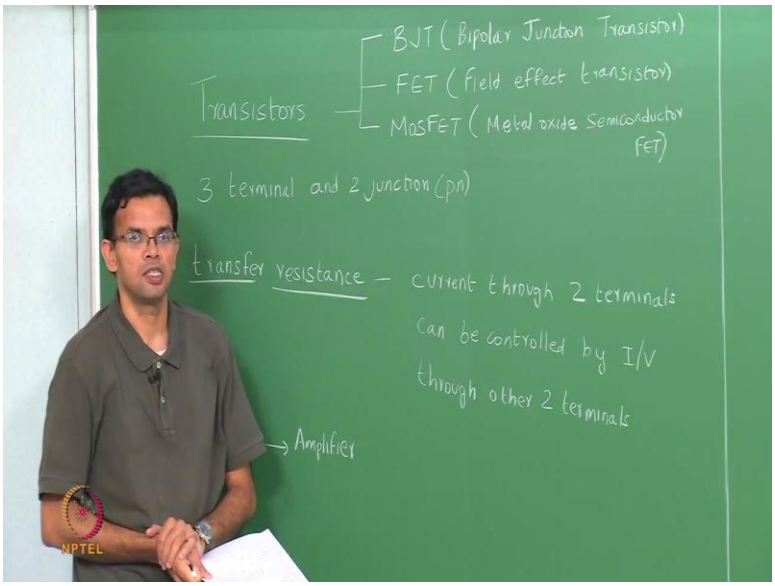
Fundamentals of electronic materials, devices and fabrication

Dr. S. Parasuraman
Department of Metallurgical and Materials Engineering
Indian Institute of Technology, Madras

Lecture – 13
Transistors

Last couple of classes we have looked at p-n junctions. A p-n junction is a 2 terminal device. We have 1 terminal on the p, 1 terminal on the n. And, it is a single junction that is the interface between p and n. Today, we are going to start looking at Transistors.

(Refer Slide Time: 00:38)

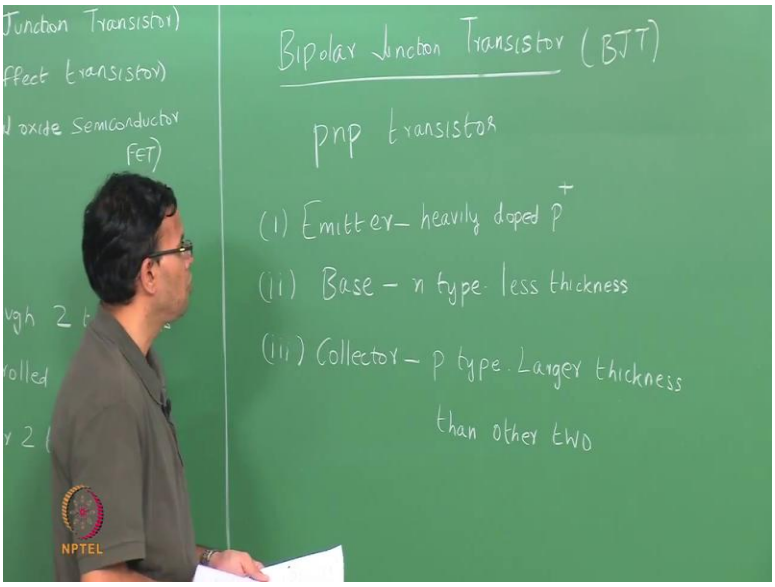


A Transistor is a 3 terminal and 2 junction device. When we mean a junction here we mean a p-n junction, so that you have 2 p-n junctions in the device. The name Transistor comes from the term transfer resistance, this means that the current through 2 terminals can be controlled by controlling the current or the voltage through other 2 terminals, controlled by either the current I or the voltage through the other 2 terminals. When we look at examples of transistors this point would be made clear. A Transistor can also act as an amplifier, in that you can take a small signal between a pair of terminals and then amplify it, so that the output signal between another pair of terminals is higher. So we will first start by looking at Bipolar Junction Transistors. So, we will first start by looking at BJT, which is your Bipolar Junction Transistor. From there will move on to

Field effect transistors and then finally a specific type of field affect transistor called a MOSFET, Metal Oxide Semiconductor Field Effect Transistor.

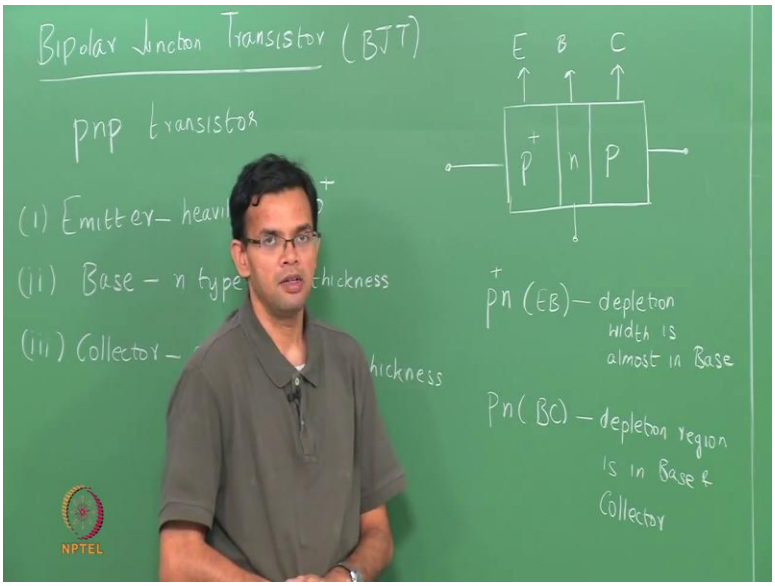
We will start with the BJT then go on to FET and then MOSFETS. So, will spend most of the time looking at MOSFETS, because MOSFETS are what are used in the current IC industry, but we will talk about the other 2 to form the basis of understanding MOSFETS. So, Let us start by looking at a Bipolar Junction Transistor.

(Refer Slide Time: 04:29)



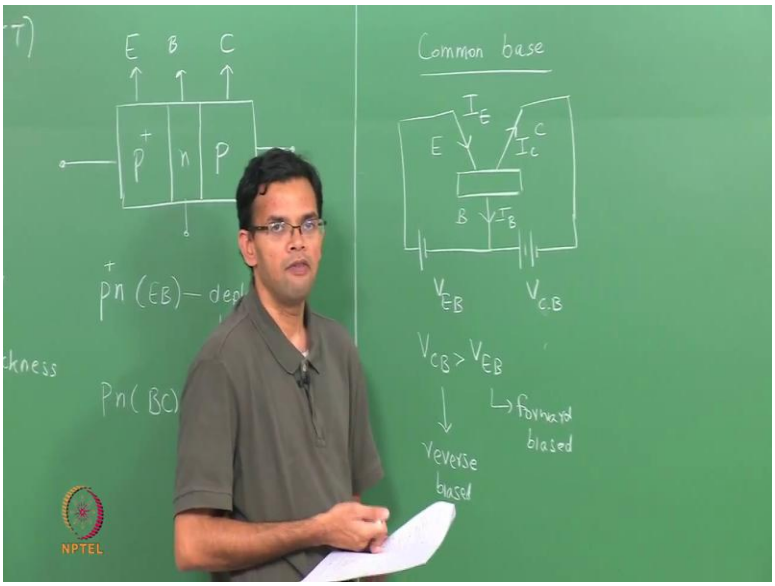
So, We will first start with an example of a pnp device. So, you have 3 regions, you have 2 of them are p-type, then you have a central region that is n-type. So, we have a pnp transistor. So, There are 3 differently doped regions in this transistor. The first region is called the Emitter, it is usually heavily doped. In this case, it is heavily doped p plus. We then have a Base; the Base here is n-type pnp. The base usually has a lesser thickness compared to the other 2 regions. And then finally, we have a Collector, in this case the collector is also p-type and it usually has a larger width or a larger thickness than the other 2 regions. If I were to show a schematic of a pnp transistor,

(Refer Slide Time: 06:51)



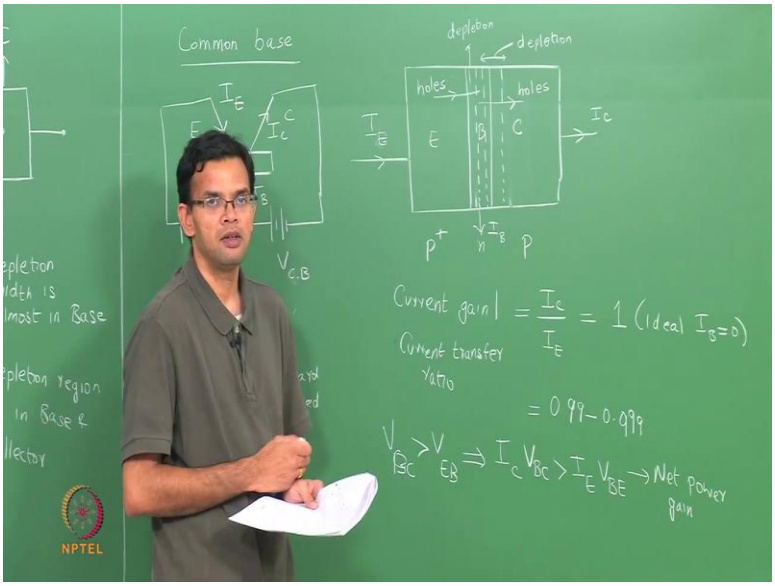
I have an emitter region that is p^+, I have a base region that is n and I have a collector region that is p. So, I will mark this E, to mean the emitter. This is your base, that is the collector. So we have 2 junctions, 1 junction between the emitter and the base, 1 junction between the base and the collector and you have 3 terminals. This makes your transistor a 3 terminal 2 junction device. So, If you look at the 2 junctions, the first junction is your p^+n which is your emitter base. The p region is heavily doped, so the depletion width is almost entirely in the base. You also have the junction between the base and the collector which is a simple p and n. In this case the depletion region is in both the base and the collector. So, is in the base and the collector. So, There are various configurations in the case of a pnp transistor, how we connect these different terminals. First configuration we are going to look at is called a Common base. This case base is common between both the emitter and the collector.

(Refer Slide Time: 09:04)



So, Let me draw the electronic the configuration for the common base. So this is the base, you have an emitter and you have a collector. So, The emitter based junction is forward biased and the collector based junction is reverse biased, this is c. In this particular example, $V_{CB} > V_{EB}$, so let me just write down this is forward biased and this is reverse biased. We can also define 3 currents, 1 is an emitter current I_E, then you have a base current and then finally you have a collector current. How does a pnp transistor work? Let me again draw a schematic of the transistor, in this I will also mark the depletion regions.

(Refer Slide Time: 11:12)



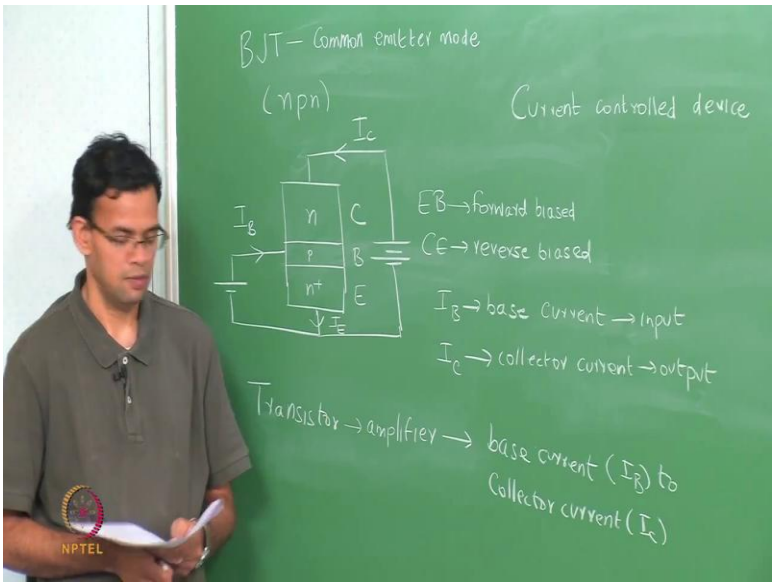
We just redraw this; I have 3 regions the emitter, the base and the collector. So the emitter is p^+, the base is n and the collector is p. So, The emitter based junction you have a $p^+$and then n. The depletion width is almost entirely on the base side, on the other hand the base and the collector the depletion width is in both regions. So, you have 2 junctions which is why you have 2 depletion regions. So, We have the emitter current, we have the base current and you have the collector current. The emitter based junction is forward biased, so we have holes that are injected from the emitter into the base. So, These holes then go through the base where they are minority carriers because the base is n-type, but the width of the base region is very small. Some of these holes can get recombined. So, there form the base current, but a majority of the holes go through the base and then they go to the collector which is reverse biased. You have holes it go through the base and they form your collector current.

So, we have a base emitter current that is because of the injection of holes, some of these holes recombine in the base region they form the base current and the remaining holes that go from the base to the collector constitute the collector current. In the case of a Bipolar Junction Transistor we can define a current gain another name for it is also the current transfer ratio; this is nothing but the ratio of the collector current to the emitter current. Now, in an ideal case if no holes are lost in the base due to recombination, this current gain should be equal to 1. This is if it is ideal and there is no base current, but usually some of the holes will always be lost in the base due to recombination. This current gained typically is around 0.99, so be up to 0.999. So, The current gain depends upon the thickness of the base region, so typically we want a very thin base so not many holes are lost due to recombination.

Instead of a pnp, if you had an npn transistor the argument is entirely the same, except that here you have injection of electrons instead of holes. The transistor action here arises because there is a net power gained, because we said that V_{BC} which is your base collector voltage is higher than V_{EB} which is your emitter based voltage. If you look at the net power, power is nothing but the voltage into the current. So you have I_C, see that is greater than the emitter V_{BE} so, you have a net power gain, new transistor. So, In this particular mode we operated the BJT in a common base configuration.

Let us look at one more configuration, we look at a configuration where we have the emitter being set as common.

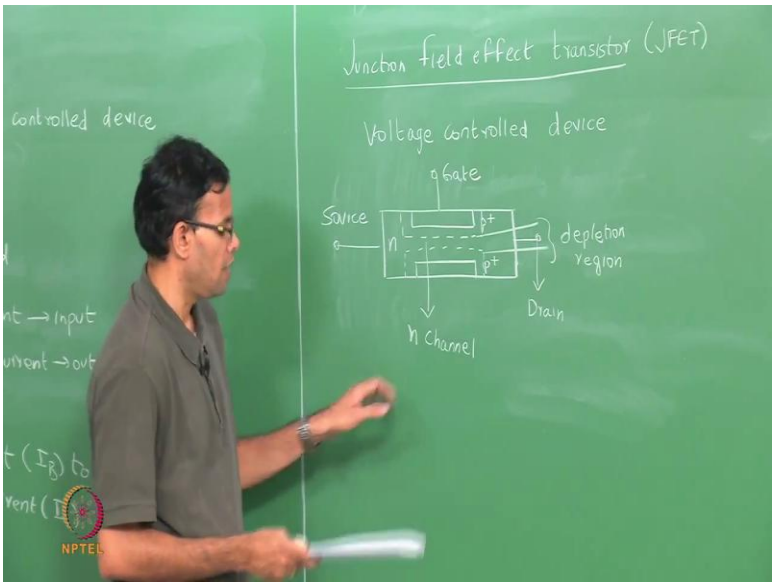
(Refer Slide Time: 17:04)



So, Now, you have your Bipolar Junction Transistor, but it is in common emitter volt. So if you redraw this and just for example, I will choose an npn transistor instead of a pnp. So, In this particular case, I have an emitter that is n^+ that is my emitter, I have a thin base region that is p-type and then have a collector that is n-type. So, in this particular case I have the emitter and the base to be forward biased and I have the collector and the emitter because it is a common emitter mode to be reverse biased. So, In this particular case I_B which is the base current is your input and I C which is the collector current is the output.

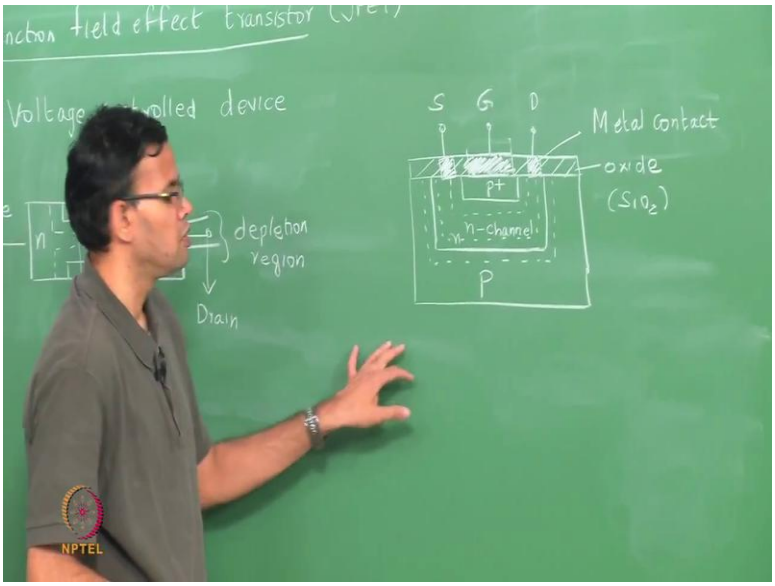
In a common emitter configuration, it is easy to see how your transistor acts as an amplifier. The base current is typically very small, while the collector current is much larger. It is almost equal to the emitter current. So, you are taking a very small current and then amplifying it into the collector current so that your transistor works as an amplifier. So, It takes the small base current and gives out a larger collector current. So, A bipolar junction device is essentially a current controlled device. So, You can say that the current from the emitter to the collector, so I_E and I_C are controlled by your base current, but overall it is a current controlled device. The next thing we are going to look at is your Field effect transistor; we are going to look at a Junction Field effect transistor.

(Refer Slide Time: 21:44)



So, we are going to look at a Junction Field effect transistor called JFET. Both the JFET and the MOSFET which is what we look after this are voltage control devices. So, Let me draw a schematic of a JFET. Remember, once again we need 2 p-n junctions we also need 2 terminals. So, in this case I start off with an n-type material, I have 2 heavily doped p-type materials or p-type layers till in this n. So, Both of these are heavily doped p, so I am going to call them p^+. So, because they are heavily doped and you have a p and n you have a p-n junction and the depletion width is almost entirely on the n side. So, If were to draw a depletion region, so the dotted line represent the depletion region and between the depletion region you have an n channel. So, We can again define 3 terminals, you have one that is called a Source, you have one that is connected to the p^+ it is called your Gate and then you have a drain, we just mark it here it is drain. So, you have a source and you have a Drain and then you have a gate. A more practical way the device will look I will just draw that.

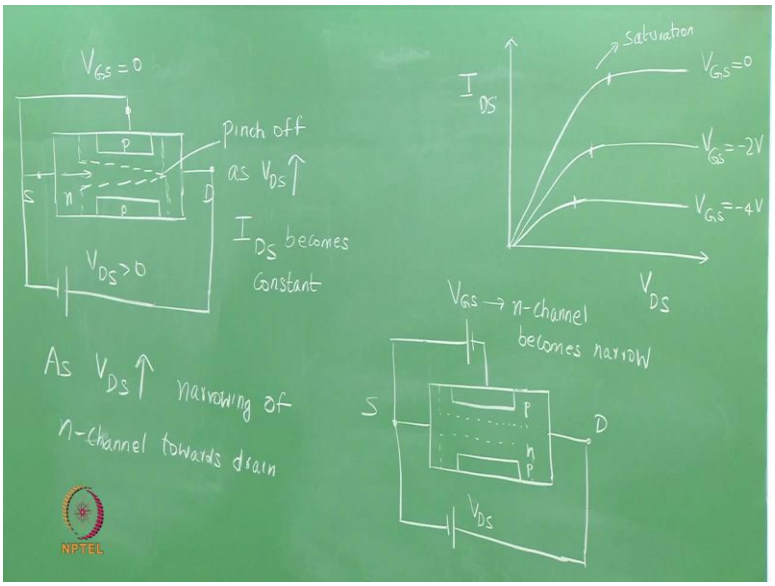
(Refer Slide Time: 24:37)



So, I have a p-region, within that p-region I define an n region and then again I have another region which is heavily p^+. So, we can define 3 terminals usually there is an oxide layer on top, So, if your material is silicon which is almost always the case where oxide is SiO_2. We can form 3 terminals, one is your Source, then you have a Gate, then you have your Drain. So, The shaded regions essentially mean metal so that you have a good contact with your semiconductor, so I will just say a metal contactor. Once again in this device you have 2 depletion regions. So you have 2 junctions, one between the p^+ and the n one between the n and p, so that you have 2 depletion regions and you have an n channel. So, This schematic represents more of how the actual device will look; this particular diagram is something that I am using in order to explain the functioning of a JFET. So, we will use this to explain how the current and voltage behavior works, but please keep in mind a more practical device will have a difference schematic. We will consider the behavior of a JFET. Once again you have a current between a source and the drain and this current will depend upon the voltage that is applied at the gate, so that is your transistor action where the current between 2 terminals depends upon the voltage or the current between the other terminals.

So, let us first look at the current between the source and drain when your gate is short circuited.

(Refer Slide Time: 28:11)



So, Let me redraw the schematic of the device. So, you have a source and you have a drain, this my drain, that is my source, I have my 2 p regions this is n. I have my 2 p regions so that I have my depletion regions and then I have my n channel. So, We are going to apply a bias between the source and the drain, so I will bias the drain positive with respect to the source let me call this $V_{DS} > 0$, which means the drain is bias positive with respect to the source so that electrons can move from the source to the drain. So, Electrons can move and this constitutes your current. We will also set the gate at a 0 potential, so that V_{GS} is 0, as we start to increase this value which is your drain and source voltage we are going to have a current it keep on increasing the value, there is going to be more current. But if you look at the 2 junctions, as we increase the value of the drain source voltage there is going to be narrowing of the n channel, this is because this drain is reverse biased with respect to the gate so that as V_{DS} increases there will be narrowing of the n channel towards the drain.

I will just show that schematically on the same figure so that if you look at the channel, the channel is wider near the source and starts to narrow near the drain. If you increase the value of V_{DS} further there is going to be further narrowing until both the depletion regions meet, in which case we will have a pinch of region. We can show that is schematically on the same plot so that your n channel as essentially pinched off here at high values of V_{DS}. When pinch off occurs the current through this device essentially becomes the constant because it depends upon the resistance of this n channel. This is

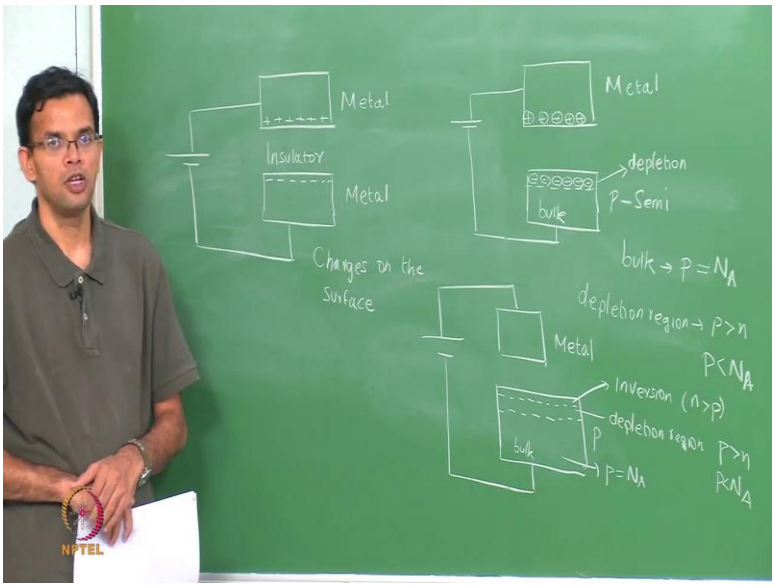
again assuming we have a pinch of region that is small in width. So, When once pinch off occurs the current I between the drain and source, so I call it I_{DS} becomes a constant. We can plot an I-V characteristics for this a JFET and if we do that I have current, so I_{DS} which is the current between the drain and source versus V_{DS}, which is again the voltage between the drain and source. Right now, I am shorting the gate so that the gate and source voltage is 0. As you increase the voltage initially your current starts to rise because you have electrons going from the source to the drain, but as the current starts to rise and as the voltage increases the width of the channel will also decrease and ultimately, you will have pinch off and when that occurs the current is a constant. So, your I_{DS} initially increases until pinch off occurs and then your current is a constant.

So, this value where your pinch off essentially occurs is your saturation and this plot is when your gate and source are shorted, so that there is no bias between the gate and the source. So, What will happen if I now have a bias between the gate and the source? So, let me redraw this and introduce a bias between the gate and the source. So, Again, I have my n-region, I have my 2 p-regions, that is my drain, that is the source and the drain that is positively biased with respect to the source. And now I am going to bias the gate negative with respect to the source V_{GS}. So, When we do this now again you have a depletion region, but if you look at the gate and the source, the gate is reverse biased with respect to the source so that the depletion region is higher or in other words the channel is narrower. Once you have a bias the n channel becomes narrow and if we keep on increasing this value the channel will become narrower and narrower. What this means is that if you plot your I-V characteristics you are going to find that pinch off becomes easier higher this voltage between the gate and the source.

So, We were to plot this, so this is for $V_{GS} = 0$. If you have another value V_{GS} is -2V. Once again your current for saturation is lower because you have a narrower channel and also the saturation occurs earlier if you increase the voltage even more, so V_{GS} is -4V the current will again go down and saturation also occurs earlier. In this particular case, which is your JFET you have a situation where the current between the source and the drain is depends upon the width of the n channel is effected by the voltage between the gate and the source V_{GS}, so higher that voltage smaller is the width of the channel and then lower is the current. So, This is again an example for transistor action, where the current between 2 terminals is affected by the voltage between the other terminals. This

is an example of a JFET, so where you already have a channel that is created within the material. The next thing we are going to look at is a MOSFET, so we have a metal oxide semiconductor which creates your n channel.

(Refer Slide Time: 38:15)



So, the Metal Oxide Semiconductor Field Effect Transistors are what there are commonly used in the current micro fabrication industry. So, before we understand MOSFETS, let us just look at metal oxide and semiconductor junction and see how we form the channel in that particular case? Right now, we will only look at this part. Once we understood that we will put it together in order to look at the MOSFET. So, consider a parallel plate capacitor formed between 2 metals, so you have metal 1 and you have another metal plate and then you apply a potential. So, Between the 2 metals you have an insulator, so this insulator could be air or vacuum or it could be an oxide layer or some other layer which acts as an insulator. So, you have 2 metal plates connected to a potential. This case you have positive charge on one plate and the negative charge on the other. So, this acts as a capacitor and since these are metals the charges will reside on the surface.

So, Let me first take this device or let me take this arrangement and replace one of the metals with a p-type semiconductor. And I have a metal, but instead of the other metal I have a p-type semiconductor. So, Once again there is an insulator layer between these 2, connect the metal to positive and the p-type to negative so, we have a positive charge on

the surface of the metal, but the charge density in the case of a semiconductor is lower. So, We have seen this example earlier in the case of schottky junctions. So you have a negative charge on the semiconductor, but it is not only at the surface, but it also extends a small distance into the bulk. In the case of a semiconductor we have a depletion region and then we also have the bulk. So, within the bulk of the semiconductor it behaves as your regular n-type, so $P = N_A$, but in the depletion region you have less number of holes less than N_A, the material is still p-type. You still have p greater than n, but the value of P is less than N_A and this is because you have biased the semiconductor negative so that you are pulling the holes away from the semiconductor or you are pushing electrons.

Now, if you keep increasing the bias, we are going to get more negative charge in the semiconductor and you are going to find that at one particular point, the number of electrons will be more than the number of holes, so that you create a region where you have n-type conductivity as opposed to p-type, this thing is called Inversion. So, let me just draw there again. Once again I have a Metal; I am going to draw my semiconductor slightly bigger so I can show the different regions that is my p-type. So, I have the bulk of the semiconductor, this is the bulk. I have a depletion region where the semiconductor is still p-type, but the concentration of the holes is less than N_A, so I have a depletion region and then finally, I have a region which is closer to the surface where I have more electrons than holes so that I have inversion.

So, In the bulk, I have $P = N_A$ that is in the bulk. In the depletion region, I have $p > n$ so that it is still a p-type, but it is less than N_A and finally, I have an inversion region where $n > p$. So, In the case of a p-type semiconductor, I have formed an n channel by applying a negative bias to the semiconductor. So, this is the principle of your Metal oxide semiconductor which we will use in the Transistor. The difference between this and JFET, is that in the case of a JFET you already had an n channel that was present and by applying a voltage we shrunk or decrease the width of the channel and control the current. In the case of a MOSFET, the channel is formed by applying an external potential so that the channel is not immediately there and by increasing the potential you can increase the width of the channel and control the current.

So, In a next class, we are going to look at the working of a MOSFET. So, once again we can draw this current versus voltage characteristics. We will also look into some more detail on the formation of the depletion and the inversion region and also calculates their

widths.

So, next class we will look at MOSFETS in detail.