Fundamentals of Semiconductor Devices Prof. Digbijoy N. Nath Centre for Nano Science and Engineering Indian Institute of Science, Bangalore

Lecture - 35 Gradual Channel Approximation: Derivation of I-V characteristics

Welcome back guys; if you recall in the last lecture, we had introduced MOS transistor, MOSFET. We qualitatively discussed about the working of the MOS transistor, how the channel gets pinched-off and how the current saturates, why it behaves like a transistor and how it acts like a switch, its extension of the MOS a capacitor with source and drain contacts.

And we are discussing about n channel transistor on a p-types substrate, but the p channel transistor and n-types substrate can also be discussed with appropriate changes in the polarities anyways. So, I have discussed you know, I have told you about the shape of the I_D -V_D curves, the plots, we call them the output characteristics.

And I told them that are initially are linear, and then the saturate. So, today we shall basically derive the simple mathematical expression to you know to summarise those I_{D} - V_{D} characteristics. And many other things it will take help of some small mathematics a simple mathematics, most of these are borrowed from the MOS capacitor expressions only. So, you will see the familiarity there.

And let us part already you know a lab, because now in MOS transistor and please remember that MOS transistor is a very vast topic, there are many advanced topics also associated in MOS transistor, it will be huge. So, in this course we will restrict ourselves to some simple concepts ok. We will not go to very advanced concepts. So, we shall try to list down what are the important topics in MOS transistor MOSFET that we are going to cover in the next few lectures ok. So, we will come to the slide, and I will show you what are the topics that we will go through ok.

(Refer Slide Time: 02:07)

1.	Current-voltage relations (gradual channel approximation)	
2.	Substrate-bias & threshold voltage	
3.	Sub-threshold conduction & sub-threshold slope	
4.	Short-channel effects	
	i) scaling	
	ii) charge-sharing between source-drain	
	iii) DIBL/punch-through	
	iv) channel length modulation	

So, let us come to the board here again. So, these are let me come to the highlighter here, the laser point is here. So, these are essentially the topics four topics that we will learn in MOS transistor thing. First thing is a current-voltage relation.

 $\mathsf{NOSFET}-\mathsf{basic working}$

(Refer Slide Time: 02:22)

And by current-voltage relation I mean this particular relation, this particular currentvoltage relation, we have already discussed the qualitative aspects, we will do the quantitative aspect. We will try to derive a mathematical expression for this particular I_D - V_D that is called current-voltage relation. We will derive a mathematical expression, which is based on something call gradual channel approximation - GCA gradual channel approximation. The second thing we will learn is substrate-bias and the effect of substrate-bias and threshold voltage.



(Refer Slide Time: 02:55)

What I mean by substrate-bias is that you see this transistor here, there is a substrate-bias option that is there it is a contact is a terminal that is connected to the body, this is called the body of the device. So, this is a p-substrate. So, even if a terminal coming out of the p-substrate, typically this is shorted to the source and grounded ok. Typically, this is grounded please remember that, so typically this is grounded, what happens is that, typically this is grounded you know typically this is call V_{BS} , this is a substrate voltage is typically grounded it is a 0 ok. This is essentially shorted here and then is grounded.

But it may not be grounded also. If it is not grounded, what is the effect of this threshold of this you know this substrate-bias, this is call substrate-bias right. If you do not ground it, if you in fact apply some voltage to here, what will be the effect of substrate-bias on the threshold voltage; how will the threshold voltage shift? So that part also we have to we will take a look that is a second point here ok, the very important point, because you will see that you will be able to get more flexibility using substrate-bias.

So, let me come back to the point again. So, second thing is the substrate-bias and the effect of substrate-bias and threshold voltage. The third point is very critical, and it is call sub-threshold conduction and sub-threshold slope. What it means is that before you turn

on the channel completely when you what I mean is that before you reach strong inversion, there is week inversion. And there is a small amount of conduction that you cannot actually observe in a linear scale that is call sub-threshold conduction it happens when there is weak inversion before strong inversion is happen.

So, you know I_D you know V_G characteristics basically when your gate voltage is slightly below your threshold voltage, you will have a small conduction that is called sub-threshold conduction. It is very important, because this will define your on-off ratio and how steep your slope is will tell you how you know how sharply, you can turn on the on the device and off the device ok.

And the fourth and the final thing is that we will discuss some short-channel effects. This become prominent, when your device becomes very small, which means your channel length becomes very small ok, by very small I mean less than 1 micron typically, but it can be is an very high scale of 100 nanometre and so on. So, short-channel effects have many things out of which we will discuss four. One is the scaling aspect; when you scale down the device, how does it affect the device performance. When I say scale down, you shrink down all the device dimensions.

Remember short-channel effects come, because the devices are basically made smaller and smaller which means they are being scale down. And this is following Moore's law that you know every 18 months your devices number will basically almost double, so there will be more than density of your devices that are per unit area, a device will become smaller and that is effect on the short-channel effects.

You know something called charge-sharing between the source and the drain when the channel the source and the drain are very far apart. What I mean is that, if this is source, this is drain, this distance between the source and drain, this is the channel, this is the channel ok, it is much larger then you do not care about it that. When this channel becomes very small, the source and drain depletion region approach each other, and then there will be some charge-sharing will affect the on and off state kind of you know device the threshold voltage and other things.

And DIBL is drain-induced barrier lowering, punch-through this is basically again a phenomenon that happens when your source and drain comes very close to each other, because your channel is very, very small or scale down ok. In the final I think we will

study is the channel length modulation is essentially not something very important compared to DIBL, but we will still go through it nevertheless. So, the first thing that we will do today in this lecture is to derive the current-voltage characteristics using the gradual channel approximation ok.

(Refer Slide Time: 06:42)



What does the gradual channel approximation say, so I told you at the beginning of last lecture that now it is a two-dimensional picture where this your current is going to your, your, because applying a voltage here, you will have a potential variation along this direction, which will assumed to be linear which means along this y-direction there will be some voltage that is changing, because of this. But the band bending is happening in this direction if you remember from the MOS capacitor I told you this is the x direction for example. But, so I can nomenclatures can vary for example, I might call this as x to make it simpler, and this is y, but that does not matter.

So, in this direction, in this perpendicular direction, you will have a band bending here, you will have a channel that we formed and the band bending over here ok. And in this direction, you will have a potential that we applying here. So, there will be gradient of the potential. Now, the gradual channel approximation will tell you that the variation of potential in this direction matters actually in this direction. The variation of potential along this direction can be ignored ok. The variation in this direction can be ignored,

which means we can assume that the channel that is formed is formed at an infinitesimally thin like a layer here.

So, there is no variation along this direction and the potential varying along this direction of the channel, where there the current is flowing is what matters, the potential varying over here does not matter. This is the simple you know it is gradual channel approximation that will hold true for many of the long channel cases, and we can get expressions that are fairly accurate. Although, it will break down when your device becomes shorter ok, and when your device pinches off for example, ok.

So, what I will use is that if I take a small slice of material over here of thickness dx ok, and a this is the depletion by the way. If I take a small slice of thickness here dx, then the differential resistance associated with that will be given by dx by the conductivity you know way ok. So, this is the charge $Q_n(x)$ is the; what is Q_n ? Q_n is actually the inverted electron charge, which is a function of position x, this is I am calling as x now, this I am calling as x ok. As a function of x, as a function of x, your inversion charge will change, because I showed you a pinch off you know eventually here your charge will become 0, it will be your triangle profile.

So, $Q_n(x)$ is essentially the electron charge as a function of x. Then this is a mobility, I will assume that mobility is constant everywhere although it may not be the case every, every time. And Z is the width of the device so that does not matter so much is basically normalizing the current. So, essentially the differential resistance associated with this particular differential element dx is the dx that element here divided by the conductivity, which is the resistance here ok.

So, what are we doing, why are we doing this? We are trying to do this to essentially we have the current versus the voltage that we applying. Now, let us take a minute and look very carefully. The total charge, the total charge in the semiconductor here, the total charge in the semiconductor here is call Q(s) Q semiconductor charge, which will also vary as a function of x. The total charge in the semiconductor is the inversion charge here plus the depletion charge here.

There will be some depletion charge, you also applying a voltage here. So, the depletion you see the depletion width gradually increases to this side. This side the depletion width is more than this side, why, because I am applying a positive voltage, so the depletion is

wider this side and then this side ok. So, the depletion charge also we will say function of x. So, depletion charge as a function of x plus the channel charge as a function of x, we will give you the total semiconductor charge as a function of x ok. Till now, it is clear? Ok.



(Refer Slide Time: 10:22)

Now, if you recall this expression from your MOS capacitor, the gate voltage that you apply part of it, part of it drops on the semiconductor, which is this; part of it sorry part of it drops on the oxide, which is this. Part of it drops on the semiconductor, which is the band bending, of course, at strong inversion this becomes you know two times that the φ_s (phi s) which is the $E_A - E_F$, but in, in MOSFET we will only talk about you know the channel is inverted is already, because the channel is to exist. If the channel does not exist, then there is no point discussing this.

So, essentially this will be you know like $2\varphi_F$ already. So, this is the voltage and the oxide, this is the voltage drop in the semiconductor. This is the flat band voltage that you have to account for because your metal and semiconductor work functions are different, your oxide charges might be there. So, this is the threshold voltage expression from the MOS capacitor which we are very familiar with.

Now, you remember that the total charge in the semiconductor over here, this expression over here is equal to the charge on the inversion channel, and the charge on the depletion region. Now, you juggle around here you put this value over here; you put there, there and you do some mathematics on there. You will get an expression for $Q_n(x)$, which is the charge mobile carrier's, charge of the mobile electrons that is the inversion charge, this is the inversion charge. Remember the current-voltage characteristics that I showed you here, this current is carried exclusively by the electrons that are formed here ok.

So, the inversion charge that has formed here, the electrons that have form the inversion channel here are alone responsible for the current this. So, we are interested in the electron charge not in the depletion charge ok that is why you are looking at the electrons charge here. So, this is the electron charge there, and that electron charge is the function of x, because I told you along the channel the potential is going to vary that is given by this expression this V_{FB} , this is should be subscript, sorry this should be V_{FB} ok.

So, this is given by you know the this is basically the capacitance of the oxide, which is ϵ/d (epsilon by d). This is the band bending on the surface minus this is the gate voltage minus the flat band voltage the whole thing is same product is here, the minus the depletion charge. Now, what you do? Remember that this band bending, this band bending is essentially φ_F , which is the $E_A - E_F$. This band bending that is happening on the semiconductor, this band bending is essentially 2 times the band two times $E_A - E_F$ plus there is a V_x here. Why? Because if you look at this, if only MOS capacitor was there then yes, the band bending and this interface is equal to $2\varphi_F$ that you know, but because there is a drain voltage that you are applying, there is a voltage that is changing here that voltage must be added here. At any point x, the voltage that you applying here will become will be V(x) eventually at this point voltage will be V_D; at this point voltage will be 0.

So, at any point here voltage will be V_x ; that voltage V_x has to be added to the total band bending here, because applying, in MOS capacitor this was not there know now it is there. So, you have to apply that voltage actually, so that is why would have to apply that voltage V_x to the band bending it is happening. And the depletion charge, if you recall the depletion charge that is formed here, this depletion charge is if you recall the MOS capacitor lecture, it is square root of the thing you know this expression that is there except that instead of 2 phi F, which was there in the MOS capacitor in MOS capacitor you only have this much, these much. Here instead of 2 φ_F , you have to have $2\varphi_F$ plus V_x , because this V_x is also being added to $2\varphi_F$ along the channel right. This, this potential along the channel V_x is added to the band bending in this direction right, so that is why it comes here as bracket missing here, sorry about that, so this will come here. So, this expression of the depletion charge, you put here, and this expression you put here, you do some juggle around.

(Refer Slide Time: 14:15)



And then you get this expression for $Q_n(x)$, which is the mobile electron charge in the channel. And this mobile electron channel charge is alone responsible for $I_D V_D$. So, we are only interested in this charge. This is given by this huge looking expression nothing very fancy, nothing very fearful. Now, to get your current-voltage, you basically take this expression into account that the potential that is dropping across this differential small length dx is given by that this is you know I_D this is actually this shall it be an I superscript D. So, let me let me correct this again right here, so that you know you do not have this problem here. So, this looks like that this looks like that.

So, you see your sorry you have a, I will take the marker here the laser pointer here ok. So, essentially the voltage that is dropping across, the voltage that is dropping across your differential element dx here will be equal to dV which is equal to the current that is flowing the current has to be continuous everywhere I_D times the resistance here right. And dR if you recall from this previous expression over here, this is the dR right. So, you substitute that here you substitute that here right. And this Q_n essentially you take this Q n; this Q_n you put it here ok.

And then your I_D this, this expression if you put here then I_D times, this expression dx. So, this can be taken the other side, so, it will become like that. Now, integrated over this entire channel lying from 0 to L, you will get the expression this expression ok.



(Refer Slide Time: 15:52)

This is the expression for your final current as a function of voltage, V_D is there, V_G is there, of course, there will be something call theta. The theta is called the body factor, and this is again N_A . So, this N_A is the doping concentration in a substrate. This is given by you know this particular quantity. This theta can be actually neglected because if your doping is light in the substrate, N_A is the doping in the substrate, if doping is like in the substrate. And your oxide is thin because this is epsilon by this epsilon will be denominated be will be here. So, D is here is a thin oxide is thin which means that D is small, this is large and then if the doping is small, this is small, then this quantity will be very small and this part sometimes can be simplified ok. This part can be simplified, so that we will do it now.

$$I_D = \frac{Z\mu_n C_{ox}}{2L} (V_G - V_{th}) V_D$$

So you see initially it is a linear, it will vary after that there will be non-linear region. And once the channel pinches off which means towards the gate end of the drain or drain end of the gate, you know gate is like with respect to the drain end of the gate is pinches off. So, the, the number of electrons arriving at that point will be important of the fast how fast, they have swept away if we recall in the last lecture. So, there will the so the current will saturate which is call the body factor.

And in the linear regime, your gate voltage actually the drain, the drain voltage that you are applying is less than this particular expression, because you are applying a very small drain voltage here. And in that case, you can simplify this expression. If your, if your drain voltage is very small, then you can do some Taylor series expansion, and you can do some simplification.

What you will get, this Z is by the way like the width of the device some textbooks call it W also, this is the width of the device periphery of the device. So, that current is flowing in this regime the linear regime in this regime specially not even here, because here it is approaching nonlinearity. Over here the current that is flowing is the width of the device by the length of the device, this is the channel length L, mobility C_{ox} mobility, I am assuming the same everywhere. C_{ox} is basically the oxide capacitance.

 V_G-V_t , this quantity V_G-V_t is called the gate overdrive ok, always remember that V_G-V_t is called the gate overdrive. So, V_G-V_t so of course that gate voltage has to be more than threshold voltage. If your gate voltage is not is less than threshold voltage, then everything here will fail, because the channel will not be formed. The premise, the foundation is that gate voltage is more than threshold voltage, only then your channel is formed right. So, gate voltage is more than threshold voltage by some amount, so that amount is this time the drain voltage.

$$I_{Dsat} = \frac{Z\mu_n C_{ox}}{2L} (V_G - V_{th})^2$$

So, you see this current varies linearly the drain voltage. So, if the drain voltage is in this regime, the current is varying linearly that is what is happening. Also the current varies linearly gate voltage. So, if you stepped on the gate, these are different gates. You if you take one fix drain voltage along there you will see there the current is also changing

along with the gate voltage ok. So, this quantity is very useful in understanding the linear regime.

Then we will come to an the highly non-linear I am not deriving the expression. So, here essentially if your drain voltage is much higher such that the current saturates there, then what you will do ok. Then what will you do?

(Refer Slide Time: 18:52)

I-V relation: Saturation	CENSE
 □ GCA becomes invalid when the channel gets pinched-off near the drain □ I_D saturates at V_{D,sat} □ Put Q_n(x) = 0 and V(x) = V_{D,sat} □ Assume that voltage drop across the oxide is very low compared to 2φ_F (i.e. we can neglect the -Q_g/C_{ox} term in the V_{th} expression) □ Θ <<1 (Why?) 	
$V_{D,sat} = VG - \left[V_{FB} + 2\phi_F + \theta\sqrt{2\phi_F}\right] \approx V_G - V_{th}$ $I_{D,sat} = \frac{Z \mu_n C_{ox}}{2L} \left(V_G - V_{th}\right)^2$	25

So, in the once you have the channel becomes pinched-off which means I will come back here, once your channel becomes pinched off here, it was actually highlighted in this diagram, yeah. So, once your channel becomes pinched off here, this pinched off here know I told you that gate becomes negative with respect to the drain. So, this particular region becomes pinched-off. Once the gate becomes pinched-off here, then your gradual channel approximation no longer hold true ok. Your gradual channel approximation will become invalid, when the channel gets pinched-off near the drain.

Now, I_D will saturates at $V_{D,sat}$, which means this is the $V_{D,sat}$ right. This is the locus of $V_{D,sat}$ by therefore, different V_G your $V_{D,sat}$ also goes, for example, discuss this is the $V_{D,sat}$. So, this point will be $V_{D,sat}$. So, this curve, this is the $V_{D,sat}$ So, this point is a $V_{D,sat}$. For this point, this is the $V_{D,sat}$. So, therefore, location of $V_{D,sat}$ is basically $V_{D,sat}$ is the voltage at which the current saturates. The current at the voltage at which the current saturates is called $V_{D,sat}$, and the $V_{D,sat}$ depends this is $V_{D,sat}$ here, this is $V_{D,sat}$ for this. So,

 $V_{D,sat}$ itself moves position as your gate bias keeps increasing. So, your $V_{D,sat}$ this point right or this point essentially at that point the current saturates.

Now, what happens at the current saturates? I told you that your gate will become pinched-off, which means the channel charge Q_n . Remember Q_n is the channel charge as a function of x at x equal to L, which means at x equal to L means at this point L, at this point x equal to L, when your channel becomes pinched-off the current saturates, which means you get this kind of I D this kind of, this kind of profile here a saturated. So, at x equal to L at this point, your total your not, total your mobile electron charge will become 0 when your channel pinches off here.

So, what you do is that you put mobile electron charge x equal to this total electron mobile transition, the, the, the that the inversion charge as 0 at x equal to L. And at that point, you will put V equal to V_D , sat where a current saturates ok. Assume that a voltage drop oxide is very low compared to the band bending that is happening. So, we can neglect this term in the V_{th} expression. And this <u>theta</u> the body term can be neglected, because if the oxide is thin and the doping is low, then this quantity will become very small.

So, with this condition, you can get and simplify the expression of $V_{D,sat}$ as this ok, we will do some juggling around, you can get the V_D . You basically put here in this expression this expression. In this expression, you put this as 0 x equal to L and this will basically will $V_{D,sat}$, then this 0 equal to this, this, this is $V_{D,sat}$, this is $V_{D,sat}$. So, if you do some juggling around, what you will get is that this expression, you will get. This is a $V_{D,sat}$ expression you will get, which is almost equal to if you do some Taylor series expansion, you will get V_G-V_t , actually this is V_t threshold voltage by the way. So, you will you will get V_G-V_t . Assuming that the voltage drop on the oxide is negligible. So, $V_G-V=$ is basically the gate overdrive.

And you plug that in the expression for this particular expression, we plug it here. What you will get? You do some simplification; you will get this. So, the saturation current which is this current please remember this, this is the saturation current this, this, this not this this is the saturation current. So, saturation current depends as the W the width of the device, mobility oxide capacitance by 2 times the length, this is the gate overdrive whole square V_{G} - V_{t} whole square. It does not depend on drain voltage. You

see this current does not depend on the drain voltage right. So, the expression also shows you that the drain current saturation does not depend on the drain voltage, but depends on the square of the gate voltage square of the gate voltage, so that is your basic expression for IV in saturation.

Now, remember these are simplifications that are done with some assumptions like for example the voltage dropping in the oxide is very negligible. The band bending is much higher in the voltage and oxide, θ (theta) is less than 1 which means the doping is light. The oxide is thin, but if all these conditions are not mapped, for example, the drop voltage in the oxide cannot be neglected, you know the, the doping is not very light, the body factor cannot be neglected, because the oxide has thicker. In all these cases, you cannot have this simplified expressions like this for example, you have to use this particular expression and plug exactly the different values.

On this you have to plug the different values, and you can get the I_D V_D ok. This is the expression, this is a mega expression from which we are deriving things from which we are deriving things. So, please be advised that this is an important expression, and all the simplified things are actually done from here. In some textbooks though, this particular expression that you see here in the linear regime, which is this; some text books also have a term here minus $V_D/2$, what I am trying to say is that let me actually, it depends on the simplification that you are doing actually it depends on the simplification that you are doing actually it depends on the simplification that you are doing.

So, in for example, if you do not in this particular expression, in this particular expression some of the terms you might expand in some other way, some of the terms you might neglect. So, different textbook do some different kind of simplifications. So, in some text books, this expression the linear regime here, this expression is slightly modified ok. Just keep in mind this is I_D equal to Z by Lµ and C_{ox} . Here what they do is that (V_G - V_{th}) - $V_D/2$ into this V_D ok. So, this expression is exactly the same except that there is a $V_D/2$ terms here minus $V_D/2$ term here that can give you a little bit more accuracy. And this is more important, when you go near this part, when you are approaching the non-linear regime ok, close to 0 this is fine ok. So, please remember that ok. So, some of the some of the text books also do that ok.

(Refer Slide Time: 25:17)



So, so, so just nothing if we recap once again, there is nothing actually very scary or fanciful about this. What we did is that we took a differential element on the channel, I told that that the resistance associated with this differential element is basically given by this expression. This is the charge in the channel; this is the mobility; this is the width. Remember the current is exclusively carried by the mobile charge $Q_n(x)$. So, we are interested in this; we are not interested in the depletion charge we are not interested in. We are only interested in the channel charge the Q_n ok. I told you that the total semiconductors charge is equal to and this is a function of x by the way, because along this direction x is varying and because you are applying potential here at any point a potential is the x. So, the x is varying along this direction ok.

But all semiconductor charge is total depletion charge, and the total charge in the inversion layer, we are only interested in this part. So, what we go ahead and do is that we recall the threshold voltage expression, we recall this expression, which I have written just now this, this, this expression just now. And we substitute one from the in the other, and we get an expression for the total this is the total channel charge, this is the total channel inversion channel charge ok, the strong inversion.

Assume the strong version is already, there is no question if there is no inversion there. And this particular band bending that happens you know if you remember band bending that happens, this band bending. This band bending that happens whatever, that band bending was actually $2\phi_F$ at strong inversion, but because there is also linearly varying charge now this V(x) we applying a drain voltage, this drain voltage does not a MOS capacitor, but now it is there. So, it at they are also any point x that also will become you know plus V(x) right and so that is why essentially this is the this quantity. Similarly, the depletion charge was actually this quantity.

But now, because there is a V(x) there, you have to also at that here. So, it because like that. You plug in this and this in this expression, and then you get this particular mega expression. Now, it does not you do a small integration to find that the current that current you will basically with this expression this currents expression. You can do some Taylor series expansion some simplification to basically make it simpler looking. Anyways with some simplification you can get this expression for linear. I told you in some text books, there is a V_D minus 2 term here like this. So, you have a linear equation and a expression for saturation current that gives you this expression ok.

So, with this we conclude the I-V relationship. And in next class we will study substratebias and the effect of substrate-bias and the threshold voltage ok. So, we moving good at good speed. So, let us we will end the class here. We have discussed basically the difference between the difference between the linear and the saturation region. We have also derive the expression for the current for the linear and the saturation regime, these are approximation up to the main expression that we have derived for the current for various numerical for very simple problems. We can try to solve some problems later in the in the course. We can use this expression to find out the linear and the saturation current density ok.

So, the current saturation MOSFET that is very important saturates from the gate and you know pinches-off because you go to higher drain voltage that is what has happened. So, in the next class, we will take into account the substrate effect, the effect of the substrate-bias how it affects the threshold voltage, and then we will go to sub-threshold, we will go to short-channel device, and we will finish a MOS transistors, and next may be 2-3 or 3-4 lectures ok.

Thank you.