MOS TRANSISTOR THEORY

Introduction

A MOS transistor is a majority-carrier device, in which the current in a conducting channel between the source and the drain is modulated by a voltage applied to the gate.

Symbols

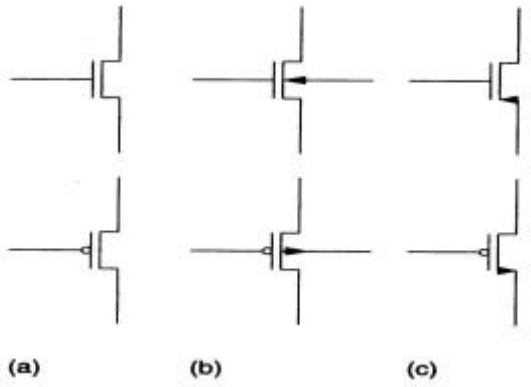

Figure 1 : symbols of various types of transistors.

NMOS (n-type MOS transistor)

(1) Majority carrier = electrons

(2) A positive voltage applied on the gate with respect to the substrate enhances the number of electrons in the channel and hence increases the conductivity of the channel.

(3) If gate voltage is less than a threshold voltage Vt , the channel is cut-off (very low current between source & drain).

PMOS (p-type MOS transistor)

(1) Majority carrier = holes

(2) Applied voltage is negative with respect to substrate.

Relationship between Vgs and Ids, for a fixed Vds:

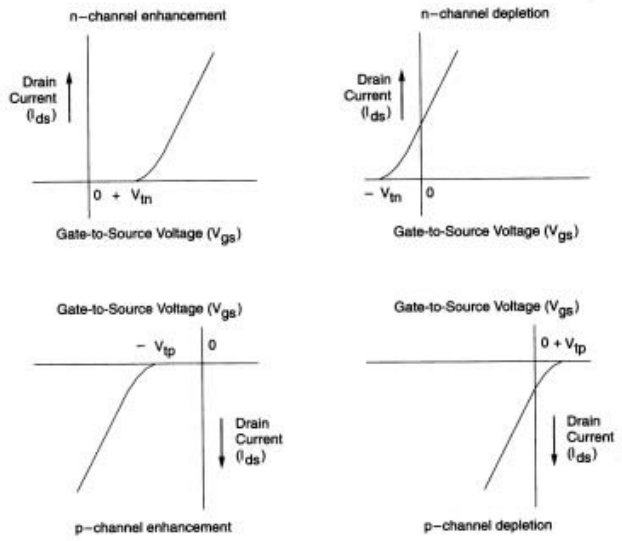

Figure 2: graph of Vgs vs Ids

Devices that are normally cut-off with zero gate bias are classified as **"enhancement-mode "**devices.
Devices that conduct with zero gate bias are called **"depletion-mode"**devices.
Enhancement-mode devices are more popular in practical use.

Threshold voltage (*Vt*):
The voltage at which an MOS device begins to conduct ("turn on"). The **threshold voltage** is a function of
(1) Gate conductor material
(2) Gate insulator material
(3) Gate insulator thickness
(4) Impurity at the silicon-insulator interface
(5) Voltage between the source and the substrate Vsb
(6) Temperature

MOS equations (Basic DC equations):

Three MOS operating regions are:Cutoff or subthreshold region, linear region and saturation region.
The following equation describes all these three regions:

$$I_{ds} = \begin{cases} 0; & V_{gs} - V_t \le 0 \quad cut\text{-}off \\ \beta\left[(V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right]; & 0 < V_{ds} < V_{gs} - V_t \quad linear \\ \frac{\beta}{2}(V_{gs} - V_t)^2; & 0 < V_{gs} - V_t < V_{ds} \quad saturation \end{cases}$$

where β is MOS transistor gain and it is given by $\beta=\mu\varepsilon/t_{ox}(W/L)$
again 'μ' is the mobility of the charge carrier
'ε' is the permittivity of the oxide layer.
't_{ox}'is the thickness of the oxide layer.
'W' is the width of the transistor.(shown in diagram)
'L' is the channel length of the transistor.(shown in diagram)

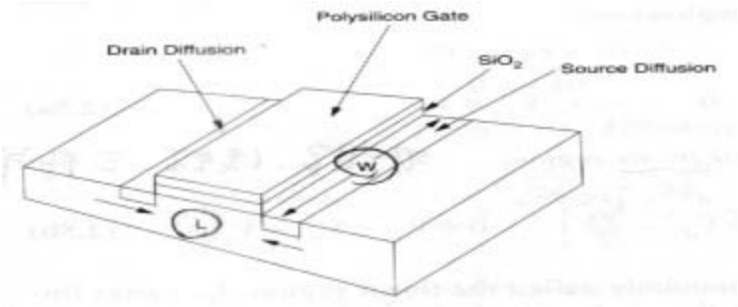

Diagram just to show the length and width of a MOSFET.

The graph of Id and Vds for a given Vgs is given below:

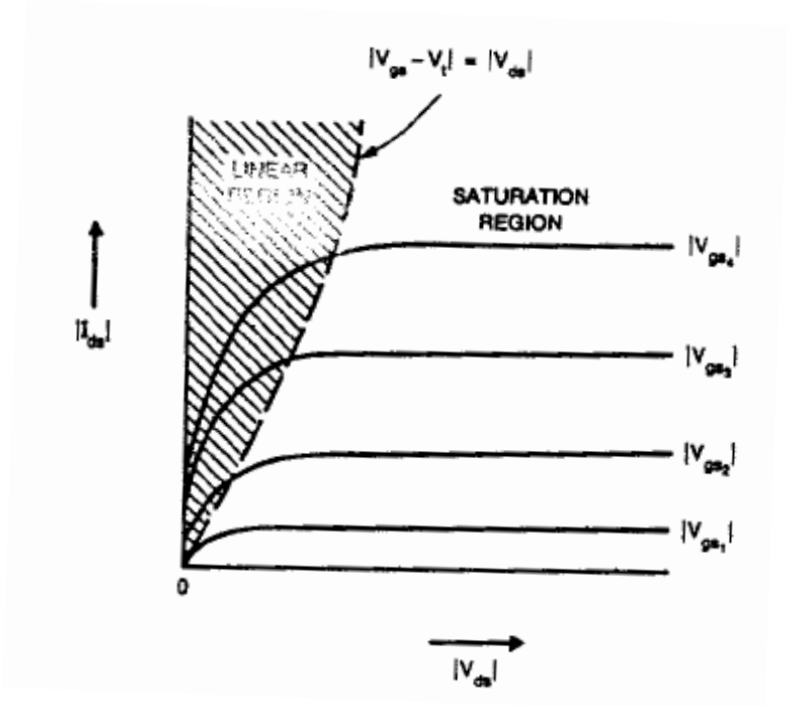

Figure 3: VI Characteristics of MOSFET

Second Order Effects:
Following are the list of second order effects of MOSFET.

- Threshold voltage – Body effect
- Subthreshold region
- Channel length modulation
- Mobility variation
- Fowler_Nordheim Tunneling
- Drain Punchthrough
- Impact Ionization – Hot Electrons

Threshold voltage – Body effect

The change in the threshold voltage of a MOSFET, because of the voltage difference between body and source is called body effect. The expression for the threshold voltage is given by the following expression.

$$Vt=V_{t(0)}+\gamma[(V_{sb}+2\Phi_F)^{1/2}-(2\Phi_F)^{1/2}]$$

where
Vt is the threshold voltage,
$V_{t(0)}$ is the threshold voltage without body effect
γ is the body coefficient factor
Φ_F is the fermi potential
V_{sb} is the potential difference between source and substarte.

If Vsb is zero, then Vt=Vt(0) that means the value of the threshold voltage will not be changed. Therefore, we short circuit the source and substrate so that, Vsb will be zero.

Subthreshold region:
For Vgs<Vt also we will get some value of Drain current this is called as Subthreshold current and the region is called as Subthreshold region.

Channel length modulation:
The channel length of the MOSFET is changed due to the change in the drain to source voltage. This effect is called as the channel length modulation. The effective channel length & the value of the drain current considering channel length modulation into effect is given by,

$$I_{ds} = \frac{\beta}{2}((V_{gs} - V_t)^2(1 + \lambda V_{ds}))$$

$$L_{eff} = L - \sqrt{2\varepsilon_o \frac{\varepsilon_{Si}}{qN}(V_{ds} - [V_{gs} - V_t])}.$$

Where λ is the channel length modulation factor.

Mobility:

Mobility is the defined as the ease with which the charge carriers drift in the substrate material. Mobility decreases with increase in doping concentration and increase in temperature. Mobility is the ratio of average carrier drift velocity and electric field. Mobility is represented by the symbol μ.

Fowler Nordhiem tunneling:

When the gate oxide is very thin there can be a current between gate and source or drain by electron tunneling through the gate oxide. This current is proportional to the area of the gate of the transistor.

Drain punchthough:

When the drain is a high voltage, the depletion region around the drain may extend to the source, causing the current to flow even it gate voltage is zero. This is known as Punchthrough condition.

Impact Ionization-hot electrons:

When the length of the transistor is reduced, the electric field at the drain increases. The field can be come so high that electrons are imparted with enough energy we can term them as hot. These hot electrons impact the drain, dislodging holes that are then swept toward the negatively charged substrate and appear as a substrate current. This effect is known as Impact Ionization.

MOS Models

MOS model includes the Ideal Equations, Second-order Effects plus the additional Curve-fitting parameters. Many semiconductor vendors expend a lot of effects to model the devices they manufacture.(Standard : Level 3 SPICE) . Main SPICE DC parameters in level 1,2,3 in 1μn-well CMOS process.

CMOS INVETER CHARACTERISTICS

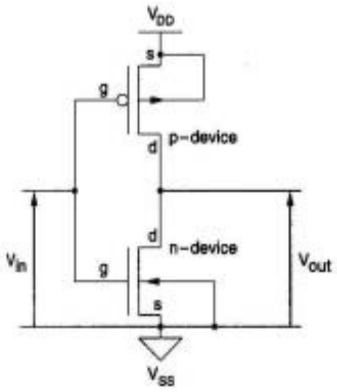

Figure 4: CMOS Inverter

CMOS inverters (Complementary NOSFET Inverters) are some of the most widely used and adaptable MOSFET inverters used in chip design. They operate with very little power loss and at relatively high speed. Furthermore, the CMOS inverter has good logic buffer characteristics, in that, its noise margins in both low and high states are large.

A CMOS inverter contains a PMOS and a NMOS transistor connected at the drain and gate terminals, a supply voltage VDD at the PMOS source terminal, and a ground connected at the NMOS source terminal, were VIN is connected to the gate terminals and VOUT is connected to the drain terminals.(given in diagram). It is important to notice that the CMOS does not contain any resistors, which makes it more power efficient that a regular resistor-MOSFET inverter. As the voltage at the input of the CMOS device varies between 0 and VDD, the state of the NMOS and PMOS varies accordingly. If we model each transistor as a simple switch activated by VIN, the inverter's operations can be seen very easily:

MOSFET	Condition on MOSFET	State of MOSFET
NMOS	Vgs<Vtn	OFF
NMOS	Vgs>Vtn	ON
PMOS	Vsg<Vtp	OFF
PMOS	Vsg>Vtp	ON

The table given, explains when the each transistor is turning on and off. When VIN is low, the NMOS is "off", while the PMOS stays "on": instantly charging VOUT to

logic high. When Vin is high, the NMOS is "on and the PMOS is "off": taking the voltage at VOUT to logic low.

Inverter DC Characteristics:

Before we study the DC characteristics of the inverter we should examine the ideal characteristics of inverter which is shown below. The characteristic shows that when input is zero output will high and vice versa.

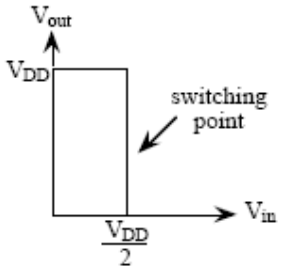

Figure 5: Ideal Characteristics of an Inverter

The actual characteristics is also given here for the reference. Here we have shown the status of both NMOS and PMOS transistor in all the regions of the characteristics.

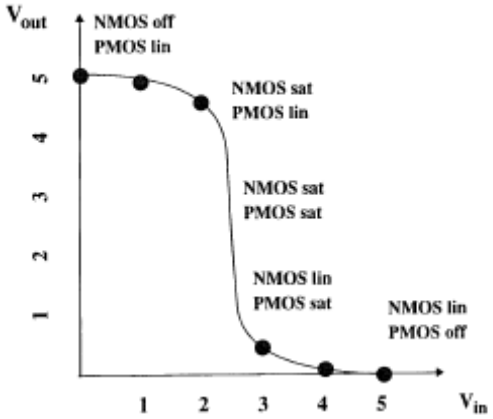

Figure 6: Actual Characteristics of an Inverter

Graphical Derivation of Inverter DC Characteristics:

The actual characteristics is drawn by plotting the values of output voltage for different values of the input voltage. We can also draw the characteristics, starting with the VI characteristics of PMOS and NMOS characteristics.

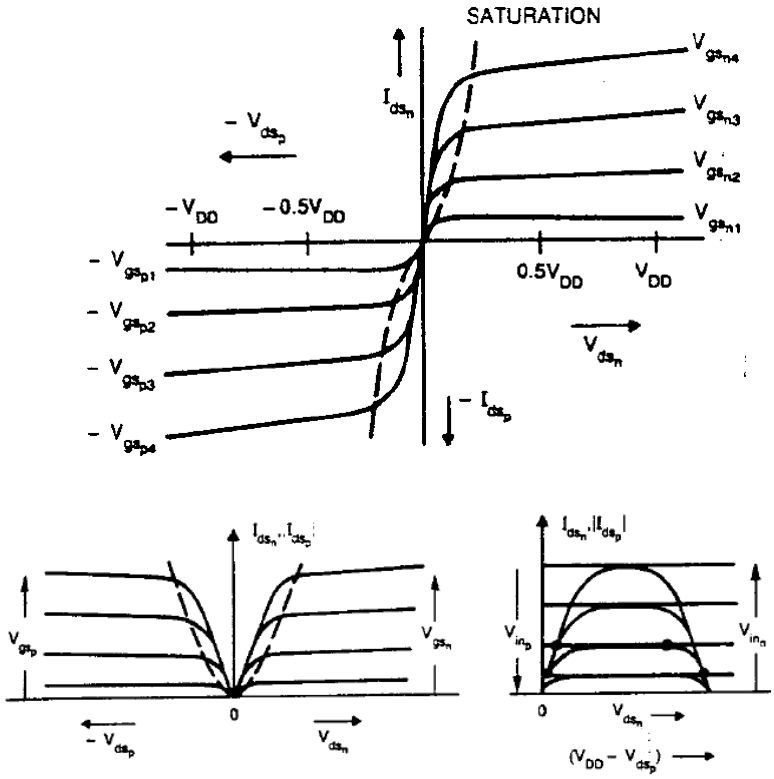

Figure 7a,b,c: Graphical Derivation of DC Characteristics

The characteristics given in figure 7a is the vi characteristics of the NMOS and PMOS characteristics (plot of Id vs. Vds). The figure 7b shows the values of drain current of PMOS transistor is taken to the positive side the current axis. This is done by taking the absolute value of the current. By superimposing both characteristics it leads to figure 7c. the actual characteristics may be now determined by the points of common Vgs intersection as shown in figure 7d.

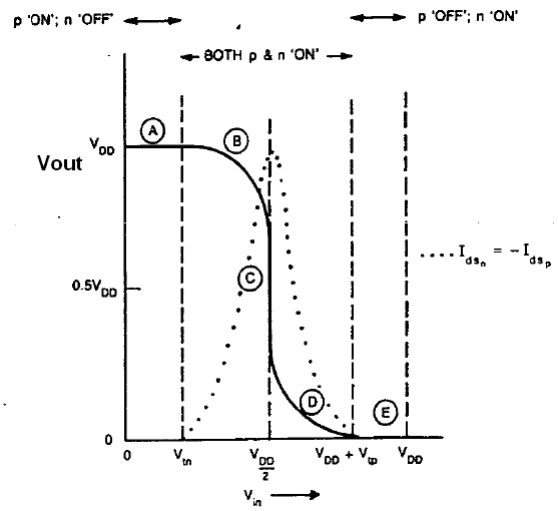

Figure 7d: Cmos Inverter Dc Characteristics

Figure 7d shows five regions namely region A, B, C, D & E. also we have shown a dotted curve which is the current that is drawn by the inverter.

Region A:
The output in this region is High because the P device is OFF and n device is ON.
In region A, NMOS is cutoff region and PMOS is on, therefore output is logic high.
We can analyze the inverter when it is in region B. the analysis is given below:

Region B:
The equivalent circuit of the inverter when it is region B is given below.

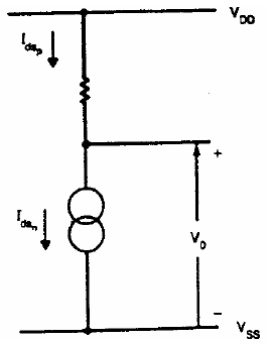

Figure 8: Equivalent circuit in Region B

In this region PMOS will be in linear region and NMOS is in saturation region.

The expression for the NMOS current is $I_{ds_n} = \beta_n \frac{[V_{in} - V_{t_n}]^2}{2}$,

The expression for the PMOS current is $I_{ds_p} = -\beta_p \left[(V_{in} - V_{DD} - V_{t_p})(V_O - V_{DD}) - \frac{1}{2}(V_O - V_{DD})^2 \right]$

The expression for the voltage Vo can be written as

$$V_O = (V_{in} - V_{t_p}) + \left[(V_{in} - V_{t_n})^2 - 2\left(V_{in} - \frac{V_{DD}}{2} - V_{t_p} \right) V_{DD} - \frac{\beta_n}{\beta_p}(V_{in} - V_{t_n})^2 \right]^{1/2}$$

Region C:

The equivalent circuit of CMOS inverter when it is in region C is given here. Both n and p transistors are in saturation region, we can equate both the currents and we can obain the expression for the mid point voltage or switching point voltage of a inverter. The corresponding equations are as follows:

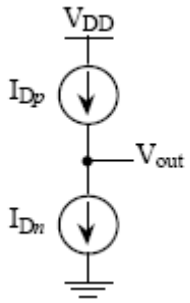

Figure 9: Equivalent circuit in Region C

The corresponding equations are as follows:

$$I_{ds_p} = \frac{1}{2}\beta_p(V_{in} - V_{DD} - V_{t_p})^2$$

$$I_{ds_n} = \frac{1}{2}\beta_n(V_{in} - V_{t_n})^2$$

By equating both the currents, we can obtain the expression for the switching point voltage as,

$$V_{in} = \frac{V_{DD} + V_{t_p} + V_{t_n}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}}.$$

Region D: the equivalent circuit for region D is given in the figure below.

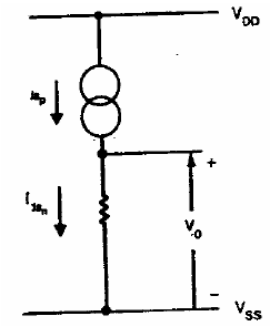

Figure 10: equivalent circuit in region D

We can apply the same analysis what we did for region B and C and we can obtain the expression for output voltage.

Region E:
The output in this region is zero because the P device is OFF and n device is ON.

Influence of βn/βp on the VTC characteristics:

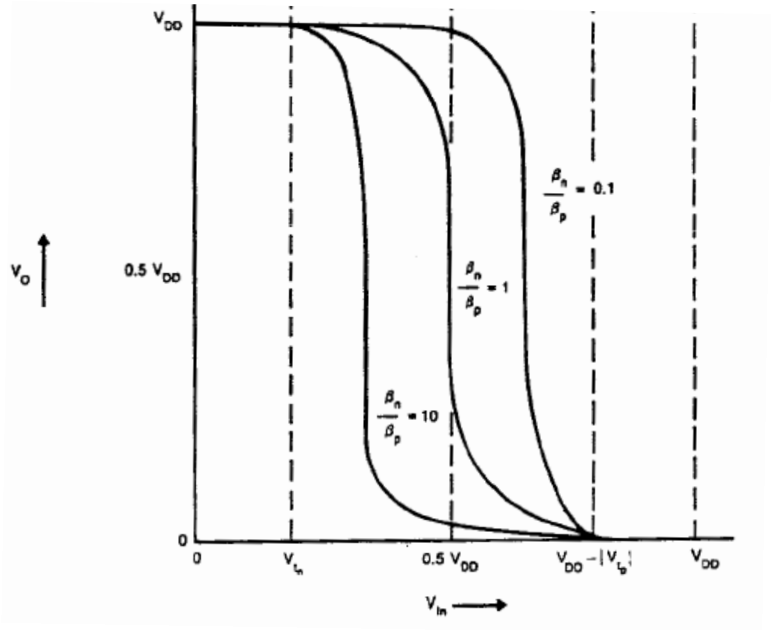

Figure 11: Effect of βn/βp ratio change on the DC characteristics of CMOS inverter

The characteristics shifts left if the ratio of βn/βp is greater than 1(say 10). The curve shifts right if the ratio of βn/βp is lesser than 1(say 0.1). this is decided by the switching point equation of region C. the equation is repeated here the reference again.

$Vm=Vsp=V_{DD}+Vtp+Vtn(\beta n/\beta p)1/2 / 1+(\beta n/\beta p)1/2$

Noise Margin:

Noise margin is a parameter related to input output characteristics. It determines the allowable noise voltage on the input so that the output is not affected.

We will specify it in terms of two things:

- LOW noise margin
- HIGH noise margin

LOW noise margin: is defined as the difference in magnitude between the maximum Low output voltage of the driving gate and the maximum input Low voltage recognized by the driven gate.

$NM_L=|V_{ILmax} - V_{OLmax}|$

HIGH noise margin: is defined difference in magnitude between minimum High output voltage of the driving gate and minimum input High voltage recognized by the receiving gate.

$\mathbf{NM_H=|Voh_{min} - V_{IHmin}|}$

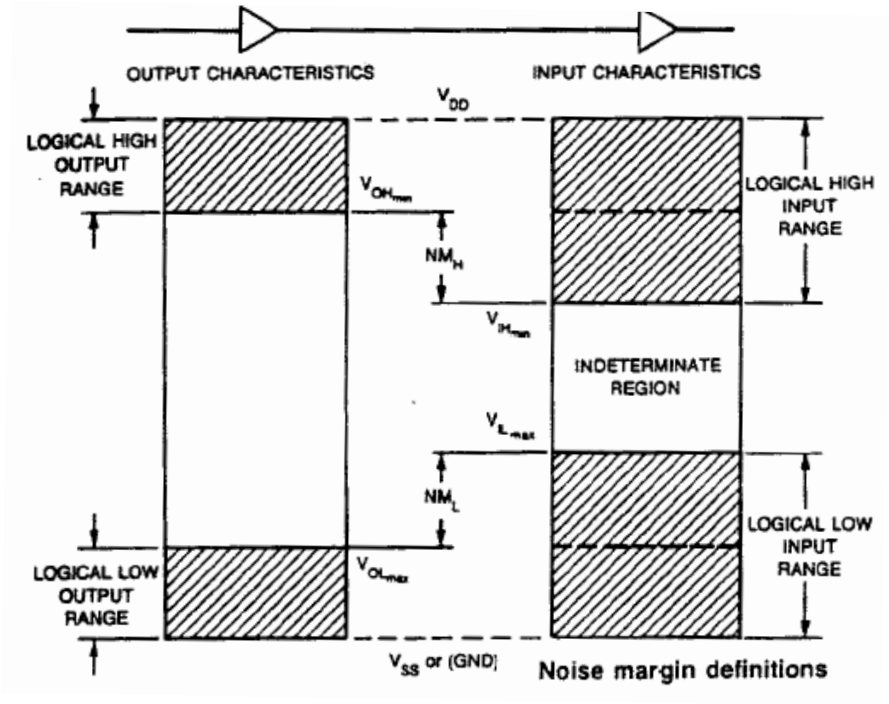

Figure 12: noise margin definitions

Figure shows how exactly we can find the noise margin for the input and output.
We can also find the noise margin of a CMOS inverter. The following figure gives the idea of calculating the noise margin.

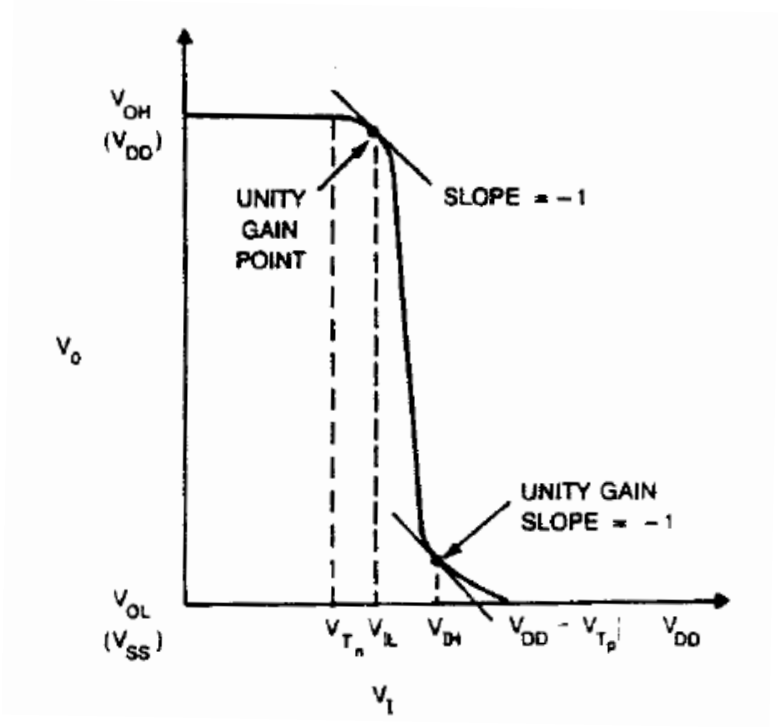

Figure 13: CMOS inverter noise margins

Static Load MOS inverters:

In the figure given below we have shown a resistive load and current source load inverter. Usually resistive load inverters are not preferred because of the power consumption and area issues.

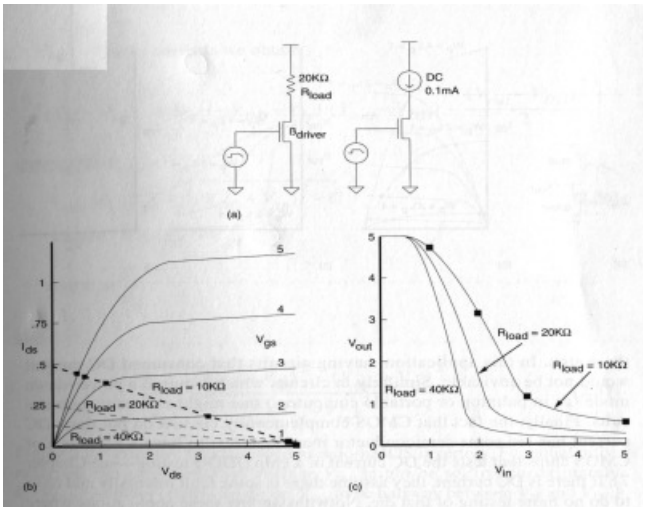

Figure 14: static load inverter

Pseudo-NMOS inverter:

This circuit uses the load device which is p device and is made to turn on always by connecting the gate terminal to the ground.

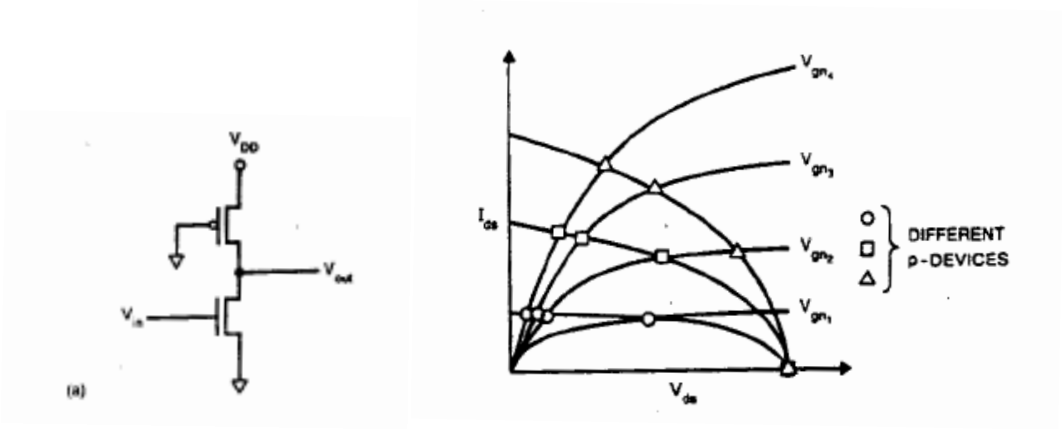

Figure 15: Pseudo-NMOS inverter

Power consumption is High compared to CMOS inverter particularly when NMOS device is ON because the p load device is always ON.

Saturated load inverter:

The load device is an nMOS transistor in the saturated load inverter. This type of inverter was used in nMOS technologies prior to the availability of nMOS depletion loads.

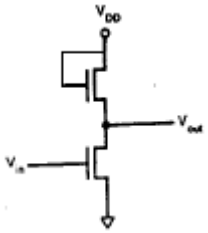

Figure 16: Saturated load inverter

Transmission gates:

It's a parallel combination of pmos and nmos transistor with the gates connected to a complementary input. After looking into various issues of pass transistors we will come back to the TGs again.

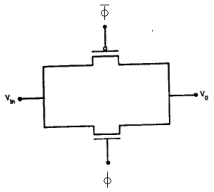

Figure 17: Transmission gate

Pass transistors:
We have n and p pass transistors.

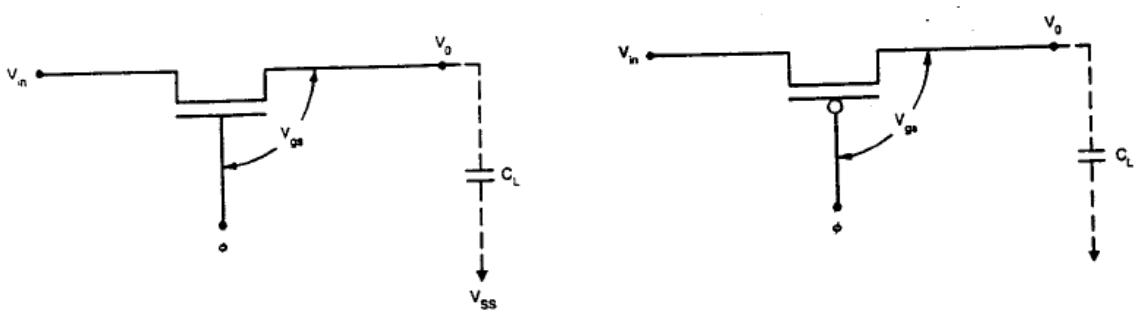

Figure 18: n and p pass transistors

The disadvantage with the pass transistors is that, they will not be able to transfer the logic levels properly. The following table gives that explanation in detail.

Transmission characteristics of *n*-channel and *p*-channel pass transistors

DEVICE	TRANSMISSION OF '1'	TRANSMISSION OF '0'
n	poor	good
p	good	poor

If Vdd (5 volts) is to be transferred using nMOS the output will be (Vdd-Vtn). **POOR 1 or Weak Logic 1**
If Gnd(0 volts) is to be transferred using nMOS the output will be Gnd. **GOOD 0 or Strong Logic 0**
If Vdd (5 volts) is to be transferred using pMOS the output will be Vdd. **GOOD 1 or Strong Logic 1**
If Gnd(0 volts) is to be transferred using pMOS the output will be Vtp. **POOR 0 or Weak Logic 0.**

Transmission gates(TGs):
It's a parallel combination of pmos and nmos transistor with the gates connected to a complementary input. The disadvantages weak 0 and weak 1 can be overcome by using a TG instead of pass transistors.

Working of transmission gate can be explained better with the following equation.

When Φ='0' n and p device off, Vin=0 or 1, Vo='Z'
When Φ='1' n and p device on, Vin=0 or 1, Vo=0 or 1 , where 'Z' is high impedance.

One more important advantage of TGs is that the reduction in the resistance because two transistors will come in parallel and it is shown in the graph. The graph shows the

resistance of n and p pass transistors, and resistance of TG which is lesser than the other two.

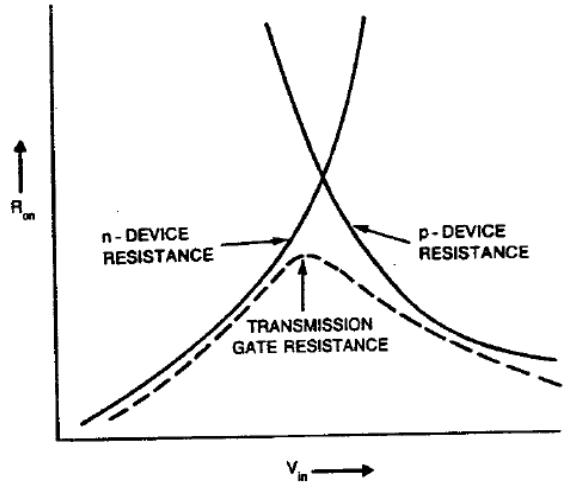

Figure 19: Graph of resistance vs. input for pass transistors and TG

Tristate Inverter:

By cascading a transmission gate with an inverter the tristate inverter circuit can be obtained. The working can be explained with the help of the circuit.

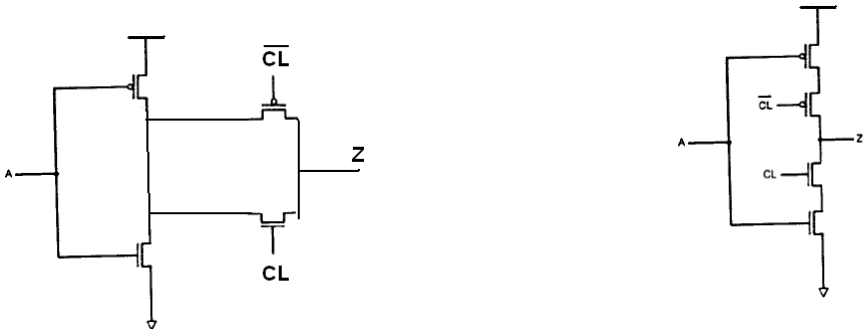

Figure 20: Tristate Inverter

The two circuits are the same only difference is the way they are written. When CL is zero the output of the inverter is in tristate condition. When CL is high the output is Z is the inversion of the input A