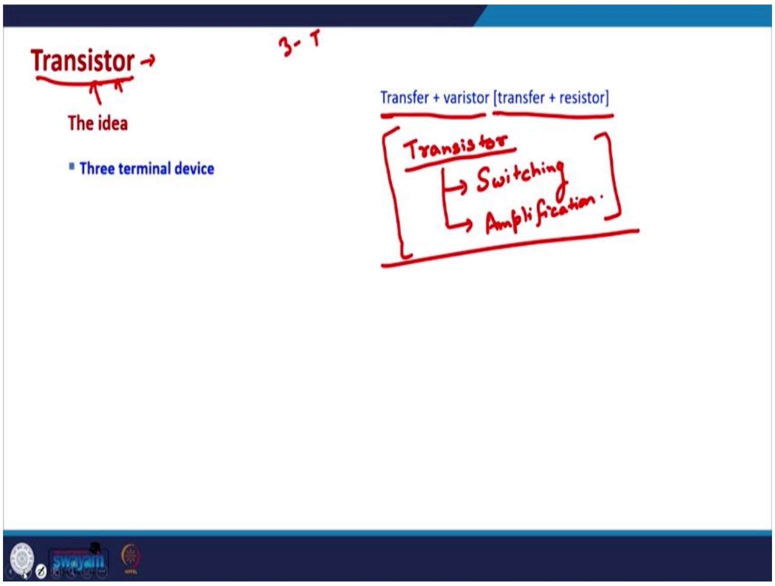
Physics of Nanoscale Devices
Prof. Vishvendra Singh Poonia
Department of Electronics and Communication Engineering
Indian Institute of Technology, Roorkee

Lecture - 01
Introduction

Hello everyone. Welcome to the opening lecture on the course on Physics of Nanoscale Devices and this lecture is mostly going to be about the Introduction to this course. And in this lecture I will try to put forward the perspective that we will be taking while reading this course while sort of studying this course why this course is there.

So, as the name of the course suggests that this course is about electronic devices and specifically the nanoscale devices and the electronic device that we are concerned about or we are mostly concerned about is the transistor.

(Refer Slide Time: 01:05)



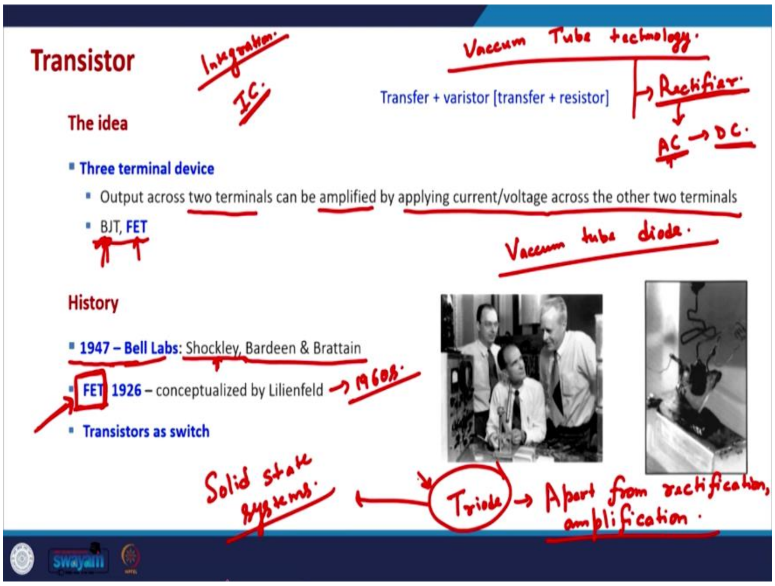
Let me take a moment and let me allow you to think what is a transistor. So, if you think about a transistor, you I guess you would think that it is a 3 terminal device it is a 3 terminal device. And if we sort of try to understand the name transistor the name sort of says that it is or the or there are two versions in which this name can be understood one is the transfer varistor and second is the transfer resistor.

So, the transistor is a 3 terminal device in which the current is transferred from a low resistor regime to high resistor regime. So, that was actually that was the initial understanding of the transistor, but modern transistors are quite different than that. And so, the way to sort of define modern transistor is that transistor is a device in a very simple terms transistor is a device that can achieve switching and amplification.

So, transistor is a 3 terminal device that can achieve switching and amplification and these two operations are extremely crucial for all the electronic devices for the; for let us say for the cell phones, for the computers that we use that just by various combination of these operations we can achieve a lot of computational tasks.

And as all of us know that transistor is essentially at the foundation of the electronics industry it is at the foundation of I would say the electronic electronics revolution that the 20th century so on. And it will also not be an exaggeration that that the development of transistor technology also draw the revolution in computing technologies. So, let us maybe have a quick look at the history of the transistor.

(Refer Slide Time: 03:24)



So, the idea of the transistor is that it is a three terminal device and it can achieve two operations, it can achieve switching and it can achieve amplification or in other words that output across the two terminals can be amplified by applying current or voltage across the other two terminals.

So, the output across two terminals can be switched on and off or amplified by the third terminal in other words. And there are two types of transistors I guess most of us know about this one is the BJT the bipolar junction transistor and second is the FET. So, let us and I must say that the importance of transistor cannot be overestimated in electronics and computing technologies, it is actually at the foundation of these two technologies ok.

And that is why the transistor is also by some people it is also call as the most important invention of the 20th century. Because it is because of the transistor that we have our cell phones and computers all those devices with us ok.

So, the history of that transistor if we look a little bit at the history of the transistor the first demonstration of transistor was done in 1947 at Bell Labs and it was done essentially by Bardeen and Brattain and very quickly Shockley was also in the same lab and he demonstrated a slightly different kind of transistor. Bardeen and Brattain demonstrated the point contact transistor and Shockley demonstrated the which is nowadays known as the NPN transistor.

So, this is the moment this is the late 1947 when the transistor was officially demonstrated, but there is a large or I would say a long history behind the development of transistor; transistor essentially stands on the foundations of the electromagnetic theory, the theory of electricity and the theory of magnetism.

And the I would say that the formal study of electricity and magnetism started around 1600s. So, it took around 350 years since the formal study of electricity and magnetism that we had the invention of the transistor in 1947. And the just precursor of the transistor is the vacuum tube technology and that is essentially this is just a history bit I would say vacuum tube technology is actually considered to be the starting point of electronics engineering or electronics as a discipline.

And why did we need vacuum tube technologies? Why did we need vacuum tubes? The reason was that that since Maxwell gave the formulation of electromagnetism, the exact mathematical formulation of electromagnetism. And then Hertz demonstrated the existence of the physical existence of electromagnetic waves after that Marconi and JC Bose they invented a radio system in which the electromagnetic signals can be transmitted and received.

So, in those receivers we actually it was needed to convert the AC signal to DC signal. So, the rectifier was needed and not only in the radio waves in not only in the radio wave technology the rectifier was needed even rectifier was needed to convert the AC signal to DC signal actually the AC current to DC form of the current ok.

Because the by that time, by early by late 19th century and early 20th century the production of AC type of current was quite common and it could be sent to large distances it many countries or at least the US had a good network of AC distribution and, but many appliance many electronic components needed DC current and it was difficult and expensive to produce and send the DC current DC form of current and.

So, that is why this conversion from AC to DC was quite crucial and this was achieved using rectifiers and the first device out the first electronic device that achieved rectification was the vacuum tube diode. So, that is essentially the beginning of the electronics engineering and soon after the vacuum tube diode a vacuum cube triode was proposed and in vacuum tube triode apart from the rectification the amplification could also be achieved.

So, apart from rectification amplification could also be achieved in the triodes vacuum tube triode. So, triode was a three terminal device and the current between the two terminals could be controlled by applying a voltage or electric field on the third terminal and this is actually the precursor to transistors actually. Soon the computer and television also were invented and initially the computer and televisions specifically the computers were based on just diodes and triodes.

And as soon as the systems started became larger it was extremely difficult to maintain vacuum tube based computing systems. In fact, they were very huge in size maintenance was extremely costly and because very frequently there used to be malfunctions in the tubes and it was extremely difficult to detect the non functioning tube as well. So, also the heating was a big problem the space was a problem because they used to take a lot of space.

So, there were many, I would say limitations of vacuum tube technology specially while making a computer that people faced. So, people always thought about making more compact systems more less energy intensive systems and that is why this idea of making sort of a triode based on semiconductors based on solid state systems or making a triode like device using solid state systems was always in mind since these diodes and triodes started becoming popular in electronics equipments ok.

So, after World War II finally, in 1947 although the work was there it was since long time that people were trying to make a solid state transistor, but soon after World War II and soon after World War II in bell labs this was demonstrated a point contact or NPN light transistor was demonstrated in 1947.

And I must also add that it was not only a scientific achievement it also led to a huge economic boom for specially for US in the beginning and then other countries. So, this NPN or BJT like transistor was demonstrated in late 50s and up to in initial part of 60s 1960s it was quite common, but as soon as the system started became as soon as the computers were made out of this transistor the need was felt to put more number of transistors on a single sort of single chip on a single platform and that is where this idea of integration came about.

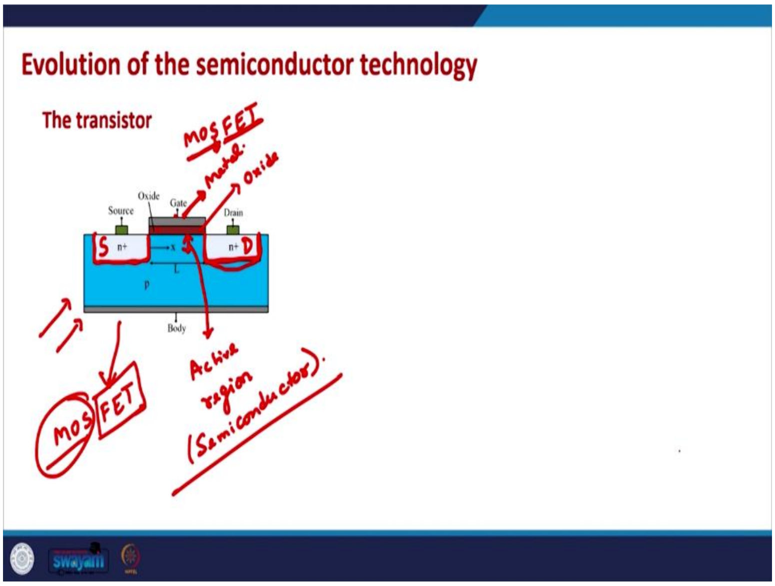
Actually, it was only after integration only after many different electronic equipments can be made on a single chip only after that the electronic equipments could be scaled faster because again it was even after the invention of transistor it was a challenge to connect various transistors and to achieve or define the functionality out of them. And in that sense this transistor BJT was not so appropriate it had many, it had. So, for example, the base current was there which was a bit of a problem and.

So, that is why this field effect transistor in which this the third terminal of the transistor the control terminal of the transistor was using electric field to control the current and voltage across with two other terminals this was then invented in 1960s early 1960s. although the idea of field effect transistor was quite old it was in 1926 that this field effect transistor was conceptualized by Lilienfeld.

But it was finally, invented in 1960s and after 1960s, the invention of IC integrated circuit that also happened and this IC technology was defined up to 1970s and in 19s starting from 1970s this growth of electronics has been I would say tremendous both in terms of application and in terms of economics.

So, this finally, nowadays we deal with this kind of mostly we deal with field effect transistors although BJT are also in use, but they have very limited applications and. So, this is going to be the central topic of our course ok.

(Refer Slide Time: 14:46)



And with this bit of history let us see how this semiconductor technology evolved. So, this is a one kind of field effect transistor and this is known as MOSFET ok. So, MOSFET means metal oxide semiconductor field effect transistor. The name field effect transistor means that we use electric field in order to control the voltage or current across the two other terminals of the transistor.

Metal oxide semiconductor means that we have a metallic gate, the gate is metallic oxide means just below the metal we have a very thin oxide although that is not very clear from this picture, but this oxide is typically very thin. And, below oxide just below oxide we have the semi conductor where the actual active region of the device is.

So, this is the active region of the device and this is made out of semiconductor material ok. So, we have a stack sort of we have a stack of semi conductor oxide metal and in this apart from this stack we have the contacts, the left contact and the right contact, these are also semi conducting contacts.

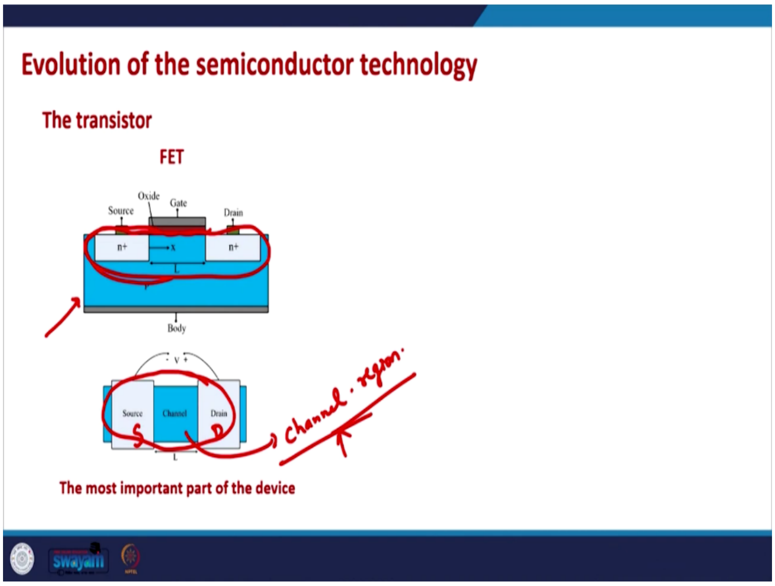
And so, the left contact by convention is termed as source contact and the right contact by convention is known as the drain contact. So, in this way we have source, drain in between we have a semiconductor on top of that semiconductor we have an oxide and a metallic gate.

So, this entire structure is known as the metal oxide semiconductor field effect transistor. So, metal oxide semiconductor is here from here by and by this name FET we mean that by applying a voltage on the metallic terminal the gate terminal so called gate terminal we can control current and voltage in the between the source and the drain.

So, the three terminals of the device are the source the drain and the gate. The active region so to say where the current actually flows is between the source and the drain and this current and the voltage or this current particularly can be controlled by the electric field that we apply on the gate that we produce by applying a voltage on the gate. And that is why this device is known as the MOSFET device Metal Oxide Semiconductor Field Effect Transistor ok.

So, this is the device that is the modern transistor, modern nowadays the transistor is slightly modified, but this is the essentially the functionality of modern transistors is quite similar to this MOSFET transistor ok.

(Refer Slide Time: 18:08)

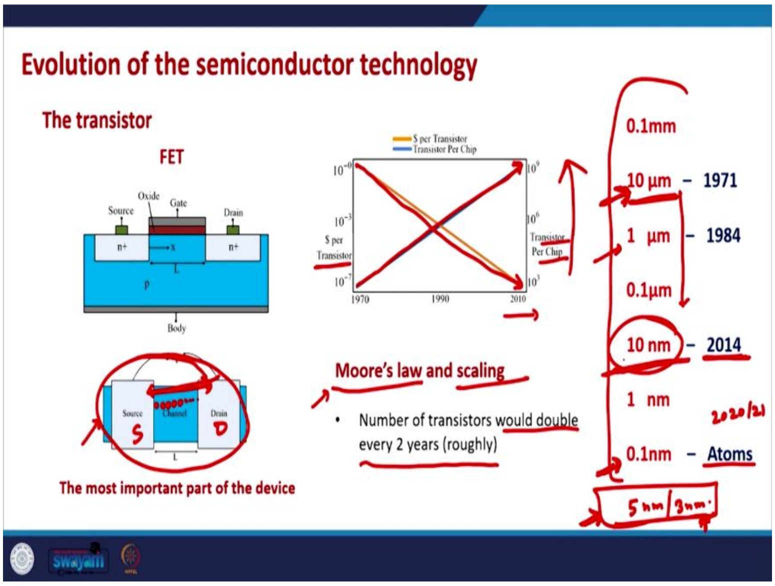


So, let us see how this has evolved over the years and as I told you that the most important part of the device is or this is the active region of the device this is the most important part of the device. It is essentially the region of the device between the source contact and the drain contact and the region between the source and the drain is known as the channel region and that is the generally that can be considered to be the most important part of the device.

Although the gate is also quite important because gate controls the electric field in the channel and gate controls or gate sort of switches on or off the current in the channel. So, that is, it is not that that is not important. But generally, a very interesting physics takes place in the channel region and that is what we will try to understand in this course. We will try to understand the physics of the channel region because of the voltage applied on the source on the drain or on the gate ok.

So, as we just saw that this device was finally, invented in 1960s and initially the size and then after integration in 1970s, late 1970s or 80s this was actually quite common and various inventions or computing systems made on top of this.

(Refer Slide Time: 19:34)



And very soon this active region or this channel region of the device was getting scaled down basically. The reason was that as the time was progressing because of the demand of high performance because of the demand of putting more transistors in a single device so that we can have more functionality in a single device.

This transistor was being made smaller and smaller and this trend of making the transistor small and smaller, this trend of scaling the transistor size down is captured by what is known as the Moore's law. So, Moore's law essentially states that the number of transistors in a unit space or you know on a unit area of the chip would double almost every 2 years roughly 18 months to 24 months.

So, that is the progress that is the sort of the scale of the progress of transistor technology that the number of transistors from the same area of the chip would almost double every 2 years. And so, Gordon Moore move was essentially the co founder of Intel and he sort of put forth this observation this Moore's law is not a fundamental law of nature. it is more like an observation about the progress of semiconductor technology semi conducting industry.

And this can be essentially plotted in this way that as this graph shows that from 1970s to up to 2010 or this is true even in up to 2020 that the number of transistors per chip is linearly increasing is increasing continuously since that time or not linearly it is actually number of transistors is increasing exponentially this is the y axis is exponential is exponential at the units are exponential.

Because every 2 years almost every 2 years the transistors on the same chip are doubling so, it is like a geometric progression it is an exponential growth and the cost per unit transistor per transistor is declining. So, that is essentially the trend of the electronics industry since last around 50 years now, 50-60 years I would say.

That the number of transistors in the same area of the chip is continuously increasing that is giving us small functionality that is enabling us to do more computations more we can perform more tasks with the same device as we could do let us say few years back and the cost per transistor is reducing continuously. So, that is the roughly the Moore's law. And Moore's law captures the progress in electronics industry it also captures the scaling in the transistors ok.

The thing is that this is essentially the trend. In 1970s the channel length was around was roughly 10 micrometer this length, the length between the source and the drain was roughly 10 micrometers of size. In 1984 or in 1980s it was around 1 micrometer and this trend continued in 2014 we had transistors with channel length or with the active region of length of tens of nanometers.

And now that is an interesting thing in 2020 ,2021 or in now in 2022 the semi conductor technology node that we are working with is around 5 nanometer or even 3 nanometer. These technology nodes this modern technology nodes they do not mean that this the length of the channel is actually 3 nanometer 5 nanometer, this is not the in specially in last since last few years last few years this the technology of node is not exactly equal to the channel length.

And it essentially this number here captures the sort of the performance and the heat budget of the system. It essentially means that as we are going down with technology node we are having better performance and we are able to put more transistors in a same area, in the same area of the chip, but still what. So, the dimensions at what we are nowadays operating since last few years let us say last around 6 to 7 years is tens of nanometers typically.

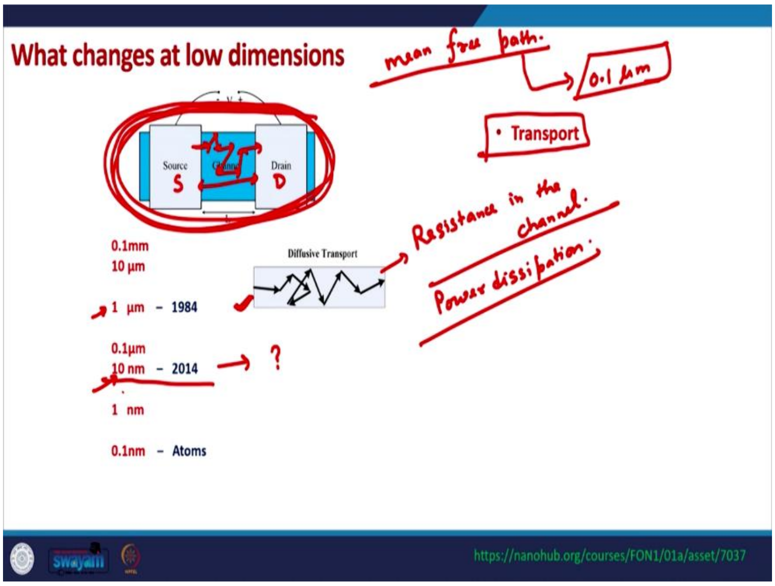
And if we look at the dimension of atoms it is few angstroms. So, the number of atoms in the channel,in modern devices if we have let us say 10 nanometer device and typically the size of a single atom is 1 angstrom 0.1 nanometer it means that we have around 100 atoms only 100 atoms in the channel between the source and the drain. As compared to a lot of atoms that used to be there in 1970s when the transistor was first invented.

So, the theories or the physics of transistor that was developed in 1980s 1970s to explain and to model the transistors that was based on the characteristics that these devices had at few micrometers let us say tens of micrometers. And nowadays since we are operating very close to the atomic dimensions though the physics is changing fundamental the transistor physics is changing fundamentally. And that is why this course is there actually, that is why we have this course.

The central focus of this course will be first to see how the physics changes as we go down to nanoscale as we go down to 10’s of nanometres of device in our systems. And second is with this in this 10’s of nanometers of device we will see what new things that appear or what the conventional or how the conventional physics needs to be changed or where that fails ,that is essentially the large part of this course that is going to be, it is going to be on these things ok.

So, I hope this evolution of semiconductor technology and the state of the art is more or less clear.

(Refer Slide Time: 27:20)



Now, this is clear so, let us for a moment let us just focus on the this channel region of the device; this two terminal sort of two terminal sub device in the MOSFET. Sub device means it is part of a bigger MOSFET and we are just looking into the two terminal device the device between the source and the drain and the channel region.

And nowadays this channel region is extremely small, it is just tens of nanometers it is or in more technical terms it is known as the mesoscopic device. What changes at these dimensions? So, that is the first question that I put forth in this course. And I will let you think about this for a moment, just take a moment and think what would change if we go let us say from a 1 micrometer transistor to let us say a 10 nanometer transistor or let us say 20 nanometer transistor ok.

So, the first thing that changes is the way electrons transport through the transistors ok. In larger systems or in the bulk devices ,generally the picture of that the picture of the electron transport is something like this in which the electron let us say starts from the left side; it strikes with and it has many atoms in the channel region because the channel is long the channel is few micrometers and it has many atoms in between.

So, electron starts from the left side is it strikes to one atom it may be changes its direction again strikes to another atom changes its direction again strikes to another atom change its direction, another again strikes change its direction and finally, it may or may not reach to the right side of the device. So, the electron starting from the source side in this two terminal part of the device will have lot of collisions in between with the crystal atoms and it may or may not reach to the right side.

But generally, if the electron reaches to the right side this is the typically the path that the electron takes in the device. And the average distance between consecutive collisions is known as the mean free path of the electron. So, that is the distance for which the electron on an average is free between two collisions ok.

So, in bulk devices the channel length is extremely large as compared to the mean free path and in that case there are lot of collisions in the channel and electron starts from the left side goes to the right side and because of these collisions, we have resistance in the channel as well as power dissipation or heat dissipation in the channel. These are the major contributors in heat dissipation in the channel.

But, generally this mean free path is around 0.1 micrometer or so, it its actually material dependent, but typically this is the range. So, if the device size is less than 0.1 micrometer, let us say device is around 10s of nanometres then what happens to the electron transport?

So, that is the first question that is that I would let you think on and we will start with this question in the next class. And after this brief introduction to this or brief summary of this course I will also let you know about the about various things that we will cover, various ideas that we will cover in this course.

So, I thank you for your attention. See you in the next class.