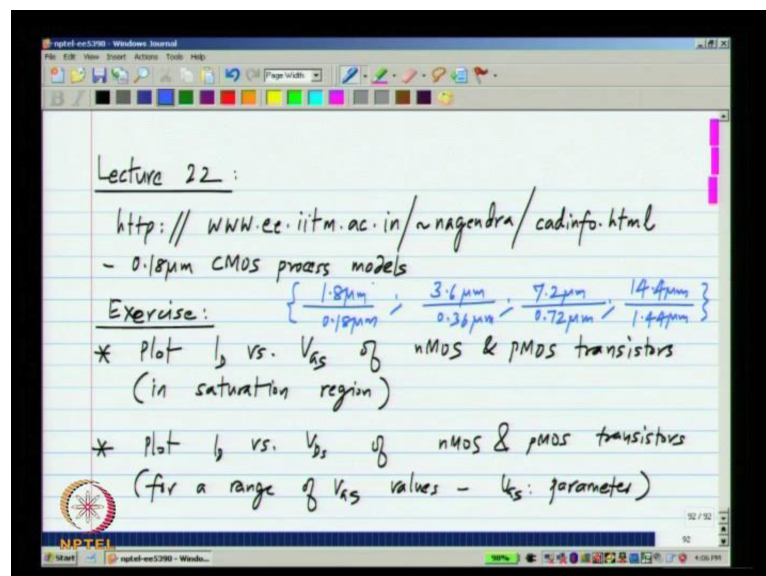
Analog Integrated Circuit Design
Prof. Nagendra Krishnapura
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 22
MOS Transistors – Parasitic Mismatch

Hello and welcome to the twenty second lecture of analog integrated circuits design. In the previous lecture, we discussed the large signal model of the mos transistor as it appears on integrated circuits. In some details, we saw that it eta square law when the VDs is greater than a certain amount. We also saw that the bulk voltages has a significant influence on the current in the mos transistor, this is because the bulk and the gate are similar in a mos transistor.

They play similar rules in an nMOS transistor, increasing the gate voltages for the bulk voltage increases the amount of charge in the channel and least shown increase in the current modeling the effect of the bulk is quite important. In a MOS transistor especially in integrated circuits, we also discussed the behavior of the MOS transistor when the VGS is below the threshold voltage and we saw that it turns out to be an exponential. Now as I also mentioned in the last class, the real model of MOS transistor is very complicated. So, you have to use a simulator to gain better understanding of the MOS transistor.

(Refer Slide Time: 01:48)



So, what I suggest is that, I take it as an exercise to get hold of some MOS transistor models. In fact, some models are available from this websites. The number of models are available, you can chose the 0.18micron pMOS process models and you take these models, get familiar with the circuits simulator. There is the number of circuits simulators available both free and commercial simulators. You can download some of the free ones from the web and the kind of exercise that I am going to mention, can be very done on a free simulator. So, please do the following, what Id verses VGS of nMOS and pMOS transistor.

When I say VGS for pMOS transistor it means the VGS of course in saturation region and similarly plot Id verses VDS of a nMOS and pMOS transistors for a range of VGS values that is with a VGS as a parameter. Now, you will have to do this kind of exercises any way when you come across the process for the first time. Now, this also helps you understand the limitation of the models that we have. The models you have on this website are quite complicated and they model the MOS transistor very elaborately.

So, when you do these two things, let us say you chose the 0.18 micron cMOS process as I mentioned earlier. When you specify a length of a process, it means usually the minimum gate length that is possible for a MOS transistor. You try this for different values of W by L. So, let us say, I will say 1.8 micron by 0.18 microns. You can double both of them, similarly double it further.

Let us say you take this four values, you could also take other things in between the point of this value that all of them have a W value ratio of 10, but as I mentioned earlier, there are lots of effect which makes the transistor characteristic deviate from square law. Now, one thing that, we did not discuss is the sort channel effect. When the channel length becomes very small then, it is definitely not a square law device it turns out when the channel length is very short, it s not a square law device. Now, again we have not discussed the device physics.

It is not a square law device, now again we have not discussed the device physics, but the square law comes from different dependence on VGS minus VT. First of all the amount of charge depends on VGST minus VT and the electric field inside the device depends on VGS minus VT and in a certain region the velocity of the carriers is directly proportional to the electric field.

So, the total current is related to the number of carriers which is related to the total amount of charge times. The velocity of the charges since each of this is depended on VGS minus VT; we get VGS minus VT square. Now, it turns out that, as you increase the electric field beyond some limits the velocity of the carriers becomes saturated, this is known as velocity saturation.

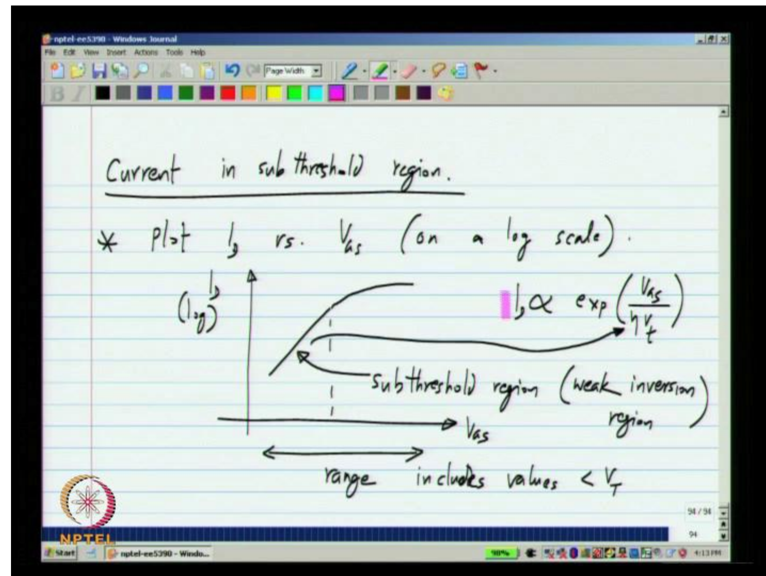
So, at this point the velocities of the carrier's no longer increase with apply of electric field. So, this means now the total current equals the product of speed which is more or less constant with times amount of charge which is proportional to VS minus VT. So, the current instead of being what you end up getting, it turns out that this is the speed of the carrier. So, the velocity saturated current is given by something proportional to be VGS minus VT and the saturation velocity of the carriers because of this reason the transistor current does not obey an exact square law dependence with VGS.

Also another thing to keep in mind is that, the regions are not appropriately divided here. We have the velocities saturation and here we do not like all physical phenomena. This happens continuously, so the bottom while is that, if you do the characterization for these sets of channels length and channel width let us say 1.8 micron by 18 micron or all the way to 14.4 microns by 1.44 microns, all of them have a W by L ratio of 10.

But, you will see that the behavior will be quite different. In fact, in particular you should try to identify these things. The reason in which the square law is not obeyed and this can be found from the plot of Id was VGS in the saturation reason and also by plotting either Id verses VGS or square root of Id verses VGS. You will also be able to identify this threshold voltage, which is used in our simple model and the Id verses VGS will give you the value of lambda.

Another thing how the transistor current depends on the VDS in the saturation region, how it depends on that in the dry out region and so on? Finally, if you absorb properly, you should see that as you increase the channel length to very large values, it will obey square law and as you reduce the value of channel length, it will not obey the square law, but you should still be able to resign circuit with this transistor, whether the current obey square law or not. So, this is one of the exercises. The other exercise is to find out the current in sub threshold region for this.

(Refer Slide Time: 09:36)



What you do is plot ID verses VGS on a log scale as you know the log scale is used when the quantity is cover a wide range and as I said, when VGS goes below the threshold voltage the current will resource small that can be meaningfully represented only on a log scale. You can plot ID verses VGS and you cover a range of VGS which includes values less than VT. Now, already said that ID will be proportional to exponential VGS divided by the thermal voltage times some factor eta, now if you recall a perfect by polar transistor will have eta equal to one.

A MOS transistor will not have a eta equal to one. In fact, all of these things you can estimate from this plot. When it is following the exponential the curve ID verses VGS, where ID is a log scale, will be straight line and this will be the sub threshold region or the weak inversion region and when you go above the threshold region, it is square law which means it is less compare to exponential and it does something like that.

So, the point of this exercise is first of all to identify the sub threshold region and secondly from this region you can get the slope eta. You can calculate the slope of this straight line and from that find the factor eta. We will discuss this a little more lately, but you can see that, if eta is close to 1, then it is better because the exponential will be really stripped, so that is the point of this exercise and do this for both nMOS and pMOS and similar rise yourself with the MOS transistor.

(Refer Slide Time: 12:05)

Now, finally the third exercise would be to find the effect of VSB, you take a MOS transistor and set with the fixed VGS and VDS and varying VSB. Let us assume VSB more than zero, so that you do not forward by the source of bulk junctions. Now, you can plot ID verses VSB, let us assume that you maintain VDS and VGS, so that it is in strong inversions and saturation. So, this is one of the exercises and secondly you can also plot ID verses VDS for different values of VSB with VSB as a parameter. Remember, earlier you plotted ID verses VDS with VGS as a parameter for different values of VGS.

(Refer Slide Time: 14:27)

You would get different saturations currents. Now, if you do the exercise correctly you will see the same effect with VSB as well. This is again to convince yourself that the back gate and the front gate qualitatively. Similarly, quantitative functions there will be different from each other. So, the previous lecture plus all this exercise should give you a very good idea to handle on the large signal characteristics of the MOS transistor.

So, now move on to small signal model with the MOS transistor because the ID is a function of VGS VDS and VSB. In the small signal case it usually is written as a function of VBS. We know that the change in ID will be a function of change in VGS, change in VDS and also change in VBS. What kind of function it will be, we know that when the changes are small it will be a linear function.

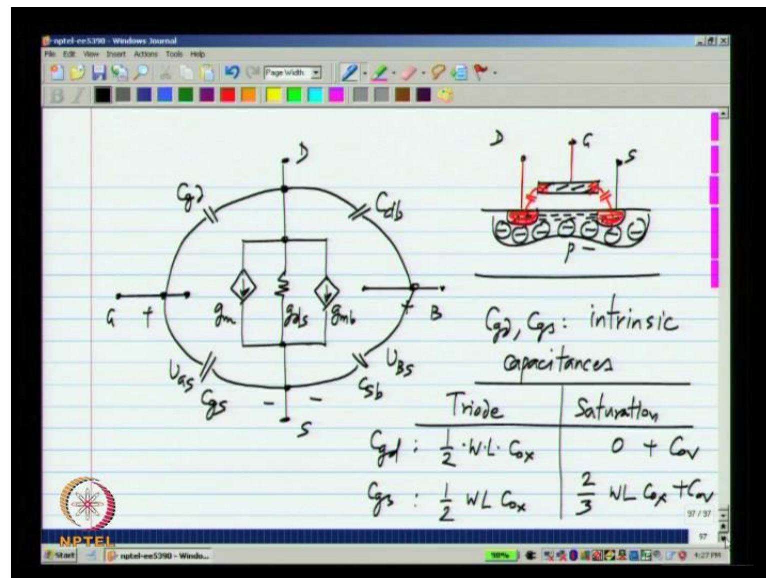
So, delta ID will be partial derivative of ID with respect to VGS at the operating point time's delta VGS plus partial derivative is respect to VDS. Again at the operating point which is very important times delta VDS plus partial derivative with respect to VBS times delta VBS. Now, it is common to denote changes by lower case letters which is what I am going to do from here onwards.

This is VGS VDS and VBS and this could be represented instead of delta ID as lower case id, this behavior is what is captured by a small signal model of the MOS transistor. So, first of all the dependence of ID on VGS, we need to have a control source. This is a linear control source, like everything in the small signal model is linear and the value of the control sources dou ID by dou VGS. So, we have four terminal drain gates bulk and source. So, between the drain and source will have a control current source and this denote this particular current and this constant aware are called gm?

Now, this current is gm times VGS, where VGS is the small signal, gets source voltage similarly, this current can be represent by another control source and this constant here dou ID by dou VBS is like gm, but from the bulk. So, it is usually denoted by gmv and in the small signal case we use VBS instead of VSB, that is the convention and this control sources gmb times VBS and finally, this current here relates the current from drain to source as a function of voltage between drain and source. We could use a control current source but it is easier to just model it as conduction because the controls in terminals on where the current is flowing are the same and this quantity is none as GDS.

So, this is the conductance which is GDS. Now, this is the DC small signal model of the MOS transistor. The left more source, it denotes the current from drain to source based on the increment in VGS, the conductance GDS denotes the dependents on a drain to source current on the drain. The small signal sense and finally, gmb times VBS denotes the dependency on brink current on the bulk source voltage and because this does not include any capacitor or charge storage effect, it is the DC small signal model of the MOS transistor.

(Refer Slide Time: 18:58)



So, this is what we will use for Dc small signal and analyses here onwards. I will simply gm here, it is understood that it multiplies VGS and will write gmb and it is assumed that it multiplies VBS and I will write GDS here. This is just to simplify the diagram and reduce plotter. Now, this is the Dc small signal models in addition to this let us take again nMOS transistor. Now, first of all there will be capacitance between the gate and channel because after all that is what the MOS transistor is. It is a capacitance that is form between the gate and the inversion layer when the transistor is on, so this contributes to some capacitance in this circuit.

Now, this is really a distributed capacitance which goes between the gate which stretches all the way from here to there and the inversion layer which also stretches all the ways form here to there. Now, we have to represent a capacitor between discreet terminals drain gate and source. So, what we will do is we will apportion the capacitance into two

parts one between the drain and gate and one between gate and source this is called Cgd and this is called Cgs due to the depletion region there will be the capacitance between drain and bulk which is Cdb and between source and bulk which is Csv. Primarily due to the depletion region below this diffusion regions. Now, these are the Cgd and Cgs intersect to the operational the mosfet as we know mosfet is nothing, but a capacitor and it controls the charge and inversion layer and in turns control the current. So, this capacitance will always be present whenever you have a mosfet without this capacitor. There is no mosfet because the Cga and Cgs what we describe so far come from the intrinsic capacitance between the gate and the inversion region.

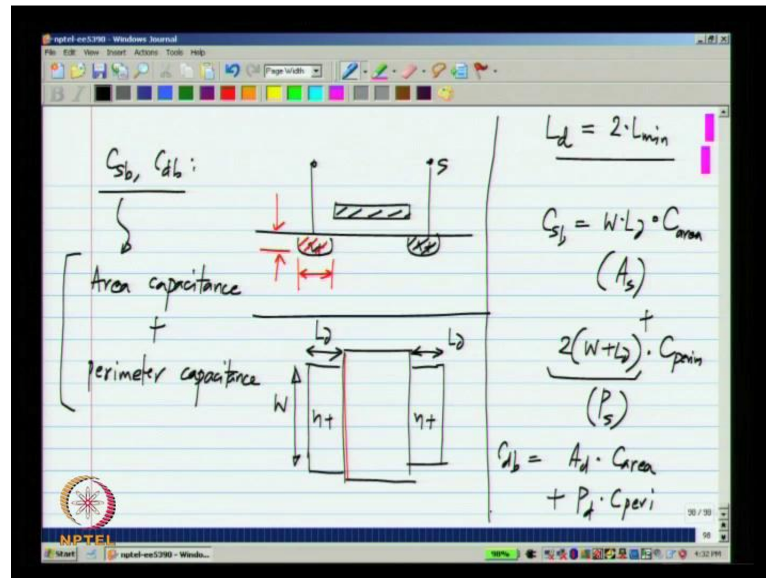
So, this is known as intrinsic capacitances and the value of these things depend on the operating point. It turns out that in try out region Cgd and Cgs are both equal to half of W times L times Cox. This W times L is nothing, but area of the gate and Cox is the capacitance per unit area of the oxide under the gate W times L times. Cox is the total capacitance of the gate, I mean try out reason this are equally opposite and in between the drain and the source both Cgd and Cgs equal half WL Cox. Now, this you can easily imagine the case where the drain and source voltages are identical then the channel will be perfectly symmetrical and then clearly the configuration between the gate and drain will be the same as that between gate and source.

So, the proportion of charges allocated to drain and source are identical as you go far away from deep region the charge distribution becomes less and less symmetrical and we will have a different charges distribution. In the extreme case saturation reason Cgd will be zero and Cgs will be two- third WL. Cox does all this come from device physics by calculating the amount of charge in the inverse layer the total charge is two third SWL Cox and all of it is the appropriation to the source. Now, nothing is appropriate to the drain because at the drain the channel is not inverted and there is pinch of layer. So, nothing is really connected to the drain. So, these are the values of Cgd and Cgs by the way this model is valid in all regions including gm gmVS and gm GDS only.

Question is what are the values of those things now? Will it be using this module mostly in a saturation region because that is why we operates our amplifier, but the model itself is valid everywhere. Now, this is about the intrinsic capacitor in a addition to this you see that you have some conducting material here. In addition to this you could also have these wires and you have the conducting gates and between these two there will be a

capacitance. This is not inherent to the operation the mosfet, but it is there simply because the two pieces of conducting material are close to each other. So, the actual Cgd and Cgs will consist of this and some overlap capacitance.

(Refer Slide Time: 25:14)



Now, again drawing a picture of mosfet you want to calculate the overlap capacitance between these two. You imagine the top view, where the gate is like that and the drain and source region are like that, now this capacitance nothing, but the fringing capacitance between this conductor which is the gate and this is which is the drain you can easily see that, the capacitance will be proportional to the width of the overlap which is nothing, but the width of the mosfet.

Something proportional to the width and it is the same thing is for Csb as well. So, I will denote this as Coe prime which is the capacitance per unit with times the width of the mosfet and exactly the same thing for Cgs as well this is plus u times W. So, those are extrinsic capacitance, but they will be there any way till you have a mosfet. So, you also have to model them, now regarding Csb and Cdb Csb is the basically the junction capacitance of the source junction. This one and Cdb is the junction capacitance of that junction. Now, exactly how much the capacitance is depends on the area of this phase. The bottom phase of this and also the side phase of this now again. If you look at the top view of the MOS transistor, this diffusion region n plus will have some length late call that Ld and this usually as the same length L d and typically as a guide line if you do not

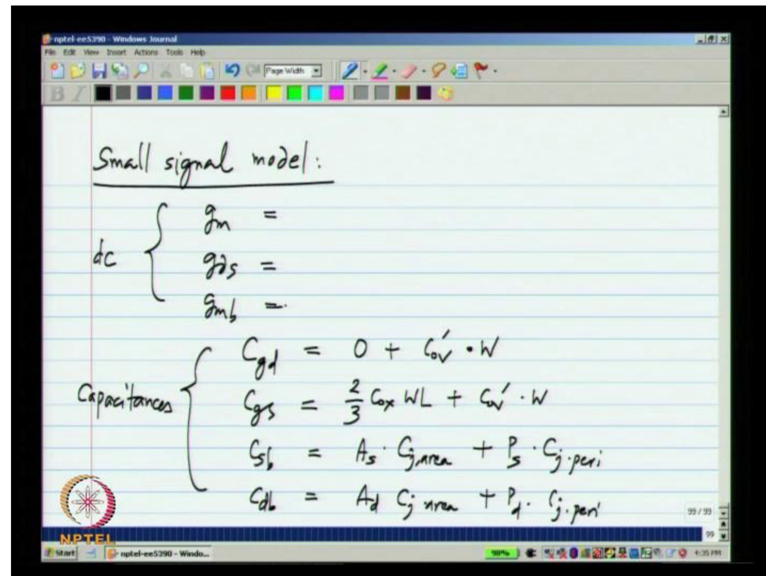
have any other information, you can assume that Ld is two times the minimum length of the gate and in a given technology and this dimension is nothing, but W as usual we do not have control over this depth, that is why we do not want any parameter describing the depth. If you know the depth, the depth times W is the side wall area, but we do not want any of that the depth is fixed.

So, the capacitance is described as being made of two pieces one which is the overlap area at the bottom which depends on W times Ld. So, each of this things consists of what are usually called the area capacitance which refers to the overlap area at the bottom plus some perimeter capacitance as I describe while defining the sheet resistance and so on. In an integrated circuit, we have control over only the two dimensions.

So, all the other dimensions are subtracted into parameters and we will describe everything into dimension that we can control which is Ld and W and so on. The area capacitance part is given by W times Ld which is usually called as the area of the source times. Some capacitance per unit area plus the perimeter of this which is two times W plus Ld times c perimeter and this parameter is usually denoted ps while you describe the mosfet in a simulator. Now, we are of course, looking at regular structure which have a rectangle area, but you could also have other odd shape.

So, then you specify the area and the perimeter directly of the source and drain junction similarly, for the Cdb we will have ad that is the drain junction area times the capacitance per unit area plus the drain junction perimeter times the capacitance per unit perimeter, now this all are detail, but you need to know these things when you want to describe in a simulator. When you module a mosfet in a simulator, you also have to put in appropriate perimeter and area values for the mosfet, so that you module all this capacitances. We later see that all this capacitances will have significant influence on the operation for the circuit. So, it is very important to understand these models and then use them properly.

(Refer Slide Time: 30:57)



So, just to summarize the small signal model of the mos transistor consist of dc parameters which are gm, gds and gmb and capacitances which are Cgd, Cgs, Csb and Cdb between every pair of terminal you have a capacitance except being source and drain directly because they are kind of isolated from each other and that capacitance generally small in most processes and in saturation region the intrinsic value of Cgd is zero, but you will have the overlap capacitance per unit length times the width of the mosfet Cgs will be two third Cs WL plus this times the width mosfet Csb and Cdv are junction capacitance and largely independent of the region of the operation of the MOS transistor.

We will of course, assume that this junction are revised by us, otherwise this capacitance equals to be very different and it is area of the source junction times the source junction capacitance may be explicitly mention, that I will say Cj area plus the perimeter of the source junction times Cj perimeter and this will be area of the drain junction times Cj area because the junction are of the same type they are same Cj area applies to both Cj and Cdb plus pd times Cj perimeter and jm, jds and gmb can be calculated in the saturation region. We can also calculate it in the try out region, but because saturation region is of greater interest what is will use.

(Refer Slide Time: 33:02)

$$I_D = \frac{\mu_n C_{ox}}{2} \cdot \frac{W}{L} (V_{GS} - V_T)^2 (1 + \lambda V_{DS})$$

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu_n C_{ox} \cdot \frac{W}{L} (V_{GS} - V_T)(1 + \lambda V_{DS})$$

$$\simeq \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T)$$

$$= \frac{2 \cdot I_D}{V_{GS} - V_T}$$

$$= \sqrt{2 \cdot \mu_n C_{ox} \frac{W}{L} \cdot I_D}$$

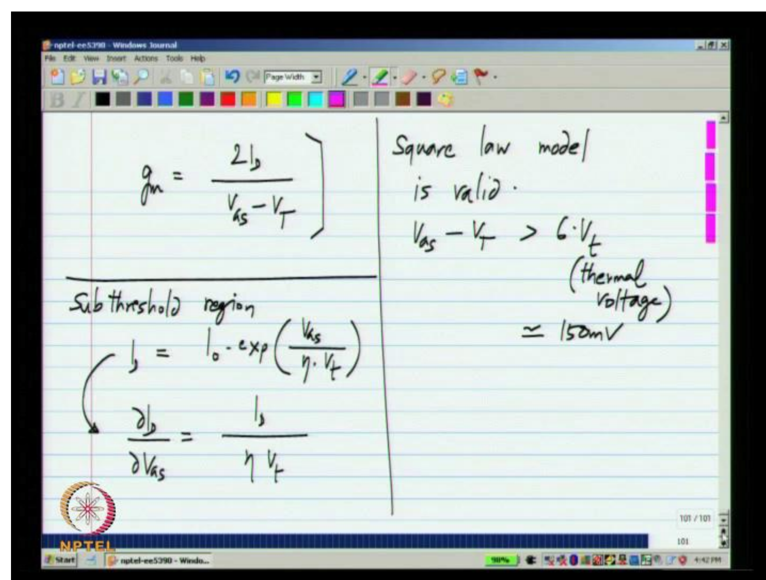
So, again will start from the square law which basically assume a strongly inverted mosfet operating with VDS greater than VGS minus VT and I will do my calculation with an nMOS transistor, but they can be easily be translated to a pMOS transistor, This is the expression for the current and the trans conductance gm is given by dou id by dou VGS which is equal to u one Cox W by L, VGS minus VT, one plus lambda VDS. Now, first of all this lambda VDS is expected to be relatively small quantity compare to one now, this is largely to except on channel length is very short.

Now, for simplicity hand calculation any times this is simply approximated into mu on C mu ox W by L VGS minus VT. Now, if you look at it can also be written as two times the drain current divided by VGS minus VT and also you can neglect this, one plus lambda VDS and write it in terms of only Id two mu on Cox W by L times Id. So, this are all just rearrangement of the same expression and here and there we will omit the one plus lambda VDS, that is we have to replace one plus lambda VDS by just one for simplicity and this is kind of thing that you have do for hand calculations. There is no point including kinds of small detail from the model itself is accurate. When you want really accurate answer you go to stimulator, but it is very important to do hands calculations because you know which circuits has to be modified when it does give you the result you want. No three expressions appear to show different dependence of a gm over the electrical quantities VGS minus VT and Id.

Now, the first once says this proportional to VGS minus VT at the second one says is inversely proportional, but there is no contradiction here. This is the function only VGS minus VT, but this is the function of Id and VGS minus VT. If we change VGS minus for VT for a given transistor Id will also change, hence you have to take that into account. So, do not get confused by this and finally the last one says, it is proportional to square root of Id, but this express only in terms of Id.

So, each of this we will use in the right contest because many times you change one parameter and you would like to find the influence of only that parameter. So, let us say you change only VGS, then the first expression is most convenient and let us say you make some manipulation by which you change only the drain current then the last one is most convenient and the second is also a quite convenient sometimes when you want to get inside into the operation of the circuits

(Refer Slide Time: 36:14)



Now, it is the second one that I want to discuss in a little more detail. It says that gm is two Id by VGS minus VT, now this seems to suggest that if I go on reducing VGS minus VT, while maintaining a constant current and how can I do this? I can go on increasing width of the transistor and reducing VGS minus VT. We know that Id is proportional to W by L and also proportional to VGS minus VT square.

So, I can reduce VGS minus VT and increase W by L and keep current the same and this thing says that as VGS minus VT approaches 0, gm approach to infinity. So, you really

think that will be the case. Some what really happens that the square law model is valid only when VGS minus VT is greater than about 6 VT, where this is the thermal voltage. Now, when VGS approaches close to VT, then the square law model itself is invalid, it is valid only when VGS minus VT is more than about 6 VT or approximately 150 mile volts.

Now, for VGS minus VT value is below this, you cannot use square law model. So, this expression which is derived from the square law model is also invalid. Now, what happens in the sub threshold region Id is given by some I naught exponential VGS divided by eta times the thermal voltage, where eta is some parameter and you have to calculate the gm from this expression and if you do that you will see that it will be Id divided by eta VT.

Now, this expression are in a similar form, the first one can be thought of Id divided by half of VGS minus VT and the second one is Id divided by eta VT. So, what happens is you imagine a case, where that transistor buyer with large VGS minus VT, then you go on increasing the WL and go on reducing the VGS minus VT, while maintaining the current to be constant. This is possible with some combination off decreasing VGS minus VT and increasing W by L.

Now, what happen is as W by L increases large value which is minus reduce the gm value will saturate. We will not go on indefinitely increasing. This happens when transistor enters sub threshold region. So, this is a an example of where you can calculate expression based on same module and then it seems to give you upset result, but it turns out that the model itself is invalid. Now, the other thing this says is that for a given current where you would operate in other regions you would operate to get the maximum gm.

Later we will see that getting a large gm is a very important thing in electronic circuits, gm is something like the gain of the transistor, so getting the large gm is quite important and you can easily see from this expression, that if you want a large gm, we should operate with L small minus VGS minus VTS possible. And in fact push the transistor into the sub threshold region. You use a very large transistor and by a sheet with their relatively small current, then you get the largest value of gm divided by Id that is you can get MOS bank for the meaning, you get the value of gm with the smallest amount of

current or for a given current you get the largest possible gm and this are all things that we will use later when scaling circuits etc. But, again something to be kept in mind because this is how the small signal parameter behaves.

(Refer Slide Time: 40:01)

So, in this I will write one particular form of this which is 2 Id by VGS minus VT. This is in strong inversion or Id divided by eta VT were this VT is different. It is a thermal voltage and this is in weak inversions

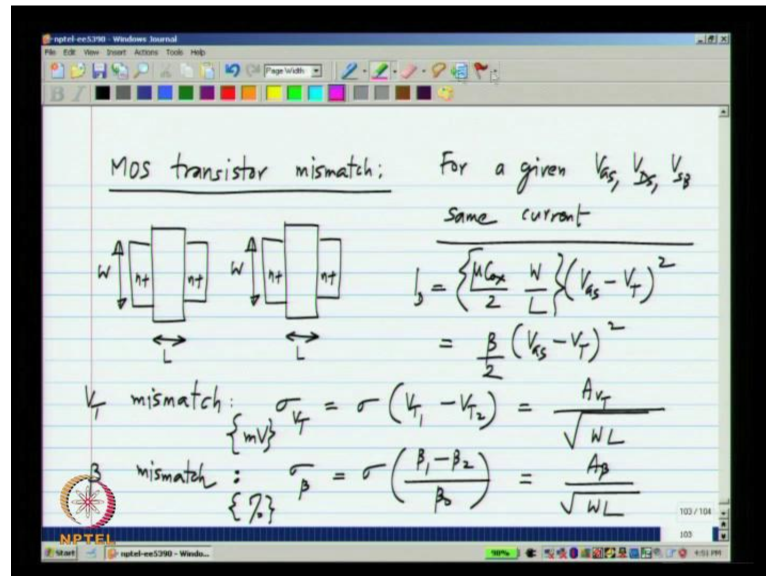
(Refer Slide Time: 40:30)

Gds which is dou Id by dou Vds is given by mu Cox by two W by L, VGS minus VT square times lambda and as before this self is the approximation by current and we will make this lambda Id. So, a good approximation for the value of Gds is lambda times Id and finally Gmb is dou Id by dou VBS. Now, VBS is not explicitly a parent in the expression, but we know that VT is a function of VBS. So, what we have is let me use chain rule dou Id by dou VT times dou VT by dou VBS, which gives me minus mu Cox W by L, VGS minus VT,1 plus lambda VDS times dou VD by dou VDS.

Now because VSB equals minus VBS, I can write this as dou VT by dou VSB and make this a plus sign and you also see that this part of the expression is nothing, but the gm of the transistor. It is gm times dou by dou VSB, so it is something proportional to gm and this dou VT by dou VSB is a number that could be of the order of let us say half to one forth or may be slightly smaller. It is not a very small number nor is it more than one, it is typically less than 1, but may be between half and one fourth or something like that, so this gmb, the trans conductance from the back gate of the bulk can be a significant fraction of gm. In some cases it is actually use for full function in circuit design, but usually the constant that you want to revise by us. The source bulk junction limits that unity of this gmb, all that it can be use in many cases and some case it is an inherent. We will later see some basic circuits, where this leads to some problems.

So, this completes the small signal model and later when we realize trans conductors and opamps and so on. With our MOS transistor, we will use the small signal module to evaluate the characteristics of those blocks. So, far we have discussed the common components like a register and capacitor and we also discuss the MOS transistor in case of resistor and capacitor. We saw that when you make physically identical components, they are nominally identical. They will be a small mismatch between them and that is related to the physical size of the device and exactly the same thing is to the case of MOS transistor as well. Now, we quickly deal with mismatch in MOS transistor.

(Refer Slide Time: 44:13)



So, let us say you make two identical MOS transistor, it will show the top view. We have some W and some L and there is the copy of this which has the same W and same L. So, this two are normally identical, that means if you buy a sheet with given VGS, VDS and VSP, you should get the same current as usual. It turns out the things that are not exactly identical. There are only nominal issues. There is a small difference between the two. Now, do you describe between these two incase of resistor and capacitor, you have a single parameter.

We have the resistances of the resistance we have the capacitance of the capacitance, where in case of mosfet, we seem to have many a parameters, but first of all we will use the square law model in saturation region which says that its mu ox by two by L, VGS minus VT square and this enter multiple factors. Let me call it beta by two VGS minus VT square, Where beta is nothing but mu Cox times W by L.

So, now there are really two parameters beta and VT. So, I will describe the mismatch and beta and VT. What we do is let us say, we make a ten thousand pairs of this transistor and measure the different in beta and measure the different in VT for each pair of transistors and somehow describe it and also related to the nominal value of beta and VT. So, the VT mismatch is described by the standard deviation of VT and in case of VT it is not normalizing to the nominal value VT, this is because nominal value VT could be 0 or a very small. It does not make a sense to normalize to nominal value of VT.

We will see that VT is nothing, but just the sigma of VT one minus VT two for each pair of transistor and this is the quantity that has dimensions of mini volts unlike the mismatch and resistor and capacitor which was that you mismatch and were dimension less is ok. And this is given by some constant that depends on some process divided by the area of the MOS transistor exactly the same way that resistor and capacitance mismatch depends on the area and data mismatch or the current factor mismatch.

This has some nominal values in multiplying this, so this is really the sigma of beta 1 minus beta 2, that is difference in beta of identical pair of transistor divided by the nominal value a 0 and this is given by a beta by square root of WL sigma VT, something that measure in mile volts and sigma beta is this something measure in percent. Sigma beta is analog to the numbers of sigma r and sigma c. So, how do we calculate the effect of this mismatches we will see that in next lecture?

Thank you