1 MOSFET Review

Device Structure

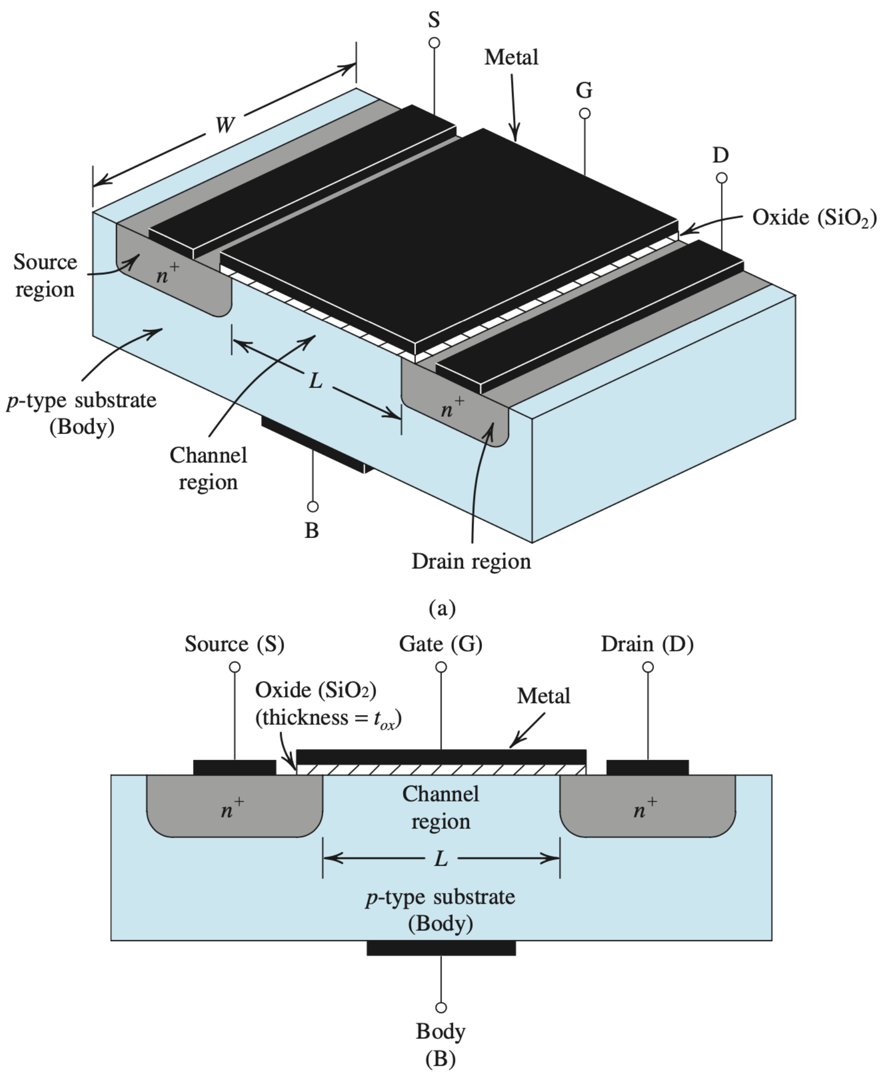


Figure 1.1 Enhancement type NMOS transistor.
The MOSFET is essentially a three terminal device. The voltage applied to the gate controlls current flow between the source and the drain. This current will flow in the longitudinal direction from drain to source in the region labeled “channel region.”

Operation with Zero Gate Voltage

With zero voltage applied to the gate, two back-to-back diodes exist in series between drain and source. These back-to-back diodes prevent current conduction from drain to source when a voltage v_{DS} is applied. In fact, the path between drain and source has a very high resistance (of the order of $10^{12}\Omega$).

Creating a Channel for Current Flow

Here, both the drain and source are grounded and a positive voltage is applied to the gate. The holes are pushed down into the substrate. The posistive gate voltage also attracts electrons from both n regions into the channel region. If a voltage is to be applied between the drain and source, current will flow through the n region.

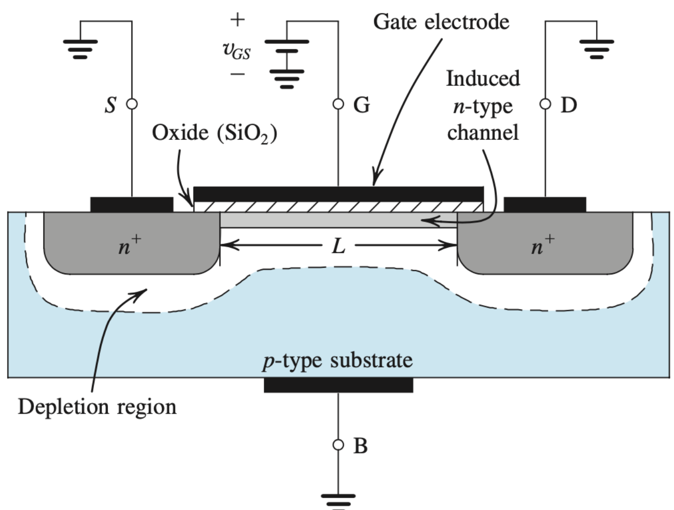


Figure 1.2 A positive voltage applied to the gate causes an abundance of electrons to form near the surface of the substrate under the gate, creating an n region.

The value of v_{GS} at which a sufficient number of mobile electrons accumulate in the channel region to form a conducting channel is called the threshold voltage and is denoted V_t
When $v_{DS} = 0$, the voltage at every point along the channel is zero, and the voltage across the oxide (i.e., between the gate and the points along the channel) is uniform and equal to v_{GS}. The excess of v_{GS} over V_t is termed the ***effective voltage*** or the ***overdrive voltage*** and is the quantity that determines the charge in the channel.

$$v_{OV} = v_{GS} - V_t \tag{1}$$

We can express the magnitude of the electron charge in the channel by:

$$|Q| = C_{ox}(WL)v_{OV} \tag{2}$$

where C_{ox}, called the ***oxide capacitance***, is the capacitance of the parallel-plate capacitor per unit gate area (in units of F$/m^2$), W is the width of the channel, and L is the length of the channel. The oxide capacitance C_{ox} is given by

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \tag{3}$$

Finally note that as v_{OV} is increased, the magnitude of the channel charge is increased proportionately. Meaning, the larger the overdrive voltage, the deeper the channel.

Applying a Small v_{DS}

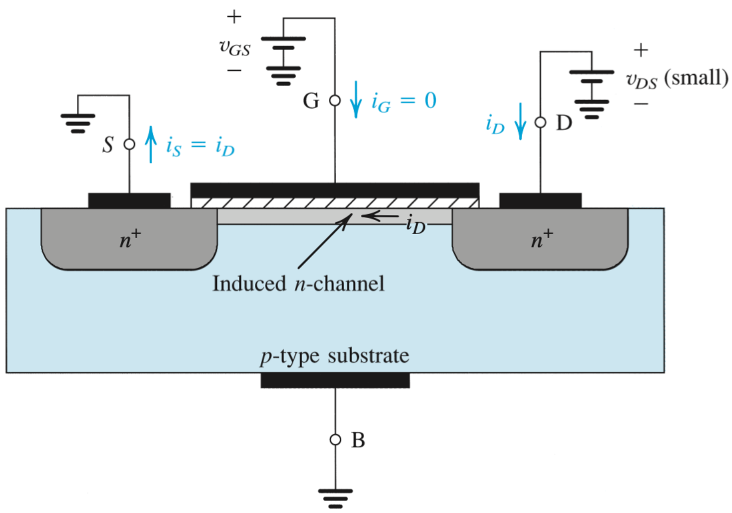


Figure 1.3 NMOS with $v_{GS} > V_t$ and a small v_{DS} applied. Of particular interest of calculating the current i_D is the charge per unit channel length, which can be found as

$$\frac{|Q|}{unitChannelLenth} = C_{ox} W v_{OV} \tag{4}$$

The voltage v_{DS} establishes an electric field E across the length of the channel,

$$|E| = \frac{v_{DS}}{L} \tag{5}$$

This electric field in turn causes the channel electrons to drift toward the drain with a velocity given by

$$\mu_n |E| = \mu_n \frac{v_{DS}}{L} \tag{6}$$

where μ_n is the mobility of the electrons at the surface of the channel. The value of i_D can now be found by multiplying the charge per unit channel length by the electron drift velocity.

$$i_D = \left[(\mu_n C_{ox}) \left(\frac{W}{L} \right) v_{OV} \right] v_{DS} \tag{7}$$

Thus, for small v_{DS}, the channel behaves as a linear resistance whose value is controlled by the overdrive voltage v_{OV} , which in turn is determined by v_{GS}

$$i_D = \left[(\mu_n C_{ox}) \left(\frac{W}{L} \right) (v_{GS} - V_t) \right] v_{DS} \tag{8}$$

The conductance g_{DS} of the channel can be found by

$$g_{DS} = (\mu_n C_{ox}) \left(\frac{W}{L} \right) v_{OV} \tag{9}$$

Note that the ***process transconductance*** parameter is given the symbol $k'_n(A/V^2)$, where n denotes n-channel

$$k'_n = \mu_n C_{ox} \tag{10}$$

The product of the process transconductance parameter k'_n and the transistor aspect ratio (W/L) is the MOSFET ***transconductance parameter*** k_n,

$$k_n = k'_n(W/L)$$

We conclude this subsection by noting that with v_{DS} kept small,the MOSFET behaves as a linear resistance r_{DS} whose value is controlled by the gate voltage v_{GS},

$$r_D = 1/g_{DS}$$

Operation as v_{DS} is Increased

We next consider a situation where v_{DS} is increased. v_{GS} is held constant at a value greater than V_t; that is, the MOSFET will be operated at a constant overdrive voltage V_{OV}. As we travel from source to drain, the voltage (measured relative to the source) increases from zero to v_{DS}. Thus the voltage between the gate and points along the channel decreases from $v_{GS} = V_t + V_{OV}$ at the source end to $v_{GD} = v_{GS} - v_{DS} = V_t + V_{OV} - v_{DS}$ at the drain end. Since the channel depth depends on the amount by which the voltage exceeds V_t, we find that the channel is no longer uniform; being deepest at the source end (where depth is proportional to V_{OV}) and shallowest at the drain end (where depth is proportional to $V_{OV} - v_{DS}$).

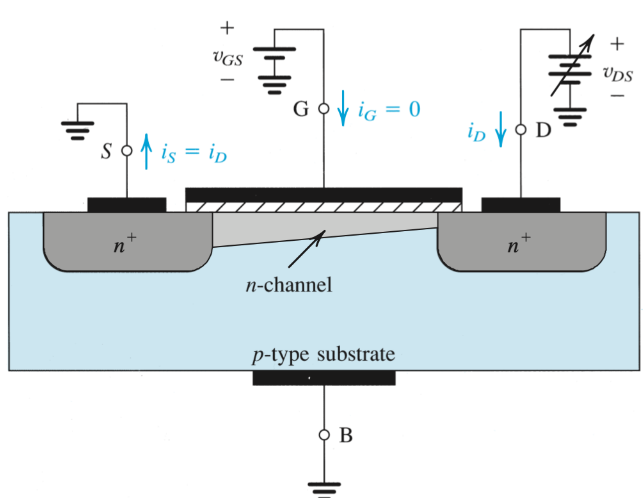


Figure 1.4

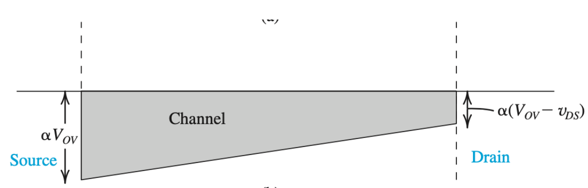


Figure 1.5

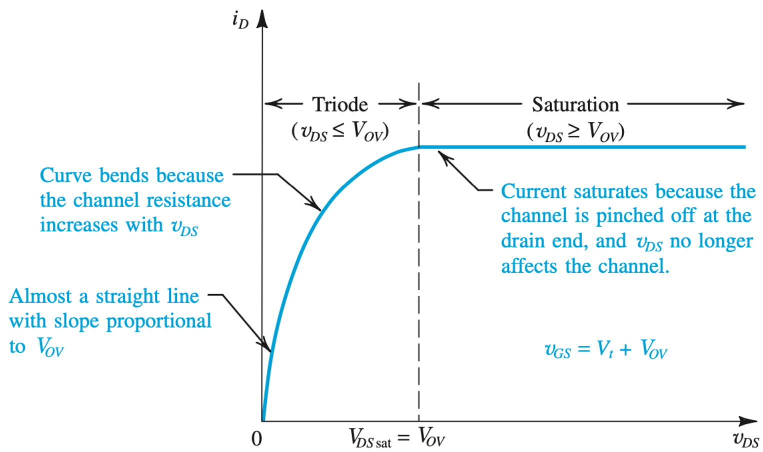


Figure 1.6

As v_{DS} is increased, the channel becomes more tapered and its resistance increases. Therefore the $i_D - v_{DS}$ curve does not continue in a straight line, but bends as shown in **Figure 1.6**. Since the charge in the tapered channel is proportional to the channel cross sectional shown in **Figure 1.5**, the relationship between i_D and v_{DS} an be found as follows

$$i_D = k_n' \left(\frac{W}{L}\right) \left(V_{OV} - \frac{1}{2} v_{DS}\right) v_{DS} \tag{11}$$

This describes the ***semiparabolic portion*** of the $i_D - v_{DS}$ curve. Notes that as v_{DS} is reduced, we can neglect $\frac{1}{2}v_{DS}$ relative to V_{OV} and can be written as the equation shown in **(7)**.

Operation for $v_{DS} \geq V_{OV}$ Channel Pinch-off and Current Saturation

The above description of operation assumed that even though the channel became tapered, it still had a finite (nonzero) depth at the drain end. This in turn is achieved by keeping v_{DS} sufficiently small that the voltage between the gate and the drain, v_{GD}, exceeds V_t. **Figure 1.7** shows v_{DS} reaching V_{OV} and $v - GD$ corrispondingly reaching V_t. The zero depth of the channel gives rise to the term ***channel pinch-off***. Increasing v_{DS} beyond this point ($i.e., v_{DS} > V_{OV}$), has no effect on the channel shape and charge.

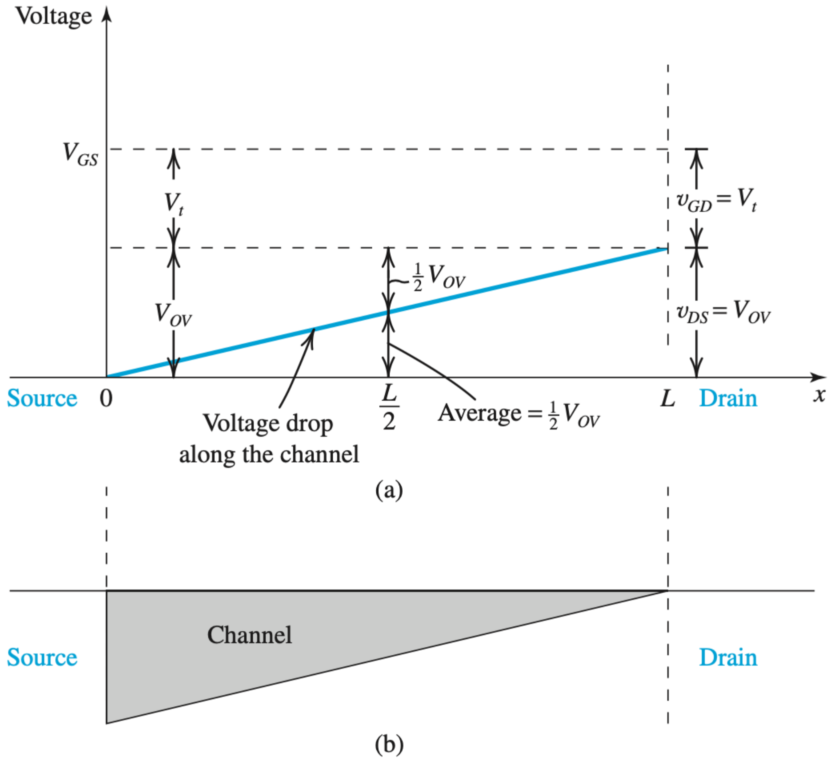


Figure 1.7

The drain current thus ***saturates*** at the value found by substituting $v_{DS} = V_{OV}$

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)V_{OV}^2 \tag{12}$$

*V_{OV} can be replaced by v_{OV} and v_{OV} by $(v_{GS} - V_t)$

The MOSFET is then said to have entered the ***saturation region***. The voltage v_{DS} at which saturation occurs is denoted V_{DSsat}

$$V_{DSsat} = V_{OV} = V_{GS} - V_t \tag{13}$$

The channel pinch off does *not* mean channel blockage. Current continues to flow through the pinched-off channel and the electrons that reach the drain end are accelerated through depletion region. Any increase in v_{DS} above V_{DSsat} appears as a voltage drop across the depletion region.

P-Channel MOSFET

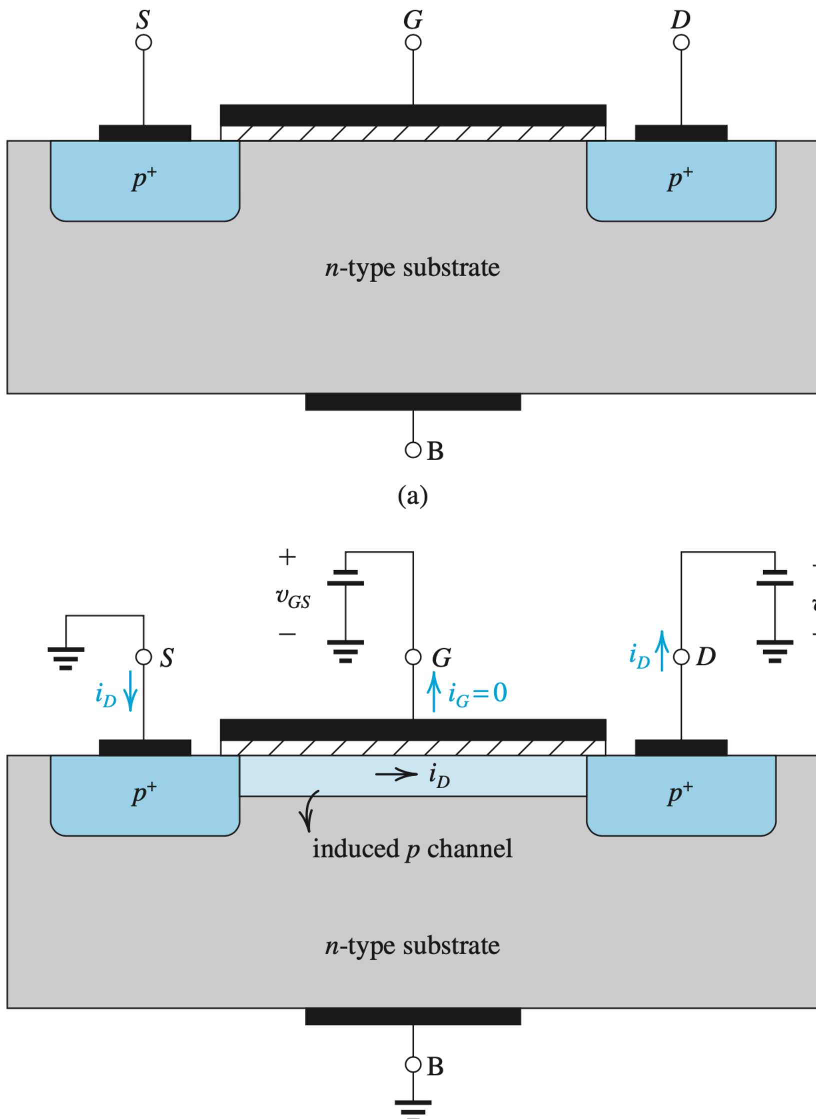


Figure 1.8 P-Channel MOSFET.

To induce current flow between source and drain, a negative voltage is applied to the gate. Once the magnitude of the negative v_{GS} is beyond that of the threshold voltage V_{tp}, the p-channel is established.

$$v_{GS} \leq V_{tp}$$

or

$$|v_{GS}| > |V_{tp}|$$

To cause a current i_D to flow through the channel, a negative voltage v_{DS} is applied to the drain. The current i_D is carried by holes and flows through the channel from source to drain.

Finite Output Resistance in Saturation

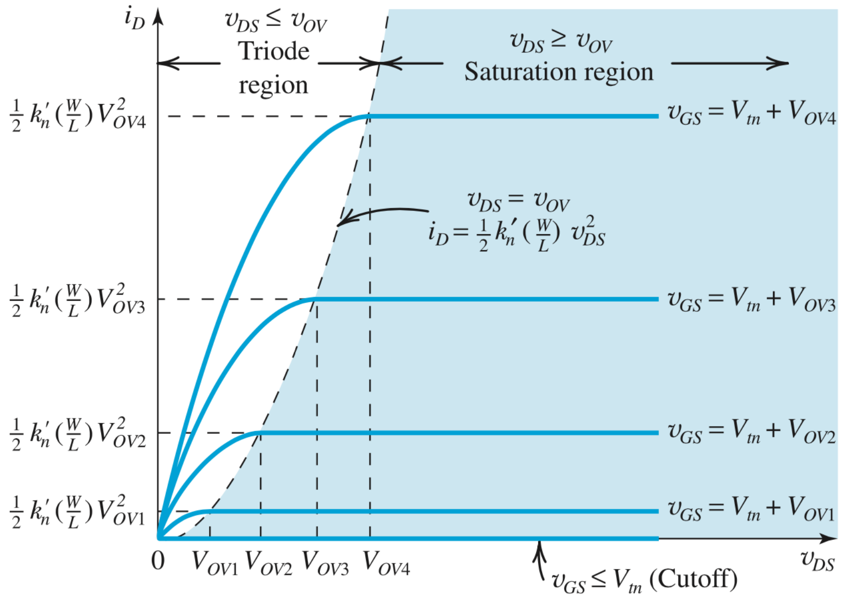


Figure 1.9 $i_D - v_{DS}$ characteristics of enhancement type NMOS transistor.

Figure 1.9 indicates that in saturation, i_D, is independent of v_{DS}, thus a change in v_{DS} in the drain to source voltage causes no change in i_D so the resistance looking into the drain of a saturated MOSFET is infinite. However, there is in fact an effect from increasing v_{DS} beyond $v_O V$ which is that the channel pinch off point moves slightly away from the drain and towards the source. The channel length is in effect reduced from L to $L - \Delta L$, a phenomenon known as **channel-length modulation**.

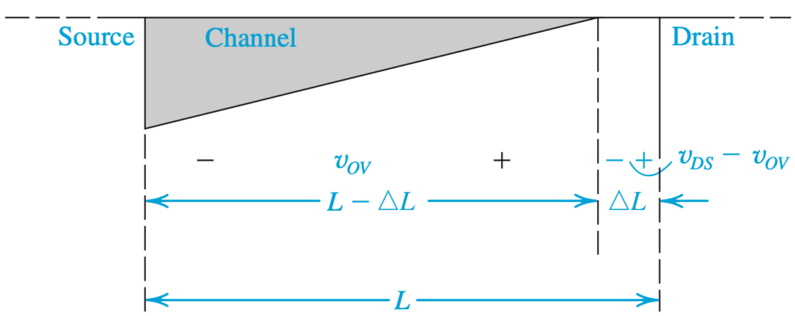
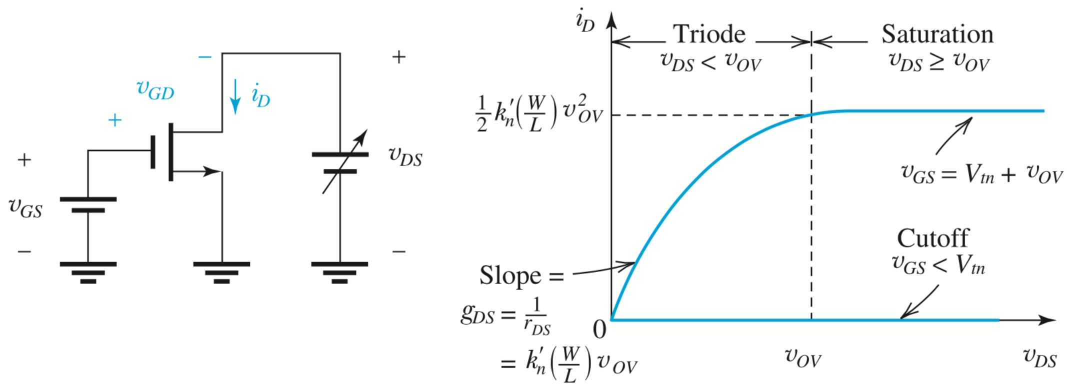


Figure 1.10

This effect can be accounted for by

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2(1 + \lambda v_{DS}) \tag{14}$$

Table 5.1 Regions of Operation of the Enhancement NMOS Transistor



- $v_{GS} < V_{tn}$: no channel; transistor in cutoff; $i_D = 0$
- $v_{GS} = V_{tn} + v_{OV}$: a channel is induced; transistor operates in the triode region or the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{GD} > V_{tn}$$

or equivalently:

$$v_{DS} < v_{OV}$$

Then,

$$i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_{tn})v_{DS} - \frac{1}{2}v_{DS}^2\right]$$

or equivalently,

$$i_D = k_n'\left(\frac{W}{L}\right)\left(v_{OV} - \frac{1}{2}v_{DS}\right)v_{DS}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{GD} \leq V_{tn}$$

or equivalently:

$$v_{DS} \geq v_{OV}$$

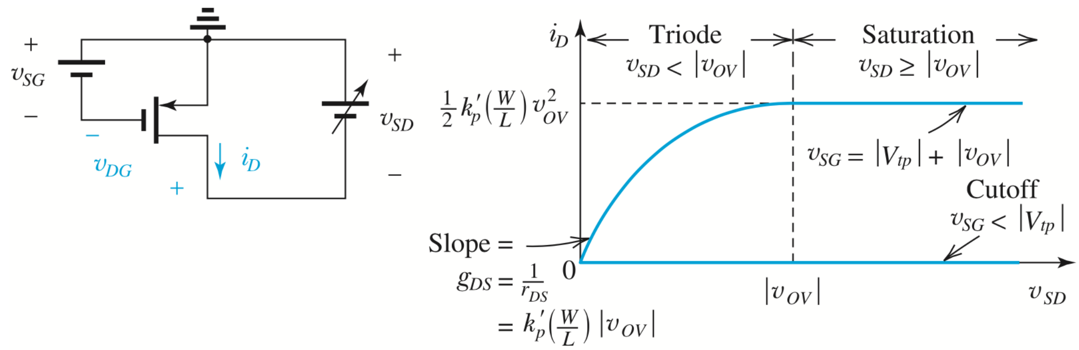
Then

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2$$

or equivalently,

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2$$

Table 5.2 Regions of Operation of the Enhancement PMOS Transistor



- $v_{SG} < |V_{tp}|$: no channel; transistor in cutoff; $i_D = 0$
- $v_{SG} = |V_{tp}| + |v_{OV}|$: a channel is induced; transistor operates in the triode region or in the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{DG} > |V_{tp}|$$

or equivalently

$$v_{SD} < |v_{OV}|$$

Then

$$i_D = k_p'\left(\frac{W}{L}\right)\left[(v_{SG} - |V_{tp}|)v_{SD} - \frac{1}{2}v_{SD}^2\right]$$

or equivalently

$$i_D = k_p'\left(\frac{W}{L}\right)\left(|v_{OV}| - \frac{1}{2}v_{SD}\right)v_{SD}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{DG} \leq |V_{tp}|$$

or equivalently

$$v_{SD} \geq |v_{OV}|$$

Then

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)(v_{SG} - |V_{tp}|)^2$$

or equivalently

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)v_{OV}^2$$

Short and Long Channel Model

2 CMOS Logic-Gate Circuits

In this section we consider the synthesis of CMOS circuits that realize combinational logic functions. In combinational circuits, the output at any time is a function only of the values of input signals at that time. Thus, these circuits do not have memory and do not employ feedback. Combinational circuits are used in large quantities in every digital system.

Switch Level transistor Model

CMOS digital circuits utilize NMOS and PMOS transistors operating as *switches*. An NMOS transistor behaves as a closed switch, exhibiting a very small resistance (R_{on} or r_{DS}) between its drain and source when its gate voltage is "high," usually at the power supply level V_{DD}, which represents a logic 1. Conversely, when the gate voltage is “low” (i.e., at or close to ground voltage), which represents a logic 0, the transistor is cut off, thus conducting zero current and acting as an open switch.

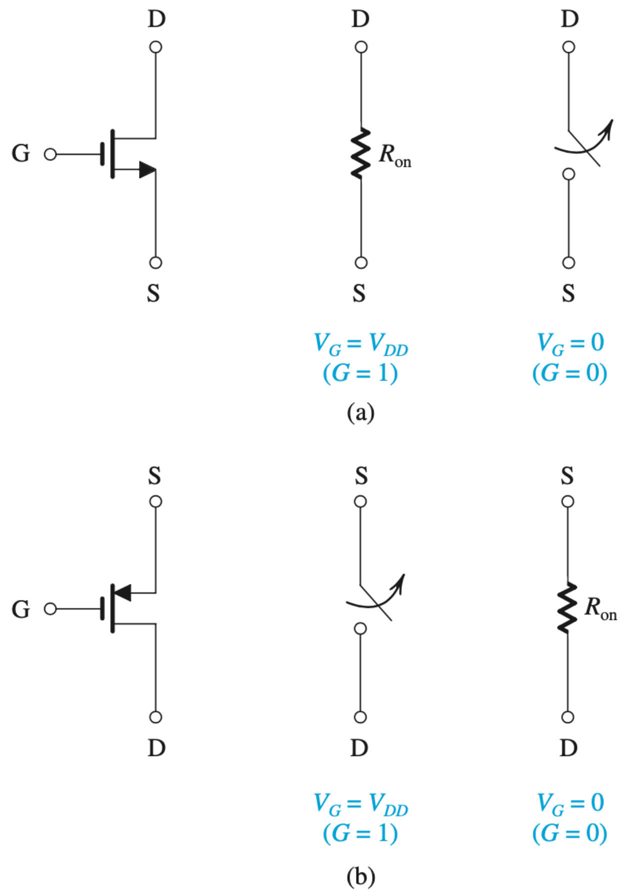


Figure 2.1 (a) NMOS and (b) PMOS transistors as a switch.

CMOS Inverter

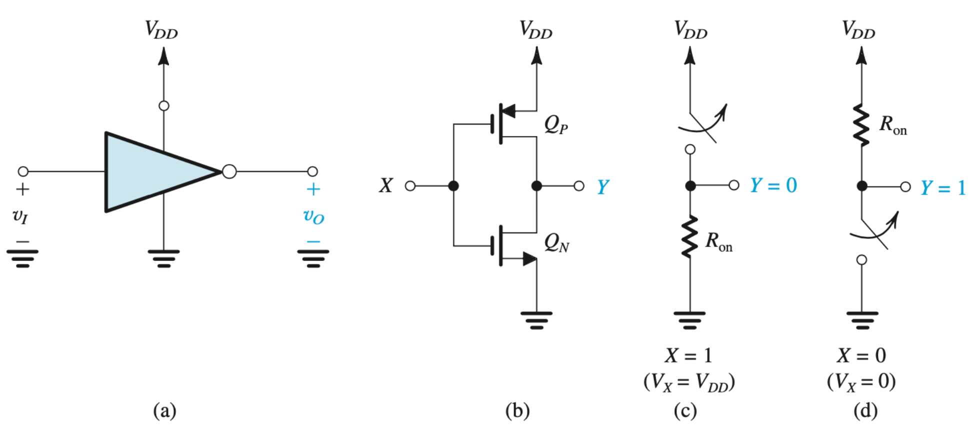


Figure 2.2 The inverted operated from a power supply V_{DD} is shown in (a). Its CMOS circuit implementation is shown in (b). (c) and (d) shows when the input $V_x = V_{DD}$ and $V_x = 0$ respectively.

Inverter Circuit Operation

We first consider the two extreme cases: when v_I is at logic-0 level, which is 0V, and when v_I is at logic-1 level, which is V_{DD} volts.

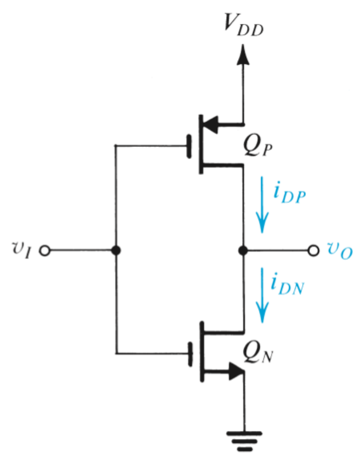


Figure 2.3 CMOS Inverter

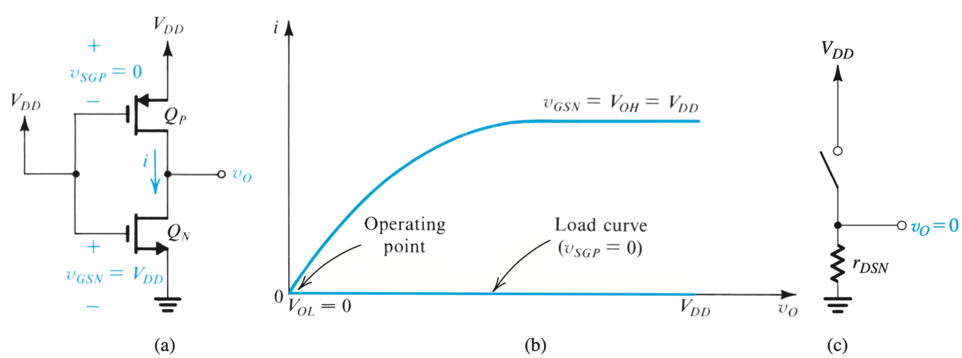


Figure 2.4 CMNOS Inverter when v_I is high.

Illustrated is the case when $v_I = V_{DD}$, showing the $i_D - v_{DS}$ characteristic curve for Q_N with $v_{GSN} = V_{DD}$. (Note that $i_D = i$ and $v_{DSN} = v_O$). Superimposed on the Q_N characteristic curve is the load curve, which is the $i_D - v_{SD}$ curve of Q_P for the case $v_{SGP} = 0V$. Since $v_{SGP} < |V_t|$, the load curve will be a horizontal straight line at zero current level. The operating point will be at the intersection of the two curves, where we note that the output voltage is zero and the current through the two devices is also zero. This means that the power dissipation in the circuit is zero. Note, however, that although Q_N is operating at zero current and zero drain-source voltage (i.e., at the origin of the $i_D - v_{DS}$ plane), the operating point is on a steep segment of the $i_D - v_{DS}$ characteristic curve. Thus Q_N provides a low-resistance path between the output terminal and ground, with the resistance obtained using

$$r_{DSN} = \frac{1}{[k_n'(\frac{W}{L})_n(V_{DD} - V_{tn})]} \tag{15}$$

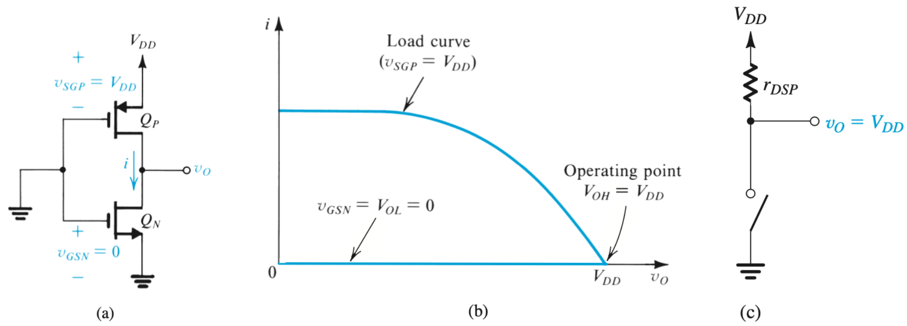
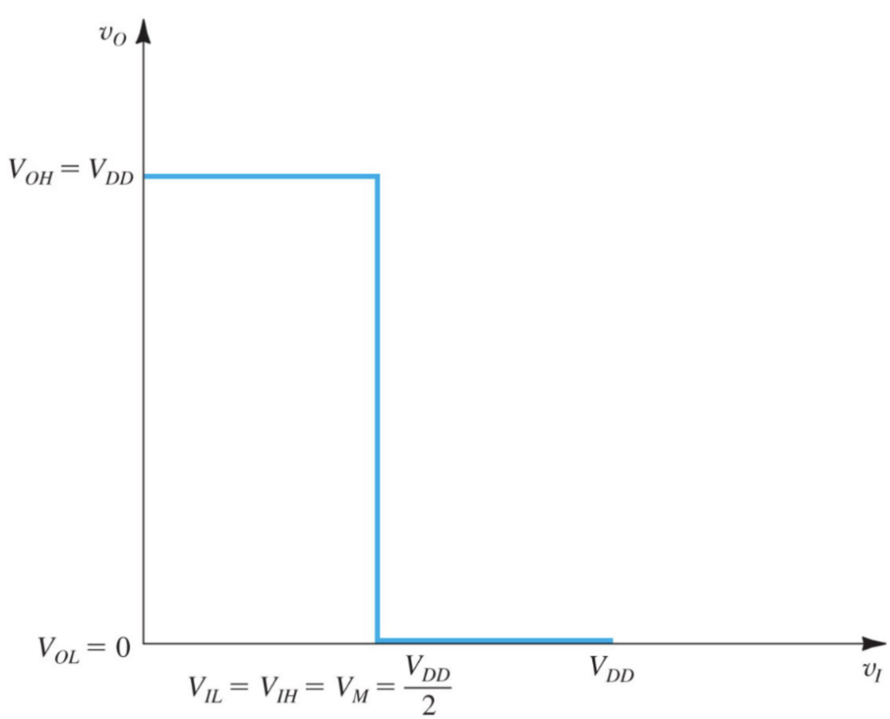


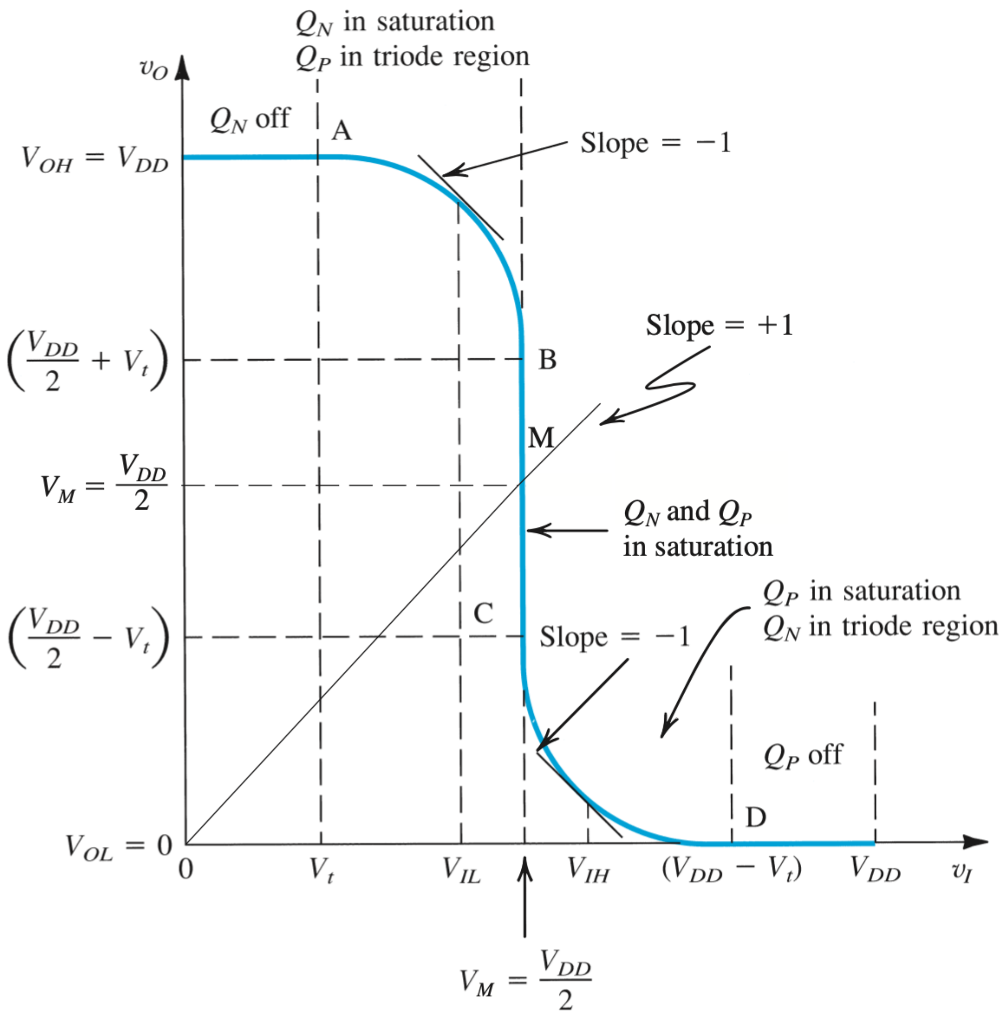
Figure 2.5 CMOS inverter when v_I is low.

$$r_{DSP} = \frac{1}{\left[k_p'(\frac{W}{L})_p(V_{DD} - |V_{tp}|)\right]} \tag{16}$$

Ideal Inverter Voltage Transfer Characteristics



- maximum output signal swing
  - $V_{OH} = V_{DD}$
  - $V_{OL} = O$
- maximized noise margins
- transition region has zero width



for $v_O \leq v_I - V_{tn}$

$$i_{DN} = k'_n \left(\frac{W}{L}\right)_n \left[(v_I - V_{tn})v_O - \frac{1}{2}v_O^2\right] \tag{17}$$

and for $v_O \geq v_I - V_{tn}$

$$i_{DN} = k'_n \left(\frac{W}{L}\right)_n (v_I - V_{tn})^2 \tag{18}$$

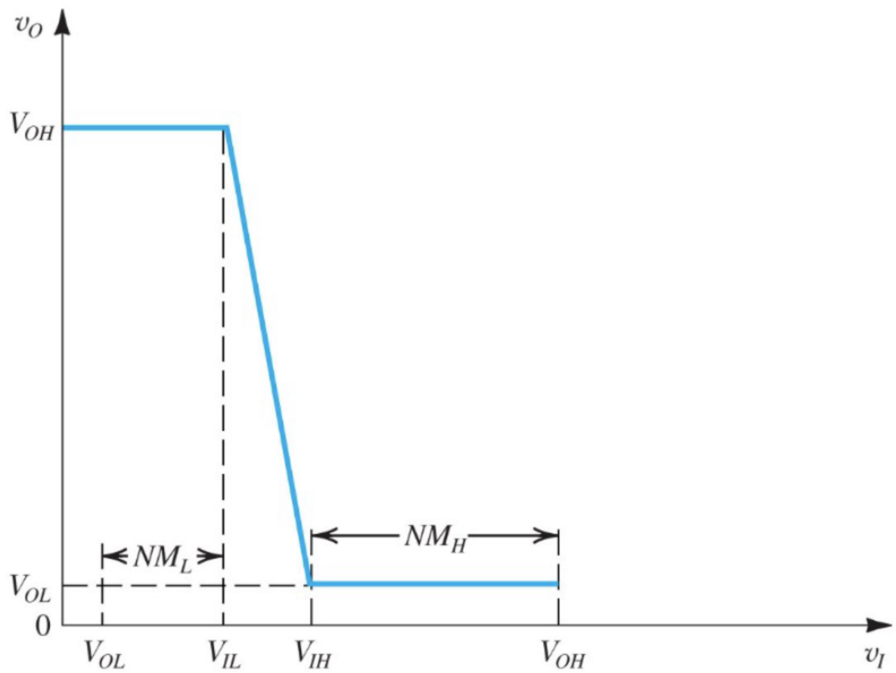
for $v_O \geq v_I + |V_{tp}|$

$$i_{DP} = k_p' \left(\frac{W}{L}\right)_p \left[(V_{DD} - v_I - |V_{tp}|)(V_{DD} - v_O) - \frac{1}{2}(V_{DD} - v_O)^2\right] \tag{19}$$

and for $v_O \leq v_I + |V_{tp}|$

$$i_{DP} = \frac{1}{2}k_p' \left(\frac{W}{L}\right)_p (V_D D - v_I - |V_{tp}|)^2 \tag{20}$$

Inverter VTC



* Straight line VTC approx.
$V_{OH} \leq V_{DD}$ $V_{OL} \geq 0$ *finite transition region: $V_{IH} - V_{IL}$ V_{OH} does not depend on v_I as long as $v_I \leq v_{IL}$ V_{OL} does not depend on v_I as long as $v_i \geq V_{IH}$ *looking to maximize noise margins

Inverter Noise Margins

$$NM_L = V_{IL} - V_{OL}$$

L: low

$$NM_H = V_{OH} - V_{IH}$$

Typical Inverter VTC

V_{OL}: output low V_{OH}: output high V_{IL}: max input interpreted as a logic 0 V_{IH}: max input interpreted as a logic 1

Ring Oscillators

Consists of an odd number of inverters in a circular chain Can be used to determine propagation delay or as a clock signal $f = \frac{1}{2Ntp}$ where n is the number of inverters $f = \frac{1}{N(t_{PLH}+t_{PHL})}$ if t_{PLH} not equal to t_{PHL}

MOSFET Amplifier as an Inverter

When $v_I = 0$, $v_O = V_{DD}$ since the transistor is off
When $v_I = V_{DD}, v_O =$ in triode region and will be modeled as a resistor
*For $v_I \leq V_t$, the NMOS is **off** $i_D = 0$ and $v_O = V_{OH} - V_{DD}$
*When v_I exceeds V_t, the MOSFET turns on an initially is in the *saturation region* (B to C)

$$i_D = \frac{1}{2}K_N(v_I - V_t)^2$$

$$v_O = V_{DD} - i_d R_d = V_{DD} - \frac{1}{2}K_n R_D(v_I - V_t)^2$$

*Beyond point C, the transistor is in the triode region

$$i_D = K_n[(v_I - V_t)v_O - \frac{1}{2}(v_O)^2]$$

$$v_O = V_{DD} - i_D R_D = V_{DD} - K_N R_D[(v_I - V_t)v_O - \frac{1}{2}(v_O)^2]$$