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

Lecture - 37
Short Channel Effects in MOSFET

Welcome back. So, if you recall in the last lecture, we had introduced the concept of sub-threshold conduction. Before that we had gone through the derivation of the I_D V_D relation in MOSFET, and also explained the effect of substrate bias. I told you that with appropriate substrate bias which is always negative with respect to the source, you can manipulate or you can tailor the threshold voltage of the device according to your need, so that is an additional nob or additional you know factor that you can take into account. The effect of substrate bias on threshold voltage, it is a beautiful expression for that.

Simple expression that shows you that that threshold voltage shift goes roughly as the square roots of the you know that is for the substrate backside voltage, we are applying right. In the end of the lecture, last lecture, we have told I told you that there is sub-threshold conduction. Sub-threshold conduction is the conduction, when which happens when the device has not reached strong inversions, so weak inversions sort of a transport. And if you plot transfer characteristics which is I_D versus V_G, then on a linear scale, you will not see that, because you will on a linear scale, you will only see the current about the sub-threshold voltage, below the sub-threshold voltage, the current will seem like 0.

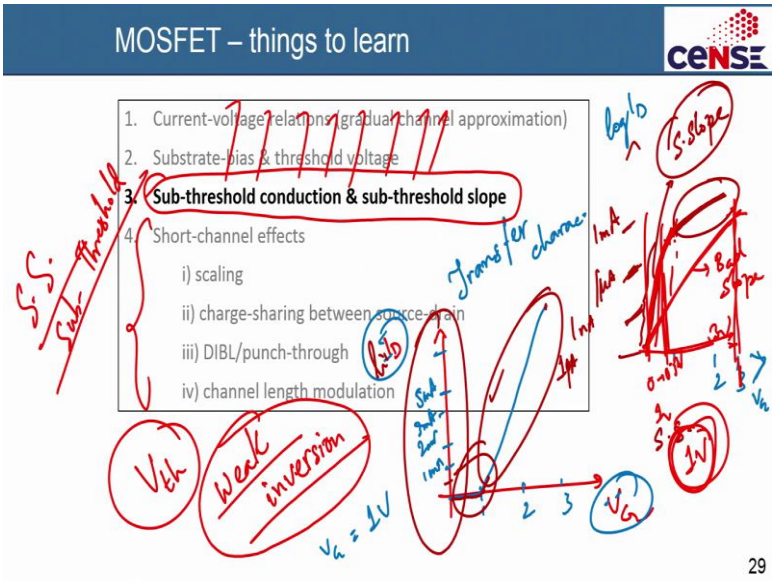
On a log scale, you will see that below the threshold. Actually, the current is rising the many orders of magnitude from like pico or nano amps to micro amp milliamps that range of current is not visible on a linear scale. So, on a log scale you can see that, so that sub-threshold slope is the slope of the I_D V_G there. And sharper that slope is that means, lower voltage you need to turn on and off the channel by many orders of magnitude which eventually define the 0's and 1's of your digital logic. The channel being on is 1; the channel being off is 0 right.

So, digital logic millions and billions of such MOSFETs are there in your cell phone that you are using now. So, if your slope sub-threshold slope is large, I mean your sub-threshold slope is not steep that means, you have to apply a larger voltage on the gate to be able to get the threshold voltage or to able to get the on off ratio that you require,

which means the battery will drain more power in your cell phone or your laptop whatever to actually make the digital logic work.

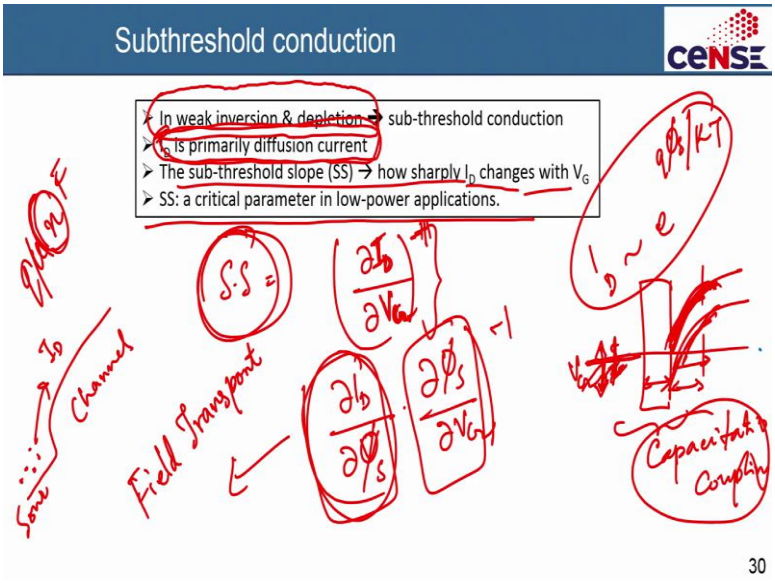
So, you need steep sub-threshold slop devices, where the sub-threshold slope is very steep which means you are able to turn on off the channel very sharply. So, at very low difference of gate voltage, you are able to do that which means your battery will have to supply lower voltage to get the same functionality in your logic circuits that means you will be able to operate much more efficiently ok, so that is why sub-threshold slope is a very important parameter. But, you will see that unfortunately sub-threshold slope in a MOSFET has a fundamental limit that you cannot exceed ok. So, the best devices will be have to will have to work at that limit ok.

(Refer Slide Time: 03:16)



So, let us come back to the slide here, where we had left. If you recall this was the slide that we have talked about last time, maybe I will erase some part of this here. So, I told you that we will discuss sub thresholds conduction again today here, we will we will I will give the expression for the limit of sub-threshold conduction. And then I will go to short channel effects and some other some other related aspects of MOSFET, and I will try to wind up the lecture of MOSFET here ok.

(Refer Slide Time: 03:39)



So, sub-threshold conduction happens in weak inversion mostly, and also in depletion that is before weak inversion has happened. In sub-threshold conduction, because the channel charge is very very low, your drift transport is negligible, because you know your drift depends on $q\mu_n$ times field. If this charge is very very low, because it is sub-threshold than your current I_D is primarily diffusion current. And the diffusion current will therefore depend, this I_D will depend exponentially to band bending the bending on the surface, which is like $q\varphi_s/k$ T, this φ_s is the band bending if you remember, this is a MOSFET diagram you know. This band bending that happens know this band bending that happens that is φ_s.

$$I_D \approx e^{q\varphi_S/kT}$$

$$SS = \left(\partial I_D / \partial V_G\right)^{-1}$$

So, essentially, I_D is the diffusion current which will depend exponentially on the band bending. And the sub-threshold slope will tell you how sharply that drain current changes with gate voltage, and it is a critical parameter in low power application. So, the sub-threshold actually is defined the slope you know is defined as the inverse of this how

the drain current is changing with respect to the gate voltage, but the slope will the inverse of it actually.

So, essentially your it should be a smaller quantity, if it is a small quantity is better which means, you know there should be a large change in the drain current with a very small change in the gate voltage below threshold below threshold ok, this is all below threshold below threshold. There should be a large change in the drain current, for a very small change in the gate voltage, I repeat that that sentence that means, this quantity will become very small, and that is a better devices steeper device steeper device.

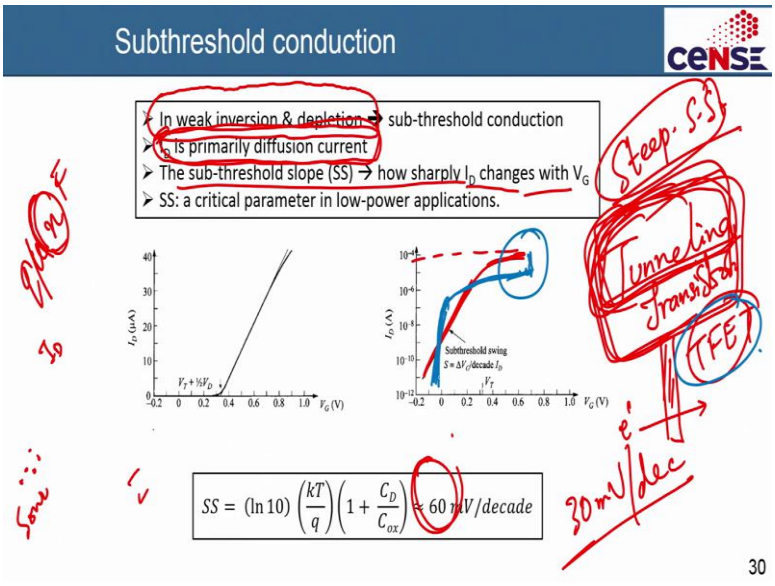
And this I_D this is inverse by the way, this I_D by V_G can be written as d I_D by d psi of s which is the band bending into $(d\varphi_s/dV_G)^{-1}$. So, this quantity the drain current changing with this band bending is essentially the field transport factor you can say ok, it is actually the field transport factor.

So, and the you know there is if you look at the source, this is the channel. And this is the source, then you have this electrons that you are injecting right. So, how much is the drain current with respect to the band bending that is happening, this quantity is captured by this. And this how much is the band bending with respect to the gate voltage you are applying, this is the gate voltage.

So, any small change in the gate voltage that you are applying, what is the rate at which you are changing the band bending with respect to the gate voltage that is essentially a capacitive coupling, it is basically a capacitive coupling, it is an interplay of capacitance across the gate capacitance, and the depletion capacitance. So, a small change in the gate voltage that you are applying here will change the band bending.

So, what is the rate at which the band bending is changing with respect to the gate voltage that you are applying that is called a capacitive coupling that it is actually manifestation of the capacitive coupling between the gate and the channel that is here. So, this two terms will give you this sub-threshold slope, and you can this expression can be derived this depletion expression. And this mathematics can be worked out in a simplistic way. We have you not deriving it here, I will show you the eventual expression ok, if you do that what will happen ok.

(Refer Slide Time: 06:52)



So, let me show you some printed figures, so that will give you a better idea, this figures are from s m g (Refer Time: 06:55), which also I will put up on the list, this is not my figure, this is figure from s m g. So, you see only a linear scale below threshold it is almost 0. But, on a large scale, you will see that the current rises many orders, when we did it here ok.

So, let me use a laser pointer again here. So, you can see that the below the threshold the threshold was at 0.3 volt around here. The slope changes, this is actually this ok, below that the slope changes. And this dramatic decrease or that increase you can say, this huge increase from many orders of magnitude happens at a very small range of like minus 0.1 volt to 0.3 volt, which is only 0.4 volt kinda, you know you have so much of change that is the sub-threshold. And it is called the sub-threshold swing or the sub-threshold slope ok.

The inverse of the sub-threshold slope is essentially the sub-threshold swing, and it so happens that this expression the sub-threshold slope is given by this quantity. If you do that maths, I had explained in a previous slide the capacitive coupling times the rate at which that drain can be changed due to band bending.

$$S.S = \ln(10)\left(\frac{kT}{q}\right)\left(1 + \frac{C_D}{C_{ox}}\right) \approx 60\ mV/decade$$

Then you get an expression like this ln(10), this comes because some of some conversion from log to linear, log scale natural, log and log base 10 and so on. So, it is a ln(10) into kT/q, which is the thermal energy point 0 to 6 electron volt at room temperature into 1 plus C_D by C_{ox}, C_D is the depletion capacitance C_D is the depletion capacitance. If you recall the capacitance in the depletion region of the semiconductor, because the bands are bending in the depletion, C_{ox} is oxide capacitance.

The best that you can have this quantity, you want this quantity to be as low as possible by the way that means, it will be very sharp transistor a sharp on and off. So, the best that you can have is that this quantity in the bracket can be one, which means this quantity can be very very less than one. If this quantity is very very less than one, then that is the best case scenario you will have, where this quantity is almost one, because any other quantity will increase that you want this quantity to be as low as possible.

So, the lowest value you can get, when you ignore this that is the best case you can get. When you ignore this, the best you will get is 60 milli volt per decade approximately. What does 60 milli volt per decade it means, it means one decade of change of current, one decade means one order of magnitude of change in current. One order magnitude or one decade change in current from this is say 10^{-11}, this is say 10^{-10}. So, this is one order. This is 10^{-10}, this is 10^{-9}. So, this is one order 10^{-9} to 10^{-8}, this is one order. So, for one on order of magnitude of change in current each order magnitude, you have to change the voltage by 60 milli volt.

So, every time you go from 10 to the power one order 10^{-10} to the 10^{-9}, you have to drop here 60 milli volt goes from 10^{-9} to 10 to the 10^{-8}, you have to go 60 milli volt. 10^{-8} to 10^{-6}, you have to do 120 milli volt, 60 again 60, because two order right. So, for example, if I want to go from 10^{-11} here, so 10^{-6} here, which is a 5 order of magnitude changed from 10^{-11} to 10^{-6} is a 5 order magnitude change of current here which means you have and this is this is you know, it will go linearly in that way, it is a large scale linear I am talking about.

So, if you change this by 5 order magnitude 10^{-11} to 10^{-6}, which is a 5 order magnitude change, then on the x-axis you have to drop the gate voltage by 5 into 60 which is 0.3 volt that means, a change of 0.3 volt on the gate voltage here will lead to a 5 order magnitude change in the drain current, and that is the best case scenario.

If your capacitance in this quantity becomes slightly comparable to 1, then this will even become worse, this will become 70, 80, 100 milli electron volt, and milli volt and so on that means, you will have to drop more voltage to have the same increase in the current per decade [FL]. If you have instead of 60, you have 100 milli volt per decade, it means for one order change in the current, you have to change the gate voltage by 0.1 volt, so that will take more charge more draining of the charge from a battery for example right on your devices. So, you want this quantity to be small.

If for example, this was 10 milli volt per decade hypothetically speaking that means, every one order by increasing the current, you only drop 10 milli volt 10 milli volt, how beautiful that would be? By the time you go from 10 milli volt to 50 milli volt, you will drop 4 or 7 magnitude that is beautiful right. But, unfortunately there is a limit to this sub-threshold slope in a MOSFET, because this is a this is a classic you know temperature dependent transport, this is a Boltzmann transport sort of a thing, so that is why you have this limit. And you cannot exceed this limit of 60 milli volt per decade, you can at best get this limit, you cannot get better than this ok.

The best of the best devices, no matter how much you scale the device, no matter what kind of geometry you use in a MOSFET, you will be able to only get maximum a best case scenario of 60 milli volt per decade that is the slope you will get, and that is the voltage that you have to apply. So, for example if I want to increase the current by 6 order magnitude, then 6 into 60 0.36 0.360 0.36 volt to I have to draw from the battery that is what it means ok. So, this sub-threshold slope is very important.

And remember this slope is only valid, below threshold, above threshold, this is the different slope, this is this which does not count, this is sub-threshold slope ok. So, because MOSFETs have this problem, so there are new kinds of devices architectures, although not adopted in the commercial you know application industry. But, there are new sort of devices that, people are trying to use to overcome this limit to overcome this limit.

For example, because this is a Boltzmann type of transport, you have this limit of 60 milli volt per decade. But, in some cases, there is also people are also using tunnelling sort of transistor. In tunnelling transistor, what happens is that you do not have this thermionic you know like a Boltzmann sort of a transport, and your off and on state of

the transistor can be defined by the tunnelling probability. You know tunnelling happens, when you have a very thin barrier of a material. And the electron here can actually tunnel through the barrier here.

So, by tailoring that you can actually block that current, and you will have an on, off ratio. And the tunnelling current, the tunnelling phenomena does not depend on the Boltzmann transport, this kind of a sub-threshold will not limit it. So, in tunnelling transport, you can get much lower than this. You can get 30 milli volt per decade for example even lower, I forgot the best number, but in tunnelling transport you can get much lower number. And so these are steep these are called steep sub-threshold slope transistor, it is a very hot area of research.

There is also negative capacitance transistors, I will not come to that here in this course, but the other kinds of transistors like negative capacitor transistors, and then there is a tunnelling transistors. This is called TFET tunnelling field effect transistor, which are widely being investigated for a long time now research in academy and industry, they are using this.

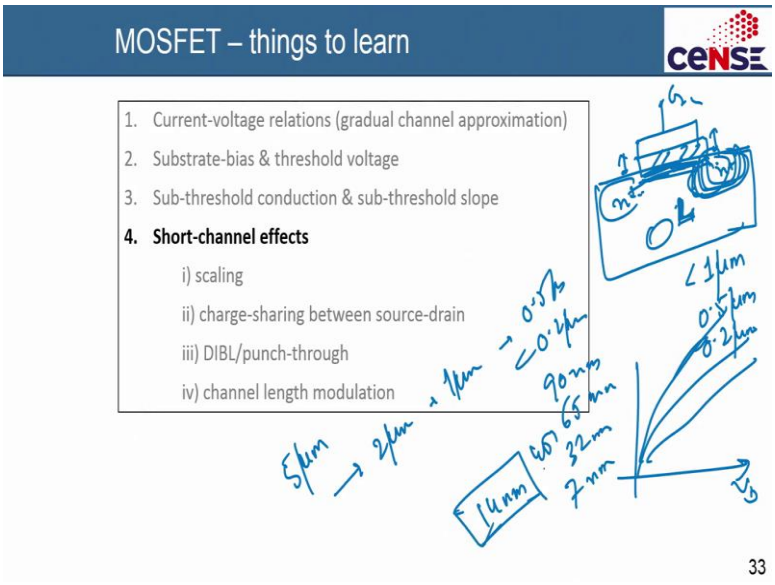
To actually, try to demonstrate much better sub-threshold slope, but as far as I understand this tunnelling transistors or steep sub-threshold transistor, this is a much this negative capacitance transistors are much more new (Refer Time: 13:51), so those are still there are undergoing research.

But, tunnelling transistors have been around for some time in research and academy industry, but tunnelling transistors have not been employed in large scale in commercial devices in cell phones, laptops, and all, because the on current of this tunnelling transistor is low. You do not only need a very steep sub-threshold slope here, you also need on current to be respectively high. The on current has to be high, for better current drive.

This tunnelling transistors will give you what it will give you know, I will tell you. It will give you a better, it will give you a much steeper slope, but the current will be very low. So, your slope will be very fast by much better, so it will be much better, but your total current output current eventually will be low. And because this is low, the current driving capability is low, your g m gain and other things also low, but in digital logic also.

You need a minimum drain current, otherwise you will have a lot of delay in the device ok, so that is the problem. So, tunnelling field effect transistors have lot of promise, but the on current is not high, because in tunnelling the current is not very high. So, although you get a better sub-threshold like this, but your current is low actually [FL], so that is one reason why you know, it is not being used commercially as of now ok. So, is a things there.

(Refer Slide Time: 15:05)



So, short channel effects; finally, short channel effects come into picture, when your MOSFET, you know if you have a source a drain, and this is the gate dielectric, then you have an the gate here right the gate here. Short channel effects come into picture, when this length, this is the length of the channel. This length becomes very small this length becomes small, and then what happens is that as you are shrinking the device you are shrinking the gate, you are shrinking the distance, you are doing everything, you know it is called scaling that is called scaling; scaling as in like you are scaling, down the dimension.

Your source-drain spacing, your gate oxide thickness, your doping, your all the geometry is being scale down to you know make the devices smaller and smaller that you know following Moore's law. So, short channel effects come in, when your length the device has become too small. When I say too small, I typically mean it is less than 1 micron or so, even at 1 micron you might start getting it; so, less than 1 micron or so. If its long

channel 5 micron 10 micron, then you do not worry. But, if you say less than 1 micron say you know 0.5 micron and so on, a 0.2 micron and so on, very small devices.

Then you have short channel effects, which are essentially detrimental to the device which are bad to the device. But, you cannot avoid them, because the source and the drain become very close to each other become very close to each other. There all kinds of leakage that may happen the drain might be able to influence the channel current, the channel current which was ideally was suppose to saturate like this know because, this is saturating, because the drain voltage here is not having an effect. But, because this source and drain are coming very close in short channel devices, you actually will not be able to get saturation, because the drain also will now start as influencing, thus current that is flowing. So, you will get something like this. This is very bad right very bad.

So, there are many phenomena associated with short channel effect will talk about scaling. Scaling actually is the process of you know scaling the device down, so it is the process essentially, it is not a short channel effect. But, scaling enables you to go to smaller devices, which will give you eventually small short channel effect. So, you have to get steps to overcome short channel effects.

Then there is charge-sharing between source and drain, this is one important thing. When the source and drain come very close, their depletion charge that is there know around the source and drain the depletion charge. The depletion charge will start getting shared between them that will lead to a lot of effect actually we will see.

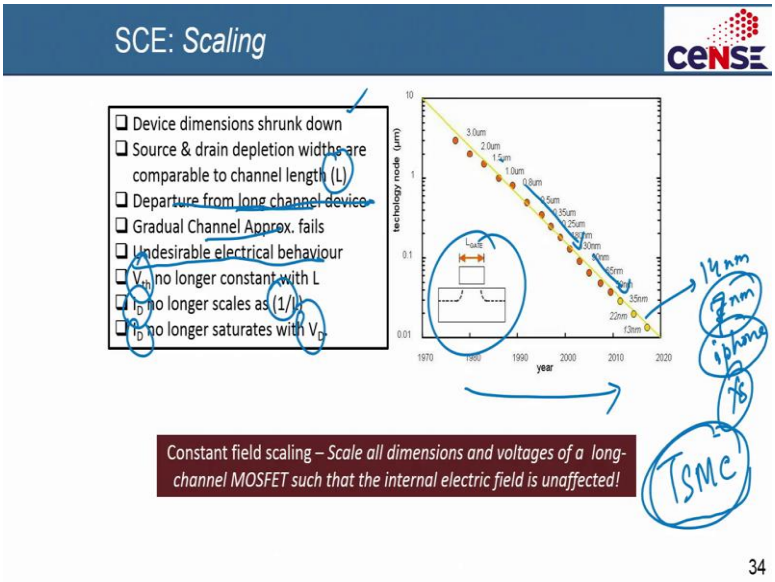
Then there is DIBL and punch-through is actually, the drain induced barrier lowering let us talking about. The drain will start influencing the channels charge, and so the what will happen is there? The drain will start having a role to play they play a role in the current is transporting here, so your current will not saturate. And then there is the you know on state, but even below threshold you have something called punch-through. So, these are bad effects that will happen. And finally, there is channel length modulation is not a very important effect, but compared to DIBL, but still will go through all this thing ok.

So, the first thing is scaling, scaling means that your devices are being shrunk down or the dimensions of the devices are being reduced, you know this is following Moore’s loss. So, initially you have say 5 micron devices, you are scaling down to 2 micron to 1

micron to 0.5 micron to 0.2 micron, then 90 nanometer, you know 65 nanometer. These are nodes, 32 nanometer, there is also 45 nanometer.

Now, for example, TSMC is a 7 nanometer node, Intel is a 14 nanometer node, all these are nodes eventually that you define. But, long channel 5 micron, 4 micron, and all are long channel devices. But, now we are very much in the short channel devices. And short channel devices have a strong effect on the device performance. So, scaling, how are you scaling the devices?

(Refer Slide Time: 18:38)



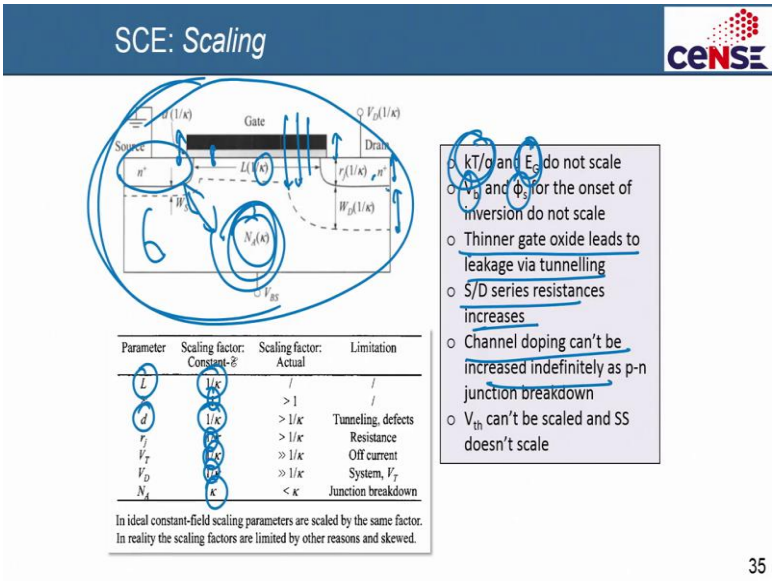
So, this is a figure, I have taken from internet, I will give the appropriate credits later on. So, it year you can see, it now as you increase as we are going to now this is little older plot anyways. This is the device size, the node 3 micron, 2 micron, 1.5 over time this is shrinking. We already at like 14 nanometer node from Intel, and 7 Nanometer node in the latest iPhone. In the latest iPhone X s, I guess. TSMC is a Taiwan semiconductor manufacturing company, I guess. They say that the apple has said that, they are already having 7 nanometer technology, made at TSMC in the iPhone X s I think, the latest X s ok. ah

So, these are these are very advanced devices that are shrinking down, and because of the shrink down, you have to take in the account many of the things. If device dimensions are shrunk down, then source drain depletion with were become comparable to channel length. Your gradual channel approximation will fail, in a long channel device equations

will no longer hold true. You will have undesirable electrical behaviour, because now threshold voltage will no longer be constant with length. Your current the drain current will not scale it 1 by L or W by L that is a problem, and also your drain current will not saturate with V_D. So, these are some of the problems that you will definitely face here ok.

So, constant field scaling, the first scaling rule that was proposed was that the electric field inside the MOSFET at every point should be constant as you scale down. So, you scale the dimensions doping, thickness, and everything else accordingly such that the electric field inside is constant that will give you a reliable device that was our people thought.

(Refer Slide Time: 20:06)



So, if you shrink down the channel by a factor of k, your thickness has to be also shrunk down the factor of k, your junction depth has to be shrunk down by factor of k, your depletion dept will therefore get shrunk down by factor of k, but for that you need a doping which is increased by a factor of k right. So, to maintain the same constant electric field, you see if your transistor length is scaled by k 1 by k, then electric field of course will remain same. Your distance that the thickness of the oxide also has to be scaled by 1 by k, this has to be everything will be scaled buy 1 by k, but doping will increase by k, only then your doping will shrink by 1 by k.

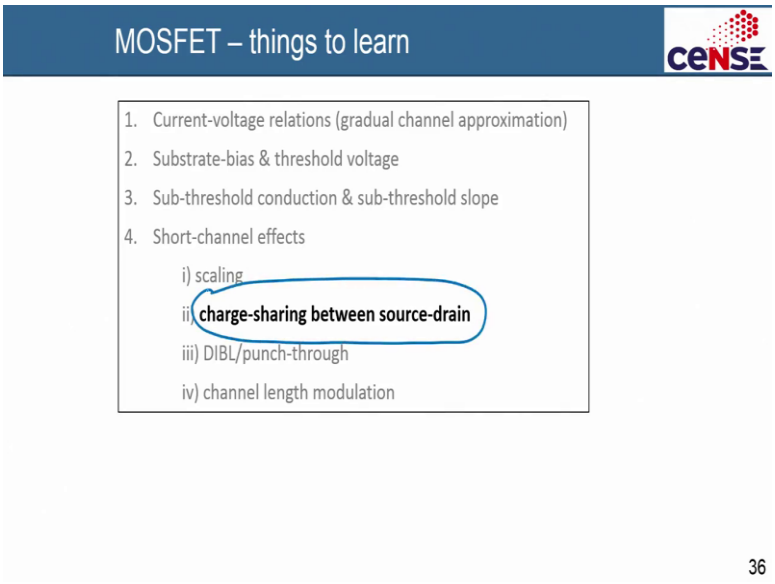
So, these are things, but you know people as were they were happy with this, but as way people were scaling more and more 80 nanometer, 90, 65 nanometer and so on. Problems started to arise, because this high electric fields, now the not high electric field. It was very difficult to maintain sustained is constant electric field scaling, because k T by q does not scale that is the thermal energy, band gap does not scale. The built in potential, and the band bending at the interface also do not scale.

And because your scaling the oxide also, because you are scaling down oxide that oxide is becoming thin. So, the gate is now leaking as a tunnelling, so that is a bad effect. Resistance is a increasing, because your decreasing you know you are in decreasing this thickness. So, resistance is also increasing. You cannot increase the channel doping, indefinitely because then this will breakdown right that is another problem.

So, there is a many other problems, and sub-threshold slope, and other things will not scale. So, people realize that you cannot maintain a constant field scaling. So, people have taken an adoption of you know field and voltage scaling sort of thing, but they scale eventually their rule is that you scale different quantities accordingly or appropriately such that the overall device works fine is not necessary that you have to scale everything by a particular factor.

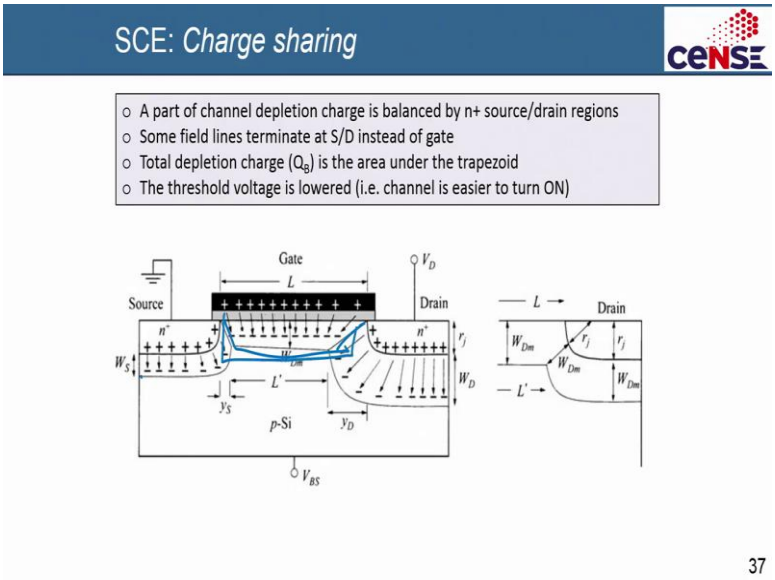
You might scale 1 by 1 factor, you might scale the other thing by another factor, gate oxide by another factor, bodies affect by other factor, you can scale down different factors. Appropriately, such that the overall device works acceptably, and simulations and computational models and other things are used. But, remember that constant field scaling that was proposed initially does not hold true, you have to take into account constant field, constant voltage, the fact that thermal energy and band gap does not scale and so many other things. So, people try to scale it in a way appropriate way these different parameters, so that your overall device performance is not compromised ok.

(Refer Slide Time: 22:17)



And next short channel effect is charge-sharing between source and drain, what does this mean?

(Refer Slide Time: 22:21)



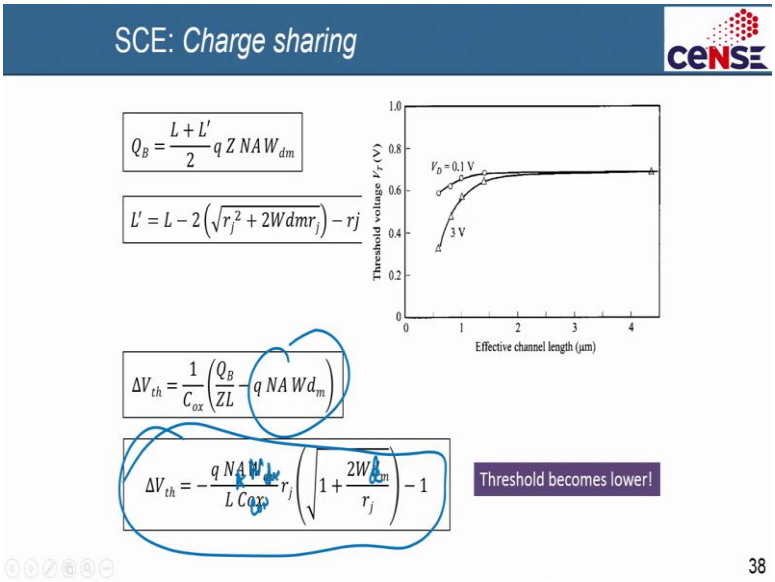
It means that when your source and drain come very close to each other, the depletion region around the source and drain, let me get the laser pointer here. The depletion laser, this is the region of the source is the depletion region of the drain, when you come very close to each other, then this depletion region will start to overlap. So, the charge-sharing

happens between the source and the drain. And this is a channel depletion, you see there depletion below the channel is happening, because you applying some gate voltage here.

The depletion below the channel which is some charge that is happening here will now be shared between the source and the drain you see this. So, this is like a trapezium, if it is a long channel, then this trapezium will look like a rectangle which is fine ok. So, it is a trapezium, because some of the electric field lines that are actually in this depletion are not terminating on the on the channel charge depletion here, but they terminating under source and drain contact, you see this field lines are terminating.

In simple language, the depletion charge below the gate should only be charged below the gate. But, no in short channel device, because the source and drain come very close to each other, their charge gets all shared up. A part of this charge will be the it was the drain depletion, a part of this charge will be shared with the source depletion, and that is why this becomes like a triangular profile.

(Refer Slide Time: 23:44)



$$\Delta V_{th} = \frac{1}{C_{ox}}\left(\frac{Q_B}{ZL} - qN_A W_{dm}\right)$$

$$\Delta V_{th} = -\frac{qN_A W_{dm}}{LC_{ox}} r_j \left(\sqrt{1 + \frac{2W_{dm}}{r_j}} \right) - 1$$

And you can do some simple math and geometry here. To essentially find out, what is the area under this trapezium that will give you essentially the charge that you have to that is the amount shared, and that is the shared you know that is the this has L, this L is the length of the gate. But, this is trapezium, so this is L dash.

So, this is L plus L dash, N is the doping, W dm is the depletion that is happening, this is subscript for example sorry it is a typo here, this W dm. Anyways that they come messages there if you do take the charge into account here, and that is the extra charge sorry that is the extra charge that you know there is the extra charge that have to now consider when you are talking about charge sharing, so this is the this is the this is the Q_B, which is the charge that is stored in this trapezium below the gate. Otherwise, it was a rectangle, if it has a large long channel device.

So, it will eventually lead to a change in the threshold voltage also, and that threshold voltage change is given by the charge that is stored this Z L comes because of the normalization factor. The charge that is stored below the gate minus the charge that is actually supposed to have been stored, below the gate due to depletion. This q N_A W d m, this q N_A W d m actually term represents the depletion charge, if you recall from a few lectures before in MOS capacitor MOSFET.

q N_A W d m essentially represents the maximum depletion charge that will be below the gate, but because of charge sharing, there will be a trapezium sort of a charge distribution here. So, what is this charge stored here minus what is should ideally store, gives you the difference in the threshold voltage, and the threshold voltage turns out to be this quantity which is and r j, here of course comes here r j is the junction depth the depth of the not junction depth this is the depth of the source and drain in plans here ok.

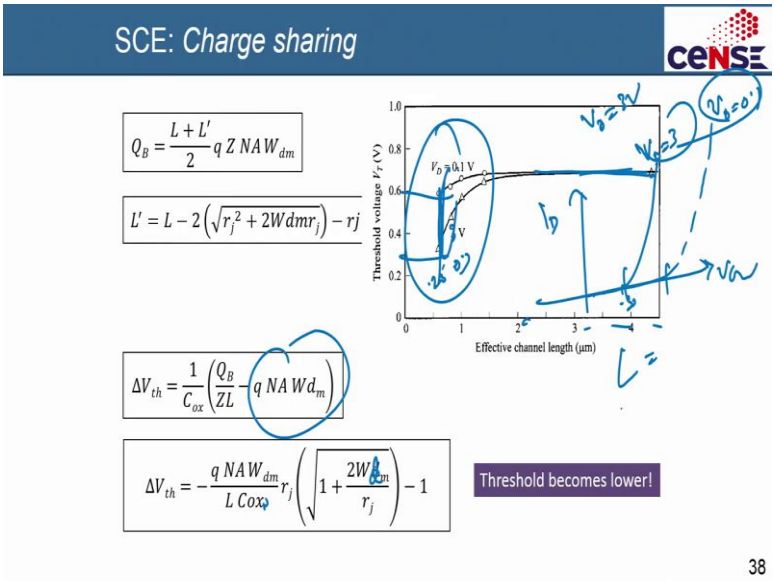
So, this is the depth of the source and drain in plan that is very important here. So, it will go as q N_A W d m by L C ox, for some reason my subscript or superscript does not come outside here C_{ox} N_A, it is also subscript. So, depends on the oxide capacitance is depends

on the doping, it depends on the depletion, it depends under source drain junction depth, this is the junction depth source-drain.

So, threshold voltage does have a change here, and is a negative change, because now you are storing less charge, then what you are suppose to store. Ideally, if I take the pen again, ideally you were supposed to have a rectangular profile, ideally you are supposed to have a rectangular profile like this. And that rectangular profile would have been this charge that rectangular profile would have been that charge.

But, unfortunately, because it is a short channel device their source and drain are coming close that is why have a trapezium profile. So, the difference of the rectangle versus this trapezium in a way is the threshold voltage. This expression is very important, by the way do not have to memorize, but you should know it this is W d m, this is C ox, this is N A, this is W d m ok, so that is what the charge sharing is (Refer Time: 26:22).

(Refer Slide Time: 26:24)

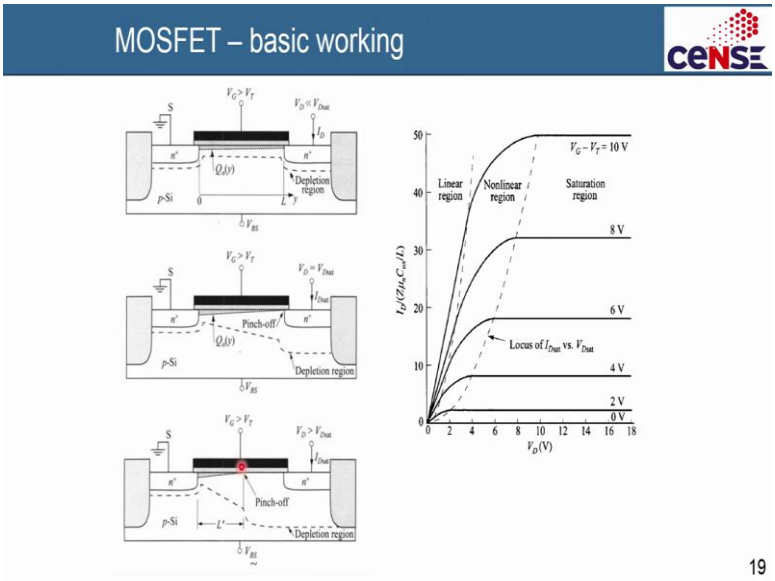


What is the manifestation as your channel length is more 2, 3, 4 micron, this is a channel length. You have no difference of the threshold voltage at different drain bias that at mean in a you know that different drain biases threshold biases, the voltage is the same. But, if you come the short channel device, then with different drain voltage at 3 volt or at 1 volt, 0.1 volt, your threshold voltage is much different. In one case is 0.6 volt, in one case it is 0.25 or 0.3 volt.

So, lot of difference in a threshold voltage that can affect the device, it excuse me it means if I plot I D versus V G, if I keep the drain voltage at 3 volt, I will get a threshold voltage like that. If I kept the drain voltage, at 0.1 volt, then you know this is 0.3 volt for example, then I will get it at like this.

You see the threshold voltage has changed because of this is V D equal to 3 volt, this is V D equal to 0.1 volt with different drain bias, you will get different threshold that is not a good idea. This threshold becomes lower alright, and that is why you are actually this is effect of the charge sharing that happens.

(Refer Slide Time: 27:46)

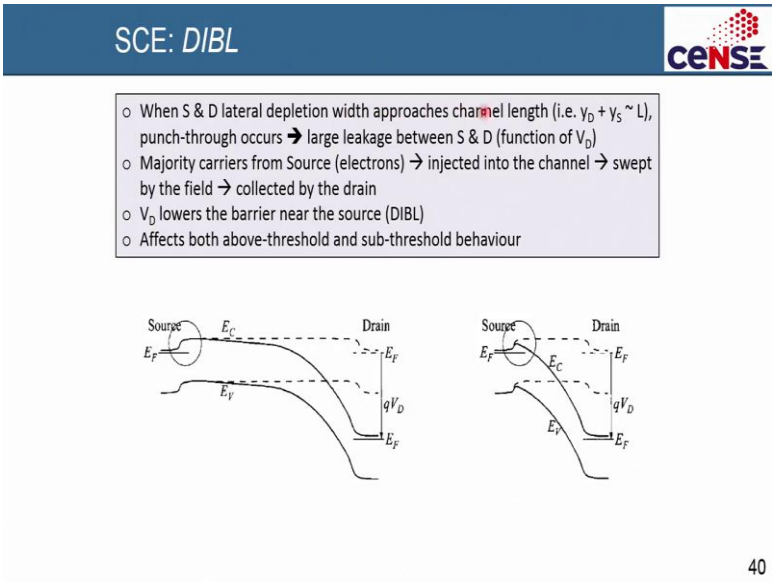


And the last thing that will study in the in this section and this channel length modulation is one important. The channel length modulation actually means that your I will come to that actually in a bit here channel length modulation means that in a MOSFET device, if you recall from very long back in a MOSFET device.

In this you know in this MOSFET device sorry in this MOSFET device, let me pull out the laser pointer here. This pinch off that is happen know with higher drain voltage, this pinch off point moves more and more. So, this channel length becomes lower and lower know that channel length modulation, typically does not come into play in a bigger channel device.

In a big channel device, this movement with respect to the channel length can be neglected. But, in a very small channel device, when the source and drain are very close, then with higher and higher pinch off I mean higher and higher drain voltage. Your pinch off point moves to the left, this channel length becomes smaller to compared to the original one. So, the current will not saturate, it will go like this like this like this that is called channel length modulation. But, that is not an important effect compared to the DIBL and punch through that will be discussing now ok.

(Refer Slide Time: 28:43)

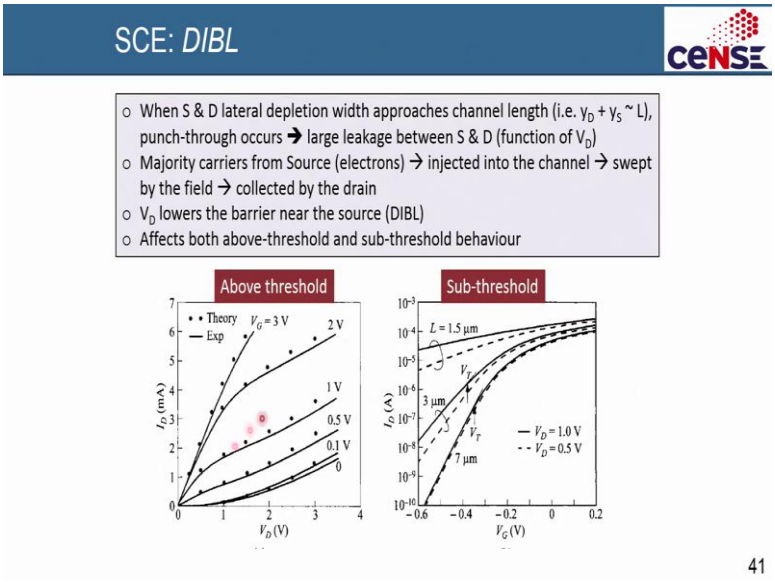


So, DIBL and punch-through, this I have told you already. DIBL and punch-through essentially, what happens is that in a long channel device your source your drain and your gate is like that. When you apply drain voltage you have band to band, but the drain voltage is not able to influence the amount of electrons that you injecting from the source that is why the current saturates. But, the moment your drain current comes very close to the source, which is this the drain comes very close to the source.

When you are apply drain voltage, and you change the band bending like that, your drain actually is able to influence the amount of electrons that you injecting from the source through the gate side to the drain through the gate and so. This barrier from the source through the gate is called the barrier, you know in injection can be influence by the drain bias which ideally should not that is why, it is called drain induced barrier lowering, it is called DIBL drain induced barrier lowering. And because of this drain induced barrier

lowering, your drain current drain voltage will now have a have an effect on the drain current, it is coming out. So, the current will not saturate.

(Refer Slide Time: 29:32)



The current will not saturate like this, you see the current will not saturate that is called DIBL for a short channel device ok. So, the source and depletion width approaches the channel length, then punch. So, one thing is that this is the DIBL that is happening ok. The barrier at the source side is lowered, but there is another manifestation of this problem by the way. This source side barrier lowering and the current increasing with increasing drain bias much to our you know nuisance that is called DIBL, and it comes at the above threshold condition, but below threshold condition also, there will be a problem.

When the source and drain come very close to each other, the source and drain lateral depth depletion depth will become comparable to the channel length. So, you if you recall this figure, your depletion from the source and drain side will become so much that this channel length will be own by the depletion from the source and drain side. In that case, what will happen? In that case, there will be a punch through that will occur, because there will be a large leakage between the source and the drain is the function of V D which means whenever your source drain source side and drain side depletion region come very close to each other, this are high field depletion region like base character depletion region of a BJT.

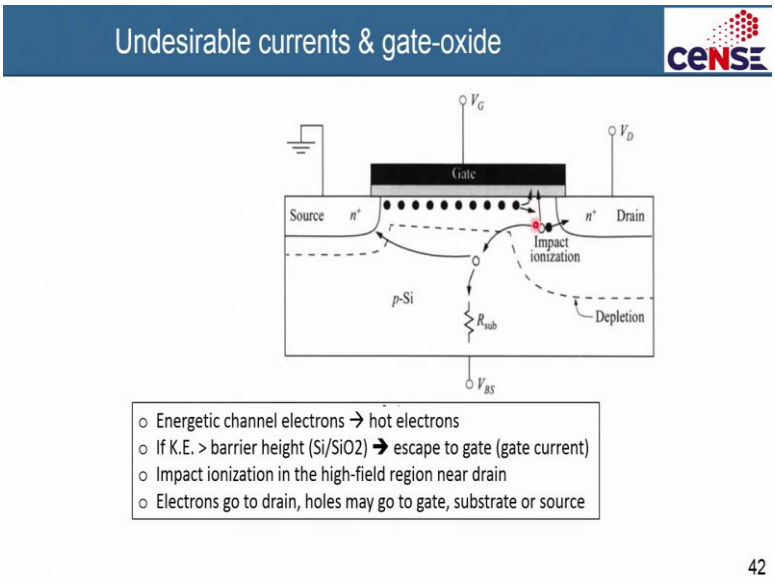
If you have very high field depletion region that come very close to each other, then electrons might be injected, maybe swept away by a very injected from the source. They might be swept away by very high electric field to the other side to the drain, because it is a very short distance. Large amount of current will increase will flow, and that is called punch-through. And as a function of drain bias the punch-through can also increase. Please remember that n plus source can inject electron to the depletion which will be sweeping the carriers very fast, but the drain is already very close to the source. So, it will it will eventually go to the drain very immediately you know, it is a punch-through, it is a very large increase in the current it will happen ok.

So, what will happen is that you will have at short channel device, you will have large current that will flow within the source and the drain, you will not be able to turn off the channel, so a threshold behaviour will suffer. What I mean is that you see this is I D, this is V G, for a long channel device like 7 micron, your current is turning off very nicely by any orders of magnitude. But, the moment you go to 1.5 micron, which is approaching short channel. At 1.5 micron the source and drain depletion layers are very close to each other, so any electron injected from the source will be swept away the field very high, so the gate will loose control of the channel.

So, you see the pinch off is not happening well, this is on, this is off. You see on and off has such a high ratio [FL] on and this is off has barely any difference. So, with short channel devices, your threshold condition will become poorer which means your gate will lose control. So, even if you are sweeping the gate voltage here, your current actually is mildly changing only. Over here the current is changing by any orders of magnitude, which are change in gate voltage which are able to turn off the channel here from 10^{-10} to 10^{-4} here right or 5 here.

But, with short channel would punch through, because you have very closely spaced source and drain very scaled device, your electrons that are injected from source will be swept away, the gate will lose control and at 1.5 micron even below, your gate voltage will have very small control on the channel, in other words the channel will not turn off. So, this is a punch-through effect in sub-threshold condition. This is a DIBL effect in above threshold condition, where the current does not saturate. These are a problems of short channel devices ok.

(Refer Slide Time: 32:38)



And apart from that there are also other things like gate oxide breakdown that you should be also aware of not much that you need to study, but you know if you have a very thin gate oxide, then there might be gate leakage. Even otherwise, there might be energetic electrons that might be injected from the gate, and energetic electrons are ejected from the gate will actually have the detrimental effect or bad effect on the gate oxide.

Apart from that there is another important thing that you should be careful. When a source and drain are closed, and the you are injecting electrons from source to the drain side, sometimes the energy of the electron might be so high, if the channel is low or a field is high that they might have the kinetic energy to surmount this barrier the surmount barrier, and leak out to the gate.

In that case, this high energetic electron will create impact ionization in the oxide. And impact ionization in the oxide if you recall, we will essentially create leakage current and then also will create device, you know more leakage actually leakage is not the big deal. The big deal is that this high energetic electrons that are escaping here, when they in a impact ionise the oxide layer, they will create lot of defects and that will create to reliability problem of the device ok.

And if impact ionization happens electrons go there, then holes may go to the substrate or it may also go to the source, there might be a parasitic channel that might form it like a BJT sort of an action. So, this is an important thing you should be aware of that it also

scaling and other things, there might be undesirable currents and the gate oxide breakdowns that might also happen in terms of reliability problems.

So, with this we end the lecture now. And we are done with a MOSFET ok, so we are more or less done with MOSFET including short channel effects including your DIBL punch-through, your substrate bias effect, your sub-threshold slope, charge-sharing, how this affects the devices. We have had a fairly you know simple description of that I hope, you are able to understand the effects.

The simple equations that I have written down without much derivation, those will be uploaded in appropriate nodes, so no worry about that. So, the equations are useful in quickly calculating the for example the change in threshold voltage, in a in a in a short channel device, in a few shrinking device channel, what is the change in the threshold voltage.

If threshold voltage changes, lot of things will change in the circuit right. So, it is very important to keep that into account, you know keep them into account. So, we are more or less done in MOSFET that is a topic the more advanced concepts in MOSFET. In terms of short channel effects, they are more advanced mathematics, they also like powers of double integral that are also more advanced topics.

Many of the things that we are not covering here, because the objective of the course is to have more basics of different courses and build up your foundation for device understanding, so that you can analyse device you can appreciate how electronic devices work, you can also get some basic understanding to scoring your gate and other exams. So, much more advanced concepts are not discussed in detail in this course.

So, we are done it MOSFET now, what remains is we will start compound semiconductor from the next lecture, you have already heard about compound semiconductor many times in this course. We will go through a couple of lectures on compound semiconductor devices, hetero structures very important. Hetero structure devices, we will study both the optic and electric, and then we will go to like things like LED’s, and solar cells, photo detectors, and then we wrap up the course, 30 hours will be completed ok. So, with that we end up the lecture today, I hope to see you in next class with a compound semiconductor ok.

Thank you.