Chapter 7

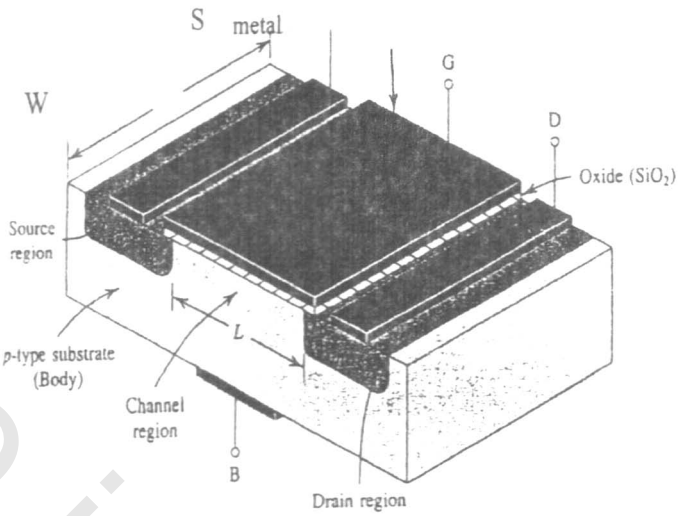
METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTOR (MOSFET)

The MOSFET is the most important device used in realizing complex digital integrated circuits such as microprocessors and high capacity semiconductor memories. MOSFET stands for Metal Oxide Semiconductor Field Effect Transistor. It is a device in which an externally applied electric field - in a direction normal to current path - controls the value of that current. MOSFET is a unipolar transistor, since the current is transported by carriers of one type only (electrons in n-channel transistors and holes in p-channel transistors.). In MOSFET, the majority dominates since the mechanism of current flow is drift, hence called majority dominated devices. In junction devices, the diffusion plays a central role, hence called injected minority carriers. BTT and other junction devices are called minority dominated devices. In MOSFET diffusion plays no part in device operation.

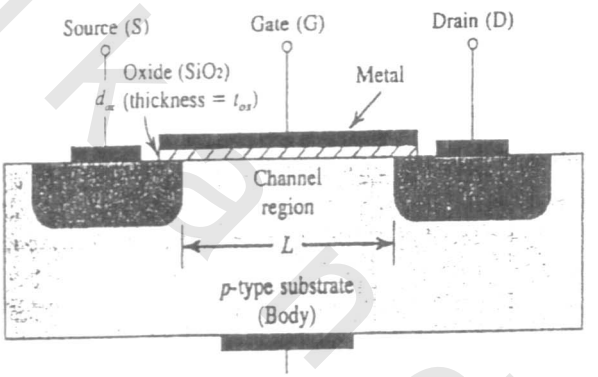
7.1 Enhancement MOSFET (E-MOSFET)

The structure of an n-channel transistor - called NMOS transistor or NMOST - is illustrated in Fig. (7.1). It is a four terminal device consisting of a p-type semiconductor substrate (body or bulk) into which two n^+ (heavily doped n) regions (the source and the drain) are formed. The gate electrode is the metal electrode on the insulating oxide. The basic device parameters which determine the transistor characteristics are the channel length L, (which is the distance between the two n^+ regions under the oxide), the channel width (the insulator thickness) d_{ox}, and the substrate doping N_A. The source electrode will be used as the voltage reference and is usually connected to the body then to ground. When no voltage is applied to the gate, the source-to-drain electrodes form two pn junctions connected back to back and no current flows between the drain and the source (the transistor is non-conducting or OFF). When a sufficiently large positive bias is applied to the gate, a surface inversion layer (or n-channel) is formed between the two n^+ regions. The minimum gate voltage required to obtain sufficient inversion is called the threshold voltage V_T. The source and the drain are now connected by a conducting n-channel, through which a large current can flow and the transistor becomes conducting. The current is enhanced, hence the name "enhancement MOSFET". The conductance of the channel can be controlled by the gate voltage. The transistor action occurs through the modulating effect of the gate voltage. Let us consider that a voltage $V_G > V_T$ is applied to the gate (Fig. 7.2a), creating a channel at the semiconductor surface (Fig. 7.2b). As current increases a voltage drop develops across the channel which decreases the gate drain voltage $V_{GD} = V_{GS} - V_{DS}$, and hence the width of the channel narrows at the drain. If a small drain to source voltage V_{DS} is applied, a drain current I_D will flow from the drain to the source through the conducting channel. Thus, the channel acts as a resistance, and I_D is linearly proportional to V_{DS}. This is the linear region of operation.

a)



b)



c)

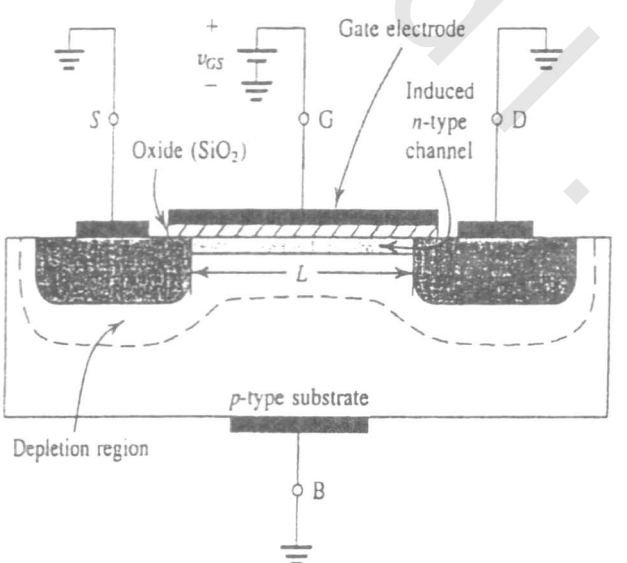


Fig. (7.1) N MOSFET

a) physical structure b) $V_{GS} < V_T$ c) $V_{GS} > V_T$

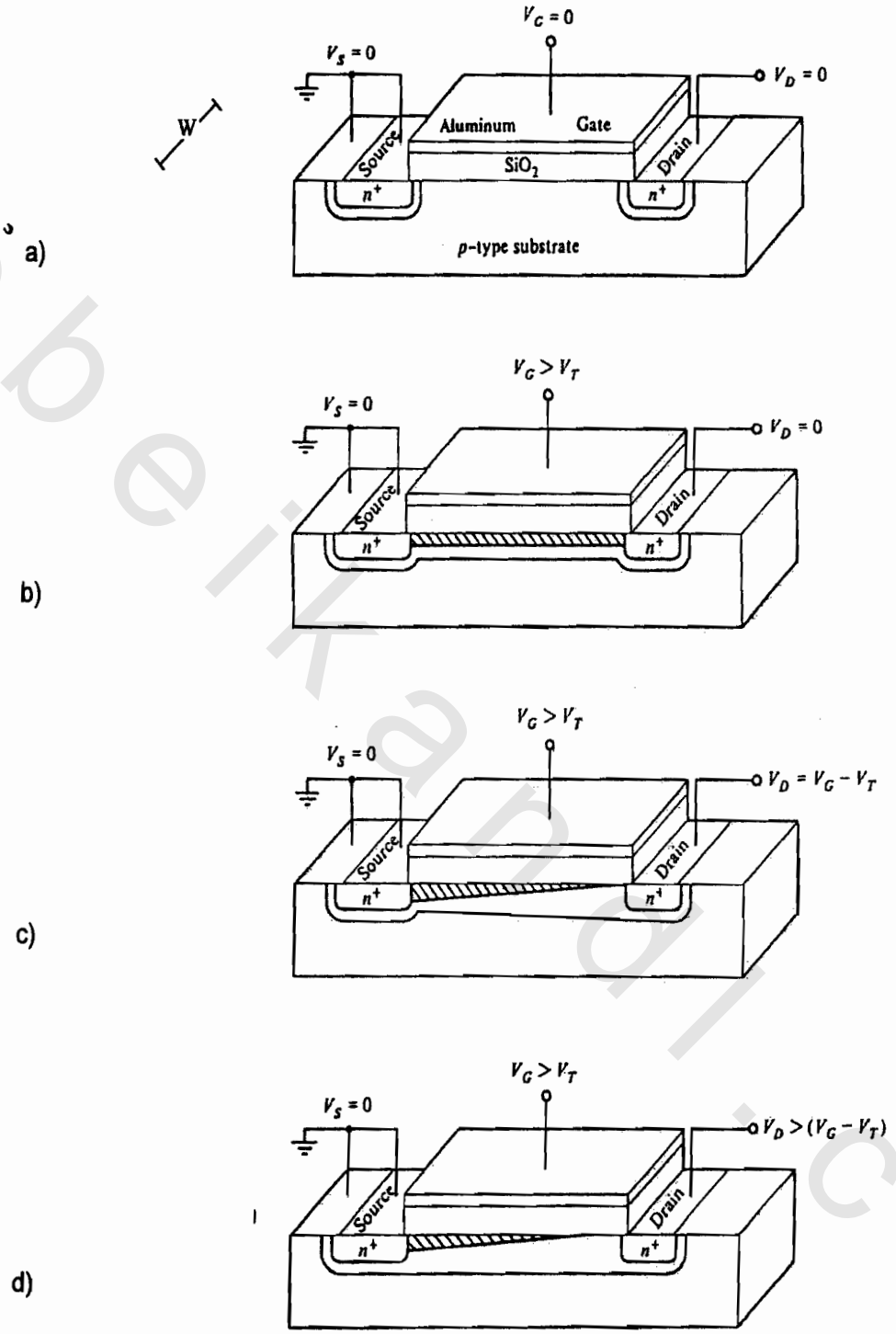


Fig. (7.2) n-channel enhancement-mode MOSFET under various bais condition

a) $V_{GS} = 0$ b) $V_{GS} > V_T$ and $V_{GD} > V_T$

c) $V_{GS} > V_T$ and $V_{GD} = V_T$ d) $V_{GS} > V_T$ and $V_{GD} < V_T$

When the drain voltage V_{DS} increases sufficiently, it will reach a value high enough to the extent that the difference in potential between the gate and the drain $V_{GD} = V_{GS} - V_{DS}$ is not sufficient to invert the surface at the drain. This is called channel pinch- off (Fig. 7.2c). For higher drain voltages beyond the pinch-off, the drain current remains essentially the same, and I_D saturates to a value $I_{D_{sat}}$ (Fig.7.2d). This is the saturation region of operation. It may be argued then that since the channel is pinched off current should cease. However, if the current tends to cease the pinch off is removed and current flows freely, but then pinch off exists again. Thus a dynamic condition prevails in which current remains constant regardless of V_{DS}. In the saturation region, the dynamic resistance is very high. A very high dynamic resistance means that an incremental increase in V_{DS} causes a negligible increase in I_D. Since the channel under saturation can be regarded as a low field resistance in series with a high field drain depletion region, any incremental increase in V_{DS} is dropped across the small depletion region, leaving essentially a constant potential $(V_{GS} - V_T) = V_{DS_{sat}}$ across the resistance of the low field channel, and hence leading to a constant drain current. The difference between V_{DS} and $V_{DS_{sat}}$ appears across the high field resistance of the small pinched off depletion region near the drain. As V_{DS} increases this high field resistance increases as well to keep the current constant.

Since V_T is the minimum difference in potential V_{GD} across the oxide required to create an inversion layer locally at the drain side, we should have

$$V_{GD} = V_{GS} - V_{DS} > V_T \quad \text{(linear and triode) region)} \qquad (7\text{-}1)$$

For the saturation region, where the channel is pinched-off at the drain, the difference in potential across the gate oxide at the drain is insufficient to form the channel there, so we have V_{GD} locally at the drain side given by

$$V_{GD} = V_{GS} - V_{DS} \le V_T \quad \text{(saturation region)}, \; V_{DS} \ge V_{GS} - V_T \qquad (7\text{-}2)$$

Usually, on the drain characteristics (I_D vs V_{DS} with V_{GS} as a parameter) one plots a locus separating the linear and triode regions from the saturation region. The equation of this locus is

$$V_{GD} = V_{GS} - V_{DS} = V_T \quad \text{(onset of the saturation locus)} \qquad (7\text{-}3)$$

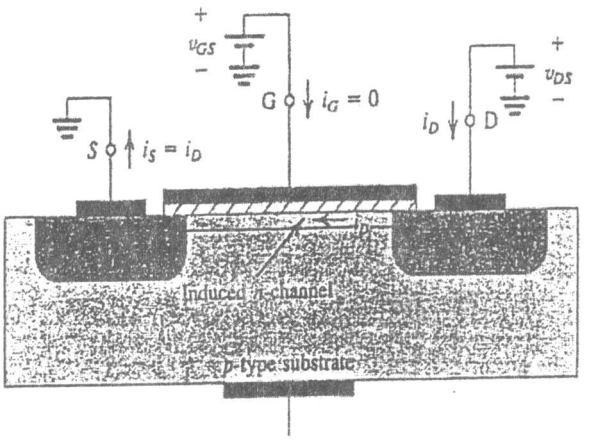
Note that this locus passes through the origin when $V_{GD} = V_{GS} = V_T$ where $V_{DS} = 0$ and, therefore, $I_D = 0$, indicating that the transistor is OFF. This cutoff region is obtained if $V_{GS} < V_T$

An overall characteristic is shown (Fig. 7.3) and the family of characteristics showing the laws $V_{GS} - V_{DS} = V_T$ or $V_{DS} = V_{GS} - V_T$ are depicted in Fig. (7.4) for NMOS and Fig. (7.5).

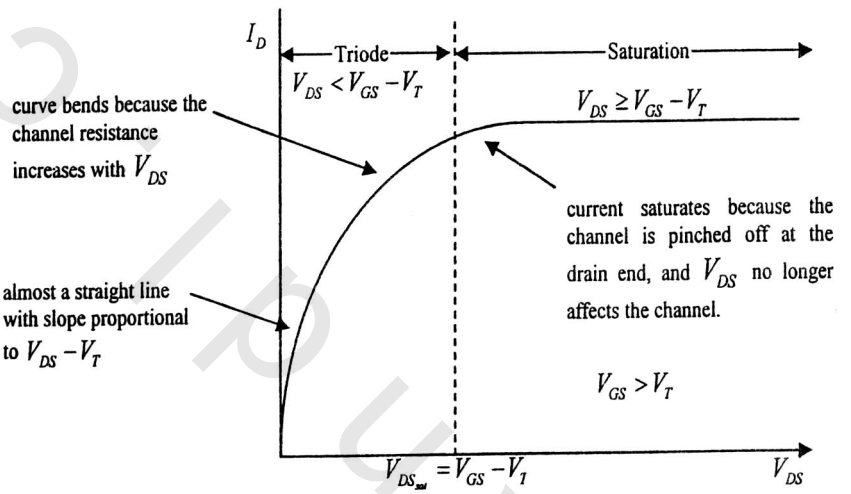
For PMOS we note that all values are negative and for saturation

$$V_{DS} \le V_{GS} - V_T \qquad (7\text{-}4)$$

a)



b)



c)

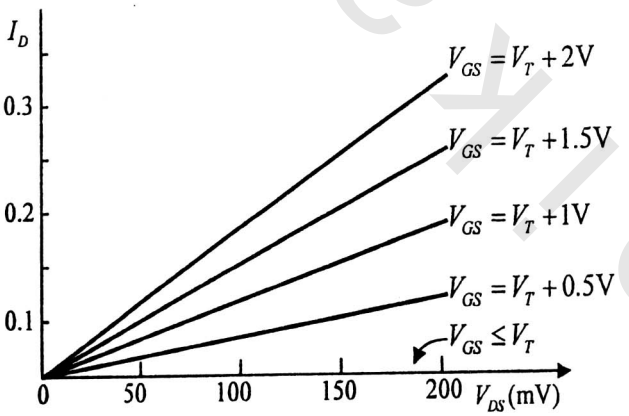


Fig. (7.3) NMOSFET

a) overall characteristics for a given V_{GS}

b) linear characteristics at small V_{DS} for different values of V_{GS}

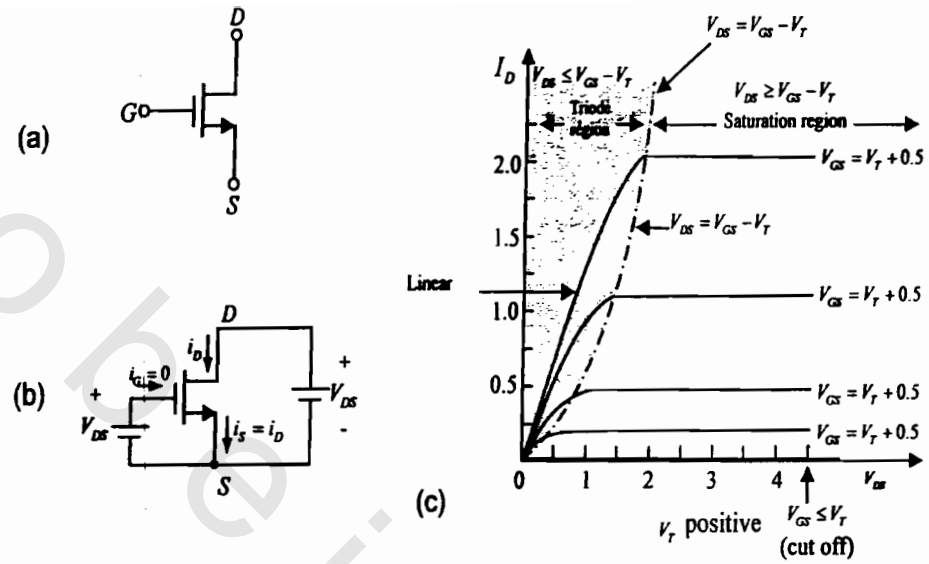


Fig. (7.4) NMOS

a) symbol b) circuit c) characteristics in saturation $V_{GS} > V_T$, $V_{GD} < V_T$

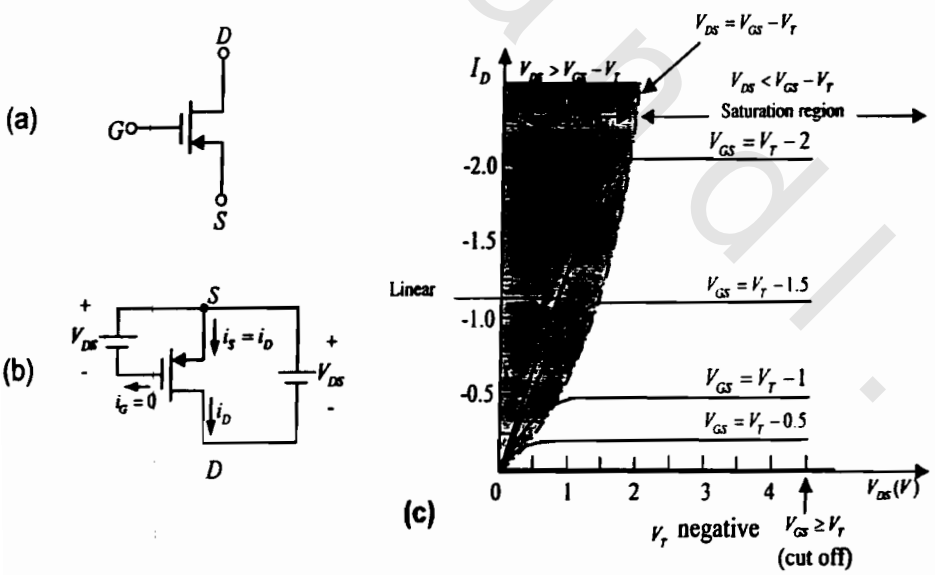


Fig. (7.5) PMOS

a) symbol b) circuit c) characteristics in saturation $V_{GS} < V_T$, $V_{GD} > V_T$

7.2 I-V Characteristics of N MOSFET (NMOST-NMOS)

We want here to find $I_D - V_D$ characteristics as a function of bias (gate) voltage V_{GS} considering the source to be grounded, and hence find the conductance of the channel. From eqns. (3-33), (3-34) we have

$$V_{GS} = V_{ox} + \phi_s$$
$$= -\frac{Q_s}{C_{ox}} + \phi_s \qquad (7\text{-}5)$$

where ϕ_s is the potential at the surface of the semiconductor. The induced negative charge Q_s in the semiconductor is composed of a mobile charge Q_n and a fixed charge in the depletion region Q_d. Substituting $Q_n + Q_d$ for Q_s we have

$$Q_n = -C_{ox}\left[V_{GS} - \left(\phi_s - \frac{Q_d}{\epsilon_{ox}}\right)\right] \qquad (7\text{-}6)$$

From eqn. (3-39) at threshold $\phi_s = 2\phi_F$. Substituting in eqn. (7-5), for $V_{GS} = V_T$, $Q_s = Q_d$ at the start of inversion

$$V_T = -\frac{Q_d}{C_{ox}} + 2\phi_F \qquad (7\text{-}8)$$

Assuming Q_d will not change much with V_G i.e., V_T remains constant

$$Q_n = -C_{ox}[V_G - V_T] \qquad (7\text{-}9)$$

with a voltage V_{DS} applied there is a voltage drop V_x (the voltage in the bulk under current at point x) from each point x in the channel to the source. Thus, the potential $\phi_s(x)$ required to achieve a strong inversion is $2\phi_F$ (eqn. 3-24) plus the voltage V_x according to the condition

$$\phi_s(x) = V_x(x) + 2\phi_F \qquad (7\text{-}10)$$

$$Q_n(x) = -C_{ox}\left[V_{GS} - 2\phi_F - V_x(x) - \frac{Q_d(x)}{C_{ox}}\right] \qquad (7\text{-}11)$$

We note from eqn. (3-38) and under $V_x(x)$

$$Q_d(x) = \sqrt{2|q|\epsilon_s N_A(2\phi_F + V_x(x))} \qquad (7\text{-}12)$$

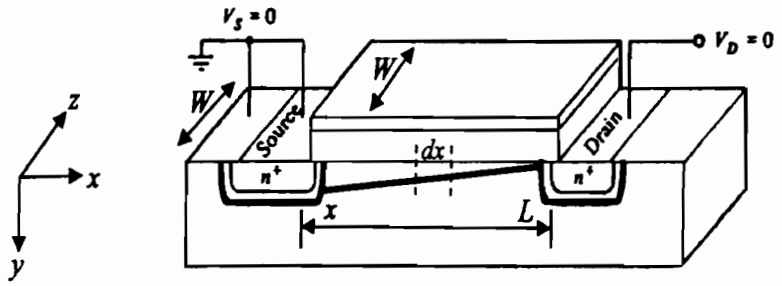
If we neglect the variation of $Q_d(x)$ with the voltage $V_x(x)$, and assuming that V_T is constant and ϕ_F will not change much with V_{GS} then eqn. (7-9) becomes

$$Q_n(x) = -C_{ox}[V_{GS} - V_T - V_x(x)] \qquad (7\text{-}13)$$

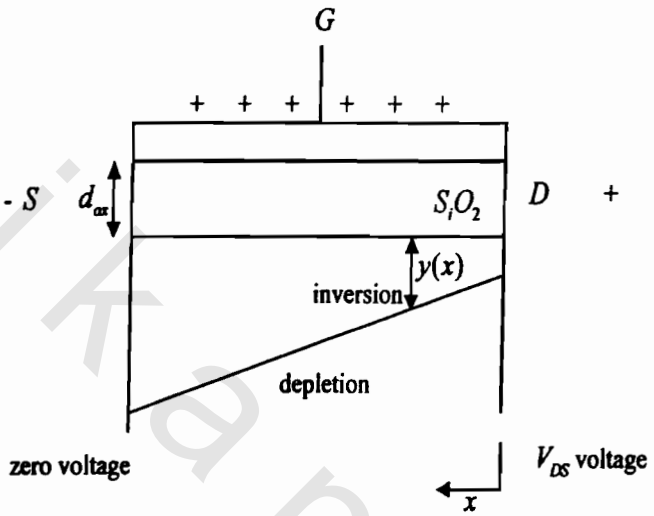
This equation describes the mobile charge in the channel at point x. The vertical potential drop across the oxide layer $V_{GS} - V_x(x)$ must always exceed the threshold voltage V_T for channel inversion to occur. Noting

$$Q_s = Q_n + Q_d = C_{ox}V_{ox} \qquad (7\text{-}14)$$
$$= C_{ox}[V_{GS} - \phi_s]$$
$$= \epsilon_{ox}\,\varepsilon_{ox}A$$

a)



b)



c)

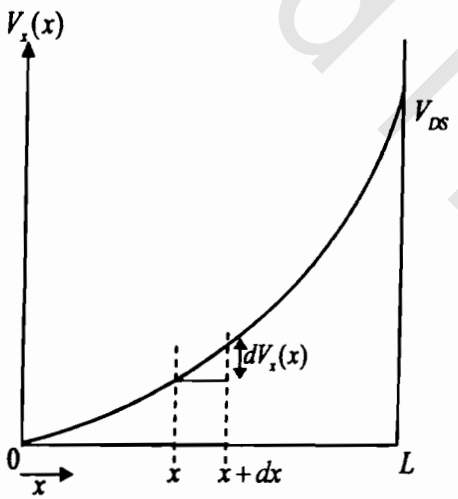


Fig. (7.6) Conduction channel in MOSFET
a) MOSFET view b) cross section c) variation of V_x along the channel

where $\bar{\varepsilon}_{ox}$ is the field in the oxide, and $\epsilon_{ox}\,\bar{\varepsilon}_{ox}$ is the displacement vector which is numerically equal to the charge density at the surface of the metal surface which is also equal to the charge density on the semiconductor side.

$$\epsilon_{ox} A\varepsilon_{ox} = C_{ox}\left[V_{GS} - \phi_s\right] \tag{7-15}$$

Noting from eqn. (7-10)

$$\phi_s = V_x(x) + 2\phi_F$$

From eqn. (7-15)

$$\varepsilon_{ox} = \frac{1}{\epsilon_{ox} A} C_{ox}\left[V_{GS} - V_x(x) + 2\phi_F\right] \tag{7-16}$$

Noting from eqn. (7-8)

$$V_T = \frac{Q_d}{C_{ox}} + 2\phi_F$$

$$2\phi_F = V_T + \frac{Q_d}{C_{ox}}$$

Eqn. (7-16) reduces to

$$\varepsilon_{ox} = \frac{1}{\epsilon_{ox} A} C_{ox}\left[V_{GS} - V_x(x) - V_T - \frac{Q_d}{C_{ox}}\right]$$

$$= \frac{1}{d_{ox}}\left[V_{GS} - V_x(x) - V_T - \frac{Q_d}{C_{ox}}\right] \tag{7-17}$$

The charge in the semiconductor $Q_n + Q_d$ must equal in magnitude the charge on the gate. Gauss's law tells us that the total flux leaving the gate normal to the surface is numerically equal to the surface charge density.

Thus

$$\epsilon_{ox}\,\varepsilon_{ox} = \varsigma_s(x) \tag{7-18}$$

where $\varsigma_s(x)$ is the surface charge density $\mathrm{C/cm^2}$

Combining eqns. (7-17), (7-18), neglecting Q_d / C_{ox} since Q_n exceeds the fixed charges Q_d

$$\varsigma_s(x) = \frac{\epsilon_{ox}}{d_{ox}}\left[V_{GS} - V_x(x) - V_T\right] \tag{7-19}$$

The thickness of the channel is $y(x)$ (Fig. 7.6). The gate area is A_g. The volume charge density $\varsigma_v(x)$ can now be expressed as

$$\varsigma_v(x) = \frac{\varsigma_s(x) A_g}{y(x) A_g} = \frac{\varsigma_s(x)}{y(x)} \tag{7-20}$$

where $\varsigma_v(x)$ is caused by trapped negative ions and mobile electrons but we will assume that the density of trapped ions is small compared to the mobile electron density. The conductivity of the channel at point x is given by

$$\sigma_n(x) = \varsigma_v(x)\mu_{ns} = \frac{\varsigma_s(x)}{y(x)}\mu_{ns} \tag{7-21}$$

$$\sigma_n(x) = \frac{\mu_{ns} \in_{ox}}{y(x) d_{ox}} \left[V_{GS} - V_x(x) - V_T \right] \tag{7-22}$$

But

$$J_n(x) = \sigma_n(x) \varepsilon(x) = \sigma_n(x) \left| \frac{dV_x(x)}{dx} \right| \tag{7-23}$$

We are assuming a variation in the channel voltage only in the x direction due to the IR drop in the channel

$$I_n(x) = J_n(x) A(x) \tag{7-24}$$

where $A(x)$ is the cross sectional area of the channel

$$A = W y(x) \tag{7-25}$$

The current must be constant (called drain current I_D)

$$I_D = W y(x) \sigma_n(x) \left| \frac{dV_x(x)}{dx} \right| \tag{7-26}$$

$$I_D dx = W y(x) \sigma_n(x) dV_x(x) \tag{7-27}$$

Substituting eqn. (7-19) in eqn. (7-24) we have

$$I_D dx = \frac{\mu_{ns} \in_{ox}}{d_{ox}} \left[V_{GS} - V_x(x) - V_T \right] dV_x(x) \tag{7-28}$$

$$I_D \int_0^L dx = \frac{\mu_{ns} \in_{ox}}{d_{ox}} \int_0^{V_D} \left[V_{GS} - V_x(x) - V_T \right] dV_x(x) \tag{7-29}$$

$$I_D = \frac{\mu_{ns} \in_{ox}}{L d_{ox}} \left[(V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2 \right] \tag{7-30}$$

Since

$$C_{ox} = \frac{\in_{ox} A_g}{d_{ox}} \tag{7-31}$$

We may define C'_{ox} as the oxide capacitance per unit area

$$C'_{ox} = \frac{\in_{ox}}{d_{ox}} \tag{7-32}$$

$$I_D = \frac{W}{L} \mu_{ns} C'_{ox} \left[(V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2 \right] \tag{7-33}$$

This is the characteristic equation for device operation in the so called triode region. It is valid only for the region of operation in which an inversion layer exists from source to drain. In other words it is valid for the region over which $V_{GS} > V_T$ and $V_{GD} = V_{GS} - (V_{DS} > V_T)$ or $(V_{GS} - V_T) > V_{DS}$ or $V_{DS} < (V_{GS} - V_T)$ (Fig. 7.3)

Note that when $V_{DS} = 0$, $I_D = 0$ in eqn. (7-33). For very small value of V_{DS}^2, i.e., $\frac{1}{2} V_{DS}^2 \ll (V_{GS} - V_T) V_{DS}$ or

$$\frac{1}{2} V_{DS} \ll V_{GS} - V_T \tag{7-34}$$

In this case we have

$$I_D = \frac{W}{L} \mu_{ns} C'_{ox} (V_{GS} - V_T) V_{DS} \tag{7-35}$$

with V_{GS} constant I_D is a linear function of V_{DS}. We call this the linear region (Fig. 7-3), where the MOSFET behaves as a linear resistor. When $V_{DS} = (V_{GS} - V_T) < \Delta V_{GS}$ where ΔV_{GS} represents the excess of V_{GS} over V_T, (called overdrive). The channel disappears near the drain since the gate-channel voltage is not sufficient to maintain strong inversion near the drain. This is called pinch off.

For $V_{DS} > \Delta V_{GS}$ the MOSFET is operating in the saturation region with the channel pinched off. Mobile charges are swept across the depletion region by the lateral electric field from the point where the channel is pinched off to the n^+ drain region. The narrowing of the channel compensates for the tendency of the current to increase with increasing V_{DS}. It results in a constant I_D. If I_D drops due to narrowing, pinch off is eliminated which tends to increase I_D. This is why I_D is saturated beyond pinch off. The saturation drain voltage is given by

$$V_{DS_{sat}} = (V_{GS} - V_T) = \Delta V_{GS} \tag{7-36}$$

The saturation current is given from eqn. (7-33) by

$$I_{D_{sat}} = \frac{W}{L}\mu_{ns}C'_{ox}\frac{(V_{GS} - V_T)^2}{2} = \frac{W}{L}\mu_{ns}C'_{ox}\frac{(\Delta V_{GS})^2}{2} \tag{7-37}$$

This is the square law characteristic for MOSFET operating in the saturation region. It tells us that with V_{GS} constant, I_D will be constant and no longer dependent on V_{DS}.

The current voltage characteristics for n-channel MOSFET are shown in Fig. (7.4). The dashed line represents the transition between the triode and saturation region.

7.3 Channel Conductance and Transconuctance

For N MOSFET we may define channel conductance g_{DS} as

$$g_{DS} = \frac{1}{r_{DS}} = \left.\frac{\partial i_D}{\partial v_{DS}}\right|_{V_{GS}=\text{const}} \tag{7-38}$$

It is a measure of the slope of the IV characteristic curves with constant gate voltage (Fig. 7.3b). It is also the reciprocal of the channel resistance r_{DS}. We find from eqns. (7-38) and (7-33)

$$g_{DS} = \frac{1}{r_{DS}} = \frac{W}{L}\mu_{ns}C'_{ox}(V_{GS} - V_T - V_{DS}) \tag{7-39}$$

In the linear case when V_{DS} is very small, the linear conductance g_{DS_o} is given by

$$g_{DS_o} = \mu_{ns}C'_{ox}(V_{GS} - V_T) = \mu_{ns}C'_{ox}\Delta V_{GS} \tag{7-40}$$

We note that the IV curves are flat in the saturation region eqn. (7-37). Application of eqn. (7-38) will yield $g_{DS} = 0$. However, we find that as V_{DS} is increased in the saturation region the depletion region around n^+ drain becomes wider. The channel length L decreases. But the current I_D increases to maintain the flow, thus, keeping $g_{DS_{sat}}$ at a non zero value (Fig. 7-7).

We may also define transconductance g_m as

$$g_m = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{V_{DS}=\text{const}} \tag{7-41}$$

It is a measure of the spacing between constant V_{GS} lines for a fixed value of V_{DS} (Fig. 7.4) or the slope of the $I_D - V_{GS}$ curve (Fig. 7.8).

Applying eqn. (7-41) in eqn. (7-33)

$$g_m = \frac{W}{L}\mu_{ns}C'_{ox}V_{DS} \tag{7-42}$$

This is valid for both linear and triode regions.

Applying eqn. (7-41) in the saturation region

$$g_{m_{sat}} = \frac{W}{L}\mu_{ns}C'_{ox}(V_{GS} - V_T) = \frac{W}{L}\mu_{ns}C'_{ox}\Delta V_{GS} \tag{7-43}$$

This is shown in Fig. (7.9)

We note that in all expressions in this section, $\mu_{ns}C'_{ox}$ is a constant-determined by the process technology used to fabricate the n-channel MOSFET. It is known as process transconductance parameter denoted by k'_n

$$k'_n = \mu_{ns}C'_{ox} \tag{7-44}$$

Thus, we may rewrite eqns. (7-33), (7-35), (7-37), (7-39) and (7-40) as

$$I_D = \frac{W}{L}k'_n\left[(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2\right] \quad \text{triode region} \tag{7-45}$$

$$I_D = \frac{W}{L}k'_n(V_{GS} - V_T)V_{DS} \quad \text{linear} \tag{7-46}$$

$$I_{D_{sat}} = \frac{W}{L}k'_n\frac{(V_{GS} - V_T)^2}{2} \quad \text{saturation} \tag{7-47}$$

$$g_{DS} = \frac{1}{r_{DS}} = \frac{W}{L}k'_n(V_{GS} - V_T - V_{DS}) \quad \text{triode region} \tag{7-48}$$

$$g_{DS_o} = \frac{1}{r_{DS_o}} = \frac{W}{L}k'_n(V_{GS} - V_T) \quad \text{linear region} \tag{7-49}$$

$$g_m = k'_n V_{DS} \quad \text{linear and triode region} \tag{7-50}$$

$$g_{m_{sat}} = \frac{W}{L}k'_n(V_{GS} - V_T) \quad \text{saturation region} \tag{7-51}$$

We may rewrite these equations with

$$\frac{W}{L}k'_n = k_n \tag{7-52}$$

It is to be noted that the drain current is proportional to the ratio of the channel width W to channel length L. This is known as the aspect ratio of the MOSFET. Thus,

$$g_{DS} = \frac{1}{r_{DS}} = k_n(V_{GS} - V_T - V_{DS}) \quad \text{triode region} \tag{7-53}$$

$$g_{DS} = \frac{1}{r_{DS}} = k_n(V_{GS} - V_T) \quad \text{linear region} \tag{7-54}$$

$$g_m = k_n V_{DS} \quad \text{linear and triode region} \tag{7-55}$$

$$g_{m_{sat}} = k_n(V_{GS} - V_T) \quad \text{saturation region} \tag{7-56}$$

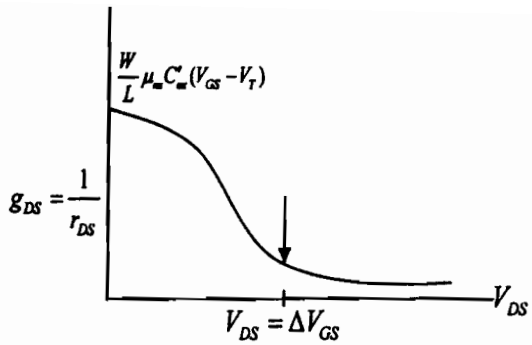


Fig. (7.7) g_{DS} for an n-channel MOSFET at constant $V_{GS}>V_T$

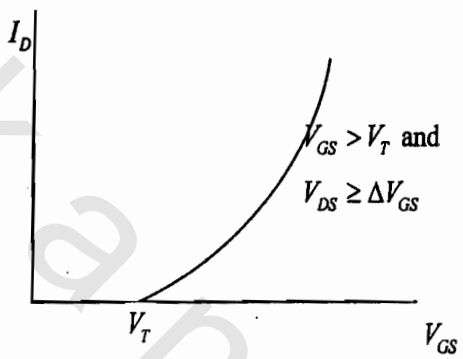


Fig. (7.8) I_D-V_{GS} characteristic for an n-channel MOSFET for constant V_{DS} in the saturation region

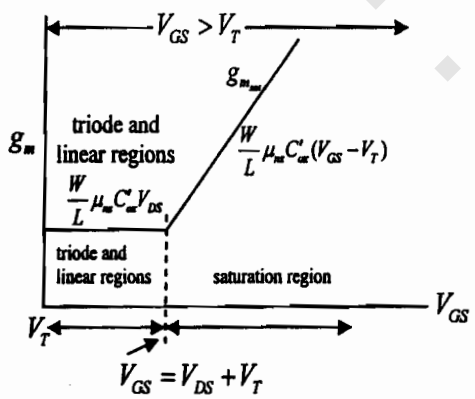


Fig. (7.9) g_m as a function of V_{GS} for a constant value of V_{DS}

Referring to the characteristics (Fig. 7.4) we have four regions: cut off ($V_{GS} < V_T$), linear region and triode region $V_{GS} > V_T$ and $V_{GD} = (V_{GS} - V_{DS}) > V_T$ or $V_{DS} < \Delta V_{GS}$ for continuous channel. If V_{DS} is sufficiently small we obtain the linear region is which we have a linear variable resistor r_{DS} controlled by V_{GS}. We may define r_{DS} in terms of over drive voltage $\Delta V_{GS} = V_{GS} - V_T$ as

$$r_{DS} = \frac{1}{k_n \Delta V_{GS}} \tag{7-57}$$

The fourth region is saturation. For saturation $V_{GS} \geq V_T$ while $V_{GD} \leq V_T$ or $(V_{GS} - V_{DS}) \leq V_T$ or $V_{DS} \geq \Delta V_{GS}$ for pinched off channel.

The boundary between the triode region and the saturation region is $V_{DS} = \Delta V_{GS}$ for which the boundary is given from eqn. (7-37) by

$$I_D = \frac{1}{2} k_n V_{DS}^2 \tag{7-58}$$

In this case the drain channel is independent of drain voltage (eqn. 7-47). The structured MOSFET behaves as an ideal current source where current is controlled by V_{GS} (Fig. 7.11) as shown by a large signal circuit model (Fig. 7.10).

The relative levels of terminal voltages for different regions are shown (Fig. 7.11)

Ex. 7.1

Redrive $I_D - V_{DS}$ characteristics of N MOSFET (NMOST) using charge control method.

Solution

Referring to Fig (7.12), consider the infinitesimal strip of the gate of distance x, from the source. The capacitance of this strip is $C'_{ox} W dx$. The effective voltage between the gate and the channel at point x needed to induce the channel is $(V_{GS} - V(x) - V_T)$.

The electron charge dQ in this strip is

$$|dQ| = C'_{ox} W dx [V_{GS} - V(x) - V_T] \tag{7-59}$$

The voltage V_{DS} produces an electric field along the channel in the negative x direction

$$\varepsilon(x) = -\frac{dV_x(x)}{dx} \tag{7-60}$$

The electric field causes the electron charge to drift toward the drain with velocity $\frac{dx}{dt}$

$$\left|\frac{dx}{dt}\right| = \mu_{ns} |\varepsilon(x)| = \mu_{ns} \left|\frac{dV_x(x)}{dx}\right| \tag{7-61}$$

The resulting drift current

$$i = \left|\frac{dQ}{dt}\right| = \left|\frac{dQ}{dx}\right| \left|\frac{dx}{dt}\right| \tag{7-62}$$

Using eqn. (7-59)

$$I_D = \mu_{ns} C'_{ox} W [V_{GS} - V(x) - V_T] \frac{dV_x(x)}{dx} \tag{7-63}$$

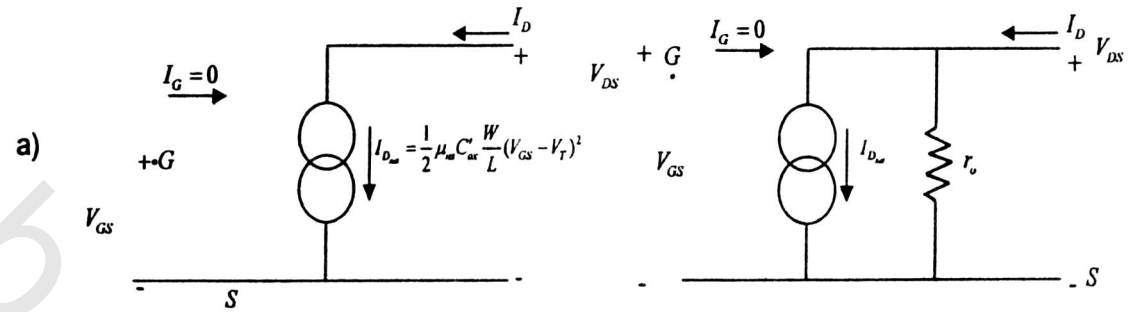


Fig. (7.10) Large signal equivalent circuit for N MOSFET at saturation $V_{GS} > V_T$, $V_{GD} < V_T$

a) without output resistance b) with output resistance

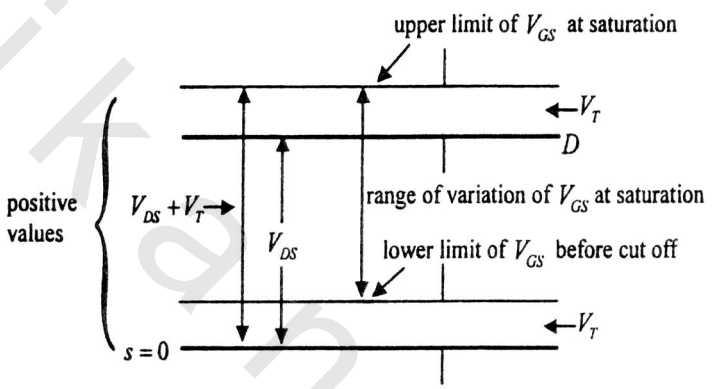


Fig. (7.11) Relative voltage levels for NMOST at saturation $V_{GS} > V_T$, $V_{GD} < V_T$ (V_T, V_{GS}, V_{GD} all positive) for a fixed value of V_{DS}, $V_{GD} = (V_{GS} - V_{DS}) < V_T$, $V_{GS} < (V_{DS} + V_T)$

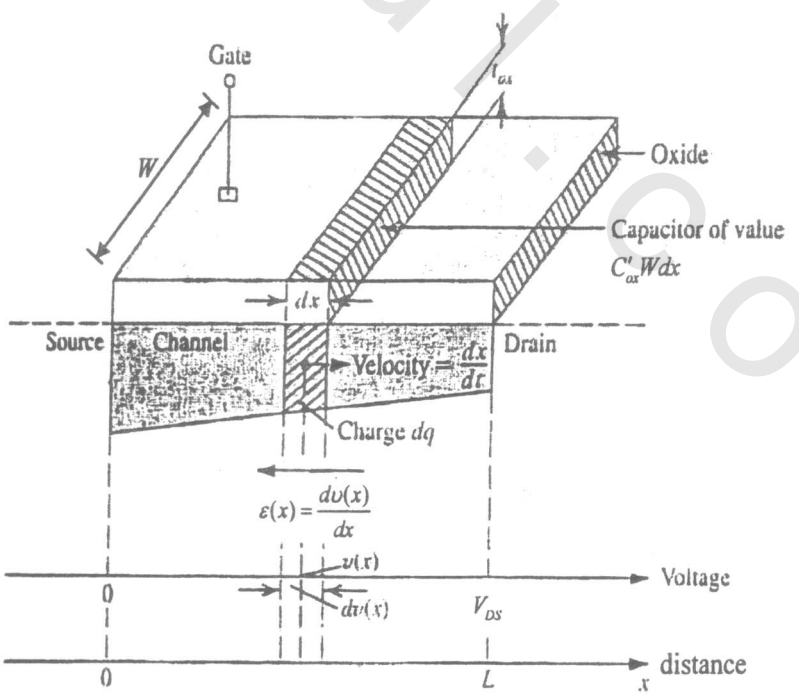


Fig. (7.12) Derivation of $I_D - V_{DS}$ characteristic for NMOST using charge control method

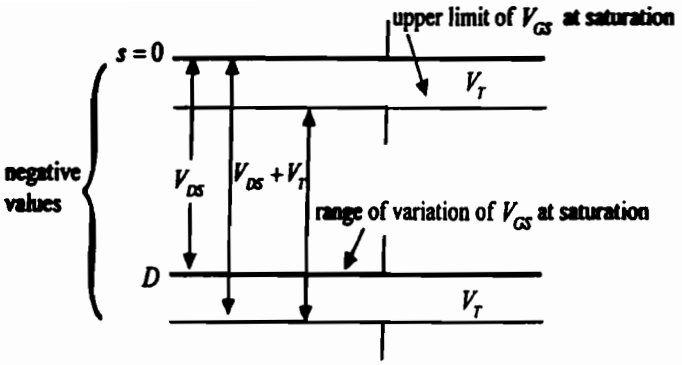


Fig. (7.13) Relative voltage levels for P MOST at saturation $V_{GS} < V_T$, $V_{GD} > V_T$ (V_T, V_{GS}, V_{GD} all negative) for a fixed value of V_{DS}, $V_{GD} = (V_{GS} - V_{DS}) > V_T$, $V_{GS} > (V_{DS} + V_T)$

$$I_D \int_0^L dx = \mu_{ns} C'_{ox} W \int_0^{V_{DS}} [V_{GS} - V(x) - V_T] dV_x(x)$$

$$I_D = \mu_{ns} C'_{ox} \frac{W}{L} [(V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2] \qquad \text{for triode region} \qquad (7\text{-}64)$$

$$I_{D_{sat}} = \frac{1}{2} \mu_{ns} C'_{ox} \frac{W}{L} (V_{GS} - V_T)^2 \qquad \text{for saturation} \qquad (7\text{-}65)$$

Eqns. (7-64) and (7-65) coincide with eqns. (7-33) and (7-37) respectively.

7.4 Characteristics of p Channel MOSFET (PMOST – PMOS)

The symbol for p channel enhancement type is shown (Fig. 7.5a). The voltage and current polarities are shown in Fig. (7.5b). Noting that the threshold voltage is negative to induce a channel we apply a gate voltage more negative than V_T i.e., $V_{GS} \le V_T$ or $|V_{GS}| \ge |V_T|$. Also V_{DS} must be negative or V_{SD} is positive which means that current I_e flows out of the drain. To operate in the triode region (continuous channel) $V_{DS} \ge V_{GS} - V_T$ where all values are negative.

Note that in NMOS $V_{DS} \ge (V_{GS} - V_T)$ for saturation, while in PMOS $V_{DS} \ge V_{GS} - V_T$ for triode operation. Similar to eqn. (7-45) but with replacing the transconductor parameter, k'_n by k'_p, we have for PMOST

$$I_D = \frac{W}{L} k'_p [(V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2] \qquad (7\text{-}66)$$

where

$$k'_p = \mu_{p_e} C_{ox} \qquad (7\text{-}67)$$

noting that V_{GS}, V_{DS} and V_T are all negative

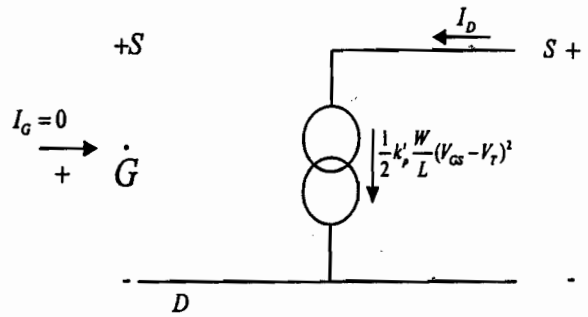


Fig. (7.14) Large signal equivalent circuit for P MOST at saturation $V_{GS} < V_T$, $V_{GD} > V_T$ (V_T, V_{GS}, V_{GD} all negative)

are all negative. Since $\mu_p < \mu_n$ it is expected that $k'_p < k'_n$.

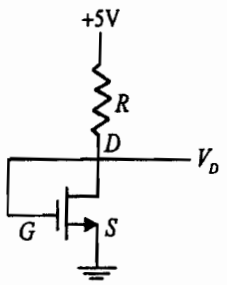
Now for saturation $V_{DS} \le (V_{GS} - V_T)$ where all values are negative. In this case, similar to eqn. (7-47)

$$I_{D_{sat}} = \frac{1}{2}k'_p\frac{W}{L}(V_{GS} - V_T)^2 \qquad (7\text{-}68)$$

Fig. (7.13) shows relative voltage levels for enhancement PMOST, and Fig. (7.14) shows large signal equivalent circuit for PMOST.

Ex. 7.2

In the circuit shown (NMOST) find R, V_{DS} at $I_D = 100\mu A$, $V_T = 1V$, $\mu_{ns}C_{ox} = 200\mu A/V^2$, $L = 1\mu m$, $W = 4\mu m$.



Solution

Since $V_{GD} = 0$ since $V_{GS} = V_{DS}$, $V_{DS} > (V_{DS} - V_T)$ and $V_{GS} > V_T$, thus the transistor is in saturation

$$I_D = \frac{1}{2}\mu_{ns}C'_{ox}\frac{W}{L}(V_{GS} - V_T)^2$$

$$= \frac{1}{2}\mu_{ns}C'_{ox}\frac{W}{L}(\Delta V_{GS})^2$$

$$\Delta V_{GS}^2 = \frac{2\times 100\times 10^{-6}}{200\times 10^{-6}\times 4} = 0.25$$

$$\Delta V_{GS} = 0.5V$$

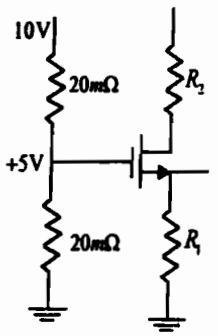
$$V_{DS} = V_{GS} = 1.5V$$

$$R = \frac{5 - 1.5}{0.1} = 45k\Omega$$

Ex. 7.3

Analyze the circuit shown to determine the voltage and current. Let $V_T = 1\text{V}$ and $k_n' \frac{W}{L} = 2\text{mA/V}^2$.

Find the condition for the transistor to be in saturation if $I_D = 1\text{mA}$



Solution

For saturation $V_{DS} \geq (V_{GS} - V_T)$

$V_G = 5\text{V}$,

$V_S = I_D R_1$

$V_{GS} = (5 - I_D R_1)$

$(V_{GS} - V_T) = (4 - I_D R_1)$

$$I_D = \frac{1}{2} k_n' \frac{W}{L} (V_{GS} - V_T)^2$$

$$= 10^3 (V_{GS} - V_T)^2$$

Thus

$$(4 - I_D R_1)^2 = 10^3 I_D$$

If $I_D = 1\text{mA}$

$$(4 - I_D R_1)^2 = 1$$

$$4 - I_D R_1 = 1$$

$$I_D R_1 = 3$$

$$R_1 = 3\text{k}\Omega$$

$$V_{GS} = 5 - 3 = 2\text{V}$$

$$V_{DS} \geq V_{GS} - V_T = 2 - 1 = 1\text{V}$$

V_{DS} must be $\geq 1\text{V}$

$$7\text{V} = I_D R_2 + V_{DS}$$

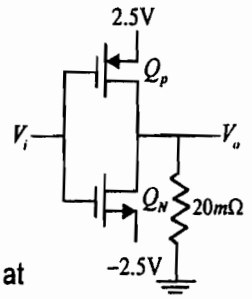
$$7 - I_D R_2 > 1\text{V}$$

$$6\text{V} \geq I_D R_2$$

$$I_D R_2 \leq 6\text{V}$$

$$R_2 = 2R_1$$

Ex. 7.4



The NMOS and PMOS transistors shown are matched $k_n' \frac{W_n}{L_n} = k_p' \frac{W_p}{L_p} = 1\text{mA/V}^2$, $V_{T_n} = |V_{T_p}| = 1\text{V}$. Find the currents and voltages for $V_i = 0\text{V}, +2.5\text{V}, -2.5\text{V}, 5\text{V}, -5\text{V}$

Solution

When $V_i = 0\text{V}$, $|V_{GS}| = 2.5\text{V}$, $|V_{os}| = V_o = 0$ due to symmetry. Then both Q_n, Q_p operate at $|V_{GS}| = 2.5 > |V_T|$, $|V_{DS}| = 0$ and hence both transistors are in saturation. Thus,

$$I_{D_p} = I_{D_n} = \frac{1}{2} \times 1(2.5-1)^2$$
$$= 1.125\text{mA}$$

When $V_i = +2.5\text{V}$, Q_p will have $V_{GS} = 0$ and thus will be cut off. Note V_o is negative. Thus V_{GD} is positive assuming $V_{GD} > V_T$ then, Q_n is in the triode region (linear region for simplicity)

$$I_{DN} = k_n' \frac{W_n}{L_n}(V_{GS_n} - V_T)V_{DS_n}$$
$$= (5-1)[V_o - (-2.5)]$$
$$= 4(V_o + 2.5)$$

Also
$$V_o = -I_{DN} \times 10\text{k}\Omega$$
$$I_{DN} = 4[-I_{DN} \times 10 + 2.5]$$
$$= \frac{2.5 \times 4}{41 \times 10^3} = 0.24\text{mA}$$
$$V_o = -2.44\text{V}\,,\ V_{DS_n} = -2.44 - (-2.5) = 0.06\text{V}$$

Which is small.

The situation for $V_i = -2.5$ is the complement of the situation $V_i = 2.5\text{V}$

If $V_i = +5\text{V}$ Q_p will remain OFF

$V_{GS_n} = 7.5\text{V}$ Q_N is ON, V_o will still almost be at -2.5V, $V_{DS_n} \cong 0$ i.e., Q_N is almost short circuited, the output is almost -2.5V. When we go positive the opposite happens V_o is almost fixed at $+2.5\text{V}$. For $V_i = -5\text{V}$ we have the complement of the situation at $V_i = +5\text{V}$.

7.5 The Transfer Characteristics

Fig. (7.15) shows the NMOS transistor as a common source amplifier with load line depicted. The Q point is determined by the biasing circuit. Operation around the Q point entails amplification of a small signal, while large excusion along the load line entails operation as a switch. Operation in the saturation region renders the transistor as a current source or transconductance amplifier.

Fig. (7.16) shows the transfer characteristic, a relation between V_o and V_i. To operate as a voltage amplifier the quiescent point Q is in the middle of the nearly linear part of the curve where dv_o / dv_i at Q is the gain. The correlation between points A, B, C, Q in Fig. (7.15) and the corresponding points in Fig. (7.16)is clear. To operate the transistor as a switch operation takes place between the two extreme points of the transfer curve.

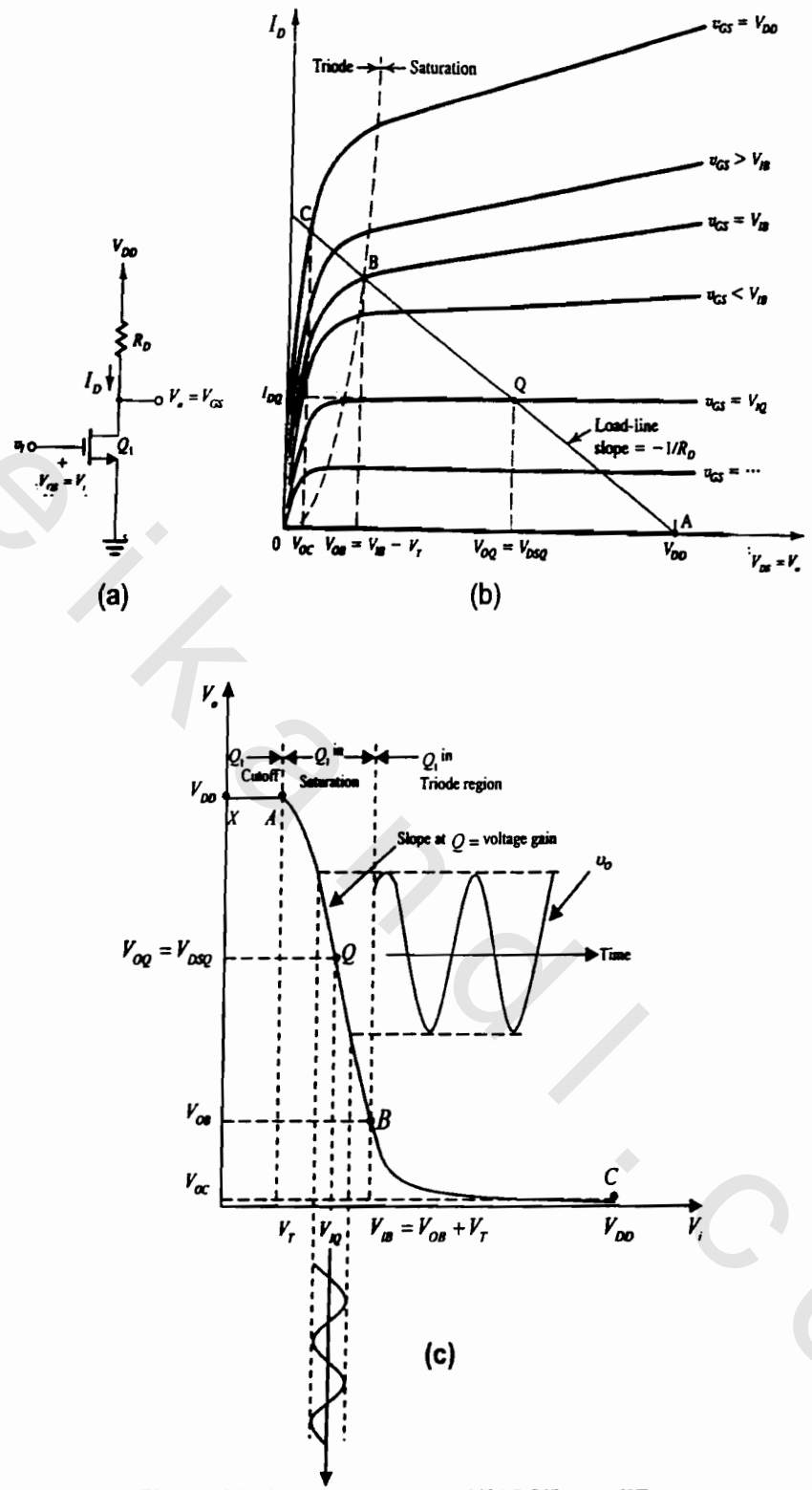


Fig. (7.15) Common source NMOST amplifier
a) circuit b) output characteristics c) transfer characteristic

In Fig. (7.16) we can recognize these regions.

a) region XA $V_i \le V_T$, $V_o \le V_{DD}$. This is cutoff

b) region AQB (the saturation region)

$V_i = V_{GS} \ge V_T$ and $V_{DS} = V_o \ge (V_i - V_T) = \Delta V_{GS}$. Since we are in saturation region.

$$I_{D_{sat}} = \frac{1}{2}\mu_{ns} C'_{ox}\left(\frac{W}{L}\right)(V_i - V_T)^2 \tag{7-69}$$

$$V_o = V_{DD} - I_{D_{sat}} R_D$$

$$V_o = V_{DD} - \frac{1}{2}\mu_{ns} C'_{ox}\left(\frac{W}{L}\right) R_D (V_i - V_T)^2 \tag{7-70}$$

The voltage gain A_v is given by

$$A_v = \left.\frac{d\upsilon_o}{d\upsilon_i}\right|_{V_Q} = -R_D \mu_{ns} C'_{ox}\left(\frac{W}{L}\right)(V_{iQ} - V_T) \tag{7-71}$$

The voltage gain is proportional to R_D, the transconductance parameter $k'_n = \mu_{ns} C'_{ox}$, the aspect ratio $\frac{W}{L}$ and the overdrive $\Delta V_i = V_{iQ} - V_T = V_{GS_Q} - V_T$

Substituting eqn. (7-71) into eqn. (7-70)

$$V_{oQ} = V_{DD} + \frac{1}{2} A_v (V_{iQ} - V_T) \tag{7-72}$$

From which

$$A_v = \frac{2(V_o - V_{DD})}{V_{iQ} - V_T} = \frac{2(V_{DD} - V_{oQ})}{V_{iQ} - V_T} \tag{7-73}$$

$$= \frac{-2(V_{DD} - V_{oQ})}{\Delta V_{iQ}} = \frac{-2V_{R_D}}{\Delta V_{iQ}} \tag{7-74}$$

From eqns. (7-71) and (7-51)

$$A_v = -g_{m_{sat}} R_D \tag{7-75}$$

From eqns. (7-47) and (7-51)

$$I_{D_{sat}} = \frac{1}{2} g_{m_{sat}} \Delta V_{GS} \tag{7-76}$$

If we consider $\Delta I_{D_{sat}} = i_d$ around Q point, from eqn. (7-41) $\Delta V_{GS} = \upsilon_{gs}$ is the small signal input voltage

$$i_d = \Delta I_D = g_{m_{sat}} \upsilon_{gs} \tag{7-77}$$

The end point of the saturation region is

$$V_{oB} = V_{iB} - V_T \tag{7-78}$$

or

$$V_{iB} = V_{oB} + V_T \tag{7-79}$$

c) The triode region $V_i \ge V_T$, $V_{DS} = V_o < (V_i - V_T)$

From eqn. (7-64)

$$I_D = \mu_{ns} C'_{ox} \frac{W}{L}\left[(V_i - V_T)V_o - \frac{1}{2}V_o^2\right] \tag{7-80}$$

$$V_o = V_{DD} - R_D I_D \tag{7-81}$$

$$V_o = V_{DD} - R_D \mu_{ns} C'_{ox} \frac{W}{L}\left[(V_i - V_T)V_o - \frac{1}{2}V_o^2\right] \tag{7-82}$$

For small V_o

$$V_o = V_{DD} - R_D \mu_{ns} C'_{ox} \frac{W}{L}(V_i - V_T)V_o \tag{7-83}$$

$$V_o = \frac{V_{DD}}{1 + R_D \mu_{ns} C'_{ox} \frac{W}{L}(V_i - V_T)} \tag{7-84}$$

From eqns. (7-39), (7-40), for small V_{DS}

$$r_{DS} = r_{DS_o} = \frac{1}{\mu_{ns} C'_{ox} \frac{W}{L}(V_{GS} - V_T)} \tag{7-85}$$

Combining eqns. (7-84), (7-85)

$$V_o = \frac{V_{DD}}{1 + \frac{R_D}{r_{DS}}} = V_{DD} \frac{r_{DS}}{r_{DS} + R_D} \tag{7-86}$$

For small V_o, MOST behaves as a resistance r_{DS} (determined by V_i) which forms a voltage divider across V_{DD}. If $r_{DS} < R_D$

$$\upsilon_o = V_{DD} \frac{r_{DS}}{R_D} \tag{7-87}$$

7.6 Current Mirror

In Fig. (7.16) the drain of Q_1 is shorted to its gate, thus Q_1 is in saturation

$$I_{D_1} = \frac{1}{2} k'_n \left(\frac{W}{L}\right)_1 (V_{GS} - V_T)^2 \tag{7-88}$$

$$I_{D_1} = I_{ref} = \frac{V_{DD} + V_{SS} - V_{GS}}{R} \tag{7-89}$$

We have two equations and two unknowns V_{GS}, R for a given $I_{D_1} = I_{ref}$

Since Q_2 is identical to Q_1 and has the same V_{GS} as Q_1, the reference current $I = I_{D_2}$ is given by

$$I = I_{D_2} = \frac{1}{2} k'_n \left(\frac{W}{L}\right)_2 (V_{GS} - V_T)^2 \tag{7-90}$$

Hence,

$$I_o = I_{ref} \frac{\left(\frac{W}{L}\right)_2}{\left(\frac{W}{L}\right)_1} \tag{7-91}$$

This circuit is called current mirror and is used in biasing the MOST. A typical circuit is shown (Fig. 7.17) where points X (Fig. 7.16) and X' (Fig. 7.17) are connected together and the current source is I_{ref}.

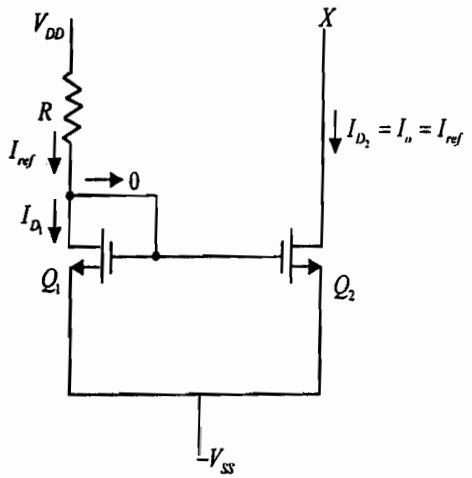


Fig. (7.16) Biasing MOSFET using a constant source (current mirror)

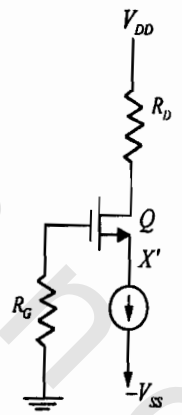


Fig. (7.17) Current mirror

7.8 Small Signal Model

Consider the situation when a small signal υ_{gs} is applied around the Q point. The total gate source voltage is v_{GS}. A small signal current i_d flows in addition to the biasing current i_d, the total current is i_D

$$\upsilon_{GS} = V_{GS} + \upsilon_{gs} \tag{7-92}$$

$$i_D = I_D + i_d \tag{7-93}$$

$$i_D = \frac{1}{2}k_n' \frac{W}{L}\left(V_{GS} + \upsilon_{gs} - V_T\right)^2$$

$$= \frac{1}{2}k_n' \frac{W}{L}\left(V_{GS} - V_T\right)^2 + k_n' \frac{W}{L}\left(V_{GS} - V_T\right)\upsilon_{gs} + \frac{1}{2}k_p' \frac{W}{L}\upsilon_{gs}^2 \tag{7-94}$$

The first term is the dc value I_D. Neglecting the third term

for small υ_{gs}

$$\upsilon_{gs} \ll 2\left(V_{GS} - V_T\right) \tag{7-95}$$

a)

b)

c)

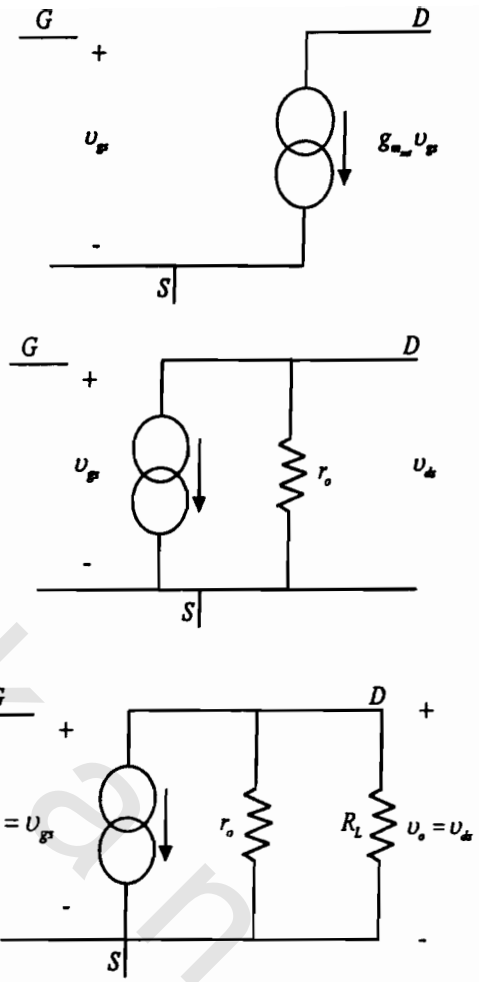


Fig. (7.18) Small signal equivalent circuit for MOSFET
a) without output resistance
b) with output resistance
c) with R_L connected

Thus, the small signal current i_d is given by

$$i_d = k_n' \frac{W}{L}(V_{GS} - V_T)\upsilon_{gs} \tag{7-96}$$

From eqn. (7-41)

$$g_m = \frac{i_d}{\upsilon_{gs}} = k_n' \frac{W}{L}(V_{GS} - V_T) = k_n' \frac{W}{L}\Delta V_{GS} \tag{7-97}$$

The small signal equivalent circuit is shown (Fig. 7.19) which describes variations around the Q point not the biasing condition. Thus

$$i_d = g_m \upsilon_{gs} \tag{7-98}$$

which is in agreement with eqn. (7-77)

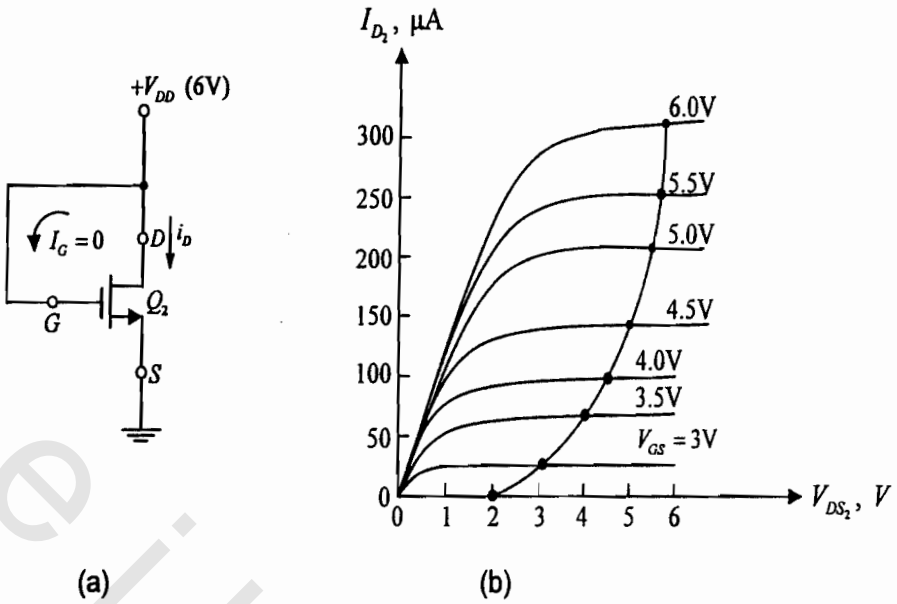


Fig. (7.19) N MOSFET as a nonlinear resistor (active load)
a) circuit b) characteristic

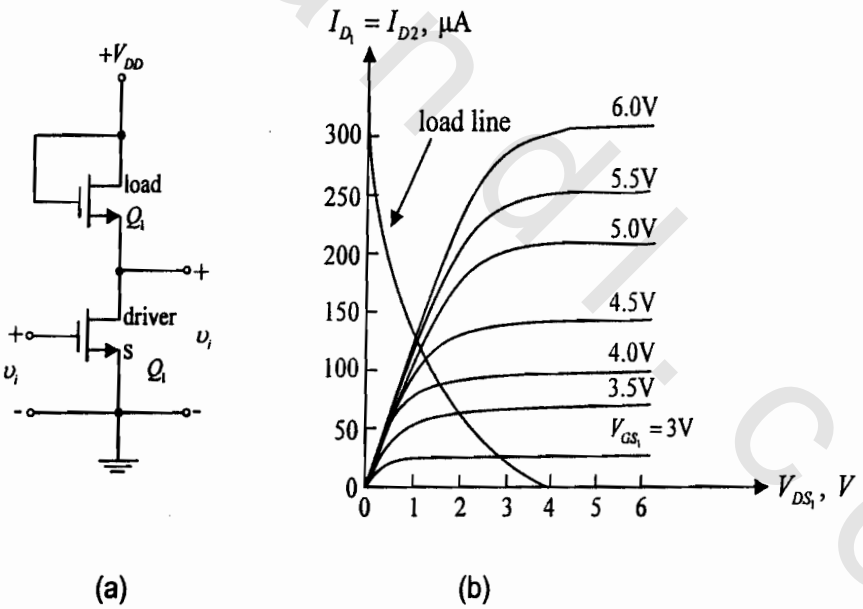


Fig. (7.20) MOSFET with active load
a) circuit b) output characteristic with active load as a load line

$$I_D = \frac{1}{2} k_n' \frac{W}{L} \Delta V_{GS}^2 \tag{7-99}$$

$$g_m = \sqrt{2k_n' \frac{W}{L} I_D} \tag{7-100}$$

$$= \frac{2I_D}{\Delta V_{GS}} \tag{7-101}$$

The voltage gain A_v is given by

$$A_v = -\frac{v_o}{v_i} = -v_{ds} / v_{gs} = -g_m R_L' \tag{7-102}$$

where R_L' is the parallel combination of r_o (the output resistance of the transistor) and R_D. For $R_D \ll r_o$

$$A_v = -g_m R_D \tag{7-103}$$

which is in agreement with eqn. (7-75)

We should note that maximum value for the ac input signal to maintain saturation for a given R_L is given by

$$v_{DS_{min}} = v_{GS_{max}} - V_T \tag{7-104}$$

$$V_{DS} - A_v \hat{v}_i = V_{GS} + \hat{v}_i - V_T \tag{7-105}$$

Where $\hat{v}_i$ is the maximum input signal. The Q point and the load line should be chosen so that the swing is symmetrical without clipping the signal and of maximum possible drive (Prob. 7.15).

7.8 Active Load

We can use MOSFET as a nonlinear resistor. In the circuit shown (Fig. 7.19a) the gate is connected to the drain. The nonlinear characteristic (Fig. 7.19b) is the locus of the points $V_{DS_2} = V_{GS_2}$. Note that this MOSFET operates in the saturation region since $\left(V_{GS_2} - V_T\right) < V_{DS_2}$. Note also that at $I_D = 0$

$$V_{DS_2} = V_{GS_2} = V_T$$

The circuit of Fig. (7.20) employs a MOSFET where Q_1 (called the driver) has output characteristics displayed in Fig. (7.20b). The load MOSFET Q_2 has the resistance chacacteristic given in Fig. (7.19b). We have

$$V_{DS_1} + V_{DS_2} = V_{DD}$$

or

$$V_{DS_1} = V_{DD} - V_{DS_2}$$

Since $I_G = 0$ for both Q_1 and Q_2 $I_{D_1} = I_{D_2}$ the load characteristic in Fig. (7.19b) is a plot of I_{D_2} versus V_{DS_2}. For $V_{DD} = 6\text{V}$, we note $I_{D_2} = 320\mu\text{A}$, $V_{GS_2} = V_{DS_2} = 6\text{V}$, $V_{DS_1} = 0$. This determines one point on the load line in Fig. (7.20b) ($I_{D_1} = 320\mu\text{A}$, $V_{DS_1} = 0$), When $I_{D_2} = 80\mu\text{A}$ $V_{GS_2} = V_{DS_2} = 4\text{V}$ so that the point $I_{D_1} = 80\mu\text{A}$, $V_{DS_1} = 6 - 4 = 2\text{V}$ is also a point on the load line in Fig. (7.20b).

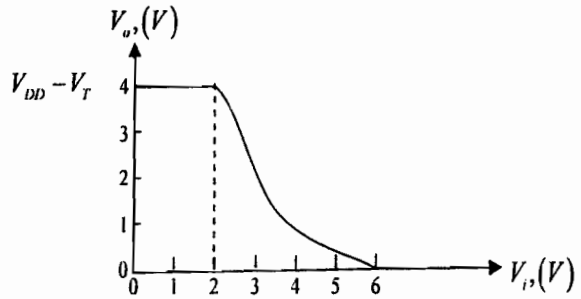
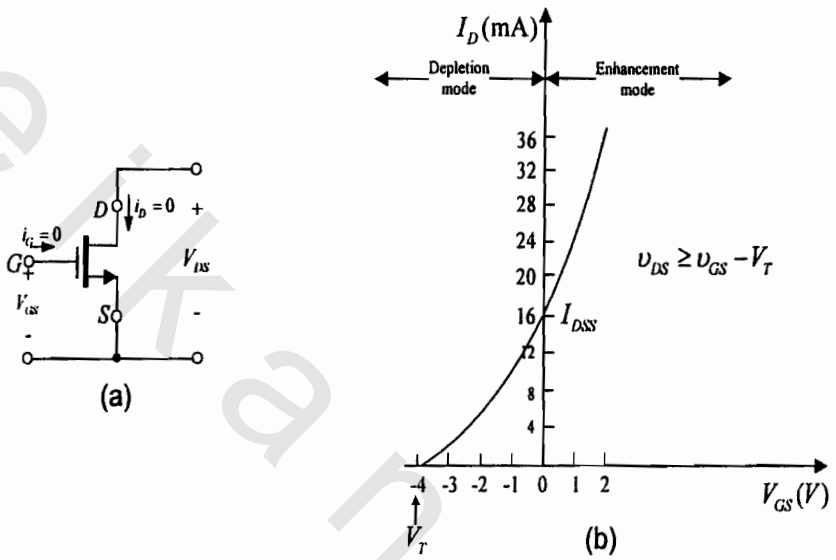


Fig. (7.21) MOSFET transfer characteristic with active load



(c)

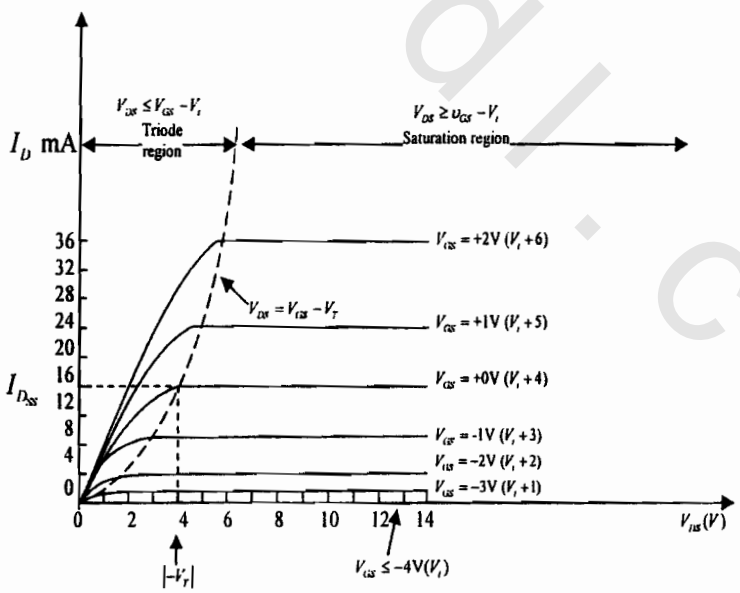


Fig. (7.22) D- MOST

a) circuit b) transfer characteristic c) output characteristics

Thus, for each value of I_{D_2} in Fig. (7.19b) for which $V_{GS_2} = V_{DS_2}$ we obtain a value of V_{DS_2}. The pair of values (I_{D_1}, V_{DS_1}) determines one point on the load line of Fig. (7.20b). Now we can determine V_{DS_1} (output) as a function of V_{GS_1} (input). For $V_i = V_{GS_1} \le V_T = 2\text{V}$, $I_{D_1} = 0$, $V_{DS_1} = 4\text{V}$. Increasing V_i to 5V results in $V_{DS_1} = 1.5\text{V}$ and is determined from the intersection of the load line and the characteristic for $V_{GS_1} = 5\text{V}$. Thus, we obtain the transfer characteristic (Fig. 7.21) as $V_o = V_{DS_1}$ versus $V_i = V_{GS_1}$. Note that for $V_i = 0$ $V_o = V_{DD} - V_T$. Active loads are usually used in integrated circuits (ICs) since they take less area than conventional resistors.

7.9 Depletion MOSFET (DMOS)

A second type of MOSFET has a narrow n-channel already embedded into the substrate. For $V_{GS} = 0$ positive V_{DS} produces appreciable drain current I_{DSS}. As V_{GS} decreases toward threshold the drain current decreases, At fixed V_{GS} increasing value of V_{DS} causes I_D to saturate as the channel is pinched off. In the depletion NMOS, V_T and V_{GS} are negative. If we apply positive gate voltage enhances the channel (Fig. 7.22).

7.10 CMOS

The combination of NMOS and PMOS transistors on the same chip are called complementary MOS (CMOS) devices. The NMOS transistor is the driver and the PMOS transistor is connected as the active load. The gates terminals of two transistors are connected together. The drain and source terminals of both transistors are series connected. When V_i is provided such that for NMOS transistor $V_{GD_n} > V_T$ the PMOS is OFF and the NMOS is ON provided $V_i > V_T$. If the input is negative such that $V_i > -|V_T|$ the PMOS is ON and the NMOS is cutoff. The situation is idealized as two switches connected in series, one is ON and one is OFF. When V_i is positive V_o is very small, while if V_i is negative or low (such that for PMOS $V_{GD_p} > |V_T|$), $V_o = V_{DD}$. The action is that of opening and closing a switch by means of the input control voltage. However, because no current exists in either state the power dissipated by the transistors is virtually zero i.e., the power is consumed in CMOS only during the switching interval. The extremely small power dissipation of CMOS circuits is the major advantage of CMOS which is the cornerstone of most digital integrated circuits, where switching between 0, 1 is the basic digital operation.

When using CMOS in analog circuits the PMOS transistor provides an active load for NMOS transistors (controlled source) (Fig. 7.23). The small signal model is shown (Fig. 7.23c). Notice the absence of $g_{m_2} \upsilon_{gs_2}$ since $\upsilon_{gs_2} = 0$ since the source and gate are at constant potential. The load resistance is r_{ds_2}. The PMOS transistor provides the dc resistance and the ac load resistor. Such resistances requires smaller area than conventional resistors.

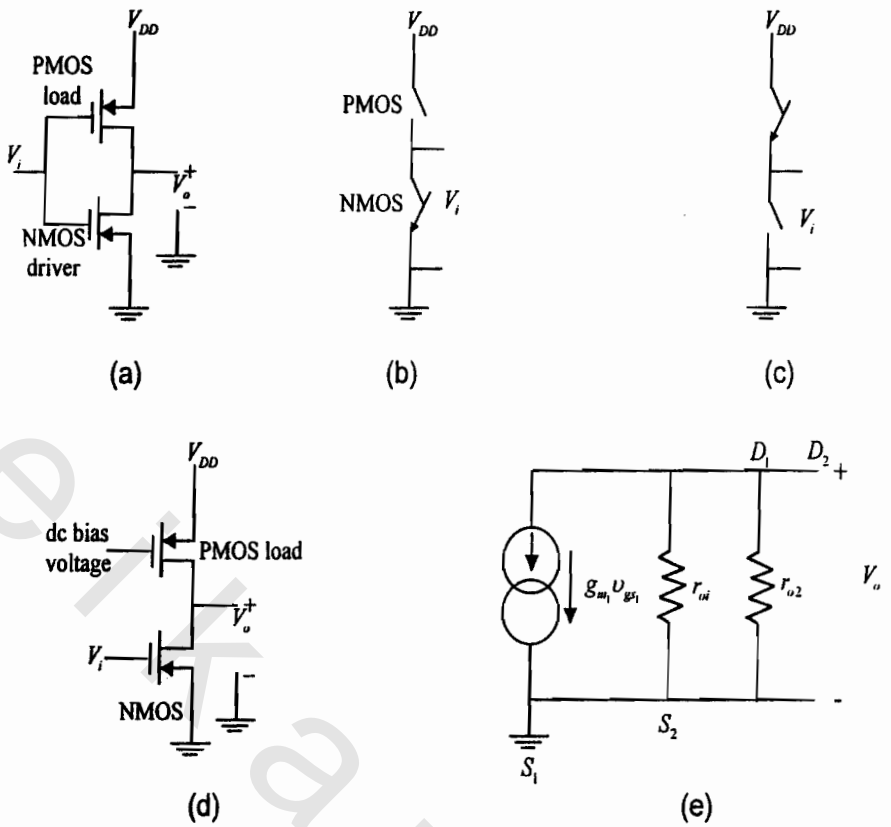


Fig. (7.23) CMOS

a) circuit b) PMOS switch OFF, NMOS switch ON
c) PMOS switch ON, NMOS switch OFF d) analog CMOS
e) equivalent circuit

Ex. 7.4

Consider CMOS common source amplifier shown $V_{DD} = 3\text{V}$, $|V_{T_n}| = V_{T_p} = 1\text{V}$, $\mu_{ns} C'_{ox} = 125\mu\text{A/V}^2$, $L_n = 0.5\mu\text{m}$, $L_p = 5\mu\text{m}$, $\mu_{ps} C'_{ox} = 50\mu\text{A/V}^2$, $I_{ref} = 100\mu\text{m}$. Find the small signal voltage gain and main points on the voltage transfer characteristic. Take $r_{o_1} = 300\text{k}\Omega$ and $r_{o_2} = 200\text{k}\Omega$

Solution

In the circuit of Fig. (7.28a), Q_2 and Q_3 are matched Fig. (7.28b) shows the $I_D - V_{SD}$ characteristic of the p channel transistor Q_2 for constant V_{SG} The value of V_{SG} is determined by passing the reference bias current I_{ref} through Q_3. Note that Q_2 is in saturation and behaves as a current source, since $V_{SD} > \left(V_{SG} - |V_{T_p}|\right)$.

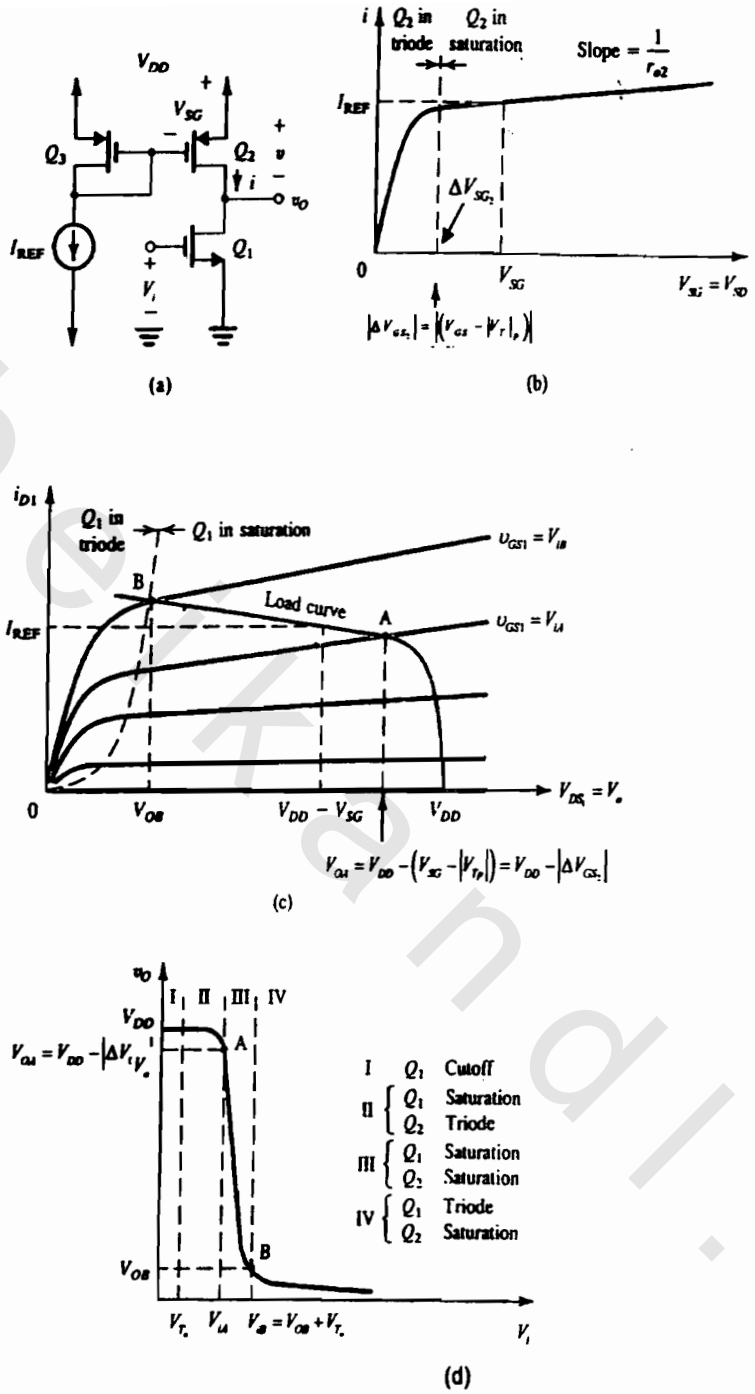


Fig. (7.24) CMOS common source amplifier

a) circuit b) characteristic of active load Q_2

c) graphical construction to determine the transfer characteristic

d) transfer characteristic

Fig. (7.28c) shows the $I_D - V_{DS_1}$ characteristic of Q_1 with the load curve of Q_2. Since $V_{GS_1} = V_i$ each of the $I_D - V_{DS}$ curves corresponds to a value of V_i. The intersection of each curve with the load curve gives $V_{DS_1} = V_o$ which yields the transfer characteristic (Fig. 7.28d). Region III is linear and has large voltage gain. In this region both Q_1 and Q_2 are saturated. At point A, Q_1 enters triode region

$$V_{o_A} = V_{DD} - \left(V_{SG} - \left|V_{T_p}\right|\right) = V_{DD} - \left|\Delta V_{SG_2}\right|$$

At B, $V_o = V_i - V_{T_n}$, Q_1 enters triode region

At any point within region III the small signa l voltage gain can be determined by replacing Q_1 with its small signal model and Q_2 with its output resistance r_{o2}

$$A_v = \frac{v_o}{v_i} = -g_{m_1} \frac{r_{o1} r_{o2}}{r_{o1} + r_{o2}}$$

Now

$$g_{m_1} = \sqrt{2k_n' \left(\frac{W}{L}\right)_1 I_{ref}}$$

$$= \sqrt{2 \times 125 \times \left(\frac{5}{0.5}\right) 100} = 0.25\text{mA/V}$$

$$A_v = -g_{m_1} \frac{300 \times 200}{500}$$

$$= -0.25 \times \frac{6}{5} \times 100 = -30$$

To determine V_{SG} of Q_2 and Q_3 corresponding to $I_D = I_{ref}$ noting that for

$$Q_3,\ V_{SD} = V_{SG} = V_{T_p} + \Delta V_{SG}$$

$$I_D = \frac{1}{2} k_p' \left(\frac{W}{L}\right)_3 \left(V_{SG} - \left|V_{T_p}\right|\right)^2$$

$$100 = \frac{1}{2} \times 50 \times 10 \times \left(\Delta V_{SG_3}\right)^2$$

where ΔV_{SG_3} is the overdrive voltage for both Q_2 and Q_3

$$\left|\Delta V_{SG_3}\right|^2 \simeq \frac{200}{500} = 0.40$$

$$\left|\Delta V_{SG_3}\right| = 0.64\text{V}$$

$$V_{SG_3} = 1.64\text{V}$$

$$V_{o_A} = V_{DD} - \Delta V_{SG_3} = 3 - 0.64 = 2.36\text{V}$$

We may obtain an expression for V_o as a function of V_i by equating $I_{D_1} = I_{D_2}$ both in saturation (Prob.7.16)

But we can conclude that

$$\frac{v_o}{v_i} = -30$$

Note that the slope of the transfer characteristic in region III is the voltage gain.

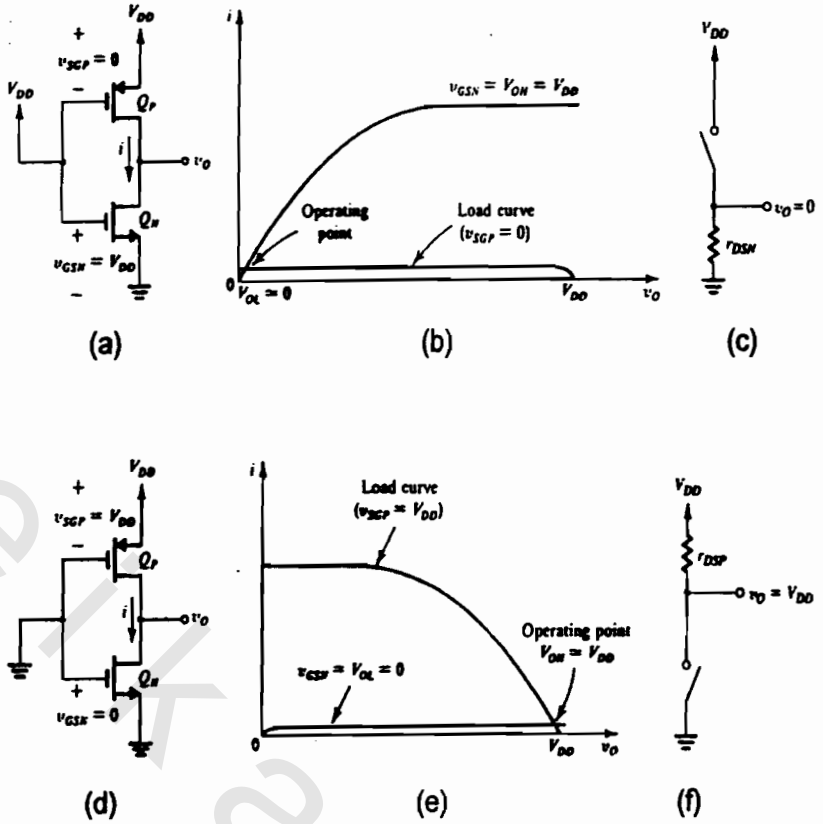


Fig. (7.25) Operation of a CMOS inverter

a) circuit with $V_i = V_{DD}$ b) graphical representation with $V_i = V_{DD}$

c) equivalent circuit for $V_i = V_{DD}$ d) circuit with $V_i = 0$

e) graphical representation for $V_i = 0$ f) equivalent circuit for $V_i = 0$

7.11 CMOS Inverter

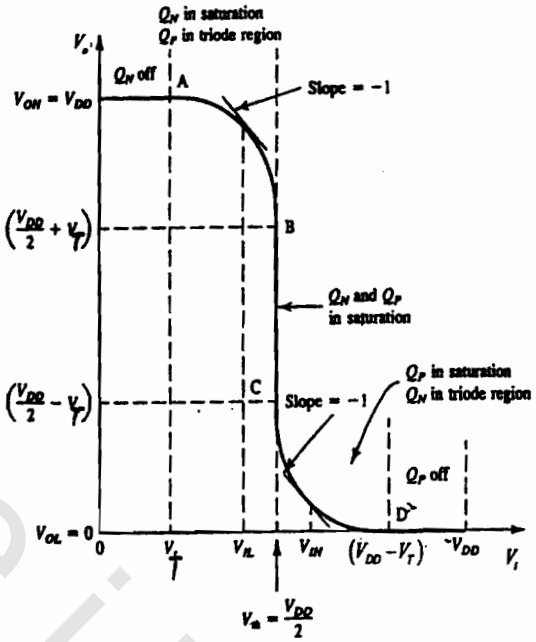
The basic in CMOS inverter is shown in Fig. (7.25a), where V_i at logic 1 i.e., $V_i = V_{DD} = V_{GS_n}$, $V_o = V_{DS_n}$. The load curve of Q_p is superimposed on the IV characteristic of Q_N for the case $V_{SG_p} = 0 < |V_{T_p}|$. Thus, this load curve is a straight line at zero current load.

The operating point is at the intersection of the two curves. Thus, the output voltage is nearly zero and the current is zero. This means that the power dissipation is nearly zero. The part of the $I_D - V_D$ curve for Q_n is nearly linear, thus, Q_n provides a low resistance path between the output terminal and ground.

$$r_{DS_n} = \frac{1}{k_n' \left(\frac{W}{L}\right)_n \left(V_{DD} - V_{T_n}\right)} \qquad (7\text{-}106)$$

Where $V_i = 0$, Q_n is operating at $V_{GS_n} = 0$, hence its $I_D - V_{DS}$ characteristic is almost a horizontal line at zero level.

a)



b)

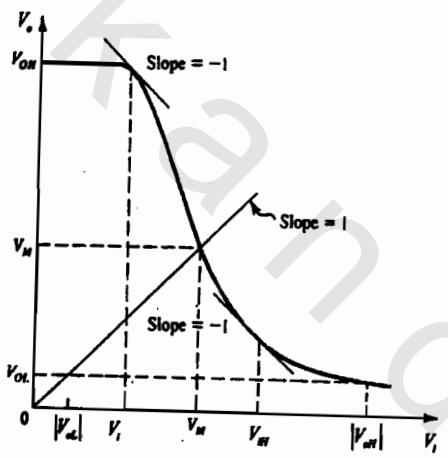


Fig. (7.26) Voltage transfer characteristic of CMOS inverter
a) critical points defined b) actual characteristic

The load curve of Q_p is shown for $V_{SG_p} = V_{DD}$. The operating point is almost at V_{DD} and $I_D = 0$ and the power dissipation is zero. Q_p provides a low resistance path

$$r_{DS_p} = \frac{1}{k'_p\left(\frac{W}{L}\right)_p}\frac{1}{\left(V_{DD} - V_{T_p}\right)} \qquad (7\text{-}107)$$

During switching Q_N can sink a large current, discharging a load capacitance, thus pulling the output voltage down toward zero. Therefore, Q_n is called pull down devices. Similarly, with the input Q_p can

source a large current charging up a load capacitance, thus, pulling the output voltage up toward V_{DD}. Hence, Q_p is called pull up device.

The voltage transfer characteristic is shown in Fig. (7.26).

For Q_n in the triode region

$$I_{D_n} = k'_n\left(\frac{W}{L}\right)_n\left[\left(V_i - V_{T_n}\right)V_o - \frac{1}{2}V_o^2\right] \text{ for } V_o \geq V_i - V_{T_n} \tag{7-108}$$

and in the saturation region

$$I_{D_n} = \frac{1}{2}k'_n\left(\frac{W}{L}\right)_n\left(V_i - V_{T_n}\right)^2 \text{ for } V_o \geq V_i - V_{T_n} \tag{7-109}$$

For Q_p in the triode region

$$I_{Dp} = k'_p\left(\frac{W}{L}\right)_p\left[\left(V_{DD} - V_i - \left|V_{T_p}\right|\right)\left(V_{DD} - V_o\right) - \frac{1}{2}\left(V_{DD} - V_o\right)^2\right]$$

for $V_o \geq V_i + \left|V_{T_p}\right|$ (7-110)

and in saturation

$$I_{Dp} = \frac{1}{2}k'_p\left(\frac{W}{L}\right)_p\left(V_{DD} - V_i - \left|V_{T_p}\right|\right)^2 \tag{7-111}$$

Usually $V_{T_n} = \left|V_{T_p}\right|$ and $k'_p\left(\frac{W}{L}\right)_n = k'_p\left(\frac{W}{L}\right)_p$ which requires $\frac{W_p}{W_n} = \frac{\mu_{ns}}{\mu_{ps}}$ which will result in equal and symmetrical current driving capability (pull up and pull down).

In the transfer characteristic the segment BC has both Q_n and Q_p in saturation. Because we neglect the finite output resistance BC is vertical (infinite slope), thus the gain is infinite. The vertical segment has a center at $V_i = \frac{V_{DD}}{2}$ and is bounded by $V_o(B) = \frac{V_{DD}}{2} + V_T$ and $V_o(C) = \frac{V_{DD}}{2} - V_T$.

To determine the noise margins of the inverter we denote two points V_{iL}, V_{iH} where V_{iL} is the maximum permitted logic 0 (low) level at the input and V_{iH} is the minimum permitted logic 1 (high) level at the input. These are defined as the two points on the transfer curve at which the incremental gain is unity (slope -1). To determine V_{IH} we note that Q_n is in the triode region and Q_p is in saturation. Thus,

$$\left(V_i - V_T\right)V_o - \frac{1}{2}V_o = \frac{1}{2}\left(V_{DD} - V_i - V_T\right)^2 \tag{7-112}$$

Differentiating relative to V_i

$$\left(V_i - V_T\right)\frac{dv_o}{Av_i} + V_o - V_o\frac{dv_o}{Av_i} = -\left(V_{DD} - V_i - V_T\right) \tag{7-113}$$

Now $V_i = V_{iH}$ and $\frac{dv_o}{dv_i} = -1$

$$V_o = V_{iH} - \frac{V_{DD}}{2} \tag{7-114}$$

Substituting eqn. (7-114) into eqn. (7-113) given

$$V_{iH} = \frac{1}{8}(5V_{DD} - 2V_T) \quad (7\text{-}115)$$

From symmetry

$$V_{iH} - \frac{V_{DD}}{2} = \frac{V_{DD}}{2} - V_{iL} \quad (7\text{-}116)$$

$$V_{iL} = \frac{1}{8}(3V_{DD} + 2V_T) \quad (7\text{-}117)$$

The margin of safety for an inverter in state (1) or state (0) is manifested when we visualize the output of an inverter being a drive input to another inverter. Thus, if the output of the driving inverter is V_{oH} (high) the difference between V_{oH} and V_{iH} is the noise margin. The driven inverter will not be affected so long as the voltage at its input does not full because of noise below V_{iH}. Thus, the noise margin for high state

$$NM_H = V_{oH} - V_{iH} \quad (7\text{-}118)$$

Similarly, if the output of the driving inverter is low at V_{oL} the driven inverter will provide a high state even if the noise disturbs V_{oL} level at the input raising it up to near V_{iL}. The noise margin at the low level is

$$NM_L = V_{iL} - V_{oL} \quad (7\text{-}119)$$

Using eqn. (7-116), (7-117)

$$NM_H = V_{oH} - V_{iH}$$

$$= V_{DD} - \frac{1}{8}(5V_{DD} - 2V_T)$$

$$= \frac{1}{8}(3V_{DD} + 2V_T) \quad (7\text{-}120)$$

$$NM_L = V_{iL} - V_{oL}$$

$$= \frac{1}{8}(3V_{DD} + 2V_T) \quad (7\text{-}121)$$

which confirms the symmetry

7.12 CMOS Logic Gates

The CMOS inverter consists of a pull down NMOS transistor and a pull up PMOS transistor. The CMOS logic gates consist of two networks, the pull down network (PDN) consisting of the NMOS transistors, and the pull up network (PUN) consisting of PMOS transistors. For three input gates shown (Fig. 7.27) the PDN will conduct for all input combinations that require $Y = 0$ and will pull the output down to ground, Simultaneously PUN will be off and V_{DD} is shut off from the ground path.

On the other hand if inputs are applied for which $Y = 1$ PUN will conduct which will pull the output node up to V_{DD}, PDN being shut off. Thus, PDN is activated when the inputs are high, and PUN is activated when the inputs are low.

In Fig. (7.28a) Q_A or Q_B conduct when A or B is high (V_A or $V_B = V_{DD}$) and will pull the output node down to ground $Y = 0$. Thus, $\overline{Y} = A + B$ The PDN in Fig. (7.28b) will conduct when A and B are both high simultaneously. Thus, $\overline{Y} = AB$. In Fig. (7.28c) $\overline{Y} = A + BC$.

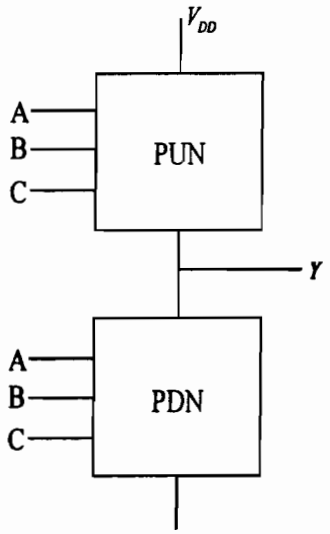


Fig. (7.27) CMOS logic gates comprrising PDN and PUN

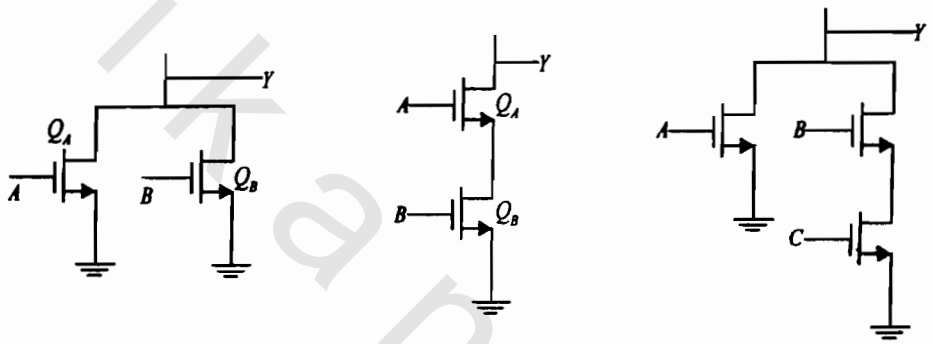


Fig. (7.28) Examples of PDN circuits

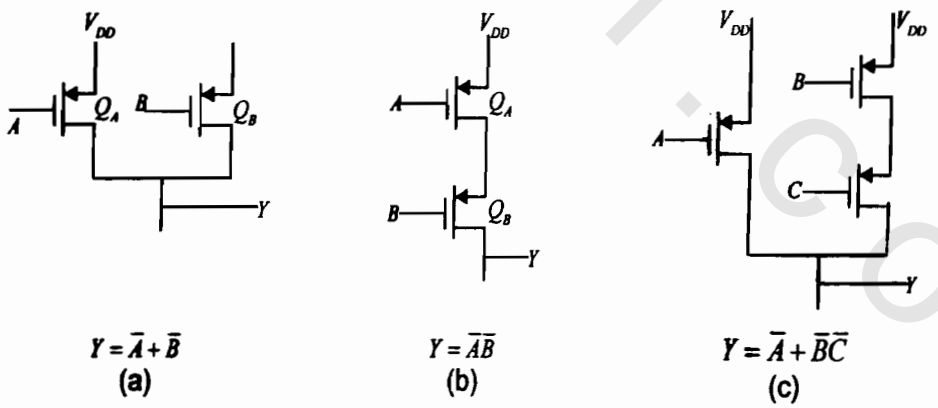


Fig. (7.29) Examples of PUN circuits

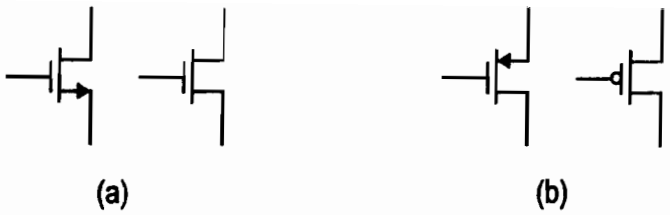
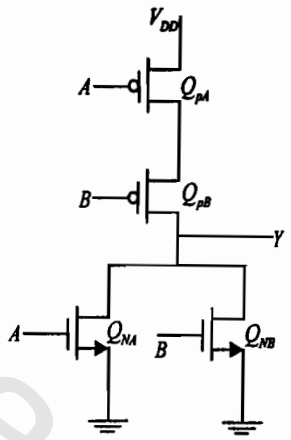


Fig. (7.30) Symbols for MOSFETs in CMOS structure
a) NMOS b) PMOS



$$Y = \overline{A+B} = \overline{A}\,\overline{B}$$

Fig. (7.31) CMOS NOR gate

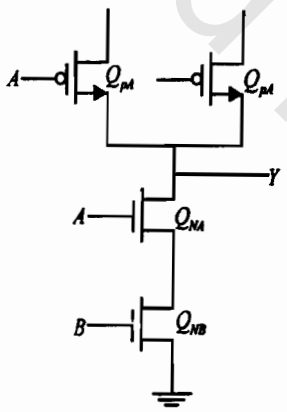


Fig. (7.32) CMOS NAND gate

Considering PUN circuits (Fig. 7.29) Y will be pulled up when $Y = \bar{A} + \bar{B}$ in Fig. (7.29b) and $Y = \bar{A}\bar{B}$ in Fig. (7.29b) and $Y = \bar{A} + \bar{B}\bar{C}$ in Fig. (7.29c)

We need now to introduce new symbols (Fig. 7.30). Not that the circle in the PMOS symbol indicates active low, i.e., it needs low voltage to be activated (conduct). The arrow is removed in the symbol since it is known that current flows from PMOS to NMOS, i.e., for NMOS D is the side closer to the PUN, while for PMOS, S is the side closer to V_{DD}. It should be noted that duality exists between PUN and PDN. Alternatively, we can also synthesize Y directly for PUN and $\bar{Y}$ directly for PDN (Prob. 7.18).

7.13 Transistor Sizing

The ON resistance of a MOSFET is inversely proportional to $\frac{W}{L}$ (eqn. 7-39). Thus, if a number of MOSFETs having ratios of $\left(\frac{W}{L}\right)_1, \left(\frac{W}{L}\right)_2, \left(\frac{W}{L}\right)_3$, .. are connected in series the equivalent series resistance is given by

$$R_{series} = r_{D_1} + r_{D_2} \tag{7-122}$$

$$= \frac{const}{\left(\frac{W}{L}\right)_1} + \frac{const}{\left(\frac{W}{L}\right)_2} + \frac{const}{\left(\frac{W}{L}\right)_3}$$

$$= \frac{const}{\left(\frac{W}{L}\right)_{eq}} \tag{7-123}$$

$$\left(\frac{W}{L}\right)_{eq} = \frac{1}{\frac{1}{\left(\frac{W}{L}\right)_1} + \frac{1}{\left(\frac{W}{L}\right)_2} + \frac{1}{\left(\frac{W}{L}\right)_3}} \tag{7-124}$$

Similarly we see that for parallel connection of transistors

$$\left(\frac{W}{L}\right)_{eq} = \left(\frac{W}{L}\right)_1 + \left(\frac{W}{L}\right)_2 + \left(\frac{W}{L}\right)_3 \tag{7-125}$$

Also it is desired to ensure symmetry to Q_n and Q_p to ensure balancing of charge and discharge capacitances and hence delay times. This is called transistor sizing.

$$k_n'\left(\frac{W}{L}\right)_n = k_p'\left(\frac{W}{L}\right)_p \tag{7-126}$$

$$\frac{\left(\frac{W}{L}\right)_p}{\left(\frac{W}{L}\right)_n} = \frac{\mu_{n_s}}{\mu_{p_s}} \tag{7-127}$$

Ex. 7.5

Show how transistor sizing is employed for the shown circuit.

Solution

In the four input NOR circuit shown the worst case i.e., the lowest current for PDN is obtained when only one of the NMOS transistors is conducting we therefore select $\frac{W}{L}$ for each NMOS transistor equal to that of the NMOS transistor of the basic inverter $\left(\frac{W}{L}\right)_n$. For the PUN, however, the worst situation and the only valid case is when all inputs are low and the four series PMOS transistors are conducting. The equivalent $\left(\frac{W}{L}\right)$ will be quarter of that for each PMOS device. We should select $\frac{W}{L}$ for for each PMOS transistor to be four times that of Q_p of the basic inverter, i.e., $4\left(\frac{W}{L}\right)_p$

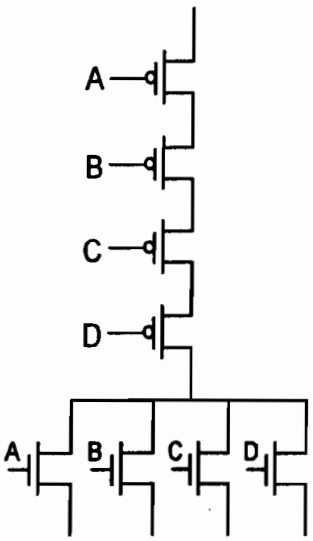


Fig. Ex. (7.5)

Problems

1- For $L = 0.4\mu m$, $d_{ox} = 8nm$, $\mu_n = 450 cm^2/Vs$, $W = 8\mu m$, $V_T = 0.7V$ and $\epsilon_{ox} = 3.45 \times 10^{-11} F/m$, find:

a. C'_{ox}, k'_n and k_n

b. V_{GS}, V_{DS} needed to operate the transistor at saturation for $I_D = 100\mu A$

c. the overdrive voltage ($V_{GS} - V_T$) required to operate the transistor at saturation

d. V_{GS} needed to cause the device to operate as 1k resistor for small V_{DS}

e. r_{DS} if the overdrive is 0.5V

2- A NMOS has $C'_{ox} = 200\mu A/V^2$, $L = 0.9\mu m$ and $W = 18\mu m$. Find:

a. I_D if $\Delta V_{GS} = 0.5V$, $V_{DS} = 1V$

b. ΔI_D if V_{DS} is doubled

3- A PMOS has $V_T = -1V$, $k'_p = 50\mu A/V^2$ and $W/L = 12$, Find:

a. range for which the transistor conducts

b. range of V_{DS} for triode operation

c. range of V_{DS} for saturation

d. ΔV_{GS} for all above cases

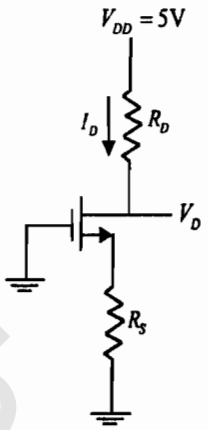
4- For $L = 0.6\mu m$, $d_{ox} = 15nm$, $\mu_n = 550 cm^2/Vs$. Find C'_{ox}, k'_n and the overdrive required to operate the transistor with $W/L = 20$ in saturation, if $I_D = 0.2mA$. Find the minimum value of V_{DS}

5- Find an expression for r_{DS} in the triode region of NMOS. Then calculate r_{DS} if $W/L = 20$, $k'_n = 50\mu A/V^2$ and the overdrive voltage is 1V

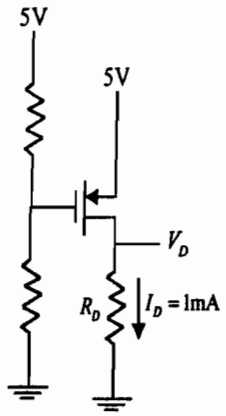
6- An NMOS has $V_T = 1V$ and 2V dc is applied to the gate. In what region does the transistor operate if $V_{DS} = 0.5V$, $V_{DS} = 1V$, $V_{DS} = 2V$ and $V_{DS} = 3V$?

7- An NMOS has $V_T = 0.5V$ and current $i_D = 150\mu A$ when $V_{GS} = V_{DS}$. Find the value of I_D for $V_{GS} = 2V$, $V_{DS} = 3V$. Also calculate the value of r_{DS} for small V_{DS} and $V_{GS} = 3.5V$

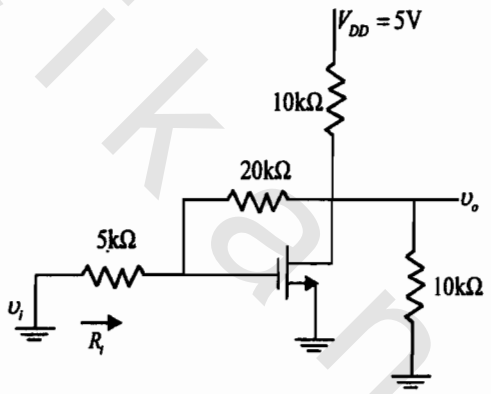
8- In a PMOS $V_T = -1V$, $k'_p = 60\mu A/V^2$ and $W/L = 15$. Find the range of V_{DS} for which the transistor conducts. Then find the range of V_{DS} for which the transistor operates in saturation. then find the value of ΔV_{GS} and V_{DS} to keep PMOS in saturation with $I_D = 75mA$.



Prob. (7.9)



Prob. (7.10)



Prob. (7.12)

9- The transistor shown has $I_D = 0.5mA$, $V_o = 0.6V$ NMOS has $V_T = 1V$, $\mu_n C'_{ox} = 150\mu A/V^2$, $L = 1\mu m$ and $W = 40\mu m$. Calculate the resistances R_D, R_S.

10- Design the circuit shown so that $I_D = 1mA$, $V_D = 3V$, $V_T = -1V$, and $k'_p \frac{W}{L} = 1mA/V^2$. What is the largest value of R_D to maintain saturation?

11- Referring to Figs. (7.14), (7.15) determine all points on the curves numerically for $k'_p \frac{W}{L} = 1mA/V^2$, $V_T = 1V$, $R_D = 20k\Omega$, $V_{DD} = 10V$.

12- In the circuit shown determine $\frac{V_o}{V_i}$ and the input resistance R_i and the largest input signal. $V_T = 1V$, $k'_n \frac{W}{L} = 0.5mA/V^2$. Also determine the biasing currents and voltages.

13- In the circuit of Fig. (7.17), calculate all currents and voltages if all transistors are identical $k'\frac{W}{L}=1\text{mA/V}^2$, $V_T=1\text{V}$, $I_{ref}=1\text{mA}$, $V_{DD}=V_{SS}=10\text{V}$, $R_G=10\text{M}\Omega$, $R_D=15\text{M}\Omega$. What is the maximum swing of the input such that the MOSFET is in saturation.

14- Analyze the common source amplifier with source resistance, common gate, common drain circuits, for A_v, R_i, R_o. Show the advantages and usage of each configuration.

15- Referring to Fig. (7.15) find the condition for maximum symmetrical swing.

16- Referring to Ex. (7.4) obtain the relation between V_o and V_i in the saturation region for both transistors Q_1, Q_2 and compare the results with A_v.

17- Derive the small signal equivalent circuit for a CMOS amplifier in which the input is applied across the common gate of both transistors.

18- Design a CMOS circuit and verify the truth table for the following case using direct synthesis and then using duality.

a) $Y=A(B+CD)$ b) $Y=A\bar{B}+\bar{A}B$

19- Discuss the effect of sizing in the above problem

20- In Fig. (7.16), (7.17) a constant current is used for biasing. How can drain current be allowed to vary as an ac signal while a constant current is used?

References

1- "Microelectronic Circuit Design", R. Jaeger, T. Blalock, 3rd edition, Mc Graw Hill, N. Y., 2008.

2- "Microelectronics, Circuits Analysis and Design", D. Neamen, 3rd edition, Mc Graw Hill, N. Y., 2007.

3- "Microelectronic Circuits", A. Sedra and K. Smith, 5th edition, Oxford University Press, N. Y., 2004.

4- "Electronic Devices and Circuit Theory", R. Boylestad, L. Nashelsky, 8th edition, Prentice Hall, Upper Saddle River , N. Y., 2002.

5- "Analysis and Design of Analog Integrated Circuits", P. Gray, P. Hurst S. Lewis, R. Meyer, Wiley, 4th edition, N. Y., 2001.

6- "Design of Analog CMOS Integrated Circuits", B. Razavi, Mc Graw Hill, N. Y., 2001.

INDEX