

Low Power VLSI Circuits and Systems
Prof. Ajit Pal
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

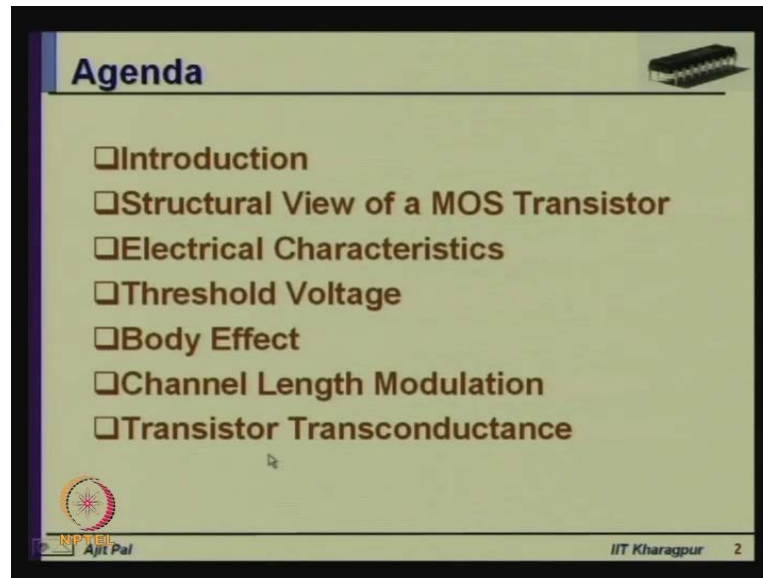
Lecture No. # 04
MOS Transistors – III

Hello and welcome to today's lecture on MOS transistors. This is the fourth lecture on the low power VLSI circuits and systems course and third lecture on MOS transistors. In the last lecture we have discussed how the operation of a MOS transistor can be visualized with the help of a very interesting model known as fluid model and fluid model as we have seen can be applied to charge controlled devices including MOS transistors. There we have seen how the MOS transistor operates and how the **the** current flow through the channel between the source and drain can be controlled with the help of gate voltage and also the drain voltage. And we have seen that there are three modes of operation.

One is the accumulation mode, another is the linear mode and third one is the saturation mode. And based on that the characteristic of MOS transistors can be divided into three regions as you have seen cut-off **cut-off** region, linear region, linear or active or you know that that partially charge controlled device. That means, partially that **that** is actually called when the inversion region is partial, weak inversion region and then the strong inversion region is represented by saturation region rather weak strong inversion mode is represented by saturation region.

And today I shall discuss the electrical characteristics of MOS transistors.

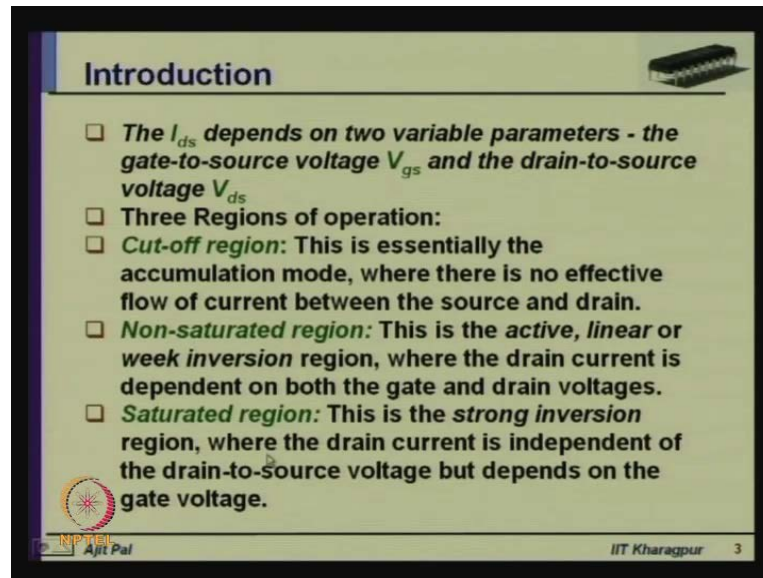
(Refer Slide Time: 03:02)



And here is the agenda of today's lecture after a brief introduction I shall **I shall** show once again the structural view of a MOS transistor and which I shall use to explain the operation and also derive expression for MOS, the **the the** drain current and particularly which is known as the electrical characteristics of MOS transistor.

Then, I shall discuss some about some of the important parameters like threshold voltage and another related subject related topic is body effect and we shall see how this **how the** substrate voltage can be changed to change the threshold voltage with the help of body effect. Then, another important phenomenon is known as channel length modulation I shall discuss about that and finally, I shall discuss about transistor trans-conductance.

(Refer Slide Time: 03:11)



Introduction

- The I_{ds} depends on two variable parameters - the gate-to-source voltage V_{gs} and the drain-to-source voltage V_{ds}
- Three Regions of operation:
- **Cut-off region:** This is essentially the accumulation mode, where there is no effective flow of current between the source and drain.
- **Non-saturated region:** This is the *active, linear* or *weak inversion* region, where the drain current is dependent on both the gate and drain voltages.
- **Saturated region:** This is the *strong inversion* region, where the drain current is independent of the drain-to-source voltage but depends on the gate voltage.

NPTEL
Ajit Pal
IIT Kharagpur 3

So, here is the introduction as I have already mentioned the drain current I_{DS} depends on two variable parameters; the gate-to-source voltage V_{gs} and drain-to-source voltage. Of course, here I did not mention about another important electrical parameter. That is your body bias that is your body-to-drain body-to-source voltage rather source-to-body voltage. That also affects the operation and with the help of which the threshold voltage can be controlled. So, essentially three voltages; gate-to-source voltage, drain-to-source voltage and the source-to-body voltage. These three voltages will control the drain current.

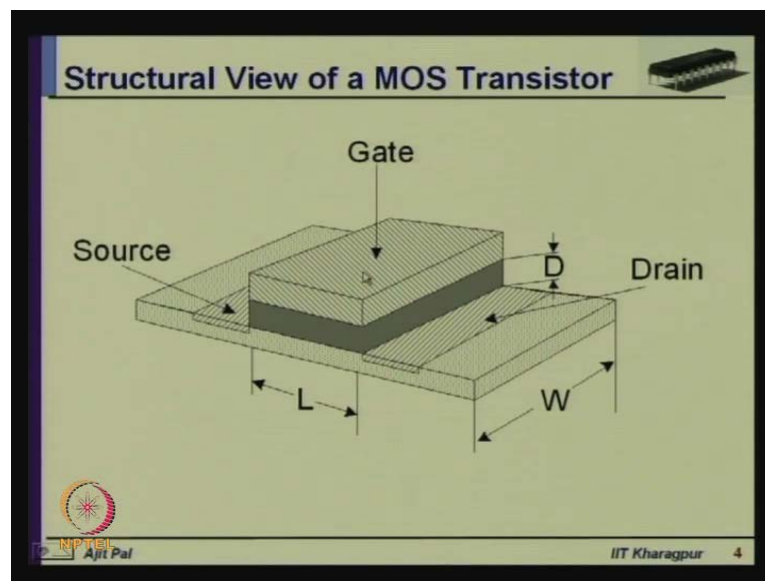
And as I have already mentioned there are three regions of operation; the cut-off region and this is represented by the accumulation mode where there is no effective flow of current between the source or drain source and drain. And we have already seen in this mode actually gate voltage is less than the threshold voltage and there is no channel formation. So, there is no inversion layer. So, there is no channel through which current can flow. So, in this particular region there is no flow of current, no effective flow of current between the source and drain.

Here I am telling effective flow. The reason for that is there is very small amount of current that flows through the device. But, for all practical purposes we can neglect it. That is why we are calling it effective flow of current I mean the no effective flow of

current. But, very small amount of current flows as we shall see that is represented by the sub-threshold leakage current. Later on we shall discuss about it in more detail.

Then, we shall consider the non-saturated region. This is the so-called active linear or weak inversion region as I have mentioned and here the drain current is dependent on both the gate voltage and drain voltages. So, **the the as** we shall see the expression for drain current will be dependent on both gate and drain voltages. Then the third region which is represented by the saturation region; this is the so-called strong inversion region where the drain current is dependent on the drain-to-source voltage. But, depends, but, does not depend on the here there is a mistake does not depend on the gate voltage **sorry** there is it does it is independent of the drain-to-source voltage, but, depends on the gate voltage. That means, in the saturation region the drain current solely depends on the gate voltage. It does not depend as we change the drain voltage as we shall see.

(Refer Slide Time: 06:01)



So, these are the three different regions. Now, I have already shown you this structural view in the last lecture and particularly there are three important physical parameters. I am calling it physical because when you do **when you do** the fabrication of a MOS transistor then you **youyouyou** will decide the length of the channel, the width of the channel and thickness of the silicon dioxide. These three are decided at the time of fabrication. So, these physical parameters will affect the operation of a MOS transistor and obviously, the electrical characteristics of the MOS transistor. So, we shall express

the drain current in terms of these three physical parameters L which is the length of the channel, W the width of the channel and D thickness of the silicon dioxide layer.

And as you can see on both sides of this channel we have got the source. I mean the **the** **the** diffusion regions, heavily diffusion regions and here also we have got another diffusion region that is called as drain. So, you have got source and drain and here is the gate on top of the silicon dioxide. We put polysilicon layer to form the gate and we can take electrical connections from source gate and drain to realize circuits. Later on we shall discuss about that.

(Refer Slide Time: 07:32)

Electrical Characteristics

$$I_{ds} = \frac{\text{Charge induced in the channel } (Q_c)}{\text{Electron transit time } (t_n)}$$

Q = CV, where C is the capacitance and V is the voltage applied across the capacitor

$$C = \frac{\epsilon A}{D}$$

where ϵ is the permittivity of the insulator in units of F/cm. For SiO₂, $\epsilon_{ox} = 3.9\epsilon_0$, where ϵ_0 is the permittivity of the free space

NPTVET
Ajit Pal

IIT Kharagpur 5

Coming to the electrical characteristics; I shall try to derive the expression for drain current in the same line of the fluid model. We have seen in the fluid model because of the voltage that we applied at the gate some charge is induced. So, inversion layer is created. Charge of opposite polarity in case of n MOS transistors it is the **it is the** electrons which are created and that charge can flow. Rather current can flow between source and drain whenever you apply some voltage between source and drain.

So, charge is I mean that in weak, in the inversion region is created because of the gate voltage that we apply. And then because of that creation of charge a conducting path is obtained which is known as channel between the source and drain and current will flow when you apply a voltage, you apply an electric field between the source and drain by applying a voltage to the drain with respect to the source.

(Refer Slide Time: 08:43)

I_{ds}

$I_{ds} = \frac{\text{Charge induced in the channel (Qc)}}{\text{Electron Transit time}(t_n)}$

$Q = C V$, C - in the capacitance
 V - Voltage.

$C = \frac{\epsilon A}{D}$ ϵ - permittivity
 $\epsilon_{ox} = 3.9 \epsilon_0$

$t_n = \mu_n E_{ds}$ μ_n - Mobility of Electron

$E_{ds} = \frac{V_{ds}}{L}$

So, let me write down the drain current expression. So, I_{ds} drain-to-source current. So, I_{ds} can be represented as charge induced in the channel **charge induced in the channel** which you can call at Q_c rather Q_c and that can be divided by the electron transit time **electron transit time** transit time is t_n . So, that means, here as you can see not only creation of charge is important or formation of inversion layer is **is** important; it is necessary that the electron will move from source to drain as we apply a voltage electron will from **from from** one end source end to the drain end and that will lead to a constant flow of current and as a consequence this electron transit time is important. So, as you know that Q that charge that is accumulated in a parallel plate capacitor is equal to C into V where C is the capacitance **C is the capacitance** and as you have seen it can be consider as a parallel plate capacitor and V is the voltage that is applied across the parallel plate capacitor.

And as you know this capacitance C is equal to ϵV ϵ , area of the plate and the D thickness of the silicon dioxide layer. So, here ϵ is the permittivity of the insulator **permittivity permittivity of the insulator** and this ϵ will be equal to you know in this case we shall be using silicon dioxide. So, ϵ_{ox} is equal to three point roughly approximately $3.9 \epsilon_0$ where ϵ_0 is the permittivity of the free space and this we shall use to represent the, to **to** find out the expression for the current.

Now, this is the expression for capacitance. What about the electron transit time? t_n will depend on what **what** thing? First of all it will depend on two things; number one is speed and second thing is the distance. So, you can say that transit time t_n will be equal to $\mu_n E_{ds}$. That means, μ_n is the mobility of electron **mobility of electron**. Here we are considering n MOS type transistor. So, the charge carrier are electrons. That is why we are considering mobility of electron and E_{ds} is the, **is a** drain-to-source electric field. We shall find out what is the drain-to-source electric field E_{ds} it is dependent on what? It is dependent on the voltage that we apply across the across the drain-to-source. So, the E_{ds} will be equal to V_{ds} by length l . So, E_{ds} by l .

Now, based on this we can find out the **the** electron transition time. So, electron transition time will be equal to.

(Refer Slide Time: 12:42)

$$v_n = \frac{\mu_n V_{ds}}{l}$$
 Velocity of electron

$$t_n = \frac{L}{v_n} = \frac{L^2}{\mu_n V_{ds}}$$

Typical value of $\mu_n = 650 \text{ cm}^2/\text{V}$
 (at room temp.)

The Non saturated region

$$V_{eff} = (V_{gs} - V_t - \frac{V_{ds}}{2})$$

A circuit diagram of a MOSFET is shown with gate voltage V_{gs} , drain voltage V_{ds} , and channel length l . The source is labeled 'S' and the drain is labeled 'D'.

First of all let us find out the velocity of electron v_n . v_n is equal to $\mu_n V_{ds}$ by l and therefore, t_n electron transition time, this is the velocity of electron and t_n will be equal to electron transition time that will be equal to l by v_n , length by v_n that will be equal to l^2 by $\mu_n V_{ds}$. So, we have got the **the** expression for the electron transition time and **and** it has been found that typical value **typical value** of μ_n **μ_n** that is the mobility of electron is equal to 650 centimeter square by volt and that is at room temperature.

So, we have got the expression, we have got the **the** value the **the** charge that we can obtain by C into V and we can we have, also found out the electron transition time t . Now, we can find out the current I_{ds} now to do that we shall divide the entire region into three parts as we have seen. Number one is cut-off region when the voltage is less than the threshold voltage, there is no channel formation. So, there is no current. So, we are not writing any expression for current for that.

Now, we shall write down the expression for the non-saturated region. Non-saturated or we also call it linear or weak inversion region. In this weak inversion region, we can find out the **the the the** current. First of all let us find out the effective voltage that is being applied between the source and drain. As you have seen, we are applying a voltage V_{ds} between the source and drain **sorry** on the gate I mean V_{ds} between the source and drain. Let me draw the diagram here little that will make it easier to understand. So, here is your source and here is your drain. This is the channel then you have got the silicon dioxide and top of that you have polysilicon.

So, what you are doing? You are applying a voltage V_{ds} , positive voltage. So, this is your V_{ds} between the source and drain and you are all supplying a voltage V_{gs} between source and gate. This is your source drain and this is your gate. So, V_{gs} , now although you are applying a voltage V_{gs} between the source and gate; the effective voltage across the channel is not same because this drain voltage that you are applying is interacting with the voltage that you apply at the gate. So, although you are applying a gate voltage V_{gs} here, the effective **the effective** voltage in this channel region will be different.

For example in here the voltage you know that the **the the the** voltage V_{ds} will be **will will be** across this length L and here you can say the voltage is 0 and here the voltage is V_{ds} . So, if you take the average value of 0 and V_{ds} this will be equal to V_{ds} by 2. Therefore, the effective gate voltage will be equal to V_{gs} that you apply here minus V_t threshold voltage as we know, **you the** you have to subtract that threshold voltage because you know the before that there will be depletion region. You will you will require a voltage V_t only after that the inversion region starts and as a consequence the effective gate voltage will be V_{gs} minus V_t minus V_{ds} by 2. We have taken the average value. Of course, it will be **it will be** varying across the channel, but, we are taking the average value. This is the effective gate voltage.

(Refer Slide Time: 17:55)

The image shows handwritten mathematical derivations on a light blue background. At the bottom left, there is a small circular logo with a star and the text 'NPTEL' below it.

$$Q_c = C \cdot V = \frac{\epsilon_{ox} \cdot W \cdot L}{D} \left[V_{gs} - V_t - \frac{V_{ds}}{2} \right]$$

$$I_{ds} = \frac{Q_c}{t_n} = \frac{\epsilon_{ox} \cdot W \cdot \mu_n}{D \cdot L^2} \left[V_{gs} - V_t - \frac{V_{ds}}{2} \right] V_{ds}$$

$$= \frac{W \cdot \mu_n \cdot \epsilon_{ox}}{D \cdot L} \left[V_{gs} - V_t - \frac{V_{ds}}{2} \right] V_{ds}$$

$$K = \frac{\mu_n \cdot \epsilon_{ox}}{D}$$

$$I_{ds} = \frac{K \cdot W}{L} \left[V_{gs} - V_t - \frac{V_{ds}}{2} \right] V_{ds}$$

$$C_g = \frac{\epsilon_{ox} \cdot W \cdot L}{D}$$

$$I_{ds} = \frac{C_g \cdot \mu_n}{L^2} \left[(V_{gs} - V_t) - \frac{V_{ds}}{2} \right] V_{ds}$$

Now, the current, we can derive the expression for current we can expression for current. For that purpose first we shall find out the Q c. This Q c charge induced in the channel is equal to C into V. That means, C is equal the Q c is equal to C into V that is equal to C. What is C value of C? Epsilon, this is your oxide into into area that is equal to W into L because we have seen that parallel plate capacitor has an area W width W and L is the length. So, area is equal to W into L by D. This is the this is the capacitance and your effective voltage will be equal to V g s minus V t minus V d s by 2. So, this is the this is this is a charge that will be induced in the channel.

So, this is the charge and now we know the we already know the electron transit time. What was the electron transit time that you find out? This is equal to L square by mu v d s. So, this we can substitute. That means, I d s I d s is equal to Q c by t n that is equal to epsilon oxide into W into L by D into V g s minus V t minus V d s by 2 and it will be divided by L square. This is the expression for electron transit time. This is the L square t n is equal to L square by mu n d V s.

So, what will happen; here you will have L square and mu will be here, mu n will be here in the numerator it will go to the numerator and here you will have V d s. So, this is the this is the expression that we get. This is mu n and this l will cancel with square. So, ultimately we get an expression that is equal to W L. It was L square W L by d, the current expression that we get is equal to W mu n. This is mu n epsilon o x. This is your

permittivity of electron into D into L and here you get $V_g - V_t - V_d$ by 2 into V_d . Now, this is the **this is the this is the** current that flows through the electron. So, that will be the current that will be flowing through the electron.

Now, you can express in terms of two other important parameters; one is known as K . K is a constant which depends on the fabrication. So, what is the value of K ? K is equal to we can express it this current in terms of K which is equal to μ_n . This μ_n and $\epsilon_0 \times D$. You know when you fabricate; these three parameters are kept usually constant. We usually change the length, this L is changed, width is changed, but, these three are usually kept constant. That is why this is sometimes is replaced by K . That means, if we substitute this value of K we get I_d is equal to $K \frac{W}{L} (V_g - V_t - V_d)^2$ into V_d .

Another we can also represent it in terms of what is known as the gate capacitance C_g . What is the value of C_g ? C_g is equal to this is your $\epsilon_0 \times$ into W/L by D . That you have already seen that is a capacitance. This **this** is the, you have already seen $\epsilon_0 \times A$ by D here A is equal to W by L . So, we can also express I_d in terms of this capacitance. So, in terms of K constant in terms of C_g . So, I_d we can write as I_d is equal to $C_g \mu_n \frac{W}{L^2} (V_g - V_t - V_d)^2$ into V_d . So, this is the current that we get whenever the circuit is in the linear region.

It may, you may note that we are calling it linear region, but, it is not exactly linear. Why it is not exactly linear because you can see, if we assume that V_d by 2 is too small compared to $V_g - V_t$. Only then it becomes proportional to V_d . That means, if we neglect this part only then it becomes linear and that is a reason why the **this the** linear region is not exactly linear, but, it approximately linear, particularly when the drain voltage is very close to 0. Only then it is linear so; that means, V_d is negligible compared to $V_g - V_t$ then is linear.

So, we have derived expression for the non-linear region.

(Refer Slide Time: 24:08)

1 Saturated Region

$$V_{ds} = V_{gs} - V_t$$

$$I_{ds} = \frac{kW}{L} \left[(V_{gs} - V_t) - \frac{V_{ds}}{2} \right] \cdot (V_{ds} - V_t)$$

$$= \frac{kW}{L} \left[(V_{gs} - V_t)^2 - \frac{(V_{gs} - V_t)^2}{2} \right]$$

$$= \frac{kW}{2L} (V_{gs} - V_t)^2$$

$$I_{ds} = \frac{C_{ox} \mu_n}{2L^2} (V_{gs} - V_t)^2 = \frac{\mu_n W C_{ox}}{2L} (V_{gs} - V_t)^2$$

$C_{ox} = C_{ox} \cdot W \cdot L$

NPTTEL

Now, let us consider the saturated region. So, you can say the second part is your saturated region. So, in the saturated region what is the K difference between the linear region and saturated region? One important difference is in the saturated region, the current is independent of the drain voltage and when it happens? When the voltage drop across the channel is equal to the effective gate voltage. That means if the drain voltage is such which is equal to V_{ds} is equal to $V_{gs} - V_t$. That means, this occurs when V_{ds} is equal to $V_{gs} - V_t$. This is the effective gate voltage which is equal to the drain voltage beyond which, beyond this point the drain current is independent of V_{ds} . That means, this is the point from where we can consider we can say that saturated region has started below that it is linear or non-saturated region.

So, what we can do, we can obtain the expression for drain current in the saturated region by substituting V_{ds} is equal to $V_{gs} - V_t$. So, we can say that I_{ds} is equal to $\frac{kW}{L} \left[(V_{gs} - V_t) - \frac{V_{gs} - V_t}{2} \right] \cdot (V_{gs} - V_t)$. So, this is equal to $\frac{kW}{L} \left[\frac{V_{gs} - V_t}{2} \right] \cdot (V_{gs} - V_t)$. So, this is equal to $\frac{kW}{2L} (V_{gs} - V_t)^2$. So, it will be equal to $\frac{kW}{2L} (V_{gs} - V_t)^2$.

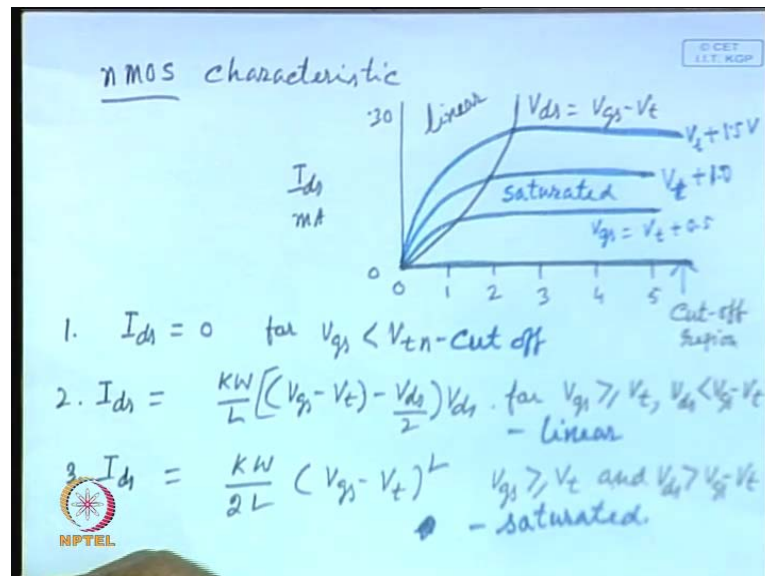
So, this will be equal to $\frac{kW}{2L} (V_{gs} - V_t)^2$ because half this if we $V_{gs} - V_t$ square minus $V_{gs} - V_t$ square by 2 will become $V_{gs} - V_t$ square by 2. So, this will be equal, this two we have taken here. So, $\frac{kW}{2L} (V_{gs} - V_t)^2$.

So, here indeed we find an expression where it is independent of the drain voltage. So, but, it is there is a **it is a there is a** square law dependence on $V_{gs} - V_t$. That means, this saturation current will increase at the square law as the gate voltage is increased. So, that is proportional to $V_{gs} - V_t$ square. So, this is the expression for drain current and also you can express it in terms of the capacitance that will be equal to $C_g \mu_n$ by $2L$ square into $V_{gs} - V_t$ square.

So, this is the expression for that. That in terms of C_g gate capacitance and of course, you can **you can** represent it with the help of another parameter. There is another parameter which is known as C_{ox} **C o x** and C_g these two are different this C_{ox} is the unit gate capacitance. That means, capacitance of unit area. So, what is a relationship between C_g and C_{ox} ? C_g will be equal to C_{ox} into W that area you have to multiply W into L . So, C_{ox} into W into L if we substitute it here. We can write it in this way, this L it will become L^2 this L will cancel out. So, it will be μ_n will be there in any case then w will be there and it will be C_{ox} by 2 into L and here it will be $V_{gs} - V_t$ square. So, this is in terms of the unit gate capacitance C_{ox} .

So, we have seen the expression for saturation current.

(Refer Slide Time: 29:01)



Now, let us represent the n MOS electrical characteristics **n MOS characteristics**. Let us now plot that saturation and non-saturation current. We have seen that we can divide into three regions. Let us assume it varies from 1 2 3 4 5; 1 volt 2 volt 3 volt 4 volt and 5 volt

and this I_{ds} usually varies in the range of in the milliamper region. So, it is milliamper. So, 0 to let us assume it is 0.30 milliamper. So, usually it will vary in this region.

Now, as you have seen there are three regions; I_{ds} is equal to 0 for V_{gs} less than V_t . We can write V_{tn} if we consider it n MOS transistor. So, which part will represent that? This line, that means, this line will correspond to this particular line will correspond to this region. That means, irrespective of the drain voltage current is 0, I_{ds} is 0. So, this represents the cut-off region **cut-off region**. So, this **this** is your cut-off region.

Now, non-linear, linear region I_{ds} is equal to we have already seen the expression. We can write it $K \frac{W}{L} \mu_n (V_{gs} - V_t - \frac{V_{ds}}{2}) V_{ds}$. This part as we have seen is separated by a line. I mean this corresponds to V_{ds} is equal to $V_{gs} - V_t$. That means this part is your linear region. So, obviously, there will be different curves for different gate voltages.

Say this is this corresponds to V_t . Let us assume V_{gs} is equal to V_t plus 0.5 volt. Let us assume this corresponds to V_t plus 1 volt 1.0 volt. This will correspond to let us assume this is V_t plus 1.5 volt. So, **the** in this part the current is constant **this part current is constant**. So, this is a saturation region. So, this is the linear region, this is the cut-off region and this is the saturation region **saturated region** and in the saturated region the current expression is given by and this is for this non-saturated region is for V_{gs} is greater than equal to V_t and V_{ds} is less than $V_{gs} - V_t$. That means, V_{ds} has to be less than $V_{gs} - V_t$ in this region and V_{gs} has to be greater than V_t .

So, for these two we call it linear region. This is cut-off **this is your cut-off** region **cut-off**. This corresponds to cut-off region. This corresponds to linear non-saturated or weak inversion region and this correspond the third region this is your 1 2 and 3 I_{ds} is equal to, we have seen the expression is $K \frac{W}{L} \mu_n (V_{gs} - V_t)^2$ and this happens when V_{gs} is greater than equal to V_t and also V_{ds} has to be greater than $V_{gs} - V_t$. So, this is your saturated region. This corresponds to saturated.

So, we have seen, we have derived the expression for drain current for two regions; linear and saturated regions and we have defined the regions based on the drain voltage and gate voltage. Here, we have not taken into account the body voltage with respect to the source. We have assumed that the body is connected to the source.

(Refer Slide Time: 34:22)

$$v_n = \frac{\mu_n V_{ds}}{L}$$
 Velocity of electron

$$t_n = \frac{L}{v_n} = \frac{L^2}{\mu_n V_{ds}}$$

Typical value of $\mu_n = 650 \text{ cm}^2/\text{V}$
 (at room temp.)

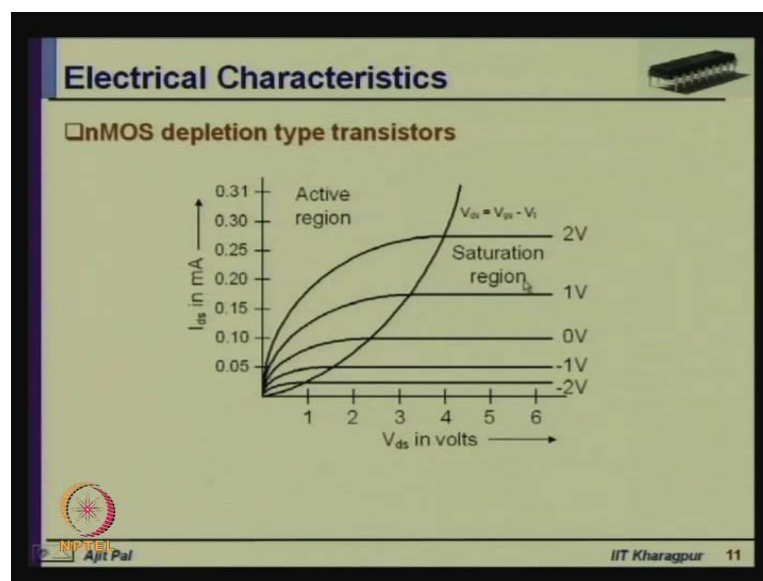
The Non saturated region

$$V_{eff} = (V_{gs} - V_t - \frac{V_{ds}}{2})$$

The diagram shows a cross-section of a MOSFET with labels: S (Source), G (Gate), D (Drain), and B (Body). The source and body are connected to ground. The gate is connected to V_{gs} and the drain is connected to V_{ds} . The channel length is labeled as L .

That means, if we consider this diagram, this simple diagram we can say this body is connected to ground. That means, this is connected to ground. Source is connected to ground and body is also connected to ground and usually this is done you know by having a region, this is called P plus and this P plus is actually used to connect the substrate to ground or to some other voltage. Later on we shall see how we can really apply this voltage to have different threshold voltages. So, for the time being let us assume that this drain, that source to substrate voltage is 0.

(Refer Slide Time: 35:14)

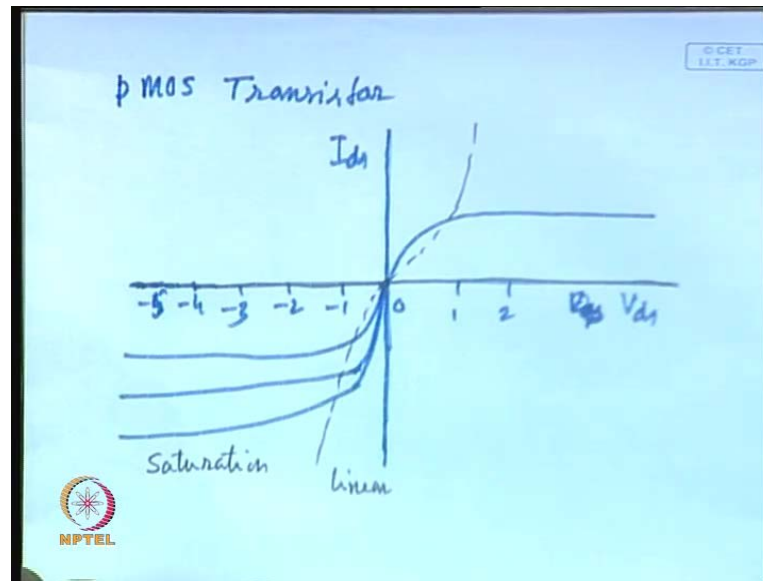


Now, we have considered the n MOS depletion type, we have n enhancement type transistor. Let us now focus on n MOS depletion type transistor and as we know in case of n MOS depletion type transistor, a channel formation is done by implanting suitable type impurity in the channel region. So, a conducting layer already exists and as a result you do not require a gate voltage to **to** form a channel. So, current can flow even when the gate voltage is 0 and you can **you can** stop the flow of current by applying a negative gate voltage and that is precisely shown in this diagram. You can see here even when the gate voltage is 0; there is a flow of current and if you want to stop it you have to apply negative voltage. That means, minus **minus minus** 1 volt minus 2 volt and so on.

Otherwise, there is no difference in the characteristic curve. I mean it is very similar to the n MOS depletion type. I mean enhancement type transistor except that these gate voltages are different for different currents. That means, normally in case of enhancement type transistor this will start when the gate voltage is slightly more than the threshold voltage. But, here as you can see this is minus 2 volt. So, in case of depletion mode of transistors only difference that we will get is the for different gate voltages you will get different curves. That means, you have to apply a negative voltage to stop the flow of current and if you apply positive gate voltage then current will increase and obviously, as we know that it will increase following square law $V_g - V_t$ square.

So, we have seen the characteristic curve for n MOS type transistors both enhancement type and depletion type. What about the P type transistors?

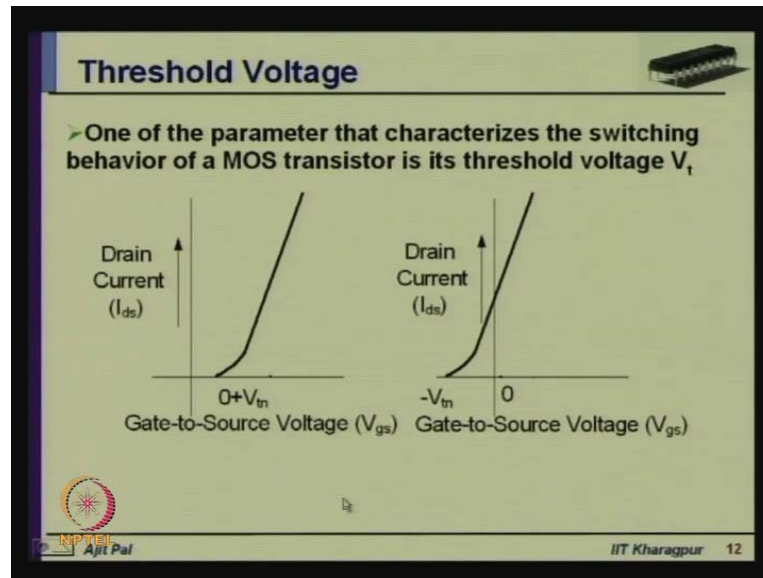
(Refer Slide Time: 37:18)



Actually that can be drawn say P MOS transistors. We shall not derive expression for that I shall simply show you the characteristic curve, how it will look like. So, you have got four quadrants. You can see here, you have got four quadrants. This is your I_{ds} and this is your V_{gs} . This is positive and this is negative. For n MOS transistors, the curve is on this side. On the other hand for P MOS transistors the curve will be on the negative side. That means, you will be applying your the voltage, that drain voltage **sorry** this is your V_{ds} . I have written wrongly V_{ds} . So, here it will be negative here also V_{ds} , but, here it is say 0 1 2 here it is minus 1 minus 2 minus 3 minus 4 minus 5 and so on

That means you will apply negative gate, negative voltage to the drain with respect to source and current will flow in the opposite directions. And these are for different **different** gate voltages and more negative they correspond to more negative gate voltages. And just like your n MOS transistors here also, there is a there are three regions cut-off. So, just like your n MOS **n MOS** transistor here also, you have got linear region, this is a saturation region and this line obviously, will represent the cut-off region. This is a cut-off region for p MOS transistor, this is a cut-off region for n MOS transistor. So, on; that means, this **this this** the **the** characteristic curve for p MOS transistor will be on this quadrant both for enhancement type and depletion type.

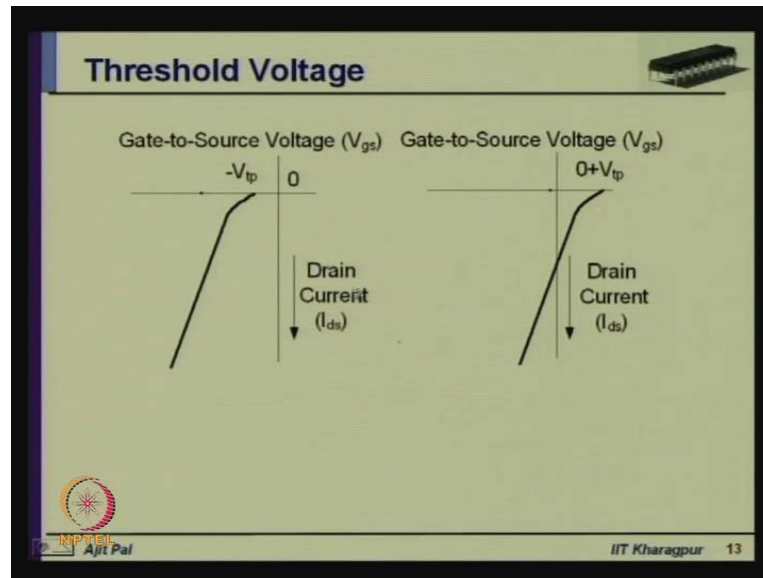
(Refer Slide Time: 39:52)



Now, let us focus on the threshold voltage. This threshold voltage has been found to be one of **one of** the most crucial parameter which controls the operation of a MOS transistors **MOS transistor**. So, here you can see this corresponds to n MOS enhancement type transistor. You have to apply a threshold voltage I mean V_{tn} to start the flow of current and characteristic curve will be somewhat like this. That this is the drain current will increase and as you know this essentially represents the **the the the** drain current in the saturation mode. That means, in the saturation mode, although I have drawn it linearly, but, it will be quadratic. As we have seen this current is proportional to I_{ds} is proportional to $V_{gs} - V_{tn}$ square. So, it increases at this rate.

So, far as the depletion type transistor is concerned; curve is somewhat similar, but, it will be shifted towards the **towards the** negative quadrant. As you can see you have to apply a minus V_{tp} , a threshold voltage which is the minus V_{tp} in this case to stop the flow of current and if you increase the **the** I mean negative to less negative to 0 and then positive; then the drain current increases following the same nature of curve and of course, the **the** amount of current will be dependent on the gate voltage. So, you can see **the the thethe** these two essentially represent the **the** drain current with respect to gate-to-source voltage for enhancement type transistor and for depletion type transistor.

(Refer Slide Time: 41:42)



Similarly, you can consider the gate-to-source voltage for p MOS type transistors. This corresponds to the characteristic curve for p MOS type transistor. Here you can see the variation of drain current with respect to the **gate volt** gate-to-source voltage. And a negative voltage is required to for the flow of current and you can see the curve is somewhat similar, but, inverted in this case because in the it is in the negative quadrant. Similarly, this is this **this** curve correspond to **corresponds to** the p MOS depletion type transistor and here as you can see you have to apply a positive voltage to stop the flow of current and as you **as you** make it smaller and smaller and when it becomes 0; this side is the gate-to-source voltage. So, when it becomes 0 still there is a flow of current and as you make it more and more negative; the current increases in case of p MOS depletion type transistor.

So, these four; these two and these two represent the **the** variation of gate drain current with respect to the variation of gate-to-source voltage and you can see the threshold voltage actually plays **plays** acts somewhat like the cutting voltage of you know cutting voltage of bipolar junction transistors. As you know in case of bipolar junction transistors flow of current is dependent you know, you require a cutting voltage. For example, for silicon the base **base** voltage should be more than 0.6 or 0.7 volt. That is known as cutting voltage. Here also here also it is somewhat similar to that, but, we call it threshold voltage not cutting voltage.

(Refer Slide Time: 43:39)

Threshold Voltage

$$V_t = V_{t0} + \gamma \left(\sqrt{|-2\phi_b + V_{sb}|} - \sqrt{|2\phi_b|} \right)$$

V_{t0} = Threshold voltage at $V_{sb} = 0$

$$\gamma = \frac{\sqrt{2q\epsilon_{si}N_A}}{C_{ox}} = \text{Substrate bias coefficient}$$
$$\phi_b = \frac{KT}{q} \ln \left(\frac{n_i}{N_A} \right) = \text{Fermi band potential}$$

- n_i is the carrier concentration of the intrinsic silicon
- N_A is the carrier concentration of the substrate

MPTET
Ajit Pal
IIT Kharagpur 14

Now, one **one** very important as I told threshold voltage plays a very important role and on what factors or parameter it depends? You can really derive an expression for threshold voltage in terms of various parameters, but, here I have simply taken the expression without and I am not going to derive it and as you can see V_t , the threshold voltage is equal to V_{t0} plus γ into $\sqrt{|-2\phi_b + V_{sb}|}$ minus $\sqrt{|2\phi_b|}$ absolute value of that and minus absolute value of $2\phi_b$.

Here this γ is known as substrate bias coefficient. Why substrate bias coefficient? As you can see this V_{sb} **V_{sb}** is actually the source-to-body voltage that you are applying **source-to-body voltage that you'll be applying**. That means, if we consider a **a** curve here you can apply a voltage between the source and body. That means, **the** this particular source and this body you can apply positive voltage to the body with respect to the source and that will actually increase the threshold voltage as we shall see.

So, by changing the body bias, substrate body bias the threshold voltage can be controlled and in this expression you have got V_{sb} . That is why it is called substrate bias coefficient γ . And γ is equal to $\sqrt{2q}$. q is the charge of electron ϵ_{si} which is the permittivity of silicon and n_a **n_a** is the carrier concentration in the **in the in the** substrate. So, n_a is the carrier concentration of the substrate and C_{ox} **C_{ox}** is the unit **unit** capacitance of the gate region.

So, this is the expression for gamma. So, it will depend on these parameters and phi b is the fermi band potential which is equal to K t by q. K is the Kelvin constant and t is the temperature in absolute value of temperature not in degree or Fahrenheit this is the absolute value of temperature and Q is the charge of electron and n i is the carrier concentration of the intrinsic silicon. So, intrinsic silicon will have some carrier concentration. This is the n i and n a is the carrier concentration of the substrate as I have already told.

So, based on that you can find out the threshold voltage by substituting different parameter values for a particular fabrication technology. Obviously, So, far as this part is concerned Q epsilon s I, K these are constants. On the other hand the C o x is the unit gate capacitance or these carrier concentrations you can will vary from process one process technology to the other process technology and as a consequence this threshold voltage will be different.

(Refer Slide Time: 47:03)

Threshold Voltage

$$V_t = V_{t0} + \gamma \left(\sqrt{-2\phi_b + V_{sb}} - \sqrt{2\phi_b} \right)$$

➤ **Example**

$$\phi_b = \frac{KT}{q} \ln \left(\frac{n_i}{N_A} \right) = 0.026 \ln \left(\frac{1.45 \times 10^{10}}{10^{16}} \right) = -0.35$$

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3.97 \times 8.85 \times 10^{-14}}{500 \times 10^{-8}} = 7.03 \times 10^{-8} \text{ F/cm}^2$$

$$\gamma = \frac{\sqrt{2q\epsilon_{si}N_A}}{C_{ox}} = \frac{\sqrt{2 \times 1.6 \times 10^{-19} \times 10^{16} \times 11.7 \times 8.85 \times 10^{-14}}}{7.03 \times 10^{-8}} = 0.82$$

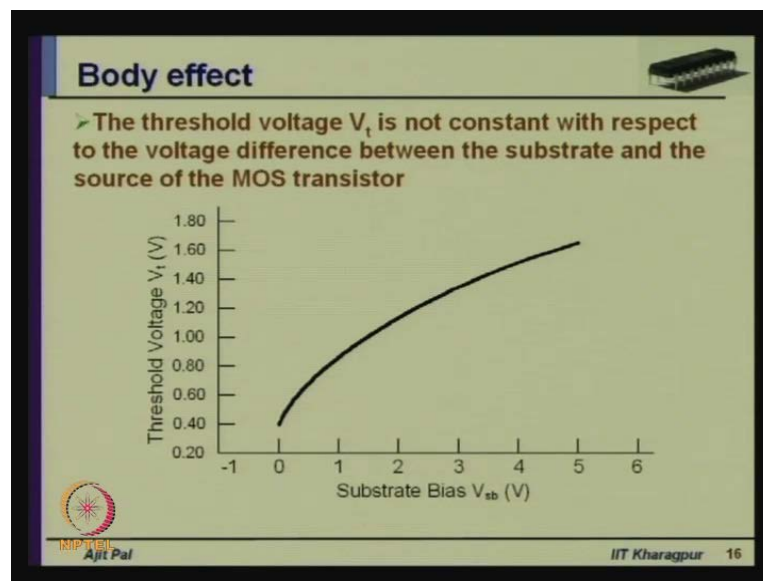
$$V_t = V_{t0} + \lambda \sqrt{-2\phi_b + V_{sb}} - \sqrt{2\phi_b} = 0.4 + 0.82 \sqrt{0.7 + V_{sb}} - \sqrt{0.7}$$

Now, let us take up an example. Let us see the, we can calculate the threshold voltage and we can see how exactly the threshold voltage varies as you vary the substrate bias phi b is equal to you can substitute different parameter values. This n i is equal to 1.45 into ten to the power ten centimeter. I mean per centimeter cube and n a is equal to 10 the power 16. So, 0.26 actually this K t by Q as you know is 26 millivolt. That is the K t

by Q value and this is the ratio of the n_i by n_a . Obviously, the intrinsic silicon has much lower carrier concentration compared to the carrier concentration of this substrate.

That means as you know substrate is uniformly doped to have more carrier concentration compared to the intrinsic substrate and ϕ_b value is calculated. See this is the value of C_{ox} . This ϕ_b is minus 0.35 and C_{ox} is equal to ϵ_{ox} by t_{ox} and this is the value 7.03×10^{-8} farad per centimeter square and you can substitute the various values to get γ . γ is equal to, becomes 0.82. Here the value of the charge of electron the permittivity of silicon dioxide n_a , these are all substituted and C_{ox} these are all substituted and by substituting various values we get V_t is equal to roughly 0.4. Assuming that 0.4 is the threshold voltage whenever the body bias is 0; that means, source is connected to the body and this part is **is** will this is the 0.82 this is your γ into $\sqrt{0.7 + V_{sb}}$. We have to take the absolute value and minus $\sqrt{0.7}$. Now, you can actually vary V_{sb} to get different **different** value of threshold voltage.

(Refer Slide Time: 49:38)



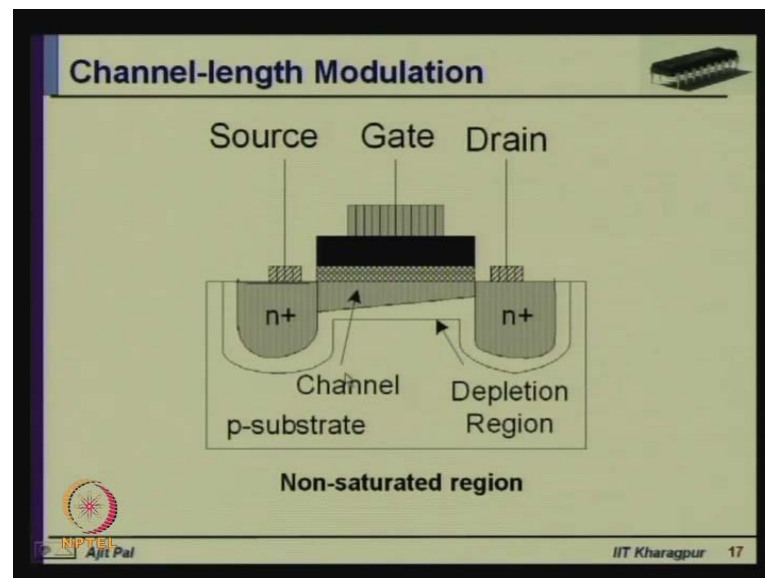
So, threshold voltage v_t is not constant with respect to the voltage difference between the substrate and the source of the transistor. As you can see using that expression how the threshold voltage can change as you as you change the substrate body bias. So, you can see here, when the substrate bias is 0, as you have seen the threshold voltage is 0.4 volt. But, as you increase the substrate bias say from say 0 to 1 volt to 2 volt to 3 volt, it

can it can reach say 1.6. So, over a large range 0.4 to 1.6, a large change in the threshold voltage that can take place as you vary the threshold, the substrate bias with and threshold voltage will change and later on we shall use this and see how the threshold voltage can be controlled to reduce what is known as leakage current.

Because the leakage current is heavily dependent on the threshold voltage and later on we shall see instead of this positive bias, we shall usually apply negative body bias to low to **to to** increase this. I mean **with** this is called the reverse body bias we shall do that to increase the threshold voltage such that the leakage current is lowered. That means, larger the threshold; voltage smaller is the leakage current. Smaller the threshold voltage larger is the leakage current. Later on we shall discuss more about it.

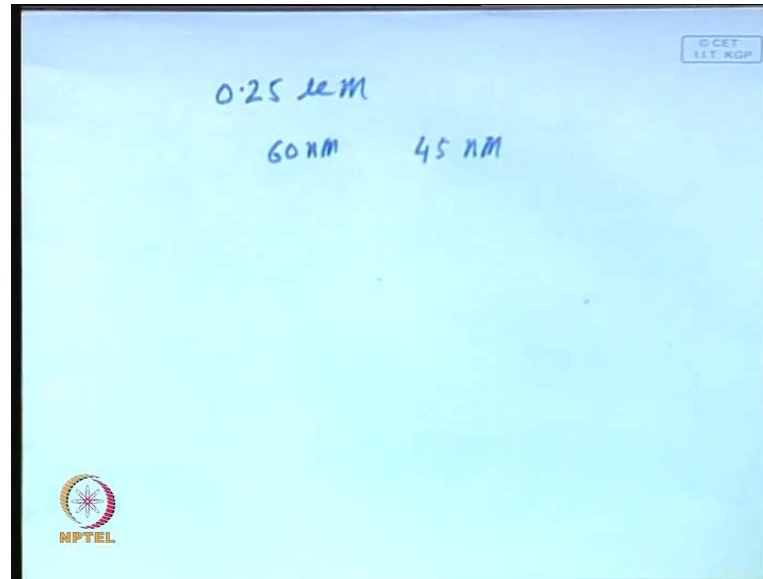
So, this is a body effect

(Refer Slide Time: 51:30)



Now, we shall consider another very important parameter that is your channel-length rather phenomenon, channel-length modulation. Here in our expression for drain current, we have assumed that this l is constant fixed. It does not really change, but, in reality we will see this l is dependent on the drain voltage. However, this dependence is not visible until the channel-length is small. That means, it is, it becomes predominant in short channel devices.

(Refer Slide Time: 52:16)



When the channel-length is say 0.25 micron **0.25 micron** micro millimeter then this variation is so small that we do not really bother about it. But, whenever you are considering say sixty nanometer **nanometer** or say 45 nanometer technology instead of 250 nanometer technology, then this channel-length variation becomes visible or it cannot be neglected any longer.

First we shall show you why this variation (()) take place. You can see the channel region is not uniform throughout the channel the reason for that is here the drain voltage is 0 here the drain voltage is high and because of the interaction between the gate voltage and drain voltage; **the dense** the **the** thickness of the channel region is not near the source and it is very narrow near the drain. And as we further increase the drain voltage what happens at some point of time the very close to the drain there is no carrier. That means, it becomes pinch-off. If we increase the drain voltage further, we will find that **that that** pinch-off point moves towards source. So, here there is small region where there is no charge carrier. So, essentially channel-length is now becoming smaller than the **the** physical length of the device.

So, effective channel-length is smaller and this is the, that is why we call it channel-length modulation. Essentially it is dependent on the drain voltage.

(Refer Slide Time: 53:57)

Channel-length Modulation

$$L_{eff} = L - \Delta L$$

$$I_{ds(sat)} = \frac{1}{\left(1 - \frac{\Delta L}{L}\right)} \cdot \frac{\mu_n C_{ox}}{2} \cdot \frac{W}{L_n} (V_{gs} - V_{th})^2 \quad \Delta L \propto \sqrt{V_{ds} - V_{dsat}}$$

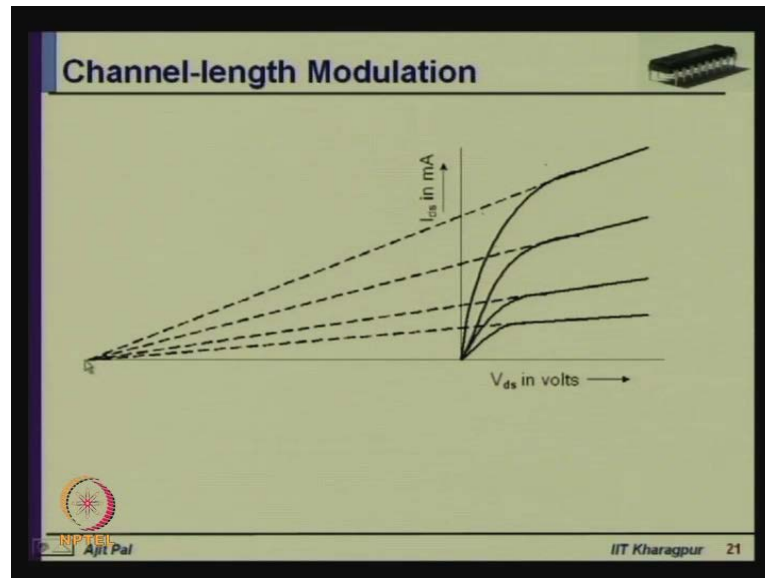
$$1 - \frac{\Delta L}{L} \approx 1 - \lambda V_{ds} \quad I_{ds(sat)} = \frac{\mu_n C_{ox}}{2} \cdot \frac{W}{L_n} (V_{gs} - V_{th})^2 (1 + \lambda V_{ds})$$

NPTTEL Ajit Pal IIT Kharagpur 20

How do you take into account? We can take into account by substituting L effective which is equal to L minus ΔL . So, n minus ΔL is actually we can represent it by this **we can represent it by this** $1 - \frac{\Delta L}{L}$ this is the expression where we have substituted effective channel-length and we have got this expression. And this **this** variation is proportional to $V_{ds} - V_{dsat}$. Essentially, when the saturation starts that is your $V_{gs} - V_{th}$ and this is the V_{ds} . So, this change in the channel-length is proportional to the drain voltage. When the drain voltage is large only then this reduction occurs and this can be represented, this change $1 - \frac{\Delta L}{L}$ by L can be represented by one minus λV_{ds} where λ is **the is** known as channel-length modulation coefficient.

So, we can now express I_{ds} in term in terms of this λ and it becomes like this, $\mu_n C_{ox}$ by $2 W$ by L_n into $(V_{gs} - V_{th})^2$ into $1 + \lambda V_{ds}$. So, you can see here, it is no longer one. It is here normally it is V_{ds} . So, what it is $1 + \lambda V_{ds}$. So, it varies with drain voltage.

(Refer Slide Time: 55:21)



And this can be depicted pictorially with the help of this diagram as you can see in the saturation region, **the** although I have drawn it in a little exaggerated form, it will not be the slope will not be so high. What you can see the drain current is varying even in the saturation region because of channel-length modulation effect and you can see here it is becoming the it is no longer parallel, but, it is there is a point where it is touching and this is how the characteristic can be represented.

(Refer Slide Time: 55:58)

$$g_m = \left. \frac{\delta I_{ds}}{\delta V_{gs}} \right|_{V_{ds} = \text{const}} \quad I_{ds} = \frac{Q_c}{t_{sd}} \quad \delta I_{ds} = \frac{\delta Q_c}{t_{sd}}$$

$$t_{sd} = \frac{L^2}{\mu_n V_{ds}} \quad \delta I_{ds} = \frac{\delta Q_c}{L^2} V_{ds} \mu_n \quad \delta Q_c = C_g \delta V_{gs}$$

$$\delta I_{ds} = \frac{\mu_n C_g}{L^2} V_{ds} \delta V_{gs} \quad \text{or} \quad g_m = \frac{\delta I_{ds}}{\delta V_{gs}} = \frac{C_g \mu_n V_{ds}}{L^2}$$

$$V_{ds} = (V_{gs} - V_t) \quad g_m = \frac{\mu_n \epsilon_{ins} \epsilon_0 W}{D L} (V_{gs} - V_t)$$

Now, another very important parameter that is known as transistor trans-conductance for bipolar junction transistors which is a current controlled device it is represented by gain of the transistor beta the collector current by base current. On the other hand in case of MOS transistors it is a voltage controlled device and similar parameter is known as trans-conductance. That means, rate of change of drain current by rate of change of gate voltage. So, here you have seen the **the** drain current is controlled by the gate voltage.

So, this trans-conductance is represented by $\frac{\Delta I_D}{\Delta V_{GS}}$ change of change of drain current by change of gate voltage ΔV_{GS} keeping the drain-to-source voltage constant. So, that we can derive as we know I_D is equal to Q_C by $t_s D$. So, ΔI_D is equal to ΔQ_C by $t_s D$ and this is the $t_s D$. That expression we have already derived I am not going into the details of that. So, $\frac{\Delta I_D}{\Delta V_{GS}}$ is equal to $\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{th})$ and g_m is equal to $\frac{\Delta I_D}{\Delta V_{GS}}$, as we have seen which is equal to $C_{ox} \mu_n (V_{GS} - V_{th})$.

And finally, by substituting V_{DS} is equal to $V_{GS} - V_{th}$, we get g_m is equal to $\mu_n \epsilon_0 \frac{W}{L} \frac{q}{D} (V_{GS} - V_{th})$. So, you can see the **the** gain which is represented by transistor trans-conductance. More the gain it is better. We can see it is dependent on the mobility of the device, width of the device, thickness of the silicon dioxide length of the channel and the effective gate voltage which is equal to $V_{GS} - V_{th}$. So, this is the parameter and we can, how we can increase the value of g_m is clear from this expression which parameter we can really; that means, if the width is more **more** current will flow. So, that is why g_m will be more. If the thickness is small of the silicon dioxide gate voltage you will have more control and obviously, g_m will be large. Similarly, the length is small more current will flow. So, g_m will be increase. **So,**

(Refer Slide Time: 58:24)

Summary

➤ **Electrical Characteristics of MOS transistors**

$I_{ds} = 0$ for $V_{gs} < V_t$

$I_{ds}(\text{lin}) = \frac{\mu_n C_{ox} W}{2 L} (2(V_{gs} - V_t)V_{ds} - V_{ds}^2)$ for $V_{gs} \geq V_t$ and $V_{ds} < V_{gs} - V_t$

$I_{ds}(\text{sat}) = \frac{\mu_n C_{ox} W}{2 L} (V_{gs} - V_t)^2$ for $V_{gs} \geq V_t$ and $V_{ds} \geq V_{gs} - V_t$

➤ **Threshold Voltage** $V_t = V_{t0} + \gamma (\sqrt{|-2\phi_b + V_{sb}|} - \sqrt{|2\phi_b|})$

➤ **Body Effect**

➤ **Channel Length Modulation** $I_{ds}(\text{sat}) = \frac{\mu_n C_{ox}}{2} \cdot \frac{W_n}{L_n} (V_{gs} - V_{t0})^2 (1 + \lambda V_{ds})$

➤ **Transistor Transconductance** $g_m = \frac{\mu_n \epsilon_{ox} \epsilon_0 W}{D L} (V_{gs} - V_t)$

NPTEL
Ajit Pal
IIT Kharagpur 23

So, here we can summarize. We have discussed about three regions of the operation of MOS transistor. We have derived the expression of threshold voltage and explained what is known as body effect and we have also considered channel-length modulation phenomenon and then explain what is known as transistor trans-conductance. So, with this, let us stop here. In the next lecture we shall discuss how a transistor can be used as a switch. Thank you.