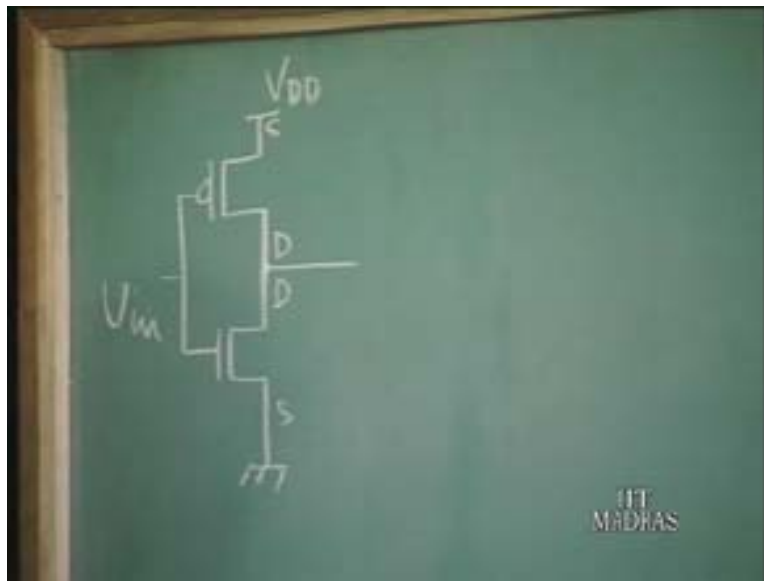


**Digital Integrated Circuits**  
**Dr. Amitava Dasgupta**  
**Department of Electrical Engineering**  
**Indian Institute of Technology, Madras**  
**Lecture No # 26**  
**CMOS inverter**

Today we shall continue our discussion on the CMOS inverter. Last class we have already seen the structure of the CMOS inverter which consist of an NMOS transistor and a PMOS transistor as a load and the gates of the two transistors are shorted at the inputs. The drains of the two transistors are also shorted; this is a PMOS transistor so this is the drain. The source of the NMOS transistor is grounded and the source of the PMOS transistor is connected to the power supply. The output is obtained at this point which is the common point where the drains of the two transistors are shorted. So this is the structure of a CMOS inverter.

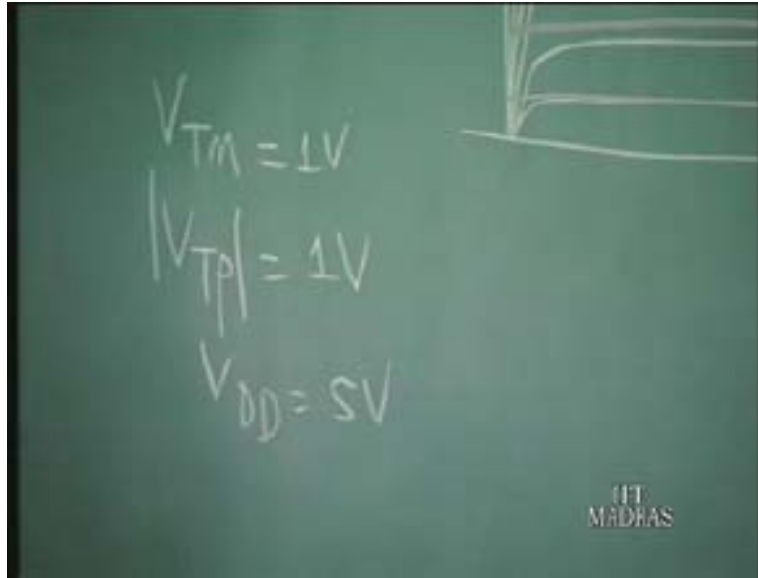
(Refer Slide Time: 01:50)



In today's class we shall try to obtain the transfer characteristics, input output characteristics of this inverter and we shall follow our usual procedure that is the graphical way of obtaining the input output characteristics. What we have done also for other MOS logic circuits is basically we have to draw the current voltage characteristics of the driver transistor which is something like this. This is a NMOS transistor characteristic and then the low transistor characteristics must now be drawn with  $V_{DD}$  as the origin.

This is the  $V_{DD}$  point say, so this is the origin point and now we have to draw the characteristics of the low transistor. Now the characteristics of low transistor, I will

draw this way. Suppose that the threshold voltage of the NMOS transistor  $V_{Th}$  which we call  $V_{Th}$  is equal to 1 volt and the  $V_{Tp}$  which is minus 1 volt.  
(Refer Slide Time: 03:59)

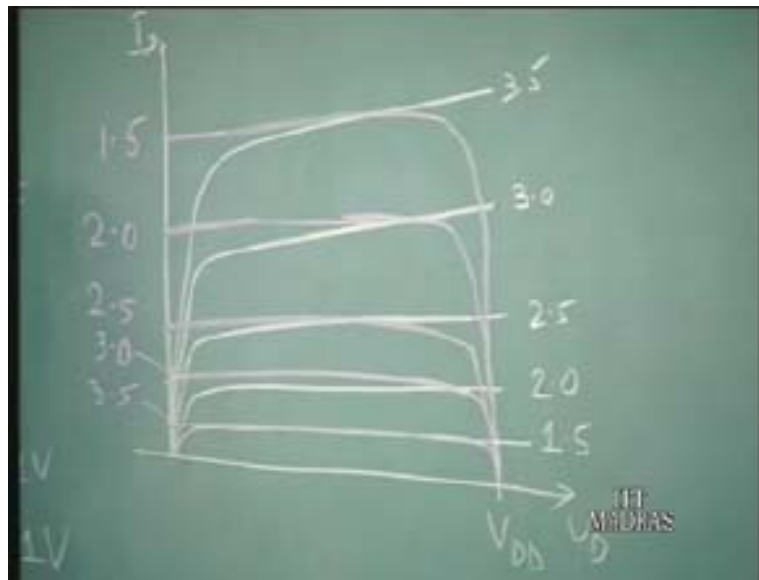


So let us write like this, modulus of  $V_{Tp}$  1 volt also and  $V_{DD}$  is equal to 5 volts. Let us take this particular case and these characteristics, so 1 volt is the threshold voltage. So let this be for say 1.5 volt say 2 volts, 2.5, 3, 3.5 and so on. This is for the NMOS transistor. So these are the characteristics for the NMOS transistor, gate voltage of 1.5, 2, 2.5, 3, 3.5.

Similarly we have the characteristics for the PMOS transistor. So this is the characteristics of the PMOS transistor when the gate voltage is minus 1.5 volts. Now gate to source voltage is minus 1.5 volts, now the source voltage of this PMOS transistor is at 5 volts. If the gate to source voltage has to be minus 1.5 volts so the input voltage or  $V_G$  must be equal to 3.5 volts. So this characteristic here is for  $V_{input}$  of 3.5 volts.

Next one, the  $V_{Gs}$  is minus 2 volts which means that the gate voltage  $V_{Gs}$  is going to be minus 2 volts when the gate voltage is 3 volts. So 3 minus 5 is minus 2 volts. So this voltage is 3 volts so this is what I am writing here, this is the input voltages so this is 3 then the next characteristic curve here is for 2.5 volts, next curve will be for 2 volts, next curves will be for 1.5 volts and so on.

(Refer Slide Time: 06:02)



These are the different curves. Now the operating point is going to be basically the intersection of these two curves. Now let us draw the input output characteristics. This is the output voltage, this is input voltage on the x axis. Now when the input voltage is zero volt say, now this characteristic for the NMOS transistor is going to be along the x axis because the gate voltage is less than the threshold voltage and so the input voltage is along this x axis and the gate to source voltage of the PMOS is minus 5 volts.

So the point of intersection is actually going to be written at  $V_{DD}$ . so when the input voltage is zero, output voltage is  $V_{DD}$ . Now this output voltage will remain at  $V_{DD}$  till the input voltage reaches  $V_{Th}$ . When the input voltage increases to  $V_{Th}$ , still the NMOS is not carrying any current so the characteristics are along the x axis. So point of intersection of the two characteristics is again at this point. The output voltage is like this, so up to  $V_{Th}$  the output voltage remains at  $V_{DD}$ .

Suppose the input voltage is increased above  $V_{Th}$ . What happens is this curve say for example it has gone up to 1.5 volts here, when it goes up to 1.5 volts here, so this is the curve for the NMOS transistor. For the PMOS transistor the gate to source voltage is actually minus 3.5 volts. So the characteristics is going to be some way here, so 1.5 volts here. This is the characteristics for the PMOS transistor. The point of intersection

again is very close to  $V_{DD}$ . so as you keep increasing the input voltage, the point of intersection is still very close to  $V_{DD}$ .

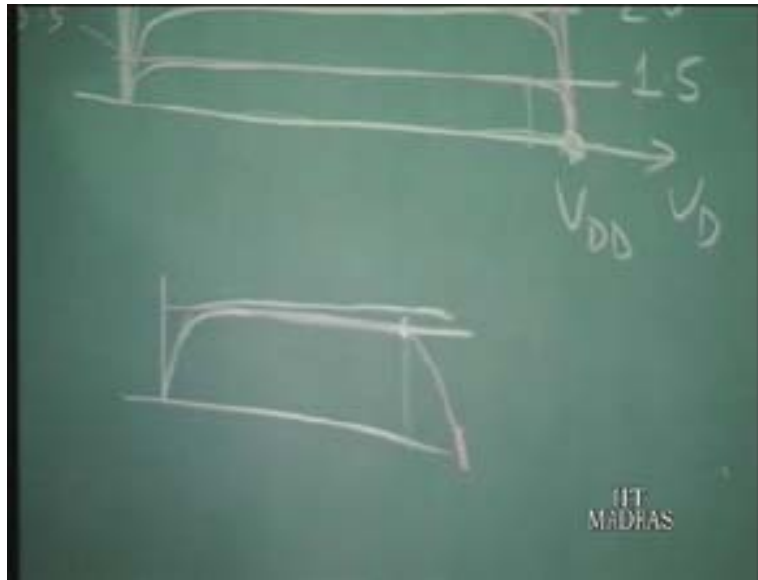
So 1.5, 2 volts so this is the characteristics for the NMOS transistor. For 2 volts, the PMOS transistor characteristics is here. So again the point of intersection is very close to  $V_{DD}$ . So this is 2 volts for the PMOS, this is 2 volts for the NMOS, point of intersection is here.

Now if we increase to two point volts, we find that the NMOS characteristics is this, the PMOS characteristics is this. So if it was just less than 2.5 volts, the point of intersection would be right here of the two curves. Now if it is slightly greater than 2.5 volts, the point of intersection say suppose if you take the 3 volt curve, the 3 volt curve for the NMOS is here this one and the 3 volt curve for the PMOS is this one.

The point of intersection you see is right here (Refer Slide Time: 9:08). So for the two volt curve the point of intersection is here, when the input voltage is 2 volt, the point of intersection is here that means the output voltage is very close to  $V_{DD}$ . whereas when the input voltage has increased to 3 volts, the point of intersection has shifted here.

There is a large drop in the output voltage, when the input voltage is changed by a small amount. What we observe is that the output goes like this and then there is a sharp fall, so up to this point it falls like this, slowly let us call this point  $V_A$  and then there is a sharp fall. See what happens is when the input voltage is increasing above  $V_{Tn}$ , this NMOS transistor is in the saturation medium. At the point of intersection the NMOS transistor is in the saturation region while the PMOS transistor is in the linear region at this point. At this point the PMOS transistor in the linear region of the characteristic. Then we have a region when both the transistors are in the saturation region. So the NMOS curve if I just draw it like this it goes like this, whereas the PMOS curve it goes like this.

(Refer Slide Time: 10:47)



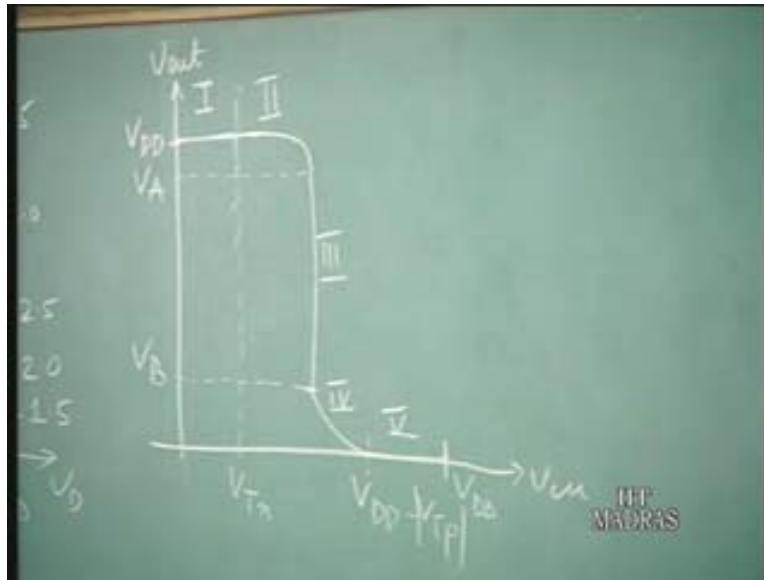
What we see is when we reach that condition, when both the transistors are in saturation the output voltage falls from here to here. Slight increase in this curve, the next characteristic the point of intersection shifts from here to here. So there is a very large gain in this region of the characteristics, when both the transistors are in the saturation. This region refers to the region, when the output voltage changes from  $V_A$  to  $V_B$  say when both the transistors are in the saturation.

Now what happens is after this when we keep increasing the input voltage say if you go to 3 volts, the point of intersection has shifted here. So in this region 3 volt, this is the characteristics of the NMOS transistor and for the PMOS transistor this is the characteristics. So the point of intersection here. We can see that the PMOS is in saturation but the NMOS at this point of operation is in the linear region, so this is in the linear region.

Now if we increase the input voltage still further 3.5 volts, this is the characteristics and for the PMOS 3.5 volt this is the characteristics. So at this point, this is the operating point which is still further close to zero and when the input voltage reaches  $V_{DD} - V_{Tp}$ , this is  $V_{DD} - V_{Tp}$ . This is  $V_{DD}$  that means the gate to source voltage is minus  $V_{Tp}$  that means that the PMOS transistor actually cuts off. PMOS transistor is off, this is conducting so the output voltage will go to zero. So at this point, the output voltage goes to zero and then this remains zero till we go to  $V_{DD}$ .

Basically if you look at this input output characteristics where there are 5 regions, let me call this 1, this region 2, this region 3, this region 4 and then we have another region 5. So there are basically 5 different regions of this curve.

(Refer Slide Time: 12:43)



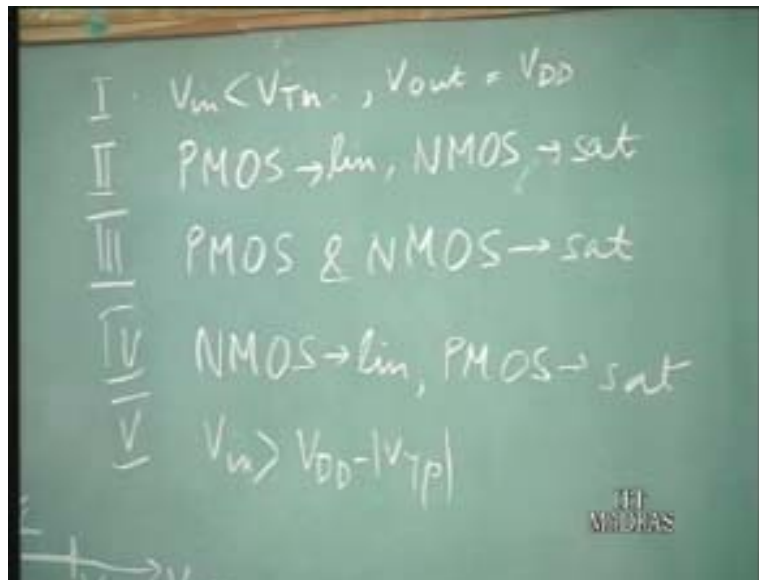
Now if you look at the different regions, number one for the first region, this region  $V_{in}$  is less than  $V_{TN}$  which means that the NMOS transistor is off, it is not conducting. The PMOS transistor however is conducting so the output voltage is going to be  $V_{DD}$ . Now if you look at region 2; NMOS transistor has come into conduction. If region 2 is say 1.5 volts, the NMOS transistor has come into conduction. So this is the point of intersection, this is the operating point. NMOS transistor is in the saturation region operation.

NMOS transistor is in the saturation region whereas the PMOS transistor is in the linear region of conduction. PMOS is the low transistor which is in the linear region of operation, NMOS is in the saturation.

This is the region 3 of the characteristics, when there is a sharp change in the output voltage for a small change in input voltage then both the PMOS and NMOS transistors are in the saturation region, this can be seen here. This is the region 3 so as we increase the voltage, initially we have this. This is the operating point where the NMOS is in saturation, PMOS in linear then both are in saturation, this is the region 3 and then if we increase the input voltage still further what happens is the point of intersection moves here where the NMOS is in the linear region and the PMOS is in saturation.

So in this point 4 NMOS linear, PMOS saturation and number fifth region where  $V_{in}$  is greater than  $V_{DD} - V_{TP}$  where the n NMOS transistor is conducting fully, the PMOS transistor is off, the output voltage  $V_{out}$  is equal to zero.

(Refer Slide Time: 15:15)



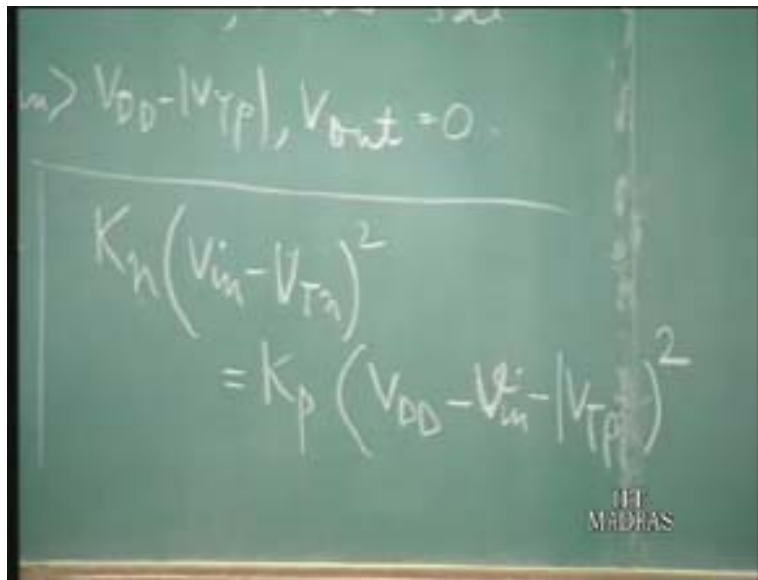
Now so this is the characteristic of a CMOS inverter. We see that it is very close to the ideal characteristics of an inverter in the sense that the output voltage is equal to the supply voltage, when the input is low. This equal to  $V_{DD}$ . whereas when the input voltage approaches the supply voltage in this region, the output voltage goes to zero. In between very close to the mid-point of the characteristics, the midpoint of zero to  $V_{DD}$ . there is a sharp drop in the output voltage from very close to  $V_{DD}$  to zero. So there is the very high gain region of this inverter. So this is the characteristics of a CMOS inverter and we see that this is very close to the ideal input output characteristics of an inverter.

Now an important design consideration for this inverter is this input voltage here. Let us call this by the logic threshold voltage. Logic threshold voltage in the sense that if you now plot line that is where  $V_{input}$  is equal to  $V_{output}$ . that is basically a line with a slope equal to 1, along this line always  $V_{input}$  is equal to  $V_{output}$ . So this line will most probably cut the characteristics along this region where the output voltage falls very sharply. So at this voltage, if you were to find the point in these characteristics where the input voltage is equal to output voltage, it would be along this region.

At this input voltage, the input voltage is equal to output voltage. The logic threshold is defined as that voltage where the input voltage is equal to the output voltage. Now the importance of the logic threshold is that at any input voltage greater than the logic threshold voltage, the output will be less than the logic threshold voltage in vice versa. So that is an important voltage for any logic family and for a good design this logic threshold voltage should be the equal to  $V_{DD}$  by 2 so that the noise margins are almost the same. The high noise margin and the low noise margin are almost the same.

So we shall try to find out what this  $V_{TL}$  depends on. Now in this region of the characteristics, we have already seen that the region 3 here, both the PMOS and the NMOS are in saturation. So in order to find out the logic threshold voltage, we will write down the equations for the two currents. Basically the PMOS current and the NMOS current and they must be the same. Now for the NMOS, the current is equal to  $K_n (V_{in} - V_{Tn})^2$ ; this is the current in the saturation. See both the transistors are in saturation so this is the NMOS current and this is equal to  $K_p$ , so  $K_p$  refers to the PMOS transistor,  $K_n$  refers to the NMOS transistor.  $V_{Gs}$  minus  $V_{Tn}$  whole square,  $V_{Gs}$  the modulus of that, the gate to source voltage is actually equal to  $K_p (V_{DD} - V_{in} - V_{Tp})^2$  minus  $V_{Tp}$  square.

(Refer Slide Time: 19:34)



$$K_n (V_{in} - V_{Tn})^2 = K_p (V_{DD} - V_{in} - V_{Tp})^2$$

So this is the equation which is satisfied in that region where both the transistors are in saturation. Now we can solve this equation. So what do we get?  $V_{in} - V_{Tn}$  is equal to  $V_{in} - V_{Tn}$  is equal to root of  $K_p$  by  $K_n$   $V_{DD} - V_{in} - V_{Tp}$ . So I can multiply this side with root  $K_n$  by  $K_p$  and bring it this way. So  $V_{in}$  into one plus root of  $K_n$  by  $K_p$  is equal to  $V_{DD} - V_{Tp}$  plus  $V_{Tn}$  into root  $K_n$  by  $K_p$ . So  $V_{in}$  or this is actually equal to the logic threshold. So this is equal to  $V_{DD} - V_{Tp}$  plus  $V_{Tn}$  into root  $K_n$  by  $K_p$  divided by one plus root of  $K_n$  by  $K_p$ .

(Refer Slide Time: 21:36)



$$V_{in} - V_{Tn} = \sqrt{\frac{K_p}{K_n}} (V_{DD} - V_{in} - |V_{Tp}|)$$

$$V_{in} \left( 1 + \sqrt{\frac{K_n}{K_p}} \right) = V_{DD} - |V_{Tp}| + V_{Tn} \sqrt{\frac{K_n}{K_p}}$$

$$V_{in} = \frac{V_{DD} - |V_{Tp}| + V_{Tn} \sqrt{\frac{K_n}{K_p}}}{1 + \sqrt{\frac{K_n}{K_p}}}$$

This is the expression of the input voltage at which both the transistors are in saturation so we have a sharp fall and as we said this is usually equal to the logic threshold voltage. If you look at these expressions and we find that if  $K_n$  is equal to  $K_p$ , this term root  $K_n$  by  $K_p$  will be one. So these two terms will cancel, this denominator will be 2 and of course if  $V_{Tn}$  is equal to  $V_{Tp}$ . Let us say if  $K_n$  is equal to  $K_p$ ,  $V_{in}$  at this point is equal to  $V_{DD}$  minus  $V_{Tp}$  plus  $V_{Tn}$  divided by 2.

So if  $V_{Tp}$  is equal to  $V_{Tn}$ ,  $V_{in}$  will be  $V_{DD}$  by 2. From this we see that if the threshold voltages of the two transistors are the same. That is  $V_{Tn}$  is equal to modulus of  $V_{Tp}$  then  $V_{in}$  is equal to  $V_{DD}$  by 2 provided root  $K_n$  by  $K_p$  that is provided root  $K_n$  by  $K_p$  is one. That is  $K_n$  is equal to  $K_p$ . Of course if  $K_n$  is not equal to  $K_p$  which may be the case, the  $K_n$  and  $K_p$  we know are dependent on the W by L ratios, it is also dependent on mobility.

(Refer Slide Time: 22:55)

The image shows a chalkboard with handwritten equations for the CMOS inverter threshold voltage  $V_m$ . The first equation is:

$$V_m = \frac{V_{DD} - |V_{Tp}| + V_{Tn} \sqrt{\frac{K_n}{K_p}}}{1 + \sqrt{\frac{K_n}{K_p}}}$$

Below this, it is noted that  $K_n = K_p$ . The second equation, which is the simplified result, is:

$$V_m = \frac{V_{DD} - |V_{Tp}| + V_{Tn}}{2}$$

A small logo for "IIT MADRAS" is visible in the bottom right corner of the chalkboard image.

Since the mobility in the case of NMOS transistor is different from that of a PMOS transistor, mobility for electrons is more than that of holes.  $K_n$  by  $K_p$  may not be the same, if we use the smallest size transistors for both NMOS and PMOS in which case this root  $K_n$  by  $K_p$  may not be equal to one, in that case the threshold voltages of the two transistors may be slightly adjusted so that we still get  $V_{DD}$  by 2 or else what we can do is we can adjust the sizes of the transistors so that  $K_n$  made equal to  $K_p$  and so again if  $V_{Tn}$  is equal to modulus of  $V_{Tp}$ , you still get  $V_n$  is equal to  $V_{DD}$  by 2. So this logical threshold voltage can be adjusted so that you get  $V_{DD}$  by 2 which is the desired voltage.

For the CMOS inverter we have an input output characteristics which is very close to the ideal inverter characteristics and again this logic threshold can be adjusted to the  $V_{DD}$  by 2. So that we have almost equal noise margins and again we get a very good characteristic. Now we have also seen in this while we derive this characteristic that there are two regions where say region 1 and region 5 where there is basically no current flowing through the inverter. If we come back to this inverter circuit, now for a current to flow from the supply to ground both the transistors must be conducting.

So if any situation, one of the transistors is off, it is not conducting then there cannot be a current and that situation occurs when the input voltage either is less than the threshold voltage of this NMOS transistor or the input voltage is increased to a value where the gate to source voltage of the PMOS transistor is such that if the PMOS does not conduct. So in region 1 and region 5 there is no conduction. Usually the CMOS is in one of these regions and so there is no power dissipation as such. There is no static power dissipation because there is no current flowing because the power will be equal to the current in to voltage. So that is one of the major advantages of CMOS that is there is no power dissipation, no static power dissipation.

But in the other regions say if we have 2, 3, 4 regions there is going to be a current flowing and the transistor may be in this regions only when the output is changing from one state to the other. That is in the transition from one state to the other. Otherwise there is going to be no power dissipation. Now if you look at this characteristic and if you plot the current flowing in these regions, we will find that the current is going to be maximum when in the region 3.

For example if we come back to this characteristic, so when the input voltage is 1.5 volts, the point of intersection is at this point. So what is the current? The current is equal to this value here, when the input goes to  $T_2$  volts, the point of intersection is somewhere here so the current has increased from here to here, 2.5 volts the point of intersection is here. So the current has again increased but when you go to 3 volts, the point of intersection comes here. So the current which is the y axis, the point of intersection has reduced so it goes up from 1.5 to 2 to 2.5 and then again it goes down to 3, 3.5 and again it goes to zero. This region say around 2.5 where both the transistors are in the saturation, the current is maximum.

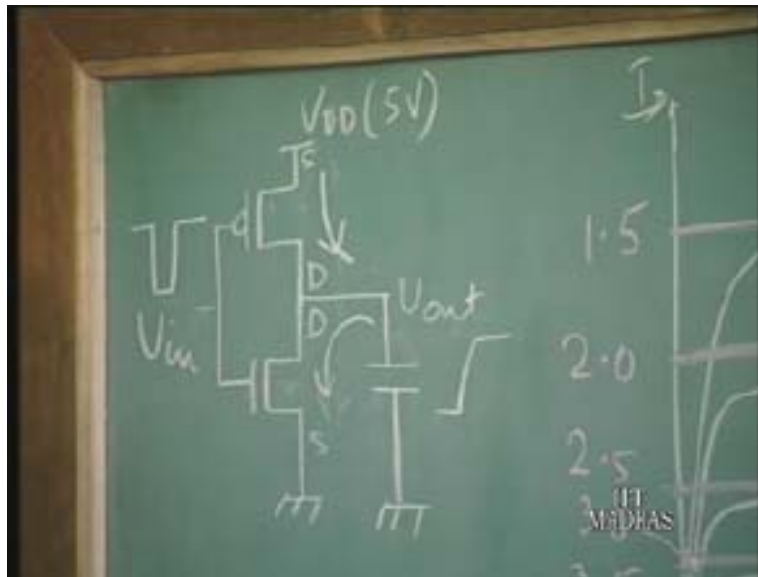
So if we plot the current, the current will be something like this, it reaches a maximum and this point and again it falls to zero (Refer Slide Time: 27:40). So if the current is zero on this side, the current is maximum at this point, now again the current goes to zero. So this is how the current is going to vary in the CMOS inverter. Now what we are interested is the power dissipation, how to calculate the power dissipation? Of course in the static condition, we have said that there is no power dissipation because the CMOS will be either in this 1 to 5 condition but when it is switching or changing state, there is going to be a power dissipation.

Now let us see what is the power dissipation. Let us consider again that this inverter has the capacitive load which is usually the case in the sense that the output of this CMOS may be connected to the input of another CMOS and then this capacitance is basically the input capacitance of next CMOS.

Since it is the CMOS, we know that there is no resistance, the input resistance is extremely high so there cannot be a current flowing in the normal static condition but because of the presence of the capacitance there will be a current flowing basically to charge or discharge the capacitance. Now suppose the input is high and it goes and makes a transition like this from high to low. What happens is there is a current flowing, the output voltage was low, it has to go high; the output voltage would go from low to high. So what happens is when the input goes from high to low, this NMOS transistor is switched off that the PMOS transistor is turned on, PMOS is conducting, NMOS is off. A current will flow from  $V_{DD}$  through the capacitance, basically to charge this output capacitance, so the output capacitance will get charged and the output voltage will rise from 0 towards  $V_{DD}$ .

Now in the next period when this input goes from low to high again what happens is this NMOS turns on and the PMOS turns off. Then what happens is the current flows. This capacitance basically discharges through the NMOS transistor, so there is no current flowing through the power supply as such. In one cycle as you see what happens is in one transition, there is a current flowing from the power supply to charge the output capacitance that is when the NMOS is off and the PMOS is on. So NMOS turns off, PMOS turns on, the output voltage was low so a current flows from  $V_{DD}$  to charge the output capacitance, the output voltage rises. Now what happens is giving the other transition that when the input goes from low to high, this NMOS turns on, PMOS turns off, so this output capacitance is discharged.

(Refer Slide Time: 31:22)



In one half of the cycle we see that a current is flowing from the power supply to charge the capacitance. So that is a current flowing through the power supply. The next half this capacitance is discharged to ground so the current is not flowing directly from the power supply to ground, in one half the energy is stored in the capacitance. If the current flows into the capacitance and then that capacitance is discharged to ground in the next half cycle. You have to calculate the power dissipation. What we do is we will find out the energy transferred during each transition. The energy transferred  $E$  let us say will be given by  $V_{DD} \int I dt$  during the transition time, so zero to say  $t_{01}$  or  $t_{10}$  that is the  $t_{dL}$  time delay or time required for the output to charge. That is the energy transferred to the capacitance.

Now  $I$  is nothing but  $C \frac{dv}{dt}$  current flowing into the capacitance  $C \frac{dV_c}{dt}$ , capacitance voltage. If we substitute here  $E$  is equal to  $V_{DD} \int C \frac{dV_c}{dt} dt$  zero to  $V_{DD}$ , the capacitance gets charged up to  $V_{DD}$ . Now  $C$  is a constant, so what we get is equal to  $C V_{DD}^2$ . So

this is the energy transferred from the power supply during each switching, during each cycle we can say.

(Refer Slide Time: 37:21)



If the input is changing so that each cycle that is the energy transferred or the energy transferred from the power supply  $CV_{DD}$  squared. Now if you have an input square wave or and the frequency is  $f$ , you have a clock so if the frequency of the clock is  $f$  that means there will be  $f$  such transitions or  $f$  such cycles per second. So per second will be  $f$  such events taking place and since in each event you have an energy transfer of  $CV_{DD}$  square. The total power dissipation can be given as  $P$  is equal to  $CV_{DD}$  squared into  $f$  (Refer Slide Time: 32:23). This is the power dissipation of the CMOS inverter and this power dissipation depends on the frequency. If the input frequency is high so the power dissipation will increase, so it is linearly dependent on frequency.

So it is actually sort of myth or whatever call it that the CMOS participation is very low. In fact the CMOS will have a large participation if the frequency is very large, it is linearly dependent on frequency. Only thing is that the static power dissipation is zero that is if the input is not changing state, the power dissipation is zero but if there are transitions then there is going to be the power dissipation which is linearly proportional to the frequency. So as we increase the frequency of CMOS circuits the clock frequency is increased, the power dissipation also increases.

In fact it is interesting to note that if you operate commonly available CMOS logic gates available in the market sufficiently at high frequencies, the power dissipation may in fact be greater than LSTTL. For example if the frequency is increased above 5 megahertz or so, it has been shown that some of this CMOS gates has larger power dissipation than the LSTTL gates. So it is sort of myth to say that CMOS power

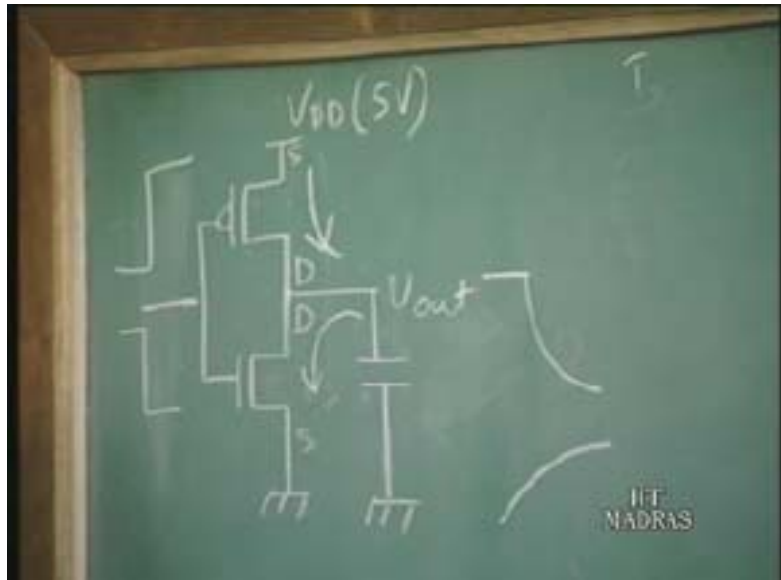
dissipation is extremely small but of course it is true. Actually the static power dissipation is basically a wastage of power, you are not doing anything.

So in the CMOS that way it is very efficient in the sense that the power dissipated is actually useful power in the sense it is used for switching purpose, there is no power dissipation when circuit is not changing state. Another important thing is that the power dissipation is proportional to  $V_{DD}$  squared. This is a matter of concern also in modern VLSI circuits for example whereas the complexity of the circuits increasing that means there are more number of gates so the participation obviously will go up, this is for one gate. So if you have more number of gates there will be more power dissipation. At the same time the frequency is going up, circuits with high and high frequency are being made so the power dissipation goes up.

So the only way to reduce the power dissipation, of course you can reduce the capacitance. Capacitance is proportional to area, so if you reduce the area or reduce the dimensions of the individual devices  $C$  will go down. Another way to reduce the power dissipation is reduce the  $V_{DD}$  square. This is one of the reasons why nowadays the power supply voltages are going down. Another reason of course is because these devices are smaller, we cannot have high voltages because electric fields will be too high. For one of the reasons is also that the power dissipation has to be reduced. We see that the power dissipation of a CMOS is dependent on the frequency.

So we have looked at the static input output characteristics, we had looked at the power dissipation. The next thing which we shall look at is the propagation delay of this CMOS circuit. Again we come back to this diagram of a CMOS inverter. Again we take a transition that the input goes from low to high suddenly, the output is going to go from high to low. basically what happens is as the input goes from low to high say zero to  $V_{DD}$ , this output which was at  $V_{DD}$  is going to decay to towards zero. Now what happens is when this input is close to  $V_{DD}$ , NMOS is fully on whereas the PMOS is off because this is  $V_{DD}$ , the gate to source voltage here is going to be zero in this case so it is going to be off. What happens is this capacitance discharges through the NMOS transistor. We can forget about the PMOS transistor now because it is off.

(Refer Slide Time: 39:55)



What we have is, the capacitance discharging through the NMOS transistor. When we have the other transition that is when the input goes from high to low, the output will go from low to high. What happens is the NMOS transistor is off, the PMOS transistor is on that discharges the capacitance. So when you have the input varying like this, one of the transistors is only conducting. So when the output is falling that is the capacitance discharging, NMOS is on, PMOS is off. So you can consider only the NMOS transistor.

On the other hand when the input falls to zero, NMOS is off, the PMOS is on, the output capacitance gets charged. So we need to only consider the PMOS transistor and we can neglect the NMOS transistor, we need not consider the NMOS transistor during this period. Now how do we calculate the times? Let us first consider the fall time. Fall time is the time when the output is falling that is this region, we are considering this case that is the input has increased from zero to VDD say and the output is falling from VDD to ground. So as the input has gone up, the NMOS transistor turns on and that is basically discharging the capacitance.

The current flowing through the NMOS has to be considered to consider the discharge. Now we know that  $I$  is equal to  $C \frac{dv_c}{dt}$  so this is the load capacitance say  $C_L$ . So we can write  $t_f$  is equal to  $\int dt$  is equal to  $C_L \frac{dv_c}{I}$ , output is falling from  $V_{OH}$  say output is high whatever is the output voltage initially to somewhere  $V_{OL}$ , so low voltage. This is the fall time.

(Refer Slide Time: 41:35)

$$I = C_L \frac{dv_c}{dt}$$

$$t_f = \int_{V_{OH}}^{V_{OL}} \frac{C_L}{I} dv_c$$

Now this is  $C_L$  so initially what happens is the input is high. So this NMOS input is high, output is also high. So  $V_D$  the drain voltage  $V_{DS}$  might be greater than  $V_{GS}$  minus  $V_T$ . So it means that this transistor is in the saturation region. So this transistor is initially in saturation and then as the capacitance is discharging, the output is falling then it comes to the linear region.

So initially the transistor is in the saturation and then it goes in the linear region. We have to consider both the cases. So first  $V_{OH}$  to say let us call it  $V_{G'}$  where  $V_{G'}$  prime is nothing but  $V_G$  minus  $V_{Th}$  where  $V_G$  is the gate voltage or the input voltage. So we have to take  $dv_c$ .  $I_{sat}$ ,  $I_{sat}$  is nothing but  $K_n V_{G'}^2$ ,  $V_{G'}$  is  $V_G$  minus  $V_T$  whole squared.  $V_c$  is actually the drain voltage also, voltage across the capacitance is also the drain to source voltage of the MOSFET plus so you have  $C_L$  plus  $V_{G'}$  to  $V_{OL}$ . So in this range the MOSFET is in the linear region.

(Refer Slide Time: 45:04)



$$\begin{aligned}
 &= C_L \left[ \int_{V_{OH}}^{V_{OH}-V_t'} \frac{dV_c}{K_n V_c^2} + \int_{V_{OH}-V_t'}^{V_{OL}} \frac{dV_c}{K_n [2V_t' V_c - V_c^2]} \right] \\
 &= \frac{C_L}{K_n V_t'} \left[ \frac{V_{OH}-V_t'}{V_t'} + \frac{1}{2} \ln \frac{2V_t' - V_{OL}}{V_{OL}} \right]
 \end{aligned}$$

IIT  
MADRAS

You have  $K_n$  twice  $V_{G, \text{prime}} V_c$  minus  $V_c$  square. So this equation has to be solved to get the fall time. I will just write down the solution of this equation. We get the solution as  $C_L$  by  $K_n V_{G, \text{prime}} V_{OH}$  minus  $V_{G, \text{prime}}$  by  $V_{G, \text{prime}}$  plus half. So this is for this part then this part is half log twice  $V_{G, \text{prime}}$  minus  $V_{OL}$  by  $V_{OL}$ . So once we solve this differential equation, you get a relation like this.  $C_L$  by  $K_n V_{G, \text{prime}}$  into  $V_{OH}$  minus  $V_{G, \text{prime}}$  by  $V_{G, \text{prime}}$  plus half  $\ln$  twice  $V_{G, \text{prime}}$  minus  $V_{OL}$  by  $V_{OL}$ . So this is the expression of the fall time. This is the fall time of the output waveform. So the point here is that when you change the input voltage, if there is an instantaneous change you don't get an instantaneous change at the output. It will take some time because the capacitance has to discharge. Now we see that of course for a good circuit this fall time must be as small as possible, you must try to reduce this fall time.

Now we see that the fall time is actually proportional to  $C_L$  by  $K_n$ . so  $C_L$  by  $K_n$  determines the fall time. Now what is this  $C_L$ ? It is the load capacitance. Suppose we are considering this inverter is driving a similar inverter then the  $C_L$  is nothing but the input capacitance of an inverter. So we can say  $C_L$  is equal to  $C_G$  where  $C_G$  is the gate capacitance or the input capacitance of this inverter and this input gate capacitance consist of a large number of components but this can be shown to a first approximation that  $C_G$  is nothing but  $C_{OX}$  that is the oxide capacitance per unit area into area.

What is this input area? If you look at the input of this CMOS, you have at NMOS and a PMOS. So  $C_{OX}$  is nothing but the oxide capacitance that is  $\epsilon_{OX}$  by  $T_{OX}$  where  $T_{OX}$  is the oxide thickness and area is the area of the NMOS and the PMOS. So this one will be equal to  $C_{OX}$  into  $W_n L_n$  plus  $W_p L_p$ .

(Refer Slide Time: 47:09)

$$C_L = C_G$$

$$\approx C_{ox} \cdot \text{Area}$$

$$= C_{ox} [W_n L_n + W_p L_p]$$

This  $W_n$  is the width of the NMOS transistor,  $L_n$  is the length of the NMOS transistor,  $W_p$  is the width of the PMOS transistor and  $L_p$  is the length of the PMOS transistor. So this  $C_L$  in this expression which we have for the fall time is actually equal to this  $C_{ox}$  into  $W_n L_n$  plus  $W_p L_p$ .  $K_n$  is of course we know as  $\frac{1}{2} \mu C_{ox} W$  by  $L$ . So both this  $C_L$  is dependent on the dimensions of the device that is  $W$  and  $L$  of the transistors,  $k$  is also dependent on the  $W$  and  $L$  ratios; it is also dependent on mobility. Both of them are dependent on  $C_{ox}$ ,  $C_L$  is also proportional to  $C_{ox}$ ,  $K$  is also proportional to  $C_{ox}$ . So in fact this ratio is independent of  $C_{ox}$  but it is dependent on the dimensions of the device.

In the next class what we shall do is we shall find out how the delays depend of the dimensions. In fact we shall see that the length of the device is much more important for this delay because if you look at this expression again just this  $C_L$ , the load is proportional to  $L$ . So  $C_L$  increases with increase in  $L$  whereas  $K_n$  is proportional  $W$  by  $L$  so  $K_n$  reduces with increase in  $L$ . So this is proportional to  $L$  this is inversely proportional to  $L$ . So  $C_L$  by  $K_n$  will be proportional to  $L$  square. On the other hand this is proportional to  $W$ ,  $K_n$  is also proportional to  $W$ . It is proportional  $W$  by  $L$ , this also proportional to  $W$ .

This ratio is independent of  $W$ . It is actually the  $L$  or the length of the device which is more important which is very easily understood because you see that as the widths increase, the capacitance increase but the  $W$  by  $L$  ratio increase which means that you have larger current. So you have a larger current to drive a larger capacitance so there

is no problem but if the  $L$  increases, the capacitance increases but the current goes down because the current is proportional  $W$  by  $L$ . So you have a smaller current to drive a larger capacitance that is to charge a larger capacitance. Smaller current charging a larger capacitance so the delay is going to increase. So we shall take up in the next class, the effect of the  $W$ 's and  $L$ 's the widths and links on the delays of the CMOS inverter.

Thank you.