**Unit 2**

**2.1 n-MOS INVERTER**

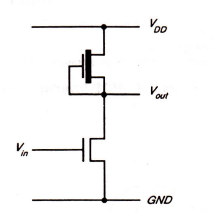
A basic requirement for producing a complete range of logic circuits is the inverter. This is needed for restoring logic levels, for *Nand* and *Nor* gates, and for sequential and memory circuits of various forms . In the treatment of the inverter used in this section, the authors wish to cknowledge the influence of material previously published by Mead and Conway.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 *Vvv·* 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 shown in fig.2.1

Fig 2.1 NMOS Inverter

• 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 *2.2*give points on the transfer characteristic, which is of the form shown in Figure 2.3.

• 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.

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

(.fl

, 

**The point** at which *Vout* = *Vin* is denoted as *Vinv* and it will be noted that the transfer characteristics and *Vinv* can be shifted by variation of the ratio of pull-up to pull down resistances (denoted *Zp.u/Zp.d.* where *Z* is determined by the length to width ratio of the transistor in question).

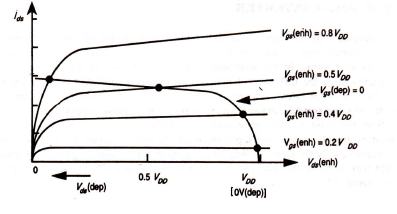
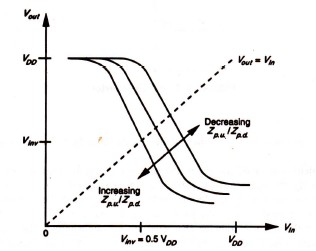
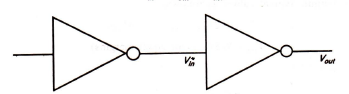


FIGURE 2.2 nMOS Inverter transfer characteristic FIGURE 2. 3 nMOS Inverter transfer characteristic.

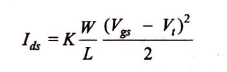
**2.2 DETERMINATION OF PULL-UP TO PULL-DOWN RATIO *(ZP.U/ZP.D.)* FOR AN NMOS INVERTER DRIVEN BY ANOTHER NMOS INVERTER**

Consider the arrangement in Figure 2.8 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 further assume that in order to cascade inverters without degradation of levels we are aiming to meet the requirement.



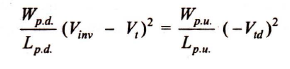
**FIGURE 2.4 nMOS inverter driven directly by another inverter**

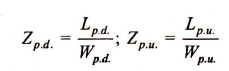
For equal margins around the inverter threshold, we set *Vinv* = 0.5 *VDD·* At this point both transistors are in saturation and

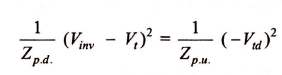
****

In depletion Mode transistor

and in the enhancement mode

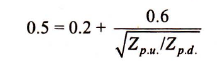
Equating (since currents are the same) ,we have





 where



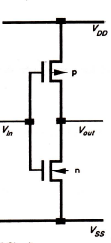
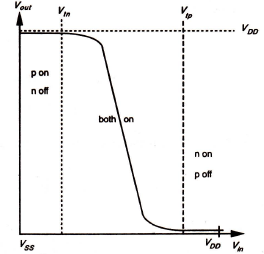
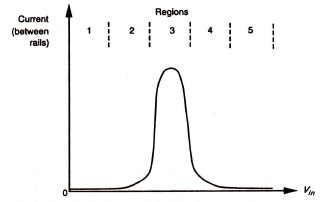




and thus 

for an inverter directly driven by an inverter.

**2.3 CMOS Inverter and its characteristics**



(a)Circuit (b) Transfer characteristic (c) CMOS inverter current versus Vtn

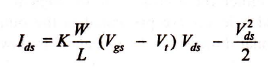
FIGURE 2.5 Complementary transistor pull-up (CMOS).

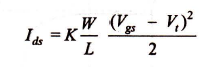
(a) No current flow either for logical 0 or for logical 1 inputs.

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

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

The general arrangement and characteristics are illustrated in Figure 2.5. We have seen (equations 2.4 and 2.5) that the current/voltage relationships for the MOS transistor may be written



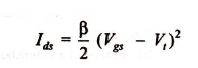
in the resistive region, or

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

The factor *WIL* is, of course, contributed by the geometry and it is common practice to

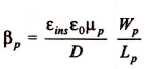
write

so



so that, for example

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



where *Wn* and *Ln WP* and *LP* are the n- and p-transistor dimensions respectively. With regard to Figures 2.5(b) and 2.5( c), it may be seen that the CMOS inverter has five distinct regions of operation. Considering the static conditions first, it may be Seen that 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 *VDD* through the p-transistor. A good logic 1 output voltage is thus present at the output. **In *region 5****: Vin.* = logic1, 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 *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.

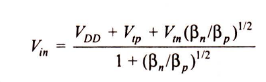
***Region 3:***is the region in which the inverter exhibits gain and in which both transistors are in saturation.

The currents (with regard to Figure 2.5(c)) in each device must be the same since the transistors are in series, so we may write





Where and

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

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 and if *,* then from equation



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

the point at which



since only at this point will the two β factors be equal. But for βn = βp ,the device geometries

must be such that



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



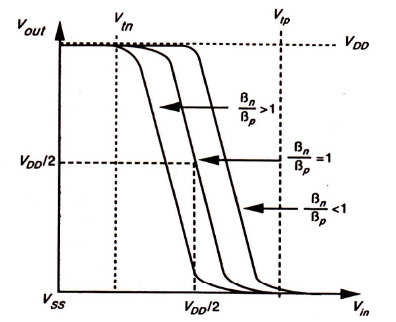
However, it must be recognized that mobility µ is affected by the transverse electric field in

the ·channel and is thus depeqdent on *Vgs* (and thus on *Vin* in this in case). It has been shown

empirically that the actual mobility is

When ɸ is a constant approximately equal to 0.05, *Vt* includes any body effect, and *µz* is the

mobility with zero transverse field. Thus a β ratio of 1 will only hold good around the point

of symmetry when *Vout* = *Vin* = *0.5VDD*



**FIGURE 2.6 Trends In transfer characteristic with β ratio.**

**2.4 LATCH-UP IN CMOS CIRCUITS**

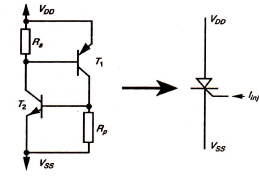
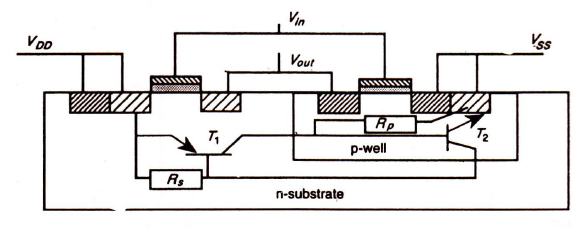
A problem which is inherent in the p-well and n-well processes is due to the relatively large number of junctions which are formed in these structures and, as mentioned earlier, the consequent pre·sence of parasitic transistors and diodes. Latch-up is a condition in which . the parasitic components give rise to the establishment of low-resistance conducting paths between *VDD* 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 2.7, which shows the key parasitic components associated with a p-well structure in which an inverter circuit (for example) has been formed There are, in effect, two transistors and two resistances (associated with the p-well and with regions of the substrate) which form a path between *VDD* and *VSS.* If sufficient substrate current flows to generate enough voltage across *RS* to turn on transistor T1, this will then draw · current through *Rp* 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 β1 x β2 > 1, latch-up may occur. Equivalent circuits are given in Figure 2.8. With no injected current, the parasitic transistors will exhibit high resistance, but sufficient substrate current flow will cause switching to the low-resistance state as already explained. The switching characteristic of the arrangement is outlined in Figure 2.9. Once latched-up, this condition will be maintained until the latch-up current drops below *II.* It is thus essential for a CMOS process to ensure that *V1* and *11* are not readily achieved in any normal mode of operation Remedies for the latch-up problem include

1. an increase in substrate doping levels with a consequent drop in the value of *RS*

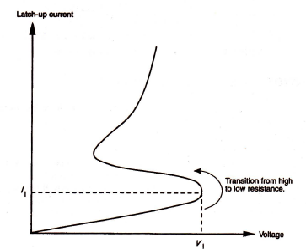
2. reducing *RP* by control of fabrication parameters and by ensuring a low contact

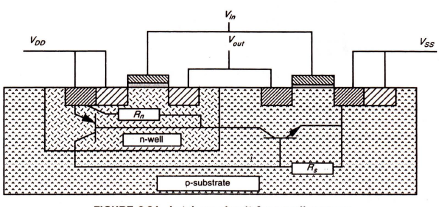
resistance to *Vss;*

3. other more elaborate measures such as the introduction of guard rings.



**FIGURE 2.7Latch-up effect in p-well structure. FIGURE 2.8 Latch-up circuit model.**





**FIGURE 2.9 Latch-up current versus voltage. FIGURE 2.10 Latch-up circuit for n-well process**

* **Stick Diagrams:** Stick diagrams may be used to convey layer information through the use of a color code .VLSI design aims to translate circuit concepts onto silicon. Stick diagrams are a means of capturing topography and layer information using simple diagrams. Stick diagrams convey layer information through colour codes (or monochrome encoding). They act as an interface between symbolic circuit and the actual layout. they show all components/vias. They show relative placement of components. They reach one step closer to the layout . They help plan the layout and routing . but they donot show (1)Exact placement of components (2)Transistor sizes (3)Wire lengths, wire widths, tub boundaries (4)Any other low level details such as parasitics. There are some layers which are commonly used in fabrication.(1) substrate , (2) Well (3) diffusion (4) Polysilicon(5) silicon dioxide(6) supply rails (metallic layer) (7) contacts.

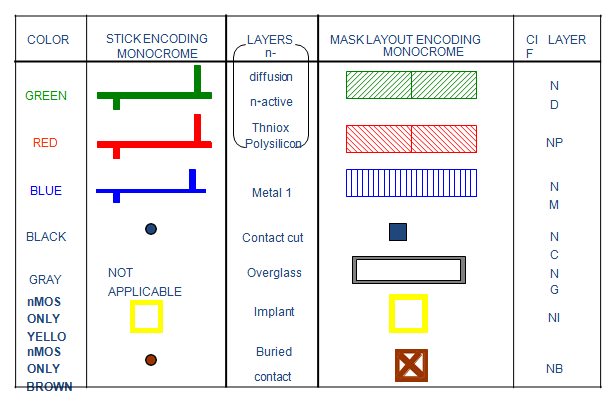
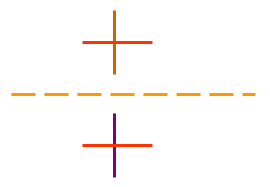
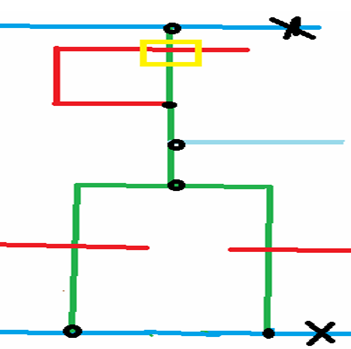
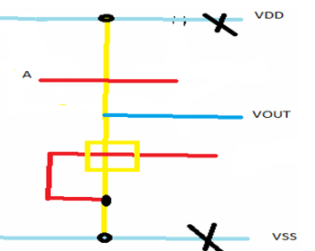
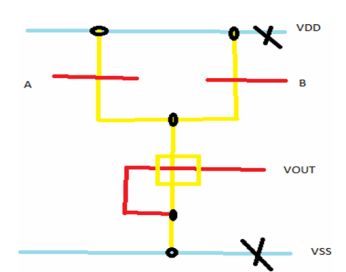
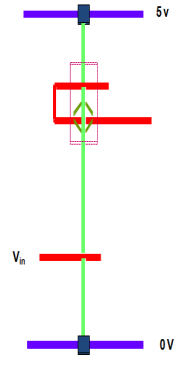
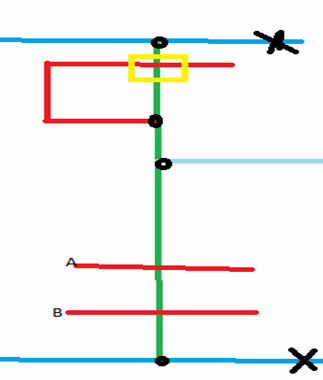
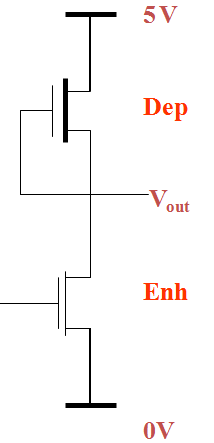
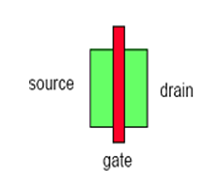
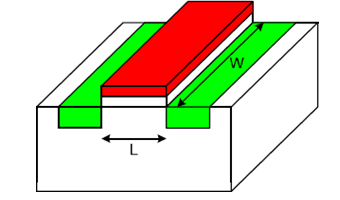


Figure 2.11 Encodings for simple NMOS process

* Rule 1.When two or more ‘sticks’ of the same type cross or touch each other that represents electrical contact
*  Rule2 (a).When two or more ‘sticks’ of different type cross or touch each other there is no electrical contact as.
* Rule2 (b)If electrical contact is needed, we have to show the connection explicitly as
* Rule 3 (a). When a poly crosses diffusion it represents a transistor
* Rule 3(b)If contact is shown on crossing ,it is not a transistor
* Rule 4. In CMOS a demarcation line is drawn to avoid touching of p-diff with n-diff. All pMOS must lie on one side of the line and all nMOS will have to be on the other side as

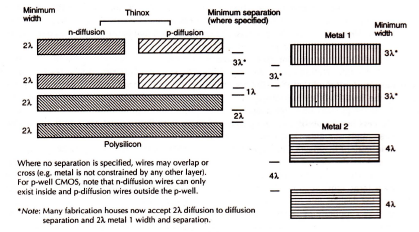


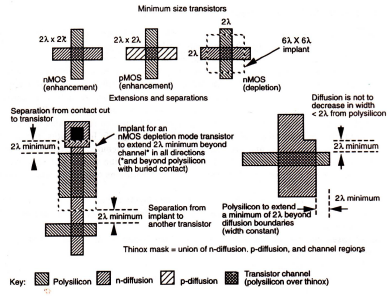


**Lambda-based Design Rules**

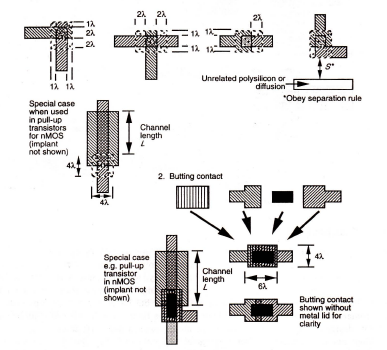
In general, design rules and layout methodology based on the concept of λ provide a process and feature size-independent way of setting out mask dimensions to scale. All paths in all layers will be dimensioned in λ units and subsequently λ can be allocated an appropriate value compatible with the feature size of the fabrication process. This concept means that the actual mask layout design takes little account of the value subsequently allocated to the feature size, but the design rules are such that,· if correctly obeyed, the mask layouts will produce working circuits for a range of values allocated to λ.For example, λ can be allocated a value of 1.0 µm so that minimum feature size on chip will be 2 µm (2λ). Design rules, also due to Mead and Conway, specify line widths, separations, and extensions in terms of λ, and are readily committed to memory. Design rules can be conveniently set out in diagrammatic form as in Figure 2.12 for the widths and separation of conducting paths, and in Figure 2.12 for extensions and separations associated with transistor layouts. The design rules associated with contacts between layers are set out in Figures 2.13 ,2.14 and 2.15 and it will be noted that connection can be made between two or, in the case of nMOS designs, three layers. In all cases, the use of the design rules will be illustrated in layouts resulting from exercises worked through in the text.



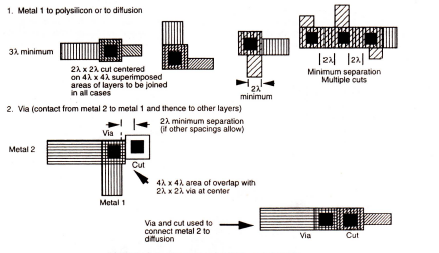






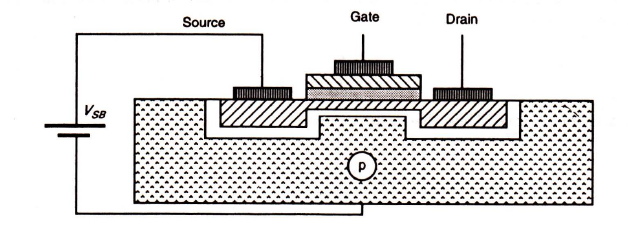




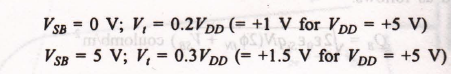
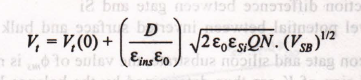




**Body effect**



* Increasing *V*sB *causes the channel to be depleted of charge carriers and thus the threshold* voltage is raised.Change in *Vt is given by Where is a constant which depends on* substrate doping so that the more lightly doped the substrate, the smaller will be the body effect.

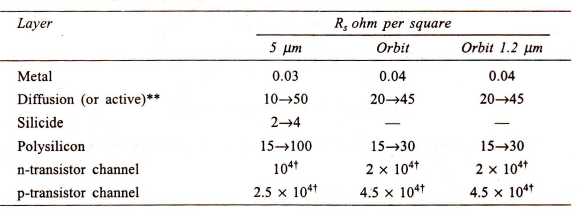




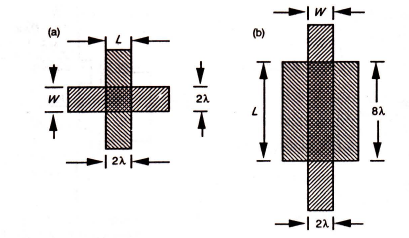
Sheet Resistance:

* The simple n-type pass transistor of Figure (a) has a channel length *L =2 λ. and a* channel width *W =2 λ. . The channel is, therefore, square and channel resistance (with or* without implant)
* The length to width ratio, denoted *Z, is 1: 1 in this case. The transistor structure of*
* ·Figure (b) has a channel length *L = 8λ. and width W = 2 λ... Therefore Z=L/W =4. sheet resistance*
* *R = ZRs = 40 K ohm*

**Sheet resistance of different layers:**

****

**Resistance calculation for transistor channels**

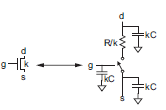
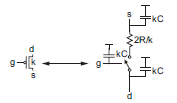


**AREA CAPACITANCES OF LAYERS**





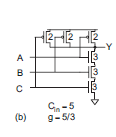
**Gate delays**

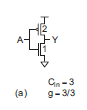
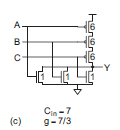


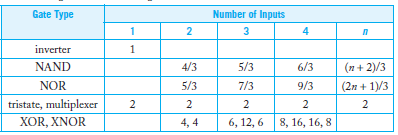
Gate delays

Linear Delay Model *d* = *f* + *p,---(1) p= parasitic delay f* = *gh---(2)*

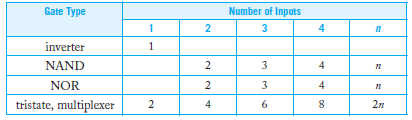
* *p* is the parasitic delay inherent to the gate when no load is attached, The parasitic delay of a gate is the delay of the gate when it drives zero load. It can be estimated with RC delay models. The parasitic delay also depends on the ratio of diffusion capacitance to gate capacitance. *f* is the effort delay or stage effortthat depends on the complexity and fan out of the gate.
* *g=*logical Effort
* Logical effort of a gate is defined as the ratio of the input capacitance of the gate to the input capacitance of an inverter that can deliver the same output current
* *h=*electrical effort where *h=Cout/Cin*

**Logical effort**



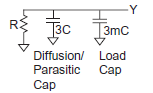


**Parasitic delay**



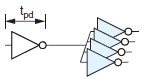
**Delay estimation**

* Q1. Estimate *t*pdfor a unit inverter driving *m* identical unit inverters in following figure

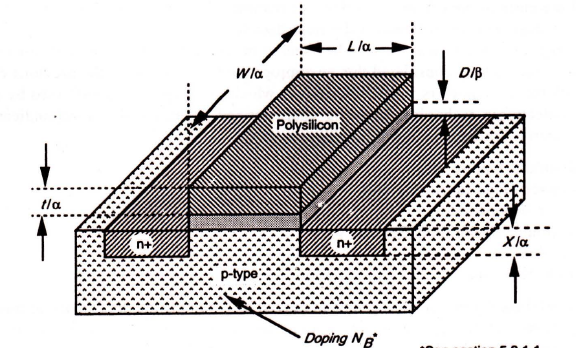


*t*pd= (3 + 3*m*)*RC*.

Q2 .If a unit transistor has *R* = 10 k􀀼 and *C* = 0.1 fF in a 65 nm process, compute the delay,in picoseconds, of the inverter in following figure with a fanout of *h* = 4.



*t*pd =(3 + 3*h*)(1 ps) = 15 ps.

**SCALING MODELS AND SCALING FACTORS**

**Scaled nMOS transistor**

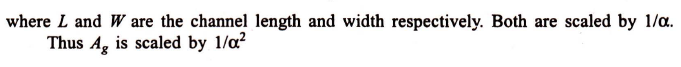
* The scaling down of feature size generally leads to improved performance and it isimportant therefore to understand the effects of scaling. There are also future limits to scaling down .Microelectronic technology may be characterized in terms of several indicators, or figures of merit. Commonly, the following are used:
* Minimum feature size
* Number of gates on one chip
* Power dissipation
* Maximum operational frequency
* Die size
* Production cost

**Models of scaling factor**

* Constant electric field scaling model
* Constant voltage scaling model.
* Combined voltage and dimension scaling model
* In order to accommodate the three models, two scaling factors-1/α and 1/β---are used.
* 1/β is chosen as the scaling factor for supply voltage *VDD and gate oxide thickness D,*
* 1/α is used for all other linear. dimensions, both vertical and horizontal to the chip surface.
* For the constant field model and the constant voltage model, β = α and β = 1 respectively, are applied

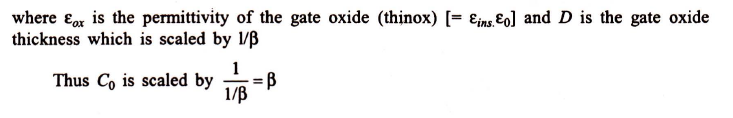
**Scaling factors**

****

Ag=L.W













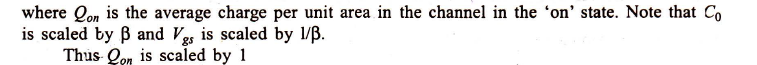






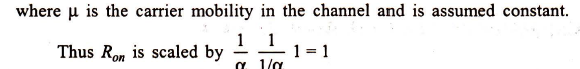


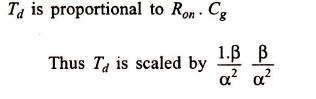






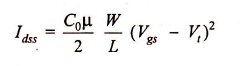


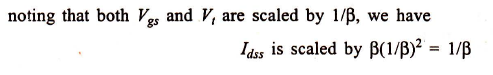






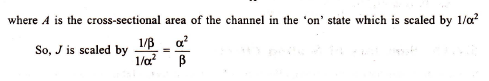




****

****

****

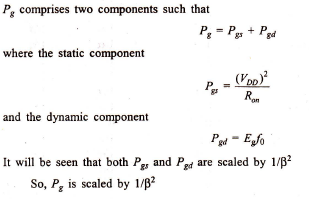
****

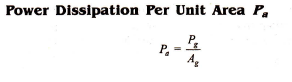
****

****

****

* **Power Dissipation Per Gate *Pg***

****



****

****

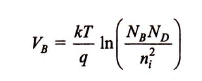
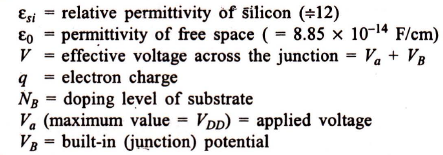
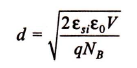
**Limitations of scaling**

* **Substrate Doping:**substrate doping impinges on many of the characteristics of transistors

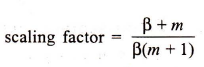
fabricated on it

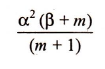
* **(a)Substrate doping scaling factors**As the channel length of a MOS transistor is reduced, the depletion region widths must also be scaled down to prevent the source and drain depletion regions from meeting. Depletion region width *d for the junctions is given by*

**Substrate doping scaling factors**



for the combined voltage and dimension scaling model applied to a transistor

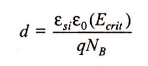
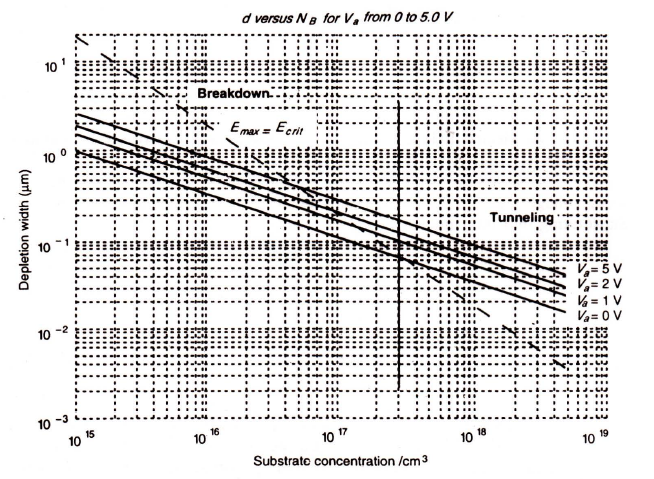


**

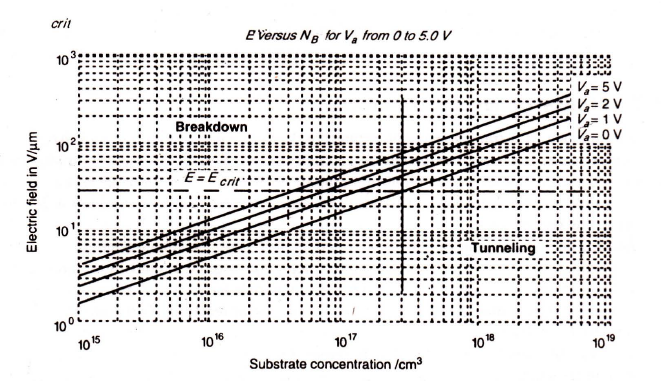
*NB should* be scaled by

* Where *m is large and β is small, the scaling*
* factor for *NB reverts to*
* **(b)Depletion width**
* *NB is increased to reduce the depletion width, but this also* increases the threshold voltage *Vt which is against the required trends for scaling down NB must be kept below*
* *At higher values of NB , the maximum* electric field which can be applied to the gate oxide is insufficient to invert the substrate so that no channel can be formed. depletion width *d and built in* potential *VB will impose limitations on scaling,*

**Depletion width**

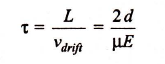
****

* The area of Figure(a), above the dashed line is the region where the increased electric field will induce breakdown. Thus, the point at which the dashed line and the *Va = 0 line intersect indicates the maximum allowable substrate doping level, which is about*
* *At higher values of NB junction tunneling will* occur. Therefore allowable values for *d fall below the dashed line and above the Va = 0 line.*
* Figure (b) shows the maximum electric field in the depletion layer versus *NB . Any* applied voltage greater than *Va = 0 will cause breakdown to occur at lower values of NB .*In the foregoing discussions, the effects of *ND have been assumed to be negligible.*



**Limits of Miniaturization**

* The minimum size of a transistor is determined by both process technology and the physics of the device itself. The reduction of device geometry currently depends mainly on alignment
* accuracy and on the resolution of photolithographic technology; the limit on feature size is now at 0.3 µm, but the increasing availability of direct write E-beam technology will allow this limit to be further reduced
* The size of a transistor is usually defined in terms of its channel length *L. As the* channel length is scaled down, the edge of the depletion region around the source comes closer to that around the drain. In order to prevent punch-through and maintain transistor action, it can be shown that the channel length *L must be at least 2d. Therefore, L is in turn* determined by the substrate concentration *NB and supply voltage VDD (which determines Va).*
* Applying the conclusions from the previous section, we may estimate the minimum possible channel length as 0.14 µm. The minimum transit time for an electron to travel from source to drain can also be calculated



* The maximum carrier drift velocity is approximately equal to *Vsat where saturation velocity (Sze, 1985), regardless of the supply voltage. Therefore the minimum* transit time may be assumed to occur for a minimum size transistor when *Va is approximately* 0 V. Transit times may be assessed. Transit times may be assessed from Figure 5.3(a) that assumes a transistor of size *L = 2d with zero space between source and drain depletion regions.*

**Transit time t versus *L.***