UNIT-I

IC Technologies

- Introduction
- MOS
- PMOS
- NMOS
- CMOS

  &
- BiCMOS

  Technologies

Basic Electrical Properties of MOS and BiCMOS Circuits

- I_{DS} - V_{DS} relationships
- MOS transistor Threshold Voltage - V_T figure of merit-ω0
- Transconductance-g_m, g_{ds};
- Pass transistor
- NMOS Inverter, Various pull ups, CMOS Inverter analysis and design
- Bi-CMOS Inverters

INTRODUCTION TO IC TECHNOLOGY

The development of electronics endless with invention of vaccum tubes and associated electronic circuits. This activity termed as vaccum tube electronics, afterward the evolution of solid state devices and consequent development of integrated circuits are responsible for the present status of communication, computing and instrumentation.

- The first vaccum tube diode was invented by **john ambrase Fleming** in 1904.
- The vaccum triode was invented by **lee de forest** in 1906.

Early developments of the Integrated Circuit (IC) go back to 1949. German engineer Werner Jacobi filed a patent for an IC like semiconductor amplifying device showing five transistors on a common substrate in a 2-stage amplifier arrangement. Jacobi disclosed small cheap of hearing aids.

Integrated circuits were made possible by experimental discoveries which showed that semiconductor devices could perform the functions of vacuum tubes and by mid-20th-century technology advancements in semiconductor device fabrication.

The integration of large numbers of tiny transistors into a small chip was an enormous improvement over the manual assembly of circuits using electronic components.

The integrated circuits mass production capability, reliability, and building-block approach to circuit design ensured the rapid adoption of standardized ICs in place of designs using discrete transistors.

An integrated circuit (IC) is a small semiconductor-based electronic device consisting of fabricated transistors, resistors and capacitors. Integrated circuits are the building blocks of most electronic devices and equipment. An integrated circuit is also known as a chip or microchip.

There are two main advantages of ICs over discrete circuits: cost and performance. Cost is low because the chips, with all their components, are printed as a unit by photolithography rather than being constructed one transistor at a time. Furthermore, much less material is used to construct a packaged IC die than a discrete circuit. Performance is high since the components switch quickly and consume little power (compared to their discrete counterparts) because the components are small and positioned close together. As of 2006, chip areas range from a few square millimeters to around 350 mm^2, with up to 1 million transistors per mm

IC Invention:

Inventor	Year	Circuit	Remark
Fleming	1904 1906	Vacuum tube diode Vacuum triode	large expensive, power-hungry, unreliable
William Shockley (Bell labs)	1945	Semiconductor replacing vacuum tube	--
Bardeen and Brattain and Shockley (Bell labs)	1947	Point Contact transfer resistance device "BJT"	Driving factor of growth of the VLSI technology
Werner Jacobi (Siemens AG)	1949	1st IC containing amplifying Device 2stage amplifier	No commercial use reported
Shockley	1951	Junction Transistor	"Practical form of transistor"
Jack Kilby (Texas Instruments)	July 1958	Integrated Circuits F/F With 2-T Germanium slice and gold wires	Father of IC design
Noyce Fairchild Semiconductor	Dec. 1958	Integrated Circuits Silicon	"The Mayor of Silicon Valley"
Kahng Bell Lab	1960	First MOSFET	Start of new era for semiconductor industry
Fairchild Semiconductor And Texas	1061	First Commercial IC	
Frank Wanlass (Fairchild Semiconductor)	1963	CMOS	
Federico Faggin (Fairchild Semiconductor)	1968	Silicon gate IC technology	Later Joined Intel to lead first CPU Intel 4004 in 1970 2300 T on 9mm^2
Zarlink Semiconductors	Recently	M2A capsule for endoscopy	take photographs of digestive tract 2/sec.

Moore's Law:

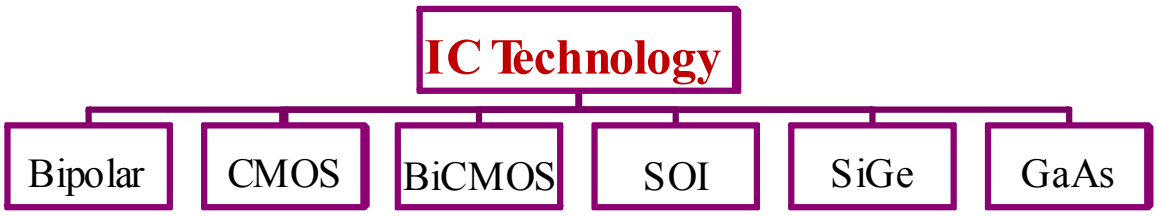
- Gordon E. Moore - Chairman Emeritus of Intel Corporation
- 1965 - observed trends in industry - of transistors on ICs vs release dates
- Noticed number of transistors doubling with release of each new IC generation
- Release dates (separate generations) were all 18-24 months apart

"The number of transistors on an integrated circuit will double every 18 months"

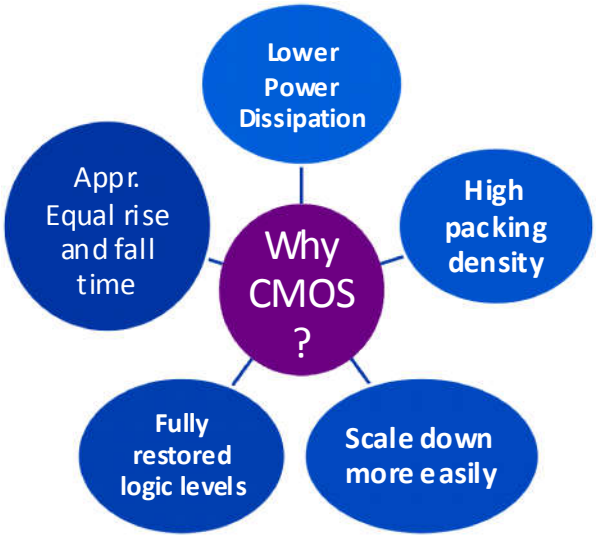
The level of integration of silicon technology as measured in terms of number of devices per IC Semiconductor industry has followed this prediction with surprising accuracy.

IC Technology:

- Speed / Power performance of available technologies
- The microelectronics evolution
- SIA Roadmap
- Semiconductor Manufacturers 2001 Ranking



Category	BJT	CMOS
Power Dissipation	Moderate to High	less
Speed	Faster	Fast
Gm	4ms	0.4ms
Switch implementation	poor	Good
Technology improvement	slower	Faster



Scale of Integration:

- **Small scale integration(SSI) --1960**

  The technology was developed by integrating the number of transistors of 1-100 on a single chip. Ex: Gates, flip-flops, op-amps.

- **Medium scale integration(MSI) --1967**

  The technology was developed by integrating the number of transistors of 100-1000 on a single chip. Ex: Counters, MUX, adders, 4-bit microprocessors.

- **Large scale integration(LSI) --1972**

  The technology was developed by integrating the number of transistors of 1000-10000 on a single chip. Ex:8-bit microprocessors,ROM,RAM.

- **Very large scale integration(VLSI) -1978**

  The technology was developed by integrating the number of transistors of 10000-1Million on a single chip. Ex:16-32 bit microprocessors, peripherals, complimentary high MOS.

- **Ultra large scale integration(ULSI)**

  The technology was developed by integrating the number of transistors of 1Million-10 Millions on a single chip. Ex: special purpose processors.

- **Giant scale integration(GSI)**

  The technology was developed by integrating the number of transistors of above 10 Millions on a single chip. Ex: Embedded system, system on chip.

- ✓ Fabrication technology has advanced to the point that we can put a complete system on a single chip.
- ✓ Single chip computer can include a CPU, bus, I/O devices and memory.
- ✓ This reduces the manufacturing cost than the equivalent board level system with higher performance and lower power.

MOS TECHNOLOGY:

MOS technology is considered as one of the very important and promising technologies in the VLSI design process. The circuit designs are realized based on pMOS, nMOS, CMOS and BiCMOS devices.

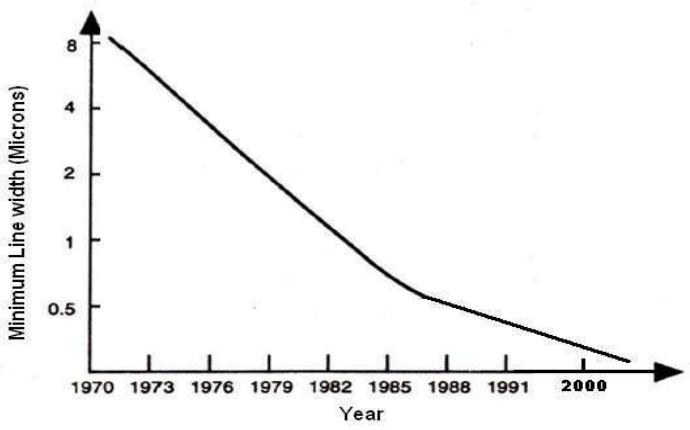
The pMOS devices are based on the p-channel MOS transistors. Specifically, the pMOS channel is part of a n-type substrate lying between two heavily doped p+ wells beneath the source and drain electrodes. Generally speaking, a pMOS transistor is only constructed in consort with an NMOS transistor.

The nMOS technology and design processes provide an excellent background for other technologies. In particular, some familiarity with nMOS allows a relatively easy transition to CMOS technology and design.

The techniques employed in nMOS technology for logic design are similar to GaAs technology.. Therefore, understanding the basics of nMOS design will help in the layout of GaAs circuits

In addition to VLSI technology, the VLSI design processes also provides a new degree of freedom for designers which helps for the significant developments. With the rapid advances in technology the the size of the ICs is shrinking and the integration density is increasing.

The minimum line width of commercial products over the years is shown in the graph below.



The graph shows a significant decrease in the size of the chip in recent years which implicitly indicates the advancements in the VLSI technology.

MOS Transistor Symbol:

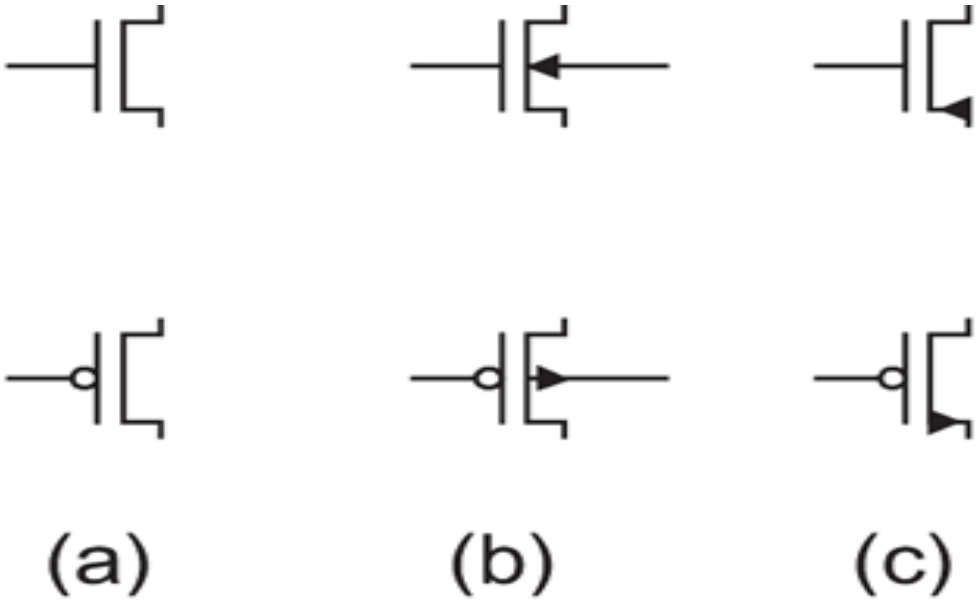


FIG 2.1 MOS transistor symbols

ENHANCEMENT AND DEPLETION MODE MOS TRANSISTORS

MOS Transistors are built on a silicon substrate. Silicon which is a group IV material is the eighth most common element in the universe by mass, but very rarely occurs as the pure free element in nature. It is most widely distributed in dusts, sands, planetoids, and planets as various forms of silicon dioxide (silica) or silicates. It forms crystal lattice with bonds to four neighbours. Silicon is a semiconductor. Pure silicon has no free carriers and conducts poorly. But adding dopants to silicon increases its conductivity. If a group V material i.e. an extra electron is added, it forms an n-type semiconductor. If a group III material i.e. missing electron pattern is formed (hole), the resulting semiconductor is called a p-type semiconductor.

A junction between p-type and n-type semiconductor forms a conduction path. Source and Drain of the Metal Oxide Semiconductor (MOS) Transistor is formed by the “doped” regions on the

surface of chip. Oxide layer is formed by means of deposition of the silicon dioxide (SiO_2) layer which forms as an insulator and is a very thin pattern. Gate of the MOS transistor is the thin layer of "polysilicon (poly)"; used to apply electric field to the surface of silicon between Drain and Source, to form a "channel" of electrons or holes. Control by the Gate voltage is achieved by modulating the conductivity of the semiconductor region just below the gate. This region is known as the channel.

The Metal–Oxide–Semiconductor Field Effect Transistor (MOSFET) is a transistor which is a voltage-controlled current device, in which current at two electrodes, drain and source is controlled by the action of an electric field at another electrode gate having in-between semiconductor and a very thin metal oxide layer. It is used for amplifying or switching electronic signals.

The Enhancement and Depletion mode MOS transistors are further classified as N-type named NMOS (or N-channel MOS) and P-type named PMOS (or P-channel MOS) devices. Figure 1.5 shows the MOSFETs along with their enhancement and depletion modes.

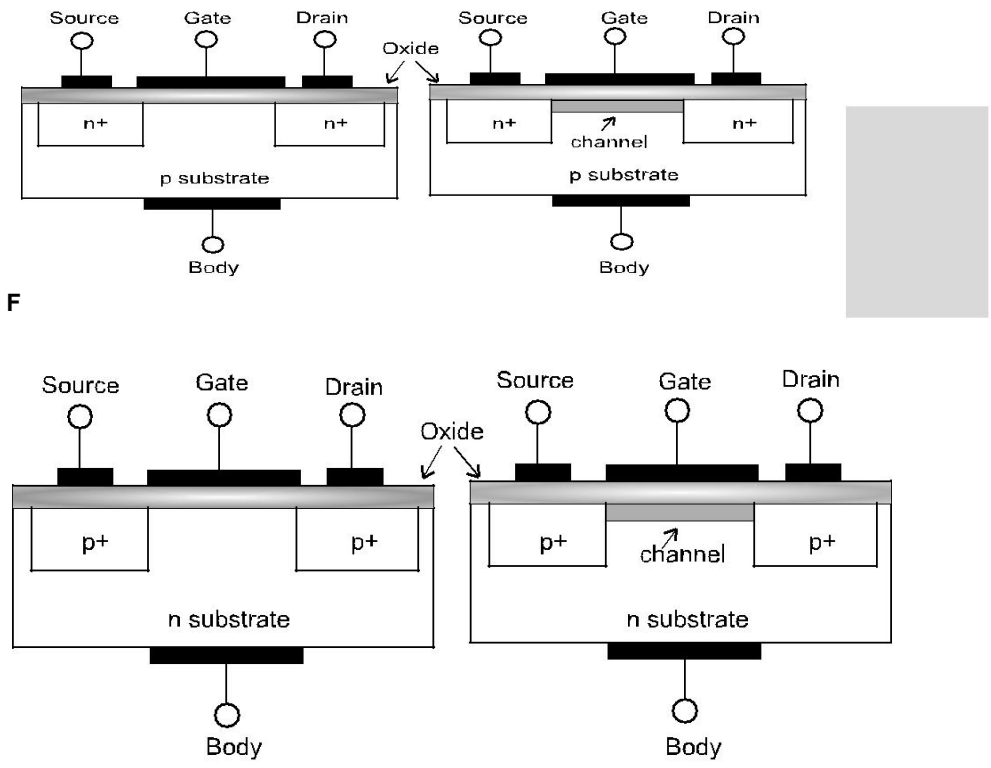


Figure 1.5: (c) Enhancement P-type MOSFET (d) Depletion P-type MOSFET

The depletion mode devices are doped so that a channel exists even with zero voltage from gate to source during manufacturing of the device. Hence the channel always appears in the device. To control the channel, a negative voltage is applied to the gate (for an N-channel device), depleting the

channel, which reduces the current flow through the device. In essence, the depletion-mode device is equivalent to a closed (ON) switch, while the enhancement-mode device does not have the built in channel and is equivalent to an open (OFF) switch. Due to the difficulty of turning off the depletion mode devices, they are rarely used

Working of Enhancement Mode Transistor

The enhancement mode devices do not have the in-built channel. By applying the required potentials, the channel can be formed. Also for the MOS devices, there is a threshold voltage (V_t), below which not enough charges will be attracted for the channel to be formed. This threshold voltage for a MOS transistor is a function of doping levels and thickness of the oxide layer.

Case 1: $V_{gs} = 0V$ and $V_{gs} < V_t$

The device is non-conducting, when no gate voltage is applied ($V_{gs} = 0V$) or ($V_{gs} < V_t$) and also drain to source potential $V_{ds} = 0$. With an insufficient voltage on the gate to establish the channel region as N-type, there will be no conduction between the source and drain. Since there is no conducting channel, there is no current drawn, i.e. $I_{ds} = 0$, and the device is said to be in the **cut-off region**. This is shown in the Figure 1.7 (a).

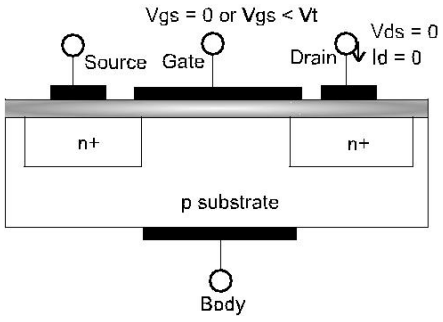


Figure 1.7: (a) Cut-off Region

Case 2: Vgs > Vt

When a minixmum voltage greater than the threshold voltage V_t (i.e. $V_{gs} > V_t$) is applied, a high concentration of negative charge carriers forms an inversion layer located by a thin layer next to the interface between the semiconductor and the oxide insulator. This forms a channel between the source and drain of the transistor. This is shown in the Figure 1.7 (b).

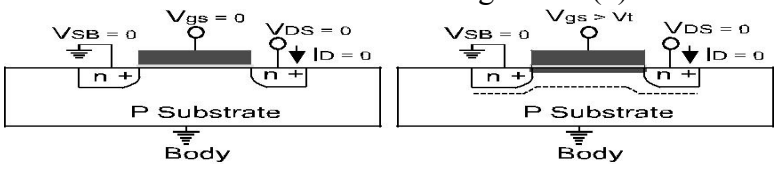


Figure 1.7: (b) Formation of a Channel

A positive V_{ds} reverse biases the drain substrate junction, hence the depletion region around the drain widens, and since the drain is adjacent to the gate edge, the depletion region widens in the channel. This is shown in Figure 1.7 (c). This results in flow of electron from source to drain resulting in current Ids.. The device is said to operate in **linear region** during this phase. Further increase in $V_{ds,}$ increases the reverse bias on the drain substrate junction in contact with the inversion layer which causes inversion layer density to decrease. This is shown in Figure 1.7 (d). The point at which the inversion layer density becomes very small (nearly zero) at the drain end is termed pinch-off. The value of V_{ds} at pinch-off is denoted as $V_{ds,sat}$. This is termed as **saturation region** for the MOS device. Diffusion current completes the path from source to drain in this case, causing the channel to exhibit a high resistance and behaves as a constant current source.

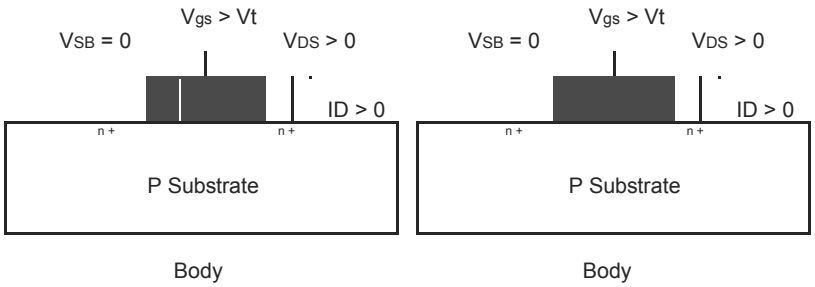


Figure 1.7: (c) Linear Region. (d) Saturation Region

The MOSFET I_D versus V_{DS} characteristics (V-I Characteristics) is shown in the Figure 1.8. For $V_{GS} < V_{t,}$ $I_D = 0$ and device is in cut-off region. As V_{DS} increases at a fixed V_{GS}, I_D increases in the linear region due to the increased lateral field, but at a decreasing rate since the inversion layer density is decreasing. Once pinch-off is reached, further increase in V_{DS} results in increase in I_D; due to the formation of the high field region which is very small. The device starts in linear region, and moves into saturation region at higher V_{DS}.

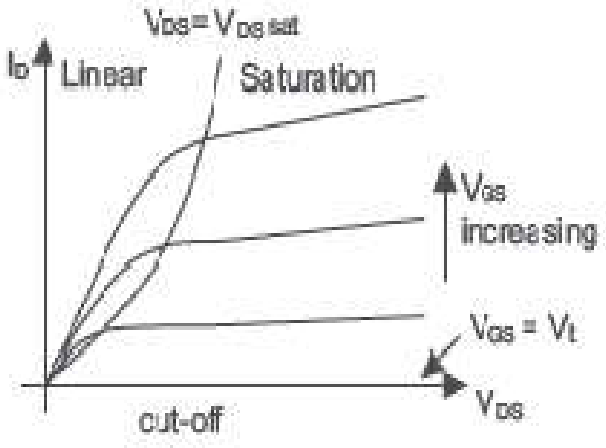


Figure 1.8: MOS V-I Characteristics

NMOS FABRICATION

The following description explains the basic steps used in the process of fabrication.

(*a*) The fabrication process starts with the oxidation of the silicon substrate.

It is shown in the Figure 1.9 (a).

(*b*) A relatively thick silicon dioxide layer, also called field oxide, is created on the surface of the substrate. This is shown in the Figure 1.9 (b).

(*c*) Then, the field oxide is selectively etched to expose the silicon surface on which the MOS transistor will be created. This is indicated in the Figure 1.9 (c).

(*d*) This is followed by covering the surface of substrate with a thin, high-quality oxide layer, which will eventually form the gate oxide of the

MOS transistor as illustrated in Figure 1.9 (d).

(*e*) On top of the thin oxide, a layer of polysilicon (polycrystalline silicon) is deposited as is shown in the Figure 1.9 (e). Polysilicon is used both as gate electrode material for MOS transistors and also as an interconnect medium in silicon integrated circuits. Undoped polysilicon has relatively high resistivity. The resistivity of polysilicon can be reduced, however, by doping it with impurity atoms.

(*f*) After deposition, the polysilicon layer is patterned and etched to form the interconnects and the MOS transistor gates. This is shown in Figure 1.9 (f).

(*g*) The thin gate oxide not covered by polysilicon is also etched along, which exposes the bare silicon surface on which the source and drain junctions are to be formed (Figure 1.9 (g)).

(*h*) The entire silicon surface is then doped with high concentration of impurities, either through diffusion or ion implantation (in this case with donor atoms to produce n-type doping). Diffusion is achieved by heating the wafer to a high temperature and passing the gas containing desired impurities over the surface. Figure 1.9 (h) shows that the doping penetrates the exposed areas on the silicon surface, ultimately creating two n-type regions (source and drain junctions) in the p-type substrate. The impurity doping also penetrates the polysilicon on the surface, reducing its resistivity.

(*i*) Once the source and drain regions are completed, the entire surface is again covered with an insulating layer of silicon dioxide, as shown in

Figure 1.9 (i).(*j*) The insulating oxide layer is then patterned in order to provide contact windows for the drain and source junctions, as illustrated in Figure 1.9 (j).

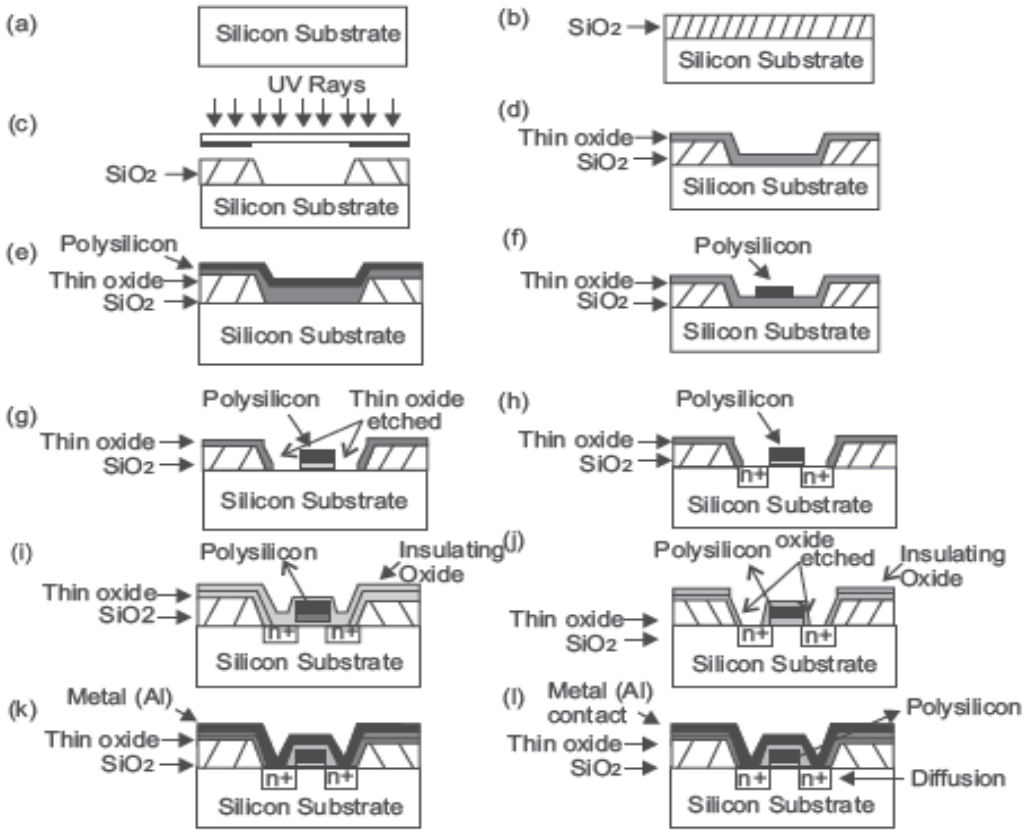


Figure 1.9: Fabrication Process of NMOS Device

CMOS FABRICATION:

CMOS fabrication can be accomplished using either of the three technologies:

- N-well technologies/P-well technologies
- Twin well technology
- Silicon On Insulator (SOI)

The fabrication of CMOS can be done by following the below shown twenty steps, by which CMOS can be obtained by integrating both the NMOS and PMOS transistors on the same chip substrate. For integrating these NMOS and PMOS devices on the same chip, special regions called as wells or tubs are required in which semiconductor type and substrate type are opposite to each other.

A P-well has to be created on a N-substrate or N-well has to be created on a P-substrate. In this article, the fabrication of CMOS is described using the P-substrate, in which the NMOS transistor is fabricated on a P-type substrate and the PMOS transistor is fabricated in N-well.

The fabrication process involves twenty steps, which are as follows:

N-Well Process

Step1: Substrate

Primarily, start the process with a P-substrate.

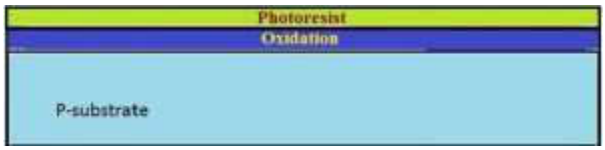
Step2: Oxidation

The oxidation process is done by using high-purity oxygen and hydrogen, which are exposed in an oxidation furnace approximately at 1000 degree centigrade.

Step3: Photoresist

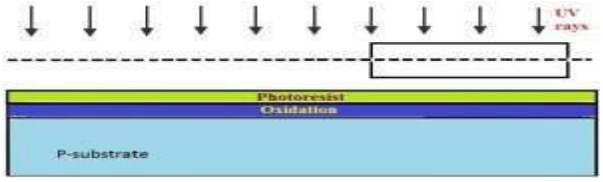
A light-sensitive polymer that softens whenever exposed to light is called as Photoresist layer.

It is formed.



Step4: Masking

The photoresist is exposed to UV rays through the N-well mask

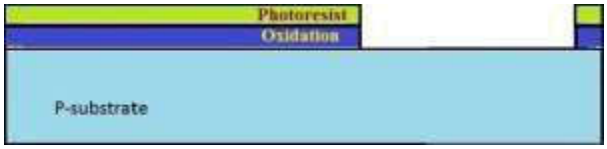


Step5: Photoresist removal

A part of the photoresist layer is removed by treating the wafer with the basic or acidic solutio n.

Step6: Removal of SiO2 using acid etching

The SiO2 oxidation layer is removed through the open area made by the removal of photoresist using hydrofluoric acid.



Step7: Removal of photoresist

The entire photoresist layer is stripped off, as shown in the below figure.

Step8: Formation of the N-well

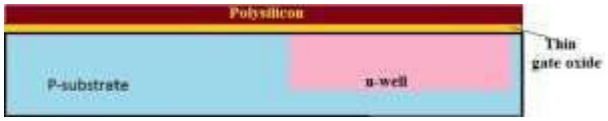
By using ion implantation or diffusion process N-well is formed.

Step9: Removal of SiO2

Using the hydrofluoric acid, the remaining SiO2 is removed.

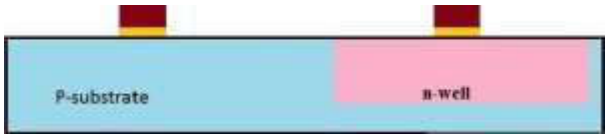
Step10: Deposition of polysilicon

Chemical Vapor Deposition (CVD) process is used to deposit a very thin layer of gate oxide.



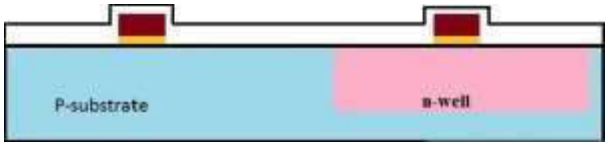
Step11: Removing the layer barring a small area for the Gates

Except the two small regions required for forming the Gates of NMOS and PMOS, the remaining layer is stripped off.



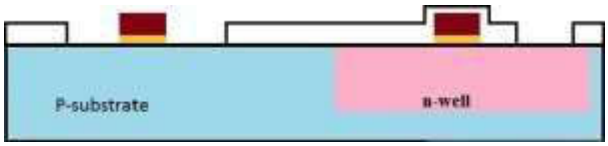
Step12: Oxidation process

Next, an oxidation layer is formed on this layer with two small regions for the formation of the gate terminals of NMOS and PMOS.

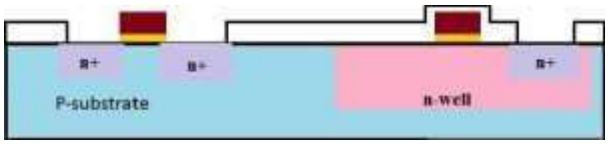


Step13: Masking and N-diffusion

By using the masking process small gaps are made for the purpose of N -diffusion.

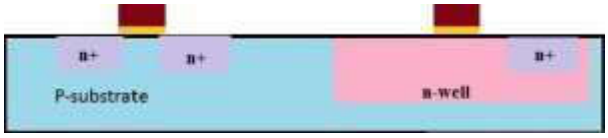


The n-type (n+) dopants are diffused or ion implanted, and the three n+ are formed for the formation of the terminals of NMOS.



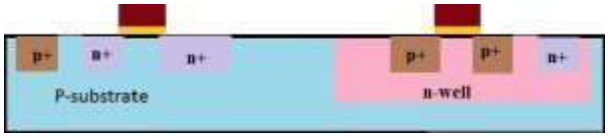
Step14: Oxide stripping

The remaining oxidation layer is stripped off.



Step15: P-diffusion

Similar to the above N-diffusion process, the P-diffusion regions are diffused to form the terminals of the PMOS.

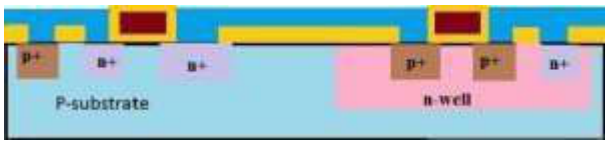


Step16: Thick field oxide

A thick-field oxide is formed in all regions except the terminals of the PMOS and NMOS.

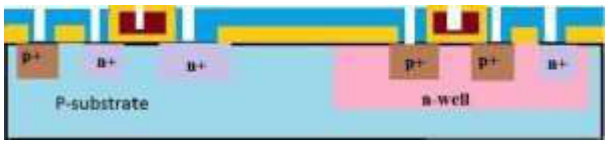
Step17: Metallization

Aluminum is sputtered on the whole wafer.



Step18: Removal of excess metal

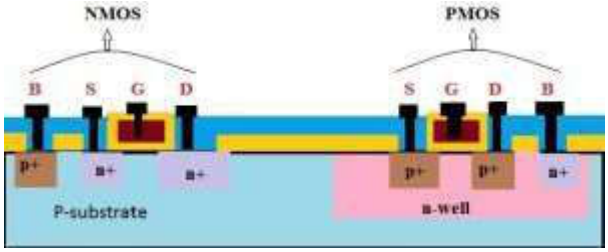
The excess metal is removed from the wafer layer.



Step19: Terminals

The terminals of the PMOS and NMOS are made from respective gaps.

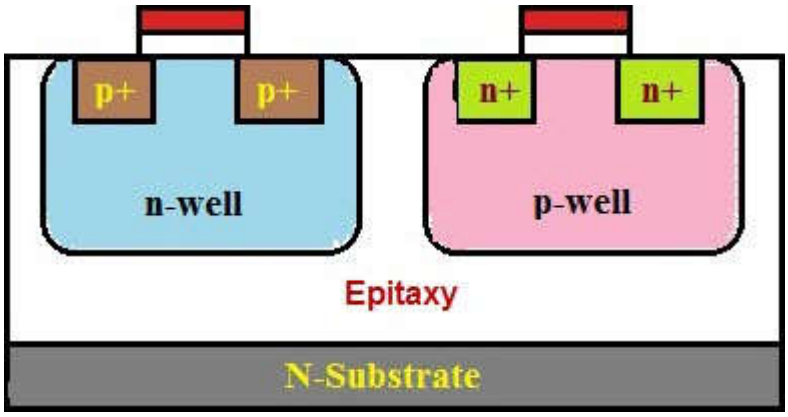
Step20: Assigning the names of the terminals of the NMOS and PMOS



Fabrication of CMOS using P-well process

Among all the fabrication processes of the CMOS, N-well process is mostly used for the fabrication of the CMOS. P-well process is almost similar to the N-well. But the only difference in p-well process is that it consists of a main N-substrate and, thus, P-wells itself acts as substrate for the N-devices.

Twin tub-CMOS Fabrication Process



In this process, separate optimization of the n-type and p-type transistors will be provided. The independent optimization of Vt, body effect and gain of the P-devices, N-devices can be made possible with this process.

Different steps of the fabrication of the CMOS using the twintub process are as follows:

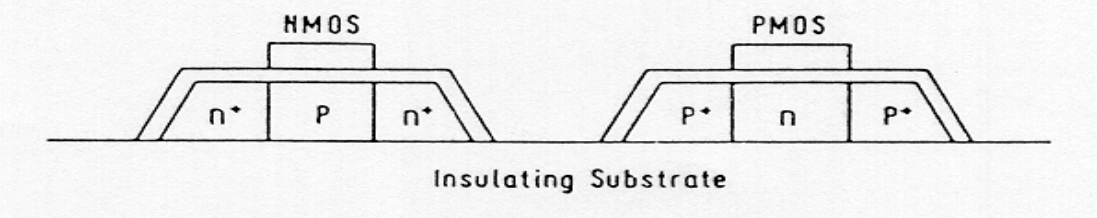
- Lightly doped n+ or p+ substrate is taken and, to protect the latch up, epitaxial layer is used.
- The high-purity controlled thickness of the layers of silicon are grown with exact dopant concentrations.
- The dopant and its concentration in Silicon are used to determine electrical properties.
- Formation of the tub
- Thin oxide construction

- Implantation of the source and drain
- Cuts for making contacts
- Metallization

By using the above steps we can fabricate CMOS using twin tub process method.

Silicon-on-Insulator (SOI) CMOS Process

Rather than using silicon as the substrate material, technologists have sought to use an insulating substrate to improve process characteristics such as speed and latch-up susceptibility. The SOI CMOS technology allows the creation of independent, completely isolated nMOS and pMOS transistors virtually side-by-side on an insulating substrate. The main advantages of this technology are the higher integration density (because of the absence of well regions), complete avoidance of the latch-up problem, and lower parasitic capacitances compared to the conventional p & n-well or twin-tub CMOS processes. A cross-section of nMOS and pMOS devicesusing SOI processis shown below.

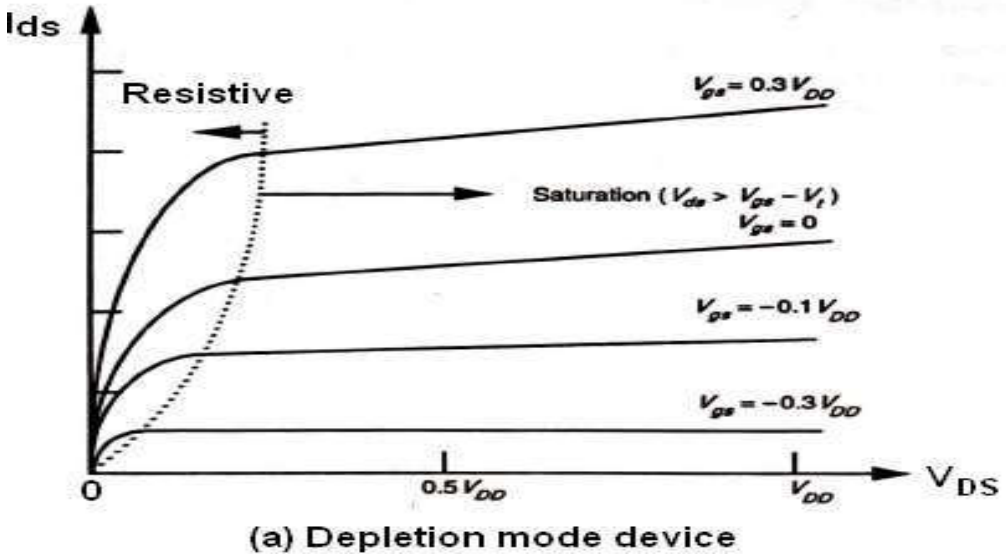


The SOI CMOS process is considerably more costly than the standard p & n-well CMOS process. Yet the improvements of device performance and the absence of latch-up problems can justify its use, especially for deep-sub-micron devices.

Basic Electrical Properties of MOS and Bi CMOS circuits

I_D-V_{DS} Characteristics of MOS Transistor :

The graph below shows the ID Vs VDS characteristics of an n- MOS transistor for several values of VGS .It is clear that there are two conduction states when the device is ON. The saturated state and the non-saturated state. The saturated curve is the flat portion and defines the saturation region. For Vgs < VDS + Vth, the nMOS device is conducting and ID is independent of VDS. For Vgs > VDS + Vth, the transistor is in the non-saturation region and the curve is a half parabola. When the transistor is OFF (Vgs < Vth), then ID is zero for any VDS value.

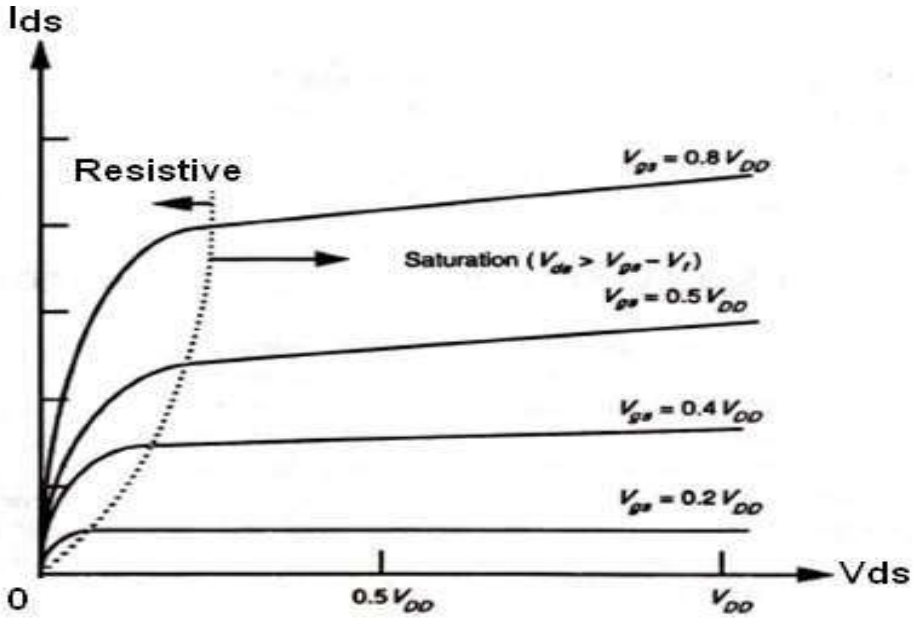


(a) Depletion mode device

The boundary of the saturation/non-saturation bias states is a point seen for each curve in the graph as the intersection of the straight line of the saturated region with the quadratic curve of the non-saturated region. This intersection point occurs at the channel pinch off voltage called VDSAT. The diamond symbol marks the pinch-off voltage VDSAT for each value of VGS. VDSAT is defined as the minimum drain-source voltage that is required to keep the transistor in saturation for a given VGS .In the non-saturated state, the drain current initially increases almost linearly from the origin before bending in a parabolic response. Thus the name ohmic or linear for the non- saturated region.

The drain current in saturation is virtually independent of VDS and the transistor acts as a current

source. This is because there is no carrier inversion at the drain region of the channel. Carriers are pulled into the high electric field of the drain/substrate pn junction and ejected out of the drain terminal.



(b). Enhance mode device

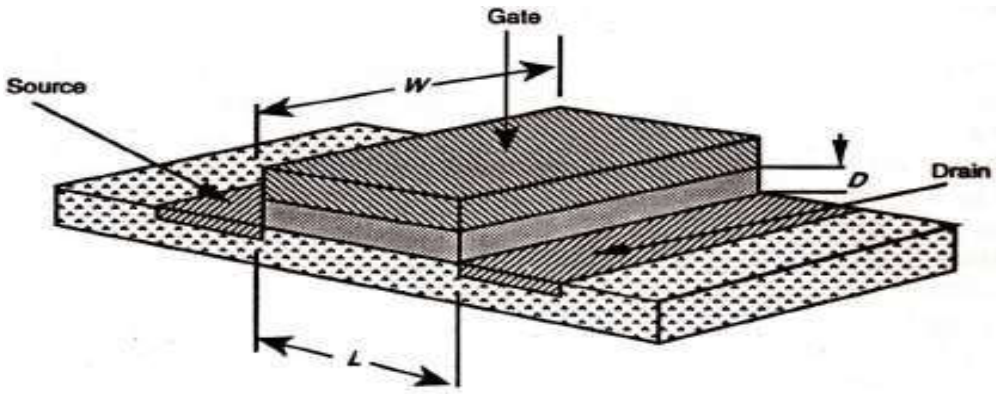
Drain-to-Source Current I_{DS} Versus Voltage V_{DS} Relationships :

The working of a MOS transistor is based on the principle that the use of a voltage on the gate induce a charge in the channel between source and drain, which may then be caused to move from source to drain under the influence of an electric field created by voltage Vds applied between drain and source. Since the charge induced is dependent on the gate to source voltage Vgs then Ids is dependent on both Vgs and Vds.

Let us consider the diagram below in which electrons will flow source to drain .So,the drain current is given by

Charge induced in channel (Qc) Ids =-Isd = Electron transit time(τ) Length of the channel Where the transit time is given by τ_{sd} = ----------

Velocity (v)



But velocity v= μEds

Where μ =electron or hole mobility and Eds = Electric field also , Eds = Vds/L

so,v = μ.Vds/L and $\tau_{ds} = L^2 / \mu.Vds$

The typical values of μ at room temperature are given below.

$$\mu_n \doteq 650 \text{ cm}^2/\text{V sec (surface)}$$
$$\mu_p \doteq 240 \text{ cm}^2/\text{V sec (surface)}$$

Non-saturated Region :

Let us consider the Id vs Vd relationships in the non-saturated region .The charge induced in the channel due to due to the voltage difference between the gate and the channel, Vgs (assuming substrate connected to source). The voltage along the channel varies linearly with distance *X* from the source due to the IR drop in the channel .In the non-saturated state the average value is Vds/2. Also the effective gate voltage Vg = Vgs – Vt where Vt, is the threshold voltage needed to invert the charge under the gate and establish the channel.

Hence the induced charge is $Q_c = E_g \varepsilon_{ins} \varepsilon_o W.L$

Where

Eg = average electric field gate to channel

ε_{ins} = relative permittivity of insulation between gate and channel ε_o=permittivity

$$E_g = \frac{\left((V_{gs} - V_t) - \frac{V_{ds}}{2}\right)}{D}$$

Here D is the thickness of the oxide layer. Thus

$$Q_c = \frac{WL\varepsilon_{ins}\varepsilon_0}{D}\left((V_{gs} - V_t) - \frac{V_{ds}}{2}\right)$$

So, by combining the above two equations ,we get

$$I_{ds} = \frac{\varepsilon_{ins}\varepsilon_0\mu}{D}\frac{W}{L}\left((V_{gs} - V_t) - \frac{V_{ds}}{2}\right)V_{ds}$$

or the above equation can be written as

$$I_{ds} = K\frac{W}{L}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right)$$

In the non-saturated or resistive region where Vds < Vgs – Vt and

$$K = \frac{\varepsilon_{ins}\varepsilon_0\mu}{D}$$

Generally ,a constant β is defined as

$$\beta = K\frac{W}{L}$$

So that ,the expression for drain –source current will become

$$I_{ds} = \beta\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right)$$

The gate /channel capacitance is

$$C_g = \frac{\varepsilon_{ins}\varepsilon_0 WL}{D} \text{ (parallel plate)}$$

Hence we can write another alternative form for the drain current as

$$I_{ds} = \frac{C_g \mu}{L^2}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right)$$

Some time it is also convenient to use gate –capacitance per unit area ,C_g So,the drain current is

$$I_{ds} = C_0 \mu \frac{W}{L}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right)$$

This is the relation between drain current and drain-source voltage in non-saturated region.

Saturated Region

Saturation begins when Vds = Vgs - V, since at this point the IR drop in the channel equals the effective gate to channel voltage at the drain and we may assume that the current remains fairly constant as Vds increases further. Thus

$$I_{ds} = K\frac{W}{L}\frac{(V_{gs} - V_t)^2}{2}$$

or we can also write that

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

or it can also be written as

$$I_{ds} = \frac{C_g \mu}{2L^2}(V_{gs} - V_t)^2$$

or

$$I_{ds} = C_0 \mu \frac{W}{2L}(V_{gs} - V_t)^2$$

The expressions derived above for Ids hold for both enhancement and depletion mode devices. Here the threshold voltage for the nMOS depletion mode device (denoted as Vtd) is negative.

MOS Transistor Threshold Voltage Vt :

The gate structure of a MOS transistor consists, of charges stored in the dielectric layers and in the surface to surface interfaces as well as in the substrate itself. Switching an enhancement mode MOS transistor from the off to the on state consists in applying sufficient gate voltage to neutralize these charges and enable the underlying silicon to undergo an inversion due to the electric field from the gate. Switching a depletion mode nMOS transistor from the on to the off state consists in applying enough voltage to the gate to add to the stored charge and invert the 'n' implant region to 'p'.

The threshold voltage Vt may be expressed as:

$$V_t = \phi_{ms} \frac{Q_B - Q_{SS}}{C_0} + 2\phi_{fN}$$

where QD = the charge per unit area in the depletion layer below the oxide Qss = charge density at Si: SiO2 interface

C_O =Capacitance per unit area.

Φns = work function difference between gate and Si

ΦfN = Fermi level potential between inverted surface and bulk Si

For polynomial gate and silicon substrate, the value of Φ_{NS} is negative but negligible and the magnitude and sign of Vt are thus determined by balancing the other terms in the equation. To evaluate the Vt the other terms are determined as below.

$$Q_B = \sqrt{2\varepsilon_0\varepsilon_{Si}qN(2\phi_{fN} + V_{SB})} \text{ coulomb/m}^2$$

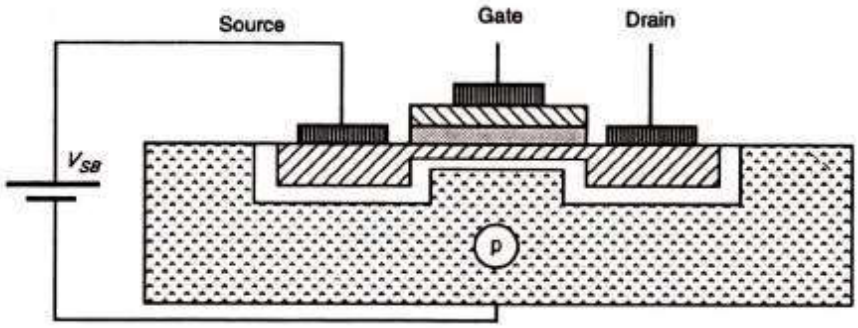
$$\phi_{fN} = \frac{kT}{q} \ln \frac{N}{n_i} \text{ volts}$$

$$Q_{SS} = (1.5 \text{ to } 8) \times 10^{-8} \text{ coulomb/m}^2$$

Body Effect :

Generally while studying the MOS transistors it is treated as a three terminal device. But, the body of the transistor is also an implicit terminal which helps to understand the characteristics of the transistor. Considering the body of the MOS transistor as a terminal is known as the body effect. The potential difference between the source and the body (Vsb) affects the threshold

voltage of the transistor. In many situations, this Body Effect is relatively insignificant, so we can (unless **otherwise** stated) ignore the Body Effect. But it is not always insignificant, in some cases it can have a tremendous impact on MOSFET circuit performance.



Body effect - nMOS device

Increasing Vsb causes the channel to be depleted of charge carriers and thus the threshold voltage is raised. Change in Vt is given by $\Delta Vt = \gamma.(Vsb)^{1/2}$ where γ is a constant which depends on substrate doping so that the more lightly doped the substrate, the smaller will be the body effect

The threshold voltage can be written as

$$V_t = V_t(0) + \left(\frac{D}{\varepsilon_{ins}\varepsilon_0}\right)\sqrt{2\varepsilon_0\varepsilon_{Si}QN_\cdot(V_{SB})^{1/2}}$$

Where Vt(0) is the threshold voltage for Vsd = 0

For n-MOS depletion mode transistors ,the body voltage values at different VDD voltages are given below.

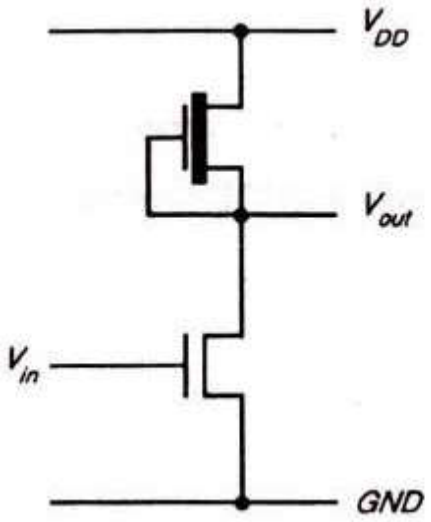
VSB = 0 V ; Vsd = -0.7VDD (= - 3.5 V for VDD =+5V) VSB = 5 V ; Vsd = -0.6VDD (= - 3.0 V for VDD =+5V)

nMOS INVERTER :

An inverter circuit is a very important circuit for producing a complete range of logic circuits. This is needed for restoring logic levels, for Nand and Nor gates, and for sequential and memory circuits of various forms .

A simple inverter circuit can be constructed using a transistor with source connected to ground and a load resistor of connected from the drain to the positive supply rail V_{DD}. The output is taken from the drain and the input applied between gate and ground .

But, during the fabrication resistors are not conveniently produced on the silicon substrate and even small values of resistors occupy excessively large areas .Hence some other form of load resistance is used. A more convenient way to solve this problem is to use a depletion mode transistor as the load, as shown in Fig. below.



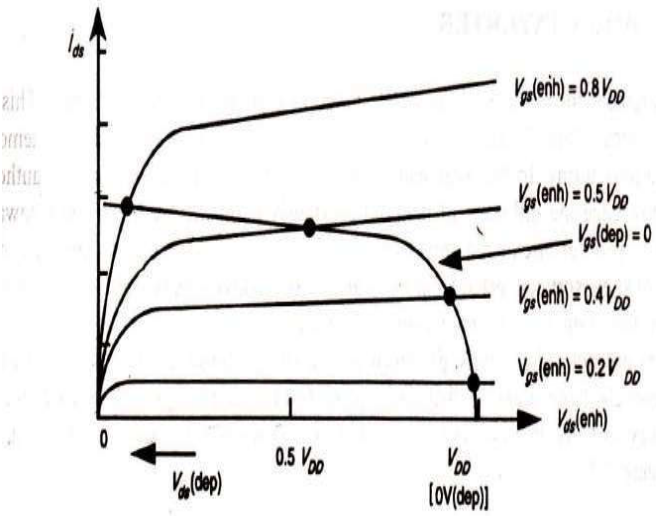
The salient features of the n-MOS inverter are

- For the depletion mode transistor, the gate is connected to the source so it is always on .
- In this configuration the depletion mode device is called the pull-up (P.U) and the enhancement mode device the pull-down (P.D) transistor.
- With no current drawn from the output, the currents Ids for both transistors must be equal.

nMOS Inverter transfer characteristic.

The transfer characteristic is drawn by taking Vds on x-axis and Ids on Y-axis for both enhancement and depletion mode transistors. So,to obtain the inverter transfer characteristic for

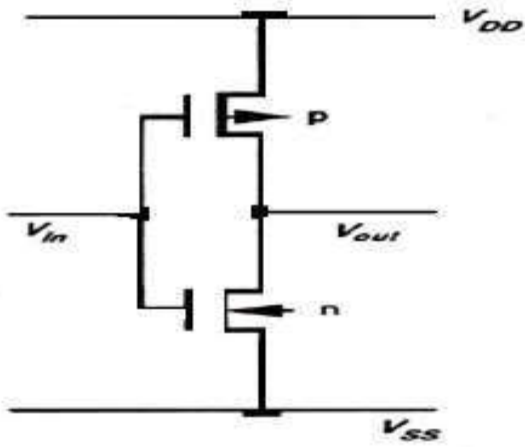
Vgs = 0 depletion mode characteristic curve is superimposed on the family of curves for the enhancement mode device and from the graph it can be seen that , maximum voltage across the enhancement mode device corresponds to minimum voltage across the depletion mode transistor.



From the graph it is clear that as Vin(=Vgs p.d. transistor) exceeds the Pulldown threshold voltage current begins to flow. The output voltage Vout thus decreases and the subsequent increases in Vin will cause the Pull down transistor to come out of saturation and become resistive.

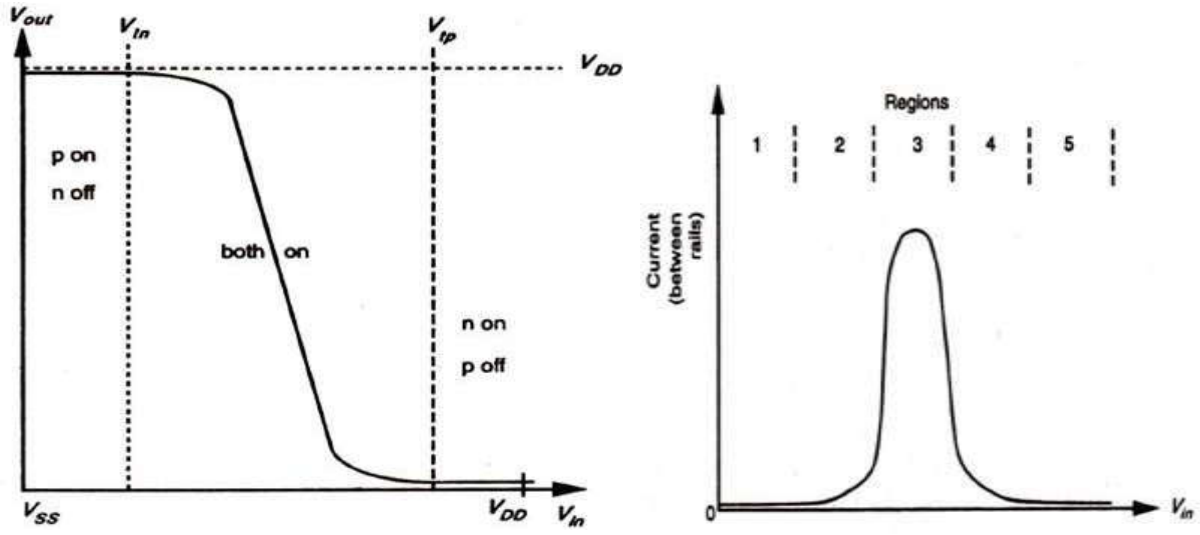
CMOS Inverter:

The inverter is the very important part of all digital designs. Once its operation and properties are clearly understood, Complex structures like NAND gates, adders, multipliers, and microprocessors can also be easily done. The electrical behavior of these complex circuits can be almost completely derived by extrapolating the results obtained for inverters. As shown in the diagram below the CMOS transistor is designed using p-MOS and n-MOS transistors.



In the inverter circuit ,if the input is high .the lower n-MOS device closes to discharge the capacitive load .Similarly ,if the input is low,the top p-MOS device is turned on to charge the capacitive load .At no time both the devices are on ,which prevents the DC current flowing from positive power supply to ground. Qualitatively this circuit acts like the switching circuit, since the p-channel transistor has exactly the opposite characteristics of the n-channel transistor. In the transition region both transistors are saturated and the circuit operates with a large voltage gain. The C-MOS transfer characteristic is shown in the below graph.

Considering the static conditions first, it may be Seen that in region 1 for which Vi,. = logic 0, we have the p-transistor fully turned on while the n-transistor is fully turned off. Thus no current flows through the inverter and the output is directly connected to VDD through the p-transistor.



Hence the output voltage is logic 1 . In region 5 , Vin = logic 1 and the n-transistor is fully on while the p-transistor is fully off. So, no current flows and logic 0 appears at the output.

In region 2 the input voltage has increased to a level which just exceeds the threshold voltage of the n-transistor. The n-transistor conducts and has a large voltage between source and drain; so it is in saturation. The p-transistor is also conducting but with only a small voltage across it, it operates in the unsaturated resistive region. A small current now flows through the inverter from VDD to VSS. If we wish to analyze the behavior in this region, we equate the p-device resistive region current with the n-device saturation current and thus obtain the voltage and current relationships.

Region 4 is similar to region 2 but with the roles of the p- and n-transistors reversed.However, the current magnitudes in regions 2 and 4 are small and most of the energy consumed in switching from one state to the other is due to the larger current which flows in region 3.

Region 3 is the region in which the inverter exhibits gain and in which both transistors are in saturation.

The currents in each device must be the same ,since the transistors are in series. So,we can write that

$$I_{dsp} = - I_{dsn}$$

where

$$I_{dsp} = \frac{\beta_p}{2} (V_{in} - V_{DD} - V_{tp})^2$$

and

$$I_{dsn} = \frac{\beta_n}{2} (V_{in} - V_{tn})^2$$

Since both transistors are in saturation, they act as current sources so that the equivalent circuit in this region is two current sources in series between VDD and Vss with the output voltage coming from their common point. The region is inherently unstable in consequence and the changeover from one logic level to the other is rapid.

Determination of Pull-up to Pull –Down Ratio *(Zp.u}Zp.d.)*for an nMOS Inverter driven by another nMOS Inverter :

Let us consider the arrangement shown in Fig.(a). in which an inverter is driven from the output of another similar inverter. Consider the depletion mode transistor for which Vgs = 0 under all conditions, and also assume that in order to cascade inverters without degradation the condition

$$V_{in} = V_{out} = V_{inv}$$

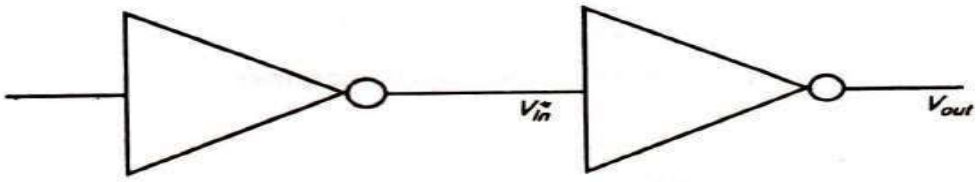


Fig.(a).Inverter driven by another inverter.

For equal margins around the inverter threshold, we set Vinv = 0.5V_{DD} · At this point both transistors are in saturation and we can write that

$$I_{ds} = K\frac{W}{L}\frac{(V_{gs} - V_t)^2}{2}$$

In the depletion mode $I_{ds} = K\frac{W_{p.u.}}{L_{p.u.}}\frac{(-V_{td})^2}{2}$ since $V_{gs} = 0$

and in the enhancement mode

$$I_{ds} = K\frac{W_{p.d.}}{L_{p.d.}}\frac{(V_{inv} - V_t)^2}{2} \text{ since } V_{gs} = V_{inv}$$

Equating (since currents are the same) we have

$$\frac{W_{p.d.}}{L_{p.d.}}(V_{inv} - V_t)^2 = \frac{W_{p.u.}}{L_{p.u.}}(-V_{td})^2$$

where $W_{p.d}$, $L_{p.d}$, $W_{p.u.}$ and $L_{p.u}$ are the widths and lengths of the pull-down and pull-up transistors respectively.

So,we can write that

$$Z_{p.d.} = \frac{L_{p.d.}}{W_{p.d.}}; Z_{p.u.} = \frac{L_{p.u.}}{W_{p.u.}}$$

we have

$$\frac{1}{Z_{p.d.}}(V_{inv} - V_t)^2 = \frac{1}{Z_{p.u.}}(-V_{td})^2$$

whence

$$V_{inv} = V_t - \frac{V_{td}}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

The typical, values for Vt ,Vinv and Vtd are

$$V_t = 0.2V_{DD};\ V_{td} = -0.6V_{DD}$$

$$V_{inv} = 0.5V_{DD} \text{ (for equal margins)}$$

Substituting these values in the above equation ,we get

$$0.5 = 0.2 + \frac{0.6}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

Here

$$\sqrt{Z_{p.u.}/Z_{p.d.}} = 2$$

So,we get

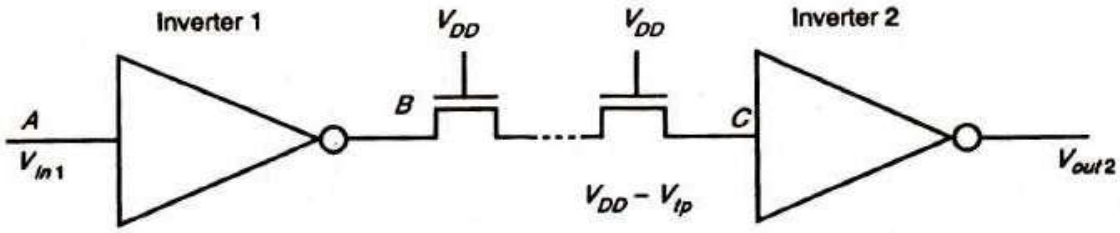
$$\boxed{Z_{p.u.}/Z_{p.d.} = 4/1}$$

This is the ratio for pull-up to pull down ratio for an inverter directly driven by another inverter.

Pull -Up to Pull-Down ratio for an nMOS Inverter driven through one or more Pass Transistors

Let us consider an arrangement in which the input to inverter 2 comes from the output of inverter 1

but passes through one or more nMOS transistors as shown in Fig. below (These transistors are called pass transistors).



The connection of pass transistors in series will degrade the logic 1 level / into inverter 2 so that the output will not be a proper logic 0 level. The critical condition is , when point A is at 0 volts and B is thus at VDD. but the voltage into inverter 2at point C is now reduced from VDD by the threshold voltage of the series pass transistor. With all pass transistor gates connected to VDD there is a loss of Vtp, however many are connected in series, since no static current flows through them and there can be no voltage drop in the channels. Therefore, the input voltage to inverter 2 is

Vin2 = VDD- Vtp where Vtp = threshold voltage for a pass transistor.

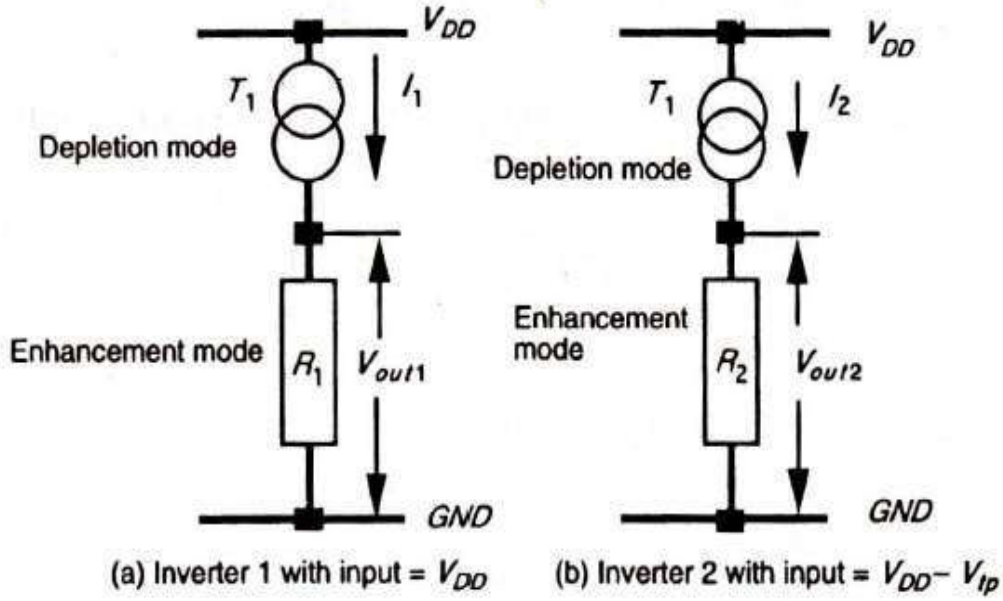
Let us consider the inverter 1 shown in Fig.(a) with input = VDD. If the input is at VDD , then the pull-down transistor T2 is conducting but with a low voltage across it; therefore, it is in its resistive region represented by R1 in Fig.(a) below. Meanwhile, the pull up transistor T1 is in saturation and is represented as a current source.

For the pull down transistor

$$R_1 = \frac{V_{ds1}}{I_{ds}} = \frac{1}{K}\frac{L_{p.d.1}}{W_{p.d.1}}\left(\frac{1}{V_{DD} - V_t - \frac{V_{ds1}}{2}}\right)$$

$$I_{ds} = K\frac{W_{p.d.1}}{L_{p.d.1}}\left((V_{DD} - V_t)\,V_{ds1} - \frac{V_{ds1}^2}{2}\right)$$

Since Vds is small, Vds/2 can be neglected in the above expression.



(a) Inverter 1 with input = V_{DD} (b) Inverter 2 with input = $V_{DD} - V_{tp}$

So,

$$R_1 \doteq \frac{1}{K} Z_{p.d.1} \left(\frac{1}{V_{DD} - V_t} \right)$$

Now, for depletion mode pull-up transistor in saturation with Vgs = 0

$$I_1 = I_{ds} = K \frac{W_{p.u.1}}{L_{p.u.1}} \frac{(-V_{td})^2}{2}$$

The product 1R1 = Vout1So,

$$V_{out1} = I_1 R_1 = \frac{Z_{p.d.1}}{Z_{p.u.1}} \left(\frac{1}{V_{DD} - V_t} \right) \frac{(V_{td})^2}{2}$$

Let us now consider the inverter 2 Fig.b .when input = VDD- Vtp.

$$R_2 \doteqdot \frac{1}{K} Z_{p.d.2} \frac{1}{((V_{DD} - V_{tp}) - V_t}$$

$$I_2 = K \frac{1}{Z_{p.u.2}} \frac{(-V_{td})^2}{2}$$

Whence,

$$V_{out\,2} = I_2 R_2 = \frac{Z_{p.d.2}}{Z_{p.u.2}} \left(\frac{1}{V_{DD} - V_{tp} - V_t} \right) \frac{(-V_{td})^2}{2}$$

If inverter 2 is to have the same output voltage under these conditions then $V_{out1} = V_{out2}$. That is

$I_1R_1=I_2R_2$, therefore

$$\frac{Z_{p.u.2}}{Z_{p.d.2}} = \frac{Z_{p.u.1}}{Z_{p.d.1}} \frac{(V_{DD} - V_t)}{(V_{DD} - V_{tp} - V_t)}$$

Considering the typical values

$$V_t = 0.2V_{DD}$$

$$V_{tp} = 0.3V_{DD}*$$

$$\frac{Z_{p.u.2}}{Z_{p.d.2}} = \frac{Z_{p.u.1}}{Z_{p.d.1}} \frac{0.8}{0.2}$$

Therefore

$$\frac{Z_{p.u.2}}{Z_{p.d.2}} \doteqdot 2 \frac{Z_{p.u.1}}{Z_{p.d.1}} = \frac{8}{1}$$

From the above theory it is clear that, for an n-MOS transistor

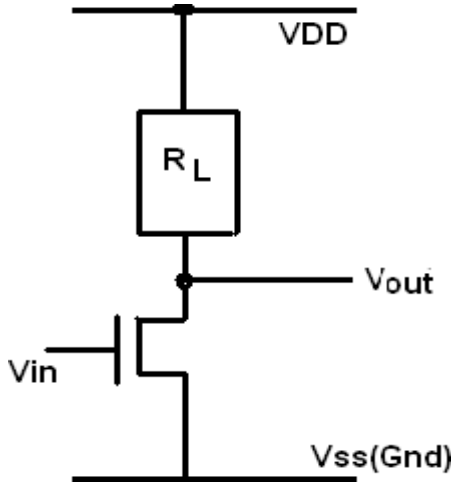
(i). An inverter driven directly from the output of another should have a $Z_{p.u}/ Z_{pd.}$ ratio of ≥ 4/1.

(ii).An inverter driven through one or more pass transistors should have a $Z_{p.u.}/Z_{p.d}$ ratio of ≥8/1

ALTERMTIVE FORMS OF PULL –UP

Generally the inverter circuit will have a depletion mode pull-up transistor as its load. But there are also other configurations .Let us consider four such arrangements.

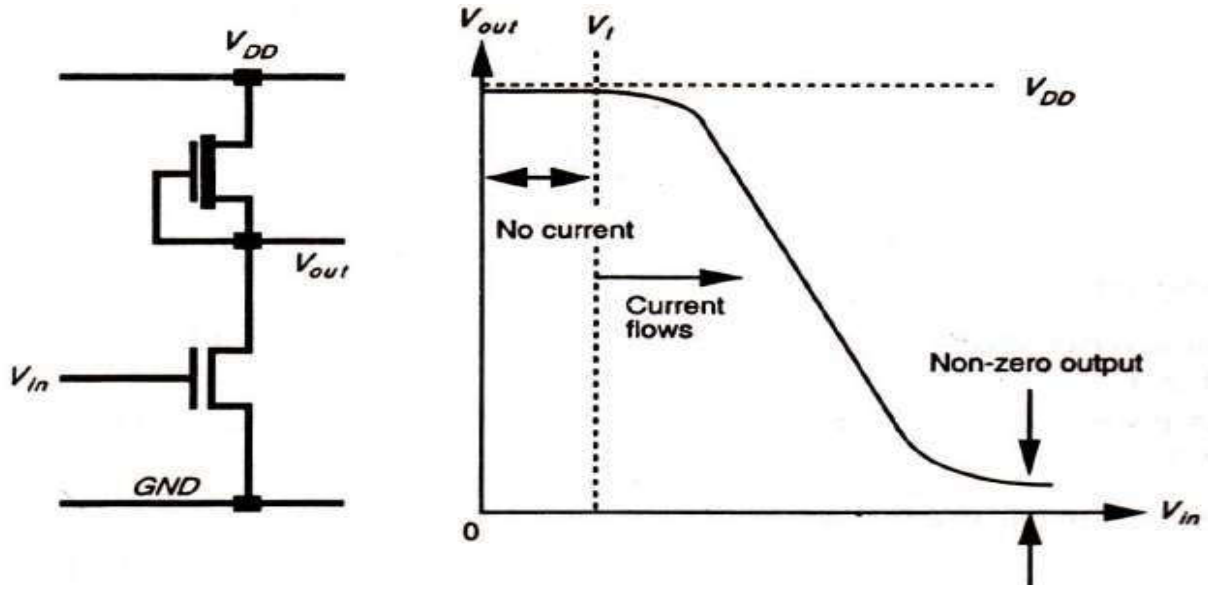
(i).Load resistance RL : This arrangement consists of a load resistor as apull-up as shown in the diagram below.But it is not widely used because of the large space requirements of resistors produced in a silicon substrate.



nMOS depletion mode transistor pull-up : This arrangement consists of a depletion mode transistor as pull-up. The arrangement and the transfer characteristic are shown below.In this type of arrangement we observe

(a) Dissipation is high , since rail to rail current flows when Vin = logical 1.

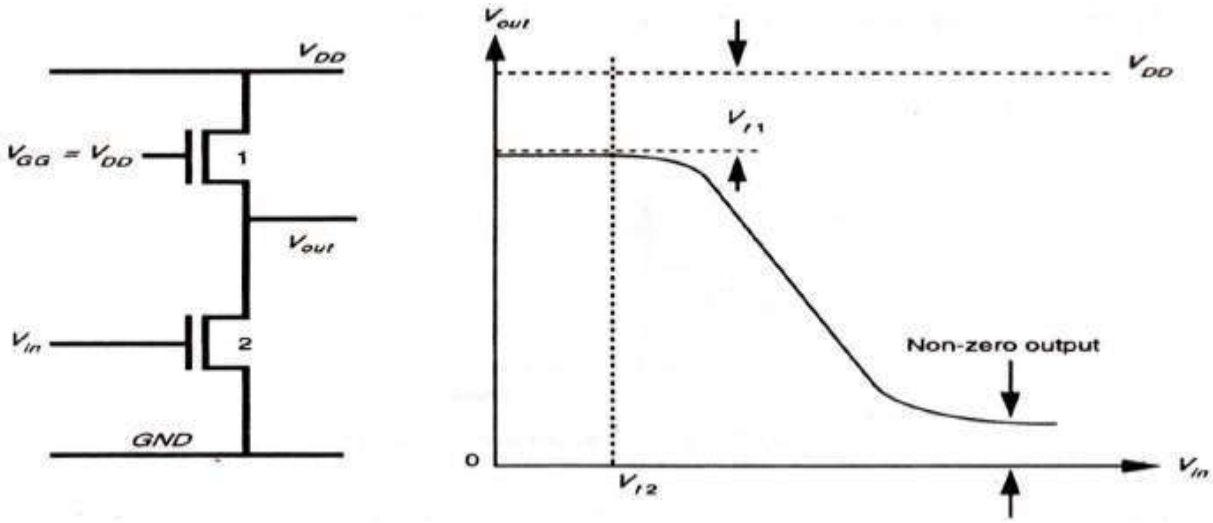
(b) Switching of output from 1 to 0 begins when Vin exceeds Vt, of pull-down device.



nMOS depletion mode transistor pull-up and transfer characteristic

(c) When switching the output from 1 to 0, the pull-up device is non-saturated initially and this presents lower resistance through which to charge capacitive loads .

(ii) **nMOS enhancement mode pull-up :**This arrangement consists of a n-MOS enhancement mode transistor as pull-up. The arrangement and the transfer characteristic are shown below.



nMOS enhancement mode pull-up and transfer characteristic

The important features of this arrangement are

(a) Dissipation is high since current flows when Vin =logical 1 (V_{GG} is returned to V_{DD}).

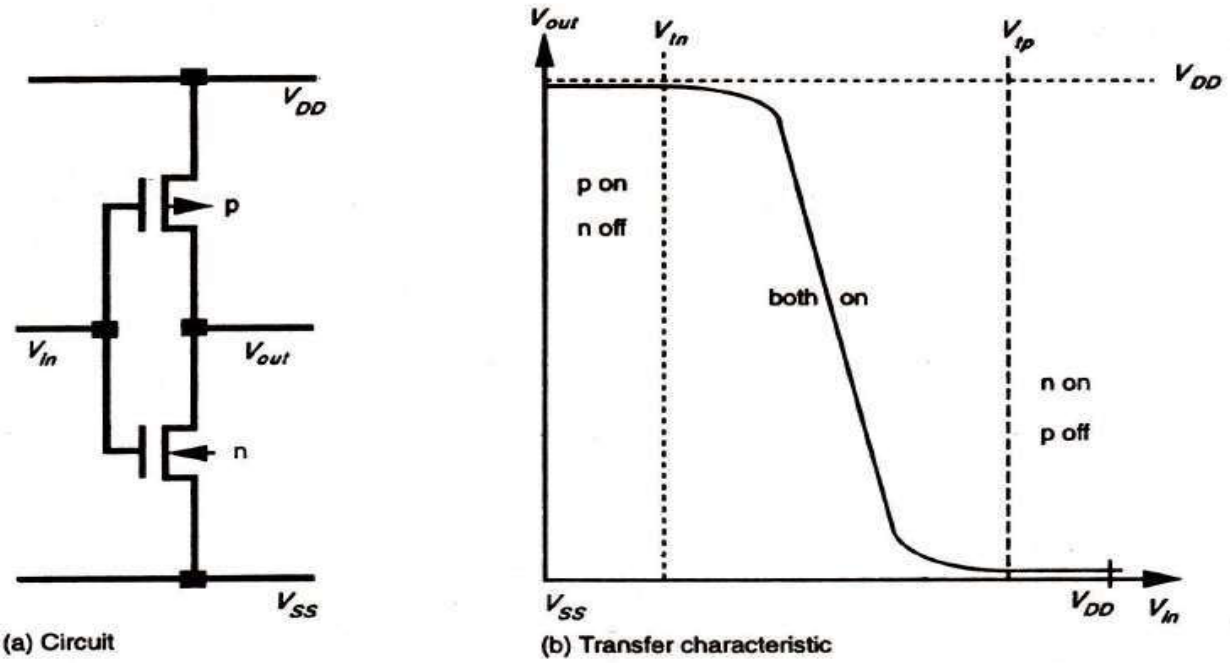
(b) Vout can never reach V_{DD} (logical I) if $V_{GG} = V_{DD}$ as is normally the case.

(c) V_{GG} may be derived from a switching source, for example, one phase of a clock, so that

dissipation can be greatly reduced.

(d) If VGG is higher than VDD then an extra supply rail is required.

(iii) **Complementary transistor pull-up (CMOS) :** This arrangement consists of a C-MOS arrangement as pull-up. The arrangement and the transfer characteristic are shown below



The salient features of this arrangement are

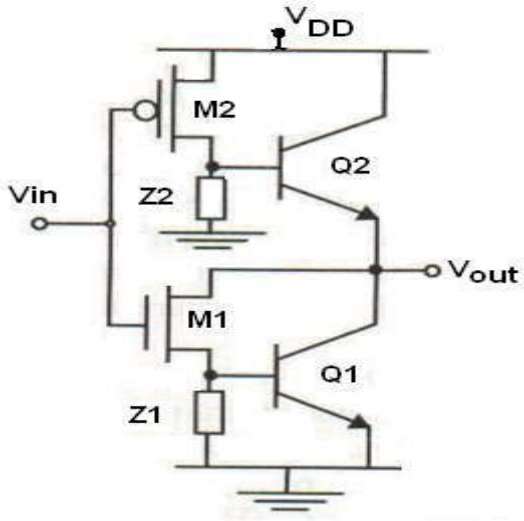
(a) No current flows either for logical 0 or for logical 1 inputs.

(b) Full logical 1 and 0 levels are presented at the output.

(c) For devices of similar dimensions the p-channel is slower than the n-channel device.

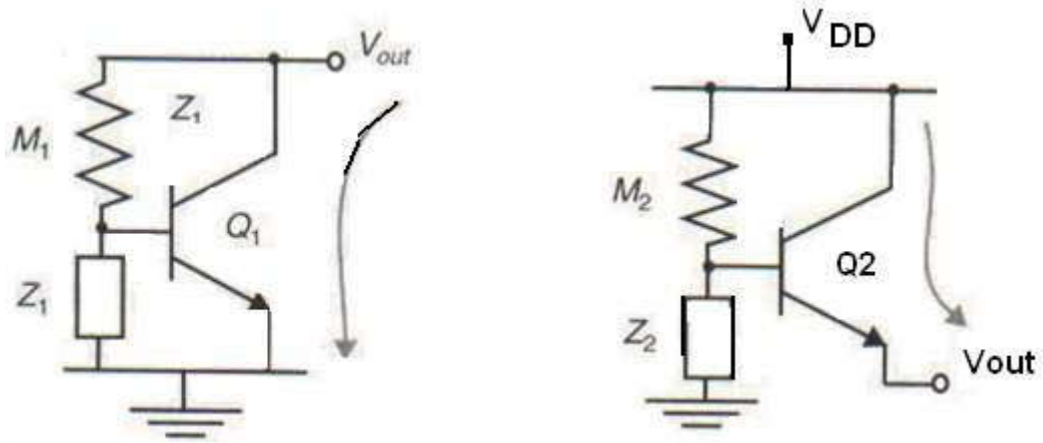
BiCMOS INVERTER:

A BiCMOS inverter, consists of a PMOS and NMOS transistor (M2 and M1), two NPN bipolar junction transistors,(Q2 and Q1), and two impedances which act as loads(*Z*2 and *Z*1) as shown in the circuit below.



When input, Vin, is high (VDD), the NMOS transistor (M1), turns on, causing Q1 to conduct,while M2 and Q2 are off, as shown in figure (b) . Hence , a low (GND) voltage is translated to the output Vout. On the other hand, when the input is low, the M2 and Q2 turns on, while M1and Q1 turns off, resulting to a high output level at the output as shown in Fig.(b).

In steady-state operation, Q1 and Q2 never turns on or off simultaneously, resulting to a lower power consumption. This leads to a push-pull bipolar output stage. Transistors M1and M2, on the other hand, works as a phase-splitter, which results to a higher input impedance.



The impedances Z2 and *Z*1 are used to bias the base-emitter junction of the bipolar transistor and to ensure that base charge is removed when the transistors turn off. For example when the input voltage makes a high-to-low transition, M1 turns off first. To turn off Q1, the base charge must be removed, which can be achieved by *Z*1.With this effect, transition time reduces. However,

there exists a short time when both Q1 and Q2 are on, making a direct path from the supply (VDD) to the ground. This results to a current spike that is large and has a detrimental effect on both the noise and power consumption, which makes the turning off of the bipolar transistor fast .

Comparison of BiCMOS and C-MOS technologies

The BiCMOS gates perform in the same manner as the CMOS inverter in terms of power consumption, because both gates display almost no static power consumption.

When comparing BiCMOS and CMOS in driving small capacitive loads, their performance are comparable, however, making BiCMOS consume more power than CMOS. On the other hand, driving larger capacitive loads makes BiCMOS in the advantage of consuming less power than CMOS, because the construction of CMOS inverter chains are needed to drive large capacitance loads, which is not needed in BiCMOS.

The BiCMOS inverter exhibits a substantial speed advantage over CMOS inverters, especially when driving large capacitive loads. This is due to the bipolar transistor's capability of effectively multiplying its current.

For very low capacitive loads, the CMOS gate is faster than its BiCMOS counterpart due to small values of *Cint*. This makes BiCMOS ineffective when it comes to the implementation of internal gates for logic structures such as ALUs, where associated load capacitances are small.

BiCMOS devices have speed degradation in the low supply voltage region and also BiCMOS is having greater manufacturing complexity than CMOS.

Assignment Questions:

1. Define threshold voltage? Drive the Vt equation for MOS transistor.
2. Explain with neat diagrams the various NMOS fabrication technology.
3. Draw and explain BiCMOS inverter circuit.
4. Discuss the Basic Electrical Properties of MOS and BiCMOS Circuits.
5. Derive the expression for estimation of Pull-Up to Pull-Down ratio of an n-MOS inverter driven by another n-MOS inverter.
6. Derive the relationship between I_{ds} and V_{ds}
7. Derive the expression for transfer characteristics of CMOS Inverter.
8. Write about BiCMOS fabrication in a n-well process with a diagram.
9. Distinguish between Bipolar and CMOS devices technologies in brief.
10. Mention about the BICMOS Inverters and alternative BICMOS Inverters.
11. Determine the pull-up to pull down ratio for NMOS inverter driven by another NMOS Inverter
12. Draw the fabrication steps of CMOS transistor and explain its operation in detail.
13. Draw the fabrication steps of NMOS transistor and explain its operation in detail.