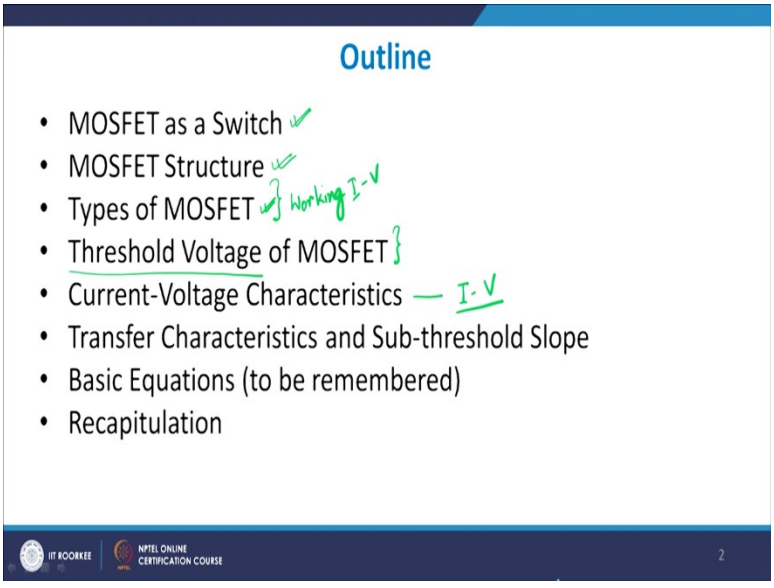
**Microelectronics: Devices to Circuits**
**Prof. Sudeb Dasgupta**
**Department of Electronics and Communication Engineering**
**Indian Institute of Technology Guwahati**
**Module 3**
**Lecture 14**
**MOS Transistor Basics-I**

Welcome to the NPTEL online course on Microelectronics: Devices to Circuits. In our previous interactions and modules, we have seen the concepts of bipolar technology or BJT, bipolar junction transistor and we also saw the various characteristics, the structure of the transistor, the types of transistor, the modes of operation of the transistor and then we saw some of its circuits implication in terms of common emitter, common base and common collector. We also saw the second order effects of bipolar transistor which were prevalent at certain conditions of bias or structure and that took care of our basic understanding of bipolar transistors right.

In our subsequent slides or subsequent interactions, we will be looking into what is known as a CMOS technology, or a complementary metal oxide semiconductor technology. So to understand that or CMOS technology, we need to first of all understand the basic concept of a transistor or a metal oxide semiconductor field effect transistor which is basically also acronym-ed as MOSFET, M-O-S-F-E-T. Right? So let us start with this lecture on MOS transistor basics one. So the lecture topic is basically MOS transistor and we start with basics one right?

(Refer Slide Time: 1:59)



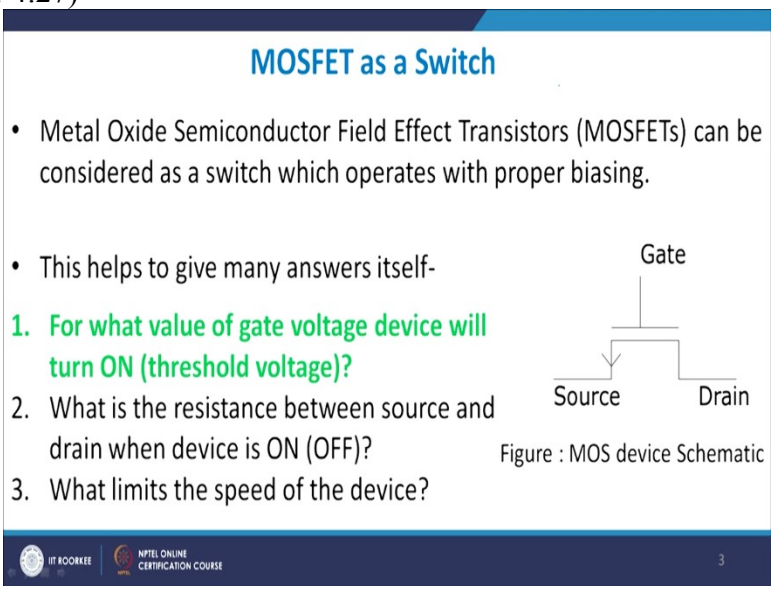
The outline of the talk is that we will be looking at MOSFET as a switch first of all, right? We will be looking at a switch, MOSFET as a switch, right? And then we will be also looking at the MOSFET structure, right? So the idea here is that we need to first of all look into the fact that why we migrated from a bipolar technology which is basically a technology in which both electrons and holes are majority current carriers and results in overall current, to a technology which is basically a unipolar or only one of the electrons, either electron or holes are the majority current carriers and they are responsible for the current. Right?

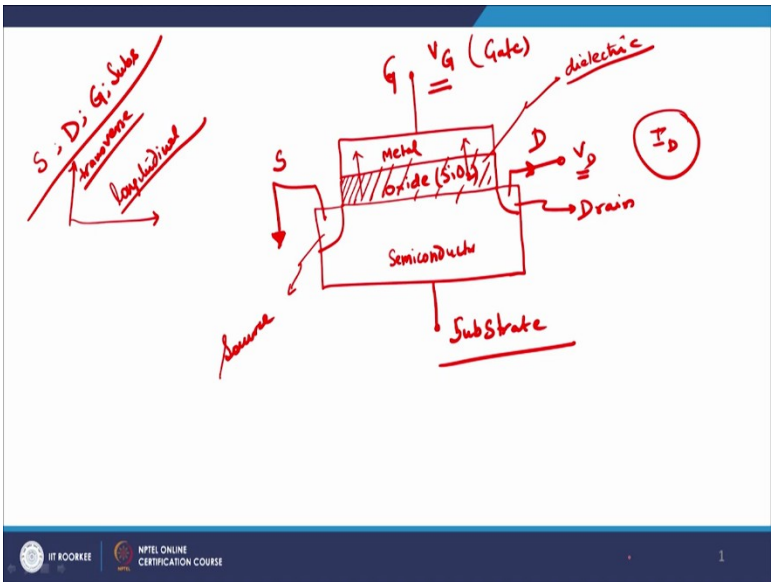
And after that, we will be looking at types of MOSFET. So once you have understood the structure, we will be (understood) understanding the types of MOSFETs and their working behaviour. So how they are working actually in terms of real IV characteristics, right? Then we will be looking into one important property of MOSFET known as the threshold voltage of MOSFET right? A very important property from the point of view of both, analog as well as digital electronics and we will see later on that this voltage is quite critical to the operation of a MOS device. We will be also looking into the current versus voltage.

So current versus voltage characteristics, which means that given a set of biases, can I predict the output current of a MOSFET? Right? Now these set of biases can be a variable quantity, can be a fixed quantity. So I can have a DC analysis in which possibly I fix the bias voltages at various points of the MOSFET and then see how much current the MOSFET is carrying, right? Or I can also have something like this that I vary the voltage and see how much variation in the current is visible to me.

We will be also looking at transfer characteristics and sub threshold slope. This is quite important from the point of view of switching. So, whenever we want to use it for high switching purposes or the purpose of high-frequency applications we need to understand what is the meaning of sub threshold slope and what is transfer characteristics. So after that we will be looking into the basic equations which will be helpful for finding out the current and then we will recapitulate the MOS device as such right. So this is the basic outline of the talk.

(Refer Slide Time: 4:27)




Let me come to you as the basic idea here. So before I even go to the MOSFET as a switch, let me give you a basic diagram of a MOS device. You must be aware from your other sources but just to give a continuity in whole idea, this is how a MOSFET primarily looks like. Right? Since I am making it as a free hand drawing, it is exactly not like this. So I have got metal here. This is a metal right? I have an oxide here which is primarily silicon dioxide in most of the cases, silicon dioxide and I have got a semiconductor here, right? I have a semiconductor here.

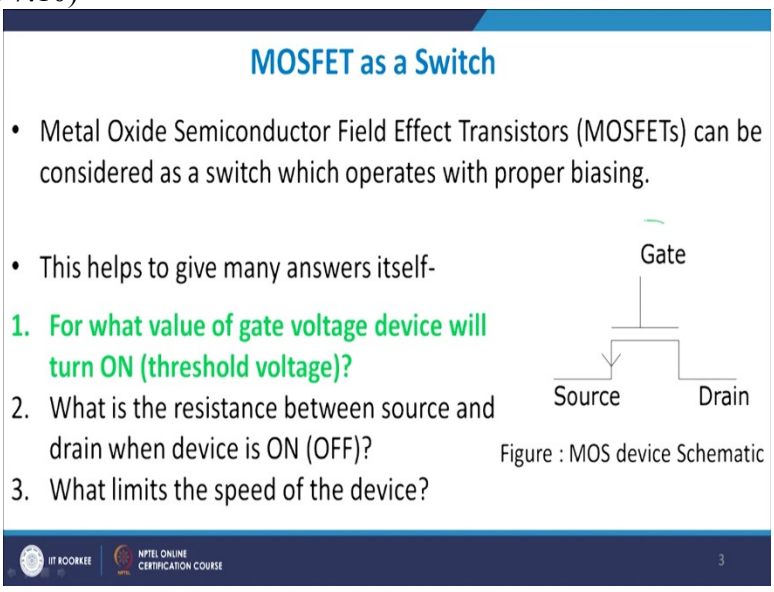
And this is basically my, let us suppose, source, this is termed as a source. Right? And we have a drain. So this is your drain. So I have a source here, I have a source and a drain here and I have a

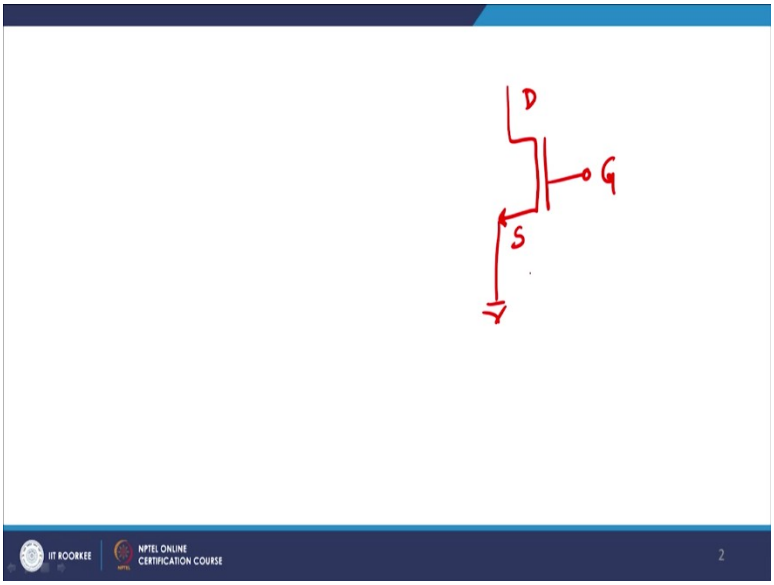
semiconductor here and this side is referred to as VG or also referred to as gate. So I have got, so basically if you look very carefully, it is basically a three terminal device at this stage. One is source, another is drain, another is gate, right? You also have fourth terminal which if possible we will come later on and this is known as substrate terminal.

So typically a MOS device has got four terminal device- source, drain, gate and substrate, right? It is a four terminal device. Typically, for all practical purposes, my source is grounded. That is the reason I have shown source to be grounded here which means that the potential on the source side is approximately equals to zero or exactly equals to zero. We give a bias on the gate side and drain side and then we try to find out the current through the drain side right. So $I_D$ is the current through the drain side which means that the current is flowing through the drain side. Right?

Now you see, since oxide is there, this is oxide layer right? So oxide is basically a dielectric. You remember, it is basically a dielectric and as a result, you won't expect to see any current flowing in the transverse direction. So we refer this as longitudinal direction, longitudinal right? And we refer to this as transverse direction. This is a typical parlance which we follow or you can call it this one vertical and this one horizontal. But in typical device physics, we convert this to be as longitudinal and this to be as transversal.. So, perpendicular to the silicon interface is basically my transversal and this is the longitudinal direction which you see.

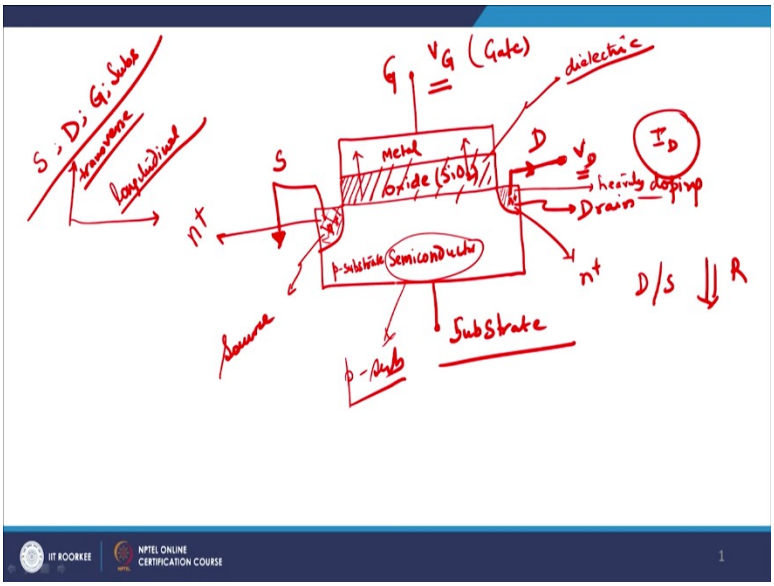
(Refer Slide Time: 7:10)






Now if you come back to, come back to its, to the operation, the way it is being transferred or written is something like this. So the methodology which we write is gate, source and drain right? So the method, so what people follow is that, if you have a transistor, then we refer to the transistor as this. This means that, this is your gate. Right? Gate. And we refer to this as source and drain. So the arrowhead right refers to the direction of the movement of holes, not electrons. Gate is there and drain is there. Now you see, we will come back to the previous slide here and see what works, how a device is turned on.

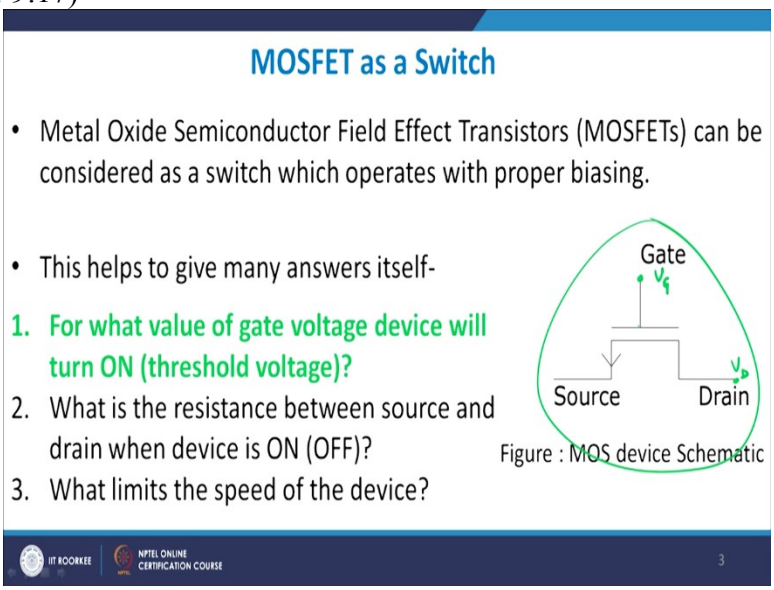
(Refer Slide Time: 7:54)




Generally, this is basically made $N^+$ and $N^+$, which means that this is practically very heavily doped, it is heavily doped right? Both sides, source and drain are very heavily doped. Heavily doped, therefore means that the number of electrons here inside the drain region and the source region are very very large as compared to the nearby regions, right? So they are very heavily doped. So therefore, since they are very heavily doped from a basic device physics, we will understand that the resistance offered by the drain and source is very very low.

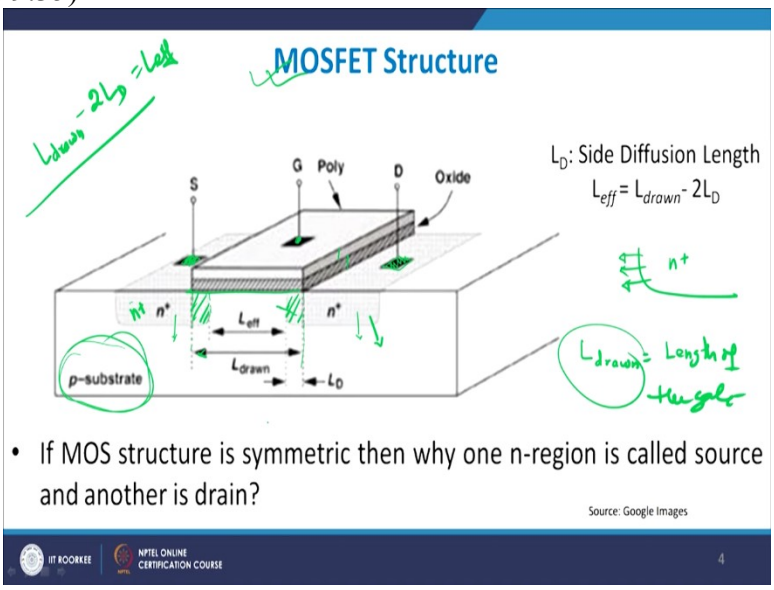
So there is no resistance or there is no potential drop inside source and drain because now you have very large number of electrons available. So the resistance is fallen down and as a result, there is no potential drop within the source or drain. But surely, you can apply a bias from an external world by applying a $V_D$ voltage and we will see how it works out, right? Let us see. So if I have this one as $N^+$ right, I have this as $N^+$. Right? The substrate which is basically my semiconductor, I use it as P type substrate. Right? I use a P type substrate, which means that typically this semiconductor is basically P type right? It is a P type substrate and I have got a P type substrate at this, at this point.

(Refer Slide Time: 9:17)




So you see this drain, you will apply a potential here, here you apply a potential $V_D$, here you apply a signal or a potential $V_G$ and generally source is grounded right. So 1, 2, 3 that will, we are not looking at substrate nowadays at this point. But we will see how it works out.

(Refer Slide Time: 9:35)




Now if you look from the basic concept, we will, if you look at the MOSFET structure now, it looks exactly like what you see in front of you. Right? So I have a P type substrate. This is a substrate. Right? This is my source. This is, this black color which you see is the metal contact. So a metal contact through which you are contacting the external world. This is my

semiconductor, heavily doped. Semiconductor, heavily doped. And then again, this black contact is my metal contact which is also metal contact here. This is my oxide layer, this is the oxide layer which you see and then the white one is basically my metal layer, right? Also referred to as poly layer and this is basically my gate junction.

So I have my gate junction, source junction, drain junction and I have got a P type substrate at this particular point. Right? Is it okay? Let us see, how we define channel length before we move forward any further. Now generally what happens is that whenever you draw, whenever you put $N^+$, for example, you have put $N^+$ region here as well as $N^+$ region here, right? Is it okay? Now obviously, because of temperature variations, because of large density potential gradient, because of large gradient of the charge carriers, some of these $N^+$ regions will shift to the left.

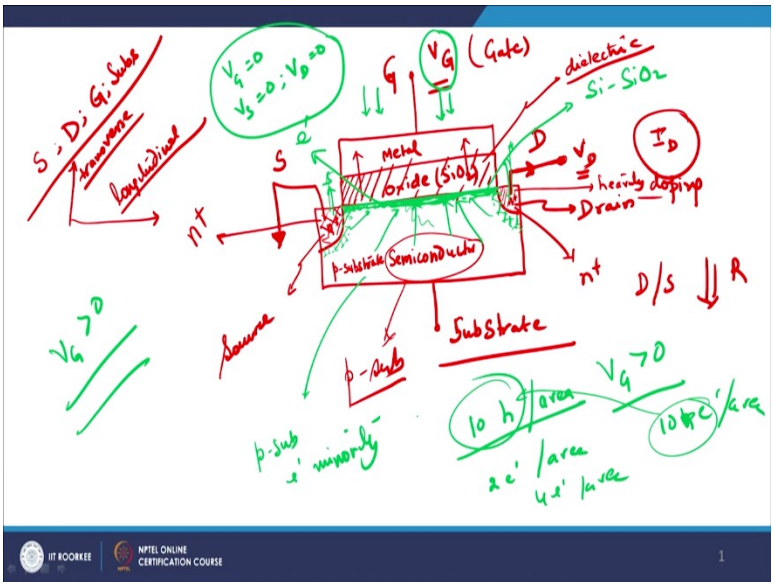
So you wanted to, you diffused it up to this much point or you doped to this much point. After some amount of time, this will actually start shifting on this direction because of sudden diffusion and as a result, you initially wanted it to be till here but you ended up till here, right? Right? Is it okay? You ended up here. Similarly, you wanted it till here, till source region, till the edge of the gate, but you ended up somewhere here. So we define 2 quantities here.

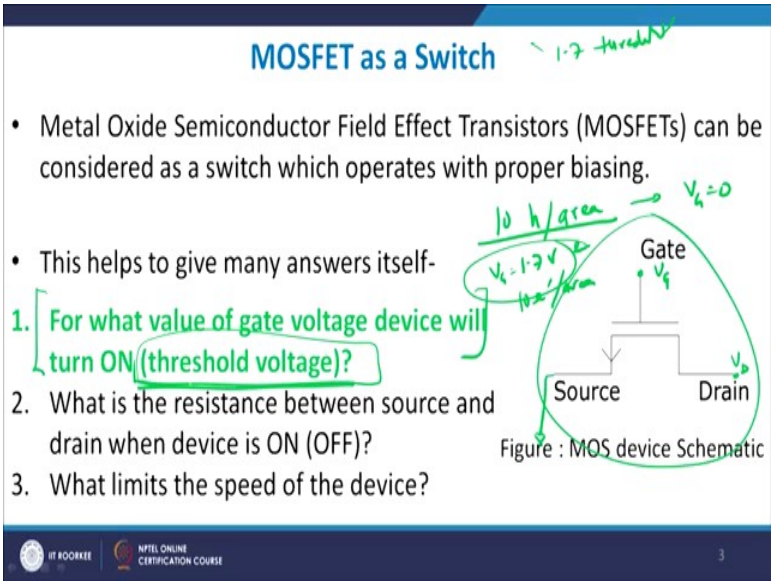
Physical drawn length which is L drawn. L drawn is a physical drawn length. Physical drawn length is nothing but the length of the gate. So if a gate length has 20 nanometer then this basically, this one is basically a twenty nanometer from this end to this end. But what has happened in the meantime is because of doping been laterally shifted, because of temperature variations and all these things, the effective channel length has actually been reduced from L effective to L drawn from L drawn which is basically my drawn length, minus two times $L_D$. And why two times? Because one on this side and one on this side. Right?

So your effective length is always smaller than your drawn length or a physical length. Fine? And this has to do with the doping or the diffusion being laterally moving across the frame. So you had a diffusion here or you had a large number of charge carriers here. It started to diffuse laterally in this direction because there is a gradient of charge. It stops at a particular point. When the excess amount of, assuming that these two are perfectly symmetrical under all circumstances, I assume that my L effective will be equals to L drawn minus $2L_D$. Right? L effective will be equals to L drawn minus 2L effective right?

And is this what you get what you get from it. So from this basic idea, let me see how can I work it as a switch, does it work as a switch.

(Refer Slide Time: 13:00)




Now you see, in our previous discussion when we were taking up this basic concept of MOS device, if you look very carefully, if you do not apply any gate voltage right so I have P type substrate here. Right? So P type substrate will have electrons as minority current carriers, we have already discussed this point. Howsoever small, they will have some amount of electron

available here. Also, this source and this source will have large amount of free electrons because it is a source and it is a drain. As a result, this forms a PN junction diode. So under no bias condition, there will be a depletion region here, assuming that my $V_G$ equals to 0, $V_S$ is of course equals to 0 and $V_D$ equals to 0 and $V_{substrate}$ is of course equals 0.

For all these three conditions of these things, I would sufficiently assume that the depletion region is exactly equal on the left and right side which means that the doping concentration on the source and drain are exactly equal. If the doping concentrations are same, you will see the depletion region to be exactly equal. Fine? So we have what is the current situation therefore? That this is basically P type substrate, this is P type substrate. I have got a depletion region here which is basically devoid of any free charge carriers, similarly I have a depletion region here devoid of free charge carriers and I have got a high $N^+$ region here and a high doping $N^+$ region here known as drain.

Now what I do is, this is quite interesting and this is where we can actually do quite amount of change. What we try to do is, we try to make our $V_G$ now greater than 0 right? And that is quite important. As you make $V_G$ greater than 0, there will be electric field in this direction, of course because of course the as reason suggest and as a result, large number of electrons from here, from source right, from drain and from substrate, minority current carriers, will all rush towards the interface, will all rush to the interface.

So please understand, that whenever my gate voltage therefore rises up, at a certain gate voltage, large number of charge carriers come near the silicon-silicon dioxide. So this is, this interface which you see in front of you is basically silicon -silicon dioxide interface. So what will happen is there will be a large number of electrons which are already stationed within the source and drain region and because of this external voltage applied on the gate side, a positive voltage, $V_G$ greater than 0, for $V_G$ greater than 0 right, they will start populating the electrons near the silicon-silicon dioxide interface.

So you have a large density of electrons here. Right? Large density of electrons. So this will be typically large density of electrons. Now these electrons are coming from where? Coming from source, and drain and the substrate right? So initially those P type, now let us suppose I am

throwing as, I am increasing the gate voltage and I am making it more and more electrons, more and more are coming here. Right? If the number of electrons per unit volume you can take or per unit area near the silicon-silicon dioxide interface, is exactly equal to the number of holes per unit area initially kept right, so suppose initially we had ten number of holes per unit area, now what we will do?

We will now increase the gate voltage and they will make you say, say there are two electrons per unit area. Right? Increase further. I get four electrons per unit area. Right? I go on increasing further till I reached to ten electrons per unit area. Agreed? Once this happens, it was happening earlier also, these ten electrons will have a chance of recombining with these ten holes and as a result, there will be no holes available with me. Now if we increase the gate voltage slightly ahead, larger than this, you will have excess of electrons there, rather than holes because whatever hole was there has been actually accommodated by virtue of recombination and as a result, that area is devoid of any free charge carriers.

Free charge carriers, they are not devoid of charge carriers but they are devoid of free charge carriers. Now what we do is that we still further increase the value of $V_G$. Then you will have now the population of electron getting more and more because recombination has ensured that the hole strength is almost zero. So, if there were ten holes available near silicon-silicon dioxide, electrons, ten electrons goes there and recombine with holes and the holes finish, right? Now if you go and increase the value of $V_G$, you will be having the excess of electrons, right?

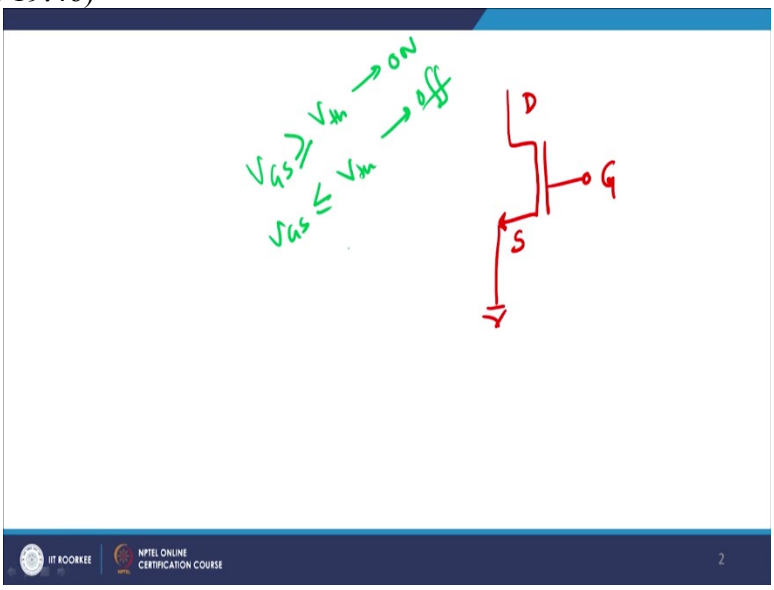
So we define a new term known as threshold voltage where we define, this, if you look at which is defined as threshold voltage, right? How do we define threshold voltage? That gate voltage at which the number of electrons per unit area or volume is exactly equals to as many number of holes were there per unit volume earlier, we define that voltage, gate voltage to be as threshold voltage. So there were let us suppose ten holes per unit area near silicon-silicon dioxide interface right, when your $V_G$ was equals to zero, now you apply $V_G$ equals to 1.7 volts and the number of electrons are ten electron per unit area near silicon-silicon dioxide, then we define $V_G$ equals to 1.7 as the threshold voltage of the device.

So how do you define threshold voltage? Threshold voltage of the device is that voltage at which it becomes as much N type or P type as it was earlier P type or N type respectively. Fine? You

got the point. And that is very very important. And therefore quite critically, the threshold voltage is also referred to as the On voltage of the MOS device. So this is the voltage when the device gets on, switched on. Right? Because you know, a large number of free charge carriers which can take part in conduction.

So On means it does not necessarily mean a current but it surely means that you will have large number of charge carriers near silicon-silicon dioxide. Now you have to apply a drain voltage in order to sweep the current and form a, sweep the charge carriers form the current. Right? So that is what I am saying? In green what is written here. So what value of gate voltage will it turn on? So the value is basically my threshold voltage, right? At threshold voltage, it will turn on the device. So if I want to switch my MOS device from on to off stage and vice versa, I just need to put the gate voltage above threshold voltage and I am, so what I do is I do something like this.

(Refer Slide Time: 19:46)

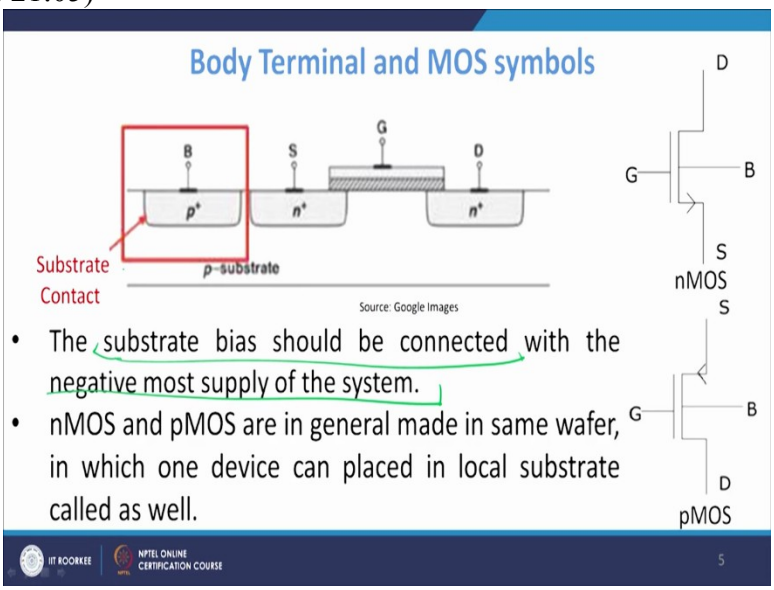



That I apply V gate to source voltage as greater than equals to threshold voltage of the device and I switch it on. If I apply $V_{GS}$ less than, equal to threshold voltage, I define it to be as a Off state. Fine? So threshold voltage is quite an important term in analog as well as in digital because it gives me an idea about how it works out. Na. Now so what is the reason between source and drain when the device is on? Well, the resistance will be very small because now you have large number of free charge carriers between source and drain.

And therefore, the resistance offered by the channel will be very small whereas when the device is in the off state, the resistance offered by the channel will be relatively very high. Right? And that is also one of the methodology by which you will know whether the device is switched on or off. So if the resistance of the channel, offered by the channel is very very high, pretty high, then you can assume it to be as a as a your if it is very high, which means that there is the drain and source are removed from each other, it is off state.

If it is low, it is in on state, right. So it is now understood and we have understood the basic feature of two terminal device. Right?

(Refer Slide Time: 21:05)

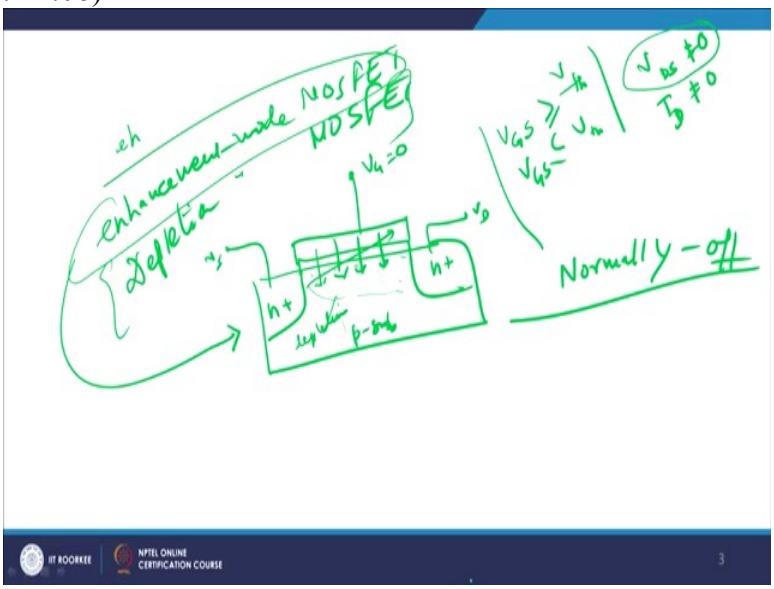



Typically the rule of thumb is, that the substrate should be always connected to the most negative terminal of the supply. So negative most terminal of the system should be connected to the substrate bias. And we will see later on why, but you can find it out yourself from many other books because once you do that, when you keep it to the negative most terminal of the battery, you don't therefore manipulate the threshold voltage once it is already been fixed by the external biases. Right? And if therefore since nMOS and pMOS are made from the same wafer, one device can be placed in local substrate called a well.

So I have a well structure, this is the well which you see in front of you. In the well, we have an $N^+$ region. So this is my nMOS and your substrate contact here which is there with me and this

is the fourth contact which is generally used for, for the forthcoming for the purpose of contacts here, right? Let me therefore come to the basic issue of of…, so let me understand, let me explain to you therefore what are the various, what are the various issues involved in MOSFET.

(Refer Slide Time: 22:08)




The first issue which we have seen is that if my gate voltage is greater than equals to threshold, the device is on and if it is less than equals to threshold, it is off, right? And it is true also that the devices are on and off but for the current to flow, you require to have $V_{DS}$. So $V_{DS}$ if it should be non-equals to 0, then only $I_D$ will be equals to non-equals to 0. Right? If $V_{DS}$ is 0, then you will not be able to have any current flow between, though the device will be in on because $V_G$ is greater than $V_{TH}$ but since $V_{DS}$ is not equals to 0 or $V_{DS}$ is equal to 0, let us suppose that $I_D$ equals to 0, primarily meaning that the device is not in the on state.

Although, it is not at least carrying the current from point A to point B. Right? So this is what we get. Let me come to the type of MOSFETs here now. And there are 2 types of MOSFETs which is with us. And one is known as an enhancement, sorry enhancement mode MOSFET. Right? And we have got a depletion mode MOSFET. Right? We will first discuss enhancement mode and then we will come to depletion mode. Fine? Let us look at enhancement mode MOSFET.

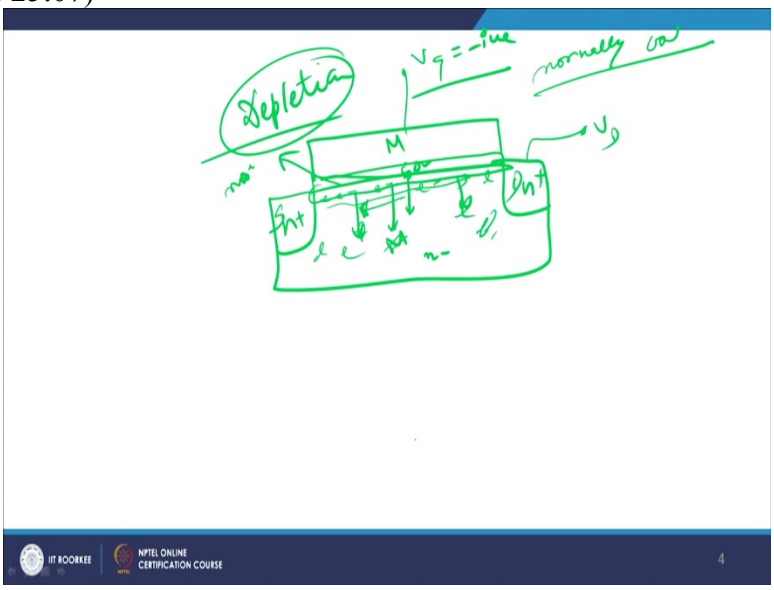
In the enhancement mode MOSFET, as the name suggests, you have to give an external potential on the gate side to enhance the number of charge carriers near the silicon-silicon dioxide

interface in order to form the channel, which means that if I have an $N^+ N^+$ region here and I have got a hole and, sorry, I have got a dielectric and a, sorry, a gate here and drain here right, drain here and this is source right and then I join these two together for understanding purposes, then if we do not apply any gate bias, $V_G$ equals to 0, this is P type substrate, nothing will happen and the reason is since there is no electric field in this direction there is no electric field, right?

There is no electric field. Then no charge carriers will be moving towards silicon-silicon dioxide interface. Right? So there will be no charge carriers. Right? And therefore this will be all this will be all depletion region. All will be depletion region. Right? All will be depletion region. Fine? Now what happens, we go on increasing $V_G$ and we are pulling the electrons towards the interface and thereby enhancing the charge carriers because it causes threshold voltage, you say that the device is fully on.

So what is an enhancement mode MOSFET? Enhancement mode MOSFET is a device which enhances the charge carriers near silicon-silicon dioxide thereby making it on, right and therefore it is also referred to as a normally off device. It is a normally off device means, normally off means, if we did not apply any bias, it will be off state, right because there is no channel charge formation taking place.

(Refer Slide Time: 25:07)




Let me come to therefore the depletion mode MOSFET. In the depletion mode MOSFET, I assume that these are all $N^+$, exactly the same as the previous case, right? And everything else

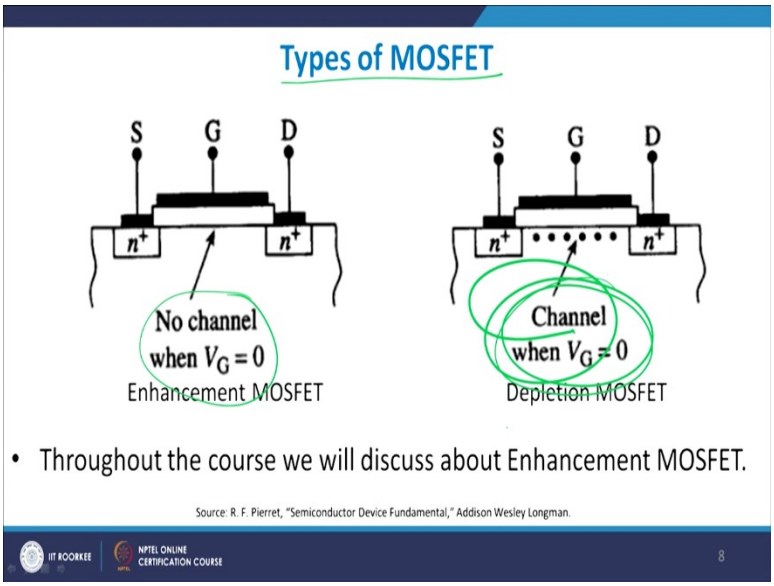
remains the same. I have an oxide layer, I have a, sorry, a metal layer, I have a silicon dioxide and a metal here. Right? Now, so we are discussing depletion mode design. What happens in this case is that, in this case let us suppose this $N^+$ $N^+$ and rather than substrate to be P type, let me make it N type. Then all will be electrons here.

All will be electrons because it is a majority current carriers. Now if we apply a drain bias $V_D$, all the electrons will be moving in this direction and therefore there will be a current $I_D$ flowing in this direction. Right? This is known as depletion mode. Why? And the reason is, this is also known as normally on. Normally on, why? Because even when you do not apply any gate bias, $V_G$, there will be large number of charge carriers between source and drain. Right? Effectively, the number of carriers will be very large.

And if we do not apply any gate bias, then all these charge carriers are already near Silicon-silicon dioxide interface and therefore they are helping you to form a current between source and drain. So how will you stop it? You have to give a gate voltage sources, such that, so it is an N type, that it should be negative in dimensions or negative in nature. If it is negative, all the electrons will be pushed backwards and therefore this whole thing will be left to the main issue right. So that will be there. Which means that if I apply a gate voltage which is negative, it will push all the electrons downwards and all the holes will come up and it will be going into off state.

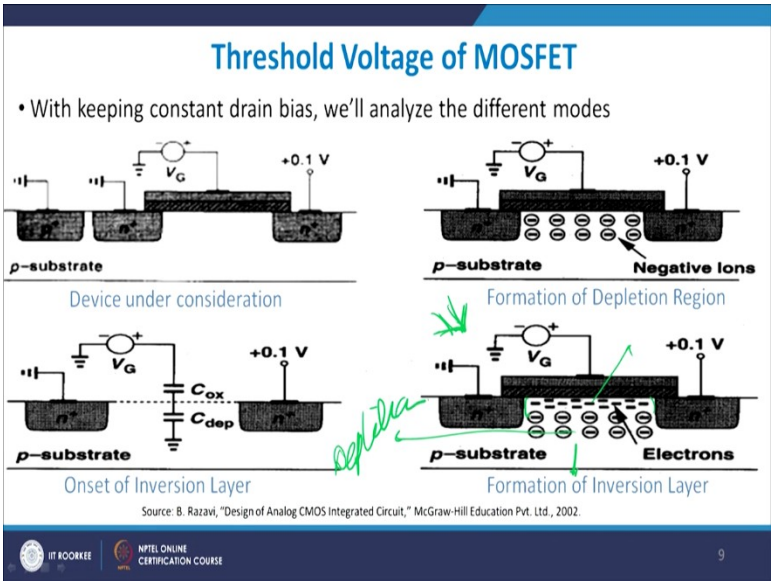
So I have got normally on and normally off device. Normally on device is depletion mode, normally off device is basically your enhancement mode structure, right? So this we have learned and this is what you see.

(Refer Slide Time: 27:02)




No channel when $V_G$ equals to 0 for enhancement mode MOSFET. For depletion mode MOSFET, when $V_G$ equals to 0, you still have some channel being formed here right? And that is quite an interesting phenomena as far as types of MOSFETs are concerned.

(Refer Slide Time: 27:14)



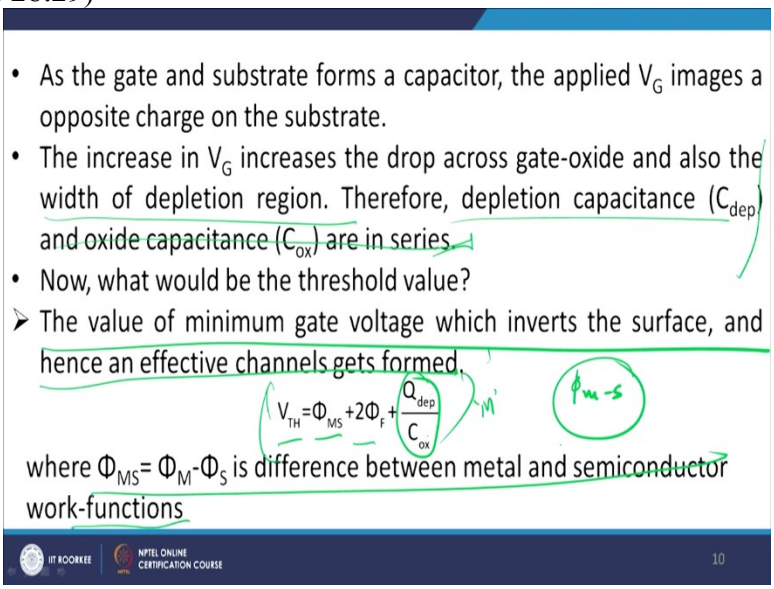

We have also seen therefore how do you define threshold voltage of the MOSFET and we see that threshold voltage is defined as that gate to source voltage at which the substrate becomes as

much N type as it was initially P type or becomes as much P type as it was initially N type, either of the 2 cases right. And you are able to solve the problem of MOSFET relatively easily. As you can see here from the fourth diagram here, fourth diagram right, the fourth diagram, as my gate voltage becomes more and more positive, it pulls electrons near.

So these are the electrons. And all my negative donor concentrations have been shifted downwards. Why? Because these are atoms right? And atoms have typically large, very small mobility and large scattering. Right? And as a result, these atoms do not get enough time to move towards silicon-silicon dioxide interface. Before even these starts to move, the electrons are very fast and they move towards the silicon-silicon dioxide interface. So that is the reason.

So you have an inversion layer here and you have a depletion layer here. Right? And that is the reason why you get a normally on device. Right?

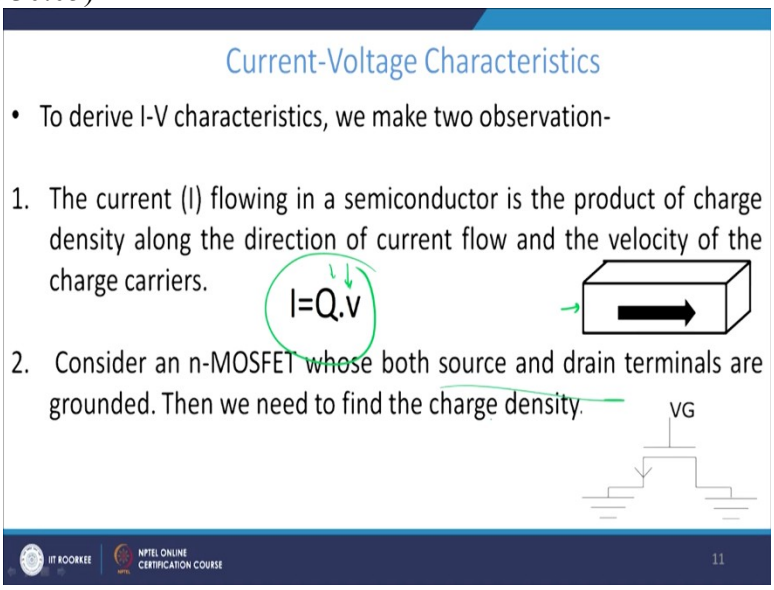
(Refer Slide Time: 28:29)




That is what I was writing here, that as we increase the value of $V_G$, the drop across the gate oxide and also the width of the depletion region increases. Therefore depletion capacitance and oxide capacitance are in series. Right? So we define threshold voltage as therefore $V_{TH}$ equals to $Phi_{MS}$ plus 2 $Phi_F$ by Q depletion by $C_{ox}$, where $Phi_{MS}$ is the metal semiconductor what function, right? $Phi_F$ is the Fermi potential and Q depletion upon C oxide is basically the potential which is basically the charge. Charge by potential is basically my potential which is available to me.

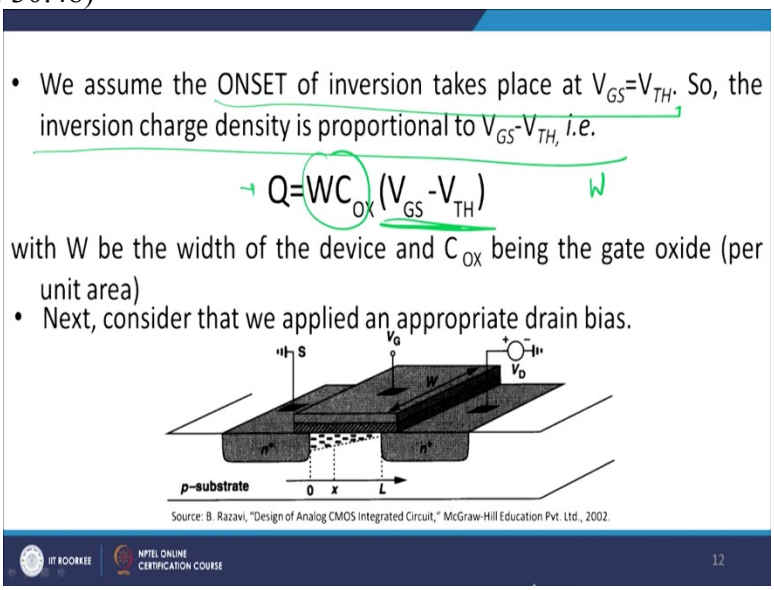
Now, so the minimum value of the gate voltage which inverts the surface and hence an effective channel gets formed, is defined as a threshold voltage. Fine? $Phi_{MS}$ is the work function difference between metal and semiconductor and that is placed here, right? So you will have typical, some value of Phi and M. What is $Phi_M$ because metal and semiconductors do not have the same bandgap and the same intrinsic carrier concentrations, so they are shifted in space domain by a large quantity. Right, so if we can find out, if we can bring somehow or the other, these two together in one domain, we will be able to distinguish or understand $Phi_{MS}$ which is basically the difference of $Phi_M$ and $Phi_S$, where $Phi_M$ is basically the work function of metal and $Phi_S$ is the work function of silicon right?

(Refer Slide Time: 30:05)




So with this, let me come to the current voltage characteristics of MOS device, right? The current voltage characteristics of the MOS device can be looked into the fact that if we apply a voltage here, then current will be equals to charge multiplied by the velocity of the charge carriers, Q into V. Suppose I assume that both my source and drain are grounded, then we require a charge density, which means that if my source and drain are grounded, either of the two are grounded, there will be no chance of movement of charge carriers in the longitudinal direction and you will not be able to explain the charge carrier formation, right? So you require to have some other technique to do it. When you have, so the onset of inversion takes place when $V_{GS}$ equals to $V_{TH}$ right?

(Refer Slide Time: 30:48)

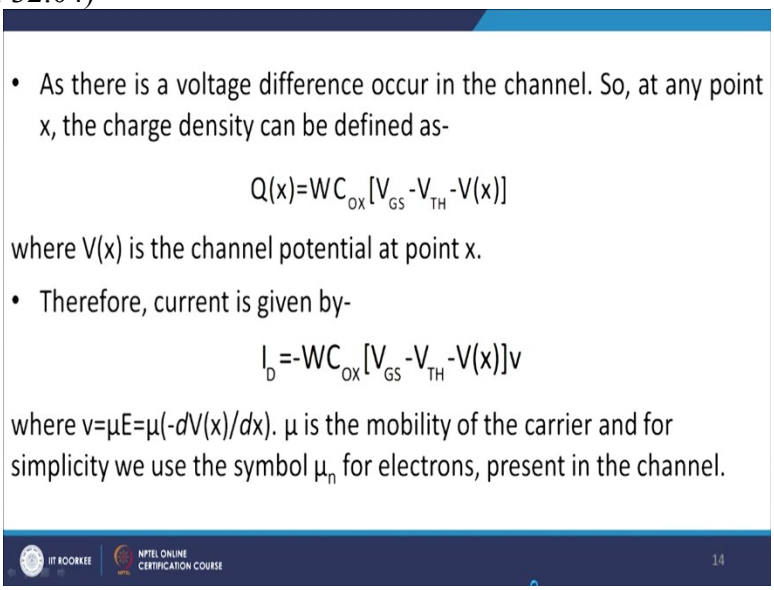


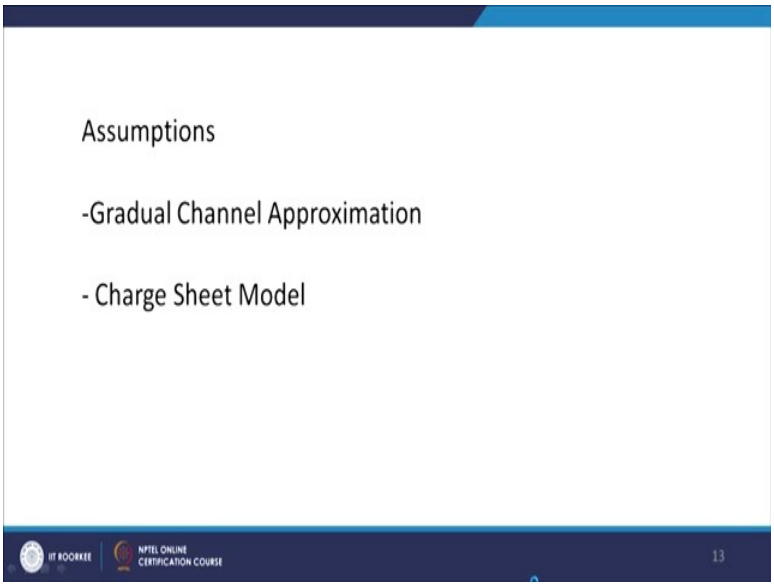

So the inversion charge density is proportional to $V_{GS}$ - $V_{TH}$. Why is it like that? The reason is, say threshold voltage is the gate voltage at which it becomes on right? Now if your $V_{GS}$, which is gate to source voltage is higher than the threshold voltage, the amount by which it is higher right? If that amount is typically very large, that will allow you to basically make the, make it much better. The device to be much better shape, right? So what we do is, at the onset of inversion, $V_{GS}$ equals to $V_{TH}$.

But as the student proceeds further, he sees that the total inversion charge density, right, is directly proportional to $V_{GS}$ - $V_{TH}$. If that is value is larger, you will be able to do anything of this which what is what? Therefore we can say that charge Q is basically W into $C_{ox}$ into $V_{GS}$ minus $V_{TH}$ right? And this is the charge there. W is the width and C oxide is the gate oxide capacitance per unit area and W is the width here. Right? Let us suppose we have applied now a appropriate drain bias.

So I have a drain here. Drain is applied with $V_D$, source is grounded and on the gate side, we again have not applied at this stage anything.

(Refer Slide Time: 32:04)




So what are the assumptions we will look? We will look at the gradual channel approximation and charge sheet model and this will be taken care of in the next lecture. Thank you very much.