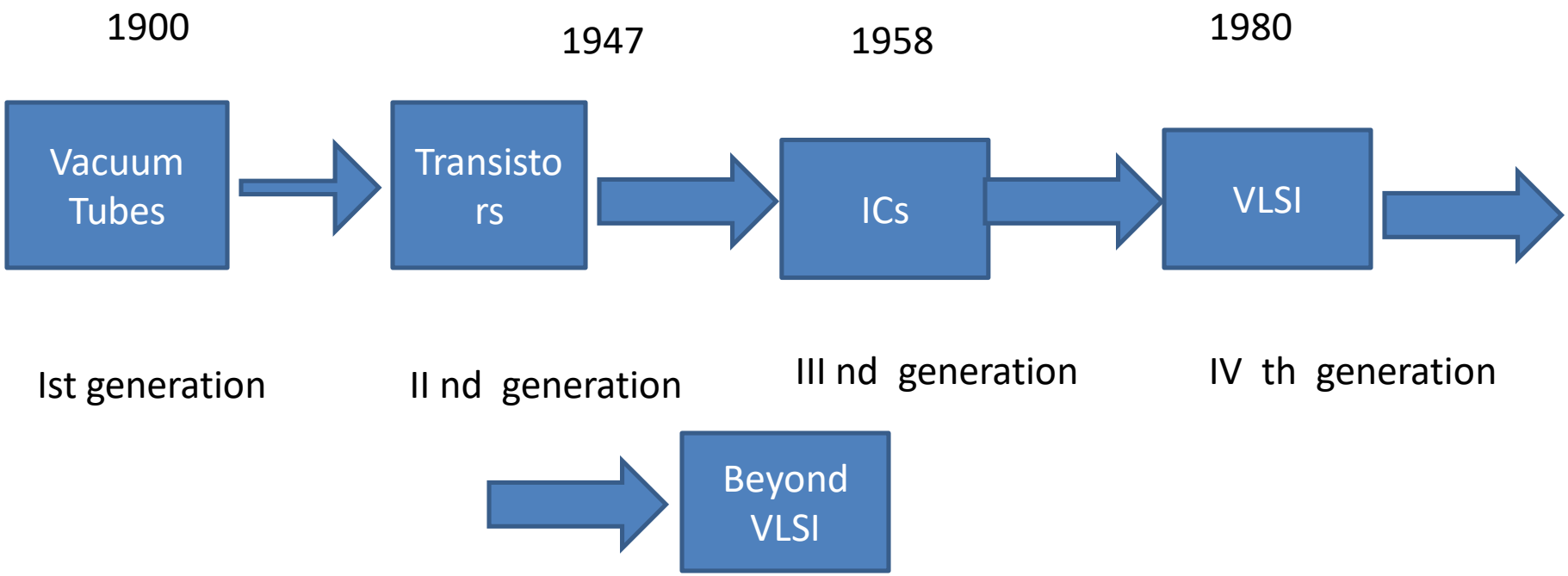
UNIT - 1

Introduction: Basic steps of IC fabrication, PMOS, NMOS, CMOS & BiCMOS, and SOI process technologies, MOS transistors - MOS transistor switches – Basic gate using switches, working polar transistor Resistors and Capacitors.

Basic Electrical Properties of MOS and BiCMOS Circuits: Working of MOS transistors – threshold voltage; MOS design equations: Ids–Vds relationships, Threshold Voltage, Body effect, Channel length modulation , gm, gds, figure of merit ω0; Pass transistor, NMOS Inverter, CMOS Inverter analysis and design, Various pull ups loads, Bi-CMOS Inverters.

Microelectronics Evolution

Q)Explain the evaluation of a VLSI design technology



Generations of Micro electronics technology

History of integrating microelectronic circuits

1947 Bipolar Transistor invented by Bardeen, Brattain and Shockley at Bell Laboratories
1958 Simultaneous Development of Integrated Circuit by Kilby at Texas Instruments & Noyce and Moore at Fairchild Semiconductor
1961 First commercial digital IC available from Fairchild Semiconductor
1967 First Semiconductor RAM (64bits) discussed at the IEEE International Solid-State Circuits Conference (ISSCC)
1968 Introduction of the first commercial IC operational amplifier the µA709 by Fairchild Semiconductor
1970 1-transistor dynamic memory cell invented by Dennard at IBM
1971 Introduction of the 4004 microprocessor by intel
1972 First 8-bit Microprocessor The Intel 8008
1974 First 1kBit memory chip, 8080 microprocessor
1978 First 16-bit Microprocessor
1984 1MBit Memory chip

Q What is VLSI

- **Very-large-scale integration (VLSI):** is the process of creating an integrated circuit (IC) by combining thousands of transistors into a single chip. VLSI began in 1970s when complex semiconductor and communication technologies were being developed. The microprocessor is a VLSI device. Before the introduction of VLSI technology most ICs had a limited set of functions they could perform. An electronic circuit might consist of a CPU, ROM, RAM and other glue logic. VLSI lets IC designers add all of these into one chip.
- **What are the four generations of Integration Circuits?**
  - SSI (Small Scale Integration)
  - MSI (Medium Scale Integration)
  - LSI (Large Scale Integration)
  - VLSI (Very Large Scale Integration)
- **What are the advantages of IC?**
  - Size is less
  - High Speed
  - Less Power Dissipation

Classification of ICS

Q)Classify ICs based on Number of Transistors per chip

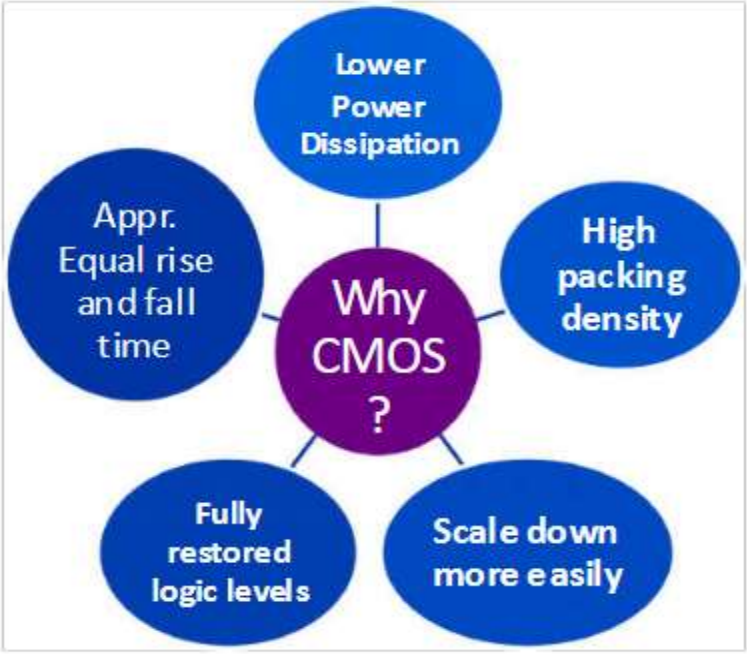
Level of integration	Design rule	No of Transistors	Year	Example
SSI	30-20	2 to 64	1960-65	Flipflops &gates
MSI	20-10	64 to2000	1965-70	Mux,adders
LSI	10-3 1/2	2000-64000	1970-78	ROM,ROM
VLSI	3 ½ -1 1/4	64000-2 million	1978-86	16.32 bit MP
ULSI	<1 1/4	>2 million	after1986	Special purpose processors
GSI		10 million		Embedded Systems,SOC

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.

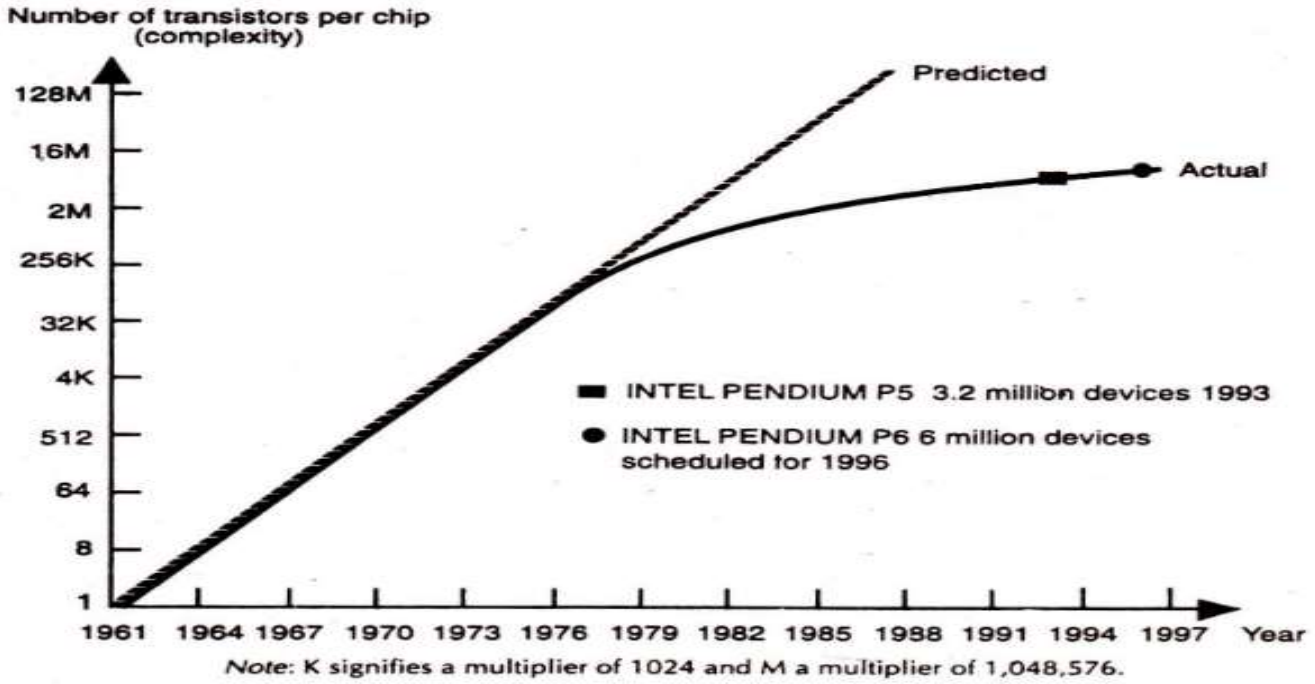
Q1)List the different IC Technologies
Q2)Compare BJT and CMOS
Q3)What are the advantages of CMOS?

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



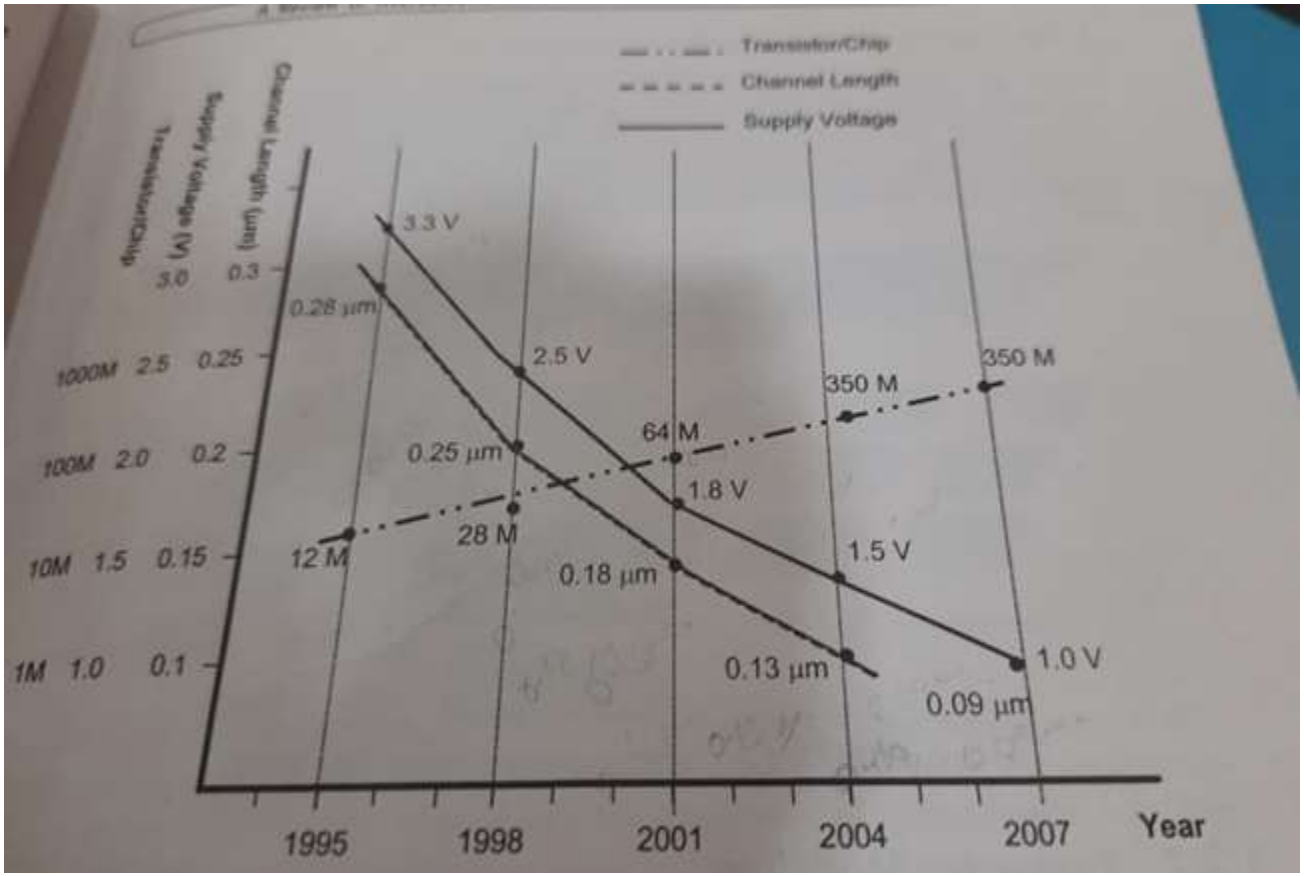
Moores Law?

Q)Explain clearly about Moore's law



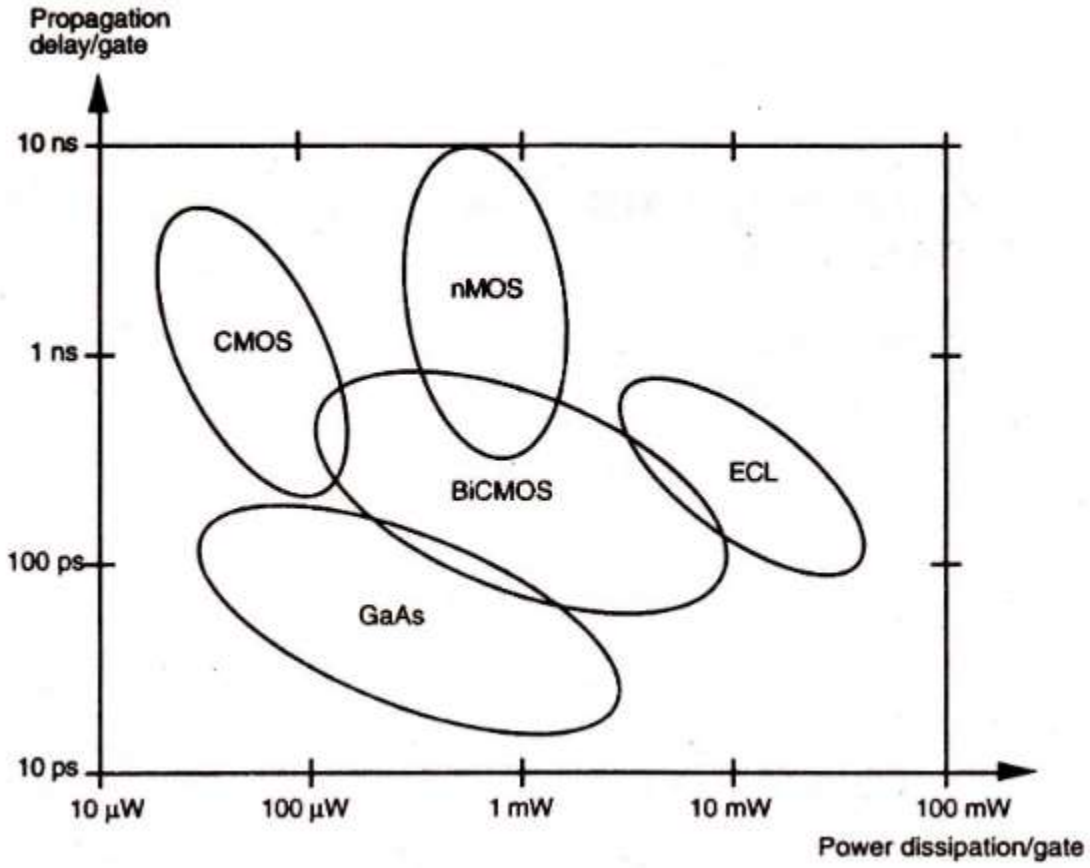
Moore's first law: Transistors Integrated on a single chip (commercial products).

Moore's law, prediction made by American engineer Gordon Moore in 1965 that the number of transistors per silicon chip doubles every year. ... Moore observed that the number of transistors on a computer chip was doubling about every 18–24 months.



System integration complexity roadmap

Q)Define speed power product



Speed/power performance of available technologies.

Why NMOS is so popular?

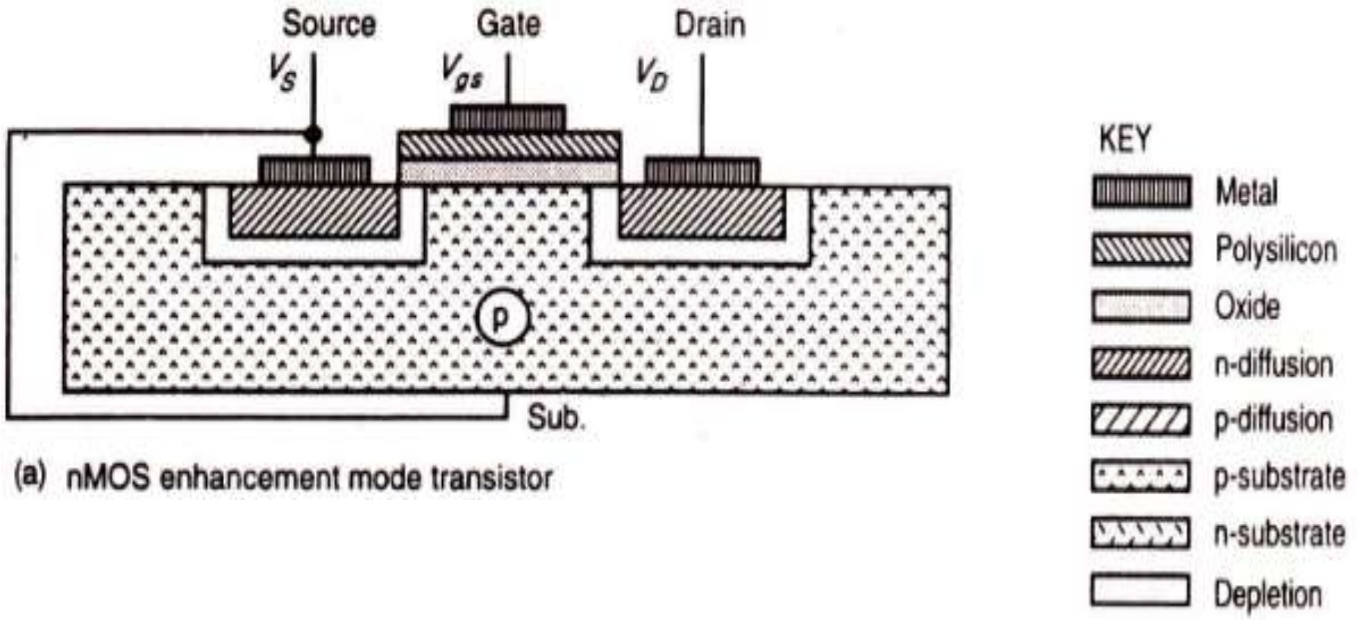
- For nMOS technology, the design methodology and the design rules are easily learned, thus providing a simple but excellent introduction to structured design for VLSI.
- nMOS technology and design processes provide an excellent background for other technologies.. In particular, some familiarity with nMOS allows a relatively easy transition to CMOS technology and design
- For GaAs technology some arrangements in relation to logic design are similar to those employed in nMOS technology. Therefore, understanding the basics of nMOS design will assist in the layout of GaAs circuits.

IC fabrication

Q)List the basic process steps for the fabrication integrated circuits

silicon wafer preparation ,
Epitaxial growth,
oxidation,
Photolithography,
Diffusion.
Ion Implantation,
Isolation technique,
Metallization,
Assembly processing &packaging

BASIC MOS TRANSISTORS

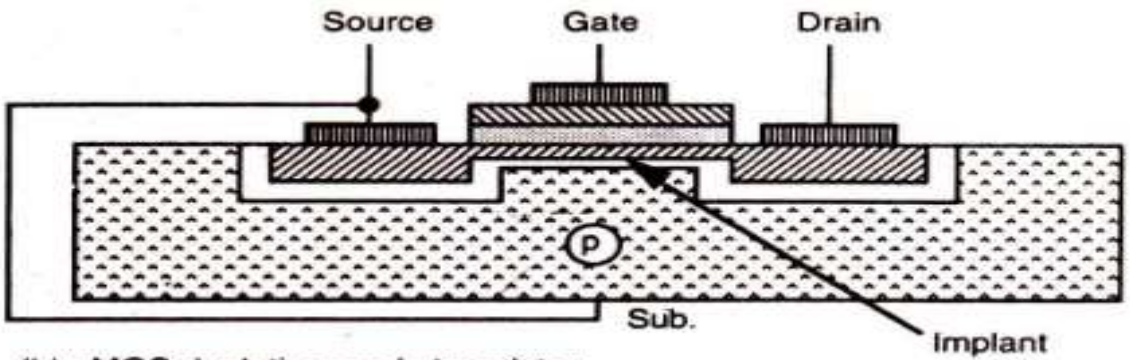


(a) nMOS enhancement mode transistor

nMOS devices are formed in a p-type substrate of moderate doping level. The source and drain regions are formed by diffusing n-type impurities through suitable masks into these areas to give the desired n-impurity concentration and give rise to depletion regions which extend mainly in the more lightly doped p-region as shown. Thus, source and drain are isolated from one another by two diodes.

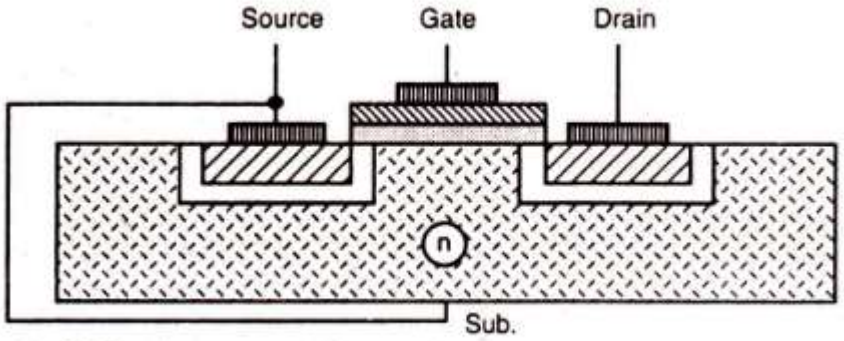
Connections to the source and drain are made by a deposited metal layer.

V_D = Vs = Vgs = 0 . If this gate is connected to a suitable positive voltage with respect to the source, then the electric field established between the gate and the substrate gives rise to a charge inversion region in the substrate under the gate insulation and a conducting path or channel is formed between source and drain.



(b) nMOS depletion mode transistor

The channel may also be established so that it is present under the condition Vgs = 0 by implanting suitable impurities in the region between source and drain during manufacture and prior to depositing the insulation and the gate.

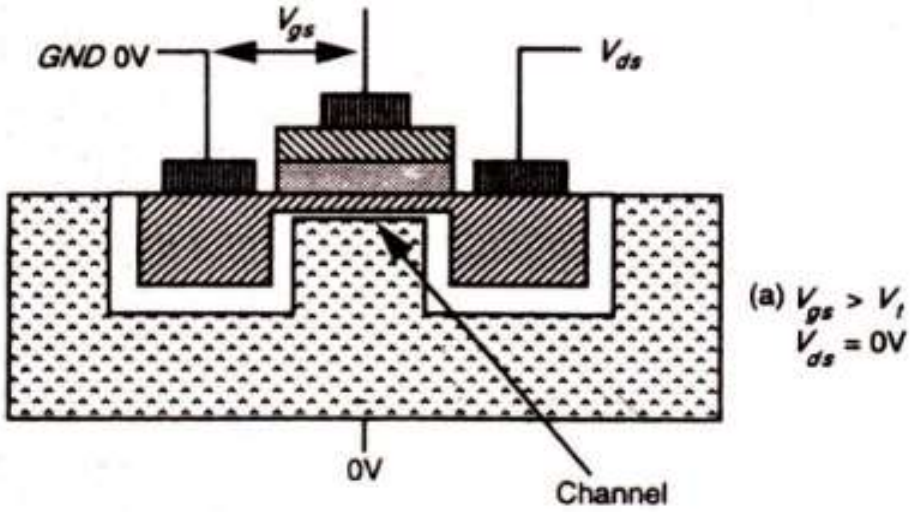


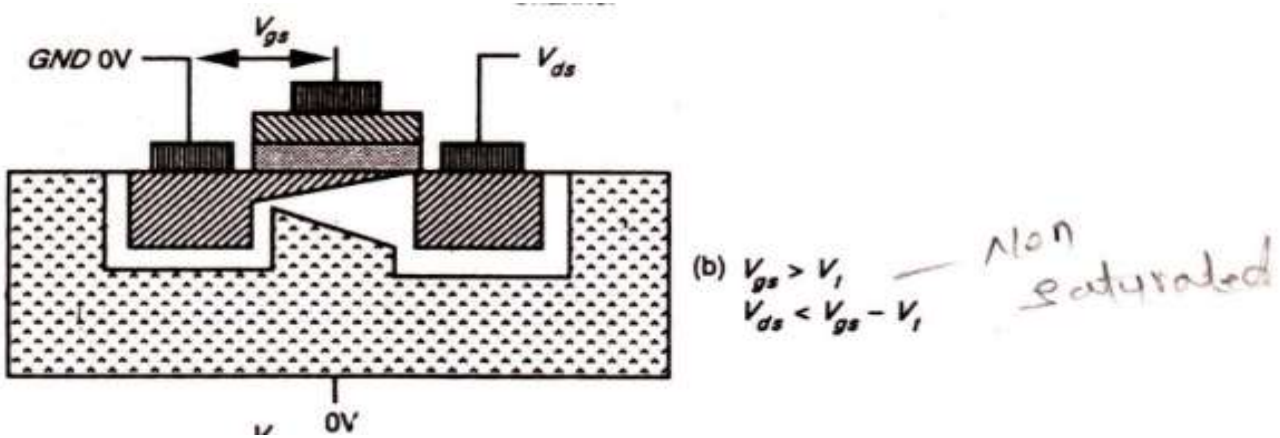
(c) pMOS enhancement mode transistor

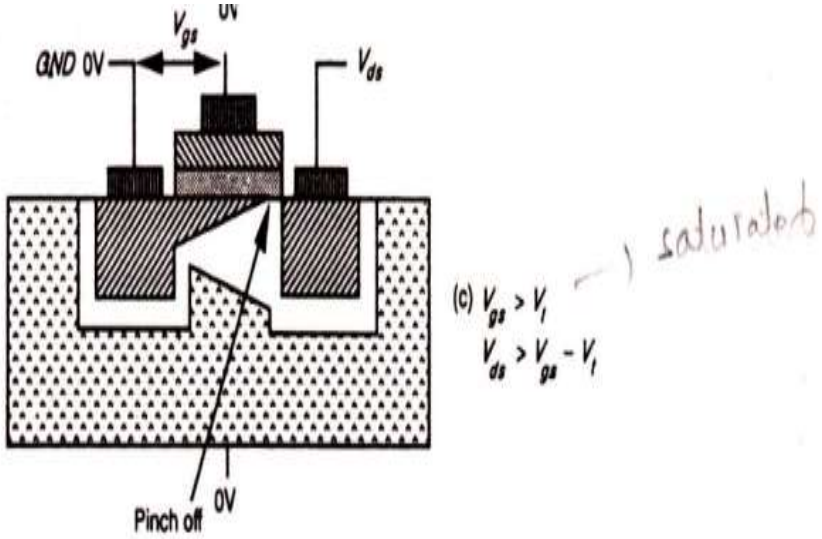
the application of a negative voltage of suitable magnitude (> |Vt|) between gate and source will give rise to the formation of a channel (p-type) between the source and drain and current may then flow if the drain is made negative with respect to the source.

DEPLETION MODE TSANSISTOR ACTION

Q)Describe the different operating regions for an MOS transistor



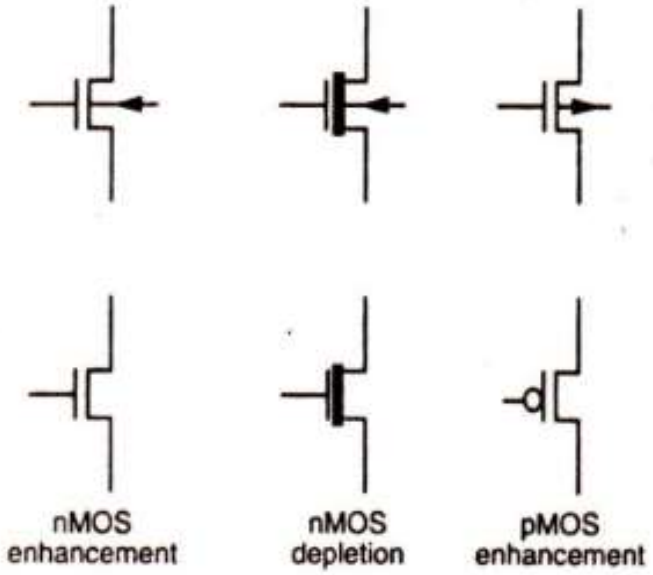




Note: Vds is the drain-to-source voltage. Substrate assumed connected to 0 V.

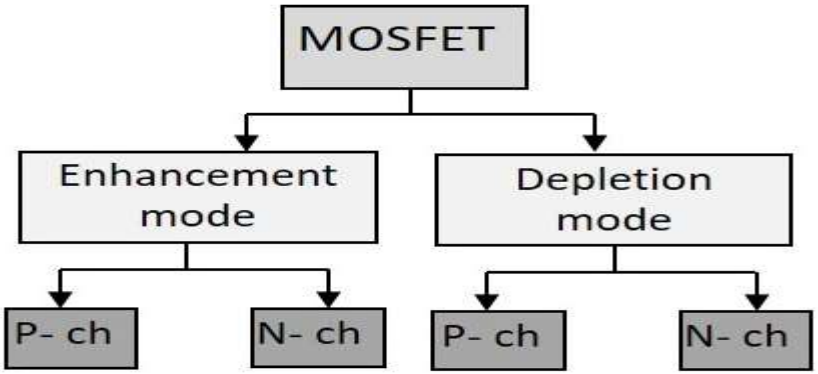
Figure:Enhancement mode transistor for particular values of Vds with (Vgs > Vt).

Q)An n-MOSFET has threshold voltage of 1v, gate voltage of 2v and drain voltage of 2.5v.Find the region of operation of the MOSFET.



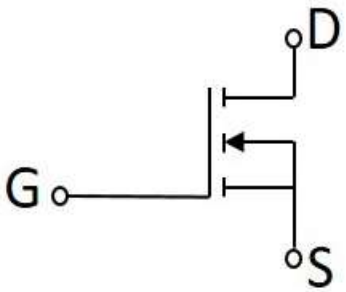
Transistor circuit symbols.

Classification of MOSFETs

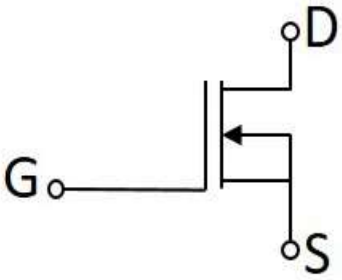


- P- ch = P- channel
- N- ch = N- channel

Symbols of N-Channel MOSFET

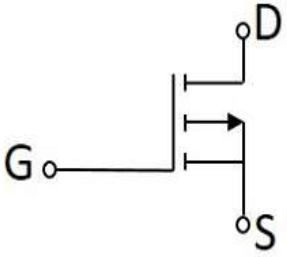


Enhancement Mode

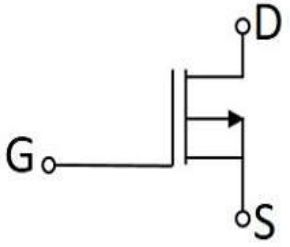


Depletion Mode

Symbols of P-Channel MOSFET



Enhancement Mode



Depletion Mode

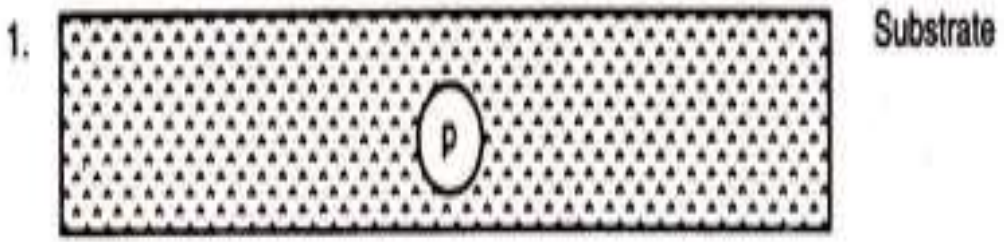
Comparison between BJT, FET and MOSFET

TERMS	BJT	FET	MOSFET
Device type	Current controlled	Voltage controlled	Voltage Controlled
Current flow	Bipolar	Unipolar	Unipolar
Terminals	Not interchangeable	Interchangeable	Interchangeable
Operational modes	No modes	Depletion mode only	Both Enhancement and Depletion modes
Input impedance	Low	High	Very high
Output resistance	Moderate	Moderate	Low
Operational speed	Low	Moderate	High
Noise	High	Low	Low
Thermal stability	Low	Better	High

nMOS FABRICATION

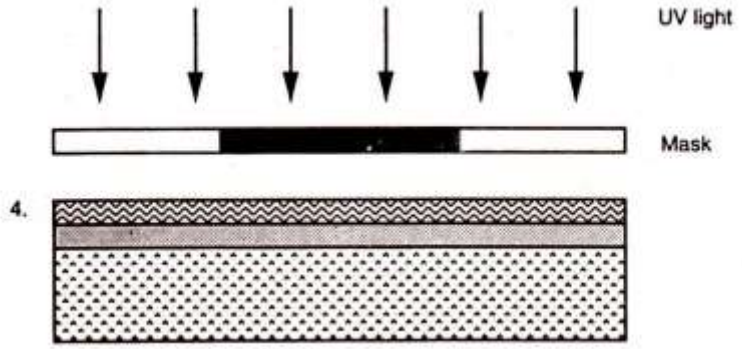
Q)Explain NMOS fabrication process flow with neat diagrams

Processing is carried out on a thin wafer cut from a single crystal of silicon of high purity into which the required p-impurities are introduced as the crystal is grown. Such wafers are typically 75 to 150 mm in diameter and 0.4 mm thick and are doped with, say, boron to impurity concentrations of 10^{15}/cm3 to 10^{16}/cm3 , giving resistivity in the approximate tange 25 ohm cm to 2 ohm cm.

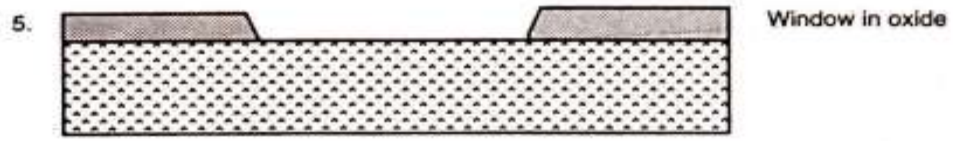


A layer of silicon dioxide (Si02), typically 1µm thick, is grown all over the surface of the wafer to protect the surface, act as a barrier to dopants during processing, and provide a generally insulating substrate on to which other layers may be deposited and patterned.

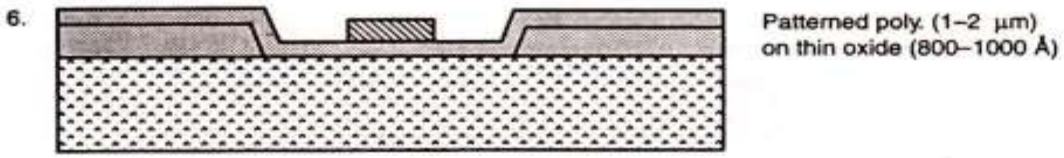
3. The surface is now covered with a photoresist which is deposited onto the wafer and spun to achieve an even distribution of the required thickness.



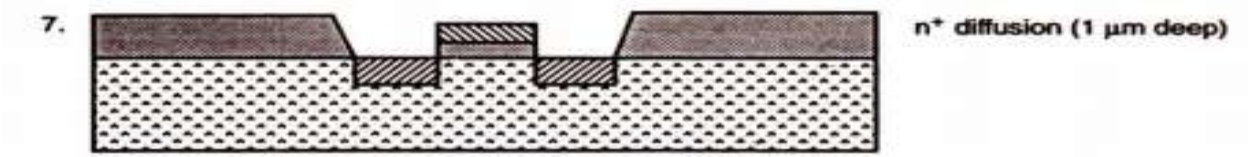
4. The photoresist layer is then exposed to ultraviolet light through a mask which defines those regions into which diffusion is to take place together with transistor channels. Assume, for example, that those areas exposed to ultraviolet radiation are polymerized (hardened), but that the areas required for diffusion are shielded by the mask and remain unaffected



5. These areas are subsequently readily etched away together with the underlying silicon dioxide so that the wafer surface is exposed in the window defined by the mask.

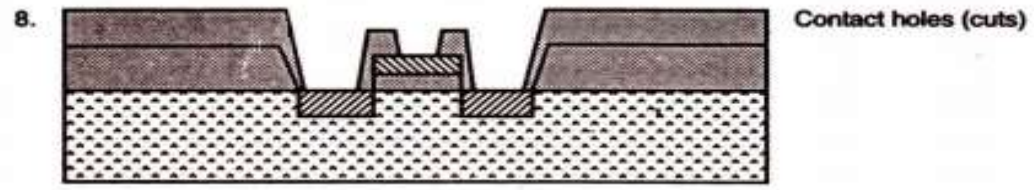


The remaining photoresist is removed and a thin layer of Si02 (0.1 μm typical) is grown over the entire chip surface and then polysilicon is deposited on top of this to form the gate structure. The polysilicon layer consists of heavily doped polysilicon deposited by chemical vapor deposition (CVD). In the fabrication of fine pattern devices, precise control of thickness, impurity concentration, and resistivity is necessary.

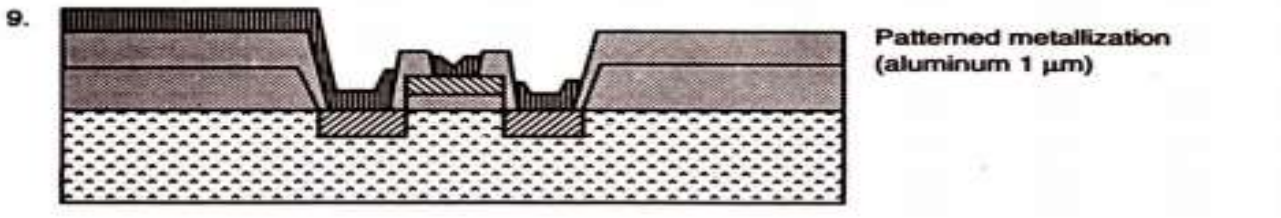


7. Further photoresist coating and masking allows the polysilicon to be patterned (as shown in Step 6) and then the thin oxide is removed to expose areas into which

n-type impurities are to be diffused to form the source and drain as shown



Thick oxide (Si02) is grown over all again and is then masked with photoresist and etched to expose selected are~s of the polysilicon gate and the drain and source areas where connections (i.e. contact cuts) are to be made.



9. The whole chip then has metal (aluminum) deposited over its surface to a thickness typically of I μm. This metal layer is then masked and etched to form the required interconnection pattern

n-type impurities are to be diffused to form the source and drain as shown. is achieved by heating the wafer to a high temperature and passing a gas containing the desired n-type impurity (for example, phosphorus) over the surface as indicated in Figure. Note that the polysilicon with underlying thin oxide act as masks during diffusion--the process is self-aligning

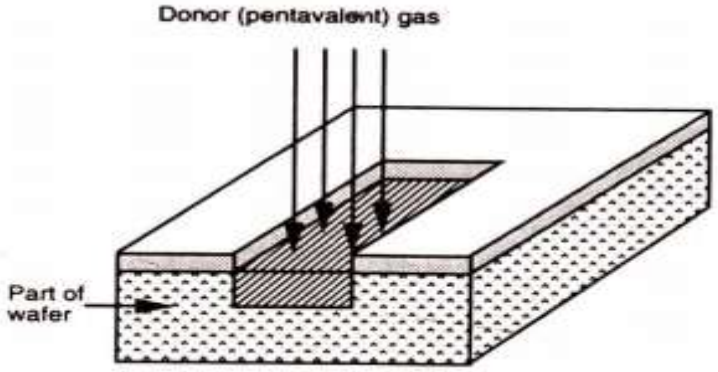


FIGURE :Diffusion process

A **mask** is a specification of geometric shapes that need to be created on a certain layer. are used to create a specific patterns of each material in a sequential manner and create a complex pattern of several layers.

Summary of An nMOS Process

• Processing takes place on a p-doped silicon crystal wafer on which is grown a 'thick' layer of Si02.

• Mask 1-Pattern Si02 to expose the silicon surface in areas where paths in the diffusion layer or gate areas of transistors are required. Deposit thin oxide over alL For this reason, this mask is often known as the 'thinox' mask but some texts refer to it as the diffusion mask.

• Mask 2-Pattern the ion implantation within the thinox region where depletion mode devices are to be produced-self-aligning.

• Mask 3-Deposit polysilicon over all (I.5 µm thick typically), then pattern using Mask 3. Using the same mask, remove thin oxide layer where it is not covered by polysilicon.

• Diffuse n + regions into areas where thin oxide has been removed. Transistor drains and sources are thus self-aligning with respect to the gate structure&.

• Mask 4--Grow thick oxide over all and then etch for contact cuts.

• Mask 5-Deposit metal and pattern with Mask 5!

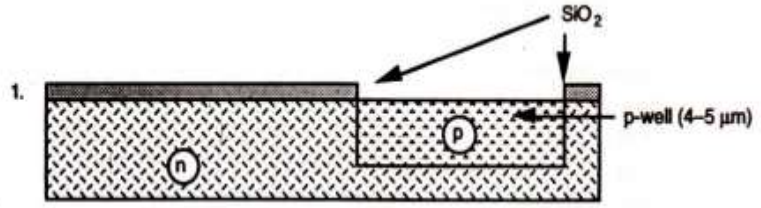
• Mask 6-Would be required for the overglassing process step.

CMOS FABRICATION

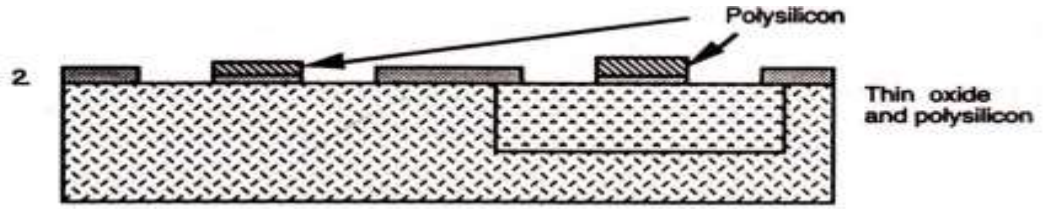
Q)Explain clearly about p-well CMOS fabrication process with neat diagrams.

There are a number of approaches to CMOS fabrication, including the p-well, the n-well, the twin-tub, and the silicon-on-insulator processes. In order to introduce the reader to CMOS design we will be concerned mainly with well-based circuits. The p-well process is widely used in practice and the n-well process is also popular,

The p-well Process



the structure consists of an n-type substrate in which p-devices may be formed by suitable masking and diffusion and, in order to accommodate n-type devices, a deep p-well is diffused into the n-type substrate as shown



This diffusion must be carried out with special care since the p-well doping concentration and depth will affect the threshold voltages as well as the breakdown voltages of the n-transistors.

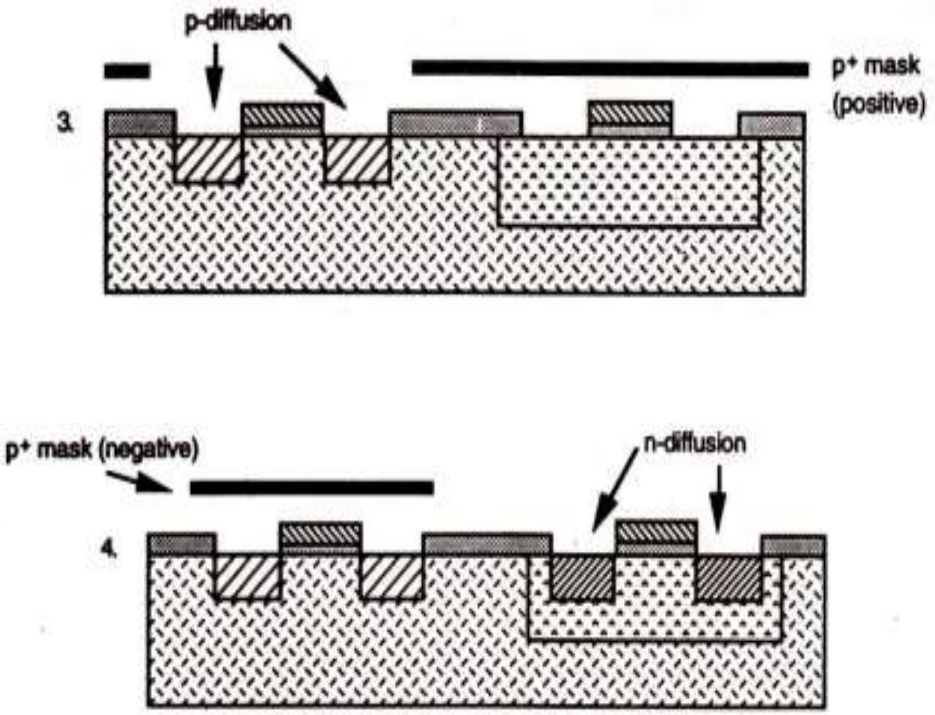


Figure CMOS p-well process steps.

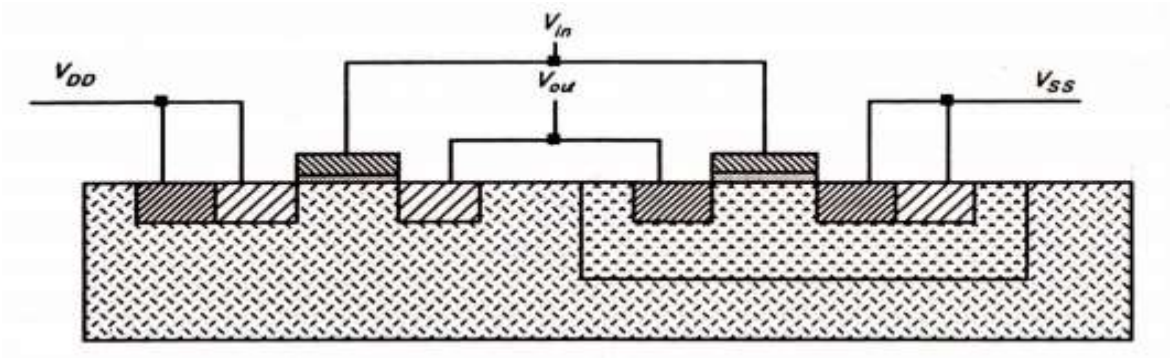
In all other respects-masking, patterning, and diffusion-the process is similar to nMOS fabrication. In summary, typical processing steps are:

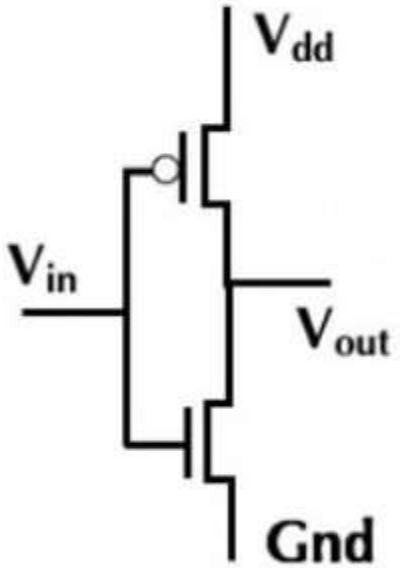
Mask 1 - defines the areas in which the deep p-well diffusions are to take place.

- Mask 2 - defines the thinox regions, namely those areas where the thick oxide is to be stripped and thin oxide grown to accommodate p- and n-transistors and wires.
- Mask 3 - used to pattern the polysilicon layer which is deposited after the thin oxide.
- Mask 4 - A p-plus mask is now used (to be in effect "Anded" with Mask 2) to define all areas where p-diffusion is to take place.
- Mask 5 - This is usually performed using the negative form of the p-plus mask and defines those areas where n-type diffusion is to take place.
- Mask 6 - Contact cuts are now defined.
- Mask 7 - The metal layer pattern is defined by this mask.

Mask 8 - An overall passivation (overglass) layer is now applied and Mask 8 is needed to define the openings for access to bonding pads.

Cross-sectional view of p-well inverter showing V_{DD} and Vss substrate connections

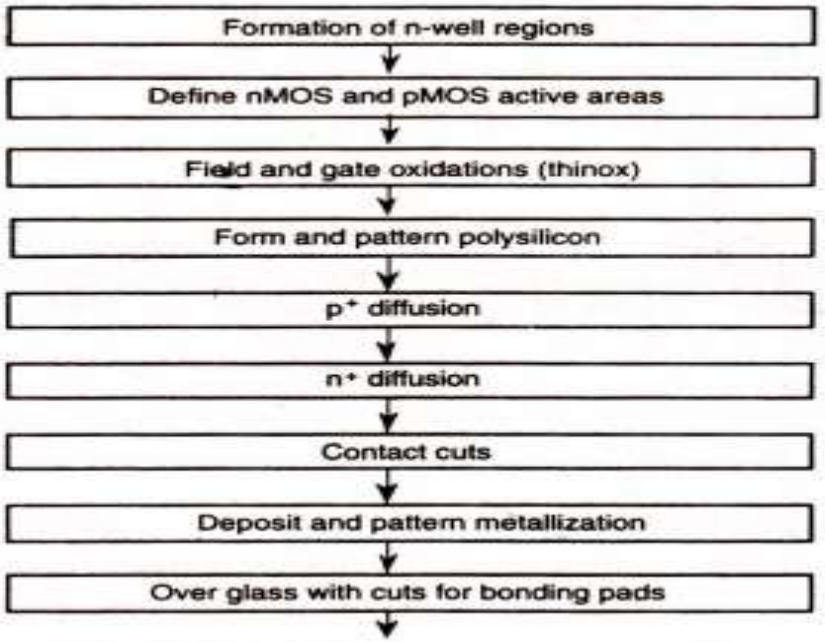




The n-well Process

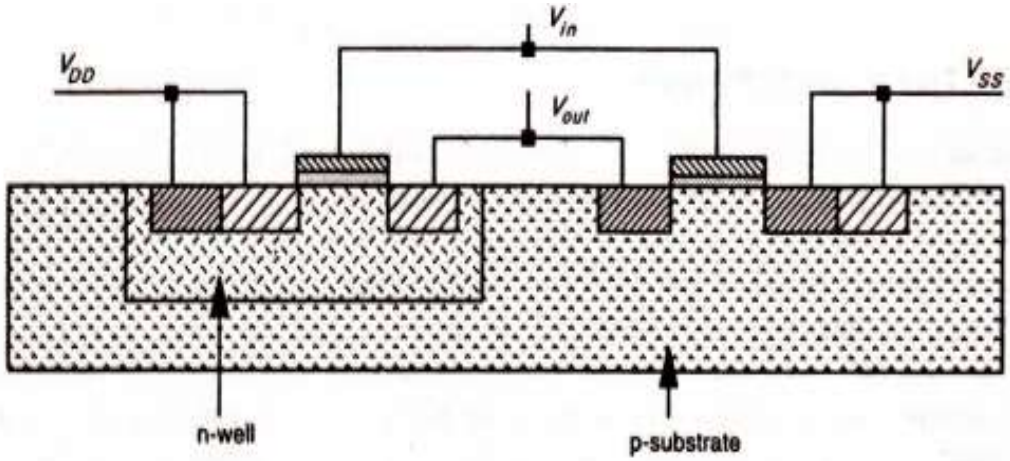
N-well CMOS circuits are also superior to p-well because of the lower substrate bias effects on transistor threshold voltage and inherently lower parasitic capacitances associated with source and drain regions.

lt will be seen that an n+ mask and its complement may be used to define the n- and p diffusions regions respectively.



Main steps in a typical n-well process~

FIGURE Cross-sectional view of n-wen CMOS Inverter.



The Twin -Tub Process

Here we start with a substrate of high resistivity n-type material and then create both .. n-well and p-well regions. Through this process it is possib!e to preserve the performance of n-transistors without compromising the p-transistors. Doping control is more readily achieved and some relaxation in manufacturing tolerances results

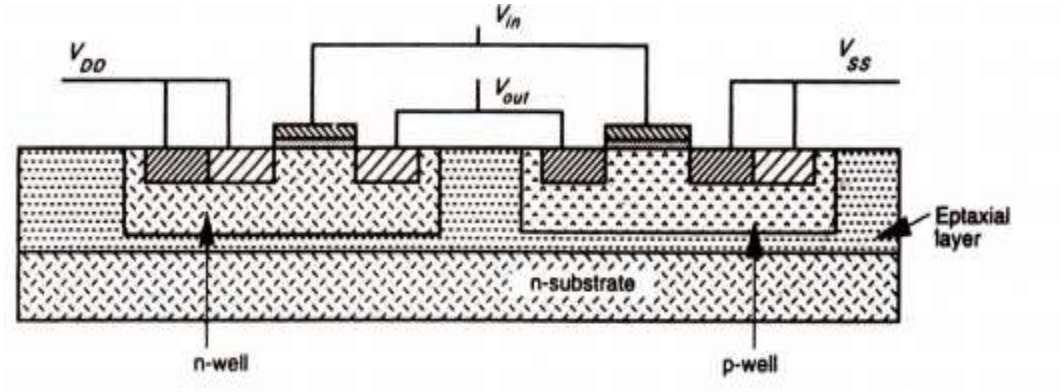


FIGURE Twin-tub structure.

THERMAL ASPECTS OF PROCESSING

The processes involved in making nMOS and CMOS devices have differing high temperature sequences as indicated in Figure .

The CMOS p-well process, for example, has a high temperature p-well diffusion process (1100 to 1250°C), the nMOS process having no such requirement.

·Because of the simplicity, ease of fabrication, and high density per unit area of nMOS circuits, many of the earlier IC designs, still in current use, have been fabricated using nMOS technology and it is likely that nMOS and CMOS system designs will continue to co-exist for some time to come.

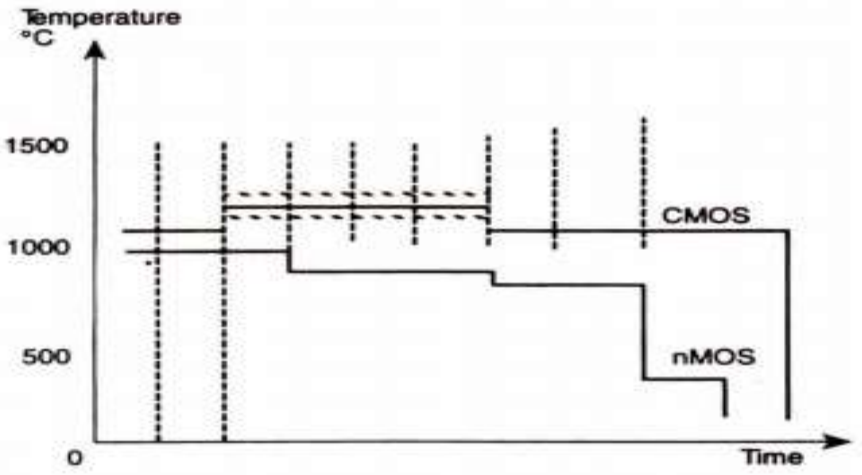


FIGURE Thermal sequence difference between nMOS and CMOS processes.

BICMOS TECHNOWGY

A known deficiency of MOS technology lies in the limited load driving capabilities of MOS transistors. This is due to the limited current sourcing and current sinking abilities associated with both p- and n-transistors and although it is possible, for example, to design so called super-buffers using MOS transistors alone, such arrangements do riot always compare well with the capabilities of bipolar transistors.

Bipolar transistors also provide higher gain and have generally better noise and high frequency characteristics than MOS transistors and it may be seen that BiCMOS gates could be an effective way of speeding up VLSI circuits. However, the application of BiCMOS in sub-systems such as ALU, ROM, a register-file, or, for that matter, a barrel shifter, is not always an effective way of improving speed.

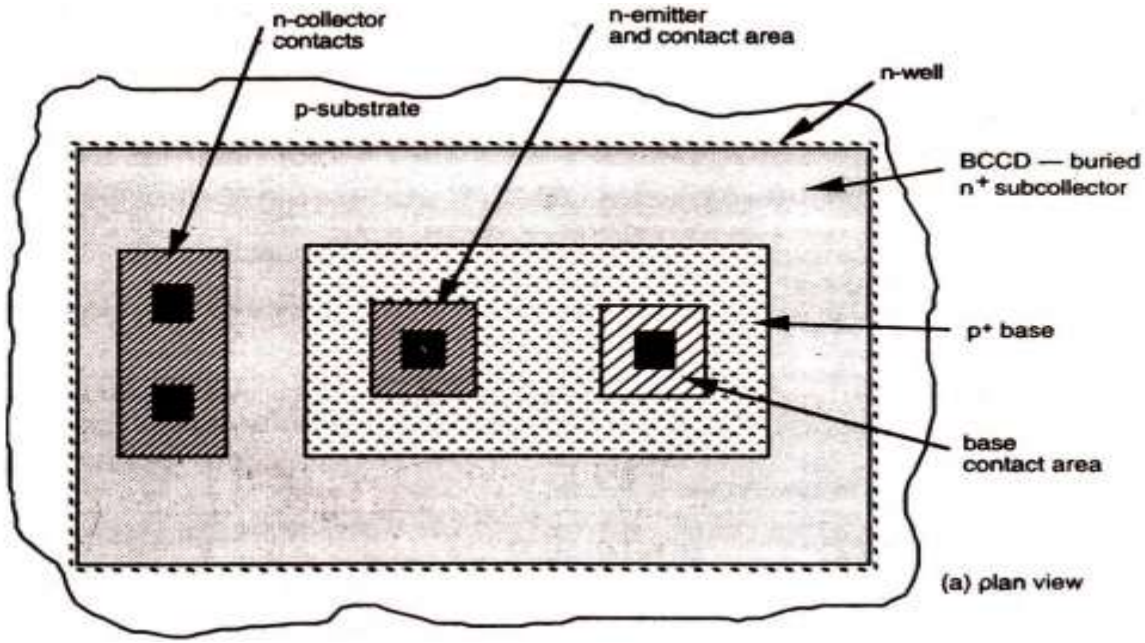
This is because most gates in such structures do not have to drive large capacitive loads so that the BiCMOS arrangements give no speed advantage.

TABLE Comparison between CMOS and bipolar technologies

CMOS technology	*Bipolar technology*
• Low static power dissipation	• High power dissipation
• High input impedance (low drive current)	• Low input impedance (high drive current)
• Scalable threshold voltage	
• High noise margin	• Low voltage swing logic
• High packing density	• Low packing density
• High delay sensitivity to load (fan-out limitations)	• Low delay sensitivity to load
• Low output drive current	• High output drive current
• Low g_m ($g_m \alpha V_{in}$)	• High g_m ($g_m \alpha e^{V_{in}}$)
	• High f_t at low currents
• Bidirectional capability (drain and source are interchangeable)	• Essentially unidirectional
• A near ideal switching device	

BICMOS Fabrication In an n-well Process

The basic process steps used are those already outlined for CMOS but with additional process steps and additional masks defining: (i) the p+ base region; (ii) n+ collector area; and (iii) the buried subcollector (BCCD).



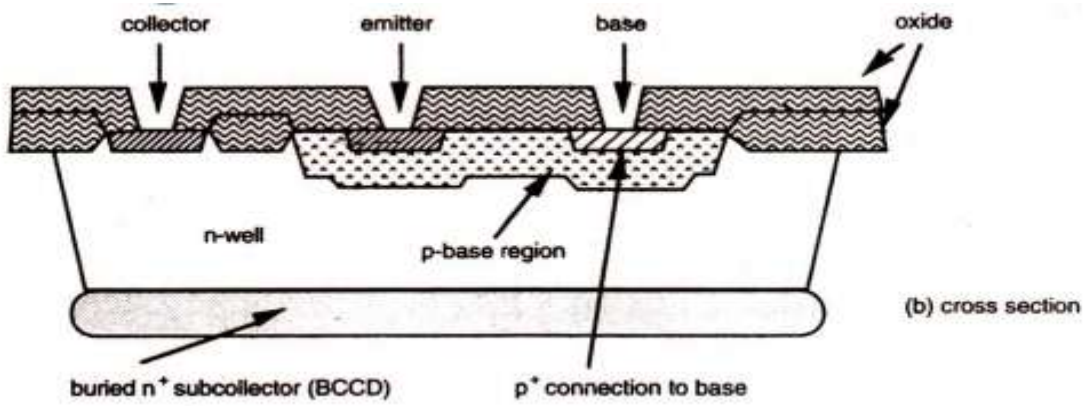


FIGURE Arrangement of BiCMOS npn transistor (Orbit 2 µm CMOS).

Table :n-well BiCMOS fabrication process steps

Single poly. single metal CMOS	*Additional steps for bipolar devices*
• Form n-well	
	• Form buried n^+ layer (BCCD)
• Delineate active areas	
• Channel stop	
	• Form deep n^+ collector
• Threshold V_t adjustment	
• Delineate poly./gate areas	
• Form n^+ active areas	
• Form p^+ active areas	
	• Form p^+ base for bipolars
• Define contacts	
• Delineate the metal areas	

Q)Draw V-I characteristics of NMOS transistor. Explain its operation. Derive the drain to source current equation in saturation and resistive region (or) Explain the electrical properties of MOS transistor in detail

DRAIN-TO-SOURCE CURRENT I_{ds} versus VOLTAGE Vds RELATIONSHIPS

The whole concept of the MOS transistor evolves from the use of a voltage on the gate to 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 V ds 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· Consider a structure, as in Figure in which" electrons will flow source to drain:

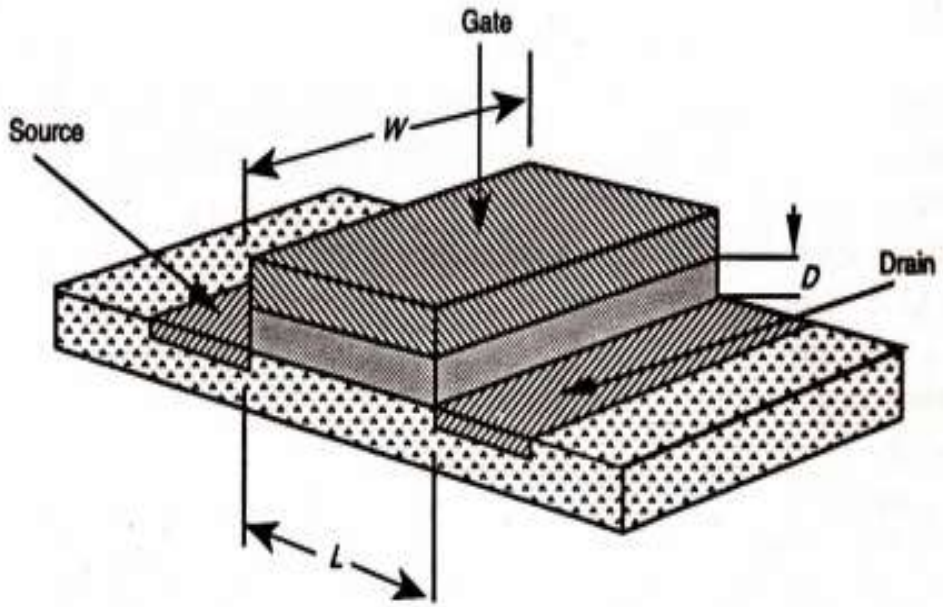


FIGURE nMOS transistor structure.

$$I_{ds} = -I_{sd} = \frac{\text{Charge induced in channel } (Q_c)}{\text{Electron transit time } (\tau)} \tag{2.1}$$

First, transit time:

$$\tau_{sd} = \frac{\text{Length of channel } (L)}{\text{Velocity } (v)}$$

but velocity

$$v = \mu E_{ds}$$

where

μ = electron or hole mobility (surface)

E_{ds} = electric field (drain to source)

Now

$$E_{ds} = \frac{V_{ds}}{L}$$

so that

$$v = \frac{\mu V_{ds}}{L}$$

Thus

$$\tau_{sd} = \frac{L^2}{\mu V_{ds}} \tag{2.2}$$

Typical values of μ at room temperature are:

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

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

The Non-saturated Region

Charge induced in channel due to gate voltage is due to the voltage difference between the gate and the channel, Vgs (assuming substrate connected to source). Now note that the voltage along the channel varies linearly with distance X from the source due to the IR drop in the channel and assuming that the device is not saturated then the average value is V ds/2. Furthermore, 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

Note that the charge/unit area = $E_g \varepsilon_{ins} \varepsilon_0$. Thus induced charge

$$Q_c = E_g \varepsilon_{ins} \varepsilon_0 WL$$

where

E_g = average electric field gate to channel
ε_{ins} = relative permittivity of insulation between gate and channel
ε_0 = permittivity of free space

(*Note:* $\varepsilon_0 = 8.85 \times 10^{-14}$F cm^{-1}; $\varepsilon_{ins} \div 4.0$ for silicon dioxide)

Now

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

where D = oxide thickness.

Thus

$$Q_c = \frac{WL\varepsilon_{ins}\varepsilon_0}{D}\left((V_{gs} - V_t) - \frac{V_{ds}}{2}\right) \tag{2.3}$$

Now, combining equations (2.2) and (2.3) in equation (2.1), we have

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

or

$$I_{ds} = K\frac{W}{L}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right) \tag{2.4}$$

in the non-saturated or resistive region where $V_{ds} < V_{gs} - V_t$ and

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

The factor W/L is, of course, contributed by the geometry and it is common practice to write

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

so that

$$I_{ds} = \beta\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right) \tag{2.4a}$$

which is an alternative form of equation (2.4).

Noting that gate/channel capacitance

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

we also have

$$K = \frac{C_g \mu}{WL}$$

so that

$$I_{ds} = \frac{C_g \mu}{L^2}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right) \tag{2.4b}$$

which is a further alternative form of equation (2.4).

Sometimes it is convenient to use *gate capacitance per unit area* C_0 (which is often denoted C_{ox}) rather than C_g in this and other expressions. Noting that

$$C_g = C_0 WL$$

we may also write

$$I_{ds} = C_0 \mu \frac{W}{L}\left((V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right) \tag{2.4c}$$

The Saturated Region

Saturation begins when $V_{ds} = V_{gs} - V_t$, 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 V_{ds} increases further. Thus

$$I_{ds} = K\frac{W}{L}\frac{(V_{gs} - V_t)^2}{2} \qquad (2.5)$$

or, we may write

$$I_{ds} = \frac{\beta}{2}(V_{gs} - V_t)^2 \qquad (2.5a)$$

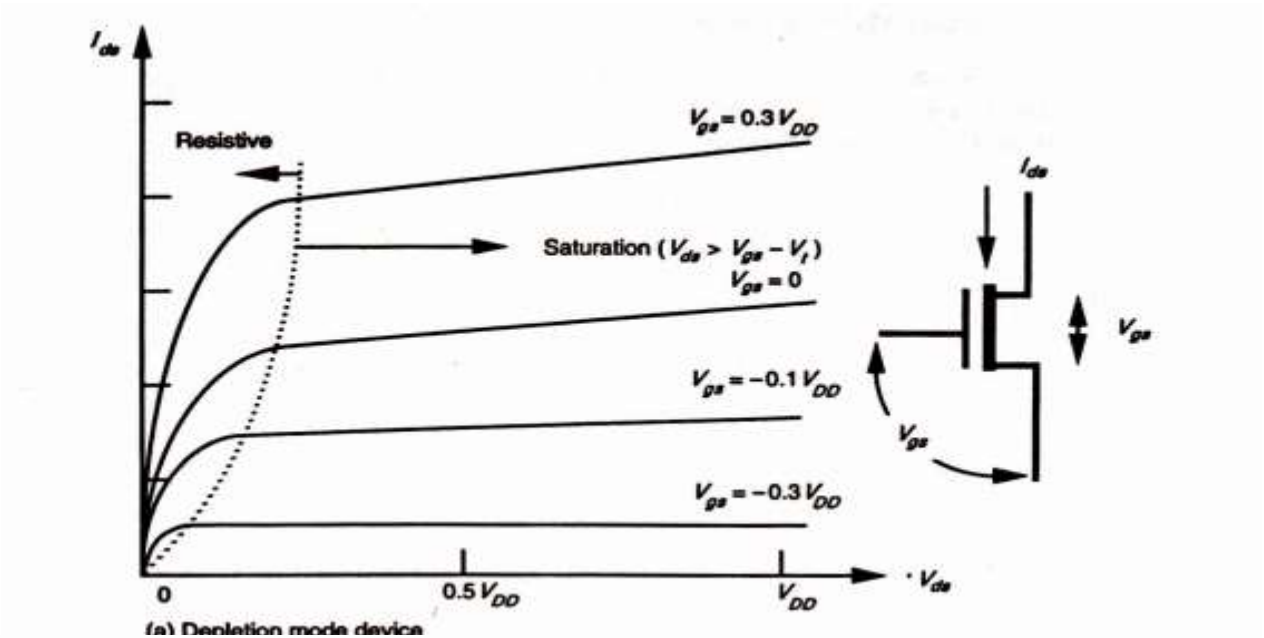
or

$$I_{ds} = \frac{C_g\mu}{2L^2}(V_{gs} - V_t)^2 \qquad (2.5b)$$

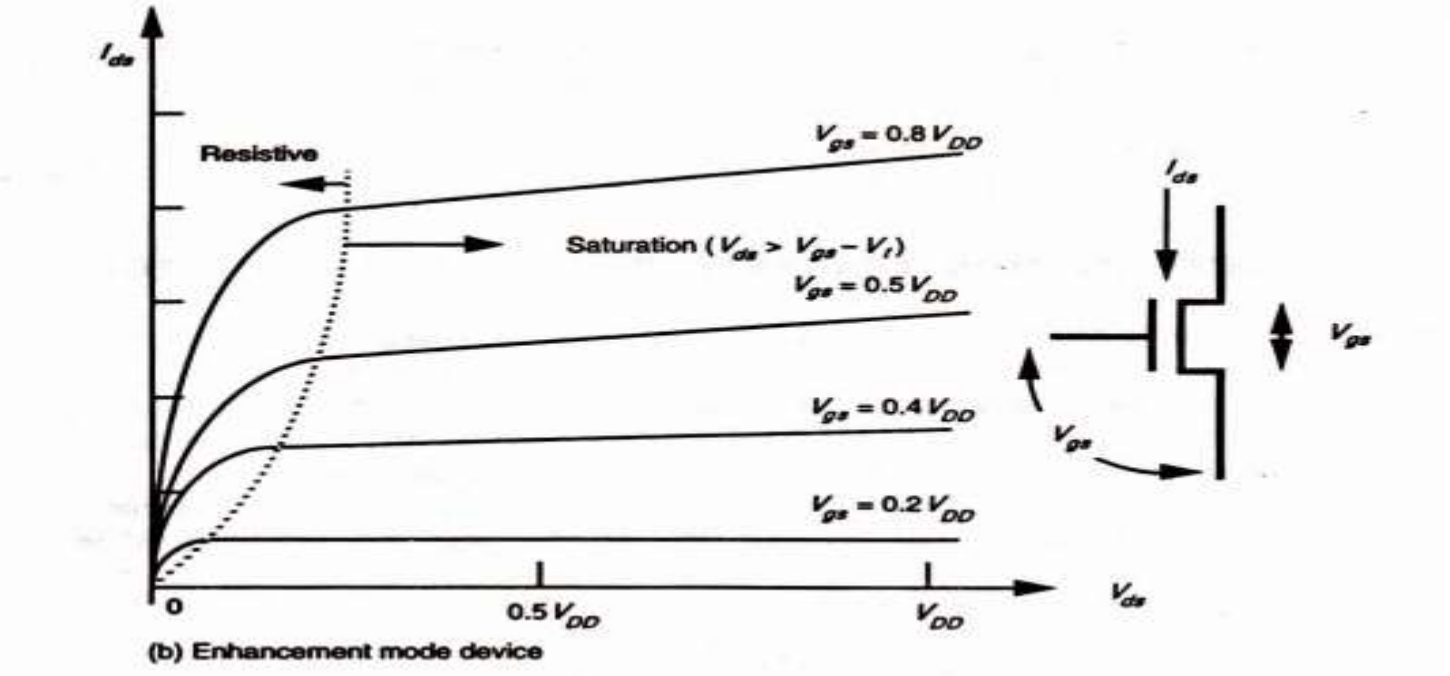
We may also write

$$I_{ds} = C_0\mu\frac{W}{2L}(V_{gs} - V_t)^2 \qquad (2.5c)$$

The expressions derived for Ids hold for both enhancement and depletion mode devices, but it should be noted that the threshold voltage for the nMOS depletion mode device (denoted as V_{td}) is negative.



(a) Depletion mode device



(b) Enhancement mode device

MOS transistor characteristics .

ASPECTS OF MOS TRANSISTOR THRESHOLD VOLTAGE Vt

Q)Define threshold voltage with suitable equation of a MOS device.

The gate structure of aMOS transistor consists, electrically, 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 sufficiamt gate voltage to neutralize these charges and enable the underlying silicon to undergo an inversion due to the electric field from the gate.

The threshold voltage V_t may be expressed as:

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

where

Q_B = the charge per unit area in the depletion layer beneath the oxide

Q_{SS} = charge density at $Si:SiO_2$ interface

C_0 = capacitance per unit gate area
ϕ_{ms} = work function difference between gate and Si
ϕ_{fN} = Fermi level potential between inverted surface and bulk Si.

Now, for polysilicon gate and silicon substrate, the value of ϕ_{ms} is negative but negligible, and the magnitude and sign of V_t are thus determined by the balance between the remaining negative term $\frac{-Q_{SS}}{C_0}$ and the other two terms, both of which are positive. To evaluate V_t, each term is determined as follows:

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

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

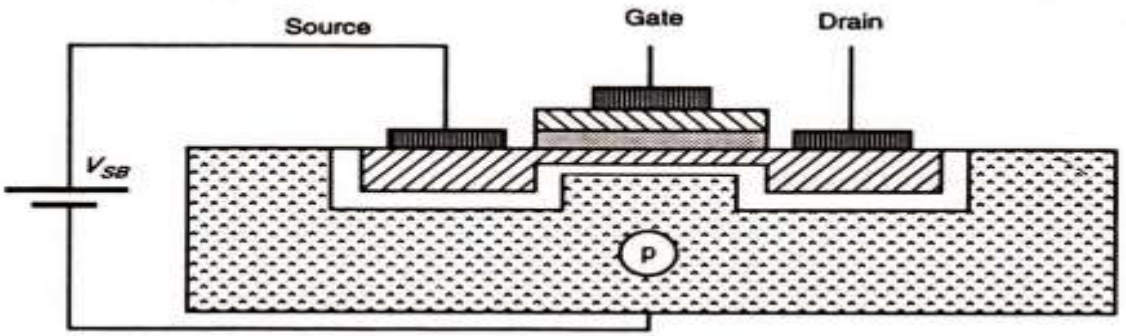
$$Q_{SS} = (1.5 \text{ to } 8) \times 10^{-8} \text{ coulomb/m}^2$$

depending on crystal orientation, and where

V_{SB} = substrate bias voltage (negative w.r.t. source for nMOS, positive for pMOS)
q = 1.6×10^{-19} coulomb
N = impurity concentration in the substrate (N_A or N_D as appropriate)
ε_{si} = relative permittivity of silicon ÷ 11.7
n_i = intrinsic electron concentration (1.6×10^{10}/cm^3 at 300°K)
k = Boltzmann's constant = 1.4×10^{-23} joule/°K

Q)Write short notes on body effect.

The body effects may also be taken into account since the substrate may be biased with respect to the source, as shown in Figure



Body effect (nMOS device shown)

Increasing V_{SB} causes the channel to be depleted of charge carriers and thus the threshold voltage is raised.

Change in V_t is given by $\Delta V_t \doteq \gamma(V_{SB})^{1/2}$ 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.

Alternatively, we may write

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

where $V_t(0)$ is the threshold voltage for $V_{SB} = 0$.

To establish the magnitude of such effects, typical figures for V_t are as follows:

For nMOS enhancement mode transistors:

$V_{SB} = 0$ V; $V_t = 0.2V_{DD}$ (= +1 V for V_{DD} = +5 V)

$V_{SB} = 5$ V; $V_t = 0.3V_{DD}$ (= +1.5 V for V_{DD} = +5 V)

} Similar but negative values for pMOS

For nMOS depletion mode transistors:

$V_{SB} = 0$ V; $V_{td} = -0.7V_{DD}$ (= –3.5 V for V_{DD} = +5 V)

$V_{SB} = 5$ V; $V_{td} = -0.6V_{DD}$ (= –3.0 V for V_{DD} = +5 V)

MOS TRANSISTOR TRANSCONDUCTANCE g_m AND OUTPUT CONDUCTANCE g_{ds}

Transconductance expresses the relationship between output current I_{ds} and the input voltage V_{gs} and is defined as

$$g_m = \frac{\delta I_{ds}}{\delta V_{gs}} \Big| V_{ds} = \text{constant}$$

To find an expression for g_m in terms of circuit and transistor parameters, consider that the charge in channel Q_c is such that

$$\frac{Q_c}{I_{ds}} = \tau$$

where τ is transit time. Thus change in current

$$\delta I_{ds} = \frac{\delta Q_c}{\tau_{ds}}$$

Now

$$\tau_{ds} = \frac{L^2}{\mu V_{ds}} \qquad \text{(from 2.2)}$$

Thus

$$\delta I_{ds} = \frac{\delta Q_c V_{ds} \mu}{L^2}$$

but change in charge

$$\delta Q_c = C_g \delta V_{gs}$$

so that

$$\delta I_{ds} = \frac{C_g \delta V_{gs} \mu V_{ds}}{L^2}$$

Now

$$g_m = \frac{\delta I_{ds}}{\delta V_{gs}} = \frac{C_g \mu V_{ds}}{L^2}$$

In saturation

$$V_{ds} = V_{gs} - V_t$$

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

and substituting for $C_g = \frac{\varepsilon_{ins} \varepsilon_0 WL}{D}$

$$g_m = \frac{\mu \varepsilon_{ins} \varepsilon_0}{D} \frac{W}{L} (V_{gs} - V_t)$$

Alternatively,

$$g_m = \beta (V_{gs} - V_t)$$

It is possible to increase the gm, of a MOS device by increasing its width. However, this will also increase the input capacitance and area occupied

The output conductance g_{ds} can be expressed by

$$g_{ds} = \frac{\delta I_{ds}}{\delta V_{gs}} = \lambda . I_{ds} \; \alpha \left(\frac{1}{L}\right)^2$$

Here the strong dependence on the channel length is demonstrated as

$$\lambda \alpha \left(\frac{1}{L}\right) \text{ and } I_{ds} \; \alpha \left(\frac{1}{L}\right)$$

for the MOS device.

MOS TRANSISTOR FIGURE OF MERIT ω_0

Q)What is the figure of merit of a MOS transistor? Mention the suitable expression for figure of merit.

An indication of frequency response may be obtained from the parameter ω_0 where

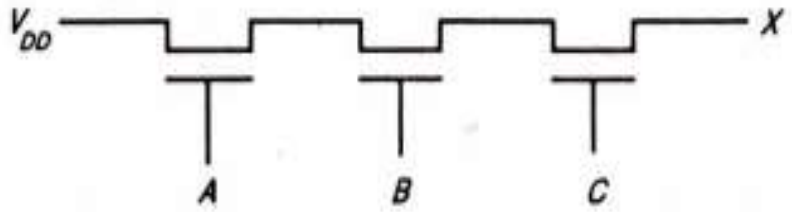
$$\omega_0 = \frac{g_m}{C_g} = \frac{\mu}{L^2} (V_{gs} - V_t) \left(= \frac{1}{\tau_{sd}}\right) \qquad (2.8)$$

This shows that switching speed depends on gate voltage above threshold and on carrier mobility and inversely as the square of channel length. A fast circuit requires that g_m be as high as possible.

$$\mu_n = 1250 \text{ cm}^2/\text{V sec}$$
$$\mu_p = 480 \text{ cm}^2/\text{V sec}$$

THE PASS TRANSISTOR

Unlike bipolar transistors, the isolated nature of the gate allows MOS transistors to be used as switches in series with lines carrying logic levels in a way that is similar to the use of relay contacts. This application of the MOS device is called the pass transistor and switching logic arrays can be formed-for example, an And array as in Figure



Note: Means must exist so that X assumes ground potential when $A + B + C = 0$.

FIGURE Pass transistor And gate

Channel Length Modulation

- I_{DS} = current from drain to source OR drain-source current
- V_{DS} = drain to source voltage
- L = length of the channel

Now for the ideal case, in the saturation region, I_{DS} becomes independent of V_{DS} i.e. in the saturation region channel is pinched off at the drain end and a further increase in V_{DS} has no effect on the channel's shape.

But in practice increase in V_{DS} does affect the channel. In the saturation region, when V_{DS} increases, the channel pinch-off point is moved slightly away from the drain, towards the source as the drain electron field "pushes" it back. The reverse bias depletion region widens and the effective channel length decreases by an amount of ΔL for an increase in V_{DS}.

Thus the channel no longer "touches" the drain and acquires an asymmetrical shape that is thinner at the drain end. This phenomenon is known as channel length modulation.

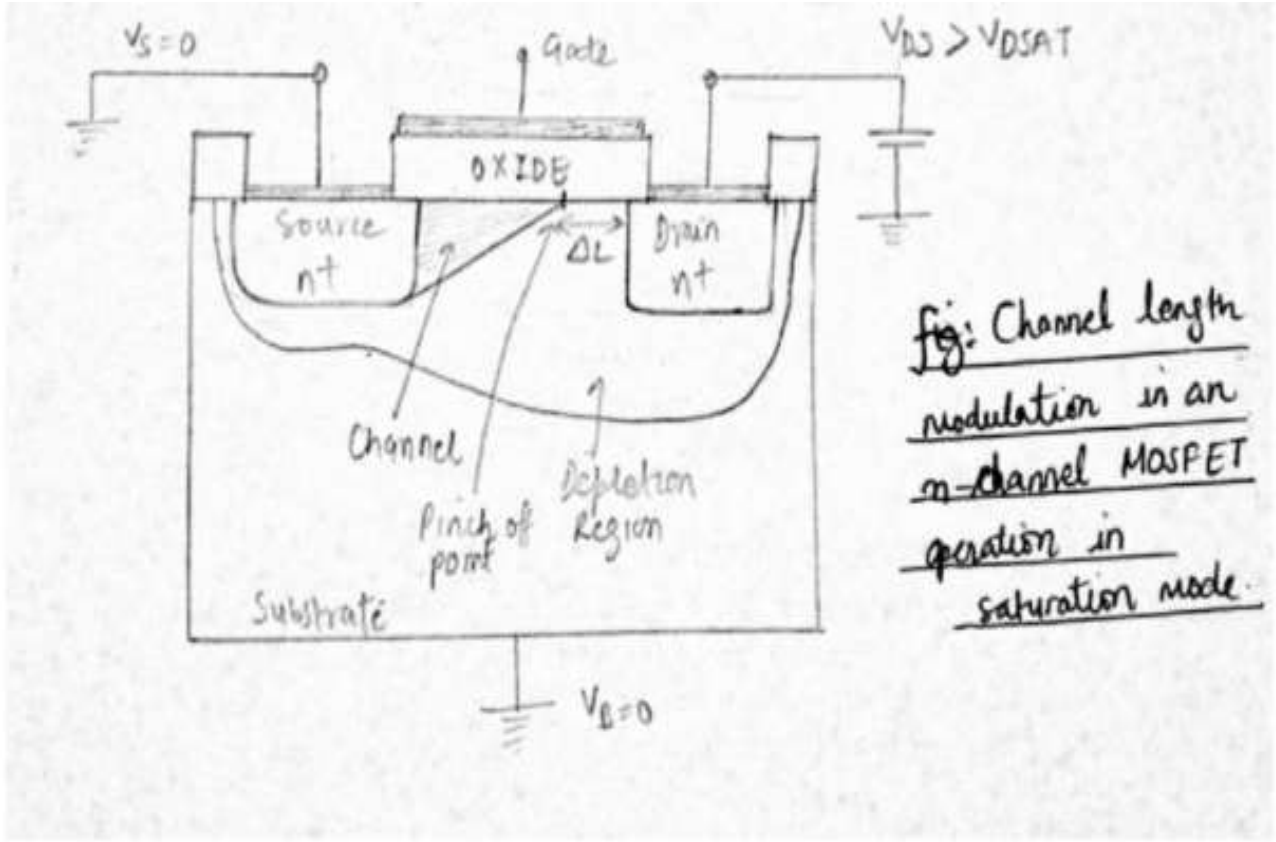
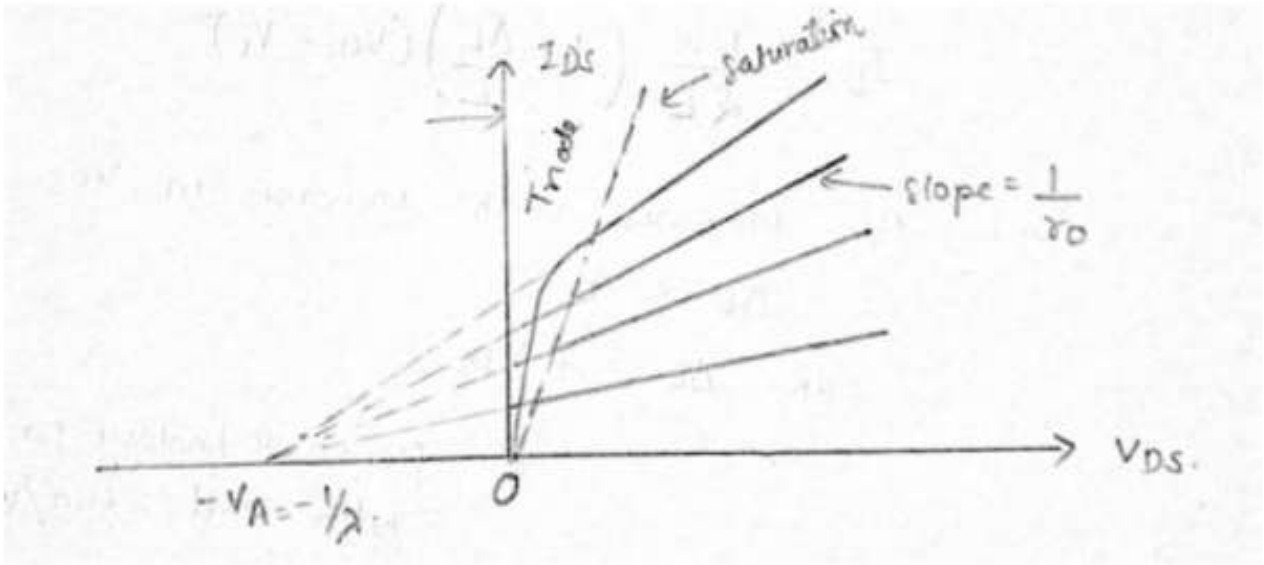


Fig: Channel length modulation in an n-channel MOSFET operation in saturation mode.

Thus channel length modulation can be defined as the change or reduction in length of the channel (L) due to increase in the drain to source voltage (V_{DS}) in the saturation region.

rge devices, this effect is negligible but for shorter devices ΔL/L becomes important. Also in the saturation region due to channel length modulation, I_{DS} increases with increase in V_{DS} and also increases with the decrease in channel length L.

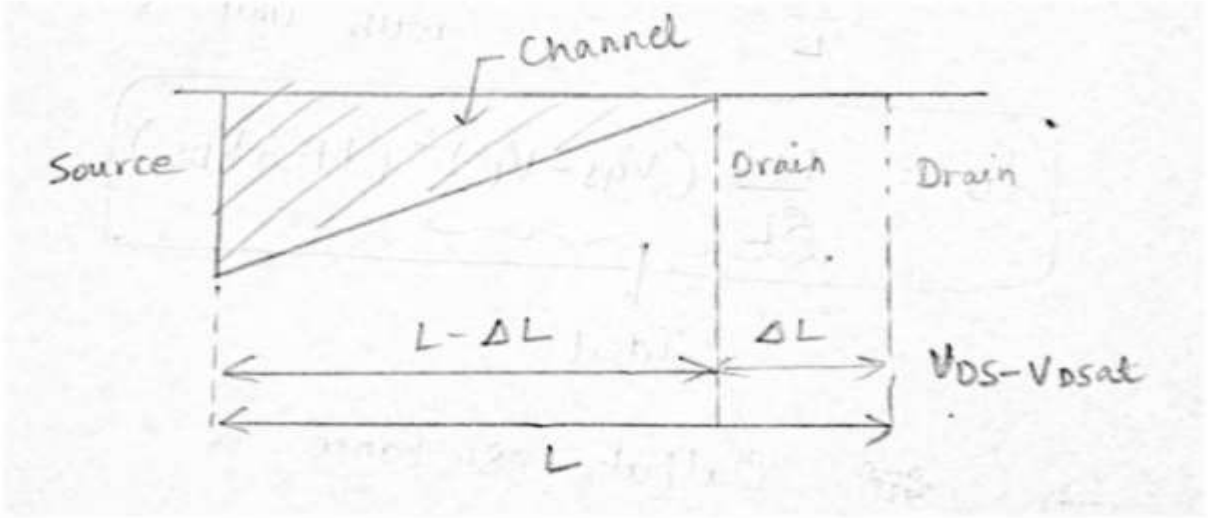
The voltage-current curve is no longer flat in this region.



The drain current with channel length modulation is given by:

$$\boxed{I_{DS} = I_D = I_{Dsat}(1 + \lambda V_{DS})}$$

DERIVATION:



To account for the dependence of I_D on V_{DS} in the saturation region, replace L by L – ΔL. We know that in the saturation region, drain to source current (I_{DS} = I_D) is given by:

$$I_D = \frac{kW}{2L}(V_{GS} - V_t)^2$$

$$I_D = \left(\frac{k}{2}\right)\left(\frac{W}{L-\triangle L}\right)(V_{GS} - V_t)^2$$

$$I_D = \left(\frac{k}{2L}\right)\left(\frac{W}{1-\frac{\triangle L}{L}}\right)(V_{GS} - V_t)^2$$

Assuming $\frac{\triangle L}{L} < 1$

$$I_D = \left(\frac{kW}{2L}\right)\left(1 + \frac{\triangle L}{L}\right)(V_{GS} - V_t)^2$$

Since ΔL increases with increase in V_{DS}

$$\triangle L \propto V_{DS}$$

OR

$$\triangle L = \lambda' V_{DS}$$

where, λ' = process technology parameter with unit *μm/V*.

$$I_D = \left(\frac{kW}{2L}\right)\left(1 + \frac{\lambda' V_{DS}}{L}\right)(V_{GS} - V_t)^2$$

therefore,

$$\boxed{I_{DS} = I_D = I_{Dsat}(1 + \lambda V_{DS})}$$

where,

$\frac{\lambda'}{L} = \lambda$ = process technology parameter with unit V^{-1}

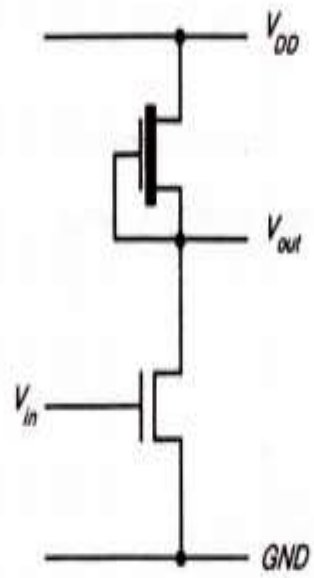
$I_{Dsat} = \left(\frac{kW}{2L}\right)(V_{GS} - V_t)^2$

THE nMOS INVERTER

A basic requirement for producing a complete range of logic circuits is the inverter. This is needed for restoring logic levels, for Nand and Nor gates, and for sequential and memory circuits of various forms.

The basic inverter circuit requires a transistor with source connected to ground and a load resistor of some sort connected from the drain to the positive supply rail Vdd· The output is taken from the drain and the input applied between gate and ground.

Resistors are not conveniently produced on the silicon substrate; even modest values occupy excessively large areas so that some other form of load resistance is required. A convenient way to solve this problem is to use a depletion mode transistor as the load, as shown in Figure .



• With no current drawn from the output, the currents Ids for both transistors must be equal.

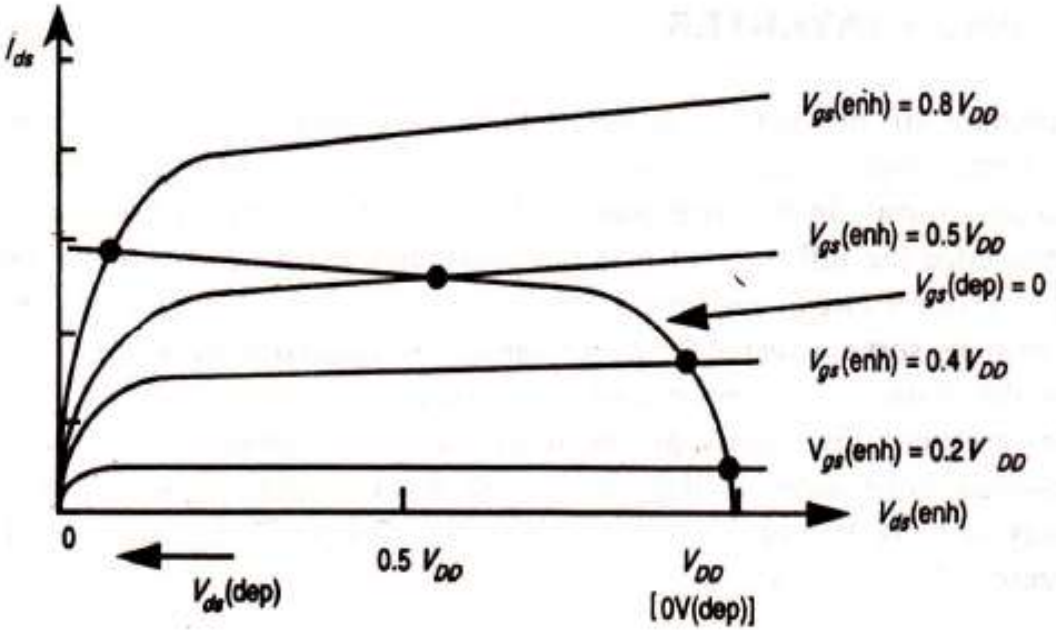
• For the depletion mode transistor, the gate is connected to the source so it is always on and only the characteristic curve Vgs = 0 is relevant.

• 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.

To obtain the inverter transfer characteristic we superimpose the *Vgs* = 0 depletion mode characteristic curve on the family of curves for the enhancement mode device, noting that maximum voltage across the enhancement mode device corresponds to minimum voltage across the depletion mode transistor.

The points of intersection of the curves as in Figure give points on the transfer characteristic, which is of the form shown in Figure.

• Note that as *Vin(=Vgs* p.d. transistor) exceeds the p.d. threshold voltage current begins to flow. The output voltage *Vout* thus decreases and the subsequent increases in *Vin* will cause the p.d. transistor to come out of saturation and become resistive. Note that the p.u. transistor is initially resistive as the p.d. turns on.



$V_{ds}(enh) = V_{DD} - V_{ds}(dep) = V_{out}$

$V_{gs}(enh) = V_{in}$ . . . intersection points give transfer characteristic

FIGURE Derivation of nMOS Inverter transfer characteristic.

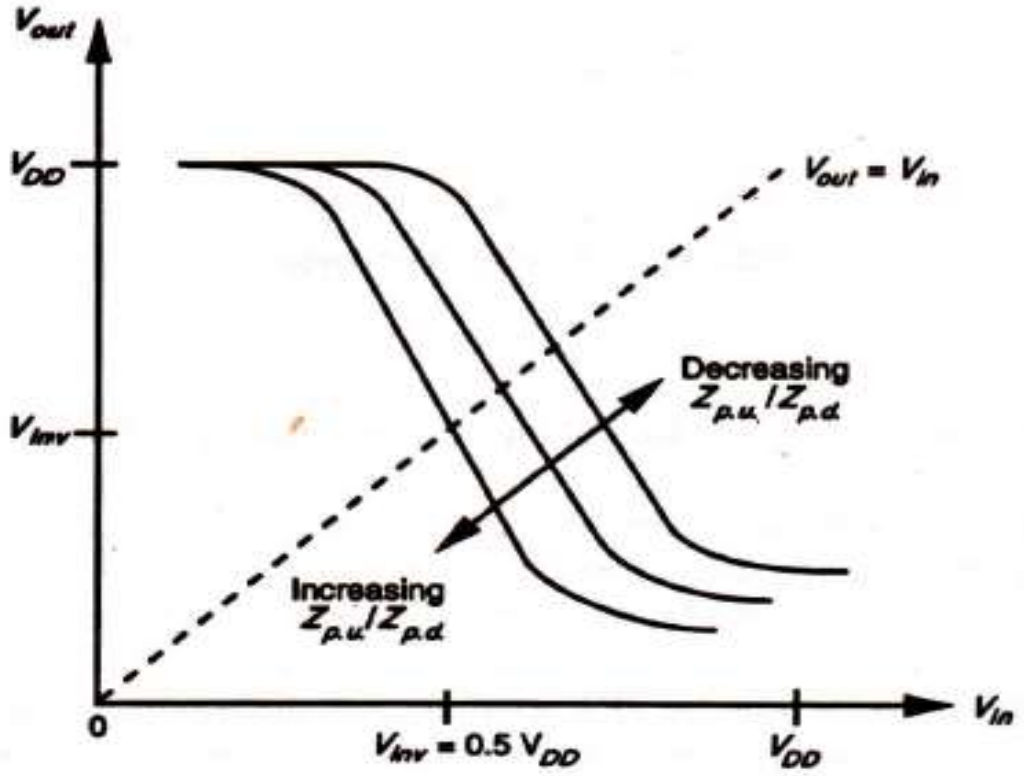


FIGURE nMOS Inverter transfer characteristic

During transition, the slope of the transfer characteristic determines the gain:

$$\text{Gain} = \frac{\delta V_{out}}{\delta V_{in}}$$

ALTERMTIVE FORMS OF PULL-UP

1. *Load resistance* R_L

This arrangement is not often used because of the large space requirements of resistors produced in a silicon substrate.

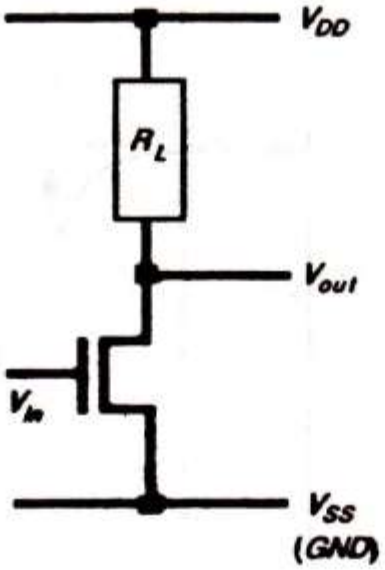


FIGURE Resistor pull-up.

2. *nMOS depletion mode transistor pull-up*

(a) Dissipation is high ,since rail to rail current flows when *Vin* = logical 1.
(b) Switchlng of output from 1 to 0 begins when *Vin* exceeds *Vt* of p.d. device.
(c) When switching the output from 1 to 0, the p.u. device is non-saturated initially
and this presents lower resistance through which to charge capacitive loads

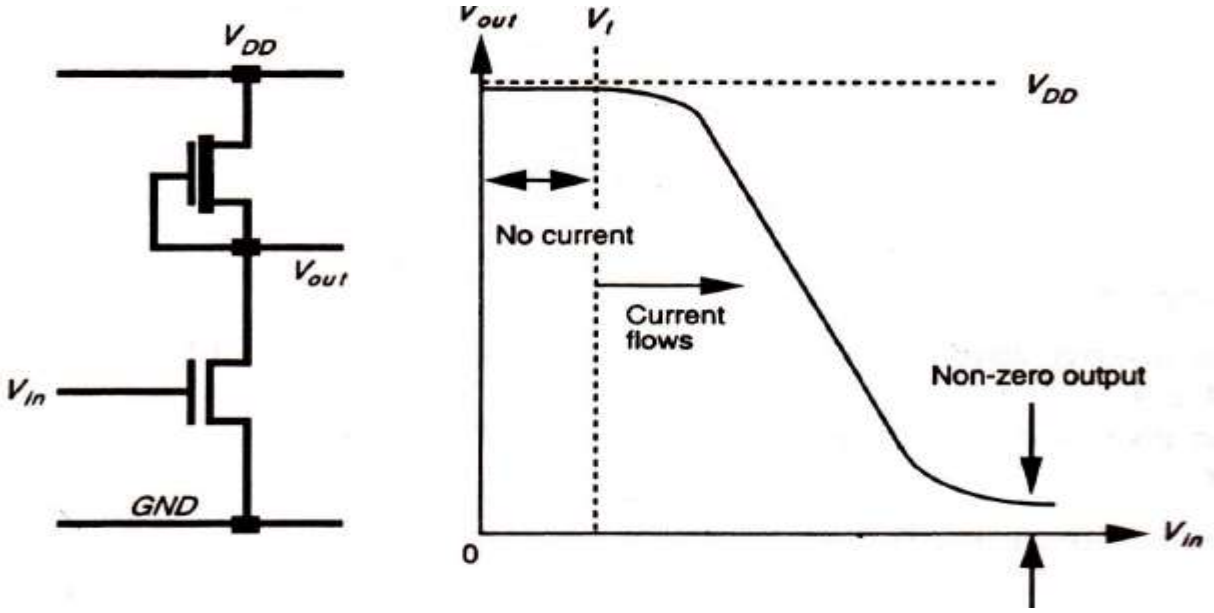


FIGURE 2.12 nMOS depletion mode transistor pull-up and transfer characteristic.

3. *nMOS enhancement mode pull-up*

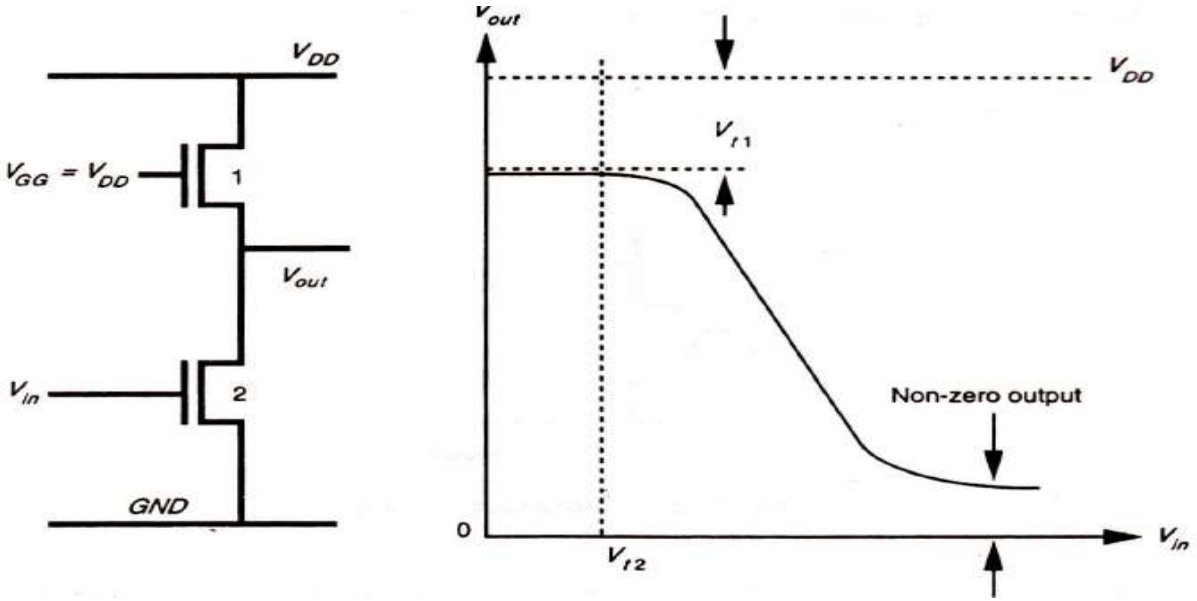


FIGURE nMOS enhancement mode pull-up and transfer characteristic

(a) Dissipation is high since current flows when *Vin* =logical 1

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

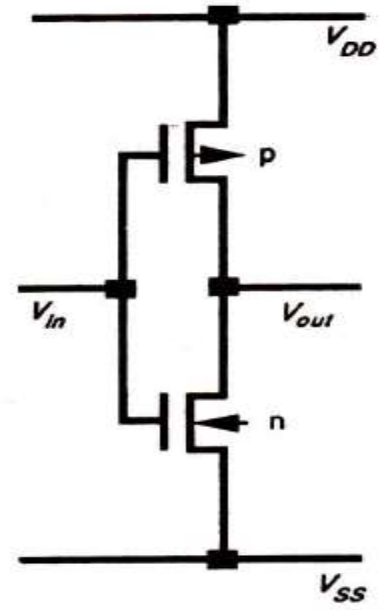
(c) V_{GG} may be derived from a switching source, for example, one phase of a clock,

so that dissipation can be greatly reduced.

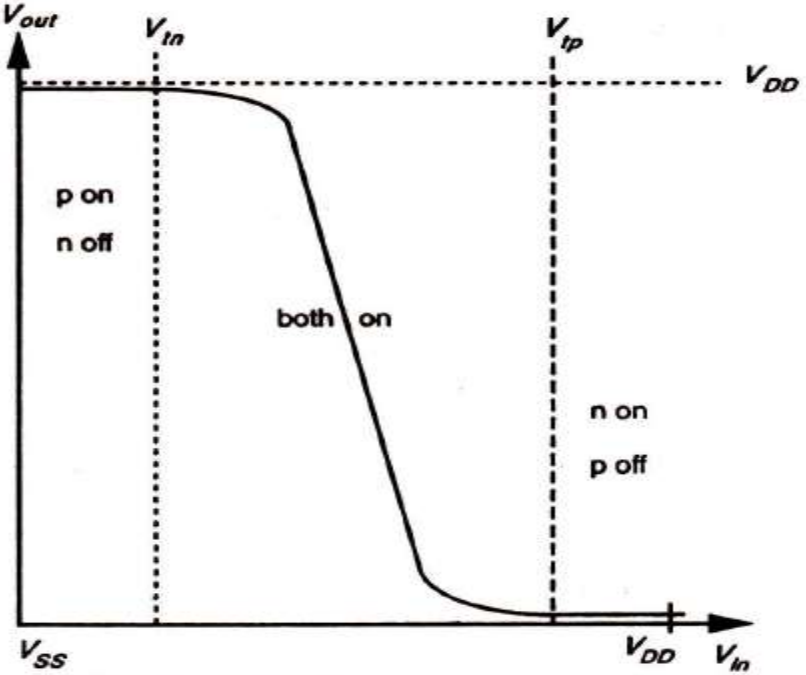
(d) If V_{GG} is higher than V_{DD} then an extra supply rail is required.

4. *Complementary transistor pull-up* (CMOS)

(a) No current flow 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.



(a) Circuit



(b) Transfer characteristic

THE CMOS INVERTER

The general arrangement and characteristics are illustrated in Figure We have seen that the current/voltage relationships for the MOS transistor may be written

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

in the resistive region, or

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

in saturation. In both cases the factor K is a technology-dependent parameter such that

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

The factor W/L is, of course, contributed by the geometry and it is common practice to write

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

so that, for example

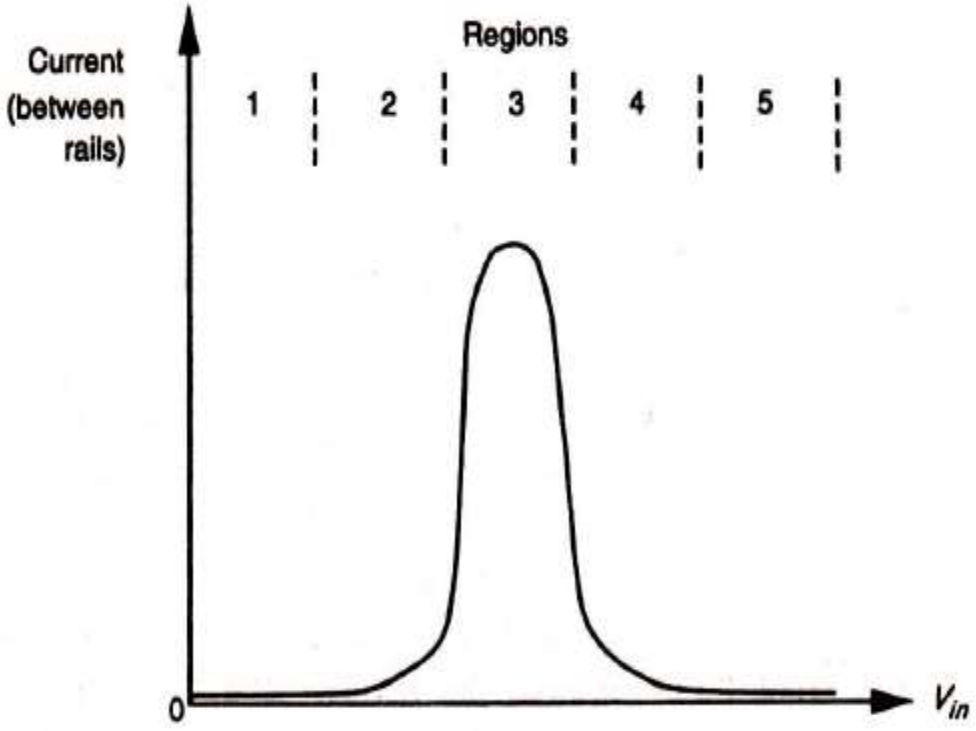
$$I_{ds} = \frac{\beta}{2}(V_{gs} - V_t)^2$$

in saturation, and where β may be applied to both nMOS and pMOS transistors as follows:

$$\beta_n = \frac{\varepsilon_{ins}\varepsilon_0\mu_n}{D}\frac{W_n}{L_n}$$

$$\beta_p = \frac{\varepsilon_{ins}\varepsilon_0\mu_p}{D}\frac{W_p}{L_p}$$

where *Wn* and L_m W_P and L_P are the n- and p-transistor dimensions respectively. it may be seen that the CMOS inverter has five distinct regions of operation.



(c) CMOS inverter current versus V_{in}

in *region 1* for which *Vin* . =
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 *V DD*
through the p-transistor. A good logic 1 output voltage is thus present at the output

In *region 5 Vin* = logic 1, the n-transistor is fully on while the p-transistor is fully off.
Again, no current flows and a good 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 V_{DD} to V_{ss}. 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 may write

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

where

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

and

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

from whence we can express V_{in} in terms of the β ratio and the other circuit voltages and currents

from whence we can express *Vin* in terms of the β ~ ratio and the other circuit voltages and currents

$$V_{in} = \frac{V_{DD} + V_{tp} + V_{tn}(\beta_n/\beta_p)^{1/2}}{1 + (\beta_n/\beta_p)^{1/2}}$$

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

.

If $\beta n = \beta_P$ and if *Vtn* = -*Vtp*, then from equation

Vin = 0.5 V_{DD}

This implies that the changeover between logic levels is symmetrically disposed about the point at which

$$V_{in} = V_{out} = 0.5\ V_{DD}$$

since only at this point will the two β factors be equal. But for $\beta_n = \beta_p$ the device geometries must be such that

$$\mu_p W_p/L_p = \mu_n W_n/L_n$$

Now the mobilities are inherently unequal and thus it is necessary for the width to length ratio of the p-device to be two to three times that of the n-device, namely

$$W_p/L_p \doteq 2.5\ W_n/L_n$$

However, it must be recognized that mobility μ is affected by the transverse electric field in the channel and is thus dependent on V_{gs} (and thus on V_{in} in this in case). It has been shown empirically that the actual mobility is

$$\mu = \mu_z (1 - \phi\ (V_{gs} - V_t))^{-1}$$

ϕ is a constant approximately equal to 0.05, V_t includes any body effect, and μ_z is the mobility with zero transverse field. Thus a β ratio of 1 will only hold good around the point of symmetry when $V_{out} = V_{in} = 0.5V_{DD}$.

The β ratio is often unimportant in many configurations and in most cases minimum size transistor geometries are used for both n- and p-devices. Figure indicates the trends in the transfer characteristic as the ratio is varied.

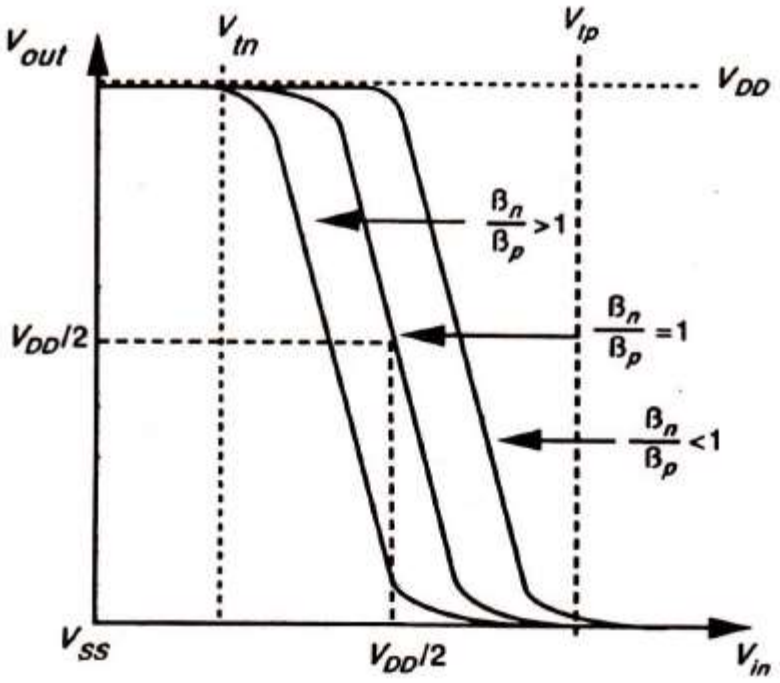


FIGURE Trends In transfer characteristic with β ratio

MOS TRANSISTOR CIRCUIT MODEL

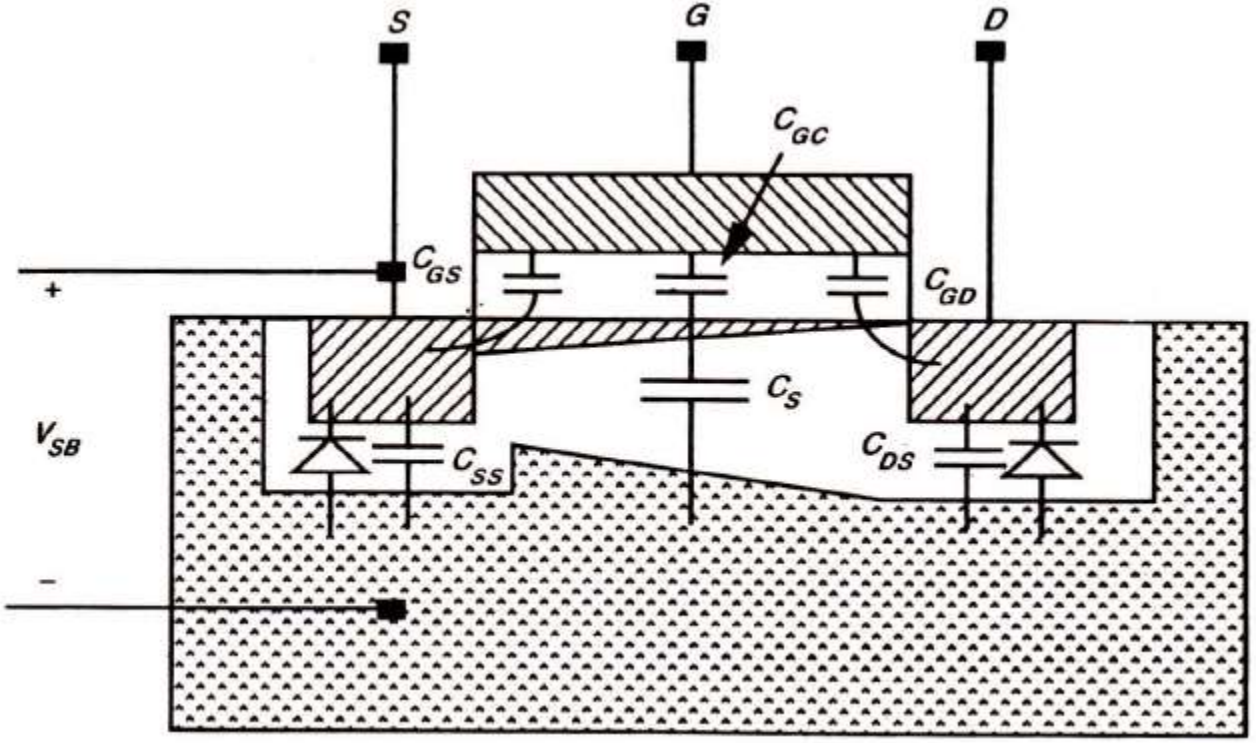


FIGURE nMOS transistor model.

Notes: C_{GC} = gate to channel capacitance
C_{GS} = gate to source capacitance
C_{GD} = gate to drain capacitance

Note that Css indicates source-to-substrate, C_{DS} drain-to-substrate, and Cs channel-to-substrate

BICMOS Inverters

It consists of two bipolar transistors *T1* and *T2* with one nMOS transistor *T3*, and one
pMOS transistor *T4*, both being enhancement mode devices. The action of the circuit is straightforward and may be described as follows:

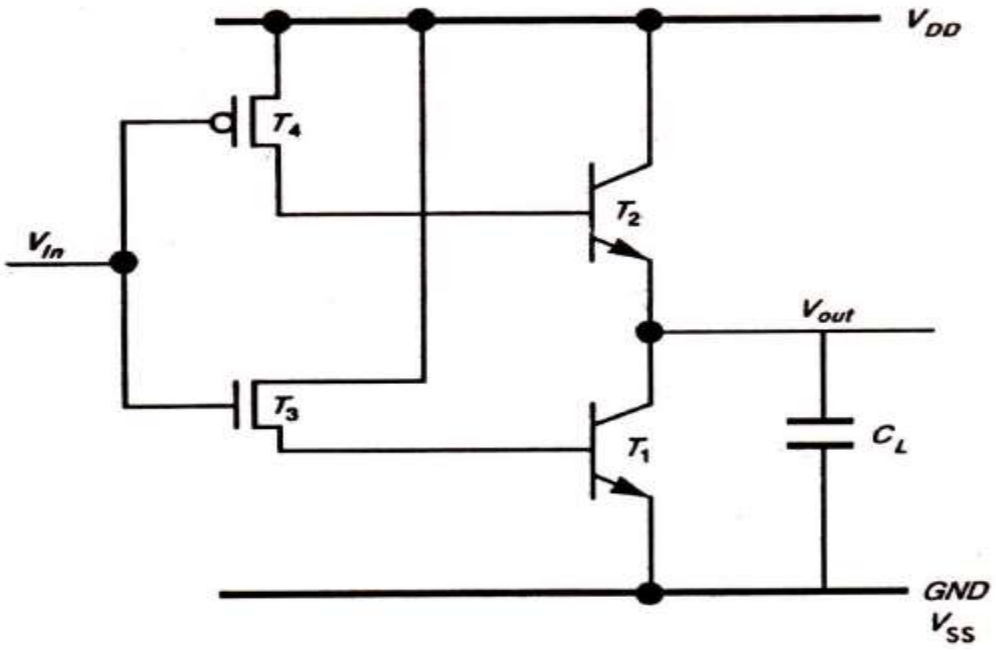


FIGURE A simple BiCMOS inverter.

• With *Vin* = 0 volts *(GND) T3* is off so that *T1* will be non-conducting. But *T4* is on and supplies current to the base of T2 which will conduct and act as a current source to charge the load C_L toward +5 volts(V_{DD}). The output of the inverter will rise to +5 volts less the base to emitter voltage V_{BE} of *T2.*

• With *Vin* = +5 volts*(*V_{DD}*) T4* is off so that *T2* will be non-conducting. But *T3* will now be on and will supply current to the base of *T1* which will conduct and act as current to the load CL discharging it toward 0 volts(GND).The output of the inverter will follow to 0 volts plus the saturation voltage $_{VCE\ Sat}$ from the collecter to the emitter of T1

- The output logic levels will be good and will be close to the rail voltages since V_{CEsat}

is quite small and V_{BE} is approximately + 0.7 volts.

• The inverter has a high input impedance.

• The inverter has a low output impedance.

• The inverter has a high current drive capability but occupies a relatively small area.

• The inverter has high noise margins.

However, owing to the presence of a DC path from *VDD* to *GND* through *T3* and T1 this
is not a good arrangement to implement since there will be a significant static current flow
whenever *Vin* = logic I.

An improved version of this circuit is given in Figure , in which the DC path through *T3* and *T1* is eliminated, but the output voltage swing is now reduced, since the output cannot fall below the base to emitter voltage V_{BE} of *T1.*

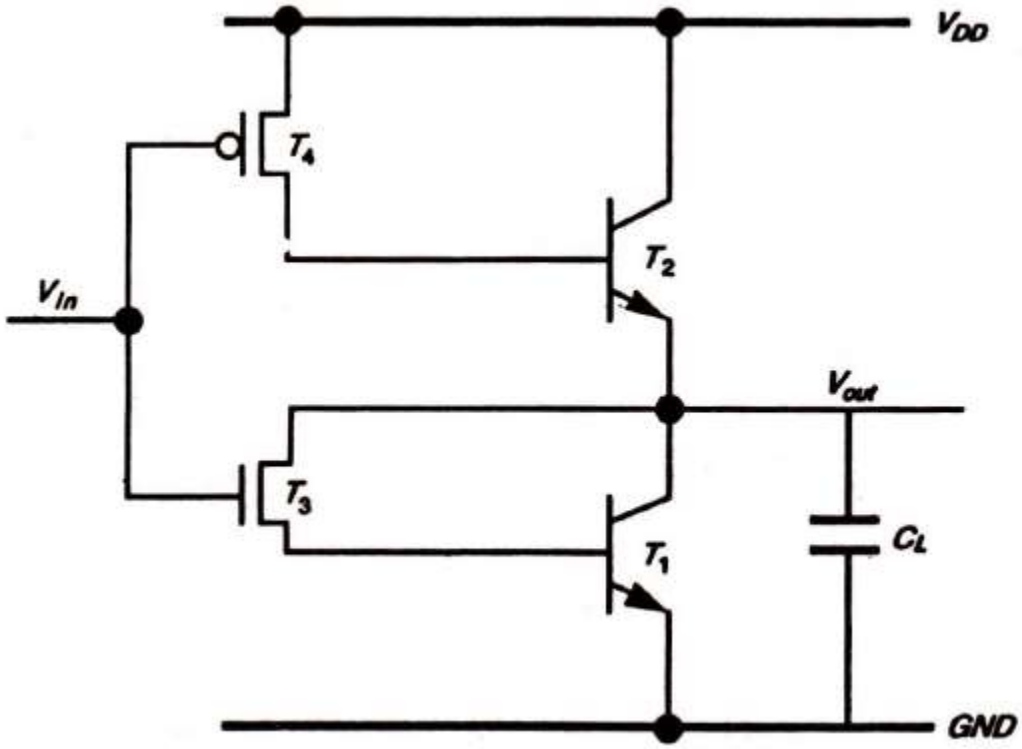


FIGURE An alternative BICMOS Inverter with no static current flow.

An improved inverter arrangement, using resistors, is shown in Figure In this
circuit resistors provide the improved swing of output voltage when each bipolar transistor
is off, and also provide discharge paths for base current during turn off.

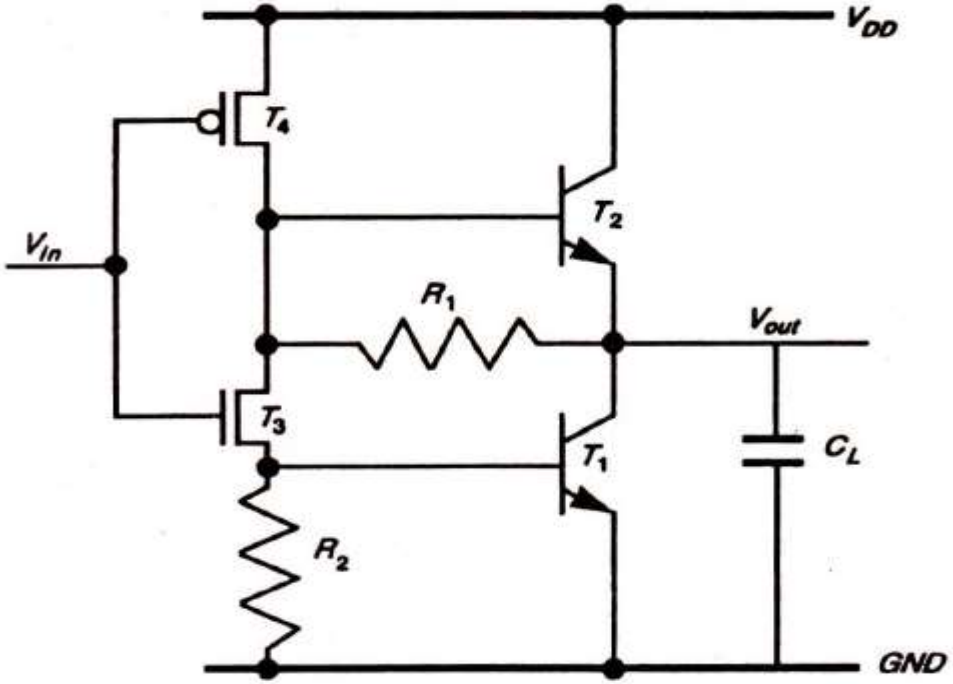


FIGURE ,An improved BICMOS inverter with better output logic levels.

The provision of on chip resistors of suitable value is not always convenient and may
be space-consuming, so that other arrangements-such as in Figure are used.

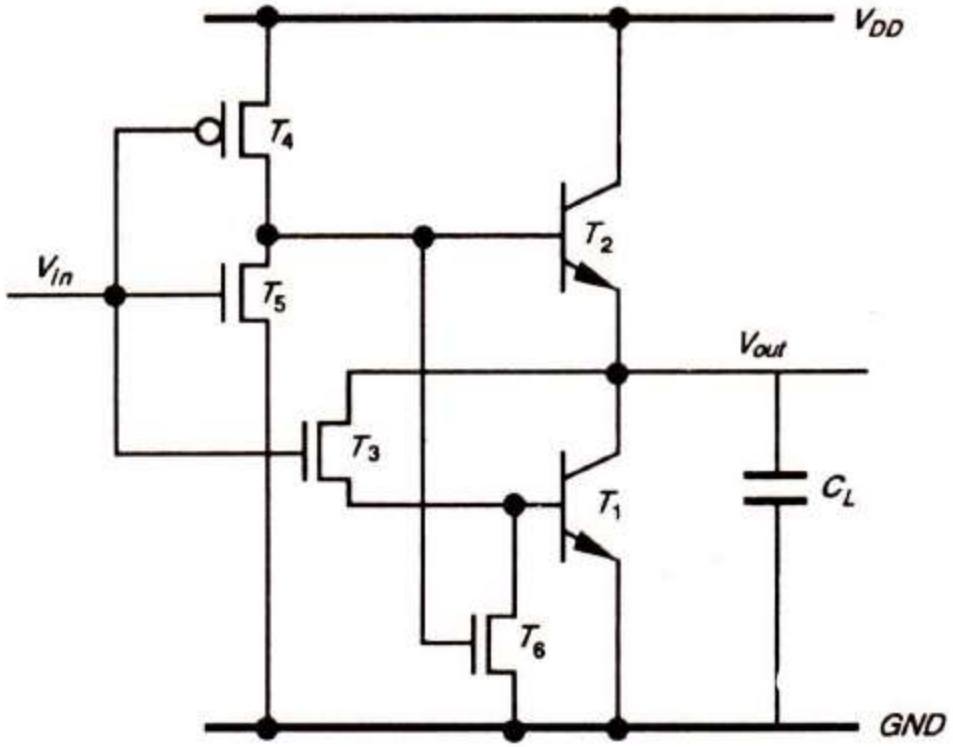


FIGURE An Improved BICMOS Inverter using MOS transistors for base current discharge

LATCH-UP IN CMOS CIRCUITS

Q)Define Latch-up in a MOS circuit. Mention any one of remedial process to reduce latch-up

A problem which is inherent in the p-well and n-well processes is due to the relatively large
number of junctions which are formed in these structures and, as mentioned earlier, the
consequent presence of parasitic transistors and diodes. Latch-up is a condition in which . the
parasitic components give rise to the establishment of low-resistance conducting paths between V_{DD} and Vss with disastrous results.Careful control during fabrication is necessary to avoid this problem

Latch-up may be induced by glitches on the supply rails or by incident radiation. The
mechanism involved may be understood by referring to Figure , which shows the key
parasitic components associated with a p-well structure in which an inverter circuit (for
example) has been formed

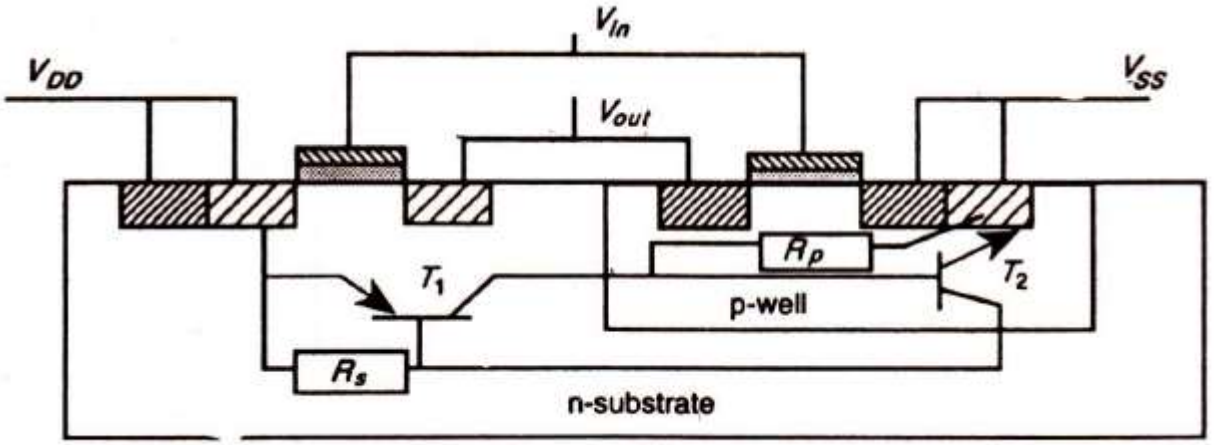


FIGURE Latch-up effect in p-well structure.

There are, in effect, two transistors and two resistances which form a path between V_{DD} and V_{SS}. If sufficient substrate current flows to generate enough voltage across R_S to turn on transistor T1, this will then draw

· current through R_P and, if the voltage developed is sufficient, *T2* will also turn on, establishing a_self-sustaining low-resistance path between the supply rails. If the current gains of the two transistors are such that $\beta_1 \times \beta_2 > 1$, latch-up may occur. Equivalent circuits are given in Figure

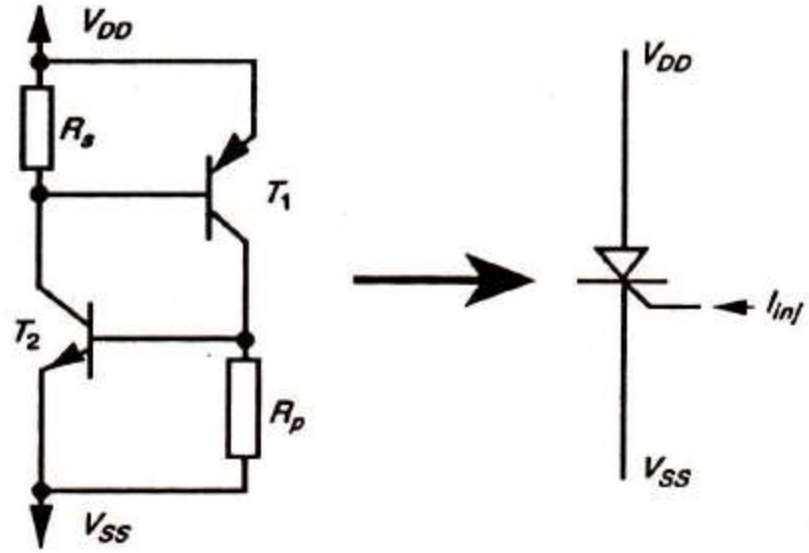


FIGURE Latch-up circuit model.

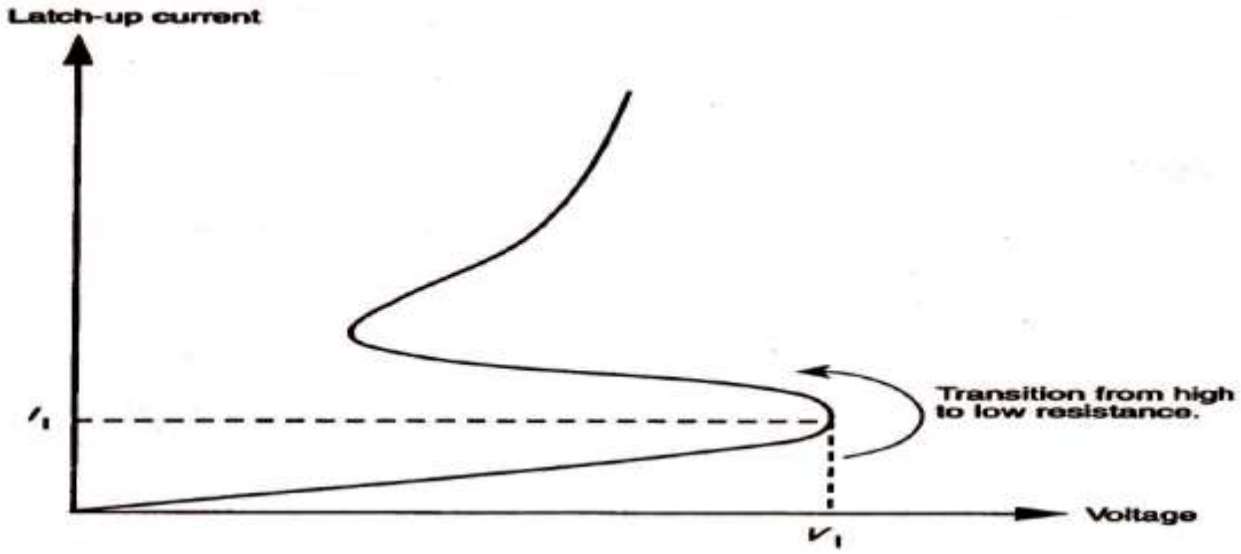


FIGURE Latch-up current versus voltage.

Remedies for the latch-up problem include:

1. an increase in substrate doping levels with a consequent drop in the value of R_S
2. reducing R_P by control of fabrication parameters and by ensuring a low contact resistance to Vss
3. other more elaborate measures such as the introduction of guard rings.

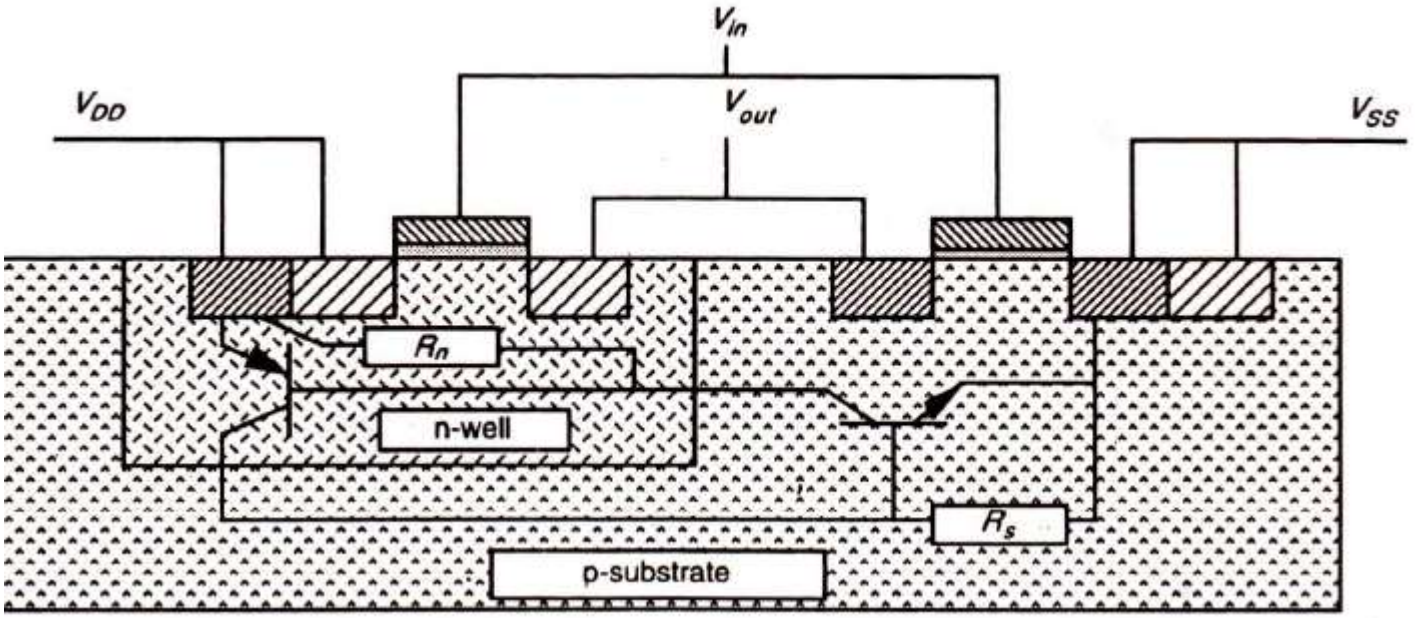


FIGURE Latch-up circuit for n-well process

BICMOS LATCH-UP SUSCEPTIBILITY

One benefit of the BiCMOS process is that it produces circuits which are less likely to suffer from latch-up problems. This is due to several factors:

- A reduction of substrate resistance *RS*
- A reduction of n-well resistance *Rw*.
- A reduction of *RS* and *Rw* means that a larger lateral current is necessary to invite latch-up and a higher value of holding current is also required.