Digital Integrated Circuits
Dr.Amitava Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture -37

GaAs MESFET
Characteristics and equivalent circuit

Till now in this course we have been discussing silicon based devices. That is bipolar logic families as well as MOS based logic families and different circuits. When we talked about bipolar logic families, we also made reference to hetero junction bipolar transistors which were made using compound semiconductors but in those devices, the circuits were not modified by the introduction of gallium arsenide devices in the sense that in the same circuits one could replace the silicon bipolar transistor with a hetero junction bipolar transistor. The circuit did not change. In today's class we are going to discuss some circuits which are specific to a certain devices made using gallium arsenide.

Before we actually go into the devices let us look at some of the properties of silicon and gallium arsenide and see why we would prefer gallium arsenide for these applications and also what are the device or which device we take for this type of logic circuits. Let us look at the properties. The properties are if you take a gallium arsenide, silicon. First let us look at the band gap. For gallium arsenide the band gap is 1.42 electron volt and it is a direct band gap material whereas for silicon it is 1.1 electron volt and it is an indirect band gap material.

(Refer Slide Time: 03:15)

The direct band gap material it means that the conduction band minima and the valence band maxima are both at the same momentum.

One can use it for optoelectronic devices. Gallium arsenide can be used for optoelectronic devices but since silicon is an indirect band gap material, it cannot be used for optoelectronic devices. That is a disadvantage of silicon. Nowadays lot of optoelectronic integrated circuits are being fabricated where you have both the signal processing circuitry as well as the optical sensor in the same chip. For that purpose gallium arsenide would obviously be an advantage. The band gap of gallium arsenide is higher than that of silicon. One advantage of that is gallium arsenide can be used at higher temperatures which is an advantage for gallium arsenide with respect to silicon.

Of course one disadvantage is if you have a higher band gap material, it is difficult to make ohmic contacts but that has been overcome through improved technology but the most important difference between gallium arsenide and silicon is in the nature of their drift velocity electric field characteristics.

(Refer Slide Time: 05:29)
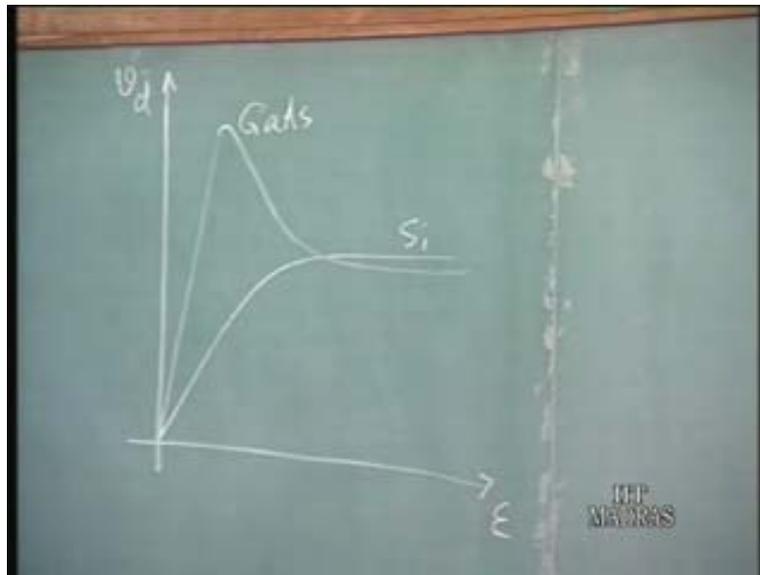
If one looks at this velocity versus electric field characteristic for electrons in gallium arsenide and silicon, we will observe that in silicon the drift velocity for electrons follows the curve something like this. Whereas for gallium arsenide the nature of the curve is something like this which means that the slope is higher initially. This is for gallium arsenide, this is for silicon, which means that the low field mobilities in gallium arsenide is higher than that in silicon. We will write the low field mobility in this table.

(Refer Slide Time: 07:10)

For gallium arsenide the low field mobility of electrons is around say 5000 centimeter square per volt second for N_D is equal to 10 to power 17 per centimeter cube. It can go right up to 8000 centimeters square per volt second for undoped gallium arsenide where as if you look at the silicon mobilities it would be around 800 centimeter square per volt second for N_D equal to 10 to the power 17 per centimeter cube. We see a distinct advantage for using gallium arsenide that is it has got a much higher low field electron mobility compared to silicon.

I have also put down the mobility for undoped samples because we also have devices like the high electron mobility transistor using gallium arsenide where the electrons actually flow in the undoped part of gallium arsenide giving rise to such high mobilities. That is why it's called a high electron mobility transistor HEMT. We see that this is the advantage which is utilized in making very fast logic circuits, the low field mobility. But what about the saturation velocity of electrons? For both of them it is of the order of 10 to power 7 centimeter per second, this is also of the order of 10 to power 7 centimeter per second.
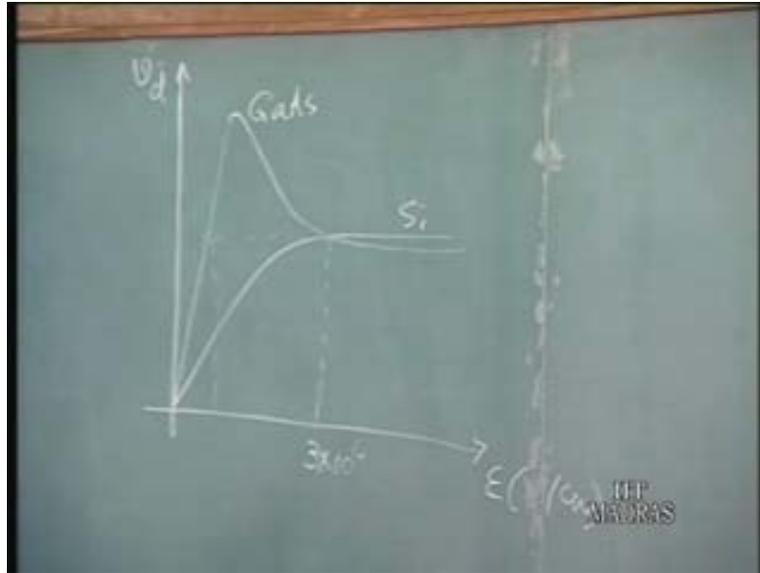
(Refer Slide Time: 07:53)

The point is because the initial low field mobility is higher for gallium arsenide, this saturation velocity is reached at a much lower electric field in gallium arsenide. If you come back to this curve again, we see that the saturation velocity of silicon is reached this is around 3 into 10 to power 4 volt per centimeter.

(Refer Slide Time: 09:01)

Whereas for gallium arsenide this saturation velocity above this electric field will be always greater than the saturation velocity. This field is much lower, it is in fact even at electric fields of the order of 10 to power 3 volt per centimeter you have velocities equal to the saturation velocity.

The point is that in order to reach saturation velocity you require lower electric field in gallium arsenide and lower electric fields means lower voltages applied across devices which means that the devices can be operated at lower voltages and if we can operate a device at lower voltage, it means you require less power. That is a distinct advantage for gallium arsenide devices also since the saturation velocity is reached at lower electric field, when you have a field effect transistor the major portion of the channel will be traveling at saturation velocity in gallium arsenide devices while in silicon devices, the carriers only very close to the drain end may be traveling at saturation velocities. That is also a distinct advantage for gallium arsenide devices. Then of course peak velocity, for gallium arsenide we see that we have a peak velocity which is around 2 into 10 to power 7 per centimeter per second which we do not have for silicon devices.

(Refer Slide Time: 10:40)

	Direct	Indirect
Low field mobility of electrons	5000 cm²/V-sec for $N_D = 10^{17}/cm^3$ 8000 cm²/V-sec for undoped	800 cm²/V-sec for $N_D = 10^{17}/cm^3$
Saturation velocity	10^7 cm/sec	$\sim 10^7$ cm/sec
Peak velocity	2×10^7 cm/sec	—

IIT MADRAS

In fact this can also be utilized if we properly design the device to achieve very high speed devices. We see that if we have sort of n channel device where the carriers are electrons, gallium arsenide has a distinct advantage over silicon in terms of speed and this can be utilized to realize devices of higher speed but if you look at the low field hole mobilities, for gallium arsenide it is around 250 centimeter square per volt second whereas this would be around 300 centimeter square per volt second again for a doping concentration of 10 to the power 17, it shows that the hole mobility is very poor in gallium arsenide.

(Refer Slide Time: 11:07)

This gallium arsenide cannot be used to make a complementary device like CMOS as we do for silicon. That is a distinct disadvantage also for gallium arsenide devices. We have to make circuits which use only n channel devices. We have to look at such circuits only, we cannot have complementary circuits. These are some of the differences. Now let us look at some more property differences between the two. Substrate resistivity that is this gallium arsenide devices can be made on undoped gallium arsenide samples. In fact they cannot only be undoped but by introducing certain deep level impurities, the resistivity of the sample can be increased very heavily.

(Refer Slide Time: 14:40)

For example you can have samples having resistivity up to 10 to power 6 to 10 to the power 8 ohm centimeter whereas in silicon this is low. Usually most gallium arsenide devices would be made on these types of substrates which are called semi insulating substrates. The advantage is this semi insulating substrate can be used to isolate individual devices, not only that since they are made on this high resistivity substrates, the parasitic capacitance from the device to the substrate is very low and because of this reduction in parasitic capacitances, the device can be operated at higher frequencies. The parasitic device to substrate capacitance can almost be eliminated. That is a distinct advantage for gallium arsenide.

Now there are some disadvantages to which we shall come to, one is the surface state. The surface states are the unsaturated tangling bonds or the surface of the material and because of that when you make a device that give rise to some surface charges. And these surface charges play a role also in device operation. These surface state densities are of the order of almost 10 to the power 12, this is per centimeter square electron volt in gallium arsenide, where this may be quite low in silicon about 10 to power 10 per centimeter square electron volt.

Now the effect of the surface states is that in gallium arsenide, number one it is difficult to make ohmic contacts because they result in what is called Fermi level pinning. Now the Fermi level is pinned at one particular level at the surface when you make a metal contact but ohmic contacts have been made using heavily doped substrates. If you make metal contact on a heavily doped gallium arsenide substrate, you do get ohmic contact so that can be overcome this way.

But the other problem is that because of this Fermi level pinning, one cannot really have inversion layers in gallium arsenide sub surface substrates, just like you have in a MOS in a silicon MOSFET because you cannot have a MOSFET as such. It's very difficult to make a MOSFET on gallium arsenide because it is very difficult to invert the surface. We cannot do that. You cannot have this enhancement type of devices, MOSFET's as you have in silicon that is one problem and we cannot have really MOSFET type of devices on gallium arsenide. Another problem of course is that if you look at a MOSFET, you require an oxide, MOS structure metal oxide semiconductor, you require a native oxide. Now gallium arsenide native oxide is very unstable, it cannot be used for any practical devices whereas for silicon it is very stable.

(Refer Slide Time: 16:40)

In fact the thermally grown oxide is used for making MOSFET's in silicon devices whereas that is not possible in the case of gallium arsenide. You see that because of these two properties that is large number of surface states and the native oxide problem, we cannot have MOSFET type of devices in gallium arsenide. Although n channel field effect transistors would be quite attractive for gallium arsenide because of the advantages with respect to mobilities. We cannot make MOSFET type of devices in gallium arsenide. The device which is used in gallium arsenide for integrated circuits is the MESFET and not the MOSFET which is also a field effect transistor but in that case we do not have a MOS structure as in a MOSFET but we have a metal semiconductor junction.

We shall now take up this particular device, we shall look at the different properties of this device and then we shall see how these devices can be used to make integrated circuits. We shall now take up this particular device, the MESFET for making integrated circuits. These devices are quite typical of gallium arsenide. In fact one can really make MESFET in silicon but nobody would make it because the MOSFET is a much better device. Since in gallium arsenide we cannot make MOSFET's, we have to go for this MESFET type of devices. Let us see how a MESFET actually looks like.
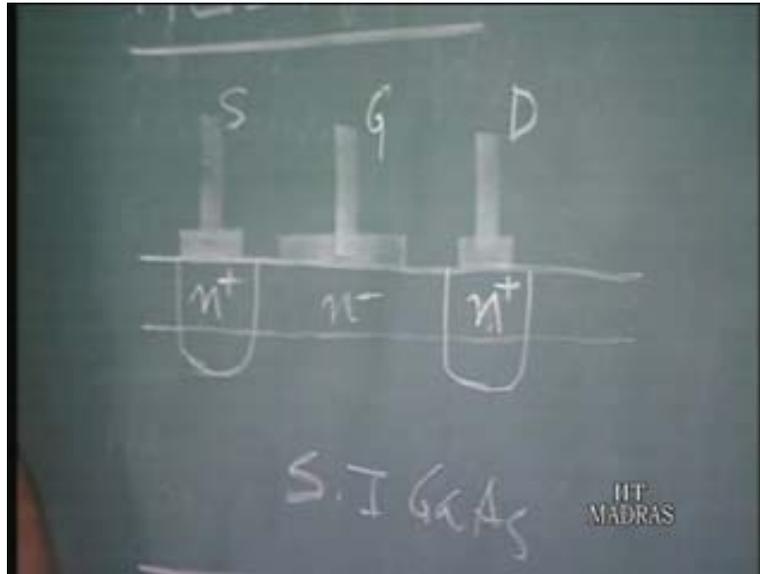
There are two ways basically which a MESFET can be made. In one case we start with a semi insulating gallium arsenide and then shallow implant of n type implant which is silicon which is an n type impurity in gallium arsenide can be made n minus I call it. This is a lightly doped n type implant in gallium arsenide at the surface and then we make heavy implant of a higher dose which we call n plus. This is to make the ohmic contacts. As we have already said that we require heavy doping in order to make ohmic contact. Now this is lightly doped region, this is the heavily doped region. When you make a metal semiconductor junction, this region is going to result in a schottky diode whereas in the heavily doped region it is going to make an ohmic contact. You have the metal here, we have a metal here and we have a metal here. This is the source, this is the drain and this is the gate.

The difference between the MESFET and the MOSFET as you can see in a MOSFET you have a MOS structure metal oxide semiconductor structure whereas here you have a metal semiconductor structure. This metal semiconductor forms a junction. This is quite similar to a JFET, junction FET where you actually have a pn junction whereas here you have a metal semiconductor junction. It is this metal semiconductor junction which will control the flow of carriers, electrons in this case from the source to the drain. The same device can also be made using an epitaxial wafer in which the starting material is you have a semi insulating gallium arsenide on which a lightly doped n gallium arsenide and on top of which you have heavily doped n gallium arsenide.

(Refer Slide Time: 22:11)
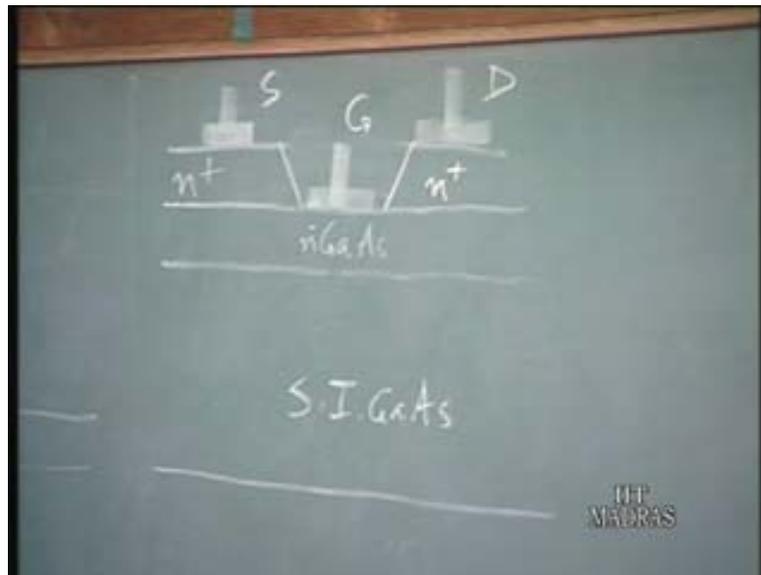
Recess etching is done on the surface, this region is etched down to reach this n gallium arsenide. You have n plus n plus here where as n is here. Now metal contact is here, metal contact here and metal contact here. Again you have the source, gate and the drain. This is the channel region, this is going to be schottky contact whereas this metal on n plus is going to be ohmic, these two are going to be ohmic. This is also a MESFET structure, this can be done using already grown epitaxial wafers where you do not have to carry out the implantations.

Let us see how this device operates that is the next thing which we are going to take up and look at the characteristics of this particular device. This metal semiconductor junction behaves almost like a pn junction or a p + n junction. This metal n junction is going to behave as a single sided abrupt pn junction that is p plus n junction. There is going to be a built in potential V_{bi} just like any pn junction and because of this built in potential there is going to be a depletion layer here on the n side. Suppose this built in potential is like this, there is a depletion layer of a particular width. The width of the depletion layer say w can be related to the built in potential by solving poisons equation we can get it. V_{bi} is equal to qN W square by twice epsilon or alternately we can write w is equal to twice epsilon silicon V_{bi} by q N where N is the doping concentration epsilon is of course the permittivity and V_{bi} is the built in potential and q is the electronic charge.

(Refer Slide Time: 27:00)
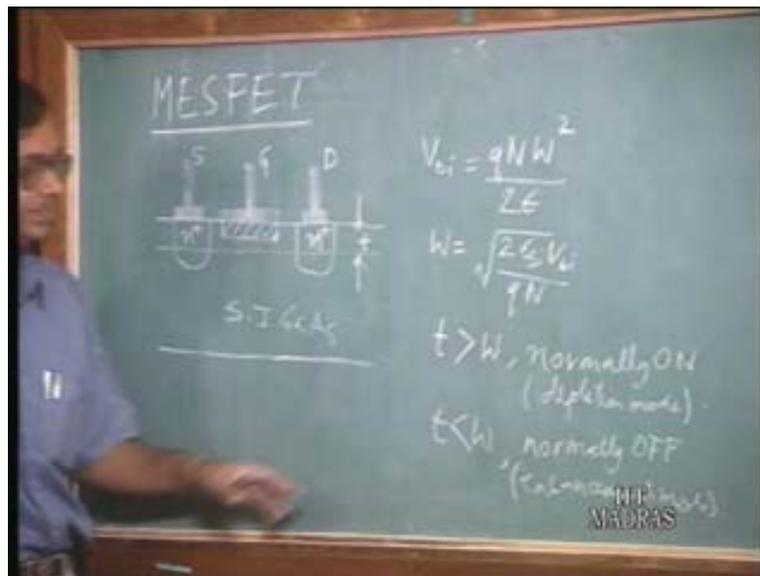
Now this is the depletion width which exists normally in the device. Now if t is the thickness of this n layer and if t is greater than w that means there is still a part of a channel which is undepleted and that means if you now apply a voltage between the drain and source, a current can flow between the drain and source, electrons can flow. There is a channel existing for current to flow. It is normally an on device, we call this normally on device or it is also called depletion mode device. On the other hand if t is less than w which means that if the depletion width due to the built in potential is much larger than the thickness of the channel that means at zero bias without applying any bias, the depletion width is completely going to cover the channel.

There is no way a current can flow between the source and the drain. If t is less than w then the device is going to be normally off and this is called an enhancement mode device. Now for these devices just in the case of MOSFET, we also define what is called a threshold voltage. The threshold voltage is that particular gate to source voltage at which the conduction just begins. That is if the gate to source voltage is greater than the threshold voltage there is conduction but if the gate to source voltage is less than the

threshold voltage there is not going to be any conduction. That is the gate to source voltage at which the depletion width just covers the channel region, the thickness t.

Now if t is greater than w that means it is not fully depleted. What we have to do is we have to reverse bias the gate junction. This junction we have to reverse bias it more to deplete the entire channel. That is we must apply a negative gate voltage to make it so that the depletion width just covers the complete channel. For a depletion mode MOSFET where t is greater than w, threshold voltage is going to be negative. Whereas for an enhancement mode MOSFET since the depletion width due to the built in potential is greater than the channel length, we in fact have to forward bias the gate voltage that is we have to apply a positive gate voltage so that the depletion width is just enough to cover the channel region t, thickness t.

For t is greater than w, the depletion mode device V_T is going to be negative and for an enhancement mode device V_T is going to be positive.

(Refer Slide Time: 27:38)
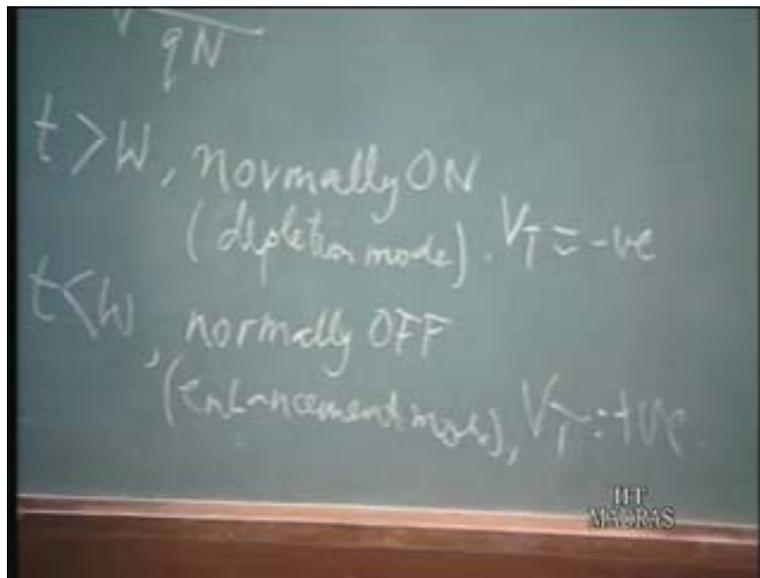
For that what we do is we define what is called a pinch off voltage V_{po} which is the voltage across the depletion region just sufficient for the depletion region to just cover the thickness of the channel region t.

(Refer Slide Time: 28:26)

$$V_{po} = \frac{qNt^2}{2\epsilon}$$

$$V_T = V_{bi} - V_{po}$$

V_{po} will be q N t square by twice epsilon and the threshold voltage V_T of this device will be given by V_{bi} minus V_{po}. That is extra voltage which is required to pinch off the whole channel. The threshold voltage is given by this. Again as I said we can have two types of devices, one is the enhancement mode device which is normally off device and depletion mode device which is a normally on device. It is very clear that if you want to make enhancement type device, since the depletion width due to the built in potential should cover the entire channel, the channel thickness t should be made quite small. For this type of device t less than w, t should be made very small or the N_D or the N, the doping concentration should also be made small.

If we come to this expression, for the same built in potential if n is small w will be large either we can make t small or w large by making N small. If you want to make enhancement mode device we have to make t small or N small, the doping concentration small. If you want to make a depletion mode device we can make the t thickness large or the doping concentration large so that the w is small.

Normally it is easier to fabricate the depletion mode devices than enhancement mode devices because you require much greater control in making enhancement mode devices. Now let us see what is the characteristic of these devices, that is the current voltage characteristic. How is the nature of these characteristics for these devices? You have the drain current on one side and the drain to source voltage on the other. Suppose we have again a depletion mode device where the threshold voltage is negative.

(Refer Slide Time: 36:16)

At this point say zero bias, the channel already exists or whatever is the bias if at a particular bias the channel exists where current can flow. That is the entire region here is not fully depleted. Now if we apply a drain to source voltage current can flow in this path and the current will be given by if I write I which can be given as A q n v where A is the cross sectional area of current flow, q is the electronic charge, n is the concentration of carriers, n is the electron concentration which can be said to almost equal to capital n which is the doping concentration and v is the velocity.

(Refer Slide Time: 31:45)
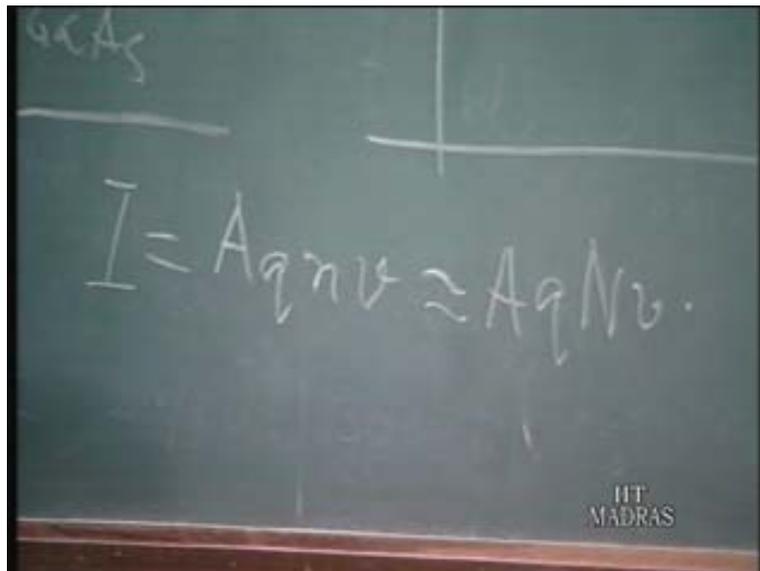
It's almost equal to A q n v, v is the velocity of the carriers. This basically behaves like a resistance and if we apply a drain to source voltage at a particular gate to source voltage then as we increase the drain to source voltage, the current is going to rise
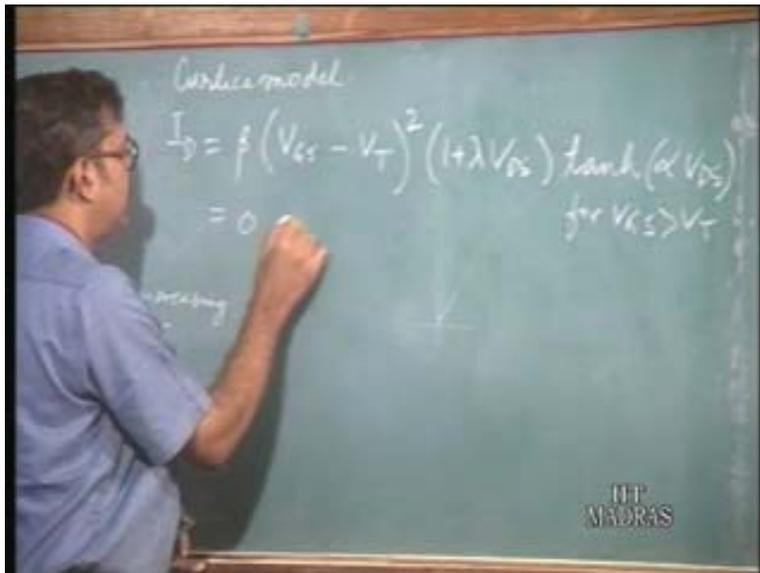
linearly initially but as we are increasing the drain to source voltage, the voltage at the drain end of the channel is increasing and this voltage is in such a direction as to reverse bias this gate to substrate or semiconductor junction at this end. The depletion width is going to keep widening at the drain end. There is going to be widening of the depletion width at the drain end.

Now since there is a widening of the depletion width at the drain end it means that if you look at this expression here I is equal to A q n v which means A that is the cross sectional area of current flow reduces. Now if we must have a constant current flowing throughout the channel which means that basically what is going to happen is if this A reduces the velocity has to keep increasing. This happens because at the drain end electric field also keeps increasing. The electric field is not constant along the current flow but the electric field is higher at the drain end which makes the velocity increase but we know that if we keep on increasing the velocity ultimately it will reach saturation.

Basically what happens is initially as we keep on increasing the drain to source voltage because of the reduction in the cross sectional area, the current actually does not rise linearly but sort of sub linearly and finally when the velocity reaches the saturation, the current also saturates and it becomes something like this. There is a slight increase in current because once after velocity saturation is reached if we keep on increasing the drain voltage, the velocity saturation point actually move towards the source end which means an effective reduction in the channel length and a slight increase in current. This is for a particular gate voltage, if we take a higher gate voltage more positive gate voltage the current is going to increase more because for a higher gate to source voltage the depletion width in the channel is going to be less but the initial cross sectional area of current flow is going to be more. This is the nature or the characteristics which we have for different gate to source voltages. In fact this nature of characteristics is valid for both enhancement as well as depletion mode devices.

If we have a depletion mode devices it means that the current flow will exist even when V_{GS} is less than zero volts. For example if the threshold voltage of a depletion mode device is minus 1 volt then you can have current at minus 0.9, this could be minus 0.8, this could be minus 0.7, minus 0.6, minus 0.5. As we are making the gate voltage approach zero, the current increases. If it is an enhancement mode device with a threshold voltage say 0.1 volt then this could be for 0.1 volt, this could be for 0.2 volt, this could be for 0.3, 0.4, 0.5. This is basically for increasing v_{GS}. This is the nature of characteristic, we see that the nature of characteristic is quite similar to that of a MOSFET. This current in gallium arsenide MESFET's we can in fact use the same relations for MOSFET but one particular model which is generally used for the current voltage characteristics is what is called a curtice model where the expression for drain current is given as I_D is equal to beta V_{GS} minus V_T square.

(Refer Slide Time: 38:35)

Curtice model

$$I_D = \beta \left(V_{GS} - V_T \right)^2 \left(1 + \lambda V_{DS} \right) \tanh \left(\alpha V_{DS} \right) \quad \text{for } V_{GS} > V_T$$

$$= 0$$

(Refer Slide Time: 38:45)

Curtice model.

$$I_D = \beta \left(V_{GS} - V_T \right)^2 \left(1 + \lambda V_{DS} \right) \tanh$$

$$= 0 \quad \text{for } V_{GS} \leq V_T$$

increasing V_{GS}.

This is similar to the square law in MOSFET's, beta V_{GS} minus V_T square then you have 1 plus lambda V_{GS}. This is to take into account, the increase in current after saturation. This is also quite similar as in a MOSFET. Then you have a term tan hyperbolic alpha V_{DS}. The tan hyperbolic is used as the argument of this alpha V_{DS} is small. When alpha V_{DS} is small, tan hyperbolic alpha V_{DS} would be almost equal to alpha V_{DS} which means that as V_{DS} increases, tan hyperbolic alpha V_{DS} would also increase almost linearly. This current would increase almost linearly but when tan hyperbolic alpha V_{DS} becomes large then this function approaches unity.

So tan hyperbolic alpha V_{DS} is going to become one which means we get this value beta V_{GS} minus V_T square into one plus lambda V_{DS} which is saturation current as in a MOSFET. The advantage here is that we can use a single expression to express the entire current voltage characteristic unlike in the square law models of MOSFET's where you have a one expression for the linear region and another expression for the saturation region and that makes it very simple for simulation purposes. This is for V_{GS} greater than V_T and this is equal to zero for V_{GS} less than equal to V_T. This is the drain current model for the MESFET. This models this current voltage characteristics. Then a little bit on the equivalent circuit of MESFET. This is the MESFET equivalent circuit, the large signal equivalent circuit what we have is a current source I_D and this is given by this relation.
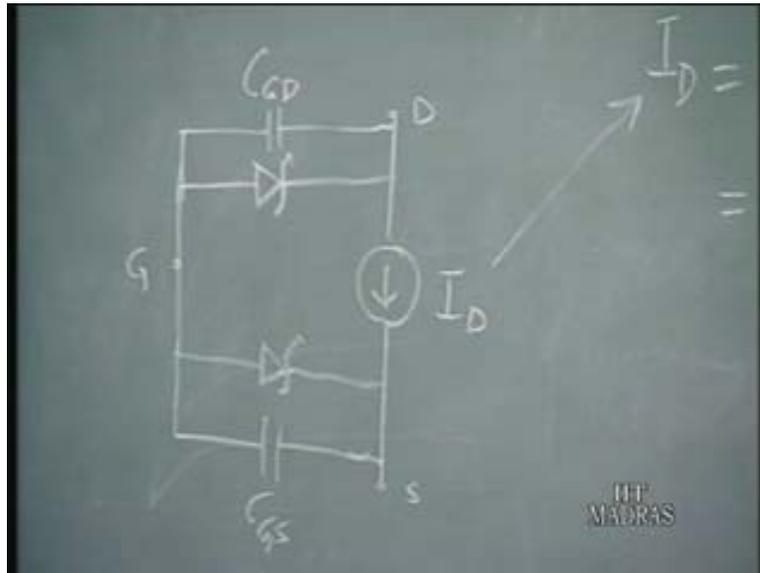
This is given by the relation of I_D here and this is the drain point and this is the source point. You have a schottky contact here (Refer Slide Time: 39:57). A current can flow, this is the schottky contact and you know that the characteristics of a schottky diode that is it is just like a pn junction diode. Once this voltage exceeds the cut in voltage there is going to be conduction. Basically you have two diodes, this is a schottky diode basically and between the gate and the drain as well from the gate to the source. These diodes will in fact start conducting once this exceeds the cut in voltage of these diodes.

There will be a gate conduction unlike a MOSFET, this is quite unlike a MOSFET because of the presence of these diodes it is also going to clamp the gate to source voltage to a particular value. That is the cut in voltages of these diodes normally in gallium arsenide devices around 0.7 volts which means that the gate to source voltage or the gate to drain voltage cannot exceed that value because once it approaches that value, there is going to be a conduction of these diodes and the current will be limited.

We also have two other elements, the capacitances between the gate to source and the gate to drain. This is expressed as C_{GD} and C_{GS}, these are basically the junction capacitance type. This is due to the depletion widths at the source and the drain ends of the channel and these can be expressed as, C_{GD} can be given as C_{GO} by one minus V_{GD} by V_{I} to the power half and C_{GS} is equal to C_{GO} by one minus V_{GS} by V_{bi} to the power half.

(Refer Slide Time: 42:45)

$$C_{GD} = \frac{C_{GO}}{\left(1 - \frac{V_{GD}}{V_{bi}}\right)^{1/2}}$$

$$C_{GS} = \frac{C_{GO}}{\left(1 - \frac{V_{GS}}{V_{bi}}\right)^{1/2}}$$

IIT
MADRAS

What is C_{GO}? That is when V_{GD} is equal to zero C_{GD} is equal to C_{GO} and when V_{GS} is equal to zero, C_{GS} is equal to C_{GO}. C_{GO} is actually the capacitance at zero bias and is given by C_{GO} is equal to where this epsilon w is the width of the device, l is the length of the device, twice I shall use another symbol say d where d is actually the depletion width at zero bias.
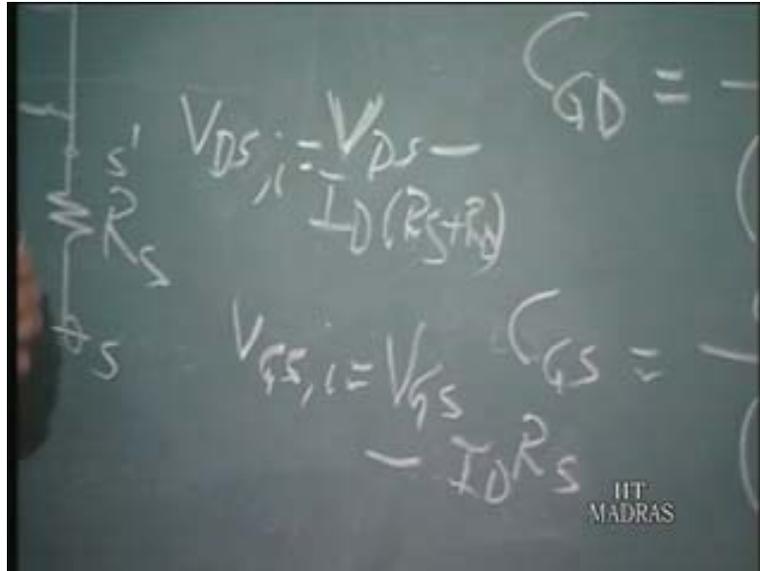
(Refer Slide Time: 44:05)

$$C_{GO} = \frac{\epsilon W L}{2d}$$

$$d = \sqrt{\frac{2\epsilon V_{bi}}{qN}}$$

IIT
MADRAS

Why do we put this two here? Because we are partitioning the total charge in the depletion layer between the source and the drain, so C_{GO} is actually epsilon a by d, the depletion width by two because half of the C_{GO} is for this gate to source and half for gate to drain and this d is nothing but the depletion width due to V_{bi} that is depletion width of V_{bi}. This is twice epsilon V_{bi} by qN. So that gives us the large signal model of the MESFET. You of course can have drain and source resistance R_D and R_S then this will become internal source point, this one will become internal drain point. You have the R_S, this is the external source point and this is the external drain point, so you have R_D and R_S in which case the internal gate to source voltage $V_{GS,i}$ will be equal to the external gate to source voltage minus I_D into R_S and the internal gate to drain voltage $V_{GD,i}$ will be equal to external gate to drain voltage minus I_D into R_S plus R_D.

(Refer Slide Time: 45:13)

The internal gate to drain voltage is equal to external gate to drain voltage minus I_D into R_S plus R_D. This is the internal drain to source voltage, this is the external drain to source voltage (Refer Slide Time: 45:26). $V_{DS,I}$ the internal drain to source voltage is equal to the external drain to source voltage minus the drop across these two resistances.

(Refer Slide Time: 45:25)

$V_{DS,i}$ equal to V_{DS} minus I_D into R_S plus R_D whereas the internal gate to source voltage between these two points is the external gate to source voltage minus the drop across this R_S. What we have to do in actual practice, if the source and drain resistance are important. We have to go to this current expression here and we have to replace the gate to source voltage and the drain to source voltage by the internal values.

(Refer Slide Time: 46:04)
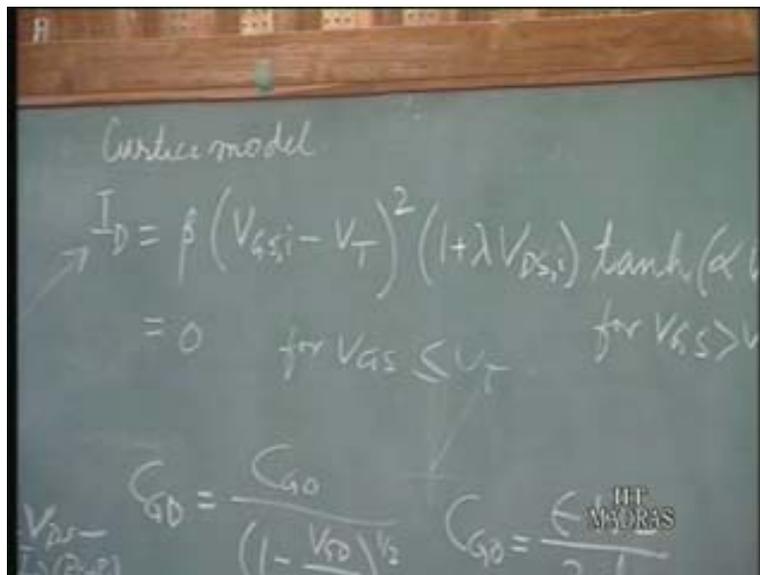
That is $V_{GS,i}$ instead of V_{GS} because the currents will be related to the internal values. This is the large signal model of the MESFET. For the small signal model of the MESFET, what we have is if we take the small signal model from the drain to source we have a current $g_m v_{gs}$.

(Refer Slide Time: 47:43)
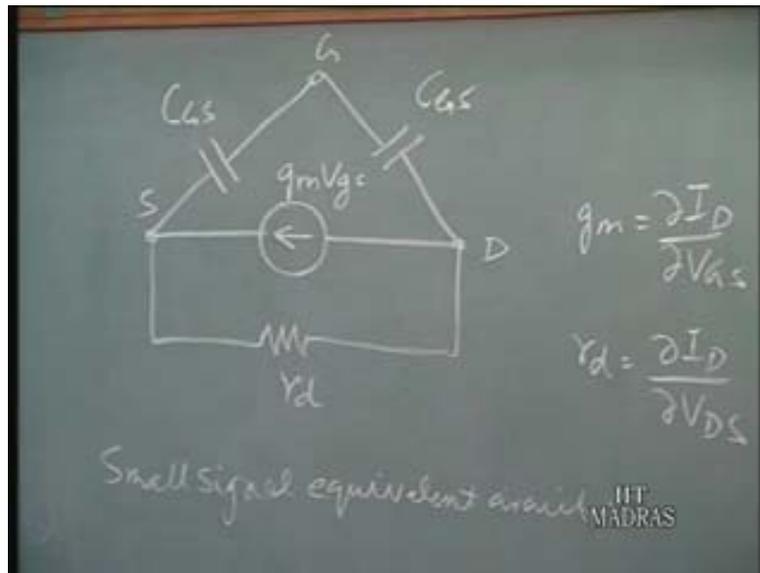
The current is dependent on the gate to source voltage and the drain to source voltage. We have a current source between drain and source driven by g_m v_{gs} where g_m is equal to del I_D by del V_{GS}. You have a drain resistance R_D where R_D is equal to del I_D by del V_{DS} and then you have the two capacitances gate to source and gate to drain. You have C_{GS} and C_{GD} which we have already seen. This is the small signal equivalent circuit of the MESFET.

In today's class what we have done is we have looked at the mesfet device in quite detail and what we shall do in the next class is we shall take this device and see how we can make different digital circuits which can be used. Now of course the circuit is going to be different from the mosfet circuits, one primary difference is that in the MESFET there can be gate conduction where as in the MOSFET we are not having and this limits the gate to source voltage. For this the circuits are going to be quite different. In the next class we shall take up the MESFET circuits for the high speed digital integrated circuits.

Thank you.