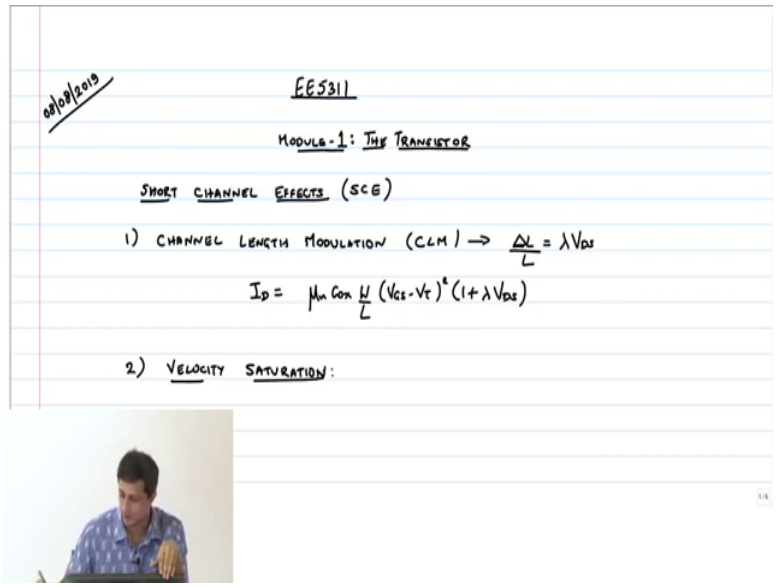
Digital IC Design
Prof. Janakiraman Viraraghavan
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 07
Velocity Saturation and Level - 1 SPICE Model

So, in the last class we introduced the concept of Short Channel Effects economies, SCE ok.

(Refer Slide Time: 00:23)

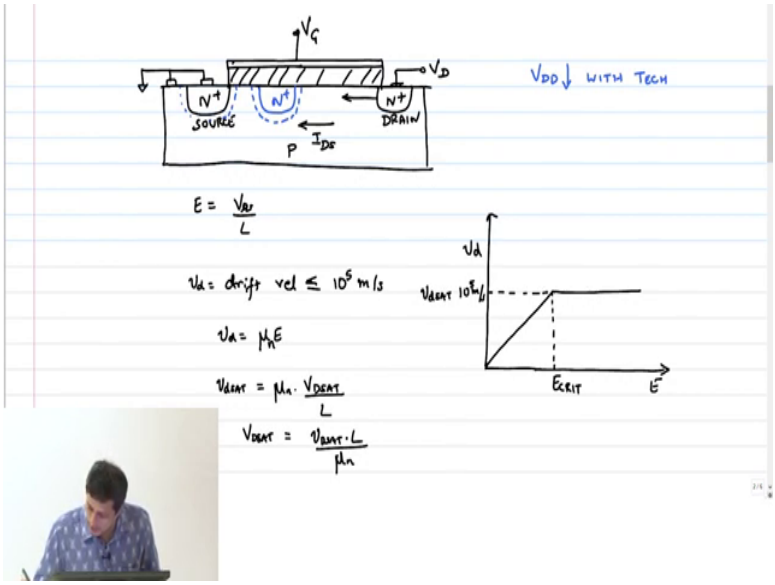


And the first thing that we spoke about was Channel Length Modulation, CLM and what was this if I have to summarize this effect the delta L by L is going to be modeled as lambda times V DS ok.

So, therefore, the equation the drain current equation gets modified as in saturation mu N Cox W by L right into V GS minus V T whole square into 1 plus lambda V DS right. This is what we said and primarily it is in saturation that you will see these effects right because V DS has

to be large ok. So, moving on with the short channel effects; in this class we will discuss what is known as velocity saturation.

(Refer Slide Time: 02:09)



So, as we discussed in last class; you have a drain I mean a source and a drain in a long channel device, they were placed very far apart upside here N plus N plus P right. So, if you applied drain potential V D here; of course, with the source and the body both grounded and you applied a V G right.

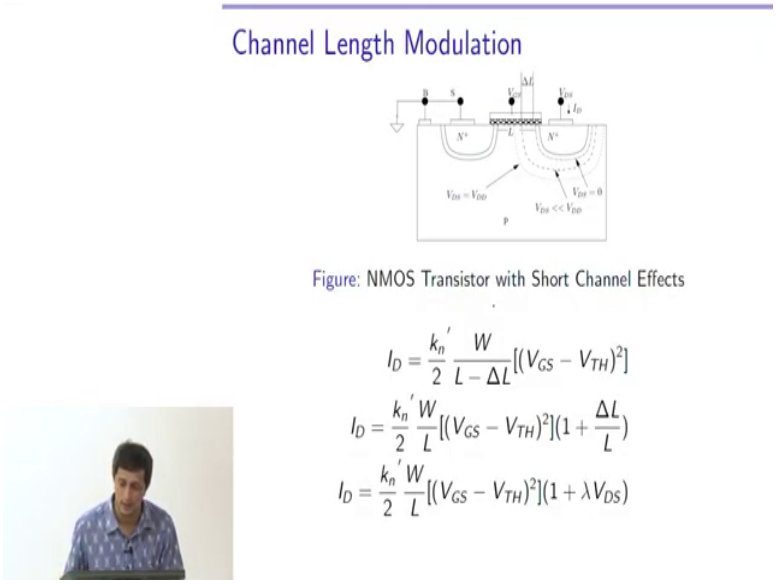
Then what would happen is the source right, this is the source drain; the source will give you some electrons and will get accelerated or it will drift with some velocity through the channel right and then get to the drain end and that is how current was generated and he derived the equation for the current right. The source basically sources the charge carrier; the drain corrects them that is why the name source and drain ok.

So, in an N MOS transistor the current because it is because of electrons; current flows from drain to source opposite of the flow of charges right. So, the I DS is basically like this and these electrons would get accelerated or would attain a larger drift velocity not accelerated in this lattice actually; it attains a larger drift velocity, if you apply a larger voltage on the drain end ok. So, basically that simply related to the electric field E and the max is V DS by L. If you just consider the two ends and the electric field is V DS by L.

Now, what happens is when you start bringing this drain very close and shrinking this transistor right and making it look like this. Of course, now we have the problem of these depletion regions eating into the channel and all that ok. Apart from that, what we had to do was when we brought these drain and drain very close to the source; I cannot continue to apply the same drain voltage max, drain voltage that I would have applied for a long channel device. Why? Because the device will break down; there is a breakdown electric field beyond which that all those properties do not hold right of a semi conductor and all that.

So, therefore, we had to lower the supply voltage also as we scaled these transistors and made them much much much smaller right. So, therefore, my V DD also started dropping with technology ok. Now, in spite of dropping the V DD; it turns out that there is only a maximum drift velocity that these electrons can attain in this lattice ok. So, there is a drift velocity V D is about 10 power 5; I think centimeter per second, these are not mistaken let me check that number; yeah it is about a meter per second ok.

(Refer Slide Time: 05:54)



The velocity cannot do a higher effect why is this? Of course, because there are atoms, there are you know a lot of things to collide with and all that stuff. So, it cannot accelerate beyond this ok; that if we have to just take this fact as a name.

Now, we have to incorporate this into our model ok, but look at our current equation; where is the drift velocity in this? We do not have a drift velocity term in any of these equations. So, there is two mobility and all that stuff right. So, somehow we have to now bring in this effect of maximum drift velocity into our current voltage equation.

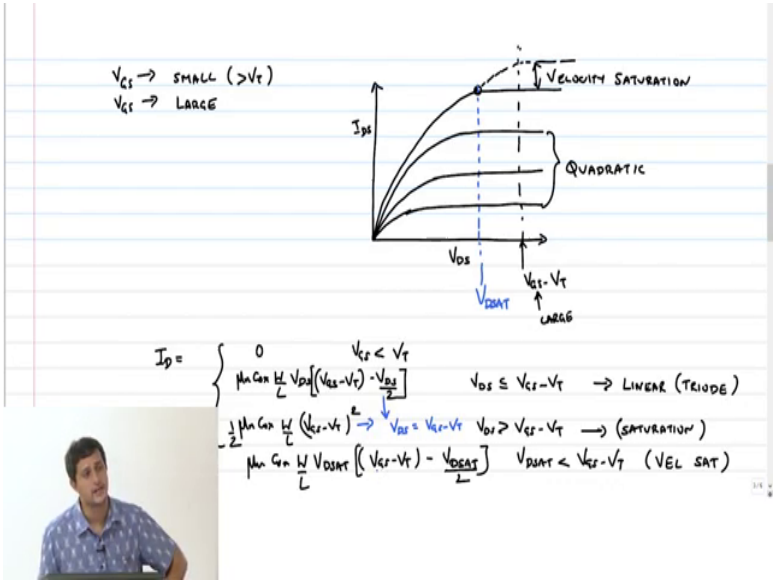
So, we have to translate this maximum drift velocity to a maximum voltage that is permissible right. So, let us look at that drift velocity is mu times E ok. So, if there was no; if

the channel was long enough right, then what how would this equation look like? V d versus E; it would basically be a straight line right, now what are we saying? This saturates right.

So, this is somewhere and power 5 meter per second ok, so; that means, there is a critical electric field beyond which this saturation takes place ok. And therefore, I will write this V d SAT; let me call this as V d SAT, the drift velocity saturation value is going to be mu n of course, into V D SAT by L ok. Therefore, V D SAT is equal to V D SAT velocity drift velocity saturation value into L divided by mu L.

So, now what are we done? We have taken the drift velocity problem and translated to a voltage equivalent value right; which means to say that if I apply a drain voltage that exceeds this V D SAT value; then I cannot get any further acceleration or any further increase in current right. Again not acceleration sorry; it is a further drift velocity which means that my current will saturate. So, now how do I incorporate this into the equation? So, we have applied a certain V GS ok; then we start increasing the V DS ok, what will happen? Sorry.

(Refer Slide Time: 09:18)



So, let us consider; let us consider two cases ok, V DS is small and then V DS is large ah; no I am considering V GS sorry large; I am going to sweep V DS ok. So, I am going to increase my V DS here and look at I DS.

So, for the small value of V GS; what will happen initially? It will be in; of course, no assume that V DS is slightly about threshold voltage that is what I mean ok. Small, but greater than V T good point. So, when I now sweep my V DS; what region will it start operating? Yeah triode region or linear region; so it will it will go like this right.

Now, where is your V D SAT right? Let us look at this, let us say my V D SAT is here; this is the critical V D SAT voltage beyond this if you cross then you will saturate correct. So, therefore if V DS is small enough then it can achieve the saturation current much earlier

remember that because at the drain end my V GS minus V T minus V should be greater than 0.

So, if V GS is small enough and the original saturation happens then the current will just saturate here right. Now, again if I take the next V GS; again which is quite small, then this will go like this. Further, if I go it will go like this and this behavior can be quadratic because we are still in the old saturation region only correct.

Now, if I apply a very large V GS; then what will happen is this point at which it saturates is going to be where? This will be; this is for a large V GS, this is V GS minus V T and assume this is large; then the current will actually go up like this and then saturates here. Normally, it should have gone on in the linear region until here and then saturated here. This effect of lowering the current is due to velocity saturation ok.

So, the point I am trying to make is; it does not mean that a transistor cannot go into the regular saturation region, if velocity saturation exists in the technology. It can; it depends on the V GS value because if that saturates first right then it cannot go it cannot even enter this velocity saturation region right. Basically, you will be drift velocity is small enough then the transistor behave is; behaves in the old fashion that we were talking about ok.

So, now the question is for a long channel device did this velocity saturation effect not exist at all? Yeah, exactly it is not that it does not exist; of course, it exists there also it just that the V DS I have to apply is so large that it is out of the range in which we are going to operate the device and therefore, we do not worry about it there that is all. There also if you increase the V DS beyond a point you will hit this 10 power 5 meter per second and then it will saturate. But that is not even in the interesting region of operation you are well above we really there and therefore, we do not worry about it in a long channel device ok.

Any questions here ok? So, how does this current look like? Now, that is just the intuitive picture that I gave you right. So, how do we incorporate this into our equation? So, the drain current is going to be the following ok. Of course, it is 0 if V GS is less than V T; now if it is

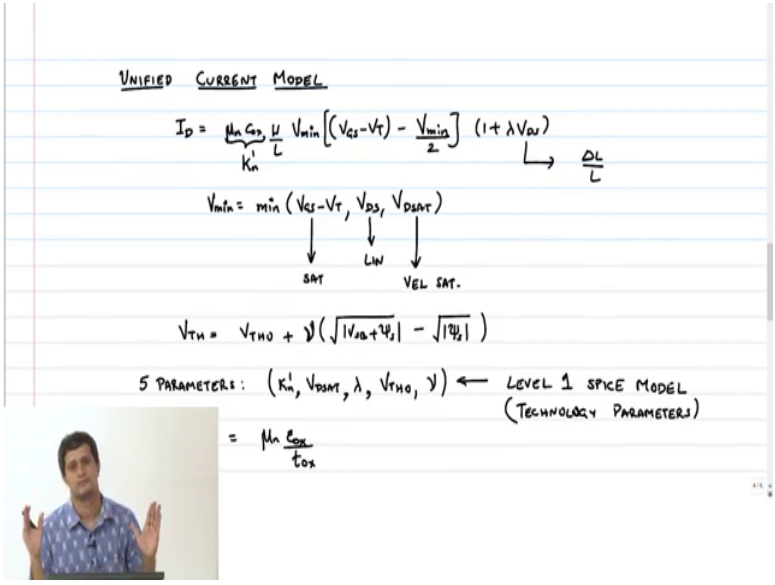
in linear region right mu N Cox W by L into V DS into V GS minus V T minus V DS by 2 ok.

I am not incorporating the channel length modulation, we can; we can add that effect later ok. If V DS is less than or equal to V GS minus V T and earlier if it had to go into saturation region; it would be W L half of V GS minus V T d whole squared if V DS was greater than V GS minus V T.

Now, somewhere in between these two regions; where you are going from linear to saturation; if this saturation drain voltage V D SAT occurs before as I am sweeping my drain voltage V DS, then the current will saturate to that value and that value is just given by; in the linear equation look at this in the linear equation you substitute V D SAT and that is the value of your current; it cannot go beyond that. So, this will become mu n; Cox, W by L; V D SAT into V GS minus V T minus V D SAT squared by 2; oh sorry V D SAT by 2; yeah I brought that out.

So, now what is the condition under which this holds? This is the condition is V D SAT is less than V GS minus V T; if that happens ok. So, now what I will do is we have to many regions of operation; linear region, saturation region, velocity saturation region and by the way; let me just put it down. This is linear right or triode or resistive region right it behaves like a voltage control resistor in that region; this is a saturation region and this region is called velocity saturation.

(Refer Slide Time: 17:50)



So, now I have three different regions of operation and this can be quite confusing and therefore, we have what is known as a unified current model; again this is an abstraction that is that works very well for this course ok. And this is more than sufficient for us because the regions in of operation are going to be sort of in the extremes. V GS will be 0 or V DD you know this kind of thing ok; we are not too worried about the details in between.

Therefore, I will write my current ok; you can just look at these equations here. This equation let me put that in blue is basically this equation with V DS equals and this equation is what? V DS is equal to V D SAT.

So, what is really happening? It is the min of three different values that is appearing in that term there right. The V DS term is getting replaced by either V DS or V GS minus V T or V D SAT; whichever is the lowest. Because you look at various regions; here V DS is less than

V GS minus V T first; so, then it becomes the linear region. Here V DS is exceeded V GS minus V T therefore, V GS minus V T comes in.

Now, the second one is V D SAT appears even before V GS minus V T therefore, it saturates even earlier. Therefore, the natural way to write this is mu n; Cox W by L into V min into V GS minus V T minus V min by 2. And now, I will incorporate even this channel length modulation straightaway into 1 plus lambda V DS. Of course, this maybe if you ask me whether this happens even in the linear region for small values of V DS; it may not, but the point is because this is a model that we are building; we need to have continuity across regions otherwise, the simulator will not be able to simulate all this correctly.

Therefore, you put this in where the impact is not very large because if V DS is very small; lambda is already a very small parameter, 1 plus lambda V DS will approximately be 1. So, it will not affect your equation there, but it helped attain continuity across regions and that is why you just put it in for all regions of operation.

Now, what is V min? Is minimum of V GS minus V T correct comma V D S comma V D SAT. So, given a problem if you have to identify which region of operation it is in; you calculate V min which is the min of these three. If this turns out to be the answer, it means the transistor is in saturation region. If this turns out to be the answer, then it is in linear region and if this turns out to be the answer; it is in velocity saturation region ok. So, this really simplifies our analysis and mathematically abstracts the whole thing out very well and quite consistently also.

It deviates a little bit from reality because what we are saying is the point at which this velocity saturation happens is like a straight line; there is one V D SAT, but that can sort of change a little bit apparently ok. If you go to a device physics course, you will see that that is not fully true, but for a purposes of this course it is more than sufficient ok.

So, now I will rewrite this yeah; V DS because that delta L by L right; see where is this lambda V DS coming from? This is basically modeling the delta L by L term and that is purely determined by the reverse bias voltage on the drain side right that; only that modulates

this the depletion region at the drain end. And therefore, that can be only 1 plus lambda V DS; it has nothing to do V DS or we V D SAT yeah.

Now, again that you know sort of too many terms. So, I will combine this and write it as K n prime ok. Then I have this V min V DS whatever and I have my threshold voltage equation what is it? It is V T H naught plus gamma times root of mod V S B plus psi S minus root of mod psi S.

So, if you look at this current equation; I have 5 parameters ok, I have 5 parameters; not just the current and threshold voltage equation. One is K n prime, second is my V D SAT, third is lambda, fourth is my V T H naught and the fifth parameter is gamma ok.

These 5 parameters characterized uniquely a particular transistor in that technology. So, what do you do? You do not worry about solving device physics here, you go and fit these parameters through measurements in a particular technology, you fit these parameters and you specify this in a particular model file. Yesterday, if you remember in this in the tutorial you were shown a model file which had like 54 parameters or something.

This is a much simpler model than that; that is it incorporates many more effects and is far more sophisticated which is understandable because reality; it is far more sophisticated. But this is effectively what is known as the level 1 spice model; just 5 parameters and if I can if I define the current in this particular way and the threshold voltage in this particular way right, then I can obtain the current is the fab gives me these 5 parameters; of course, it is a very simplified equation.

Now, I my question is the following; I combined mu n into Cox all right and what is Cox by the way? Just to remind you it is epsilon oxide right by t ox; oxide thickness. Now, why did not I combined W by L into this equation; why did not I say K n prime into V min right into V GS minus V D that is the current equation. Why did I separate this mu n Cox alone as one parameter and kept W by L outside? Yeah, you can change or I want to differentiate between two things; one is a technology parameter, other is a design parameter.

So, these level 1 spice model; these are known as technology parameters which means that once I have shrunked my technology to a new node and I have defined the feature sizes and all that stuff; these values or these parameters are not dependent on the design parameters.

So, as a designer what; what is the freedom that you have to tweak? You can tweak W, you can tweak L in analog designs ok. In digital design of course, for high speed you have to keep it as close as possible; so you; it is effectively fixed right, but you can tweak W, you can tweak L and you can tweak the voltages.

So, therefore, these parameters should not be combined with the technology parameters and sort of confuse you. You should know exactly; when I say it is a technology parameter; it means it is fixed. You cannot propose a solution that says I will change the technology parameter; you have to propose a solution where you change the design parameters and achieve the goal that you want going forward.

So, that is why in this course; we will always separate the terms out as technology parameter and design parameters and bring out the dependence very clearly on the design parameters ok. For the oxide thickness for example, that is fixed for all transistors; once you scale to a new technology in 40 nanometer; it is about 1 nanometer. 40 nanometer technology; the oxide thickness is about 1 nanometer; for are all transistors including pmos transistors right. You have another variety called a T tox oxide device where the oxide thickness is slightly larger, but that is a separate model, but even there across N MOS and P MOS; the oxide thickness is fixed.

Again V D SAT; if you scale to a new technology then you get a new maximum drain voltage that you can apply before which the velocity will saturate in the velocity will saturate in the channel right. So, that again is given to you as a parameter; now the question is for what value of channel length are they giving you this?

Because I just told you that L can be varied correct in the analog designs I can change L to even 10 micron right in a forty nanometer technology is possible. So, does this V D SAT

apply there; so the answer is no it does not. When they say they give you a model file with these technology parameters; there is a region of operation where this is valid. If you try to take this model file and apply it for some random region of operation; the fab will not give you any guarantee that it is even expected to work.

These parameters have been fit very specifically to measurements made on devices in a range of L length, width and supply voltage conditions. So, the model file they will tell you right; in the model file yesterday you saw V DD equal to 0.8 volt. What if I apply V DD equal to 1.6 volt in that? The answers can be totally wrong right. So, these model files are actually specific to the application that you are doing. Here in digital design we are fixing our L because we do not we want maximum speed therefore, this V D SAT makes a lot of sense there ok.

So, what about the you know the lambdas ok; all of these parameters have some sort of a length dependence. Because if I take the length very far away; then this does not really affect it at all. Of course, sometimes they will fit these things so that there is a dependence on the L also ok; you can fit it to that kind of an equation as well right.

Now, the threshold voltage V T H naught ok; this again has no dependence on L; L and W as you can see in this previous equation or the equations that we derived earlier. It is independent of L and W to a first order and you get one particular number that the fab says this transistor has a V D H naught of 200, 250 millivolt right; it is independent of L and W ok.

But again it does not mean that; I can take L to 10 nanometer and expect it to give me the right threshold voltage. No, it is for that region of L and W where that is valid ok. Similarly, the gamma is fit to a particular value and these 5 technology parameters put together constitute what is known as the level 1 spice model ok.