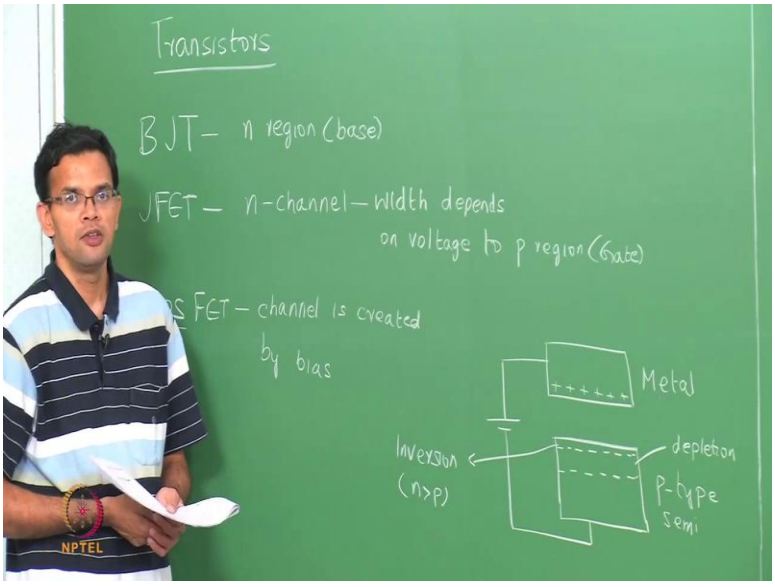
Fundamentals of electronic materials, devices and fabrication

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

Lecture - 14
MOSFETS

Let us start with a brief review of last class. Last class, we started looking at Transistors. Transistors are 3 terminal, 2 junction devices.

(Refer Slide Time: 00:36)

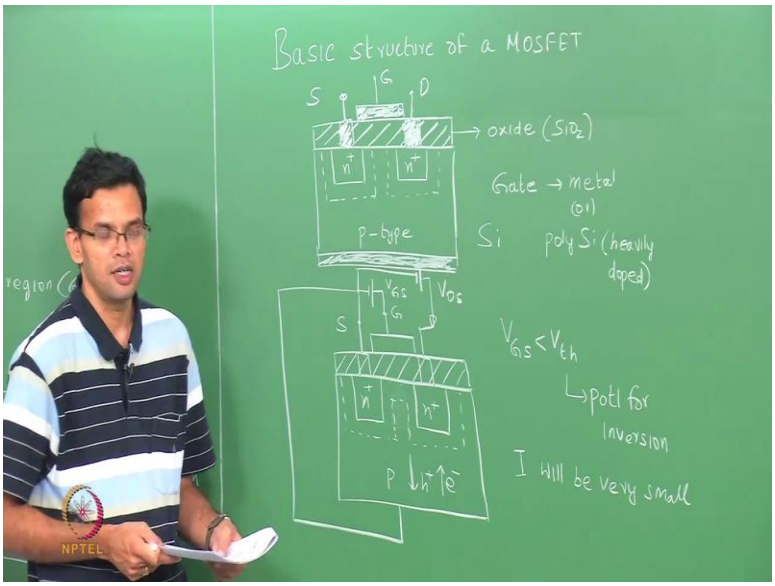


We first started by looking at bipolar junction transistors, BJT. So, in the case of PNP bipolar junction transistor, we had n-region which was your base through which the minority carriers which are your holes go as they go from the emitter to the collector. After BJT, we looked at a junction field effect transistor JFET. In the case of a JFET, we already had an n-channel and the width of the channel was controlled by applying the voltage to the p region. So, this essentially forms the gate while we had a source and the drain and the current went from the source to the drain through the n-channel. So, a BJT is a current control device and then a JFET is a voltage control device. Towards the end of last class we started looking at a Metal Oxide Semi Conductor Field Effect Transistor. So, in this particular case the channel is not there initially when the device is in

equilibrium, but the channel is created by applying a bias. So, we saw the example of a metal oxide semiconductor junction.

In this case we have a metal, we have p-type semiconductor. So, we apply a bias between these 2, the metal is connected to positive and the p-type semiconductor is connected to negative. In this particular case, you have a positive charge on the metal side, now there is a negative charge that extends through the p-type material not only at the surface, but also some with within the junction, within the bulk. So, we have 2 regions, we define one region that we call a depletion region. In the depletion region, the material is still a p-type, but the concentration of holes is less than N_A, which is your acceptor concentration and we apply a further bias. We defined a region which is your inversion region in which case you have $n > p$. So, that you have an n-channel that is created within your p-type material. So, today we will take up this further and look at MOSFET. The first thing I am going to do is to draw the structure of a MOSFET device, so that we can look at the I-V characteristics.

(Refer Slide Time: 04:35)



So, let us look at the basic structure of a MOSFET. So, I have my bulk semiconductor material, we still keep it to be p-type. There is a metallization layer, so we can form electrical contacts and then there is an oxide layer on the surface. We have 2 regions within the MOSFET which are heavily n-type dope n plus, n plus. So, these are called the source and drain and we again make electrical connections to the source and the

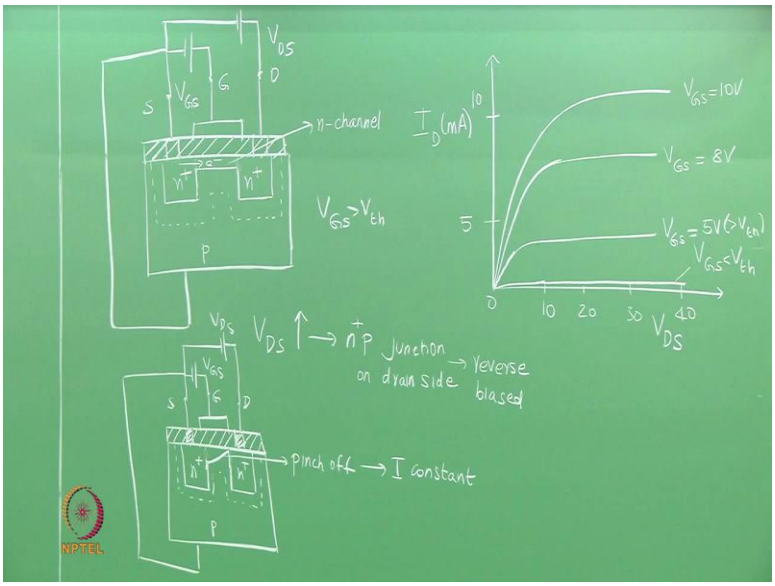
drain. So, let me call them S and D, so that they are your source and the drain. We also have a gate region through which you can apply a potential, so the 2 n^+ regions form a junction with a p-type. So, you have 2 p n junctions since, n^+ is heavily doped most of the depletion region will be on the p side. So, if these materials were silicon. So, you have p-type silicon and n-type silicon in the oxide will essentially be SiO_2, the gate material can either be a metal or it can be poly silicon, which is heavily doped, so, that it is conductive.

So, we are going to look at how n-channel is formed in this MOSFET and also the I-V characteristics of this transistor. So, let me start by looking at the working of the MOSFET. I am going to redraw this figure, but then I am also going to make electrical connections between the various terminals. So, you have a source gate and a drain terminal. So, let me redraw, we are going to keep the MOSFET as the same. So, we have a bulk P, I have 2 n^+ regions. So, I have my source and the drain and I have a gate. So, let me just mark them as G and D. So, how do we bias this MOSFET? So, let me first bias the source and the gate.

So, in this particular case the gate is connected to a positive potential and the source is connected to a negative potential, we call this V_{GS}, and then I will also bias the source and the drain. So, the drain is connected to a positive potential with respect to the source. So, we are going to apply a potential to the gate in such a way that holes will flow away from the gate towards the bulk of the p-type and electrons will flow towards the oxide layer, so we have 2 depletion regions here. We are going to bias the material in such a way that the holes will flow down an electron will flow up.

But, the potential is lower than the potential that is required for inversion. So, $V_{GS} < V_{th}$, we saw earlier that V_{th} is the potential for inversion. So, in this particular case, we are going to find that you will have a depletion region here. So, let me just erase this and join the 2 depletion regions together. So, we now have a depletion region, but the material is still p-type, so that electrons are still the minority carriers. So, current I will be very small and it will be equal to your minority carriers flowing in a p-type material. So, it will be very similar to your reverse saturation current. So, As we keep increasing the gate and source voltage, the depletion region is going to increase and ultimately, when the voltage is above V_{th}, we are going to form a channel between the source and the drain. So, let me draw that next.

(Refer Slide Time: 11:56)



So, once again I have my MOSFET, my 2 n^+ regions, I have a bulk p-region and I have my gate. So, I have my source, my gate and my drain. This is V_{GS}, this is V_{DS}. Now, I am in a situation where I have the gate source voltage greater than V threshold, so that I have an inversion region. Once again, I have a depletion region between the 2 p n junctions, but now I am at a higher potential, so that I have an n-channel that forms between the 2 n^+ regions. So, This one is your n-channel and the n-channel forms because a gate source voltage is our V_{th}.

So, Now, I can have electrons flowing from the source to the drain and the electron flow is because we have a positive potential that is applied to the drain compared to the source. In the case of a MOSFET, you will not have current conduction when the voltage is below V_{th}, but above V_{th} we now have an n-channel and we have conduction. Thus, the gate source voltage essentially acts as a controlling parameter in order to control the current in the MOSFET and this essentially is the transistor action where the voltage between 2 terminals determines the current or the voltage between other 2 terminals. You can also say that these gates were source voltage acts as a switch; so, it helps to turn off or turn on your transistor. Now, what happens as V_{DS} increases? When a V_{DS} increase, V_{DS} is the voltage between the drain and the source, so as V_{DS} increases if you look at the p-n junction on the drain side, if you look at this p n junction, n^+ is connected to positive p is connected to negative. So, that this is reverse biased. So, This essentially

means the channel starts to get narrower as V_{DS} increases and ultimately, above a certain voltage the channel just gets pinched off.

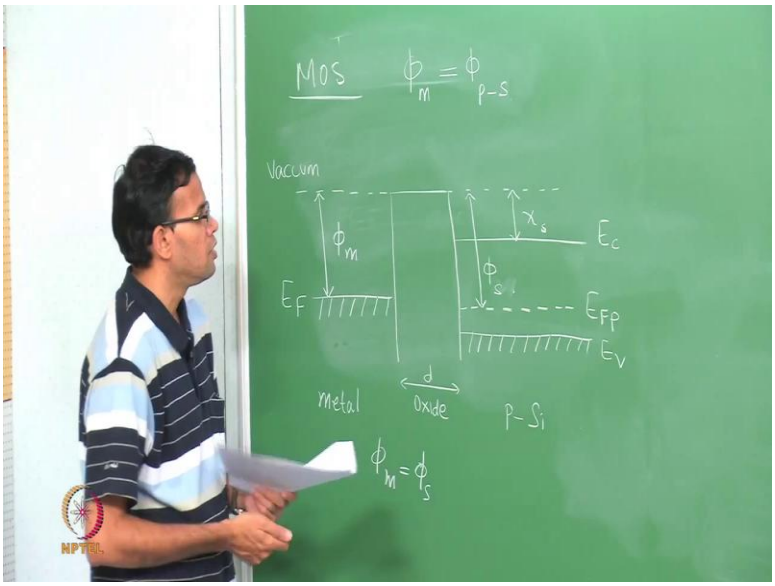
If you were to draw that, again I have my 2 n^+ regions. I have my bulk p it is my oxide layer and my gate source gate and drain. I have my gate source voltage above the threshold, so that I have n-channel, but when I start increasing V_{DS} the channel starts to narrow towards the drain region, so that ultimately you have pinch off to show this ensure the n-channel that is narrowing as we approach the drain region still have a depletion region that is surrounding it. Once pinch off occurs the current is essentially a constant because the current is determined by the resistance of the n-channel and as long as the pinch off width is small, it will not affect the total current.

So, in the case of a MOSFET, if you put all this information together we can draw current verses voltage characteristics. We will plot the current through the channel as a function of the drain source voltage for different values of V_{GS}. So, I_D is the drain current. This is the current that is flowing from the source to the drain as a function of V_{DS}, which is the voltage between the drain and the source. If the gate source voltage is below the threshold value the current is very small, so this is the line that is very close to the access. This is as long as $V_{GS} < V_{th}$. So, once V_{GS} goes above V_{th}, we are going start to see an increase in current. Ultimately, there comes a point when pinch off occurs and the current becomes the constant. So, this is for VGS equal to 5 volts which is greater than V_{th}.

Now, if you keep increasing the value of V_{GS} then you are going to have a wider channel, which means there will be more current, but eventually there will be pinch off. So, this is V_{GS} 8 volts. So, this is the I-V characteristics in the case of a MOSFET. The difference between this and the JFET where we saw earlier was, there we would shrink the n-channel by applying a potential to the gate in this case you increase the n-channel by applying a potential to the gate.

So, Let us do some calculations in order to figure out the width of this depletion region and the inversion region and how that is related to the doping level in your bulk semiconductor. So, we will again look at the metal oxide semiconductor junction and look more closely at the band picture especially, band bending.

(Refer Slide Time: 20:57)



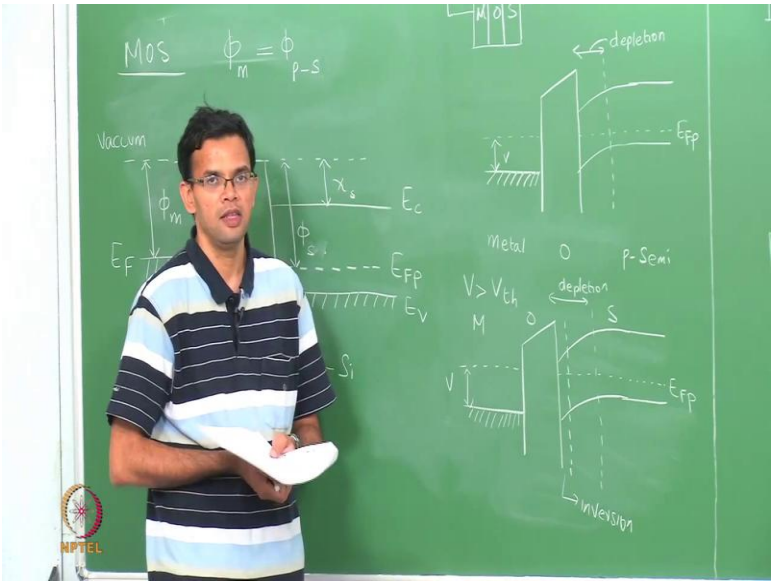
So, let me again go back to my picture of the metal oxide semiconductor and I want to draw a band diagram for this. For simplicity, I am going to say that the work function of the metal is the same as the work function of the p-type semiconductor. Instead of silicon, I will just say S. So, that it is a more generalized argument, but we will start with the work functions being the same, the advantage is that you say both work functions are the same in the absence of any external potential the fermi levels will just line up.

So, On the left, I have my metal which has a fermi level of E_F; the dotted line represents my vacuum level, so that φ_m is the work function of the metal. I then have an oxide layer of some thickness d towards the end, we will see what role this oxide or in more general in insulator place. So, we have an oxide layer and then I have a p-type semiconductor E_{FP}, so this a p-type material. So, the fermi level is close to the valence band, it is E_V that is E_C. So, For the semiconductor this will be my work function call it φ_{PS} and then this will be my electron affinity. In this diagram let me just drop the subscript P. So, I will just call this φ_S. So, just saying φ_m is equal to φ_S, but we know that we are starting with a p-type material.

So, this metal oxide semiconductor it can be biased and there are 2 ways of biasing it. In one particular way if you connect the metal to a negative and a semiconductor to a positive then your injecting holes into the semiconductor in these holes will accumulate

at the junction, but that is not what we want we want to bias in such a way that we are going to have holes moving away from the junction. So, we have an n-channel.

(Refer Slide Time: 24:12)

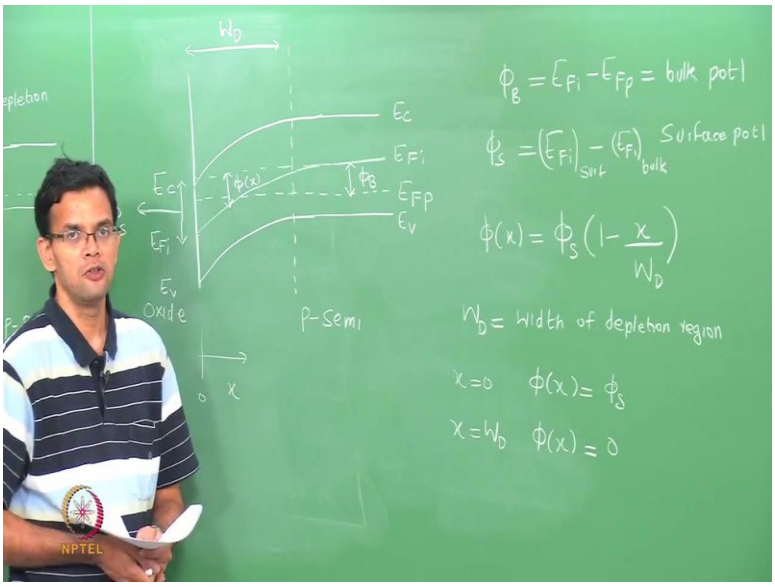


So, let us look at the metal oxide semiconductor under bias. So, I am just drawing a schematic, we are going to bias this in such a way that the metal is connected to positive and the semiconductor is connected to negative some voltage, V. So, if you were to draw the band diagram in this case we are applying an external potential. So, that the Fermi levels no longer line up.

This is my metal oxide and p-type semiconductor, this is the work function of the semiconductor this is the externally applied potential. So, the Fermi level shift by the applied potential far away from the junction you still have a p-type, but as you go closer towards the junction since you are moving electrons away your material becomes more and more intrinsic in the depletion region and ultimately, in the inversion region you have an n-type material. So, you have band bending as you go towards the junction. In this particular case, $V < V_{th}$ so that you can define a depletion region if you have a situation, where $V > V_{th}$, your bands have bent so much that towards or near the surface you have inversion. So, if you were to draw that again. Let me say metal oxide semiconductor shell V this is E_{PF}. So, far away from the junction it still behaves like a p-type.

Now, you have a depletion region, but within the depletion region there is another smaller region where N > P, so, that you have inversion. So, This is what happens to the band diagram. In the case of metal oxide semiconductor, when we try to form an n-channel. So, let us look more closely at the semiconductor side. We can put in some of these potentials and use it to calculate the width of the depletion and the inversion regions. So, we will look only at the semiconductor side of the junction.

(Refer Slide Time: 28:10)

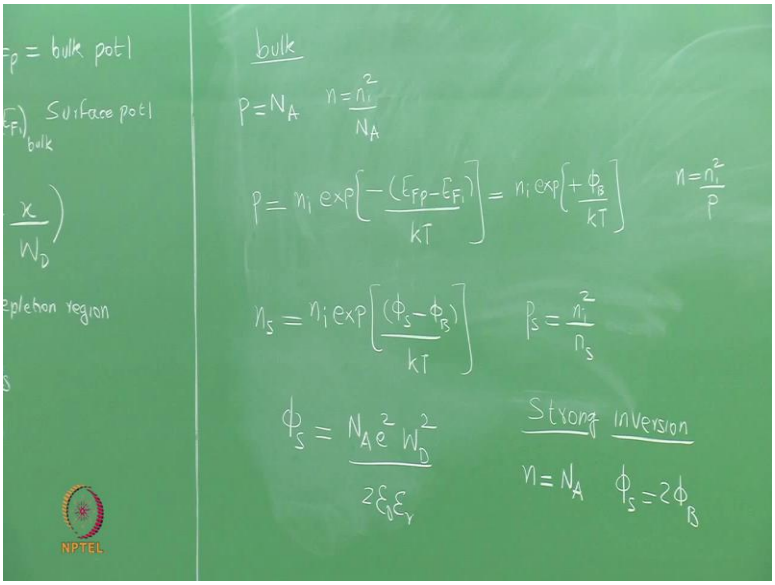


This is my oxide layer, this is my p-type semiconductor, let me mark E_{FP}. So, E_{FP} since it is p-type will be closer to the valence band. This is E_c and E_v in this diagram. Let me also mark the intrinsic Fermi level E_{Fi}, now the intrinsic fermi level as we have seen earlier is very close to the center of the band gap. It is not exactly at the center because the effective masses of the electrons and holes are different, but it is very close to the center. So, we also have drawn E_{Fi}, the gap between E_{Fi} and E_{FP}, I am going to call φ_B which is the bulk potential. Let me call this φ_B, φ_B is nothing but E_{Fi} - E_{FP} which is the bulk potential. Now, we are applied a potential such that you have an inversion layer at the surface and you also have a depletion layer.

So, I will also draw E_{Fi}. So, this is E_c E_v and E_{Fi}. So, In this case, we can also define a surface potential φ_S. So, I will call surface potential which is the difference between the intrinsic fermi level in the bulk and the intrinsic fermi level in the surface. So, φ_S is E_{Fi} at the surface -E_{Fi} in the bulk. If I were to draw this, let me just draw a dotted line, so, that

this is φ_S. So, we have a bulk potential we have a surface potential we can also define the potential at any distance x from the surface. In this particular case the surface is taken is 0 and x is the distance as we go into the material. So, φ as the function of x is nothing but $\frac{\varphi_S - x}{W_D}$, where W_D is the width of the depletion region. So, x is defined from the surface, so that when x is 0, $\varphi(x)$ is just the surface potential and when $x = W_D$, $x = 0$. So, W_D represents the start of the band bending because it is a measure of how much your intrinsic fermi level as shifted because of band bending. So, let us now relate these potentials to the concentration of electrons and holes in this semiconductor.

(Refer Slide Time: 33:22)



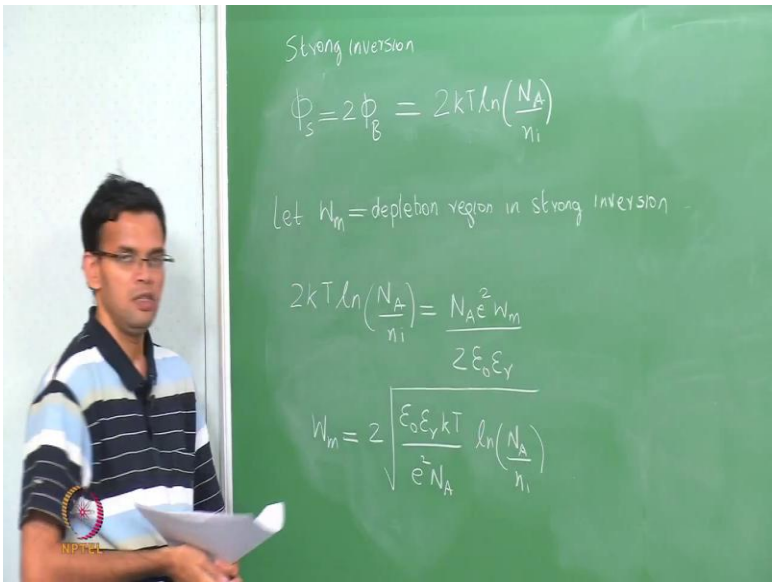
So, within the bulk of semiconductor $P = N_A$, which is your acceptor concentration N will just be $\frac{n_i^2}{N_A}$. We can also relate the position of P mean the position of the fermi level to P and if you were to that P is nothing, but n_i exponential is $n_i exp \frac{(-E_{Fp} - E_{Fi})}{kT}$. So, this we have seen before E_{FP} - E_{Fi} it is nothing but the bulk potential. So, this is $n_i exp \frac{(-\varphi_B)}{kT}$ or actually this should be $+\varphi_B$ because it is minus E_{Fp} - E_{Fi}, which is $-\varphi_B$. So, it is $+\varphi_B/k$; same way, at the surface we can also define a concentration of electrons and holes. So, if you were to do that N_S which is the concentration of electrons of the surface it is again related to the position of the fermi level. So, n_i exponential and this we can simplify to include the surface potential and the bulk potential. So, it is $exp \frac{(\varphi_S - \varphi_B)}{kT}$, so

that it is a difference between the surface potential and the bulk potential or in other words how much the fermi level has shifted with respect to the intrinsic Fermi level.

PS We can just calculate from n_i^2 and the same way here n_i can calculate from just P, this value of φ_S depends upon the concentration of acceptors within the bulk p-type material. So, we can relate φ_S to the bulk concentration and also the width of depletion region. If we do that, φ_S is just related to $\frac{N_{Ae}^2 W_D}{2\varepsilon_o \varepsilon_r}$. So, this we can get by assuming a certain distribution of electric charge within the depletion region and thus calculating the electric field and linking that to the potential. So, we have done a calculation for this when you look at the p-n junction. So, the argument here is also similar. So, we can relate φ_S, which is your surface potential to N_A, which is the concentration of acceptors within your p-type semiconductor and also W_D, which is the width of the depletion region. In the case of a MOSFET we have inversion, so that we have an n-channel that is created. We define a condition called strong inversion. In the case of strong inversion we create an n-channel, where the concentration of electrons is equal to N_A, so that we create a channel that is as strongly n-type as the bulk material is p-type.

When we have strong inversion an $N = N_A$ φ_S will be just twice the bulk potential. So, this we can understand because phi B is a location of the Fermi level with respect to the intrinsic Fermi level. In the case of p-type this Fermi level is located below E_{Fi} and now, we can make it n-type the Fermi level goes above E_{Fi} and when $N = N_A$, the distance of the Fermi level from the intrinsic will be the same. So, it is starts φ_B below E_{Fi} and when becomes n-type, it becomes φ_B above E_{Fi}, so that the total potential is just 2 φ_B. This is the particular case of inversion and it is called strong inversion. So, we will equate this in order to calculate the width at strong inversion.

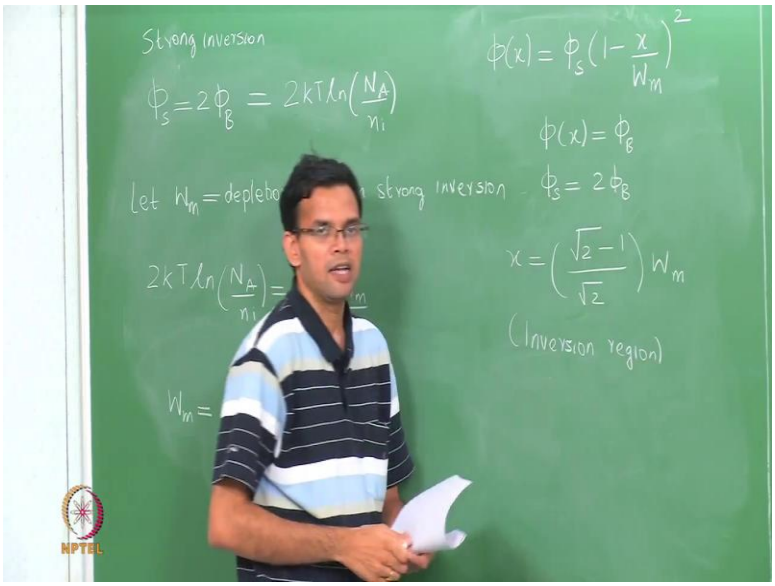
(Refer Slide Time: 39:05)



So, We have a case of strong inversion were φ_S is 2 φ_B, this in turn is related to the concentration of the acceptors lawn N_A/n_i. We also saw an expression relating φ to the width of the depletion region.

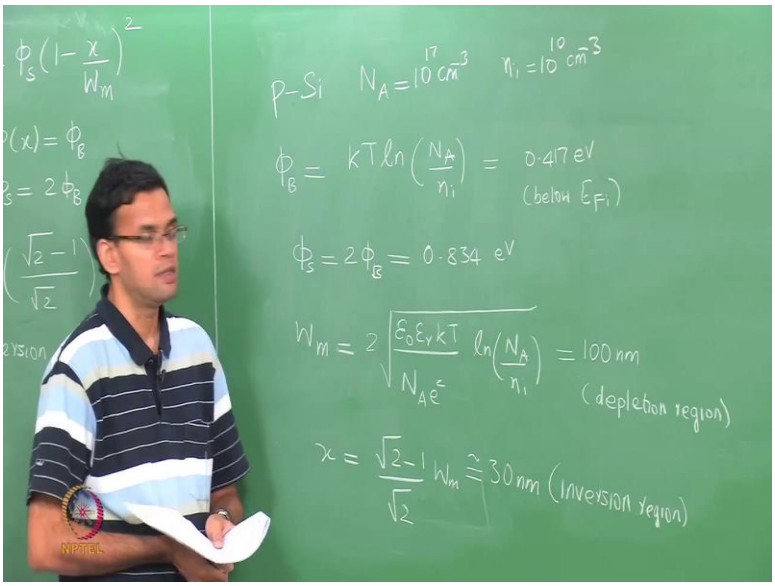
So, let W m be the width of the depletion region in strong inversion. So, in this particular case, $2kT\ln\frac{N_A}{n_i}$ is related to $\frac{N_A^2 e W_M}{2\varepsilon_o\varepsilon_r}$., where ε_{or} is the relative permittivity of your semiconductor. We can rearrange this to get the width of a depletion region that is just $2\sqrt{\frac{\varepsilon_o\varepsilon_r kT}{e^2 N_A}\ln\frac{N_A}{n_i}}$. So, this equation gives the width of the depletion region. The total width when we have strong inversion and this is related to the concentration of the acceptors within the material. We also wanted to find the width of the inversion region. The depletion region is both the inversion and the region, where we have P > N, but less than N_A.

(Refer Slide Time: 41:32)



In order do that, we start with this expression φ(x) - x and now, I am looking at a condition of strong inversion, so I will put W_m^2, we will define the inversion region as where φ(x) = φ_B. Remember, the surface potential is 2 φ_B. So, we can put this values here and simplify to get $x = \frac{\sqrt{2}-1}{\sqrt{2}} W_m$. So, We can calculate both the total width of the depletion region which is given by this formula and the width of the inversion region. Let us actually substitute some values in order to get a sense of these numbers.

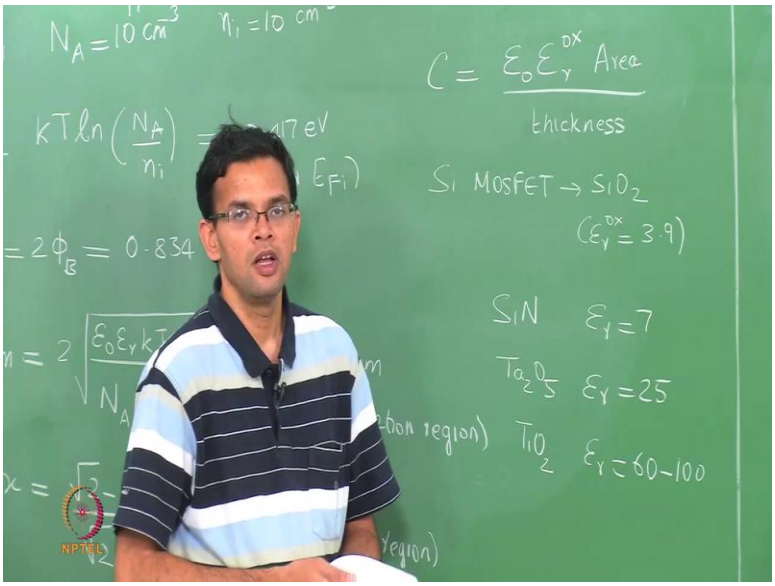
(Refer Slide Time: 42:51)



So, I have a p-type semiconductor. I am going to take silicon with N_A to be 10^{17} cm^{-3}. We can first calculate the bulk potential which is the difference between the Fermi level in the intrinsic material and the Fermi level in your p-type that is nothing, but $kT \ln \frac{N_A}{n_i}$. If you have silicon we know that n_i is 10^{10}, so that this gives you a value of 0.417 electron volts and this is below E_{Fi}. So, we now want to create n-channel in this p-type material and we want to make it a strong inversion. So, that the material is as much n-type as the bulk is p-type. In order to do that we need to have a surface potential φ_S and is 2 φ_B, this is nothing but 0.834 electron volts. We also wrote down an expression for the total width of the depletion region, so we can substitute the values here and that gives φ_N to be 100 nanometers. So, this represents the total width of the depletion region.

If you want to calculate the width of the inversion region, which is x and this one is approximately 30 nanometers, So, this one which is 100 nanometers, which is W_m; it is the total width of the depletion region. Within this depletion region you have an n-channel, which is your inversion region which is approximately 30 nanometers wide. If you increase the value of N_A so, you make your material more p-type, the total width of the depletion region will reduce correspondingly the width of the channel will also reduce. So, one last thing before we look or before we end the MOSFETS, the oxide layer in the case of a MOSFET acts as an insulator.

(Refer Slide Time: 46:07)



So, we can define a capacitance for the oxide layer or for the MOSFET then the capacitance is nothing but epsilon naught epsilon oxide times area divided by the thickness of the oxide layer. In the case of the IC industry, silicon is your material of choice for the MOSFET. So, the oxide layer that is used is always SiO_2 and SiO_2 has a relative permittivity of 3.9. Now, as the dimensions of the transistor start to come down, so, this is your scaling where you have more transistors that are packed within a given IC as the dimension starts to come down the dimensions of all the other components of your transistor will also have to come down. So, as the thickness reduces your capacitance will increase, but if your thickness becomes too small the oxide layer is very thin. So, that you can have tunneling between the made metal layers which is your gate and the n-channel and this will affect the properties of the channel. In order to prevent that the oxide layer is replaced by other hickey dielectric materials, you can still have a reasonable thickness while at the same time having a higher capacities. So, we replace the silicon dioxide with other materials with a higher ε_{or}.

So, Some of the other materials that are used are silicon nitride with a relative permittivity of 7, but metal oxides are also used. So, tantalum oxide is 25, titanium dioxide is anywhere from 60 to 100. In the recent IC industry, we replaced the metal oxide semiconductor with an insulator that has a high dielectric value.

With this we are done with the MOSFET part of the course and the transistor as well. Next, we will start to look at what happens when light interacts with your semiconductors? So, we are going to look at Optoelectronic devices like LED's and Solar cells. But before we do that, we will first treat the general interaction of light with semiconductors.