beta(-V_g q)$$

For extra electrons come in the channel will give rise to negative energy, the channel will be out of equilibrium, so the second term appears as in equation (4.38) where $U_0$ is known as single electron charging energy.

Suppose a negative voltage $(-V_g)$ ls put at the gate, that means it adds more electrons to the channel to increase the energy inside the channel by an amount $\beta(-V_g q)$, so all the energy levels float up.

Let the electrochemical potentials in figure below stay where they are before and after adding the gate voltage $(-V_g)$, because these electrochemical potentials are controlled by the contacts;

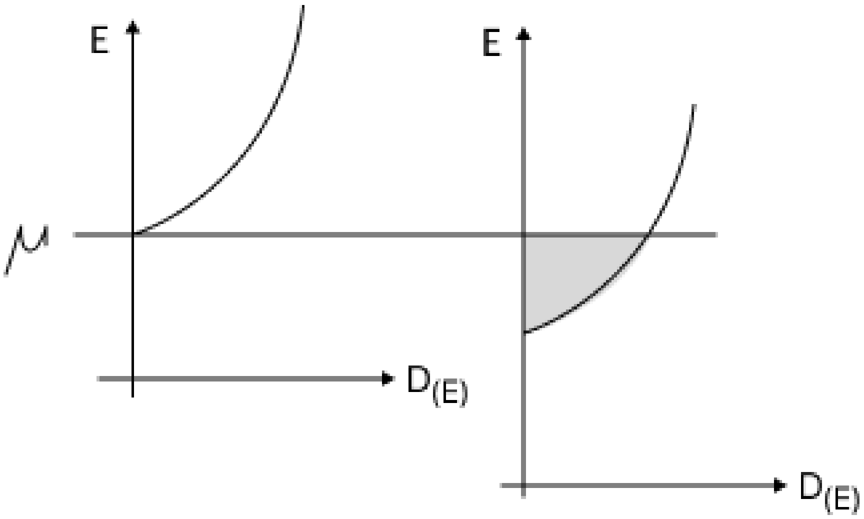


Figure 4.31: How energy levels float up, Left/ After adding $(-V_g)$ ; Right/Before adding $(-V_g)$

Before adding the gate voltage $(-V_g)$ the channel conducts well because there is a lot of states there. After adding the gate voltage $(-V_g)$, the energy levels float up and the density of states are small, and the channel does not conduct as well anymore.

To model this description mathematically, the Fermi-function can be brought in, whose mathematical form is;

$$f_0(E) = \frac{1}{1+e^x}$$

Where, $x = \frac{E-\mu}{KT}$

The Fermi-function tells how the states are occupied. At zero temperature everything below $(\mu)$ is filled, at non-zero temperature the Fermi-function describes the occupation of the levels.

The number of electrons is;

$$N_0 = \int_{-\infty}^{\infty} D(E) f_0(E) dE$$

Where, $D(E)$ tells how many states there are

$f_0(E)$ tells what fraction of the states is filled

When the gate potential changes, as a result the positive potential $(U)$ adds up inside the channel, so the density of states float up, therefore, the electron density can be written as,

$$N = \int_{-\infty}^{\infty} D(\text{È} - U) f_0(\text{È}) dE \quad (4.39)$$

By transformation of variables, let $(E' = E + U)$, then;

$$N = \int_{-\infty}^{\infty} D(E) f_0(E+U) dE \qquad (4.40)$$

There are two cases;

(1) if the density of states floats up, and the electrochemical potential happen to be below the band then the non-generative approximation $\frac{(E-U)}{KT}$ is positive then;

$$f_0(E) = \frac{1}{1+e^x} \approx \frac{1}{e^x} = e^{-x}$$

This what is called Boltzmann's approximation, the Fermi-function where the Fermi-function can be replaced by a simple exponential function;

$$f_0(E+U) = e^{-\frac{E+U-\mu}{KT}} \quad (4.41)$$

Then,

$$N = \int_{-\infty}^{\infty} D(E) e^{-\frac{E+U-\mu}{KT}} dE \qquad (4.42)$$

Pull $(e^{-\frac{U}{KT}})$ out of the integral, because it doesn't depend on energy $(E)$;

$$N = e^{-U/KT} \int_{-\infty}^{\infty} D(E) e^{-\frac{E-\mu}{KT}}\, dE = N_0 e^{-U/KT} \quad (4.43)$$

$$Log_e\left(\frac{N}{N_0}\right) = -\frac{U}{KT} \qquad (4.44)$$

When the gate voltage changes, the energy inside the channel changes by an amount $U = \beta(-qV_g)$ where the best value of $(\beta)$ can be equals to one, but practically always less than one

$$Log_e\left(\frac{N}{N_0}\right) = -\frac{\beta(-qVg)}{KT} \qquad (4.45)$$

Turn the relation to $(V_g)$;

$$Vg = \frac{KT}{q\beta} Loge\left(\frac{N}{N_0}\right) \qquad (4.46)$$

$$dVg = \left(\frac{KT}{q}\right)\left(\frac{1}{\beta}\right)\, d\left(\, Loge\left(\frac{N}{N_0}\right)\right) \qquad (4.47)$$

At room temperature,

$$\frac{KT}{q} = 25mV$$

If the electron density factor $(N/N_0)$ changed by a decade, then $Log_e(10 = 2.3)$, therefore, $\left(V_g\right)$ will change by $(25mV * 2.3 = 60mV)$, this represents the ideal case for the change of $\left(V_g\right)$ by $(60mV/decade)$. If the change per decade is greater than $(60mV)$, the work should be confined to $(\beta)$, in the opposite if the change per decade is less than $(60mV)$ that means $(\beta)$ is greater than $(one)$, and that of course impossible, therefore, either a big mistake has been made or something big has been discovered.

(2). If the density of states floats up, and the electrochemical potential $(\mu)$ happen to be inside the band of the states, where there are enough number of electrons, in such a case the non-generative approximation cannot be used any more instead the original modified form of electron density equation is used;

$$N = \int_{-\infty}^{\infty} D(E) f_0(E+U) dE$$

$$\frac{dN}{dU} = \int_{=-\infty}^{\infty} D(\,E) \left[ -\frac{\partial f_0(E+U)}{\partial E} \right] dE = -D_0 \quad (4.48)$$

$(D_0)$ is a negative quantity, because the derivative of the Fermi-function is negative, that is, as the energy goes up the derivative goes down in the energy range of interest. The quantum capacitance will be given by;

$$\frac{dQ}{dV} = \frac{d(Nq)}{d\left(\frac{U}{q}\right)} = \frac{q^2 dN}{dU} = q^2 D_0 = CQ$$

$$\frac{dN}{dU} = D_0 = \frac{CQ}{q^2} \qquad (4.49)$$

The potential energy inside the channel changes by an amount ( $dU = \beta d(-qV_g)$), this is true as long as charging energy $(U0(N-N0))$ due to the electron's interaction is ignored when there are a lot of electrons. An extra term should be added to compensate for the interaction of each electron with the rest of other electrons.

$$U = \beta\left(-qV_g\right) + U_0(N - N_0)$$

Where, $(U_0)$ is the single electron charging energy.

$(N)$ is the number of electrons increased to the value N

$(N_0)$ is the number of electrons in the neutral state.

Change in the number of electrons $(N - N_0)$ give rise to negative energy in the channel, $(U_0)$ tells how much the potential changes due to one electron in the channel, and the actual change will be given by $(U_0)$ times the change in the number of electrons. When $(N = N_0)$ the channel is neutral and all electrons will compensate the positive charge of the protons in the channel, if extra electrons come in, they give rise to negative energy (out of equilibrium case).

When the number of electrons goes up and need to be lowered still the number of electrons will go up to cause an increase in the overall potential energy. Because the electrons are treated as uncharged particles different pictures other than what is really is, are expected. Any time if there are extra electrons will make it harder for the other electrons to come in and at the same time makes the energy states to float, so the state goes down less. Back to equation (4.38);

$$U = \beta\left(-qV_g\right) + U_0(N - N_0)$$

$$\frac{dU}{d\left(-qV_g\right)} = \beta + \frac{U_0(N - N_0)}{d\left(-qV_g\right)}$$

Using the chain rule;

$$\frac{dU}{d\left(-qV_g\right)} = \beta + U_0\left(\frac{dN}{dU}\right).\left(\frac{dU}{d\left(-qV_g\right)}\right)$$

$$\frac{dU}{d\left(-qV_g\right)} = \beta + U_0(-D_0).\left(\frac{dU}{d\left(-qV_g\right)}\right) = \beta - U_0 D_0 (\frac{dU}{d\left(-qV_g\right)}$$

The quantity $\left(\frac{dU}{d(-qV_g)}\right)$in both sides of the equation, move all to the left side;

$$\frac{dU}{d(-qV_g)} + U_0 D_0 \left(\frac{dU}{d(-qV_g)}\right) = \beta$$

$$\frac{dU}{d(-qV_g)} = \frac{\beta}{1 + U_0 D_0} \qquad (4.50)$$

The change in the potential energy U is due to the change of the gate voltage $(-V_g)$ given by the nonideal factor $(\beta)$, the change has been shown is less than $(\beta)$ depending on how big $(U_0 D_0)$ is, where $(D_0)$ is the density of states at equilibrium and $(U_0)$ is the single electron charging energy. For small values of $(D_0)$, don't worry about this result, but in general the extra term should be included. Another way to write this term is to introduce the electrostatic capacitance.

Since the quantum capacitance is given by equation (4.49);

$$\frac{dN}{dU} = D_0 = \frac{C_Q}{q^2}$$

The electrostatic capacitance is given;

$$U_0 = \frac{q^2}{C_E}$$

$$\frac{dU}{d(-qV_g)} = \frac{\beta}{1 + U_0 D_0} = \frac{\beta}{1 + \frac{C_Q}{q^2} \cdot \frac{q^2}{C_E}} = \frac{\beta C_E}{C_E + C_Q} \quad (4.51)$$

Equation (92) can be visualized as an electric circuit of series capacitors as shown here below;

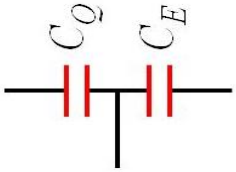


Figure 4.32: Subthreshold gate voltages

Applying the gate voltage $(\beta V_g)$ at the terminal of the electrostatic capacitor $(C_E)$, and the quantum capacitor $\left(C_Q\right)$ is connected to the ground at the end terminal of the source, and from the middle node the potential of the channel is taken. When the density of states in the channel is very low to confirm the non-generative case, and the Fermi-energy level is down in the gap, , $(D_0)$ is small then $\left(C_Q\right)$ is small, so the entire voltage will appears at $(C_E)$ , where essentially in the channel is the applied voltage on the gate $(U/q)$, because small capacitor is a kind of small conductance, or high resistor. On the other hand, the Fermi-energy level $(\mu)$ is inside the band and the density of states are very high, hence $\left(C_Q\right)$ is large, like having small resistance in $\left(C_Q\right)$. So, the potential actually got in the middle node, the channel, is very much less than the ailed voltage $\left(\beta V_g\right)$.

This is the basic framework in terms of which how the gate voltage changes the potential energy in the channel is understood by making the band to float up and down to control the conductance, which of course is the essential physics underlining the operation of the field effect transistors.

## 4.9 Wrapping up summary

This chapter (4) the model for density of states, that is, the energy band structure model is adopted, which is widely used for crystalline solids. The crystalline solid introduces the concept of effective mass and other relatedness things connected with the Drude formula, that is why this chapter started with energy momentum relation $E(P)$ in vacuum, then $N(P)$ which tells how many states are available, whose momentum is less than the value of the momentum $(P)$ in vacuum. The energy in vacuum is related to momentum by addition of the potential energy to the kinetic energy where the kinetic energy is directly proportional to the momentum squared. In complicated solid objects energy can be written in terms of momentum relations with some caveats. One of these caveats the solid could have more than one band. This is one of the early insights in solid states physics. Energy is not necessary to be proportional to the momentum squared it could be non-parabolic complicated function. Specific scheme known as the Periodic Boundary Condition $(PBC)$, for counting how many states are available whose momentum is less than $(P)$, has been adopted. $(PBC)$ is not necessary to be obeyed by real solids, because $(PBC)$ doesn't matter in big solids, and could be used whenever convenient from the mathematical point of view to lead the function $N(P)$. By the combination of $E(P)$ & $N(P)$ to get various quantities of interest, namely the density of states $D(E)$, the number of modes $M(E)$, and the electron density $n(E)$. When these quantities are ready to enter the expression of the conductivity sigma in its new prospective form,

$$\sigma = \frac{G_B \lambda}{A}$$

Hence it will be so easy to connect this new prospective, with the old prospective, the Drude formula;

$$\sigma = q^2 \left(\frac{n}{m}\right) \tau$$

Finally, many things have been prepared to introduce the third terminator the transistor which is going to be discussed in the next chapter, chapter (5).

# Chapter Five
# Nanoscale Transistor

## 5.1 introduction

The transistor is the basic circuit element from which electronic systems are built. The discovery of the transistor effect in 1947 set the stage for a revolution in electronics. The invention of the integrated circuit in 1959 launched the revolution by providing a way to mass produce monolithic circuits of interconnected transistors. As semiconductor technology developed, the number of transistors on an integrated circuit chip doubled each year. This doubling of the number of transistors per chip, driven by continuously downscaling the size of transistors, has continued at about the same pace for more than 50 years. The resulting continuous increase in the capabilities of electronic systems and the continuous decrease in the cost per function have shaped the world of today. The theory of the MOSFET (the most common type of transistor) was formulated in the 1960's when transistor channels were about 10 micrometers (10,000 nanometers) long **[65,66,67]**. *As semiconductor technology matured, transistor dimensions shrunk, new physics became important, and the models evolved*. By the end of the 20th century, transistor dimensions had reached the nanoscale, and the transistor became the first active, nanoscale device in high-volume manufacturing. The flow of electrons in modern transistors is much different from what it was 50 years ago when transistor models were first developed, but most people continue to study traditional MOSEFT theory. The goal of this thesis is to demonstrate the essential operating principles of nano transistors. Nano transistors are much different from those transistors which had been described some decades ago, note that these operating principles are remarkably simple and easy to understand. The approach is based on a new understanding of electron transport that has emerged from research on molecular and nanoscale electronics **[68],** but it retains much of the original theory of the MOSFET. In addition to describing a specific device, this research should serve as an example of how other nanodevices might be understood and modeled. This research is not meant to cover a comprehensive treatment of transistor science and technology; but it could be considered as a starting point aims to convey some important physical fundamentals, assuming an understanding of basic semiconductor physics where a new approach is presented.

## 5.2 The full current voltage characteristic

For the first time in this research the third terminal, (the gate), is going to be introduced

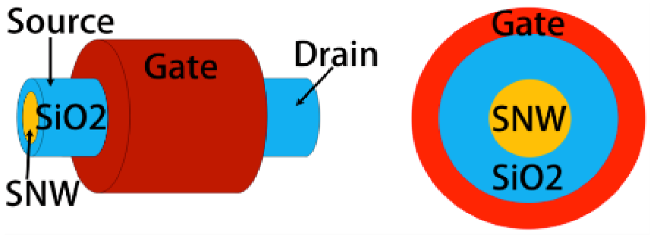


Figure 5.1: Pictorial structure of the nano transistor

The conductance is charged through the gate voltage;

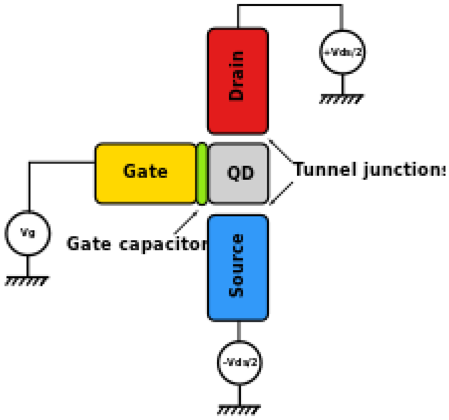
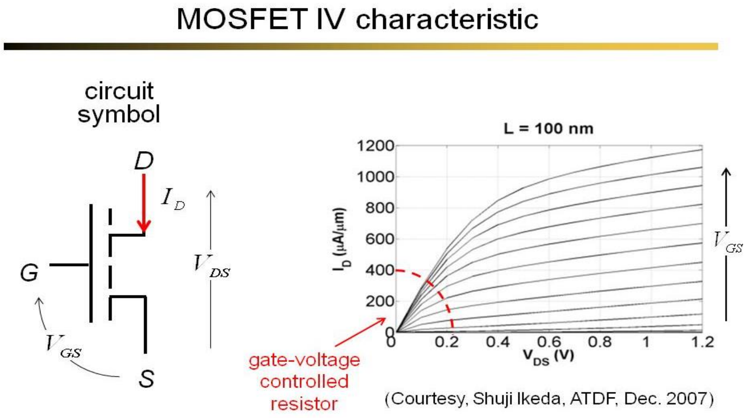


Figure 5.2: Conductance charged by gate voltage

So far everything has been discussed is about the low bias looking right around the origin of the (I-V) characteristic to see the slope which is going to be known as the low bias conductance, the shape of (I-V) curve is an important factor to get familiar of what is going on inside the transistor.



(Courtesy, Shuji Ikeda, ATDF, Dec. 2007)

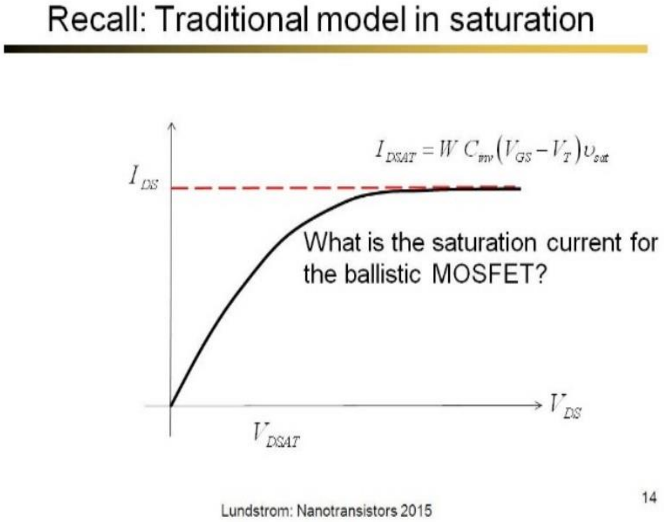
Figure 5.3: Low bias conductance around the origin of IV MOSFET characteristics

The full current against the voltage characteristics right around the origin for small bias, the slope of which gives the conductance, because the characteristics are represented by straight line.

The expression for conductance starts from the full current expression;

$$I = \frac{1}{q}\int_{-\infty}^{\infty} G(E)\big(f_1(E) - f_2(E)\big)\, dE$$

By Taylor series approximation, $\left[\frac{\partial f}{\partial E}\right]$ is deduced, hence the conductance. Using the current equation, the entire voltage current characteristics curve takes the following shape;


Recall: Traditional model in saturation

What is the saturation current for the ballistic MOSFET?

Lundstrom: Nanotransistors 2015

Figure5.4: Voltage current characteristics curve for the ballistic MOSFET

This shape drawn above is not likely to be quite like that, there are some few things to be explained, one of them is why the current saturates so perfectly when the voltage is increased beyond a certain value where the current doesn’t increase any more. The reason is, when applying a voltage, the source is taken as a reference (zero voltage), and the drain has a certain positive voltage $(V_D)$ in it, and let $(V_D)$ to grow bigger and bigger the electrochemical potential in the drain goes down accordingly, and the current flows in the window between $(\mu_1 \& \mu_2)$.

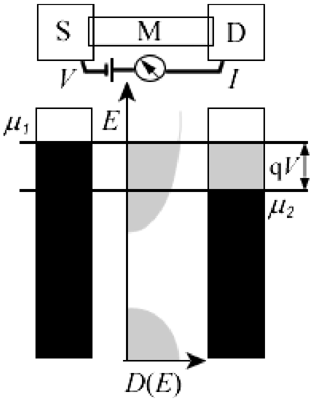


Figure 5.5: Electrochemical potential goes down according to how much the positive applied drain voltage is.

For small voltages more current flows because there are more channels available for conduction. Once ($\mu_2$) dropped below the band no states will be available, it doesn't matter where ($\mu_2$) is located if the empty gap is under the conduction band and ($\mu_1$), that is why saturation is seen in the model. If ($\mu_2$) pulled down further to get into the lower band (the valence band), then again current will start to flow, but usually the voltage range is controlled to be kept at the stage not to reach that end for the current to saturates perfectly.

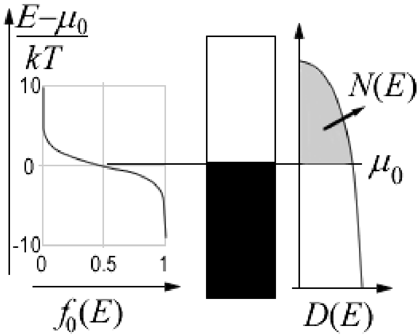


Figure 5.6: Current flow in the lower valence band when the electrochemical potential is pulled below the band gap

If the number of electrons increases in the channel it causes a negative potential, and a positive potential if it decreases. For that reason the band goes down somewhat, because the band is full of electrons and when a voltage is applied the electrochemical potential goes down leaving fewer electrons in the band and the system is out of equilibrium, so one contact wants to fill the channel the other wants to empty it, so the band states are lowered to half of their value because they are used to be full all the time, now they are filled half the time because when one electron comes in another electron goes out, the average results in half the electrons. As a result, there are fewer electrons in the channel to cause a positive potential and makes the band to go down,

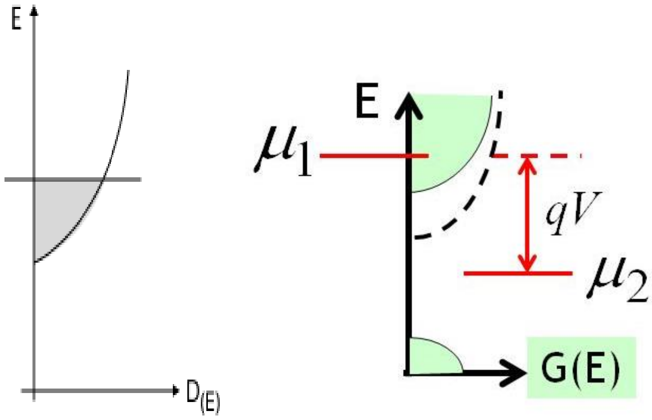


Figure 5.7: Out of equilibrium system due to the reduced number of electrons in the band

That would then give more current flow, what is expected from this effect is, more current flow compared to the current given when this effect is neglected. Mathematically how this effect is explained is by defining the channel potential $(U)$ and rewrite it in the form of something that depends on the gate voltage as shown in the next paragraph self-consistent model **[69,70].**

## 5.3 Self consistent model

The positive potential applied would make the band to go down and then more current flows over what is expected. Mathematically define the channel potential $(U)$ which depends on the gate voltage given by the following relation;

$$U = \beta\left(-qV_g\right) + U_0(N - N_0) \qquad (5.1)$$

The first term doesn't matter because of the assumption that the gate voltage is fixed, the second term stands for the discharging energy. The effect of the bias makes the number of electrons to go down by half of they are used to be, so $(N)$ goes down and $(N_0)$ becomes negative, adding this effect to the current equation then,

$$I = \frac{1}{q}\int_{-\infty}^{\infty} G(E - U)\big(f_1(E) - f_2(E)\big)\, dE \qquad (5.2)$$

The number of electrons;

$$N = \int_{-\infty}^{\infty} G(E-U)\,\frac{\left(f_1(E)+f_2(E)\right)}{2}\;dE \qquad (5.3)$$

Out of equilibrium there are two Fermi-functions, one at the right contact and the other at the left contact, therefore, the occupation of the states is average of the two functions.

These two equations can be solved self consistently, because $(U)$ depends on $(N)$, and $(N)$ depends on $(U)$. There is a numerical technique for that solution. If they are solved self consistently results in a much bigger current than what is expected to be.

If the band lowering process is ignored $(U_0 = 0)$, but when the band lowering process is taken into account much bigger current flows and $(N - N_0)$ tends to be small as if the band floats down, so the number of electrons used to be will try to go back to half of its value, but instead it goes down enough and almost stay as it is. Under these conditions again more current flows like what is shown by $(U_0 = \infty)$

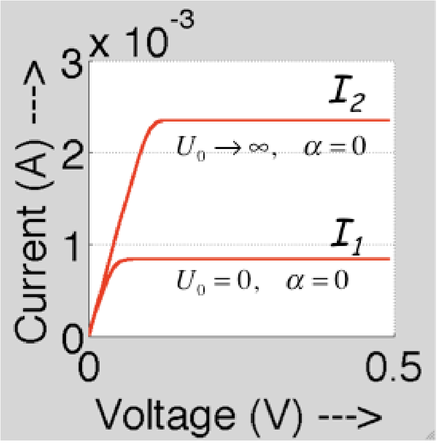


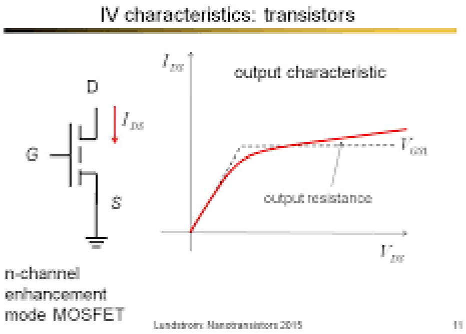
Figure 5.8: Band lowering process results in bigger current flow

## 5.4 The Role of Electrostatics

Another factor is missing in this discussion, that is why the saturation is very perfect, although this a very good thing if the transistor has got such perfect saturation at the on state, this will help the transistor user to consider the on state is independent of the voltage $(V_D)$, and this is what really good transistors are likely to be to control the current entirely by the gate voltage and is not affected by the voltage between the source and the drain. This model seems to predict what has been said so easily, but that is not true completely because an important effect of the potential inside the channel has been ignored, so the channel is not affected by the gate voltage alone but also by the drain voltage which should be included in equation (5.1);

$$U = \beta\left(-qV_g\right) + U_0(N - N_0) + \alpha\,(-qV_D) \qquad (5.4)$$

When $(V_D)$ is included in equation (5.1), the current does not saturate so nicely as before, physically this true when $(V_D)$ is increased, because the channel potential tends to go down more hence the band sinks down more, and the current keeps increasing.



IV characteristics: transistors

Lundstrom: Nanotransistors 2015

11

Figure 5.9: Physical operation IV characteristics of nano transistor

To get good current saturation in the Nano transistor, the art of technology is, $(\alpha)$ should be kept as small as possible. As long as the long devices are concerned nice saturation can be achieved, because the gate voltage controls the potential through $(\beta)$, and the drain potential is charged through $(\alpha)$, as the devices gets smaller enough and the channel is about few hundred atoms across the drain voltage $(V_D)$ get more effective in terms of controlling the channel potential, this effect is sometimes called Drain Induced Barrier Lowering $(DIBL)$ **[71].**

In order to make sure that the gate controls the potential better than the drain that means $(\beta)$ is a whole lot bigger than $(\alpha)$, the insulator should have to be very few atoms (more like ten atoms), and that is extremely hard to do, because there will be more leakage currents the electrons tunnel is through the insulator, though ten atoms across billion devices is really hard job to do to make billion Nano transistors to work, any way from the qualitative point of view the insulator should be very thin for the gate to control the channel potential much better than the drain do. On the other hand, from the quantitative point of view this model is not adequate, much deeper knowledge in some other issues is required, the most important one of them is putting the channel potential as a single number.

Practically the potential varies along the channel. In other words, the value of the potential at the source end is different from the potential at the other drain end. It changes continuously. That is, the density of states $D(E)$ or the conductance function $G(E)$ will keep going down, when the channel potential $(U)$ changes, the bottom of the band tracks that change by going down along the channel length that means the channel potential $(U)$ is not a single number.

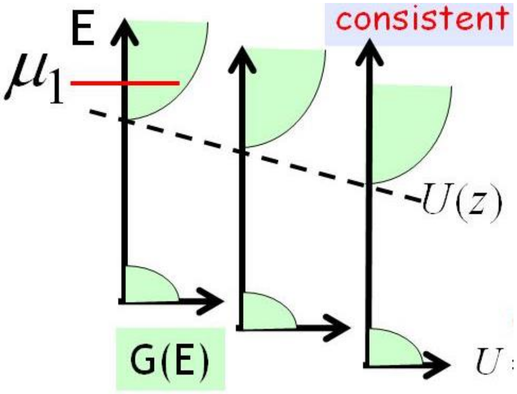


Figure 5.10: The track of the bottom of the band

## 5.5 Model extension

To calculate the channel potential U a differential equation is needed to do that calculation. That differential equation is the Poisson Equation

$$\left[\frac{d}{dz}\left(\varepsilon\frac{dU}{dz}\right) = q^2(n - n_0)\right] \qquad (5.5)$$

Doing a serious device simulation Poisson Equation should be used which tells the potential $(dU/dZ)$ as a function of the electron density $(n - n_0)$, and what have been done is a poor algebraic version of equation (5.4),

$$U = \beta(-qVg) + U_0(N - N_0) + \alpha\,(-qVD) \qquad (5.4)$$

From equation (96) a single number $(U)$ is got, just to give a rough idea of what are those factors in the equation look like, to understand the real observed characteristics or the physical features well. One thing still missed in the model shown by equations (5.2,5.3), that is, the current and the number of electrons equations, were based on the elastic cross transport without any change of energy, to give a good approximation for ballistic devices.

For diffusive transport the electron losses momentum at each stage as it moves down along the potential gradient from the source to carry on to the drain and will not be confined to the energy channel like what happens in the ballistic case. If the model allows electrons to travel along the channel, they start at an energy, where if the current is calculated results in much lower current than what is actually should be, because in practice, some of the elastic processes allow an electron to relax, so the current increases because an electron from the source goes down the energy slope, then it is so hard for that electron to get back to the source, and continue to proceed on, to the drain. This description can be done in a more convincing qualitative way by segmenting the long device into a lot of little ones in series. What is usually done for device modeling techniques is the usage of the drift diffusion equation, which is written as, current directly proportional to the slope of the electrochemical potential;

$$I = -\frac{\sigma(Z)d\mu}{qdZ} \qquad (5.6)$$

Usually the electrochemical potentials are considered at the source and the drain ends without considering what is going on in between. If the electrochemical potential in between is assumed to be $(\mu)$, and it changes along the way it could be written as $\left(\frac{d\mu}{dZ}\right)$ as shown above in the drift diffusion equation (5.6) that is widely used in for device simulation, and it should be coupled by the continuity equation in one dimension, and the current does not change with length;

$$\frac{dI}{dZ} = 0 \qquad (5.7)$$

The equations (5.5, 5.6, 5.7) are old equation generally used, and they are not based on new prospective sight of view, the old prospective conductivity $(\sigma)$ is given by the Drude formula, whereas the new prospective has added new understanding of the conductivity $(\sigma)$ plus a new understanding of an extra ballistic resistance. The conductance from the freshman's physics or the old prospective is given by the usual Ohm's law,

$$G = \frac{\sigma A}{L}$$

What had been seen in the new prospective, the conductance is given by,

$$G = \frac{G_B \lambda}{1 + \lambda} \rightarrow \left(\frac{\sigma A}{L + \lambda}\right) \quad (5.8)$$

When the length of the conductor turns to zero, the conductance doesn't go to infinity but rather goes to a fixed number, if this relation is turned around the resistance can be written as,

$$R = R_B \left(\frac{L + \lambda}{\lambda}\right) = \frac{R_B L}{\lambda} + R_B$$

The first term of the equation$\left(\frac{R_B L}{\lambda}\right)$, follows Ohm's law, the second term of equation $(R_B)$ is constant, in the early days, the major controversy was about what this constant is. To understand where this constant comes from just look back to equation (5.8) to see its origin. This constant resistance $(R_B)$ is associated with the contacts interfaces, since it is independent of the length, the middle part of the model is the channel and has a resistance that it is directly proportional to the length of the channel $(L)$. This is what the new prospective added to the old understanding is this important idea of interface resistance. Although it is very well known always there are some contact resistances, but they are considered as an undesirable thing to get rid of it by good engineering efforts. What is not recognized there is a fundamental limit, where, there is always the ballistic resistances appearing at the interfaces which cannot be avoided whatever how good technology efforts are done to get rid of them. But still an answer will be needed where these interface resistances come from? the answer to this question is found in the next chapter, and that is why when the drift diffusion equation is used the user should be careful to add these interface resistances.

### 5.6 Wrapping up summery

In this chapter the third terminal of the transistor (the gate) has been introduced to the ordinary two terminal device (the resistor) to give the essence of the transistor action. When the gate terminal is introduced the potential energy inside the channel is controlled by this gate terminal, so that conduction band energy $(E)$, can be moved up or down inside the channel by an amount $(U)$ due the electron charging or electron to electron interaction to change the resistance of the channel. When a positive voltage $(V_D)$ is applied to the drain, all the energy level in the drain $(\mu_2)$ are pushed down by an amount $(-qV_D)$ relative to the electrochemical potential of the source $(\mu_1)$, at the same time the conduction band energy levels inside the channel move down by an amount $(-U)$ due to (the negative electron charging), if the gate voltage $\left(V_g\right)$ is positive, or move up by an amount $(+U)$ due (the positive electron charging),therefore, $(U)$ depends on the gate potential ,the drain potential, and the change in electron density inside the channel. This electron charging (or the electron -electron interaction) should be included in the current equation (5.2), brought here below;

$$I = \frac{1}{q}\int_{-\infty}^{\infty} G(E-U)\big(f_1(E) - f_2(E)\big)\, dE \qquad (5.2)$$

The conductance function $G(E)$ is shifted by an amount $(U)$, this $(U)$ is a function of $(N)$; see equation (5.4), brought here below;

$$U = \beta(-qVg) + U_0(N - N_0) + \alpha\,(-qVD) \qquad (5.4)$$

There is another expression for the number of electrons as a function of $(U)$, where $(N)$ depends on the density of states and how it is occupied, see equation (5.3), brought here below;

$$N = \int_{-\infty}^{\infty} G(E-U)\,\frac{\big(f_1(E) + f_2(E)\big)}{2}\; dE \qquad (5.3)$$

Since $(N)$ is a function of $(U)$, and $(U)$ is a function of $(N)$, these two equations (5.3&5.4) in general can be solved self consistently. The simultaneous solution of these two equations can be done in straightforward way numerically, and sometimes in simpler cases can be solved even analytically. All the devices simulation usually involves this self-consistent calculation. One part of it focuses on the density of states and the other part describes the electrostatics of the transistor problem. $(U_0)$ is the single electron charging energy, it tells how much the potential energy changes by adding one electron, may be the potential changes by $(1\mu eV)$, but for big devices still this could be sizable under operation conditions, because the change $(N - N_0)$ might reach a high number in the order of 10,000 or 100,000 electrons which causes a change of (10mV or 100mV). For small devices one electron makes much bigger difference in terms of what potential is. The potential energy could be significantly affected because $(N - N_0)$ could be relatively big, and that leads to a new class of phenomena called the single electron charging effect or Colombo blockade, these are important when $(U_0)$ is relatively large number in the order of $(KT)$ or more.

To understand the operation of the transistor the model $(U)$

$$U = \beta(-qVg) + U_0(N - N_0) + \alpha\,(-qVD) \qquad (5.4)$$

Could be used to see how the gate voltage controls it, and how current voltage characteristics look like etc. But if the model is extended beyond that the potential inside the channel should not be taken as a single number but varies specially. This part should not be ignored for more serious model and should be included by replacing the algebraic $(U - equation)$ to a differential equation using Poisson's equation to describe the electrostatics of the transistor,

$$\left[\frac{d}{dz}\left(\varepsilon\frac{dU}{dz}\right) = q^2(n - n_0)\right] \qquad (5.5)$$

Sticking to the basic electrostatics keeps the subject away from Poisson equation corrections based on the exchange and correlations to avoid the various quantum effects that scientists are talking about. For the current equation (5.2), the electron number equation (5.3), and so on .. etc.

$$I = \frac{1}{q}\int_{-\infty}^{\infty} G(E - U)\big(f_1(E) - f_2(E)\big)\,dE \qquad (5.2)$$

$$N = \int_{-\infty}^{\infty} G(E - U) \, \frac{\big(f_1(E) + f_2(E)\big)}{2} \; dE \qquad (5.3)$$

There would something like the drift diffusion equation, says the charging current is directly proportional to the slope of electrochemical potential inside the channel, by which the current can be calculated, in connection with the continuity equation, which says; in one dimensional device the current does not change spatially, it is constant.

$$I = -\frac{\sigma(Z)d\mu}{qdZ} \qquad (5.6)$$

$$\frac{dI}{dZ} = 0 \qquad (5.7)$$

All these equations (*Poisson, drift diffusion, and the continuity*) equations are based on the semi-classical transport Boltzmann's Transport Equation $(BTE)$, which is corner stone of all semi-classical theory people are using long time ago before nanoelectronics evolved. From nanoelectronics a new thing has been gained, that the conductance not more than four decades was,

$$G = \frac{\sigma A}{L}$$

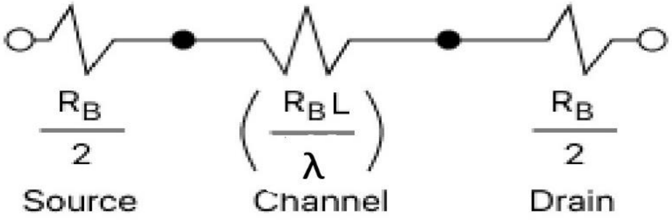
But today as L approaches zero the conductance turns to a constant known as the ballistic conductance $(G_B)$, so $(L + \lambda)$ is needed to be substituted for the length $(L)$ to give the conductance,

$$G = \frac{G_B \lambda}{L + \lambda} \rightarrow \left(\frac{\sigma A}{L + \lambda}\right) \qquad (5.8)$$

By inverting the resistance could be written as;

$$R = R_B \left(\frac{L + \lambda}{\lambda}\right) = \frac{R_B L}{\lambda} + R_B$$

Therefore, part of the resistance is directly proportional to the length of the channel $(L)$, the other part is the constant ballistic resistance. The origin of this ballistic resistance comes out of the Boltzmann's Transport Equation $(BTE)$ in the sense that it all hidden there, because people were unaware of that till nanoelectronics emerged to understand that the first part of the resistance which is directly proportional to the channel length could be viewed as the channel resistance, whereas, the constant part $(R_B)$ can be associated with the interface resistances at the contacts.



For the device simulation, the drift diffusion equation is used. The important piece of new knowledge that should be added the interfaces resistances. If using $(BTE)$, and did it correctly, the interfaces resistances $(R_B/2)$ would come out automatically, when using the drift diffusion equation addition of the interface's resistances should have to be remembered. The important thing that everyone must realize is; always there is a contact or interface resistance for any conductor especially short ones, and there is a fundamental limit to that. To get rid of the contact resistances to improve the quality of the conductors was the usual thinking in the past, but today the fundamental limit represented by the ballistic resistance is discovered, hence the roadmap to the future devices should have to take these fundamental limits into account. Still there is something has not been discussed till now, that is how to realize that the extra resistance ($R_B$) is associated with the interfaces, so as to understand that, the potential change inside the device should be studied. This what is going to be done in the next chapter (6).

# Chapter Six

# Potential Change inside the channel

## 6.1 Introduction

So far, a channel with a voltage across it has been discussed to write the current and the resistances in the channel for small voltages as shown here below;

$$I = G_B V$$

$$G_0 = G_B \left( \frac{\lambda}{1+\lambda} \right)$$

$$R_B = \frac{1}{G_B}$$

$$R_0 = \frac{1}{G_0} = R_B + \frac{R_B L}{\lambda}$$

The second term of $(R_0)$, which is directly proportional to the length, can be associated with the central region (the channel), the longer the channel the bigger the resistance. The first term is a constant part (the ballistic resistance), can be associated with the contacts, but this decision is not quite obvious, because resistances used to be associated with heat when a current flow through them, but that thinking is very misleading if adopted to clarify that $(R_B)$ can be associated with the contacts. For small devices the heating is not in the channel, it is entirely in the contacts, because the channel is assumed to be elastic and has got the characteristic of Landauer resistance. If the resistance is assumed to be everywhere, this point of view goes against something felt instinctively it is not going with the intuitive feeling. Let us increase the resistance of the channel by making a hole in the middle of the channel, but that doesn't mean heating will be associated with that resistance occurred in the hole, because the hole haven't got the degree of freedom needed to dissipate heat from the electrons going into the atoms to let them vibrate so as to convert their thermal energy into heat. Therefore, heating occurs somewhere else. It is quite wrong to associate that heat location with the location of the resistance; the right criterion is to follow the voltage to locate where the voltage drop takes place. When a current run through a set of resistances in series, the voltage drops across each one of them is directly proportional to that resistance. By measuring the voltage at the interfaces, a voltage drops are found there, and that is why $(R_B)$ can be located at the interfaces.

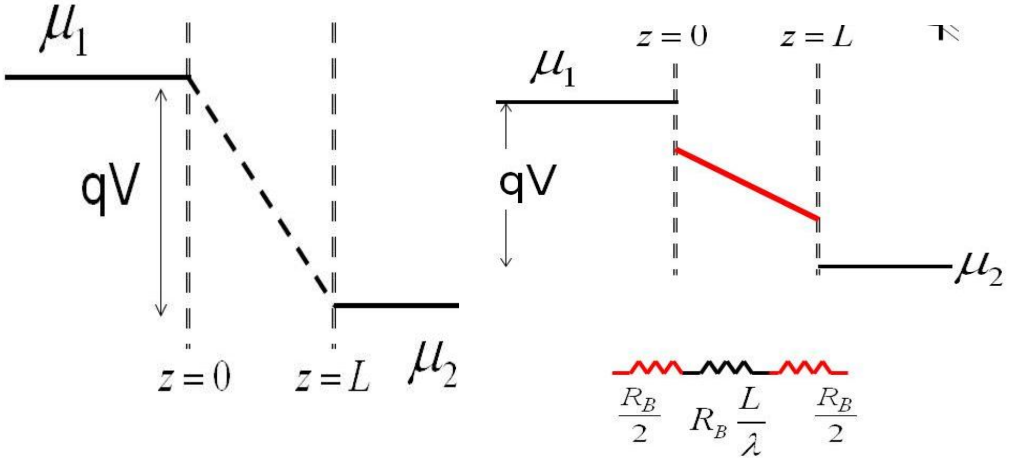


Figure 6.1: Resistance associated with voltage drop

The contacts have two different values of the electrochemical potentials $(\mu_1 \;\&\; \mu_2)$, because the source potential is negative relative to the drain potential, the negative side will be higher up, and they are separated by the applied voltage by an energy amount $(qV)$ , then the electrochemical potential vary spatially inside the channel from one contact to the other. The actual voltage drop looks like the bold line shown in the figure above, it is linear, but doesn't go directly from $\mu_1$ to $\mu_2$, rather there is a drop at the interfaces, these drops are associated with the ballistic resistance $(R_B)$.

What is drawn is the electrochemical potential which causes a lot of confusion and comes up with a deep question, namely what is meaning of voltage? Simply the answer will be, it is an electrostatic potential, because the electrochemical potential is a deep subtle concept, normally felt uncomfortable due to its statistical mechanics concept, but that answer is not quite right if the current is described by the electric field, where the electric field is directly proportional to the slope of the electrostatic potential energy $\left(\frac{dU}{dZ}\right)$ ,

$$J = \sigma_0 E \rightarrow J = \frac{\sigma_0}{q}\frac{dU}{dZ} \quad (6.1)$$

On the other hand, the correct answer that all experts agree with is that the current is directly proportional to the slope of the electrochemical potential ($\mu$), and not to the slope of electrostatic energy potential $\left(\frac{dU}{dZ}\right)$;

$$J = \frac{\sigma_0}{q}\frac{d\mu}{dZ} \quad (6.2)$$

This result needs a little more elaboration. Consider a channel structure with the density of states distributed at low temperature as shown by the figure below, this is what goes into the Fermi-function,

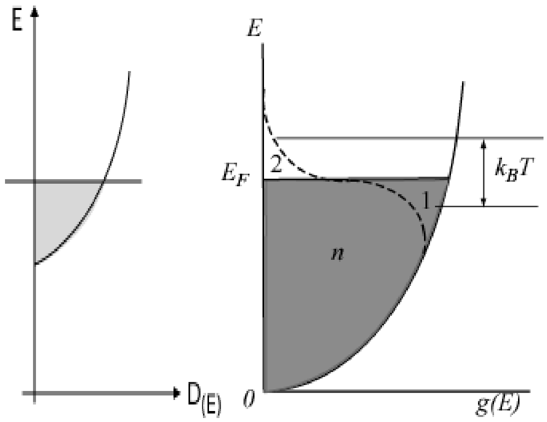


Figure 6.2: Left/ A channel structure with the density of states distributed at low temperature Right/What goes into Fermi function

Suppose there is an electrostatic potential that varies across the channel, the lowest energy point of the band is expected to go down continually,

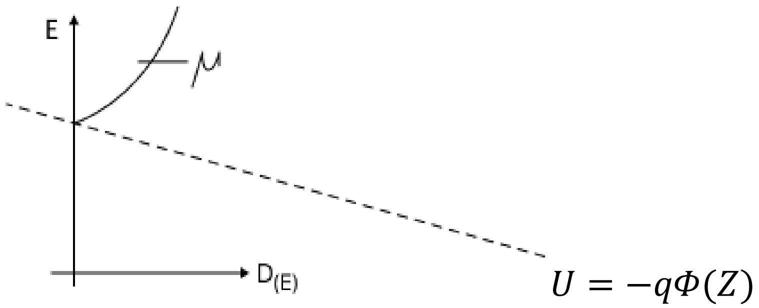


Figure 6.3: An electrostatic potential that varies across the channel

because there is a positive electrostatic potential and negative electron charge, that means negative energy $U = -q\Phi(z)$. The density of states plot would look as shown here below;

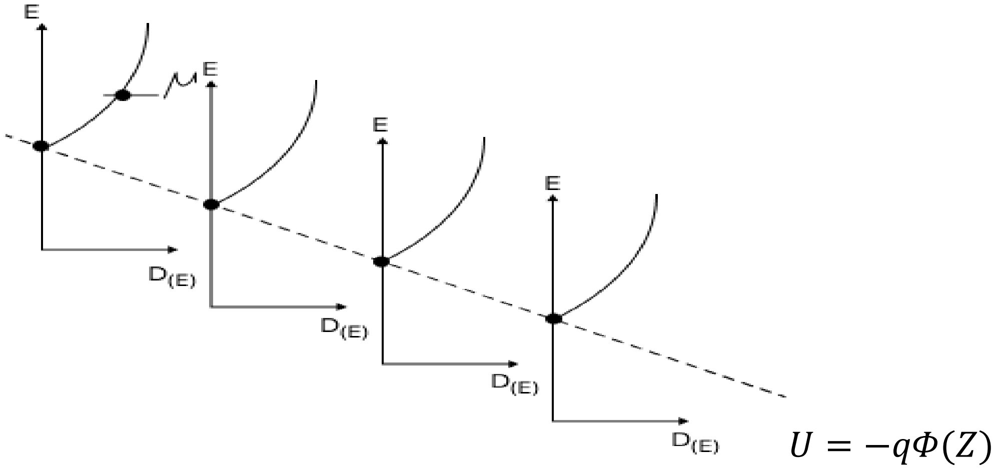


Figure 6.4: The density of states plot follows the negative energy $U = -q\Phi(z)$.

In homogenous conductors whose electron density is the same everywhere, and the electron density depends on the energy range between the bottom of the band and the electrochemical potential μ. The current equation; $\left(J = \frac{\sigma_0}{q}\frac{dU}{dZ}\right)$, being proportional to the electric field is not quite true unless the electrochemical potential goes down also, or only if $U = -q\Phi(Z)$ is parallel to $\mu(Z)$. Like what has been shown here below,

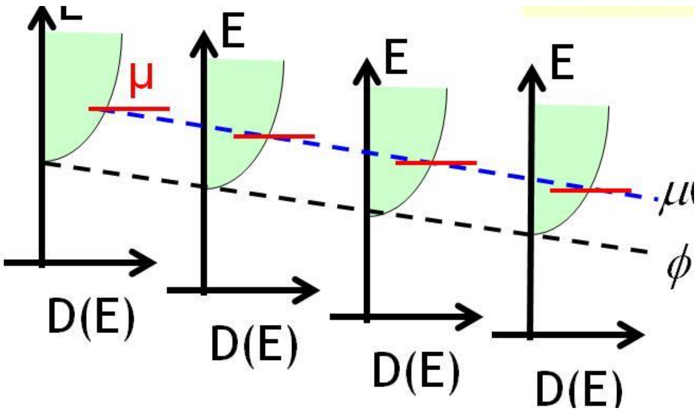


Figure 6.5: Energy range between the bottom of the band and the electrochemical potential μ.

For an inhomogeneous conductor there are more electrons in one side than in the other side. At equilibrium the electrochemical potential is constant with higher electron density on the right side than in the left side and there is no current flow because the system is at equilibrium. The explanation of this condition is by the addition of the diffusion current to the drift current as follows;

$$J = -\frac{\sigma_0}{q}\frac{dU}{dZ} - q\frac{dn}{dZ} \quad (6.3)$$

Where the first term is the drift current while the second term is the diffusion current represented by the slope of the electron density. From the following figure the electron density is seen due to the smaller density of electron at the left side than in the right side.

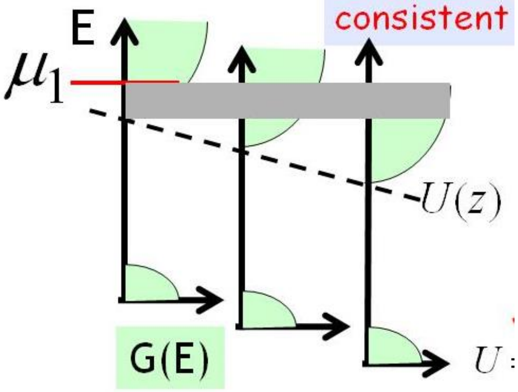


Figure 6.6: The electron density is much smaller at the left side than in the right side.

The diffusion current can be rewritten in the following form;

$$-q\frac{dn}{dZ} = -\frac{\sigma_0}{q}\frac{d(\mu - U)}{dZ}$$

Where $(\mu - U)$ is the distance between the electrochemical potential and the electrostatic potential that varies across the channel in figure above, which is directly proportional to the electron density $(-\frac{dn}{dZ})$. That is just as $\left(\frac{dU}{dZ}\right)$ in the drift current, could be written as $\left(\frac{d(\mu-U)}{dZ}\right)$ in the diffusion current.

To add the diffusion current to the drift current results in the following correct expression for the current;

$$J = -\frac{\sigma_0}{q}\frac{dU}{dZ} - \frac{\sigma_0}{q}\frac{d(\mu - U)}{dZ} \quad (6.4)$$

$$J = \frac{-\sigma_0}{q}\frac{d\mu}{dZ} \quad (6.5)$$

If the first term out of this equation is used, the current is directly proportional to the electrostatic potential, could that be right or wrong? The answer of this question could be completely right at the same time it could be completely wrong, that means there is something important has been missed. But there is nothing controversy about the correctness of the usage of that term to calculate the current, because all scientists and experts agree that is the correct answer. Most of the experts avoid talking about the second term of the equation, because $(\mu)$ is changing continuously and very difficult to define under none equilibrium conditions. At equilibrium conditions $(\mu)$ is the same everywhere (constant), exactly like what heat is doing, it the same everywhere at equilibrium, but when the temperature is varying the concept gets more subtle.

In conclusion at equilibrium $(\mu)$ is the same everywhere, when it is out of equilibrium it is very difficult to define, because it leads to more subtle issues, and that why most of the experts try to avoid it, instead they prefer to talk about the electrostatic potential.

Considering the electrostatic potential, it doesn’t quite show the discontinuity at the contacts associated with the ballistic resistances. That part of physics is going to be lost, because it will be smoothed out as shown in the following figure;

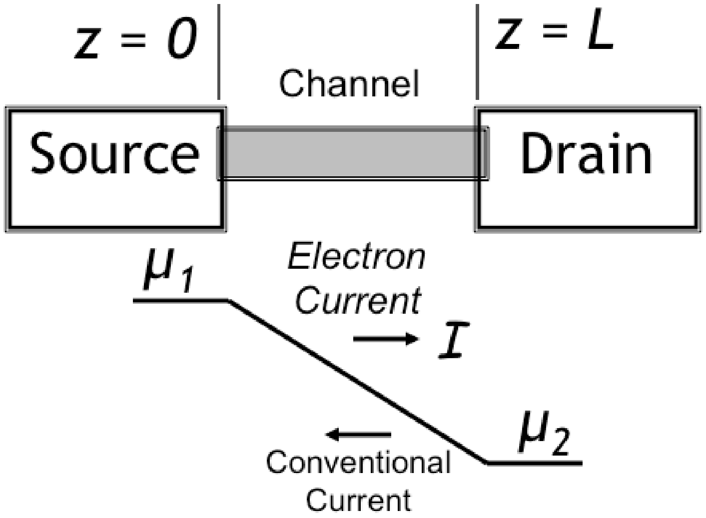


Figure 6.7: With ballistic resistances, the electrostatic potential at the contacts smoothed out

The actual quantitative calculation based on quantum transport formalism; the electrochemical potential is discontinuous while the electrostatic potential is smoothed out. The electrochemical potential inside the channel should be defined with care, this was done on the bases of quantum transport based on NEGF calculations on (*McLennan et Phys-Rev B 4313846(1991)).*

But in this dissertation the electrochemical potential will be defined in the semi classical context in terms of Boltzmann's Equation to show the concept of the Quasi- Fermi levels **[72,73,74,75].**

There are many electrochemical potentials when the system is out of equilibrium, in this context as if there are right moving electrons and left moving ones, with different electrochemical potentials represented by $(\mu^{+})$ and $(\mu^{-})$ respectively, known as the quasi Fermi levels. If a single electrochemical potential is drawn in between it represents the average of the two. This is something has been described during the last half of last century.

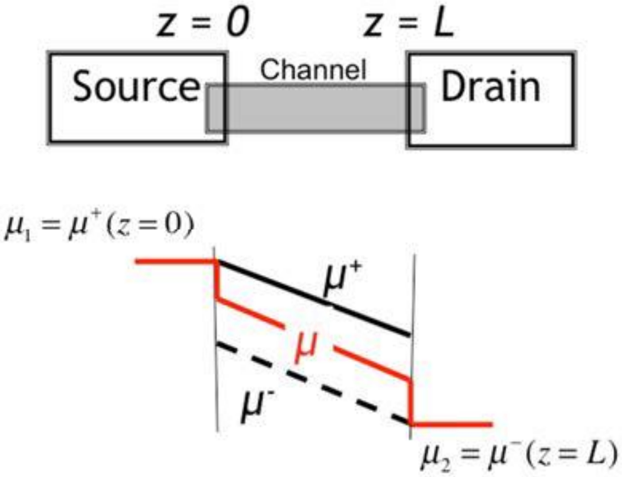


Figure 6.8: The quasi Fermi levels with different electrochemical potentials represented by $(\mu^+)$ and $(\mu^-)$.

After the enormous progress in spintronics the discussion of these issues became much more important, specifically in TOPOLOGY INSULATORS, in such materials there is something called SPIN-MOMENTUM LOCKING in which electrons moving to right direction their momentum is considered to spin upwards, while those moving to the left direction are spinning downwards, this is what is known as the spin-momentum locking.

Because in the ordinary conductors as the electrons are moving to the right direction, they could have either to spin up or down, not like the topological insulator materials which lock together in each direction, that is, all of them spin up or down according to what direction they are moving to. That is why right moving carriers $(\mu^+)$ are different from the left moving carriers $(\mu^-)$. This difference appears as SPIN POTENTIAL, and that is why they are different from one another in their quasi Fermi levels. The spin can be measured using a magnetic probe depending on the direction of the magnet. While looking for the up spin or the down spin in a certain region of the channel the (spin potential) decreases rapidly in the right direction with the up spin $(\mu^+)$, and the spin potential increases rapidly in the right direction with the down spin $(\mu^-)$

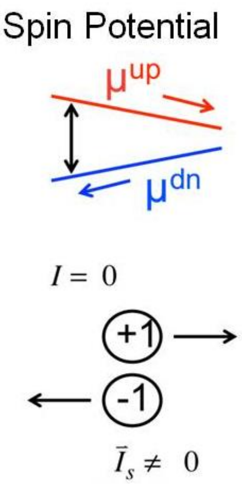


Figure 6.9: Spin potential behavior

This phenomenon can be explained by saying the current depends on the spin potential gradient, that means the up-spin carriers diffuses to the right direction and the down spin carriers diffuse to the left direction. This phenomenon is so difficult to be explained in terms of the electrostatic potential, because there is only one electrostatic potential either to slope to right direction or to the left. It is very hard to understand how one type of carriers could be going right and the other type of carriers going left. To study nanoelectronics devices it very important to understand what potential means, this idea of quasi Fermi levels, and the different species can have different quasi Fermi levels.

Considering the new boundary conditions, if required to locate where the resistance is, it is very important not to think of the heat to be associated with the resistance. That will not get you to anywhere, so it must be better to follow the voltage drop. In this case the knowledge of what does the voltage mean should be clearly understood by looking to the electrochemical potential and not to the electrostatic potential, in the sense that how does it vary across the device. Start by noting that the contacts have got two different values of electrochemical potentials $(\mu_1 and\ \mu_2)$, due to the negative voltage in the source relative to the drain, the negative side (the source) is higher up, and separated from the other side (the drain) due the applied voltage by an energy amount of $(qV)$.

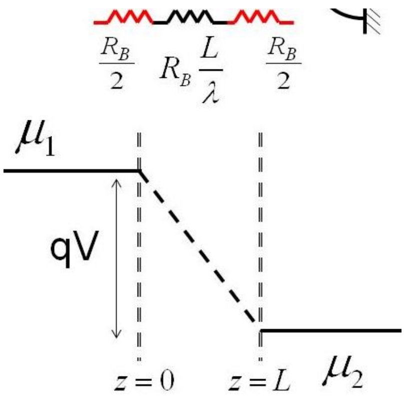


Figure 6.10: Spatial variation of the electrochemical potential inside the channel

But how the electrochemical potential varies spatially inside the channel?

Let $(\mu)$ be the electrochemical potential, because, the following diffusion equation, includes both the drifting diffusion as shown by the deduction of equation (105) above;

$$I = -\frac{\sigma_0 A}{q}\frac{d\mu}{dZ} \quad (6.6)$$

This basic equation says the current is directly proportional to the slope of the electrochemical potential. To solve this equation, the continuity equation, $\left(\frac{dI}{dZ} = 0\right)$ is needed. This continuity equation is a $(1D)$ equation and the current at steady state $\left(\frac{dI}{dZ} = 0\right)$ should be spatially constant. That is to say, the current is constant everywhere, at different points along the channel, whatever comes in per second must be leaving per second. Since the current is constant spatially, the out front $(\frac{\sigma_0 A}{q})$ is constant, to indicate that the slope of the electrochemical potential $\left(\frac{d\mu}{dZ}\right)$, must be constant, and goes in straight line from $(\mu_1\ to\ \mu_2)$ over a distance $(L)$, with a negative sign because it is going down. Therefore, the slope can be written as;

$$-\frac{d\mu}{dZ} = \frac{\mu_1 - \mu_2}{L}$$

Substituting this relation in equation (106), to calculate the current;

$$I = -\frac{\sigma_0 A}{q} \frac{(\mu_1 - \mu_2)}{L} \quad (6.7)$$

Using equation (100) the current can be calculated as;

$$I = \frac{\sigma_0 A}{q} \frac{(\mu_1 - \mu_2)}{(L + \lambda)} \quad (6.8)$$

Equations (107) and (108) are the same except that λ is missed in equation (107), because the discussion is about the ballistic transport, and the diffusion equation is used for the discussion. This reasoning could be right, and in that sense basically the wrong equation is used. Therefore, a different equation should be looked for, but that is may not be the case. The equation may be the right one, but the boundary conditions are wrong, because of the addition of the ballistic resistor $(R_B)$ in the diffusive equation (6.6). To get rid of the ballistic resistor out of the standard diffusion equation is by modifying the boundary conditions**.**

## 6.2 The New Boundary Conditions

In the previous paragraph the boundary conditions used are, $\mu(z = 0)\ is\ \mu_1$, and $\mu(z = L)\ is\ \mu_2$, then straight line is drawn between ($\mu_1$ & $\mu_2$) as required by the diffusion equation (6.6) and the continuity equation $\left(\frac{dI}{dZ} = 0\right)$. The new boundary conditions suggested are,

$$\mu(z = 0) = \mu_1 - \frac{qIR_B}{2} \quad (6.9)$$

$$\mu(z = L) = \mu_2 + \frac{qIR_B}{2} \quad (6.10)$$

A straight line is drawn inside the channel, the line is still straight line as required by the standard diffusion equation (6.6), but depending on the amount of current, it starts somewhat below $\mu_1$, and ends slightly above ( $\mu_2$). The new slope;

$$\frac{d\mu}{dZ} = \frac{\mu_1 - \mu_2 - qIR_B}{L}$$

Substitute in the standard diffusion equation (6.6), where $(R_B)$ is ballistic resistance;

$$I = \frac{\sigma_0 A}{qL}\big((\mu_1 - \mu_2) - qIR_B\big)$$

$$I + \frac{\sigma_0 A}{qL}(qIR_B) = \frac{\sigma_0 A}{qL}(\mu_1 - \mu_2)$$

$$I\left(1 + \frac{\lambda}{L}\right) = \frac{\sigma_0 A}{qL}(\mu_1 - \mu_2)$$

$$I = \frac{\sigma_0 A}{q(L + \lambda)}(\mu_1 - \mu_2)$$

This is the correct answer, therefore, when the boundary conditions are modified as shown, the right answer out of the diffusion equation is got with no need to change the equation itself.

This how the boundary conditions are represented pictorially; to justify the answer,

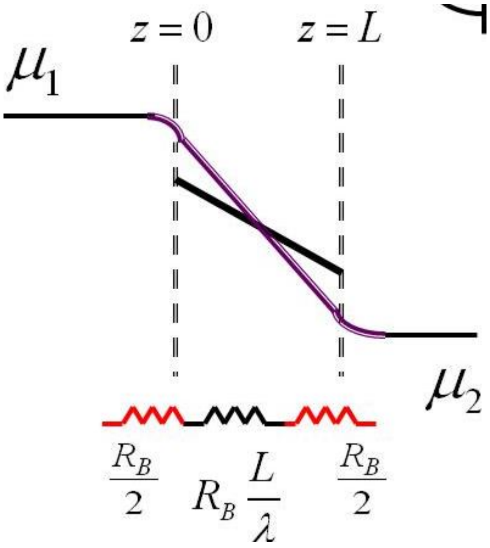


Figure 6.11: Pictorial representation boundary conditions

### 6.3 Quasi Fermi Levels (QFL)

The new boundary conditions have been introduced in the last paragraph for the electrochemical potential to solve the standard diffusion equation (6.6) to get the correct answer for the current calculation. The new boundary conditions include an extra interface resistance. There are two quasi Fermi levels, one for the right moving carriers and the other for the left moving carriers.

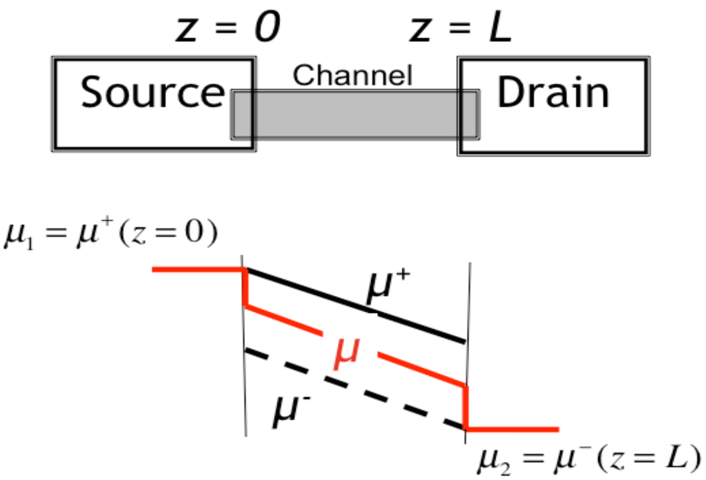


Figure 6.12: Two quasi Fermi levels, one for the right moving carriers and the other for the left moving carriers.

From the figure (6.12) the correct boundary conditions of the quasi levels are;

$$\mu^+(z=0) = \mu_1 \quad (6.11)$$

$$\mu^-(z=L) = \mu_2 \quad (6.12)$$

The electrochemical potential is the average of the two. The approximate boundary conditions of this average are;

$$\mu(z=0) = \mu_1 - \frac{qIR_B}{2} \quad (6.9)$$

$$\mu(z=L) = \mu_2 + \frac{qIR_B}{2} \quad (6.10)$$

To determine the correct boundary conditions for these two quasi Fermi levels, for the purpose of connecting them later.

(1) Start from the ballistic end where the length of the channel is very much less than the mean free path ($\lambda$). Therefore, the middle region resistance is very small, and the entire resistance is just the interface resistance ($R_B$). In this case the electrochemical potential inside the channel will not be able to

change and expected to be flat, but half way in between the two quasi Fermi levels, because the quasi Fermi level are separate, the right moving carriers quasi Fermi level is connected to the electrochemical potential ($\mu_1$), and the left moving carriers quasi Fermi level is connected to the electrochemical potential ( $\mu_2$).

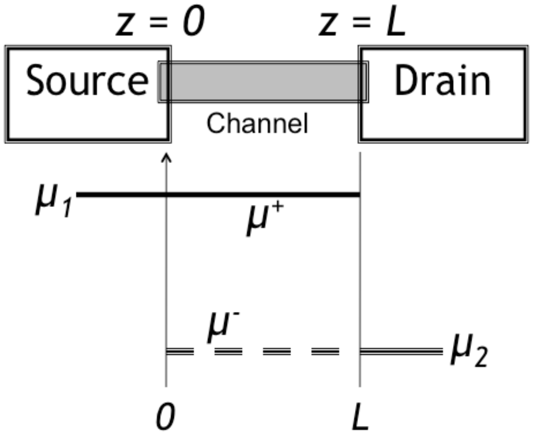


Figure 6.13: Quasi Fermi levels are connected to the electrochemical potential ($\mu_1$), and to the electrochemical potential ( $\mu_2$).

To understand this, the electrochemical potential originates from the Fermi function which possesses the property when high above the the electrochemical potential μ equals to zero, that is to say, all the states far above the electrochemical potential are completely empty, and far below the electrochemical potential μ, Fermi function is equal to one, that means all the states far below the electrochemical potential are completely filled.

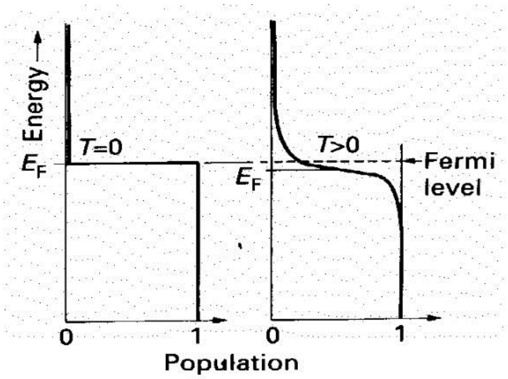


Figure 6.14: Transition change in Fermi function at different temperatures

At low temperature $(T = 0)$ the transition from one to zero is quite abrupt (Brickwall). The energy range over which this transition change take place, is in the order of few $(KT)$ for $(T > 0)$.

The electrochemical potential $\mu_1$ in the left hand side contact is drawn to indicate that all the states below it are completely filled, the same with the electrochemical potential ( $\mu_2$) in the right hand side contact. Consider the energy range $(qV)$ between ($\mu_1$ & $\mu_2$), the Fermi function below $(\mu_1)$ is equal to one in the left hand side contact, and equal to zero above ( $\mu_2$) in the right hand side contact.

For the energy range $(qV)$, all the right moving carriers states are completely filled and all the left moving carriers states are completely empty, between them is the channel to act as a highway.

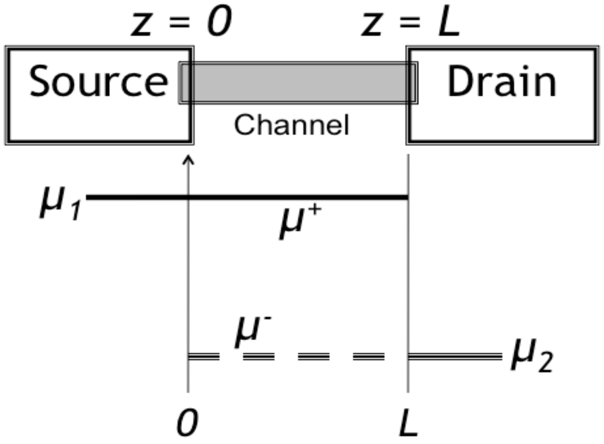


Figure 6.15: In Ballistic Transport, Right moving carriers and left moving carriers makes the channel to act as a highway

Then all the carriers in the left hand side contact are trying to get to the right hand side contact, at the same time there are no carriers from the right hand side contact trying to get to the left hand side contact. All the lanes from left contact to the right contact are expected to be full of carriers moving towards the right contact ,that is $(\mu^+)$ , all the lanes from right contact to left contact are completely empty ,that is $(\mu^-)$.

This situation can be describe as the BALLISTIC TRANSPORT, since the length of the channel is much less than the mean free path. In this context when an electron gets into the highwayit has no chance to turn around, because it is ballistically transported and it keeps going in straight line till it gets to the other end.

In the DIFFUSIVE TRANSPORT conductors, the situation is different, because some of the electrons have the chance to turn around and come back to the left hand side contact, where they originate. At this left hand side end the traffic from left hand side to the right hand side is very condense. As the electrons go along the channel from left to right they reduce in number, because some of them start to turn round and go back. The states in the other end start of being empty, they gradually build up due to the electrons turned around from the other lanes. In terms of the quasi Fermi levels what is expected is the right moving carriers starts from the left contact at $\mu_1$ and gradually reduce in number, hence the quasi Fermi level gradually goes down, and the left moving electrons start from $\mu_2$ and gradually they increase in number, hence the quasi Fermi level build up.

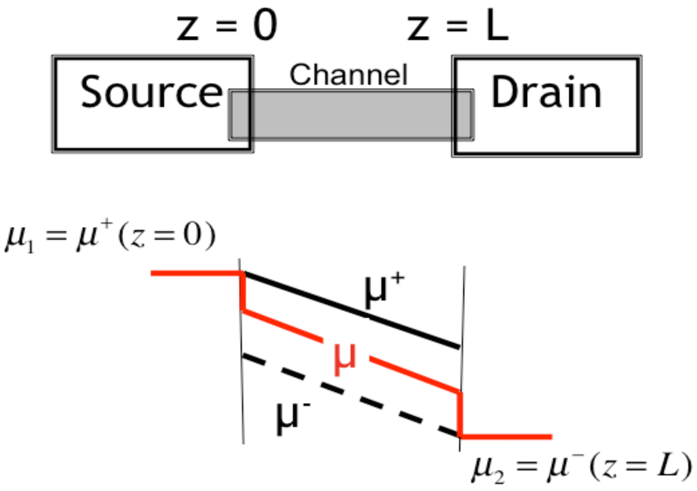


Figure 6.16: In Diffusive Transport, Right moving carriers and left moving carriers makes the channel to act as a highway with some electrons having chance to turn around to the left handside

The boundary condition of the electrochemical potential starts at the left contact end at $\mu_1$ to follow the quasi Fermi level $(\mu^{+})$, therefore, the states inside the channel are filled up to $\mu_1$ as well. The boundary condition of electrochemical potential starts at the right contact at $\mu_2$ to follow the quasi Fermi level $(\mu^{-})$, therefore, the states inside the channel are empty at $\mu_2$ as well. Then;

$$(\mu^{+} at\ z = 0) = \mu_1$$

$$(\mu^{-} atz = L) = \mu_2$$

The key point here, $(\mu)$ tells the degree of filling the states, and not to be considered as the average energy of the electrons, because the average energy of the electrons goes down due to some dissipation. This dissipation is not necessarily happening in the channel, but somewhere else, mostly in the contacts. The inclination of the quasi Fermi level doesn't imply any energy dissipation but telling that the density of electrons will keep going down as expected from the common sense. Two quasi Fermi levels are expected to describe the occupation of the different electrochemical potentials in either of the right moving carriers' lanes or the left moving carriers' lanes.

(2) To connect the quasi Fermi levels boundary conditions to the deduced modified new boundary conditions to show that the quasi Fermi levels boundary conditions are nothing more than the modified boundary conditions obtained, or the quasi Fermi levels boundary conditions implies the modified boundary conditions.
Recall the diffusion equation (6.6), and use the set of boundary conditions to get a result including the ballistic resistance;

$$\mu(z = 0) = \mu_1 - \frac{qIR_B}{2}$$

$$\mu(z = L) = \mu_2 + \frac{qIR_B}{2}$$

Also recall the boundary conditions of the quasi Fermi levels

$$(\mu^{+} at\ z = 0) = \mu_1$$

$$(\mu^{-} at\ z = L) = \mu_2$$

To deduce the modified boundary condition from the quasi Fermi levels boundary conditions a relation of current to the separation between the two quasi Fermi levels is needed;

$$I = \frac{G_B}{q}(\mu^{+} - \mu^{-}) \qquad (6.13)$$

When the ballistic conductance is seen within the diffusive equation, there will be a miss understanding, because everyone may think that this equation only applies to the ballistic case, but this a general equation can be applied to both ballistic and diffusive transports.

Turn the equation around to get;

$$(\mu^{+} - \mu^{-}) = \frac{qI}{G_B} = qIR_B \quad (6.14)$$

Then the separation of quasi Fermi levels is directly proportional to the current. If there is no current the two quasi Fermi levels will collapse into one level to tell the two quasi Fermi levels are equally filled, and there will be no current flow unless one of the two quasi Fermi levels is (less or more) than the other.

To accept the current equation (6.13), without knowing where this current equation comes from, these quasi Fermi levels boundary conditions shown by equations (6.11 & 6.12), give the boundary conditions for $(\mu)$ shown by equations (6.9) & (6.10), quite simply as follows:

The average of the two quasi Fermi levels can be given as;

$$\mu = \frac{\mu^{+} + \mu^{-}}{2} \qquad (6.15)$$

Since from equation (6.11&6.12), $(\mu_1 = \mu^{+} and\ \mu_2 = \mu^{-})$, and from equation (6.14) the difference between the two quasi Fermi levels is given by;

$$(\mu^{+} - \mu^{-}) = qIR_B$$

Then this can be substituted in equations (6.9 & 6.10) respectively;

$$\mu(z=0) = \mu = \frac{\mu^{+} + \mu^{-}}{2} = \mu^{+} - \frac{(\mu^{+} - \mu^{-})}{2} \quad (6.16)$$

$$\mu(z=L) = \mu = \frac{\mu^{+} + \mu^{-}}{2} = (\mu^{-}) + \frac{(\mu^{+} - \mu^{-})}{2} \quad (6.17)$$

These are the relations which directly give the boundary conditions for

$$\mu(z=0)=\mu_1=\mu^+$$

$$\mu(z=L)=\mu_2.=(\mu^-)$$

$$qIR_B=(\mu^+-\mu^-)$$

So, the boundary conditions automatically give the modified boundary conditions in one line.

### 6.4 Current from The Quasi Fermi Levels

The relation which relates the current to the separation of the quasi Fermi levels is given by equation (6.13) as;

$$I=\frac{G_B}{q}(\mu^+-\mu^-)$$

The aim now is to explain where this equation comes from, after it has been taken for granted it is a correct, and used without having to be proved, for the purpose to connect the two quasi Fermi levels boundary conditions to the modified boundary conditions, to show how the quasi Fermi levels imply the modified boundary conditions. The origin of the current is the flow of electrons across the channel. Starting by the current of electrons from left to right to be expressed as;

$$I^+t=q\,\frac{D(E)}{2}\,f^+(E)\qquad(6.18)$$

Where, $I^+$ is the right moving electrons

$t$ The time spent inside the channel

$q$ is the electron charge

$\frac{D(E)}{2}$ is the density of states in the right moving bound lane

$f^+(E)$ is the occupation of the right moving states

The time spent inside the channel for ballistic transport equals $\left(\frac{L}{\tilde{V}}\right)$, where $(\tilde{V})$ is the average velocity of the electron, $(L)$ is the channel length. Substitute the time in equation (6.18), and rearrange:

$$I^{+} = q\ \frac{D(E)\tilde{V}}{2L}\ f^{+}(E)dE \quad (6.18)$$

$\left(\frac{D(E)\tilde{V}}{2L}\right)$ is defined as $(M(E)$, the number of modes

$$I^{+} = \frac{q}{h}\ M(E)\ f^{+}(E)dE \quad (118')$$

Similarly, with the left moving electrons current;

$$I^{-} = \frac{q}{h} M(E)\ \ f^{-}(E)dE \quad (6.19)$$

The overall current in the energy range $dE$ is the difference between the two currents;

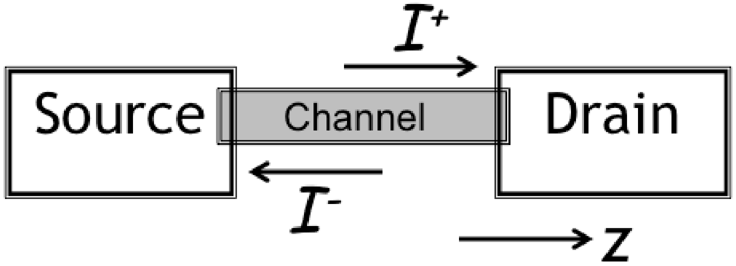


Figure 6.17: The overall current in the energy range $dE$

$$I = I^{+} - I^{-} = \frac{q}{h}\ M(E)\ \ (f^{+}(E) -\ f^{-}(E))\ dE \quad (6.20)$$

Using the general Taylor series expansion to get from the occupation, expressed in terms of the Fermi function, to the quasi Fermi levels, then $(f^{+}(E) - f^{-}(E))$will be changed to the derivative

$$I = I^{+} - I^{-} = \frac{q}{h}\ M(E) \left[-\frac{\partial f_0}{\partial E}\right] dE(\mu^{+} - \mu^{-}) \qquad (6.21)$$

Integrate over energy for the total current;

$$I = I^{+} - I^{-} = \frac{q}{h}(\mu^{+} - \mu^{-}) \int_{-\infty}^{\infty} M(E) \left[-\frac{\partial f_0}{\partial E}\right] dE$$

Let, $\int_{-\infty}^{\infty} M(E) \left[-\frac{\partial f_0}{\partial E}\right] dE = M_0$ ,then ;

$$I = I^+ - I^- = \frac{qM_0}{h}(\mu^+ - \mu^-) \ \rightarrow I = \frac{G_B}{q}\ (\mu^+ - \mu^-) \qquad (6.22)$$

Therefore, the ballistic conductance is;

$$G_B = \frac{q^2}{h} M_0 \quad (6.23)$$

Basically $G_B$times the quasi Fermi levels separation holds at any point, and the current should have to be the same everywhere as required by the continuity equation to imply that the quasi Fermi levels difference must be the same everywhere i.e. quasi Fermi levels are parallel straight lines.

Usually the current is written as the conductance times the electrochemical potentials voltage difference;

$$I = \frac{G_0}{q}\ (\mu_1 - \mu_2) \quad (6.24)$$

Also, the current is related to the separation of the quasi Fermi levels;

$$I = \frac{G_B}{q}(\mu^+ - \mu^-)$$

To connect these two current expressions, equate both relations to find the ratio between ( $\frac{G_0}{G_B}$ );

$$\frac{G_0}{G_B} = \frac{(\mu^+ - \mu^-)}{(\mu_1 - \mu_2)}$$

Recall equation (100);

$$G_0 = G_B \frac{\lambda}{L + \lambda} \quad (6.25)$$

Then;

$$(\mu^+ - \mu^-) = \ (\mu_1 - \mu_2) \frac{\lambda}{L + \lambda} \quad (6.26)$$

For ballistic transport $(L)$ is much less than the mean free path, and the quasi Fermi levels separation is equal to the applied energy potential $(qV)$ i.e. The separation between the electrochemical potentials at the contacts. For effusive conductors length of the channel $(L)$ is longer than the mean free path, and the quasi Fermi levels separation is much smaller than the applied energy potential difference $(qV)$, and that is why the quasi Fermi level separation is forgotten, and no one needs to worry about it.

## 6.5 Landauer Formalism

Historically Landauer Formulae are very important to develop all the points of view to see how these formulae connect to what has been discussed. So far channels with a continuous scattering processes everywhere, has been discussed.

For Landauer resistance the case is completely different because the channel is largely ballistic where the scattering is localized somewhere. This helps very much to clarify certain ideas.

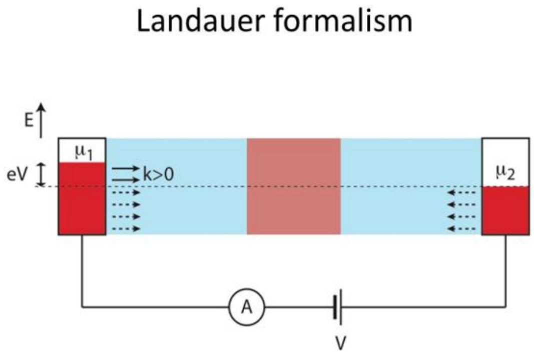


Figure 6.18: Localized barrier where scattering take place.

The channel has distributed scattering. The scattering processes are described by the mean free path $(\lambda)$, that is how far an electron goes before it is turned around. Consider a localize scatterer (barrier), i.e. The channel is basically ballistic and no scattering anywhere except one point where the acatterer is localized.

Under these conditions, two things will be discussed;

1. The average current flow
2. The shape of the potential profile

The conductance is given by the ballistic conductance times the transmission coefficient (Ŧ).

$$G_0 = G_B(\text{Ŧ}) \quad (6.27)$$

Where, (Ŧ) is the transmission coefficient, that describes the localized scatterer or the fraction of electron that crosses the scatterer to get to the other side.

For example, if $(\text{Ŧ}) = 0.3$, means 30% leaves to the other side from the total 100% coming in, and the current flow expected is not more than 30%, so the conductance decreases to 30% of its value.

To compare, $G_0 = G_B(\text{Ŧ})$ with $G_0 = G_B\frac{\lambda}{L+\lambda}$, as if (Ŧ) for the scatterer replaces ($\frac{\lambda}{L+\lambda}$) the quantity of distributed scattering processes, i.e. ($\frac{\lambda}{L+\lambda}$) can be interpreted as (Ŧ). Hence a channel of length $(L)$, and mean free path $(\lambda)$, with distributed scattering processes, the fraction (Ŧ) $is$ the ratio of the incident electrons from left to right. But that is something different from what the localized scatterer whose transmission probability is (Ŧ), in that case the electrochemical potential of the quasi Fermi levels, the upper quasi Fermi level drops sharply from $\mu_1 to (\text{Ŧ})$, and at the other end the lower quasi Fermi level steps up from $\mu_2 = 0$ to $\left(1 - (\text{Ŧ})\right)$, bear in mind that $\mu_1\ is\ 100\%\ full, while\ \mu_2\ is\ 100\%\ empty$, here below find the plot of the state's occupation;

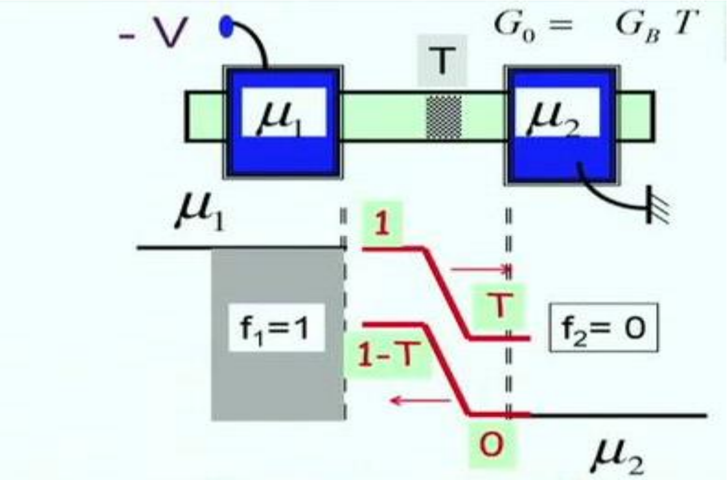


Figure 6.19: The states occupation in the presents of scatterer.

The states in $\mu_1$ contact are all filled, so the occupation is (1), and in $(\mu_2)$ contact the occupation is zero. In the left side contact the right moving states are all occupied, because $(\mu^{+})connect\ to\ \mu_1$, all the states are filled. When the scatterer is crossed the occupation drops sharply, and the occupation is much less than one, and the occupation will go from $(1\ to\ (Ŧ))$ inside the channel. When the right moving carriers arrived at the other end, they face a situation that all the states are completely empty above $(\mu_2)$ in the range of interest. The right moving carriers can't distribute themselves below $(\mu_2)$ because all the states are fully occupied there, so this range doesn't matter to them. Above $(\mu_2)$ there is no proper energy distribution, so a virtual $(\mu)$ is chosen not to describe the energy distribution of electrons rather than just to represent the degree of filling the states. $(\mu)$ is not a proper electrochemical potential, because the actual energy distribution needs an inelastic process. In elastic process some electron may go up to a level equals $(\mu_1)$ but still cannot be considered an equilibrium condition, because the equilibrium condition needs an inelastic process. The degree of filling the states could be turned to look like Fermi function to change at low temperature from zero to one.

In brief, when the right moving carriers cross the barrier, lanes at the other end in the right contact appear very empty and the virtual $(\mu)$ gives the degree of filling the states. There is no proper electrochemical potential.

The left moving carriers quasi Fermi level steps up sharply from $(\mu_2 = 0)$ to $\mu = (1 - Ŧ)$, after crossing the barrier they see the channel occupied up to $(\mu = 1 - Ŧ)$. When they reach the left contact they see that it is completely filled, they can't climb to $(\mu_1)$ level, so they stay at the level somewhere between $\mu_1 and\ zero\ at\ (1 - Ŧ)$, therefore, the problem is very complicated and people feel not quite comfortable with it because it needs a lot of debate, better not to go deep into that question at all and to be left for those who are interested in the subject as a future work. It is quite obvious the quasi Fermi levels are deep subtle concepts need statistical mechanic to get it right.

Remember that the electrochemical potential $\mu$ is not describing the energy distribution of electrons. The occupation $(f)$ is related to the electrochemical potential $(\mu)$ by the following equation;

$$(\mu - \mu_2) = qVf$$

Recall the current equation (113)

$$I = \frac{G_B}{q}(\mu^+ - \mu^-) = \frac{G_B}{\cancel{q}}\,\cancel{q}Vf$$

Where, $(\mu^+ - \mu^-)$ is a constant everywhere with a difference equal to $(\text{Ŧ})$. The current is directly proportional to the voltage, and can be given by;

$$I = VG_B(\text{Ŧ}) = VG_0 \rightarrow R = R_B/(\text{Ŧ}) \quad (6.28)$$

Where, $G_B$ is the ballistic conductance.

$(\text{Ŧ})$ is the transmission coefficient

The inverse of $G_B$ is the ballistic resistance $R_B$ . The resistance of the scatterer is $R_s$ gives rise to a lot of extra resistance. The voltage drop across the scatterer is a fraction of the total voltage drop $(\mu_1 - \mu_2)$. If the two quasi Fermi levels have been drawn, any one of them can be used to give the same answer, to find Landauer formulae, because in either case the drop across the scatterer is the overall voltage times $(1 - \text{Ŧ})$.

$$\Delta V = V(1 - \text{Ŧ})$$

$$R_s = \frac{\Delta V}{I} = \frac{V(1 - \text{Ŧ})}{VG_B(\text{Ŧ})} = \frac{R_B(1 - \text{Ŧ})}{(\text{Ŧ})} = \frac{R_B}{(\text{Ŧ})} - R_B \quad (6.29)$$

Equations (6.28 & 6.29) give the overall resistance of the device R, and the scatterer resistance $R_s$ respectively, known as Landauer formulae;

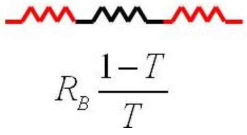
$$R = R_B/(\text{Ŧ})$$

$$R_s = \frac{R_B(1 - \text{Ŧ})}{(\text{Ŧ})}$$

They differ by an amount $R_B$,

$$R - R_s = R_B$$

Therefore, there is an extra part $R_B$ to $R_s$which appears at the interfaces. Taking the average of the two quasi Fermi levels there will be an extra drop at the source contact and another drop at the drain contact as well, therefore, half of $R_B$is associated with either of the two ends. The middle part $R_s$is associated with the scatterer.



For the right moving carriers' quasi Fermi level all the resistance $R_B$ appears at the drain contact $(\mu_2)$, the same at the other end for the left moving carriers quasi Fermi level, all the resistance $R_B$ appears at the source contact $(\mu_1)$. This is a very subtle condition needs more discussion. In this context professor Supriyo Data of Prude University USA said:

*The overall resistance is* $\frac{1}{(\text{Ŧ})}$ *, whereas the scatterer resistance is* $(\frac{1-\text{Ŧ}}{\text{Ŧ}})$*, these are the Landauer formulating the sense that Landauer had appreciated all the sublet issues involved, although the interface resistance took more discussion and people credit Imref* with not recognizing this interface resistance, and in this context,*

$$R_s = \frac{R_B(1-\text{Ŧ})}{(\text{Ŧ})}$$

*Is the original Landauer formula whereas today people use,*

$$R = R_B/(\text{Ŧ})$$

*To interpret their results, because usually they look for the total resistance rather than looking for scatterer resistance.*

---

**The Quasi Fermi Level is called (Imref) which is Fermi spelled backwards, it is a term used in quantum mechanics and specially in solid state physics.*

## 6.6 The Electrostatic Potential

Quasi Fermi Level is rather difficult concept to be understood, because it needs a deeper knowledge of statistical mechanics. Quasi Fermi Levels are very difficult because it depends on the electrochemical potential, which is so difficult to define out of equilibrium, that is why it leaves uncomfortable feeling to push some of the researchers to think of something more concrete. For that purpose, two approaches are suggested;

a) The discussion of the subject should be carried through the study of electrostatic potential instead of studying it through the electrochemical potential, although electrostatic potential hides some of the information explored by using the electrochemical potential, as a result some of the physics will be obscured.

b) Practical measurement of the Quasi Fermi Level to make use of to justify why to look for the average instead of looking to the Quasi Fermi Levels.

a) Let’s start with the study of electrostatic potential. Consider the figure below,

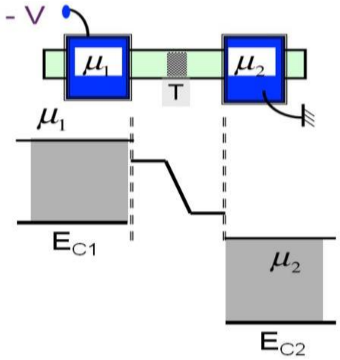


Figure 6.20: Electrostatic potential inside the channel between the tops of the source band and the drain band is separated by (qV)

The bottom of the band $(EC_1)$ represents, where the density of states ends. Assume there is only one band under consideration at each end. When a voltage is applied, the bottom of the bands at the source and the drain are separated by an amount $(qV)$, the same as the electrochemical potential do at the top of the bands. All the energy levels in the positive drain side go down with respect to the energy levels in the negative source side. The electron densities are the same in both contacts. Since the bottom of the band generally follow the electrostatic potential (except for minus sign) as shown in the introduction of this chapter (6).

According to what has been said; initially anyone may think that the electrostatic potential inside the channel between the bottom of the bands resemble what the electrochemical potential is doing inside the channel in between the top of the bands. If it did that means the electron density remained unchanged.

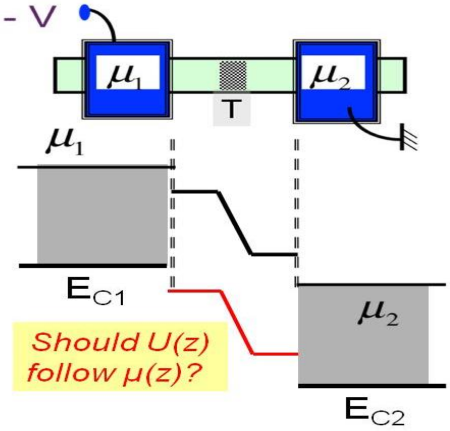


Figure 6.21: Electrostatic potential inside the channel between the bottoms of the source band and the drain band resembles the electrochemical potential in between the tops of the bands

The electrostatic potential is seen to be changing along the barrier, so the gradient of which gives rise to an electric field. This electric field goes from the positive drain side to the negative source side, as a result charges create a residual dipole of around the barrier, to accommodate the extra electrons on the drain side and the deficient of the electrons in the source side. Therefore, the electrostatic potential curve cannot quite follow the electrochemical potential, it should have to be a little different. If the channel is very conductive, the density of states will very high, then the electrostatic potential approximately follows the electrochemical potential. To prove that let's write down the number of electrons as follows;

$$N = \int_{-\infty}^{\infty} D(E-U) f(E-\mu) dE \quad (6.31)$$

Where, $D(E-\mu)\ is\ the\ density\ of\ states$

$f(E-U) is\ the\ filling\ factor\ or\ the\ fraction\ of\ occupation$

The density of the states is shifted by the electrostatic energy (U). If (U) is high, the density of states is moved up, if $(U)$ is low the density of states is moved down. By the change of variables;

$$N = \int_{-\infty}^{\infty} D(E) f(E - \mu + U) dE \quad (6.32)$$

The argument f now picks all the change into the Fermi function rather than into the density of states. By Taylor series expansion;

$$\delta N = \delta(\mu - U) \int_{-\infty}^{\infty} D(E) \left(-\frac{\partial f_0}{\partial E}\right) dE \quad (6.33)$$

The change in the number of electrons, $(\delta n)$ due to the change in $\delta(\mu - U)$ are related by the integral $(\int_{-\infty}^{\infty} D(E) \left(-\frac{\partial f_0}{\partial E}\right) dE = D_0)$. The electron density depends on the distance from the electrochemical potential at the top of the band to the electrostatic potential at the bottom of the band, $(\mu - U)$, and it is related by the average density of states, $D_0$;

$$\delta n = \delta(\mu - U) D_0 \quad (6.34)$$

If the medium is so conductive with a high density of states, that means, $D_0$ is high, so for a given change of electron density, small change in $\delta(\mu - U)$ is needed, all the extra electrons create an electric field to be accommodated with only slight change in $\delta(\mu - U)$; which cannot be noticed. That is to say, in a very conductive medium the change in $(U)$ is approximately equals the change in $(\mu)$.

For semiconductors the density of statuses not high enough but still equation (134) holds. From the mathematical point of view the actual change in potential $(U)$ can be calculated using Poisson's equation;

$$\nabla^2(\delta \mathrm{U}) = -\frac{q^2 \delta \mathrm{n}}{\varepsilon} \qquad (6.35)$$

Substitute $(\delta n)$ from equation (6.34) into equation (6.35),

$$\nabla^2(\delta \mathrm{U}) = -\frac{q^2 \mathrm{D}_0 \delta(\mu - \mathrm{U})}{\varepsilon} \qquad (6.36)$$

The factor $(\frac{q^2 \mathrm{D}_0}{\varepsilon})$ has got the dimension $(m^{-2})$ ,and it can be defined as the inverse of the screening length squared $(\frac{1}{\lambda_s^2})$ , then rewrite equation (6.36) as,

$$\nabla^2(\delta U) = -\left(\frac{1}{\lambda_s^2}\right)\delta(\mu - U) \ \ (6.37)$$

By simple algebra,

$$\delta U\left(\nabla^2 - \frac{1}{\lambda_s^2}\right) = -\frac{\delta\mu}{\lambda_s^2} \qquad (6.38)$$

Where, the screening length $\lambda_s = (\frac{\varepsilon}{q^2 D_0})^{\frac{1}{2}}$.

This differential equation (138) defines how $(\delta U)$ respond to the screening length, hence to the change in electrochemical $(\mu)$. If the screening length approaches zero implies a very high density of states $(D_0)$ , and the inverse of the screening length approaches $(\infty)$, thus $(\nabla^2)$ can be dropped, therefore,

$$\delta U = \delta\mu$$

This the case for the very conductive medium which has been discussed already. But in case the screening length is long enough, then equation (6.38) should be solved by assuming that the change in μ is at one point then the corresponding $(\delta U)$ is an impulsive response of equation (6.38);

$$S\left(\nabla^2 - \frac{1}{\lambda_s^2}\right) = \delta\boldsymbol{r}. \ \ (\boldsymbol{6.38'})$$

Where, S is the screening function which is roughly the screening length responding to $(\delta\boldsymbol{r})$.

The actual response $(\delta U)$ is the convolution to of the screening function $(S)$;

$$\delta U = \delta\mu * S(Z)$$

So, the change of electrostatic potential U equals the change in electrochemical potential convoluted with $S(Z)$, the screening function, then the electrochemical potential is smoothed out curve, and the sharp edges of electrostatic potential smeared out by screening length. But this would not be the case when the medium is very conductive.

b) Practical measurement of the Quasi Fermi Level to make use of to justify why to look for the average instead of looking to the Quasi Fermi Levels.

Some of the researchers suggested that it too early to talk about the Quasi Fermi Levels before studying the accessible quantities inside the channel which can be measured from outside by using scanning tunneling probe. There are two groups of states, the right moving states with a Quasi Fermi Level ($\mu^+$), and another group of states, the left moving ones with a Quasi Fermi Level ($\mu^-$), all these are inside the channel. Let us have a probe whose potential is ($\mu_p$), to access the inside quantities from outside by a volumetric measure. Electron can tunnel through the region ($T$) could be represented by a little conductance. Two separate conductance are required, one for the right moving states ($g^+$), and ($g^-$) for the left moving states. Electrons in the probe is preferred to be coupled to the right moving states, but that wouldn't happen, they may be coupled equally, so, the two conductance are equal, but in general they could be unequal. To determine the probe potential, suppose the voltmeter draws no current to flow out into the probe, by Kirchhoff's law whatever current flows through ($g^+$), must be the negative of the current flows through ($g^-$).

$$I^+ = \left(\mu^+ - \mathrm{\mu_p}\right)g^+$$

$$I^- = \left(\mu^- - \mathrm{\mu_p}\right)g^-$$

The sum of these two currents is equal zero, then the probe potential ($\mu_p$) comes out as a weighted combination or weighed average of ($\mu^+ and\ \mu^-$).

$$\mathrm{\mu_p} =. \left(\frac{g^+}{(g^+ + g^-)}\mu^+ + \frac{g^-}{(g^+ + g^-)}\mu^-\right)$$

$$\mathrm{\mu_p} = (\alpha\mu^+ + (1-\alpha)\mu^-)$$

The factors in front add to one. If the probe couple very well to the positive right moving states, then $(g^{+})$ would be high conductance and $(g^{-})$ will be low conductance. Therefore, $(\alpha = 1$ & $(1 - \alpha) = 0)$ in that case.

On the other hand, if the probe couple better to the left moving states $(\alpha = 0$ & $(1 - \alpha) = 1)$. This is how the probe is measuring the Quasi Fermi Levels.

If $(\alpha = 1)$, the probe measures the top curve of the Quasi Fermi Level $(\mu^{+})$. When it measures (one) and seen drops to $(T)$, therefore, the voltage drop is $(1 - T)$, and the current is directly proportional to $(T)$, the resistance is $(R = \frac{1-T}{T})$.

If $(\alpha = 0)$, the probe measures the bottom curve of the Quasi Fermi Level $(\mu^{-})$. When it measures $(1 - T)$ and seen drops to (zero), therefore, the voltage drop is still $(1 - T)$.

More commonly what would happen if $\alpha = (1 - \alpha) = \frac{1}{2}$, that will correspond to the average curve $(\mu)$, still the drop will be $(1 - T)$. So, in any case the same conclusion comes out, if the probe had always fixed α with no change as the probe is moving. This provides a good understanding of what the non-invasive probe is measuring relative to the Quasi Fermi Levels when it is weakly coupled to sense what is going inside the channel without changing it.

The three-terminal circuit with a probe will be as follows,

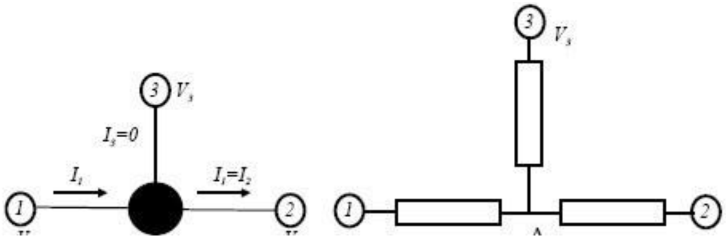


Figure 6.22: Three terminal circuit with terminal (3) as a probe potential $(\mu_p)$.Terminal (1) is the right moving states with the QFL $(\mu^{+})$. Terminal (2) is the left moving states with QFL $(\mu^{-})$. The scatterer region T is represented by separate conductance $(g^{+})$ at terminal (1) and, $(g^{-})$ at Terminal (2). These conductances are determined from terminal (3).

Late 1980's and 1990's of the last century, voltages are measured in mesoscopic physics with a probe with conductors connected directly to the channel, this make these conductors to act as scatterers in their own and change the entire distribution of what is required to be measured. Hence, they need some calculations using linear algebra (matrices) to assure that these probes were not there, as an example, *Busttiker equation, suggested by him back in 1986. So, they are not relevant now, since the non-invasive probes has already emerged.*

## 6.7 The PN junction

The PN junction is one of the primary building blocks of semiconductor devices. It is a one-way valve with P-type on one side and N-type on the other side. The P-type material has a conduction band edge and a valance band edge, it's electrochemical potential (the Fermi level) ($\mu$) is closer to the valance band edge. The same with the N-type material, except that it's electrochemical ($\mu$) is closer to the conduction band edge. At equilibrium the electrochemical potentials the same in both materials, but the P-type material must be offset from the N-type material by a built-in potential.

When a voltage is applied for forward bias, all the energy levels in the N-type material are pulled up, so it's electrochemical potential separates from the electrochemical potential of the P-type material. Around the junction, in the active region, two electrochemical potentials are found to create two separate quasi Fermi levels there, one for the N-type material and the other for the P-type material. One of the quasi Fermi Levels is controlled by the electrons coming from the N-type material, and the other is controlled by the holes coming from the P-type material.

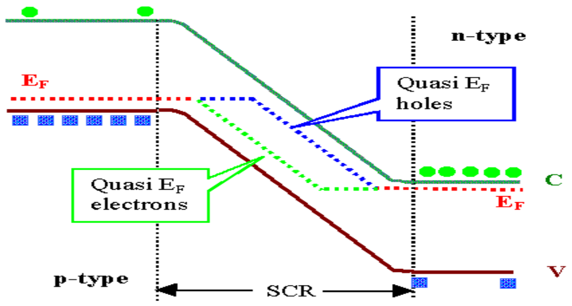
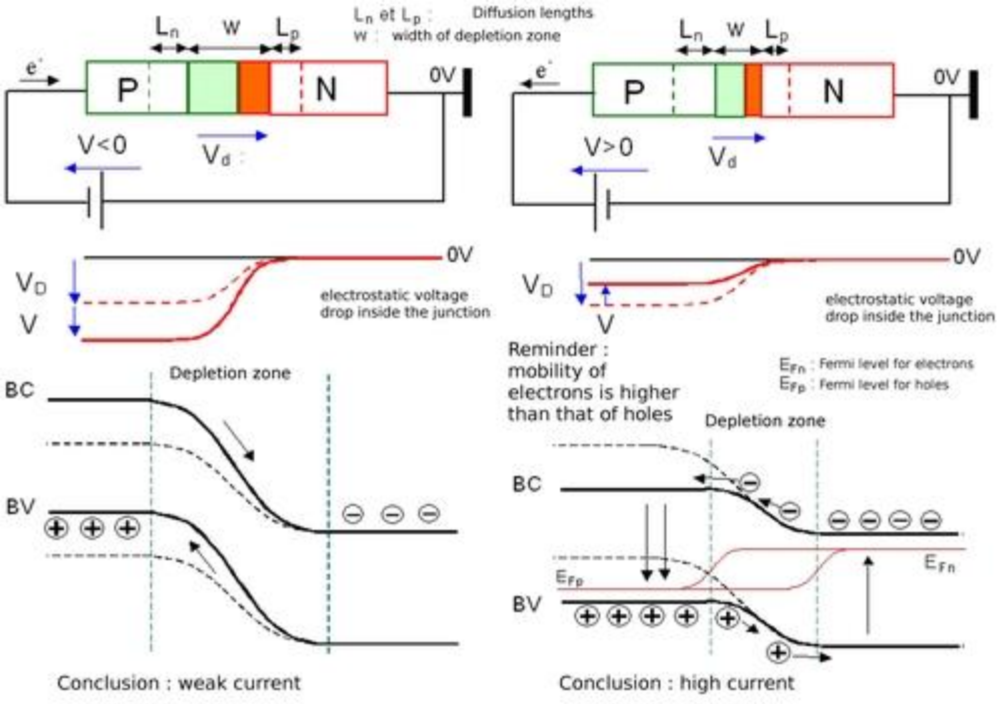


Figure 6.23: Two separate quasi Fermi levels are created in the active region around the junction

These two quasi Fermi Levels give a lot of information of how the current flows.

When the electrostatic potential is considered, at equilibrium there will be a built-in potential around the junction in the active region. This is the major difference with electrochemical potential case, because the electrochemical potential is flat, i.e. the electrochemical potential is the same everywhere, but for the electrostatics case there is a major variation around the junction region to indicate that there is a flow due to an electric field, the problem there is a built in electric field at equilibrium with net flow equals to zero (no flow).

When a forward bias voltage is applied the electrostatic potential is reduced. The change in the electrostatic potential has become less as shown in the following figure **[76],**



a) Direct bias

b) Reverse bias

Figure 6.24: Reduction of electrostatic potential under forward bias conditions

All the information given by the Quasi Fermi Levels diagrams are smeared and lost, up to now the electrostatic potential is used to study the PN junctions and the usage of the Quasi Fermi Levels is avoided, because the electrochemical potential requires additional conceptual discussions and justification, also it so difficult to define the electrochemical potential when the system is out of equilibrium. The proper justification of all what has been discussed so far involve the Boltzmann’s Equation which is going to be discussed in the following section.

## 6.8 Boltzmann's Equation

All the semi classical transport results arrived to, have their basis from Boltzmann's Equation. To justify properly everything has been discussed, start from the Boltzmann's Equation, which can be written as;

$$\left[\left(V_z \frac{\partial f}{\partial Z}\right) + \left(F_z \frac{\partial f}{\partial P_z}\right)\right] = -\left(\frac{f - \vec{\vec{f}}}{\tau}\right) \qquad (6.40)$$

It has not been used from the first begging because this equation is rather difficult to understand it mathematically. It is a partial differential equation with multivariable in it. It includes the spatial location $(Z)$, the momentum $(P_z)$, and time $(t)$. In general, the more the independent variables the more difficult to conceive and solve the equation mathematically or numerically. For this reason, this dissertation started by the derivation of everything from its fundamentals analytically, to explore the basic knowledge required in a simple way, till a reasonable standard capacity of knowledge is reached, to provide the capability of understanding everything has been discussed, to justify it directly from Boltzmann's Equation.

To introduce the idea of Quasi Fermi Levels the Fermi function should be taken out of equilibrium, and rewritten in the form;

$$f(Z, P_z, t) = \frac{1}{1 + \exp\left(\frac{E - \mu(Z, P_z)}{KT}\right)} \qquad (6.41\ )$$

Where, $\mu(Z, P_z)\ is\ the\ Quasi\ Fermi\ Level.$

$f(Z, P_z, t)\ is\ the\ occupation.$

At equilibrium $(\mu)$ is constant, then the function $(f)$ is given by the Fermi function. Out of equilibrium, $(\mu)$ is not constant, and the function just looks like the Fermi function with different $(\mu)$ at different locations, and different momentum.

The derivative $\left(\frac{\partial f}{\partial Z}\right)$ ,can be written to involve the same derivative and Taylor series expansion,

$$\left(\frac{\partial f}{\partial Z}\right) = \left(-\frac{\partial f_0}{\partial E}\right)\frac{\partial \mu}{\partial Z}$$

Similarly,

$$\left(\frac{\partial f}{\partial P_z}\right) = \left(-\frac{\partial f_0}{\partial E}\right)\frac{\partial \mu}{\partial P_z}$$

Bring in the Boltzmann’s Equation and write a similar equation in terms of $\mu(Z, P_z)$ as shown by equation (131). As $f$ depends on $(Z, P_z\,, and\ E)$, because as far as $(E)$ is concerned, a particular form is assumed.

$$V\frac{\partial \mu}{\partial Z} + F_z\frac{\partial \mu}{\partial P_z} = -\frac{\mu - \vec{\mu}}{\tau}$$

Since $\mu = \mu^{-} = constant$, the equilibrium solution is given by Fermi function, then Boltzmann’s Equation is satisfied, because

$$\frac{\partial \mu}{\partial Z} = \frac{\partial \mu}{\partial P_z} = \frac{\mu - \vec{\mu}}{\tau} = 0$$

That is what statistical mechanics is saying at equilibrium, the electrochemical potentials the same everywhere. For nonequilibrium solution $\mu(Z, P_z)$, varies with the location $(Z)$ and the momentum $(P_z)$, and can have different electrochemical potentials or Quasi Fermi Levels associated with.

In devices if the applied voltage is small compared to the thermal broadening $(KT)$, the Fermi function with varying $(\mu)$ can be assumed under conditions of nonequilibrium. While it not necessarily true with larger applied voltages, because the energy distributional may not look like the Fermi function. Such case has been mentioned in Landauer formula, where the energy distribution may not look like the Fermi function, because $(\mu)$ is not describing the energy distribution rather than just telling the degree of filling states. Therefore, $(\mu)$ is defined to give the right number of electrons, in such a case any change in the occupation $(f)$ can be related to the change in electrochemical potential $(\mu)$ by;

$$\Delta\mu = qVf$$

So, the relation goes from $(f)$ to $(\mu)$, that is, any change in $(f)$ is related to the change in $(\mu)$ through $(qV)$. Similarly, when the Boltzmann's Equation is taken from the equation in terms of $(f)$ to the equation in terms of $(\mu)$, any change in $(f)$ is related to the change in $(\mu)$ through $(qV)$ and not through these derivatives,

$$\left(\frac{\partial f}{\partial Z}\right) = \left(-\frac{\partial f_0}{\partial E}\right)\frac{\partial \mu}{\partial Z} \;\&\; \left(\frac{\partial f}{\partial P_z}\right) = \left(-\frac{\partial f_0}{\partial E}\right)\frac{\partial \mu}{\partial p_z}$$

Then;

$$\left[\left(V_z \frac{\partial \mu}{\partial Z}\right) + \left(F_z \frac{\partial \mu}{\partial P_z}\right)\right] = -\left(\frac{\mu - \vec{\mu}}{\tau}\right) \quad (6.42)$$

The energy variable $(E)$ is taken out the discussion and $(\mu)$ dominate. Another important point the middle term $(F_z \frac{\partial f}{\partial P_z})$, consists of the force $(F_z)$, that an electron feels due to the electric field, this is where the electric field often enter the Boltzmann's Equation. Sometimes under certain conditions this middle term can be ignored, because at equilibrium all the derivatives turn to zero ~~with constant μ for an elastic ballistic transport through the channel.~~ Out of equilibrium the term $(F_z \frac{\partial f}{\partial P_z})$, really is a second order, when a voltage is applied, then an electric field is applied, so another first order term $(F_z = \frac{\partial P_z}{\partial t})$, exist there , and there is another first order term $\frac{\partial f}{\partial P_z}$ already there, in total is like the second order term. Therefore, the middle term can be dropped. This is wouldn't be true if there was a built-in field in some region, or there was a magnetic field due to Hall effect. For this discussion assume we can drop that term. The rest of the equation can be written in the form;

$$V_z \frac{\partial \mu(z, P_z)}{\partial z} = -\frac{\mu(Z, P_z) - \mu(Z, P_z)}{\tau} \quad (6.43)$$

Note that $(\mu)$ at a given location can be different for different momenta. To simplify further keep track of the two Quasi Fermi Levels $\mu^+ \; and \; \mu^-$ , $(\mu^+)$ for forward moving carriers, with positive velocity, and $(\mu^-)$ for backward moving carriers, with a negative velocity;

$$\|V_z\|\frac{\partial\mu^+}{\partial Z} = -\left(\frac{\mu^+ - \vec{\mu}}{\tau}\right) \qquad (6.44)$$

$$-\|V_z\|\frac{\partial\mu^-}{\partial Z} = -\left(\frac{\mu^- - \vec{\mu}}{\tau}\right) \qquad (6.45\ )$$

Since there are two different Quasi Fermi Levels, the system is under nonequilibrium situation. The scattering processes will try to restore the system into equilibrium by collapsing $\mu^+ \ and\ \mu^-$ into single average value $\vec{\mu}$, so, this the average value of the two;

$$\vec{\mu} = \left(\frac{\mu^+ + \mu^-}{2}\right)$$

Take equations (6.44 &6.45), and by making use of the average $\vec{\mu}$, to get;

$$\frac{d\mu^+}{dZ} = \frac{d\mu^-}{dZ} = -\left(\frac{\mu^+ - \mu^-}{2\|V_z\|\tau}\right)$$

Either slope of the two Quasi Fermi Levels is given by separation of the two Quasi Fermi Level divided by the mean free path ($\lambda$);

$$\frac{d\mu^+}{dZ} = \frac{d\mu^-}{dZ} = -\left(\frac{\mu^+ - \mu^-}{\lambda}\right) \qquad (6.46)$$

Recall the current equation (6.13), which relates the current to the ballistic conductance times the Quasi Fermi Levels separation,

$$I = \frac{G_B}{q}(\mu^+ - \mu^-) \qquad (113) \rightarrow (\mu^+ - \mu^-) = \frac{qI}{G_B}$$

Replace $(\mu^+ - \mu^-)$ by $(\frac{qI}{G_B})$ in equation (6.46),

$$\frac{d\mu^+}{dZ} = \frac{d\mu^-}{dZ} = -\left(\frac{qI}{G_B\lambda}\right)$$

Turn the equation around, to write the current;

$$I = -\left(\frac{G_B\lambda}{q}\right)\frac{d\mu^+}{dZ} = -\left(\frac{G_B\lambda}{q}\right)\frac{d\mu^-}{dZ} \qquad (6.47)$$

Basically, the current is directly proportional to the slope of $(\mu^+)$, or the slope of $(\mu^-)$. To solve this current equation (6.47), for constant current to vary linearly, to get the variations of the two Quasi Fermi Levels, using the boundary conditions at the two ends;

$$\mu^+(Z = 0) = \mu_1$$

$$\mu^-(Z = L) = \mu_2$$

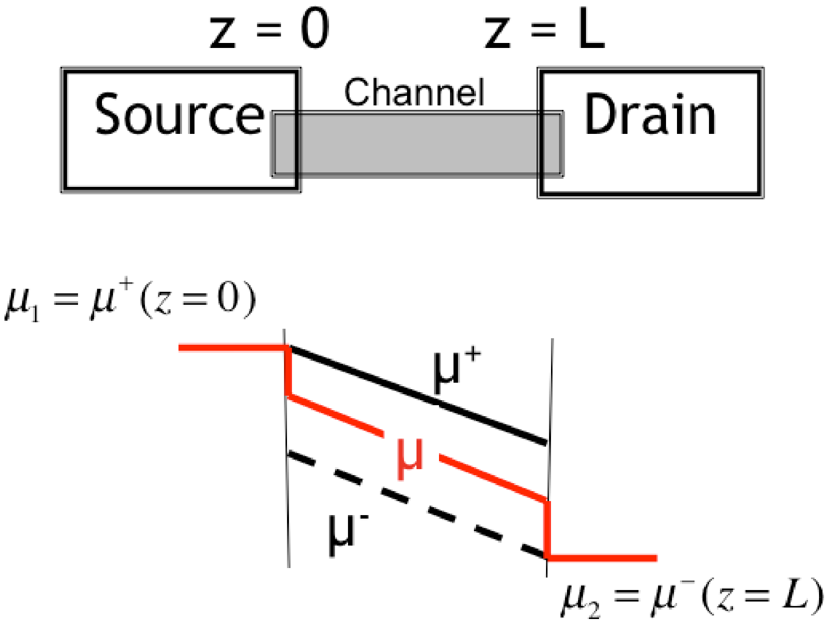


Figure 6.25: The average quasi Fermi level, for the right moving carriers and the left moving carriers.

Also, the two currents equation for the right moving electrons and the left moving electrons can be added to get a single equation for the average potential $(\mu)$,

$$\mu = \frac{(\mu^+ + \mu^-)}{2}$$

By using the actual conductivity, $\sigma_0 A = G_B \lambda$, to get the normal diffusion current equation;

$$I = \frac{\sigma_0 A}{q} \frac{d\mu}{dZ}$$

Solve this diffusion current equation using the new boundary conditions,

$$\mu(Z = 0) = \mu_1 - \frac{qIR_B}{2}$$

$$\mu(Z = L) = \mu_2 + \frac{qIR_B}{2}$$

Note that all of what has been discussed physically from their first principles, to see how they are working, can be done in a proper way mathematically using Boltzmann’s Equation with appropriate assumptions to tie them together, and derive everything step by step. That is why it is preferred not use Boltzmann’s Equation from the first beginning rather than to start with the intuitive descriptions first, because Boltzmann’s Equation gives a lot of insight and a little physical feeling towards the origin from which all these expressions come out.

# Chapter Seven

# Conclusions

## 7.1 Comments and Findings

The discussion started by a simple model consisting of a channel and two big contacts, the resistance of which has got two parts, one part is directly proportional to the length of the channel L, and the second part is an extra constant resistance independent of the channel length called the ballistic resistance. To show that this ballistic resistance is essentially associated with the contacts interfaces, a lot of very important and general concepts has been introduced.

The first of which is the tendency to associate any resistance with heat loss, but that does not give the insight about why this ballistic resistance is associated with the interfaces. The heat loss can be all over the place, largely in the two big contacts, because the heat loss is distributed spatially over an inelastic scattering length, from this point of view nobody can say that the resistance is here or there, it seems to be everywhere. The intuitive feeling of the resistance to be associated with obstacles, if a little hole was put in the channel, it will be so difficult for the electrons to get across to the other side, therefore, no doubt the resistance should go up. Then it quite clear that a criterion is needed to determine where these obstacles are, simply by following the voltage drop $(IR)$, and not to follow the heat loss $(I^2R)$. This is the first point to associate the voltage drop with the obstacle, even at atomic scale.

The second point is what is the voltage? Without answering this question in a way to associate the voltage definition with the electrochemical potential, nobody can be able to see the drops at the two ends of such device or model are associated with interface resistances. If the voltage is defined through the electrostatic potential the voltage drops at the two interfaces will be smeared out by the screening length, and all the information given by the electrochemical potential will be lost.

Fundamentally statistical equilibrium mechanics requires the electrochemical potential (ECP) to be the same everywhere at equilibrium, that is,

$\left(\frac{d\mu}{dZ} = 0\right)$ . On the other hand, fundamentally electrostatic potential (ESP) is not necessary to be constant even at equilibrium. In microscopic level all kinds of electrostatic potentials are there, even at equilibrium, this is very confusing to be associated with current flow.

Boltzmann's Equation is the key of semi-classical transport, which is what this dissertation is about, this equation can be used to model how these potentials (ECP & ESP) vary inside the device quantitatively, also there is an alternative quantum transport to Boltzmann's Transport Equation, which is the Non Equilibrium Green's Function (NEGF) including the quantum mechanics represented by Schrodinger equation instead of Newton's laws plus the thermodynamics effects. This NEGF is left as an extension to this research for future work. NEGF based calculations for the electrochemical potential and the electrostatic potential (see McLennan et al. Phys. REVB43.1384(1991). The electrochemical potential has sharp drops at the contacts while the electrostatic potential is more smoothed out.

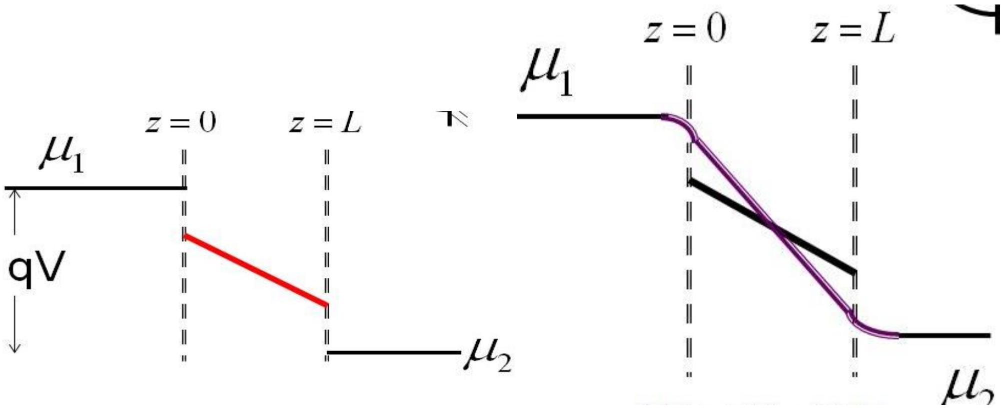


Figure 7.1: The electrochemical potential has sharp drops at the contacts (left), while the electrostatic potential is more smoothed out (right).

So, NEGF is a nice perfect way to calculate both the (ECP) and the (ESP), but still the approximate concepts like the (ECP) and the Quasi Fermi Levels are dealt with, because everybody should have to realize that by calculation, can gain a lot of insight, but a little to understand the physics behind it. Calculation is something and understanding is something else, they are not quite the same. For example, the Quasi Fermi Levels hep very much to provide the physical imagination of how to calculate what is expected from the simple model needed to explain the Quasi Fermi Levels behavior, so as not to be led astray. Another example the general approach to ignore the inelastic scattering, selection of small section of the device (channel) has been adopted leaving the inelastic scattering to be only in the contacts.

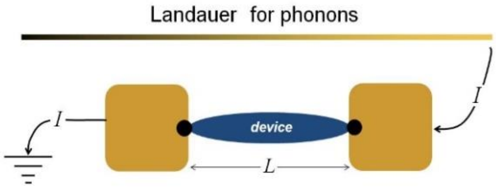


Figure 7.2: The inelastic scattering to be only in the contacts

The small section of the device could be looked at, as different independent energy channels or energy lanes. Any specified energy channel $E_1$ can be represented by a set of resistances in series, in between the two contacts. Similarly, there will be another set of resistances for the energy channel $E_2$, and so on

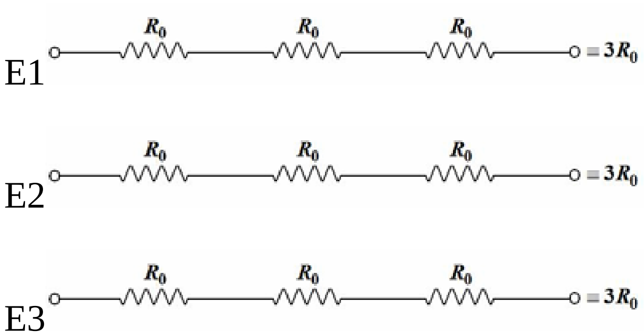


Figure 7.3: The small channel section of the device could be represented by different independent energy channels or energy lanes.

Bear in mind that the inelastic scattering is ignored.

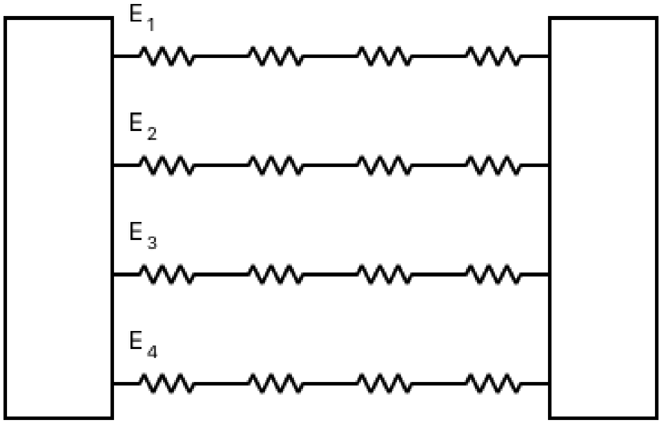


Figure 7.4: The inelastic scattering is ignored

The inelastic scattering conceptually tries to bring different energies into equilibrium. So, the inelastic scattering can be considered as a vertical resistor to connect one energy channel to the other. To include the inelastic scattering pictorially, some resistors to go between energy channel are added, and the potentials for one set of connected points can be considered one set of electrochemical potential appropriate for energies $(E_1, E_2 or E_3..)$.

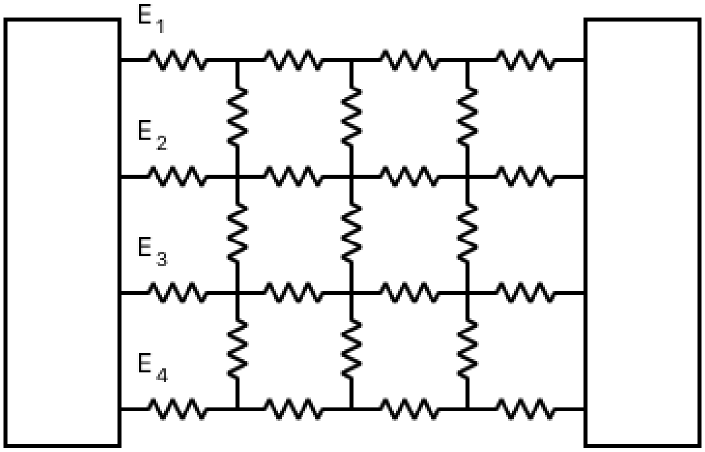


Figure 7.5: To include the inelastic scattering pictorially, the inelastic scattering considered as a vertical resistor to connect one energy channel to the other.

Similarly, another set of the connected points is another electrochemical potential appropriate for the same energies $(E_1, E_2 or E_3)$, and so on. The channel is out of equilibrium, and the distributions of the electrons is not necessarily given by the Fermi Function.

$$f(E) = \left(\frac{1}{1 + e^{\frac{E-\mu(E)}{KT}}}\right)$$

Where $\mu(E)$, is not necessarily constant, but It could be energy dependent, so that any function can be represented by Fermi function in a convenient way.

There should be different (ECP) at each energy level, so the inelastic processes cause a current flow. The model which ignores these vertical resistances might give certain current, more current may be expected if these resistances are added. It possible to have different situations to show these resistances are connected to contacts. In case all these energy levels are connected to the contacts and all the vertical resistances are added make all the energy levels at the same potential since all these vertical resistances are similar, inserting vertical connections between two points makes no difference because each resistor is connected between two equal potentials, and that is why for low bias neglecting an elastic scattering gives the exacts results given by solving Boltzmann's equation. On the other hand, a high bias nano transistor, the inelastic scatter makes a difference, because more current is given than what has been calculated otherwise. There are many ways of a complicated situations where the channels can be connected to different energies of the left-hand side contact, and on the other right side contact the channels are connected to different energies.

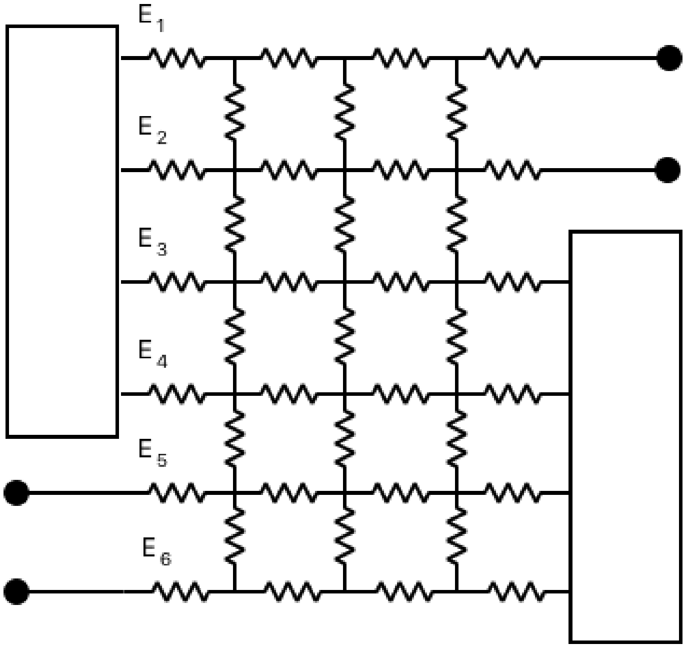


Figure 7.6: PN junction, where the left-hand side contact (the n-side) is connected to higher energy channel ($E1$) of the conduction band, and the right-hand side contact (p-side) is connected to lower energy channel ($E5$)

A more complicated situation to represent a PN junction where in the left hand side contact (the n-side) can be connected to higher energy channel ($E1$) of the conduction band, and at the right-hand side contact (p-side) is connected to different energy channel ($E5$), in that case the inelastic processes make a very big difference. If the inelastic processes are neglected there won't be any path from left to right. It is this inelastic process that make current flow possible in such structure. Making use of this concept of this electrochemical potential a pictorial view will show whether the inelastic processes will make a difference or not. A check can be carried by the comparison of the actual calculations and the experimental results to gain a good thinking idea about how these things are going.

What has just been explained is the different electrochemical potential for different energy channels, but what has been discussed in more details in this dissertation is the general idea of Quasi Fermi Levels which is a very useful concept in terms of understanding current flow in small devices. The QFL is the idea of having different potentials for right moving electrons and left moving electrons in a channel.

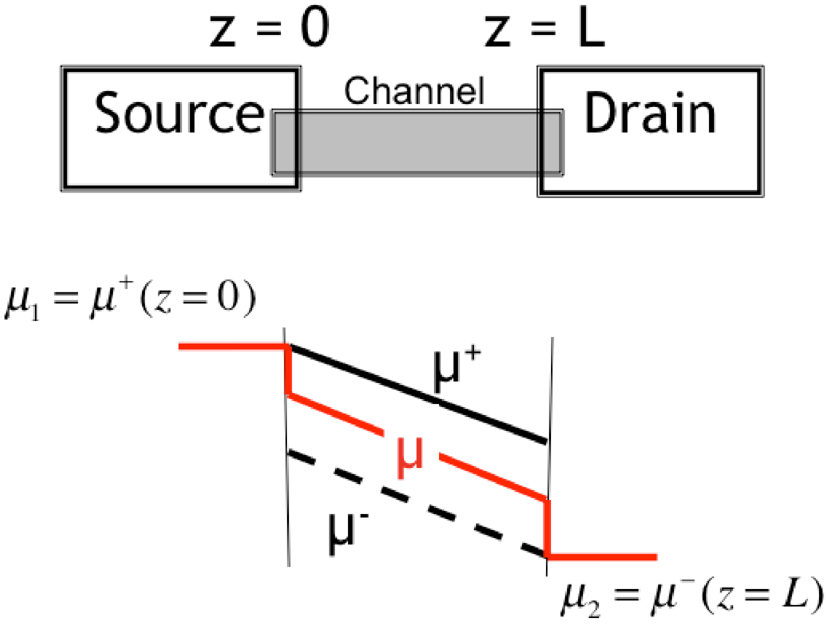


Figure 7.7: The QFL idea of having different potentials for right moving electrons and left moving electrons in a channel.

The idea of spin potentials is Avery wide area of spintronics, but a little has been discussed in this dissertation to talk about two different Quasi Fermi Levels for up spin and down spin, and the material with spin orbit coupling. What has been tried to be explained is how ($\mu^+$) gets associated with ($\mu^-$) up, and ($\mu^-$) gets associated with ($\mu^-$) down, and in general could have four QFLs. Two QFLs for up spin moving to the right and down spin moving to the left, and vice versa two QFLs for down spin moving to the right and up spin moving to the left. In general, QFL for particular momentum ($P_z$), and particular spin direction could have $\mu(Z, P_z, S)$. This gives a very useful physical picture in terms of thinking about how current flows on nanoscale. Thinking in terms of electrostatics potentials picture ($J = \sigma E = -\frac{d\Phi}{dX}$) that most people use, when intuition is guided by certain physical picture, all rich physics given by the QFLs will be missed. Therefore electrostatics picture is not rich and not general enough to capture all the new physics in the nano scale electronics being discovered in nano devices [79,80,81,82,83,84,85,86,87,88], or those which will be discovered in the coming years as researchers move along for a lot of more developments in terms of how they make contacts a kind of material to give exquisite control over different QFLs.

For quantitative benchmark Boltzmann's equation and NEGF are used, in order to have pictorial view of how current flows, control is carried over different groups of states, spins, right moving and left moving carriers, conduction bands, valance bands, and all different things needed for the concept of QFLs. In conclusion QFLs is a way of thinking that helps to understand or catalogue what could be seen.

## 7.2 Future work suggestions

Nanotechnology is extremely diverse, so that the research can be extended to include the photo electronic and thermal characteristics as well. There are a lot of suggestion for future work to continue studying the subject. One of the most important topics is the introduction of None Equilibrium Green Function method to analyze Quantum Transporting nanoscale devices to explore a number of topics within nano electronics area [89,90,91,92,93,94,95,96,97,98,99,100] for the sake of deep looking at the quantum transport to gain greater insight into the application of Schrodinger Equation and learning the basis of spintronics.

Another suggestion is to study the thermal electricity to see how current can be driven by temperature, this is something a little different from what had been discussed in driving currents by external voltage source.

A third suggestion is concerned with the study of the Graphene material since one of the clearest specialties of this material is its thermal management properties to help control the heat generated by electronics. The cooler the operating temperature of computers, tablets, or mobile phones, the better they run. The outcome of the continuous miniaturization of electronics is in need for enhanced thermal management, because there has been a large number of circuits to be packed very close together, their design should be miniaturized, and the heat should be minimized to disappear, in order for the electronics to work properly, this is one of the major strength opportunity of Graphene.

For thousands of years humans have been crafting the earth elements to shape the tools and technologies of our world. Graphene is getting to appear as a relatively new technology to realize its potential and its manufacturing applications in different areas.

Graphene or Angstrom material could change the technology of the future, only what is required is to be prepared mentally to get there. Graphene has been around ten years ago under investigation, to be applied for devices which has not been tried before to change industry for the better.

Graphene is a two-dimensional material composed of carbon atoms arranged in hexagonal manner, and it has unique properties can’t be found in any other material, it is six-sided shape which gives the material its extraordinary properties to conduct electricity and heat better than most materials. This improves electricity more efficiently from one place to the other.

There are a lot of other suggestions, but it is preferred to stick to those three mentioned suggestions above, since the subject is very wide open without limits. Recently remarkable developments happens to the field of nano electronics [101,102,103,104,105,106,107,108,109,110]. The researchers utilize quantum laws for development [111,112,113,114,115,116,117,118,119,120,121]. Many researches are done including experimental and theoretical models [122,123,124,125,126]. But till now no ultimate model describing all nono electronic phenomena has been established.

# Very important note

To access the references list, the following abbreviations should be consulted to understand what they are standing for;

**Et al**. is defined as an abbreviation for the Latin phrase **et** alia which means "and others." An example of **et**. **al**. used as an abbreviation is in the sentence, "The article was written by Smith, Jones, Paul, **et al**." which means that Smith, Jones, Paul and others wrote the article.

A **DOI**, or **Digital Object Identifier**, is a string of numbers, letters and symbols used to permanently identify an article or document and link to it on the web. A **DOI** will help your reader easily locate a document from your citation.

**URL** stands for Uniform Resource Locator, and is used to specify addresses on the World Wide Web. A **URL** is the fundamental network identification for any resource connected to the web (e.g., hypertext pages, images, and sound files). **URLs** have the following format: protocol://hostname/other_information.Jan 18, 2018

An **ISBN** is essentially a product identifier used by publishers, booksellers, libraries, internet retailers and other supply chain participants for ordering, listing, sales records and stock control purposes. The **ISBN** identifies the registrant as well as the specific title, edition and format.

The **International Standard Book Number** (**ISBN**) is a numeric commercial book identifier which is intended to be unique. Publishers purchase **ISBNs** from an affiliate of the International **ISBN** Agency.

An **ISBN** is assigned to each separate edition and variation (except reprintings) of a publication. For example, an e-book, a paperback and a hardcover edition of the same book will each have a different **ISBN**. The **ISBN** is ten digits long if assigned before 2007, and thirteen digits long if assigned on or after 1 January 2007. The method of assigning an **ISBN** is nation-specific and varies between countries, often depending on how large the publishing industry is within a country.

The initial **ISBN** identification format was devised in 1967, based upon the 9-digit **Standard Book Numbering** (**SBN**) created in 1966.

The 10-digit **ISBN** format was developed by the International Organization for Standardization (**ISO**) and was published in 1970 as international standard **ISO** 2108 (the 9-digit **SBN** code can be converted to a 10-digit **ISBN** by prefixing it with a zero digit '0').

Privately published books sometimes appear without an **ISBN**. The International ISBN Agency sometimes assigns such books ISBNs on its own initiative.[2]

Another identifier, the International Standard Serial Number (**ISSN**), identifies periodical publications such as magazines and newspapers. The International Standard Music Number (**ISMN**) covers musical scores.

**NIST** is the **National Institute of Standards and Technology**, a unit of the U.S. Commerce Department. Formerly known as the National Bureau of Standards, **NIST** promotes and maintains measurement standards. It also has active programs for encouraging and assisting industry and science to develop and use these standards.

Although it can be so used, it is a misconception that **ACME** was by the cartoonists intended an **acronym** standing for such things as "A Company Making Everything", "American Companies Make Everything" or "American Company that Manufactures Everything".

**IEEE**” stands for The Institute of Electrical and Electronics Engineers. **IEEE** style is a numbered referencing style that uses **citation** numbers in the text of the paper, provided in square brackets. A full corresponding **reference** is listed at the end of the paper, next to the respective **citation** number.N.TElectr

# References

[1] Gordon Moore,1988, Fellow computer History museum. Archived from original on January8,2015, Retrieved January 8,2015.

[2] Moore Gordon, (April,19,1965), Gramming more components on integrated circuits, Electronic magazine38(1):82-85.

[3] Moore Gordon, (Jan.,1998), Gramming more components on integrated circuits (Reprinted) proceeding of IEEE,86(1):82-85 dio:10 119/Jproc.1998-658762 Retrieved Jan.8,2015

[4] Gordon Moore at DBIP Biography server.

[5] Gordon Moore, Author profile page at ACM Digital Library

[6] Moore, G.E (1997), The microprocessors ‘Engine of technology Revolution’ Communications of ACM 40(2):112-114 dio:10 1145/25367.253746.

[7] Beaumont, Steven P. September 1996, ‘iii-v nanoelectronics’, microelectronics Engineering. 32(1): 283-295 dio:10 1016/0167-9317(95)00367-3[ISSN 01.67-9317].

[8] Colinge, Jean-Pierre, Colinge.C. A, (2005). Physics of semiconductor devices, Springer Science and Business media, p165, ISBN 9780 3872 85238

[9] EETime, 12 December 2018, ‘Transistors keep Moore law active’, Retrieved 18 July 2019.

[10] Supriyo Datta, 1995, Electronic transport in mesoscopic systems, Cambridge University Press, ISBN 0521416043.

[11] Brunglinghaus, Motion, 2011, mean free path, European nuclear society. Retrieved 8/11/2011.

[12] Jennings.S,1988, The mean free path, Journal of Aerosol science.

[13] F.U. Hill Brecht, 2001, photoemission spin polarized and angle resolved, Encyclopedia of materials science and technology.

[14] Sci-Tech Dictionary. Mc Graw Hill, Dictionary of Scientific and technical terms. 2003 Mc Graw Hill companies, Inc.

[15] The Mc Graw Hill companies, Inc,2005’ Mesoscopic physics’, Mc Graw Hill Encyclopedia of Science and Technology, Answers.com 25 January 2010.

[16] Cahy, M (2000), Quantum confinement IV: ‘Nano structured materials and devices’, preceding of international symposium. Cahy, M., Electrochemical Society Pennington, N.T Electrochemical Society, ISBN 978-1566773522. OCLC 4905 145.

[17] Harmans, C (2003) ‘Mesoscopic physics, an introduction, PDF, Open course ware TU Delft. Retrieved 14 June 2018.

[18] Science Learning Hub *,2016* .*"The microscopic scale"*. The University of Waikato. Archived from *the original* on 20 April 2016. Retrieved 31 March 2016.

[19] Jaeger, Gregg (September 2014). "What in the (quantum) world is macroscopic?". American Journal of Physics. **82** (9): 896–905. *Bibcode*:*2014AmJPh..82..896J*. *doi*:*10.1119/1.4878358.*

[20] Reif, F. (1965). *Fundamentals of Statistical and Thermal Physics* (International student edition. ed.). Boston: McGraw-Hill. p. *2*. *ISBN* *007-051800-9*. We shall call a system 'microscopic' (i.e., 'small scale') if it is roughly of atomic dimensions or smaller (say of the order of 10 *Å* or less).

[21] [21] Supriyo Datta, 29 June2012, Lessons from nanoelectronics, a new prospective on transport, publisher’s: WSPC. ISBN-13: 978-9814335294.

[22] [22] Wolfram Schommers, 2018, Basic physics of nanoscience, Traditional Approaches & new Aspects at ultimate level. University of Texas, USA.

[23] [23] Jesse Maassen, Mark Lundstrom, 7April 2015, Simple Boltzmann’s Transport Equation for ballistic to diffusive transient heat transport. Journal of applied physics.

[24] [24] Jung-Hoom, Mark Lundstrom, 2019, A drift-diffusion equation for ballistic transport in nanoscale MOSFET. Purdue University, USA. J

[25] [25] L.V. Keldysh, L. P. Kadanoff, G. Baym, 1965, “Diagram Technique for non-equilibrium processes” in Sov. Phys. JETP, Benjamin/Cummings, vol. 20, pp.1018

[26] [26] Supriyo Datta, 2015, NEGF method: A different prospective, International workshop in computational electronics (IWCE), 1-6.

[27] [27] Edward M. Purcell,1965, Electricity and magnetism. McGraw-Hill. PP.117-122. ISBN 987-0-07-004908-6

[28] [28] David J. Griffiths (1999). *Introduction to Electrodynamics*. Prentice-Hall. pp. *289*. *ISBN* *978-0-13-805326-0*.

[29] [29] Heaney, Michael B., January 2003. “Electrical conductivity and resistivity.” CRC press, pp 7-1 to 7-14, Ed: John G. Webster.

[30] [30] Data, Supriyo, 2005. Quantum Transport: Atom to transistor. Cambridge University press. P7. ISBN9 9780521631457.

[31] [31] Kittel, Charles; Herbert Kroemer, 15-1-1980. Thermal physics (second Edition). W.H. Freeman. P. 357. ISBN 978-0- 7167-1088-2

[32] [32] D. Chattopadhyay (2006). *Electronics (fundamentals And Applications)*. *ISBN* *978-81-224-1780-7*. *and* Balkanski and Wallis (2000-09-01). *Semiconductor Physics and Applications*. *ISBN* *978-0-19-851740-5*.

[33] [33] Shegelski, Mark R. A. (May 2004). *"The chemical potential of an ideal intrinsic semiconductor"*. American Journal of Physics. **72** (5): 676–678. *Bibcode*:*2004AmJPh..72..676S*. *doi*:*10.1119/1.1629090*. Archived from *the original* on 2013-07-03.

[34] [34] Calvin R. King, Jr, 2005, Introduction to microelectronics theory, Georgia Institute of technology.

[35] [35] Brennan, Kevin F. 1997, The physics of semiconductors, Cambridge University Press. New York. NY.

[36] [36] Neudeck and Pierret, 2003, “Advanced semiconductor fundamentals”, Peterson Education, Inc. Upper saddle river, NJ.

[37] [37] Socks, M. 1963, Solid state theory, New York, McGraw Hill Book Company pp. 159-160, 238-242.

[38] [38] Crouch, Stanley; Skoog, Douglas A. (2007). Principles of instrumental analysis. Australia: Thomson Brooks/Cole. *ISBN* *978-0-495-01201-6*.

[39] [39] Z. Ren, R. Venugopal, S. Datta, M. Lundstrom, D. Jovanovic, and J. Fossum, 2000, ``The ballistic nano transistor: a simulation study" *IEDM tech. Dig.*, pp.715.

[40] [40] G. Timp, J.Bude, K.K. Bourdelle, J.Garno, A.Ghetti, H. Gossmann, M.Green, G.Forsyth, Y.Kim, R.Kleiman, F.Klemens, A.Kornblit, C.Lochstampfor, W. Mansfield, S.Moccio, T.Sorsch, D.M. Tennant, W.Timp, R.Tung,1999, ``The ballistic nano-transistor", *IEDM Tech. Dig.*, pp.55-58.

[41] *[41] Srivastava, R. K.; Ashok, J. (2005). "Chapter 6". Statistical Mechanics*. New Delhi*: PHI Learning Pvt. Ltd.* ISBN 9788120327825.

[42] [42] Landauer, Rolf (August 1976). "Spatial carrier density modulation effects in metallic conductivity". Physical Review B. **14** (4): 1474–1479. Bibcode:1976PhRvB..14.1474L. doi:10.1103/PhysRevB.14.1474.

[43] LNE, 2015, Lessons from Nanoelectronics: A New Perspective on Transport, World Scientific.

[44] QTAT, 2005, Quantum Transport: Atom to Transistor, Cambridge

[45] ETMS, 1995, Electronic Transport in Mesoscopic Systems, Cambridge, (Paperback Edition).

[46] C.B. Parker, 1994, McGraw Hill Encyclopedia of Physics (2nd Edition), ISBN 0-07-051400-3

[47] Yuriy Kruglyak, 2014, Landauer-Datta-Lundstrom Generalized Transport Model for Nanoelectronics, Academic Editor: Xizhang Wang

[48] R. Landauer, 1957, “Spatial variation of currents and fields due to localized scatterers in metallic conduction,” International Business Machines Corporation. Journal of Research and Development, vol. 1, no. 3, pp. 223–231.

[49] R. Landauer, 1970, “Electrical resistance of disordered one-dimensional lattices,” Philosophical Magazine, vol. 21, pp. 863–867.

[50] pS. Datta, 2012, Lessons from Nanoelectronics: A New Perspective on Transport, World Scientific Publishing Company, New Jersey, NJ, USA, 2012.

[51] M. Lundstrom and J. Changwook, 2013, Near-Equilibrium Transport: Fundamentals and Applications, World Scientific, New Jersey, NJ, USA.

[52] J. M. Ziman, 1960, Electrons and Phonons. The Theory of Transport Phenomena in Solids, Clarendon Press, Oxford, UK.

[53] Britney Spears’ Guide to semiconductor physics, 2011, Density of states, Google’s cache of http://britneyspears.ac/physics/dos/dos.htm.

[54] Bart Van Zeghbroeck, 2004, Principles of semiconductor Devices, Colorado University – just.edu.jo.

[55] D.A.B. Miller, 2000, Semiconductor Optoelectronics Devices, Stanford University

[56] Blevins, Robert D. (2001). Formulas for natural frequency and mode shape (Reprint ed.). Malabar, Florida: Krieger Pub. *ISBN* *978-1575241845*.

[57] Tzou, H.S.; Bergman, L.A., eds. (2008). Dynamics and Control of Distributed Systems. Cambridge [England]: *Cambridge University Press*. *ISBN* *978-0521033749*.

[58] Shearer, Peter M. (2009). Introduction to seismology (2nd ed.). Cambridge: Cambridge University Press. pp. 231–237. *ISBN* *9780521882101*.

[59] NIST, 2019,*"2018 CODATA Value: conductance quantum"*. The NIST Reference on Constants, Units, and Uncertainty. *NIST*. *Retrieved 2019-05-20*.

[60] B.J. van Wees; et al. (1988). "Quantized Conductance of Point Contacts in a Two-Dimensional Electron Gas". Physical Review Letters. **60** (9): 848-850. *Bibcode*:*1988PhRvL..60..848V*. *doi*:*10.1103/PhysRevLett.60.848*. *hdl*:*1887/3316*. *PMID* *10038668*.

[61] S. Frank; P. Poncharal; Z. L. Wang; W. A. de Heer (1998). "Carbon Nanotube Quantum Resistors". Science. **280** (1744–1746): 1744–6. *Bibcode*:*1998Sci...280.1744F*. *CiteSeerX* *10.1.1.485.1769*. *doi*:*10.1126/science.280.5370.1744*. *PMID* *9624050*.

[62] S. Datta, 1997, *Electrical Transport in Mesoscopic Systems*, Cambridge University Press.

[63] R. Saito, 1998, *Physical Properties of Carbon Nanotubes*, Imperial College Press.

[64] pH.D. Cornean et al, Advances Materials (2004).

[65] K. Natori, 1994, ``Ballistic metal-oxide-semiconductor field effect transistor", *J.Appl. Phys.*, Vol. 76, pp.4879-4890.

[66] J.D. Bude, 2000, ``MOSFET modeling into the ballistic regime", *Proceedings of the International Conference on the Simulation of Semiconductor Processes and Devices (SISPAD) 2000*, pp.23-26.

[67] G. Fiori, G. Iannaccone, 2002, ``The effect of quantum confinement and discrete dopants in nanoscale 50 nm MOSFETs: a three-dimensional simulation", *Nanotechnology*, Vol. 13, p. 294-297.

[68] Supriyo Datta, 2011, Lessons from Nanoelectronics: A new approach to transport theory, Vol.1 in Lessons from Nanoscience: A Lecture Notes Series, World Scientific Publishing Company, Singapore.

[69] Mark Lundstrom, December 11, 2015, Fundamentals of nano transistor, Purdue University, west Lafayette, Indiana, USA.

[70] Kanoo, M. S. Lundstrom, 2000, electron transport in a model Si transistor, Purdue University, west Lafayette, IN 47907-1285, USA.

[71] lRamesh Venugopal, 2003, Modeling quantum transport in nanoscale transistor, Purdue University, USA.

[72] E. Przezdziecka, M. Stachowicz, S. Chusnutdinow, R. Jakela, A. Kozanecki, 2015, Appl. Phys. Lett. 106, 062106 (2015).

[73] R. Pietruszka, G. Luka, K. Kopalko, E. Zielony, P. Bieganski, E. Placzek-Popko, M. Godlewski, 2014, Mater. Sci. Semicond. Proc. 25, 190 (2014).

[74] S. Chusnutdinow, R. Pietruszka, W. Zaleszczyk, V.P. Makhniy, M. Wiater, V. Kolkovsky, T. Wojtowicz, G. Karczewski, 2014, Acta Phys. Pol. A 126, 1072 (2014).

[75] M.A. Pietrzyk, E. Zielony, M. Stachowicz, A. Reszka, E. Placzek-Popko, A. Wierzbicka, E. Przezdziecka, A. Droba, A. Kozanecki, J. Alloys Comp. 587, 724 (2014). [5] L. Sosnowski, B. Orlowski, Phys. Status Solidi A 3,1970

[76] Nelson, Jenny (2003-01-01). *The Physics of Solar Cells*. Imperial College Press. ISBN 9781860943492.

[77] J. M. Ziman, 1964, Principles of the Theory of Solids, Cambridge University Press, Cambridge, UK, View at: MathSciNet

[78] C. Kittel, 1976, Introduction to Solid State Physics, John Wiley & Sons, New York, NY, USA.

[79] N. W. Ashcroft and N. D. Mermin, 1976, Solid State Physics, Saunders College, Philadelphia, Pa, USA.

[80] K. Schwab, E. A. Henriksen, J. M. Worlock, and M. L. Roukes, 2000, “Measurement of the quantum of thermal conductance,” Nature, vol. 404, no. 6781, pp. 974–977.View at: Publisher Site | Google Scholar

[81] C. Jeong, R. Kim, M. Luisier, S. Datta, and M. Lundstrom, 2010, “On Landauer versus Boltzmann and full band versus effective mass evaluation of thermoelectric transport coefficients,” Journal of Applied Physics, vol. 107, no. 2, Article ID 023707.View at: Publisher Site | Google Scholar

[82] M. Lundstrom, 2012, Fundamentals of Carrier Transport, Cambridge University Press.

[83] C. Jeong, S. Datta, and M. Lundstrom, 2011, “Full dispersion versus Debye model evaluation of lattice thermal conductivity with a Landauer approach,” Journal of Applied Physics, vol. 109, no. 7, Article ID 073718.View at: Publisher Site | Google Scholar

[84] J. Callaway, 1959, "Model for lattice thermal conductivity at low temperatures," Physical Review, vol. 113, no. 4, pp. 1046–1051.View at: Publisher Site | Google Scholar

[85] M. G. Holland, 1963, "Analysis of lattice thermal conductivity," Physical Review, vol. 132, no. 6, pp. 2461–2471.View at: Publisher Site | Google Scholar

[86] C. Jeong, S. Datta, and M. Lundstrom, 2012, "Thermal conductivity of bulk and thin-film silicon: a Landauer approach," Journal of Applied Physics, vol. 111, no. 9, Article ID 093708.View at: Publisher Site | Google Scholar

[87] C. Gang, 2005, Nanoscale Energy Transport and Conversion: A Parallel Treatment of Electrons, Molecules, Phonons, and Photons, Oxford University Press, New York, NY, USA.

[88] C. J. Glassbrenner and G. A. Slack, 1964, "Thermal conductivity of silicon and germanium from 3∘ K to the melting point," Physical Review, vol. 134, no. 4A, pp. A1058–A1069.View at: Publisher Site Google Scholar

[89] J. B. Pendry, 1983, "Quantum limits to the flow of information and entropy," Journal of Physics A, vol. 16, no. 10, pp. 2161–2171.View at: Publisher Site | Google Scholar | MathSciNet

[90] D. E. Angelescu, M. C. Cross, and M. L. Roukes, 1998, "Heat transport in mesoscopic systems," Superlattices and Microstructures, vol. 23, no. 3-4, pp. 673–689.View at: Publisher Site | Google Scholar

[91] L. G. C. Rego and G. Kirczenow, 1998, "Quantized thermal conductance of dielectric quantum wires," Physical Review Letters, vol. 81, no. 1, pp. 232–235.View at: Publisher Site | Google Scholar

[92] M. P. Blencowe, 1999, "Quantum energy flow in mesoscopic dielectric structures," Physical Review B—Condensed Matter and Materials Physics, vol. 59, no. 7, pp. 4992–4998.View at: Google Scholar

[93] L. G. C. Rego and G. Kirczenow, 1999, "Fractional exclusion statistics and the universal quantum of thermal conductance: A unifying approach," Physical Review B: Condensed Matter and Materials Physics, vol. 59, no. 20, pp. 13080–13086.View at: Google Scholar

[94] V. Krive and E. R. Mucciolo, 1999, "Transport properties of quasiparticles with fractional exclusion statistics," Physical Review B—Condensed Matter and Materials Physics, vol. 60, no. 3, pp. 1429–1432.View at: Google Scholar

[95] C. M. Caves and P. D. Drummond, 1994, “Quantum limits on bosonic communication rates,” Reviews of Modern Physics, vol. 66, no. 2, pp. 481–537.View at: Publisher Site | Google Scholar

[96] R. S. Kim, 2011, Physics and Simulation of Nanoscale Electronic and Thermoelectric Devices, Purdue University, West Lafayette, Ind, USA.

[97] ^ *Sze, S. M. (1964).* Physics of Semiconductor Devices. *Wiley.* ISBN 978-0-471-05661-4.

[98] ^ *Sommerfeld, Arnold (1964). Thermodynamics and Statistical Mechanics. Academic Press.*

[99] ^ "3D Fermi Surface Site". *Phys.ufl.edu. 1998-05-27.* Retrieved 2013-04-22.

[100] ^ Technically, it is possible to consider the vacuum to be an insulator and in fact its Fermi level is defined if its surroundings are in equilibrium. Typically however the Fermi level is two to five electron volts below the vacuum electrostatic potential energy, depending on the work function of the nearby vacuum wall material. Only at high temperatures will the equilibrium vacuum be populated with a significant number of electrons (this is the basis of thermionic emission).

[101] Edward M. Purcell (1965). Electricity and Magnetism. McGraw-Hill. pp. 117–122. ISBN 978-0-07-004908-6.

[102] David J. Griffiths (1999). Introduction to Electrodynamics. Prentice-Hall. pp. 289. ISBN 978-0-13-805326-0.

[103] G. Stefanuci, R. van Leewuen, 2013, Non-Equilibrium Many-Body Theory of Quantum Systems, Cambridge.

[104] S. Datta,1995, Electron Transport in Mesoscopic Systems, Cambridge.

[105] S. Datta, 2012, Lessons from Nanoelectronics, World Scientific.

[106] S. Salahuddin, M.S. Lundstrom, S. Datta, 2005, "Transport Effects on Signal propagation in Quantum Wires", *IEEE Trans. Electron Dev.*, vol. 52, pp. 1734.

[107] P. Willke, T. Druga, R.G. Ulbrich, M.A. Schneider, M. Wenderoth, 2015, "Spatial Extent of a Landauer Resistivity Dipole in Graphene Quantified by Scanniag Tunneling Potentiometry", *Nature Comm.*, vol. 6, pp. 6399.

[108] M.J. McLennan, Y. Lee, S. Datta, 1991, "Voltage Drop in Mesoscopic Systems", *Phys. Rev. B*, vol. 43, pp. 13846.

[109] R. Lake, S. Datta, 1992, "Energy Balance and Heat Exchange in Mesoscopic Systems", *Phys. Rev. B*, vol. 46, pp. 4757.

[110] *Saghaei, Jaber; Fallahzadeh, Ali; Saghaei, Tayebeh (June 2016). "Vapor treatment as a new method for photocurrent enhancement*

*of UV photodetectors based on ZnO nanorods".* *Sensors and Actuators A: Physical.* *247: 150–155.* doi*:*10.1016/j.sna.2016.05.050.

[111] ^ "The microscopic scale". *Science Learning Hub. The University of Waikato. Archived from* the original *on 20 April 2016*. Retrieved 31 March 2016.

[112] Stop here

[113] Gordon Moore,1988, Fellow computer History museum. Archived from original on January8,2015, Retrieved January 8,2015.

[114] Moore Gordon, (April,19,1965), Gramming more components on integrated circuits, Electronic magazine38(1):82-85.

[115] Moore Gordon, (Jan.,1998), Gramming more components on integrated circuits (Reprinted) proceeding of IEEE,86(1):82-85 dio:10 119/Jproc.1998-658762 Retrieved Jan.8,2015

[116] Gordon Moore at DBIP Biography server.

[117] Gordon Moore, Author profile page at ACM Digital Library

[118] Moore, G.E (1997), The microprocessors 'Engine of technology Revolution' Communications of ACM 40(2):112-114 dio:10 1145/25367.253746.

[119] Beaumont, Steven P. September 1996, 'iii-v nanoelectronics', microelectronics Engineering. 32(1): 283-295 dio:10 1016/0167-9317(95)00367-3[ISSN 01.67-9317].

[120] Colinge, Jean-Pierre, Colinge.C. A, (2005). Physics of semiconductor devices, Springer Science and Business media, p165, ISBN 9780 3872 85238

[121] EETime, 12 December 2018, 'Transistors keep Moore law active', Retrieved 18 July 2019.

[122] Supriyo Datta, 1995, Electronic transport in mesoscopic systems, Cambridge University Press, ISBN 0521416043.

[123] Sci-Tech Dictionary. Mc Graw Hill, Dictionary of Scientific and technical terms. 2003 Mc Graw Hill companies, Inc.

[124] The Mc Graw Hill companies, Inc,2005' Mesoscopic physics', Mc Graw Hill Encyclopedia of Science and Technology, Answers.com 25 January 2010.

[125] Cahy, M (2000), Quantum confinement IV: 'Nano structured materials and devices', preceding of international symposium. Cahy, M., Electrochemical Society Pennington, N.T Electrochemical Society, ISBN 978-1566773522. OCLC 4905 145.

[126] Harmans, C (2003) 'Mesoscopic physics, an introduction, PDF, Open course ware TU Delft. Retrieved 14 June 2018.

## Appendix (I)

## Boltzmann Transport Equation (BTE)

The semiclassical transport results can be justified properly by Boltzmann's Transport Equation. The general Boltzmann's Transport Equation is a partial differential equation, can be written as;

$$\left[V\text{z}\frac{\partial f}{\partial z} + Fz\frac{\partial f}{\partial Pz} = -(f - f')\right] \quad (1)$$

Where, $f = f(Z, Pz, t)$

Z is the spatial location of electron

$P_z$ is the momentum of electron in Z- direction

t is the time

$V_{z.}$is velocity of the electron along the Z- direction

$F_z$ is the force of the electron in the Z- direction

$f$' is the occupation function at equilibrium state.

The origin of this equation is Newton's laws;

$$Vz = \frac{dZ}{dt}, \qquad Fz = \frac{dPz}{dt} \quad (2)$$

Equation (2) above describes the motion of a particle like an electron by Newton's laws. Where, $f = f(Z, Pz, t)$ in Boltzmann's equation can be defined by the following equation to tell the occupation of states with a certain momentum $P_z$ at a spatial location Z at a given time t;

$$f(Z, Pz, t) = \big( (Z - V\Delta t), (P - Fz\Delta t), (t - \Delta t)\big) \quad (3)$$

In equation (3) above Z, $P_z$, and t, are all independent variables. The left-hand side part must be equal to the right hand side part, because looking to the number of electrons in the present instant of time t, must have been passed the position they were in sometime $\Delta t$ ago, whenever they are now at time t, must have been at a little time ago at $(t - \Delta t)$ , therefore the momentum $P_{z,}$ now must have $(Pz -$

$Fz\Delta t$) a short time ago. This viewpoint is known as (collective description). This description evolved as shown by equation (3) above. Since the motion of each electron is shown by Newton's laws as in equation (2) above, whatever has been at time t, must have been in at time $\Delta t$ a little time ago. This statement connects the collective description viewpoint to note the only independent variable is t, so there is an important change in concept of equation (3), compared to the individual particles given by Ohm's law viewpoint, by assuming that $\Delta t$ is very small Taylor series can be applied, then;

$$f(Z,Pz,t) = f(Z,Pz,t) - \left[\frac{\partial f}{\partial t}\right] \qquad (4)$$

$$f(Z,Pz,t) = f(Z,Pz,t) - \left[Vz\frac{\partial f}{\partial z}\Delta t + Fz\frac{\partial f}{\partial Pz}\Delta t + \frac{\partial f}{\partial t}\Delta t\right]$$

Note that in equation (3) the left hand side is equal to the right hand side because of Newton's laws, and in equation (4) the two sides are equal from the mathematical view, because of Taylor series expansion with no physical contents, all the physics is found in equation (3). In equation (4), $f(Z,Pz,t)$ can be cancelled to give;

$$0 = -\left[Vz\frac{\partial f}{\partial z}\Delta t + Fz\frac{\partial f}{\partial Pz}\Delta t + \frac{\partial f}{\partial t}\Delta t\right] \qquad (5)$$

$When\ \Delta t \rightarrow \delta$ ;

$$0 = \left[\frac{\partial f}{\partial z} + Fz\frac{\partial f}{\partial Pz} + \frac{\partial f}{\partial t}\right] \quad (6)$$

Nothing evolved with time at steady state $\left(\frac{\partial f}{\partial t} = 0\right)$;

$$0 = \left[\frac{\partial f}{\partial z} + Fz\frac{\partial f}{\partial Pz}\right] \qquad (7)$$

Which is the left-hand side of Boltzmann's Equation, equation (7) is a reformulation of Newton's laws using collective variables. Boltzmann's Equation includes the entropy driven things or scattering, to describe irreversible processes. In principle Mechanics are represented by Newton's laws and the irreversible phenomenon needs an extra scattering term, which can be written as $(Sop\ f)$ as the scattering operator is acting on the function $f$, and the precise nature of the scattering operator requires a lot of discussion in general, therefore;

$$0 + (Sop\ f) = \left[\frac{\partial f}{\partial z} + Fz\frac{\partial f}{\partial Pz}\right] \qquad (8)$$

For the purpose of this thesis, the scattering operators assumed to be $-(f - f')$ to take $f$ and restore it to a local equilibrium $'$ , when $(f = f')$ there will be no scattering at all, if away from equilibrium then the scattering can be represented by a term proportional to how far the case is from equilibrium condition, this what is often referee to as the relaxation time approximation. Then the final form of Boltzmann's Equation is,

$$-\left(\frac{f - f'}{\tau}\right) = \left[\frac{\partial f}{\partial z} + Fz\frac{\partial f}{\partial Pz}\right] \quad 10$$

QED

# Appendix (II)

## Angular Averaging

To start from the basic expression of the conductance,

$$G = \frac{q^2 D}{2t}$$

Using the idea of time taken by the electron from the left contact (the source) to the right contact (the drain), to calculate the diffusive conductance then the ballistic conductance. This time has got two components one for diffusive transport directly proportional to the channel length squared, the other for ballistic transport directly proportional to the length of the channel.

*Using the diffusive part of time,*

$$G = \frac{q^2 D}{2t} = \frac{q^2 D Đ}{L^2}$$

Since for the diffusive transport Ohm's law applies without any modification, then;

$$R = \frac{\rho L}{W} \rightarrow G = \sigma W / L$$

$$\sigma = \frac{GL}{W} = \frac{q^2 D Đ L}{W L^2} = \frac{q^2 D Đ}{L w}$$

$(\frac{D}{LW})$ is the material property, which would not change by changing the material dimensions, since the density of states changes by the same amount of increment, so the ratio remains the same, since the density of states is given by energy per unit (length, area or volume) for 1D,2D and 3D respectively. Therefore, the density of states depends on (WL) for 2D.

*Using the ballistic part of time,*

$$G_B = \frac{q^2 D}{2t} = \frac{q^2 D \tilde{V}}{2\mathrm{L}}$$

From the last paragraph seen that D depends on (WL) for 2D, then the overall ballistic conductance does not depend on the length of the channel while it depends on the width only,

$$G_B = \frac{q^2 WL\tilde{\mathrm{V}}}{2L} = \frac{q^2 W\tilde{\mathrm{V}}}{2}$$

To average the velocity $\tilde{\mathrm{V}}$, consider the maximum velocity V going in different directions, what matters is the velocity from source to drain (to be considered as the z- direction) to which the average velocity is to be found by a little algebra. The answer depends on how many dimensions the average is to be taken.

For 1D conductor;

Here there is only one direction Z and there is nothing to average, it is just the maximum velocity V, so the factor of averaging is one. But for convenience the new prospective indicates that the fundamental material parameters;

a- the ballistic conductance,

$$G_B = \frac{q^2 D}{2t} = \frac{q^2 D\tilde{\mathrm{V}}}{2\mathrm{L}}$$

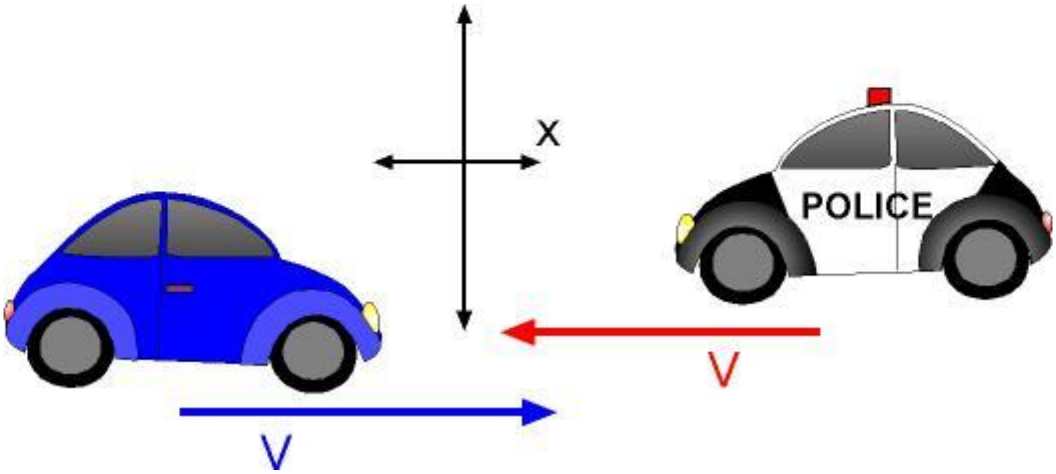
b- the conductivity multiplied by area is equal to the ballistic conductance times the mean free path,

$$\sigma A = G_B \lambda$$

c- The mean free path (λ), is defined as the ratio of the diffusion coefficient to the average velocity,

$$\lambda = \frac{2\text{Đ}}{\tilde{\mathrm{V}}}$$

Recall the elastic transport where an electron with certain energy corresponds to a certain velocity. But the velocity could be either in the positive or the negative Z- direction, and could have one of the two values (+ V) or (- V), because in elastic model an electron with a single energy at a time is considered, so one particular value of velocity, then an integral over energy is done. The average value of velocity has got two possible values (+V) or (-V), basically the average is just (V);



$$\tilde{V} = < V >$$

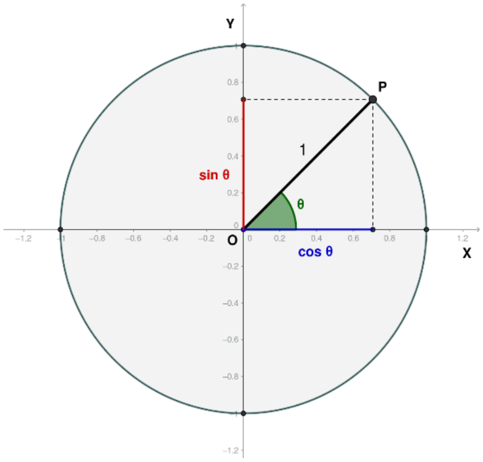
Then the average diffusion coefficient Đ entered the conductivity,

$$\sigma = \frac{q^2 Đ}{AL} D \ \ where, Đ = \ < V^2 \tau >$$

Recall ($\lambda = \frac{2Đ}{\tilde{V}}$), where λ is the mean free path for back scattering. In this context the completely turned around back scattering is interested in, because if the electrons are going straight, they are considered heading in the right direction from source to drain. Therefore, half of the electrons keep heading in the same direction, the other half turned around and come back, and the average time for getting turned around is not τ but 2τ. That is why there is a factor of two in front of the mean free path Vτ. This factor of two will be in different dimensions, right now it for 1D.

For 2D conductors,

In 2D conductor, consider the Z- direction between the source and drain. The direction of velocity can be at an angle to Z- direction, so the velocity along the Z- direction is the $V_Z Component$, which is the absolute value of $|V cos\ \theta|$.



Integrate from (-π to +π) and divide by the integral of (dθ), to get the average value of the ballistic velocity;

$$\tilde{V} = \left( \frac{\int_{-\pi}^{+\pi} |V \cos\theta| d\theta}{\int_{-\pi}^{+\pi} d\theta} \right) =$$

$$\left( \frac{V \int_{-\pi/2}^{+\pi/2} |\cos\theta| d\theta}{\int_{-\pi/2}^{+\pi/2} d\theta} \right) = \frac{2V}{\pi}$$

In 2D the average velocity for ballistic transport is given by ( $\frac{2V}{\pi}$ ), with an extra factor ($\frac{2}{\pi}$).

But for the diffusive transport the average velocity is involved in the diffusion coefficient,

$$Đ = V^2 \tau$$

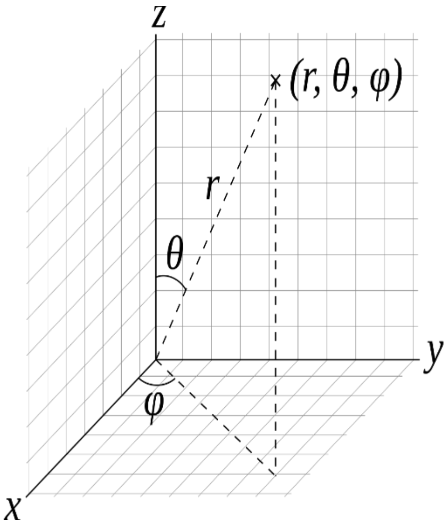
The average of that is quantity is to integrate from (0 to 2π) and divide by the integral of $(d\theta)$, to get the average value of the diffusive velocity;

$$Đ = \left( \frac{\int_0^{+2\pi} \tau V^2 \cos^2\theta \, d\theta}{\int_0^{+2\pi} d\theta} \right) = \frac{\tau V^2}{2}$$

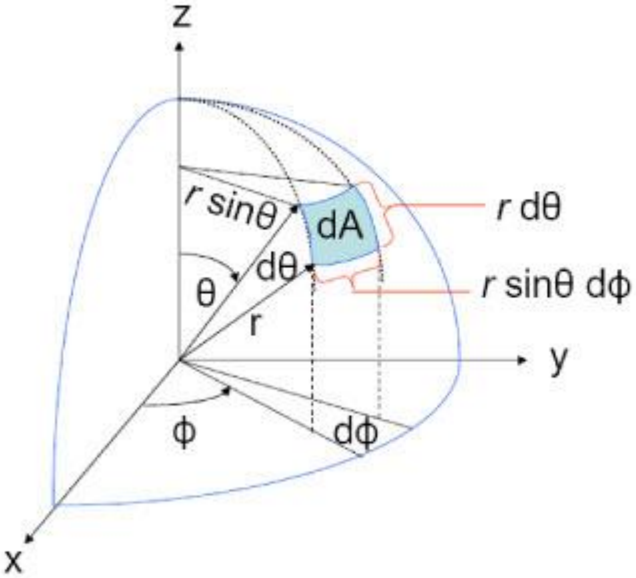
For 3D conductors:

This is the most complicated case mathematically. Here there are two angles, θ with respect to the Z-axis, when (r) is projected down to the (xy) plane, another angle Φ with the x- axis is formed. These two angles define a particular direction

in 3D.



How to integrate over an angle, define the solid angle



In 2D it quite simple, just integrate over $d\theta$, but in 3D this solid angle $(\int_0^{2\pi} d\Phi \int_0^{\pi} sin\theta d\theta)$ is needed, and the integral is done to find the average as follows;

$$\tilde{V} = \left( \frac{\int_0^{+2\pi} d\Phi \int_0^{+\pi} d\theta \, sin\theta |Vcos\,\theta|}{\int_0^{+2\pi} d\Phi \int_0^{+\pi} \sin\theta \, d\theta} \right) = \frac{V}{2}$$

The same procedure is followed for calculating the diffusion coefficient using $(\tau V^2 cos\,^2\theta)$. Pull $(\tau V^2)$ out and integrate $(cos\,^2\theta)$ over the solid angle $(\int_0^{2\pi} d\Phi \int_0^{\pi} sin\theta d\theta)$;

$$Đ = \left( \frac{\int_0^{+2\pi} d\Phi \int_0^{+\pi} \; sin\theta \, d\theta \, \tau V^2 cos^2\theta}{\int_0^{+2\pi} d\Phi \int_0^{+\pi} \sin\theta \, d\theta} \right) = \frac{\tau V^2}{3}$$

To collect all these results together, the average value of the velocity for 1D,2D, and 3D respectively;

$$\tilde{V} =< V >= V(1,\frac{2}{\pi},\frac{1}{2})$$

The average value of the diffusive coefficient for 1D,2D, and 3D respectively;

$$Đ =< \tau V^2 >= \tau V^2(1,\frac{1}{2},\frac{1}{3})$$

So, it like dividing by the number of dimensions.

In fact, there is a simple way can be used to get all these results without doing all these integrals.

The way to define the mean free path $(\lambda = \frac{2Đ}{\tilde{V}})$, λ can be expressed in all the three dimensions as;

$$\lambda = V\tau(2,\frac{\pi}{2}\,,\frac{4}{3})$$

The way the mean free path is defined Leeds to the simple expression;

$$G = \frac{G_B \lambda}{1 + \lambda}$$

Where $(\lambda = V\tau)$, as might be expected, but it is given multiplied by certain factors, depending on what number of dimensions the conductor has. The factors always greater than (one), because the mean free path is taken for back scattering where the electron turned around. Factor for 1D is (2), for 2D is (π/2) and for 3D is (4/3).

# Appendix (III)

## *Derivation of Relativistic Energy*

The distance, time, length, and momentum are all relativistic quantities, that the energy of an objected might be expected to be also relativistic, because the object energy depends on the inertial frame of reference.

Assume *work-energy principle* holds under relativity, the principle can be used to derive the energy.

By work – energy theorem, the amount of work done on particle to increase its velocity from (zero) to (V) is equal to;

$$W = \int_I^F F dx = \int_I^F \frac{dP}{dt}.dx = \int_I^F V dP \quad (1)$$

Where: I, is the initial velocity equals to zero

F, is the final velocity V

Since;

$$d(PV) = V.dP + \mathrm{P.dV} \rightarrow \mathrm{V.dP = d(VP) - P.dV}$$

Then;

$$W = \int_I^F V.dP = \int_I^F (d(VP) - P.dV) = \int_I^F d(VP) - \int_I^F P dV \quad (2)$$

Recall,

$$P = \frac{mV}{\sqrt{1 - (\frac{V}{C})^2}}$$

Evaluating the above integral,

$$W = \int_I^F (d(VP) - P.dV) = \int_I^F d(VP) - \int_0^V \frac{mV}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} dV \quad (3)$$

$$w = [PV]_0^v + \left[ m^2 C \sqrt{1 - \left(\frac{V}{C}\right)^2} \right]_0^v$$

$$W = \frac{mV}{\sqrt{1 - (\frac{V}{C})^2}} + mC^2 \sqrt{1 - \left(\frac{V}{C}\right)^2} \; - mC^2 \qquad (4)$$

Multiply the second term of equation (4) by; $\frac{\sqrt{1-\left(\frac{V}{C}\right)^2}}{\sqrt{1-\left(\frac{V}{C}\right)^2}} = 1$

$$\frac{mV^2 + mC^2 - mV^2}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} - mC^2 = \frac{mC^2}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} - mC^2 \quad (5)$$

By work – energy principle the work done on particle is equal to the kinetic energy it gains.

## *Energy momentum relation*

The relationship between energy and momentum of a particle is exactly,

$$E = K.E + mC^2 \; (i)$$

$$E = \frac{mC^2}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} \qquad (ii)$$

$$P = \frac{mV^2}{\sqrt{1 - \left(\frac{V}{C}\right)^2}} \qquad (iii)$$

Square both sides of equation (ii)

$$E^2 = \frac{(mC^2)^2}{1 - \left(\frac{V}{C}\right)^2} \quad (a)$$

Insert the term ($V^2 - V^2 = 0$) as shown here below,

$$E^2 = \frac{m^2C^2(V^2 - V^2 + C^2)}{1 - \left(\frac{V}{C}\right)^2}$$

$$E^2 = m^2C^2V^2/\gamma^2 - \ m^2C^2V^2/\gamma^2 + m^2C^4 \ \ (b)$$

$$E^2 = (mV/\gamma)^2C^2 + \ m^2C^4 - m^2C^2V^2/\gamma^2 \ \ (c)$$

$$E^2 = (P)^2C^2 + \ m^2C^2(C^2 - V^2)/\gamma^2) \ \ \ (d)$$

$$E^2 = (P)^2C^2 + \ m^2C^4(C^2 - V^2)/(C^2 - V^2) \ \ \ (e)$$

$$E^2 = (P)^2C^2 + \ m^2C^4$$

Where E is the total energy of the particle, then;

$$m^2C^4 = E^2 - (P)^2C^2$$

Implies that it will be identical in all reference frames.

# Appendix (IV)

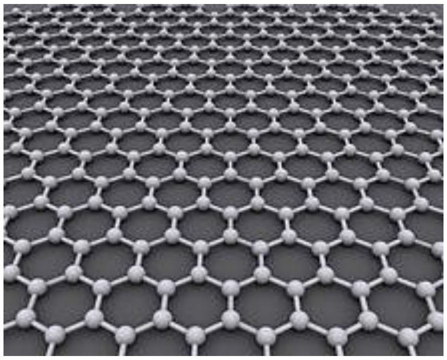
## Graphene

Scientists have been puzzling over graphene for decades. Back in 1947, Canadian physicist Philip Wallace wrote a paper about the electronic behavior of graphite that sparked considerable interest in the field. Nobel-Prize winning chemist Linus Pauling was speculating about how flat, single layers of carbon atoms would behave as long ago as 1960. In 1962, such materials were named "graphene" by German chemist **Hans-Peter Boehm**, who had spotted them under an electron microscope.

Theoretical research into graphene continued for the next four decades, boosted in the 1980s and 1990s by the discoveries of fullerenes (effectively, graphene curled up into balls) and carbon nanotubes (graphene folded into a pipe). Even so, no-one could ever actually make the stuff in practice; graphene was only produced in a laboratory in 2004, by Russian-born scientists Andre Geim and Konstantin Novoselov working at the UK's University of Manchester. They made graphene by using pieces of sticky tape to pull off flakes of graphite, then folding the tape and pulling it apart to cleave the graphite into even smaller layers. Eventually, after a great deal of work, they were amazed to find they had some bits of graphite only one atom thick graphene.

Four years later, the Manchester team managed to create a graphene transistor just one atom thick and ten atoms wide. The same year, workers at Rice University in the United States built the first graphene-based flash memory. In recognition of the huge importance of their work, Geim and Novoselov were awarded the 2010 Nobel Prizing physics.2010 Physics.

*Graphene should not to be confused with Graphite Grapheme, or Graphane.*



Graphene is an atomic-scale hexagonal lattice made of carbon atoms.

**Graphene** is an allotrope of carbon in the form of a single layer of atoms in a two- dimensional hexagonal lattice in which one atom forms each vertex. It is the basic structural element of other allotropes, including graphite, charcoal, carbon nanotubes, and fullerenes. It can also be considered as an indefinitely large aromatic molecule, the ultimate case of the family of flat polycyclic aromatic hydrocarbons.

## General properties

Graphene is an amazingly pure substance, its simple, orderly structure based on tight, regular, atomic bonding, Carbon is a nonmetal, so you might expect graphene to be one too. In fact, it behaves much more like a metal (though the way it conducts electricity is very different), and that's led some scientists to describe it as a semimetal or a semiconductor (a material mid-way between a conductor and an insulator, such as silicon and germanium). Even so, it's as well to remember that graphene is extraordinary—and quite possibly unique.

## Strength and stiffness

Graphite is horribly soft. That's because the carbon layers inside a stick of graphite shave off very easily. But the atoms *within* those layers are very tightly bonded so, like carbon nanotubes (and unlike graphite), graphene is super-strong, even stronger than diamond! Graphene is believed to be the strongest material yet discovered; some 200 times stronger than steel. Remarkably, it's both stiff and elastic (like rubber), so it can be stretched by an amazing amount (20-25 percent of its original length) without breaking. That's because the flat planes of carbon atoms in graphene can flex relatively easily without the atoms breaking apart.

No-one knows quite what to do with graphene's super-strong properties, but one likely possibility is mixing it with other materials (such as plastics) to make composites that are stronger and tougher, but also thinner and lighter, than any materials we know. Imagine an energy-saving car with super-strong, super-thin, super-light plastic body panels reinforced with graphene; that's the kind of object we might envisage appearing in a future turned upside down by this amazing material!

## Heat conductivity

As if super strength and featherweight lightness aren't enough, graphene is better at carrying heat (it has very high thermal conductivity) than any other material—better by far than brilliant heat conductors such as silver and copper, and much better than either graphite or diamond. Again, we're most likely to discover the benefit of that by using graphene in composite materials, where we could use them to add extra heat-resistance or conductivity to plastics or other materials.

## Electrical conductivity

This is where graphene starts to get interesting! Materials that conduct heat very well also conduct electricity well, because both processes transport energy using electrons. The flat, hexagonal lattice of graphene offers relatively little resistance to electrons, which zip through it quickly and easily, carrying electricity better than even superb conductors such as copper and almost as well as superconductors (unlike superconductors, which need to be cooled to low temperatures, graphene's remarkable conductivity works even at room temperature). Scientifically speaking, we could say that the electrons in graphene have a longer mean free path than they have in any other material (in other words, they can go further without crashing into things or otherwise being interrupted, which is what causes electrical resistances). What use is this? Imagine a strong, light, relatively inexpensive material that can conduct electricity with greatly reduced energy losses: on a large scale, it could revolutionize electricity production and distribution from power plants; on a much smaller scale, it might spawn portable gadgets (such as cellphones) with much longer battery life.

## Electronic properties

Electrical conductivity is just about "ferrying" electricity from one place to another in a relatively crude fashion; much more interesting is manipulating the flow of electrons that carry electricity, which is what electronics is all about. As you might expect from its other amazing abilities, the electronic properties of graphene are also highly unusual. First because the electrons are faster and much more mobile, which provides the possibility of computer chips that work more quickly (and with less power) than the ones we use today. Second, the electrons move through graphene a bit like **photons** (wave-like particles of light), at speeds close enough to the speed of light (about 1 million meters per second, in fact) that they behave according to both the theories of relativity and quantum mechanics, where simple certainties are replaced by puzzling probabilities.

After this introduction about graphene, this research is deeply concerned with of how to get an expression for conductivity that is directly proportional to

$\sqrt{n}$¸ the way to this is to start from the general standard Dude formula for conductivity given by;

$$\sigma = \frac{q^2 \tau n}{m}$$

Where, n is the density of electrons per unit (length, area, or volume),

$$\sigma = \frac{q^2 \tau N \left(\frac{1}{L}, \frac{1}{wL}, \frac{1}{AL}\right)}{m}$$

The mobility μ is given by;

$$\mu = \frac{q\tau}{m} = \frac{\sigma}{qn}$$

Where, $qn$ is the carrier density

$m$ is the mass of the electron

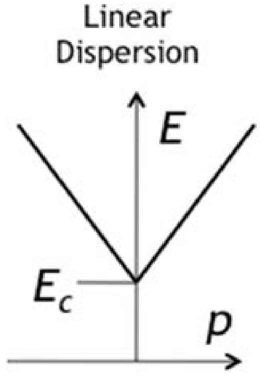
$q$ is the electron charge

$\tau$ is the mean free time

Graphene has no parabolic energy – momentum relation, it is rather linear relation,

$$E = V_0 P$$

The mass in this case should be treated with care, in the sense that mass is not constant. In graphene the mass in with the increase in energy.



Experimentally if the electrochemical potential raised to an upper energy level from where it has been before, the mobility μ keeps going down, not because the mean free time is going shorter, but the mass is increasing because the electrons in the upper electrochemical potentials are more heavier than the electrons in the lower chemical potentials. In graphene the conductivity σ is not directly proportional to the electron density n, but it is expected to be directly proportional to the square root of the electron density n, because the momentum P is related to the counted number of states by the relation,

$$N = (\frac{P}{h})^d \; (2L, \pi wL, 4\pi AL)$$

The graphene is a 2D conductor, then,

$$N = \left(\frac{P}{h}\right)^2 \; (\pi wL) \;\; (1)$$

Since $\sigma = \frac{q^2 \tau n}{m} = \frac{q^2 \tau N}{m}\left(\frac{1}{L}, \frac{1}{wL}, \frac{1}{AL}\right)$, then for 2D, $n = \frac{N}{wL}$?

$$\sigma = \frac{q^2 \tau n}{m} = \frac{q^2 \tau}{m}\left(\frac{N}{wL}\right) \;\; (2)$$

$$E = V_0 P \rightarrow \frac{dE}{dP} = v = V_0 = constant$$

$$m = \frac{p}{V_0}$$

$$\sigma = \frac{q^2 \tau n}{m} = \frac{q^2 \tau V_0}{P}\left(\frac{N}{wL}\right) \quad (3)$$

To relate P back to n, eliminate P from equation (3) using equation (1),

$$N = \left(\frac{P}{h}\right)^2 (\pi wL) \; \rightarrow P = h\sqrt{\frac{N}{\pi wL}}$$

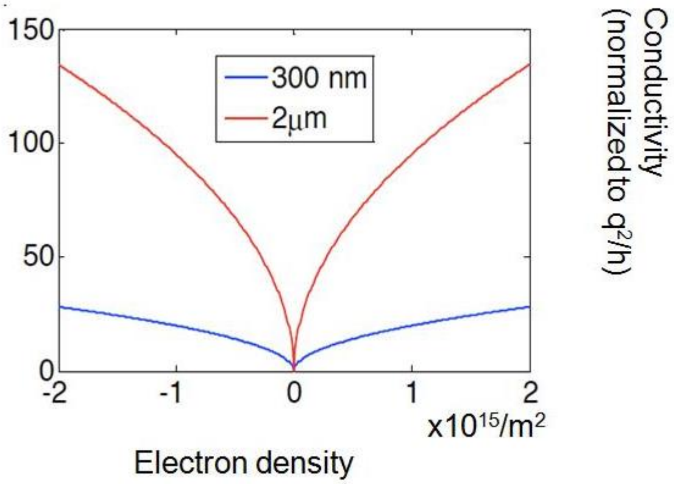
Then,

$$\sigma = \frac{q^2 \tau V_0}{h}\left(\sqrt{\frac{\pi N}{wL}}\right) \quad (4)$$

Since $(\tau V_0)$ is the mean free path,

$$\sigma = \frac{q^2 \lambda}{h}\sqrt{n\pi}$$

The degeneracy factor is taken equal to four, which is equal to spin multiplied by the number of valleys which is two for graphene. The graph hereafter is the plot of conductivity as a function of electron density for graphene with degeneracy factor of four, mfp λ equals 300 nm, and mfp equals 2μm.



*Experiment: See for example, Bolotin et al. PRL 101, 096802 (2008)*