CHAPTER 2

LITERATURE REVIEW

2.1 Transistor

Transistor is constructed with three doped semiconductor region separated by two pn junctions

2.1.1 N Metal Oxide Semiconductor (NMOS)

The NMOS is significantly faster than PMOS because electrons have two to three times' higher mobility than holes.

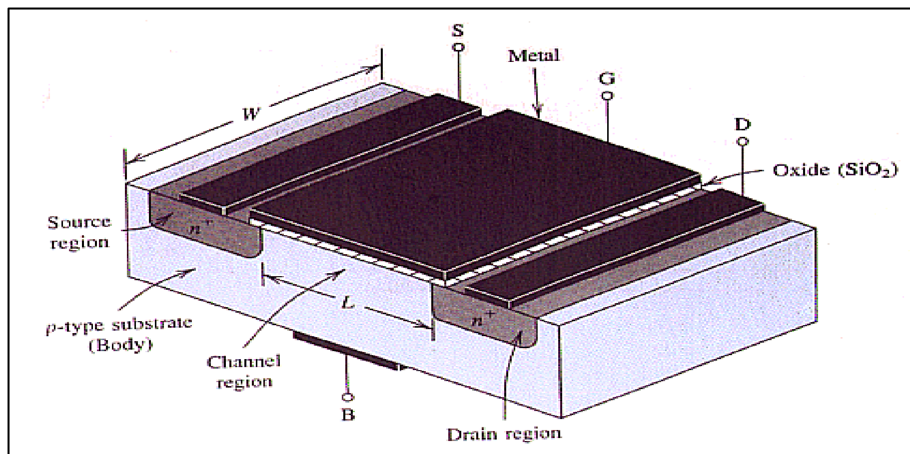

Figure 2.1: Cross Section of a NMOS [3].

Since, a polysilicon gate is often used instead of a metal gate. This is to reduce the threshold voltage value. A p+ implant is done in regions outside the transistor as shown in figure. 2.1. This serves to isolate devices.

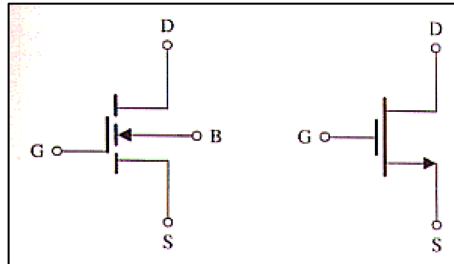

Figure 2.2: Symbol of NMOS [3].

From figure 2.2 shows the body and source contacts are tied together. This ensures that there is no current in the source-substrate diode. Since the drain is at a positive bias with respect to the source, the drain-substrate diode is also cut off. This case of isolation of the devices is a key reason why very high-density MOS circuitry can be produce [2].

2.1.2 P Metal Oxide Semiconductor (PMOS)

A p-channel MOSFET (PMOS) transistor is fabricated on an n-type substrate with p+ regions for the drain and source, and holes as charge carriers.

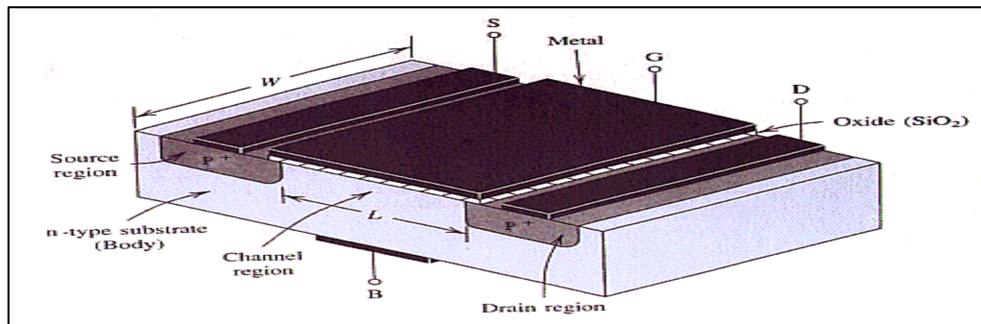

Figure 2.3: Cross Section of a PMOS [3].

The n-well acts as the body or substrate for the PMOS. In addition to creating the n-well, one needs to do a p+ implant for the source and drain of the transistor. This allows similar current values in the NMOS and PMOS devices.

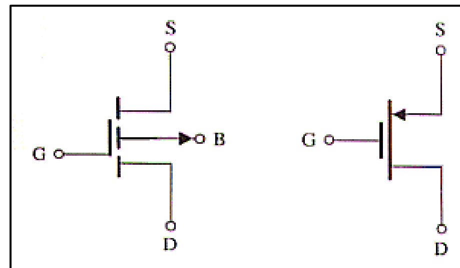

Figure 2.4: Symbol of PMOS [3].

2.2 Basic Operation of VCO

Actually, Voltage controlled oscillators (VCO) are essential building blocks of modern communication system. The VCO performance in terms of phase noise, tuning range and power dissipation determine many of the basic performance characteristic of a transceiver. An ideal voltage-controlled oscillator is a circuit whose output frequency is a linear function of its control voltage. In practice, as the oscillation amplitude increase, the stages in the signal path experience nonlinearity and eventually "saturation" limiting the maximum amplitude. The poles begin in the right half plane and eventually move to the imaginary axis to stop the growth. If the small-signal loop gain is greater than unity, the circuit speeds enough time in saturation so that the average loop gain is still equal to unity.

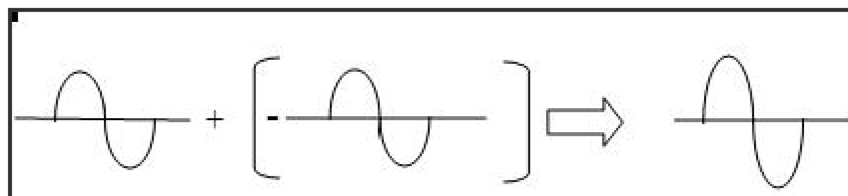

Figure 2.5: Double Amplitude [1]

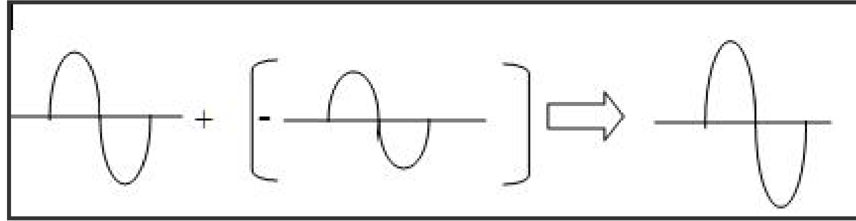

Figure 2.6: Triple Amplitude [1]

Thus:- $Vx = Vo + [\ H(jwo)]\ Vo + [H(jwo)]^2 Vo + [H(jwo)]^3 Vo + \ldots$

When $[H(jwo)] > 1 \longrightarrow Vx$ will diverge (grow) $= H(jwo) = -1$

$Vout/Vin = 1/1\text{-}1 = \infty$

2.2 Technique for design VCO

There are two techniques to design VCO using CMOS technology, firstly cascade technique and the other way is differential pair with MOS loads technique.

2.3.1 Cascade Technique

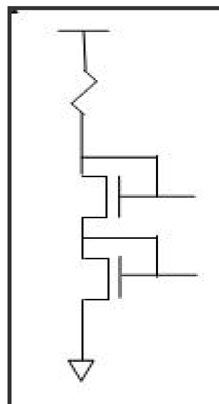
Figure 2.7: Cascade Schematic [1]

The cascade technique has two distinct advantages over the inverting. First, it provides higher output impedance similar to the cascode current sink and the cascode current mirror. Second, it reduces the effect of Miller capacitance on the input VCO which will be very important in designing the frequency behaviors of VCO. One of the most important aspects of the cascade amplifier has not yet been examined. This is because, up to this point, we assumed that the cascade VCO was driven by a low resistance source such as a voltage source. For the cascade, the design parameters are $(W/L)_1$, $(W/L)_2$, $(W/L)_3$ the DC current and the bias voltage. Typical specifications for the cascade might be V_{DD}, small-signal gain (Av), maximum and minimum output voltage swing and power dissipation.

2.3.2 Differential Pair with MOS Loads Technique

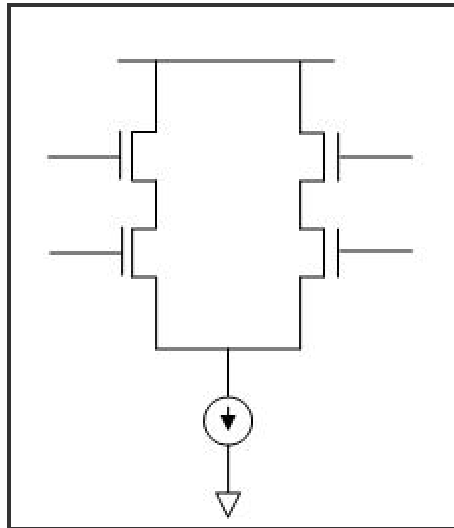

Figure 2.8: Differential Pair Schematic [1]

A differential signal is defined as are that measured between two modes that equal and opposite signal excursion around a fixed potential.

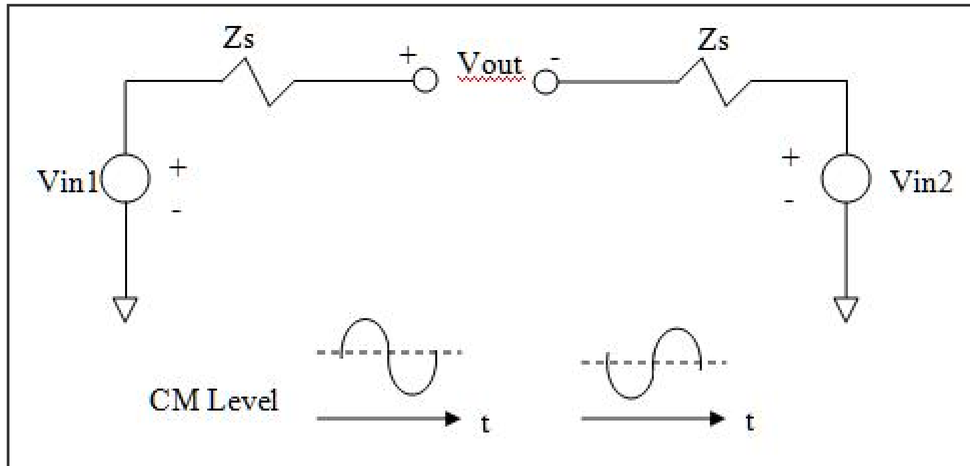

Figure 2.9: Differential pair Operational Circuit [1]

The “center” potential in differential signaling has called the “common-mode” (cm) level.

2.3.2.1 Advantages of differential pair.

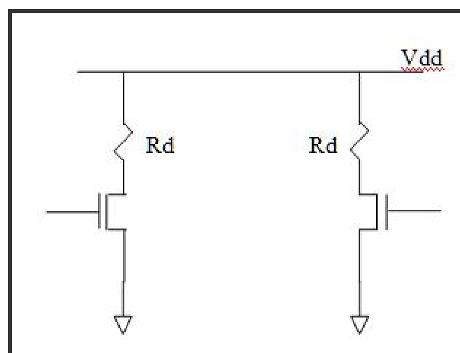

Figure 2.10: Differential pair circuit [1]

The advantages of differential pair are its higher immunity to "environmental" noise. Differential signaling is increase in maximum achievablevoltage swings also its advantages. Lastly, its simpler biasing and higher linearity.

2.4 Basic Differential Pair

Differential signaling is high rejection of supply noise higher output swings. The input cm level is excessively low, the minimum values of Vin1 and Vin2 may in fact turn off M1 and M2, leading to severe clipping at the output. Thus, it is important that the bias currents of the devices have minimal dependence on the input cm level.

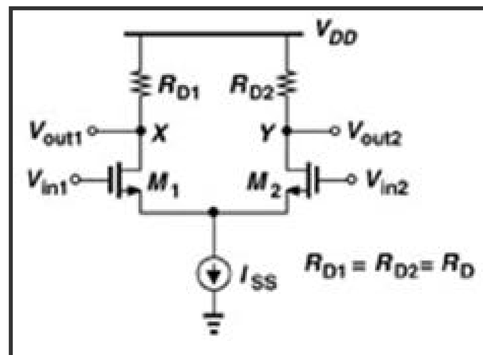

Figure 2.11 : Differential pair circuit lables [1]

Differential (Figure 2.10) emplace a current source Iss to make Id1 + Id2 independent of V_{IN} Thus, if V_{IN1}=V_{IN2}, the bias current of each transistor equals Iss/2 and the output common-mode level is Vdd-RdIss/2.

The higher the input cm level, the smaller the allowable output swings. It is desirable to choose a relatively low V_{IN}, but the preceding stage may not provide such a level easily.

2.5 Common-mode Stage with Triode Load

A MOS device operating in deep triode region behaves as a resistor and can therefore serve as the load in CS stage.

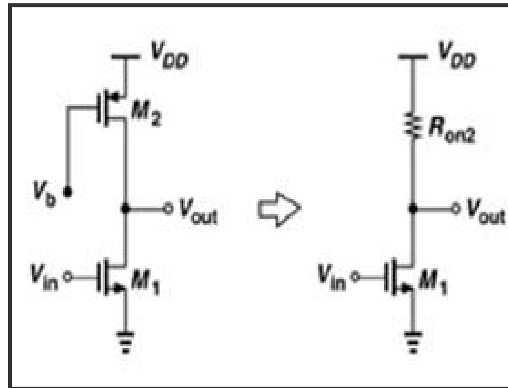

Figure 2.12: Common mode circuit [1]

In figure 2.11, such a circuit biases the gate of M2 at a sufficiently low level, ensuring the load is in deep triode region for all output voltage swings. Since, R_{ON2} can be calculated by :

$$R_{ON2} = \frac{1}{\mu_p C_{ox}\left(\frac{W}{L}\right)_2 (V_{DD} - V_b + V_{TH2})} \tag{2.0}$$

The voltage gain can be readily calculated. The principal drawback of this circuit start from the dependence of Ron upon µpCox, Vb and V_{HTP}. Since µpCox, and V_{HTP} vary with process and temperature and since generating a precise value for Vb requires additional complexity, this circuit is difficult to use. Triode loads, however consume less voltage headroom than the diode-connected devices because Vout,max = Vdd; where :

$$\text{Vout,max} \approx |\,\text{Vdd}\,| - V_{HT} \tag{2.1}$$

2.6 Low Voltage Analog

Low voltage analog are distributed into two part. Low analog voltage circuit and low analog voltage design.

2.6.1 Low Voltage Analog Circuit

In the past, low power consumption usually was less critically considered among key design specifications. But today, both increased circuit density of current fine line CMOS technology and battery-operated portable equipment necessitate low voltage low power system design [13]. For CMOS analog circuits, when the transistors operate in weak inversion region, $_{gm}/I_D$ reaches the maximum; hence the minimum power consumption can be achieved due to the small quiescent current at the expense of large silicon area and slow speed. When MOS transistors operate in strong inversion, however, although good frequency response and small area are obtained, non-optimum larger power is consumed, and V_{DS} (sat) is high. For most analog circuits, the best tradeoff among area, power and speed can be achieved when the transistors work in moderate inversion region.

2.6.2 Low Voltage Analog Design

Most designers often assume conservative ways to make the MOS transistors work in strong inversion, with power consumption and speed higher than needed, thus avoiding an optimal design. In recent years, some attempts of MOS modeling have been made to have one-equation model for all the operation regions.

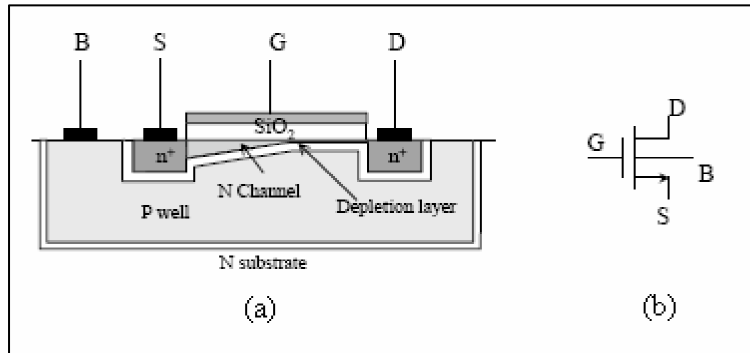

Figure 2.13: Bulk-driven MOS transistor, (a) cross section and (b) symbol of an N-channel MOSFET in P-well technology [11].

The cross section of an N-channel MOSFET structure is illustrated in Figure 2.11 (a). Unfortunately, there are also some disadvantages, as the transconductance of a bulk-driven MOSFET is substantially smaller than that of a conventional gate driven MOS transistor, which may result in lower GBW and worse frequency response. The other disadvantages are ; the polarity of the bulk-driven MOSFETs is technology related. For a P (N) well CMOS process, only N (P) channels bulk-driven MOSFETs are available. This may limit its applications; the equivalent input referred noise of a bulk-driven MOS amplifier is larger than a conventional gate driven MOS amplifier because of its smaller transconductance; and prone to turn on the parasitic bipolar transistors, which may result in a latch-up problem

2.7 Bulk Connections

Generally, two basics kinds of transistors are shown here. From review, the bulk connectionare discussed about how it is operate. This is most easily understood by understanding a cross-section of the water and transistor. Note that a layout designer can only understand this concept and cannot influence its design.

Nowadays raw silicon wafers were made of are p-type, so of the two transistors types, an NMOS transistors is the easier to design. The transistor layout is simply implemented in the bare substrate as shown figure 2.13 (a). To generated PMOS

transistors one need to create a separate bulk node and therefore need another layer. This is typically called an N-well; when implemented, it forms an island of N-type substrate. Implementing P-typed active regions within this N-well creates a PMOS transistor with a bulk connection as defined by the N-well as shown in figure 2.13 (b).

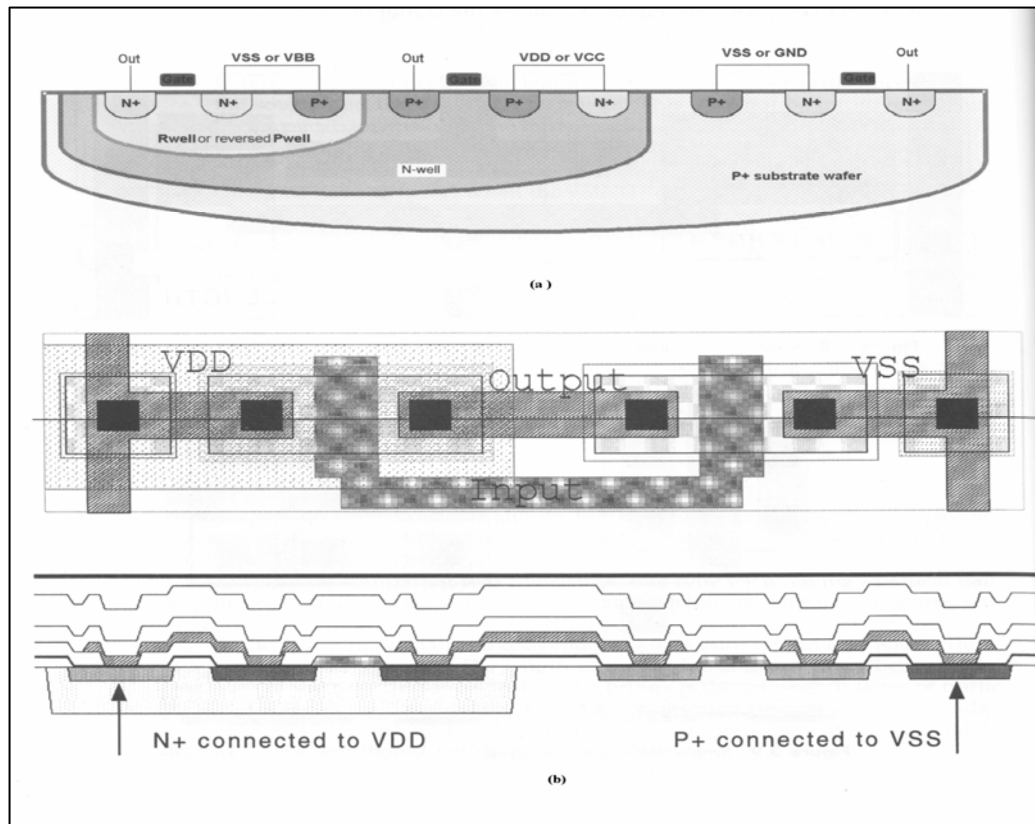

Figure 2.14: (a) wafer cross-section showing bulk connection and (b) cross-section inverter bulk connection [10].

Fgure 2.13 (a) also shows a NMOS transistor design that has a different bulk node than that of the substrate. A retrograde well (R-well) or P-type well (P-well) has been implemented in the N-well. This region creates a separate P-type bulk node for the transistors implemented within this region. This is an example of substrate connection in a DRAM process. In the case of an N-type wafer, the polarities of the transistor connections are simply the reverse of those shown previously. Figure 2.13 (b) shows substrate connections for an inverter in an ASIC process [10].

2.8 Second-Order Effects

Analysis of the MOS structure has thus far entailed various simplifying assumptions, some of which are not valid in many analog circuits. This section will determine one of second-order effects that are essential in our subsequent circuit analyses. Second order effect has four phenomena that appear in submicron devices. The effects in second order effect are body effect, channel length modulation, subthreshold modulation and voltage limitation.So at this project design, the body effect phenomena are focused to reduce the problem.

2.8.1 Body Effect

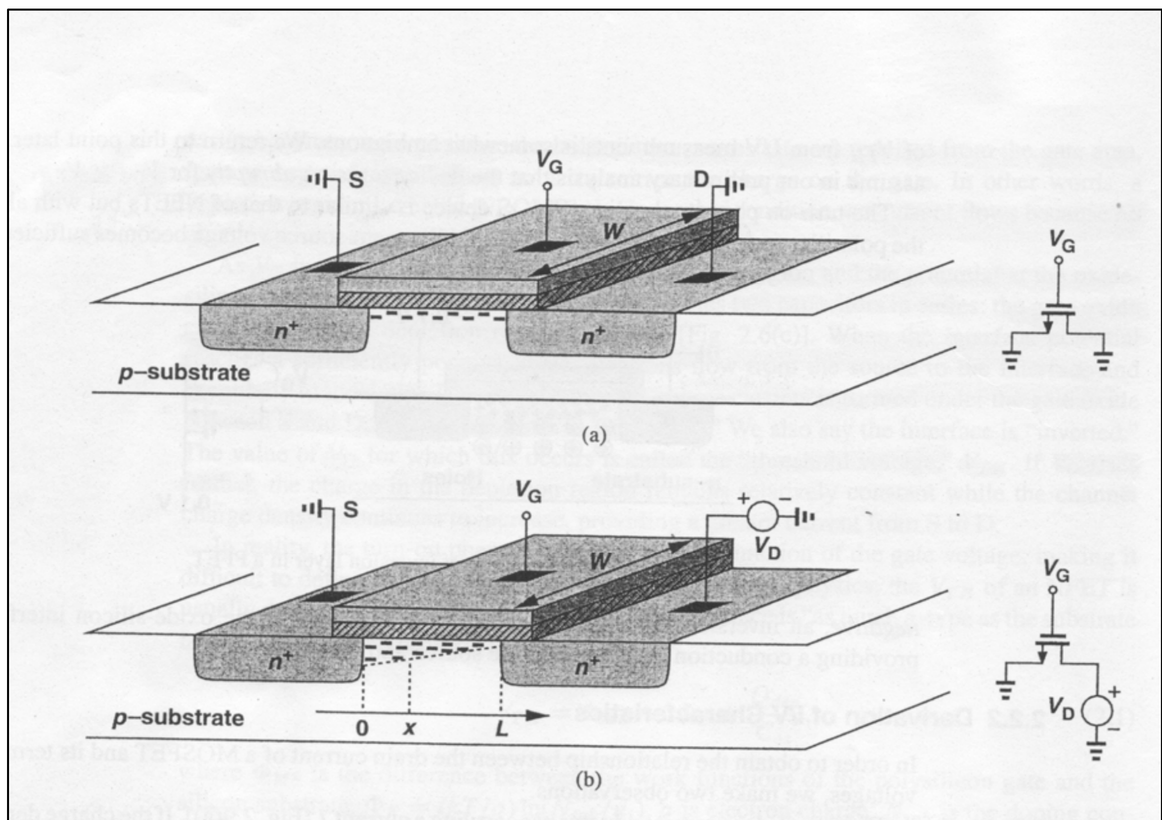

Figure 2.15: Channel charge with (a) equal source and drain voltages, (b) unequal source and drain voltage [4].

In the analysis of figure 2.14, assume that the bulk and the source of the transistor were tied to ground. Since the source and drain junctions remain reverse-biased, surmise that the device continues to operate properly but certain characteristics may change.

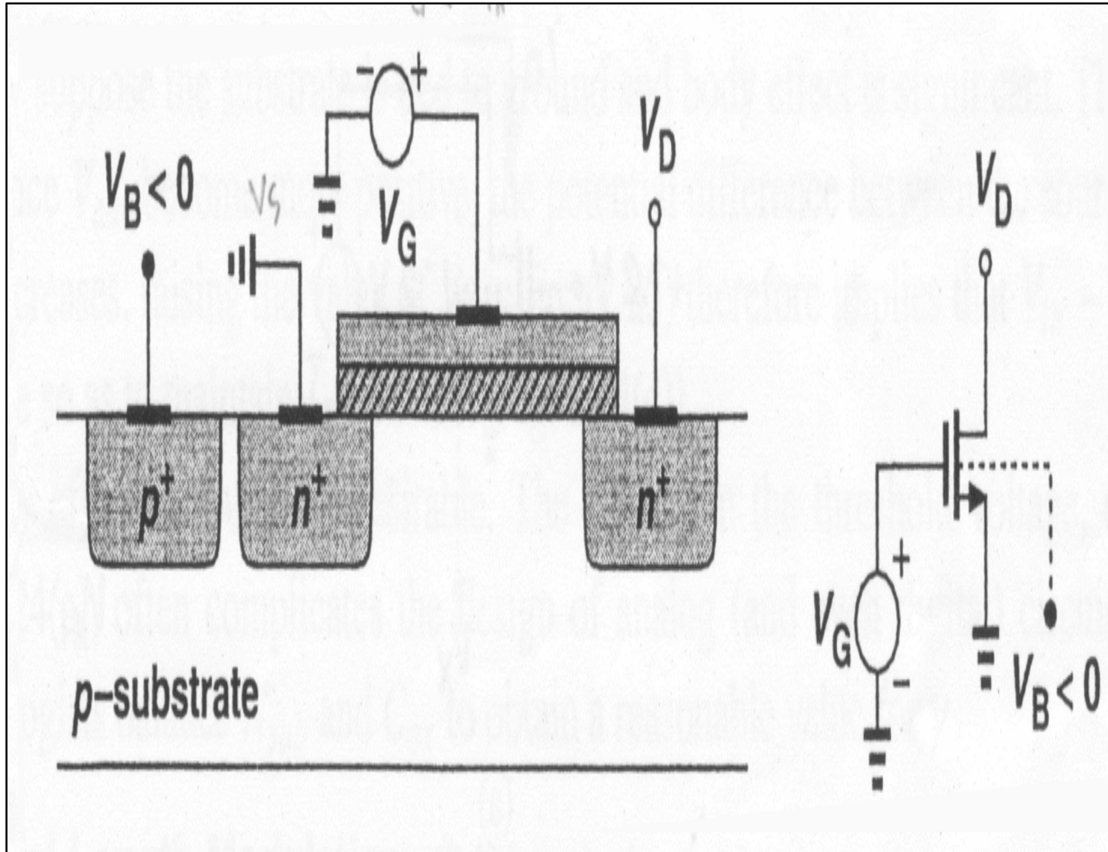

Figure 2.16: NMOS device with negative bulk voltage [4].

To understand the effect, suppose $V_S = V_D = 0$, and V_G is somewhat less than V_{TH} so that a depletion region is formed under the gate, but no inversion layer exist. As V_B becomes more negative, more holes are attracted to the substrate connection, leaving a larger negative charge behind, i.e., as depicted in figure 2.16, the depletion region becomes wider [1]

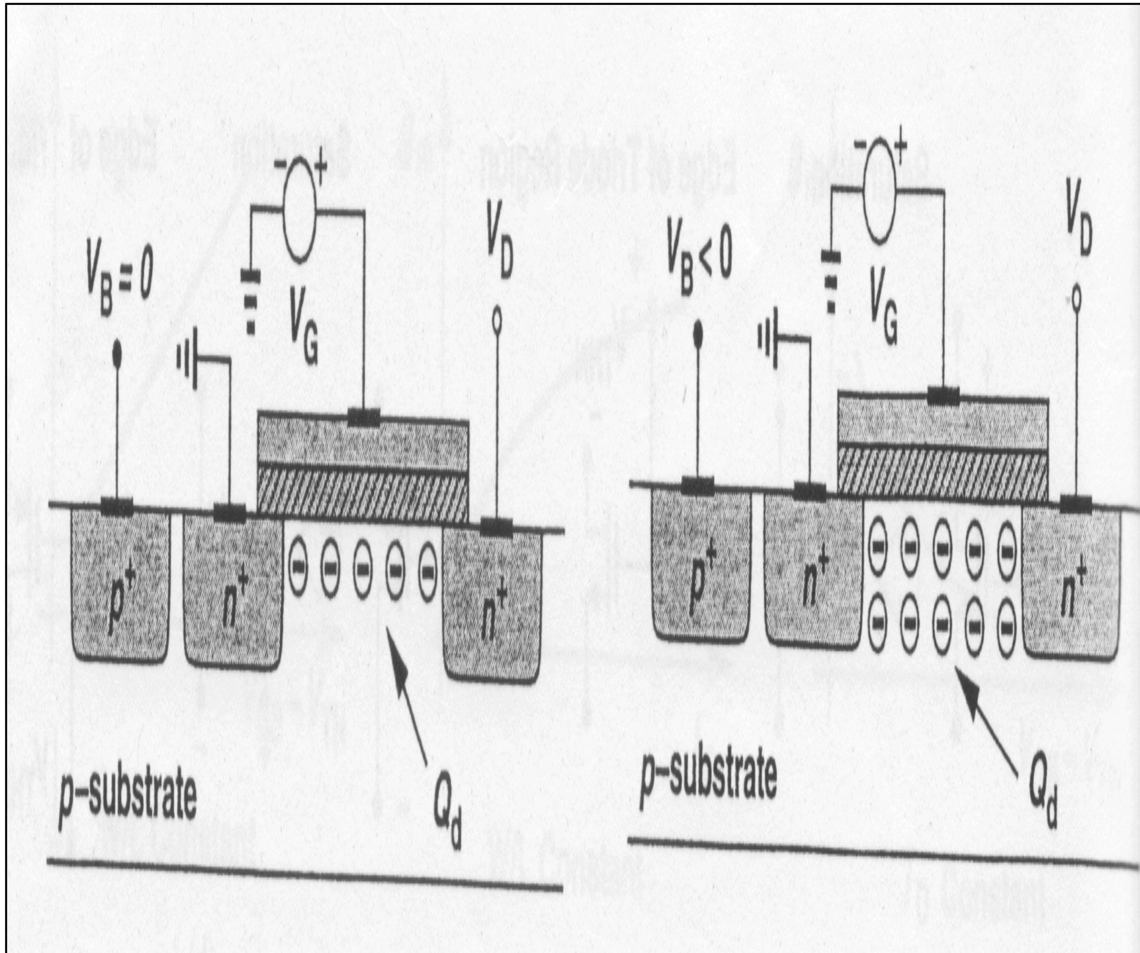

Figure 2.17: Variation of depletion region charge with bulk voltage [4].

$$V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}} \tag{2.2}$$

Now refer to Equation 2.2 that the threshold voltage is a function of the total charge in the depletion region because the gate charge must mirror *Qd* before an inversion layer is formed. Thus, as V_B drops and *Qd* increases, V_{TH} also increases. This is called the "body effect" or the "backgate effect."

It can be proved that with body effect:

$$V_{TH} = V_{TH0} + \gamma\left(\sqrt{|2\Phi_F + V_{SB}|} - \sqrt{|2\Phi_F|}\right) \quad (2.3)$$

Where V_{THO} is given by, $\gamma = \sqrt{2q\varepsilon_{si}N_{sub}} / C_{ox}$ denotes the body effect coefficient, and V_{SB} is the source-bulk potential different. The value of γ typically lies in the range of 0.3 to 0.4 V ½ [5].

2.9 Layout Design Rules

Layout rules also referred to as design rules or ground rules (IBM), and can considered a prescription for preparing the photomasks that are used in the fabrication of integrated circuits [8]. The main objective of the layout rules is to build reliably functional circuits in as small an area as possible. In general, design rules represent a compromise between performance and yield. Design rules specify to the designer certain geometric constrains on the layout artwork so that the patterns on the processed wafer will preserve the topology and geometry of the design.

2.10 Design Rule Background

Mentor graphics software has fixed the rule of design include well rule and transistor rule.

2.10.1 Well Rules

The n-well is usually a deeper implant than the transistor source/drain implants, and therefore it is necessary for the outside dimension to provide sufficient clearance between the n-well edges and the adjacent n+ diffusion. The inside clearance is

determined by the transition of the field oxide across the well boundary. Processes that use STI may permit zero inside clearance [9]. In order LOCOS process, problems such as the bird's beak effect usually prevent this. Because the n-well sheet resistance can be several KΩ per square, it is necessary to thoroughly ground the well. This will prevent excessive voltage drops due to well currents.

2.10.2 Transistor Rules

CMOS transistors are generally defined by at least four physical masks [8]. These are active, n-select, p-select and polysilicon. The active mask defines all areas where either n- or p-type diffusions is to be placed or where the gates of transistors are be placed. The gates of transistors are defined by the logical AND of the polysilicon mask and the active mask, where polysilicon crosses diffusion. The select layers define what type of diffusion is required. Frequently, design systems will only define n-diffusion and p-diffusion to reduce the complexity of the process. The appropriate selects are generated automatically. That is, n-diffusion will be converted automatically into active with an overlapping rectangle or polygon of n-select.

It is essential for the poly to be completely cross active; otherwise the transistor that has been created will be shorted by a diffusion path between source and drain [8]. Hence, poly is required to extend beyond the edges of the active area. This often termed the gate extension. Active must extend beyond the poly gate so that diffused source and drain regions exist to carry charge into and out of the channel. Poly and active regions that should not form a transistor must be kept separated; this results in a spacing rule from active to polysilicon. Figure 2.17(a) shows the mask construction for the final structures that appear in Figure 2.17(b).

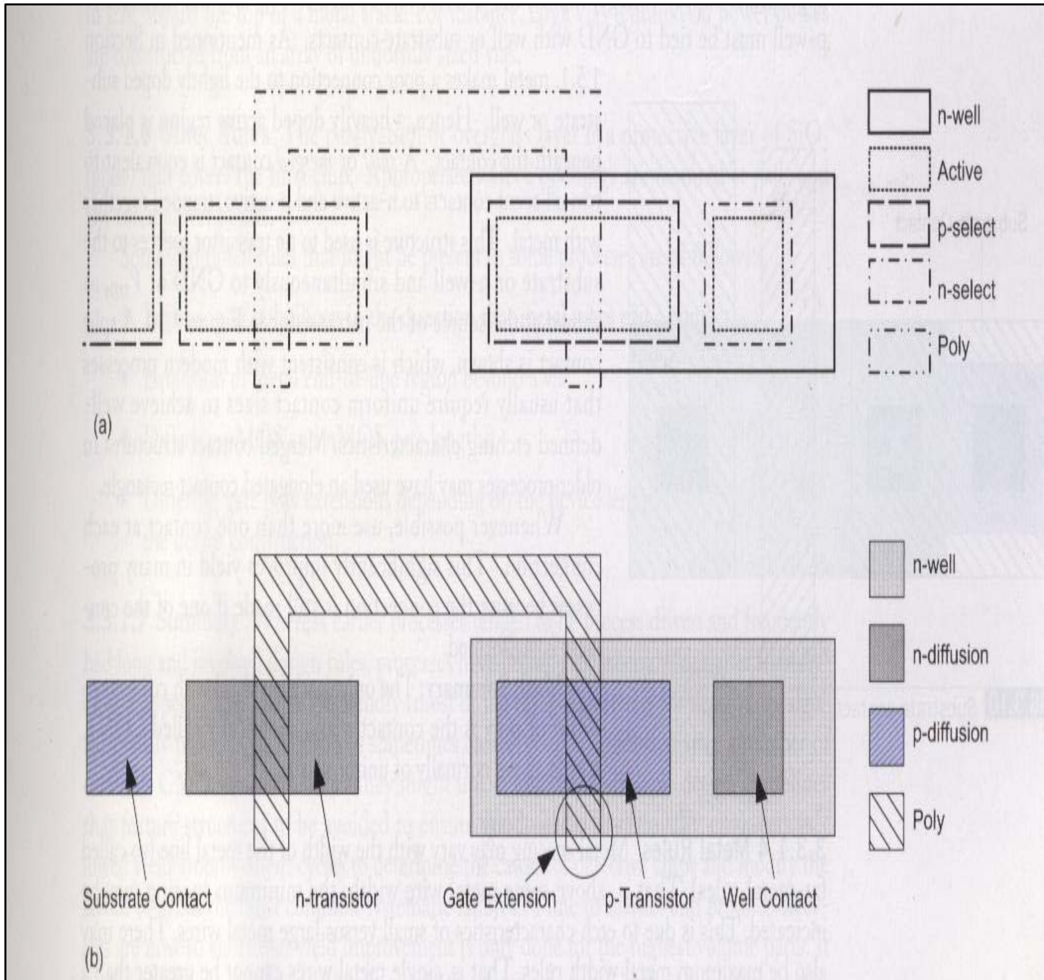

Figure 2.18: (a) and (b) show CMOS n-well process transistor and well/substrate contact construction [8].

2.10.3 Contact Rules

There are several generally available contacts as metal to p-active (p-diffusion), metal to n-active (n-diffusion), metal to polysilicon and metal to well or substrate.

Depending on the process, other contacts such as buried polysilicon-active contacts may be allowed for local interconnect. Because the substrate is divided into well regions, each isolated well must be tied to V_{DD} and the substrate or p-well must be tied to G_{ND} with well or substrate contacts.A split or merge contact is equivalent to two adjacent

contacts to n-active and p-active strapped together with metal[8]. This structure is used to tie transistor sources to the substrate or n-well and simultaneously to G_{ND} or V_{DD}, and is shown at the source of the n-transistor in Figure 2.18. A split contact is shown, which is consistent with modern processes that usually require uniform contact sizes to achieve well defined etching characteristics. Merged contact structures in older processes may have used an elongated contact rectangle.

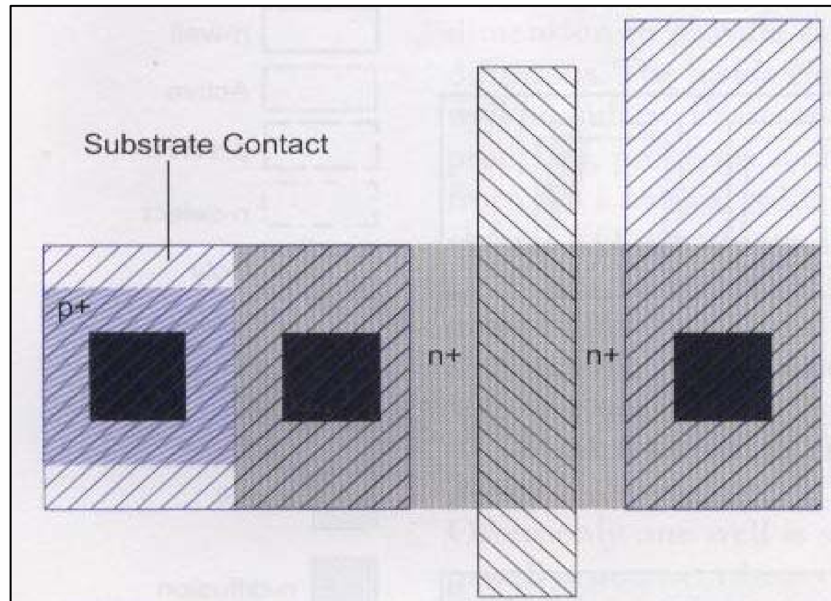

Figure 2.19: Substrate contact [8].

2.10.4 Metal Rules

Metal spacing may vary with the width of the metal line. That is above some metal wire width the minimum spacing may be increased. This is due to etch characteristics of small versus large metal wires. There may also be maximum metal width rules. That is single metal wire cannot be greater than a certain width. If wider wires are desired, they are constructed by parelling a number of smaller wires and adding checkboard links to tie the wires together. Additionally, there may be spacing rules that are applied to long, closely spaced parallel metal lines. Older nonplanarized processes

required greater width and spacing on upper-level metal wires to prevent breaks or shorts between adjoining wires caused by the vertical topology of the underlying layers. This no longer a consideration for thicker metal layers.

2.10.5 Via Rules

Processes may vary in whether they allow stacked vias to be placed over polysilicon and diffusion regions. Some processes allow vias to be placed within these areas, but do not allow the vias to straddle the boundary of polysilicon or diffusion. This results from the sudden vertical topology variations that occur at sub-layer boundaries. Most modern planarized processes allow for stacked vias, which reduces the area required to pass from a lower-level metal to a high-level metal.

2.10.6 Other Rules

The passivation or overglass layer is a protective layer of SiO_2 that covers the final chip. Appropriately sized openings are required at pads and any internal test points. Some additional rules that might be present in some processes are as extensions of polysilicon in the direction that metal wires exit a contact, extensions of metal end-of-line region beyond a via, differing pMOS and nMOS gate lengths; and differing gate poly extensions depending on the device length or the device construction.

All the MOSIS design rules are shown in appendix A.