Chapter 4: the MOS transistor

1. Introduction

First products in Complementary Metal Oxide Silicon (CMOS) technology appeared in the market in seventies. At the beginning, CMOS devices were reserved for logic, as they offer the highest density (in gates/mm^2), and the lowest static power consumption. Most high-frequency circuitry was carried out in bipolar technology. As a result, a lot of analog functions were realized in bipolar technology. The technology development, which is driven by digital circuits (in particular by flash memories), lead to smaller and faster CMOS devices. At the beginning of the seventies, 1µm transistors length was considered short. Currently, CMOS technology with 22nm channel length is available. In the last twenty years a lot of analog circuits started to be developed in CMOS technology. In fact, the technology scaling enabled CMOS devices at higher frequencies of working, also for the analog counterpart.

Today, CMOS and bipolar technologies are in competition over a wide frequency region up to 100GHz. The challenge indeed, to choice the technology that fulfills best the system and circuit requirements at a reasonable cost. Bipolar is more expensive than standard CMOS technology. Moreover, most systems and circuits are mixed signal, i.e. they include digital and analog parts. In the past, separated integrated circuits were dedicated to the analog (bipolar) and digital (CMOS) circuits. As analog circuits were also available in CMOS technology, this technology started to offer the opportunity to integrate cheap, high density and low power digital circuits, as well as analog circuits, in the same chip. This brings enormous advantages in terms of reduced costs and smaller form factors of electronic devices. Currently, CMOS technology dominates the market. Bipolar transistors field of applications is reduced to very high frequencies of working.

2. The MOS transistor structure

Fig. 4.1 shows a cross section of a typical n-type MOS (NMOS) transistor.

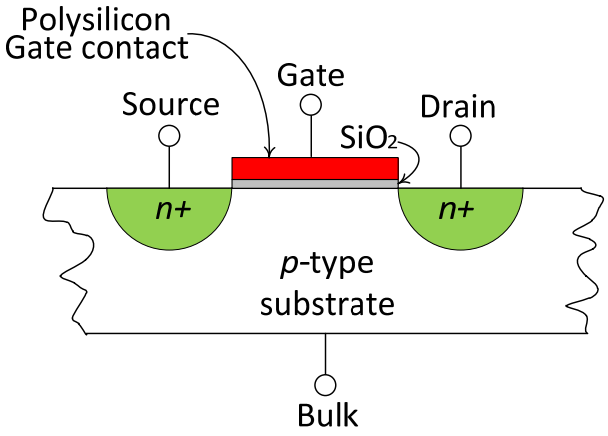

Fig. 4.1 Cross section of a typical NMOS transistor.

Fig. 4.2 shows the top view of a typical NMOS transistor. The main transistors dimensions, i.e. the channel length (L) and width (W), are indicated.

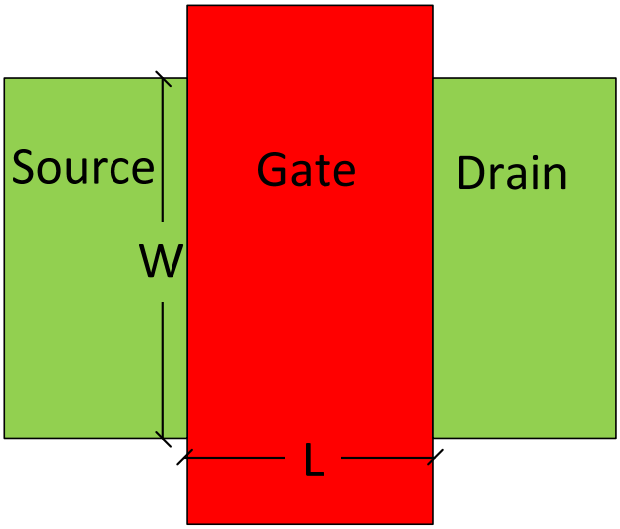

Fig. 4.2 Top view of a typical NMOS transistor.

The two deeply n-type doped active areas of Source and Drain are fabricated in a p-type substrate. A thin layer of silicon dioxide is growth in the region between the Source and the Drain. A conductive material (generally polysilicon) covers the silicon dioxide, implementing the Gate terminal. The Gate terminal regulates the current conduction between Source and Drain.

The Bulk terminal biases the n-type substrate, which is common to all NMOS transistors. The Bulk terminal is set to the lowest voltage available for the circuit, generally the ground. This makes the p-n junction between the active areas of Source and Drain, and the substrate, inversely biased. Practically, the two active areas result electrically isolated, as well as all NMOS transistors on the same substrate. Current conduction between Source and Drain is possible only when an opportune voltage is applied to the Gate.

Complementary MOS (CMOS) technology includes both NMOS and PMOS transistors. PMOS transistor has p-type deeply doped active areas of Source and Drain fabricated in a n-type substrate. In a standard CMOS technology, PMOS transistors is built in a N-well obtained in a p-type substrate. Fig. 4.3 shows a NMOS and a PMOS devices integrated on the same silicon die.

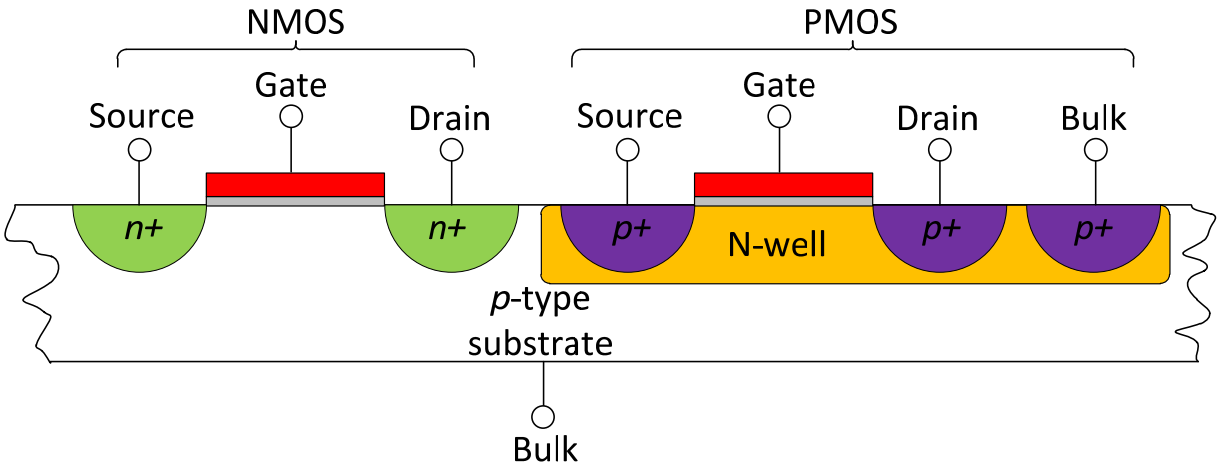

Fig. 4.3 NMOS and PMOS transistors fabricated on the same silicon die.

In digital gates, the Bulk terminals of N-Wells of PMOS transistors are connected to the supply. However, since the N-Well is isolated from the rest of the substrate, its Bulk terminal can be connected to a voltage different with respect to the supply. In analog circuits, the Bulk terminal of PMOS transistors is often connected to the Source, in order to reduce the threshold voltage shift due to the Body effect. This effect occurs when Source and Bulk are not biased to the same voltage. While this effect can be always mitigated in PMOS transistors, it cannot be avoided in NMOS transistors when the Source is connected to a voltage different from the ground, where its Bulk is bounded.

Fig. 4.4 shows the NMOS and the PMOS transistors symbols used in CMOS circuits. When the Bulk terminal has a standard connection (i.e. to the ground and to the supply for NMOS and PMOS transistors, respectively) it is not drawn.

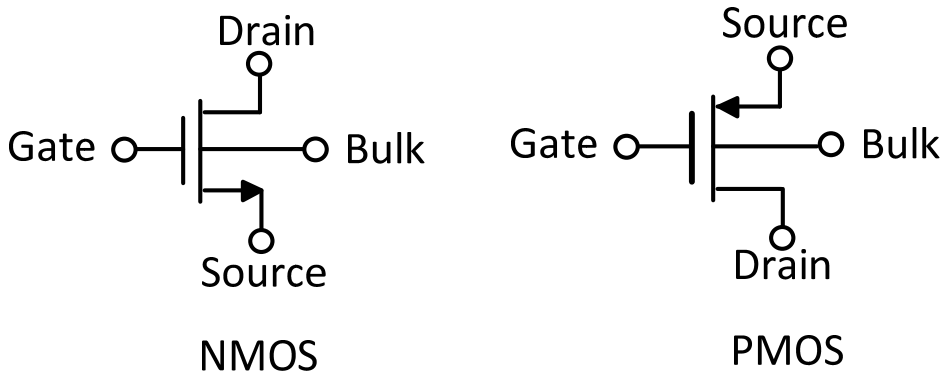

Fig. 4.4 NMOS and PMOS transistors symbols used in CMOS circuits.

3. Large signal behavior of MOS transistors

The NMOS transistor is a strongly nonlinear device. Its transfer characteristics depends on the bias conditions. In order to study the NMOS transistor behavior, four regions of operation are distinguished:

- cut-off region;
- linear or triode region;
- saturation region;
- weak inversion.

In the following, these four regions of operation and the NMOS secondary effects are analyzed in details.

3.1 Interdiction region

To derive the transfer characteristics of the NMOS transistor let start considering all its terminals grounded, as Fig. 4.5 shows.

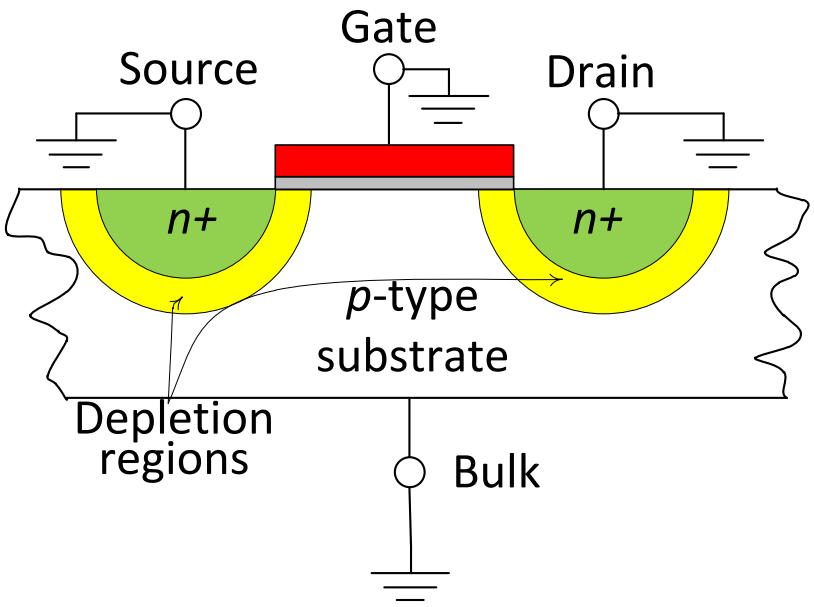

Fig. 4.5 NMOS transistor with all its terminals grounded.

The n+ islands of Source and Drain are completely enveloped by a depletion layer, as it happens in p-n junctions. If the Gate-Source voltage (V_{GS}) is null, the Source and the Drain regions are separated by a back-to-back p-n junctions. The circuit is equivalent to that one reported in Fig. 4.6.

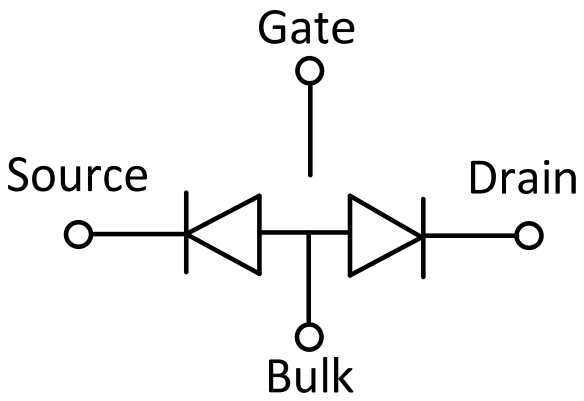

Fig. 4.6 Equivalent circuit of an NMOS transistor with all terminals biased to ground.

Gate terminal results floating. Diodes of Fig. 4.6 represent junctions formed by the n-type Source and Drain regions and the p-type substrate. Since these junctions are inversely biased, they behave like an extremely high resistance. Device is off, as current conduction between Source and Drain cannot occur. This region of operation is called cut-off region.

Now consider the Bulk, the Source and the Drain grounded, while a positive V_{GS} is applied to the Gate, as Fig. 4.7 shows.

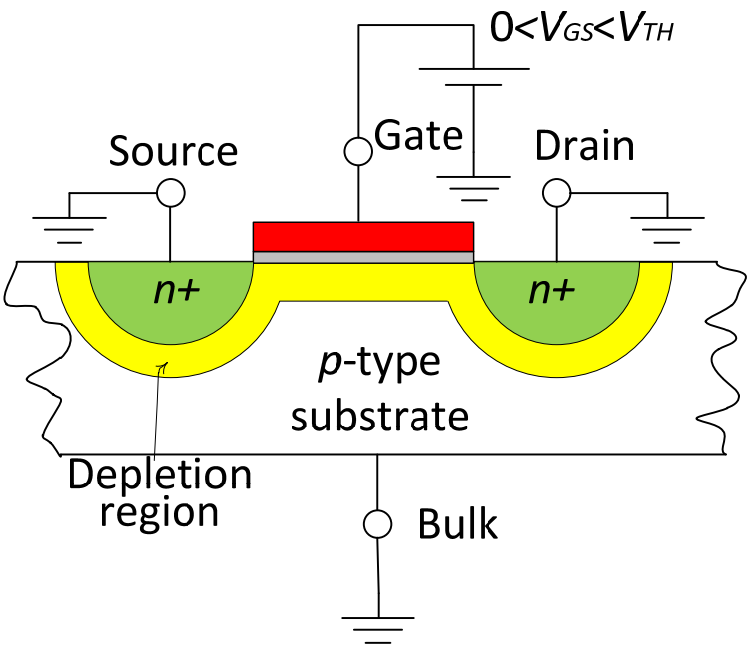

Fig. 4.7 NMOS transistor with a $0<V_{GS}<V_{TH}$.

The Gate and the substrate form a capacitor with the silicon dioxide as dielectric. Positive charges are accumulated on the Gate, while negative charges are attracted in the substrate. Initially, negative charges accumulated in the substrate are manifested by the creation of a depletion region under the channel, that excludes holes under the Gate. However, in this conditions, current cannot flows between Source and Drain, and the device is still in cut-off region.

3.2 Linear or triode region

As V_{GS} reaches a critical value called threshold voltage, V_{TH}, a thin layer of electrons at the interface between silicon dioxide and substrate is induced. The layer of electrons forms a conducting channel between Source and Drain. This phenomena is known as inversion.

When $V_{GS}>V_{TH}$, the NMOS transistor is on, i.e. the conducting channel is formed, then a current flow between Source and Drain can exist.

In order to have a current between Source and Drain, a positive V_{DS} has to be applied. The NMOS transistor situation is illustrated by Fig. 4.8.

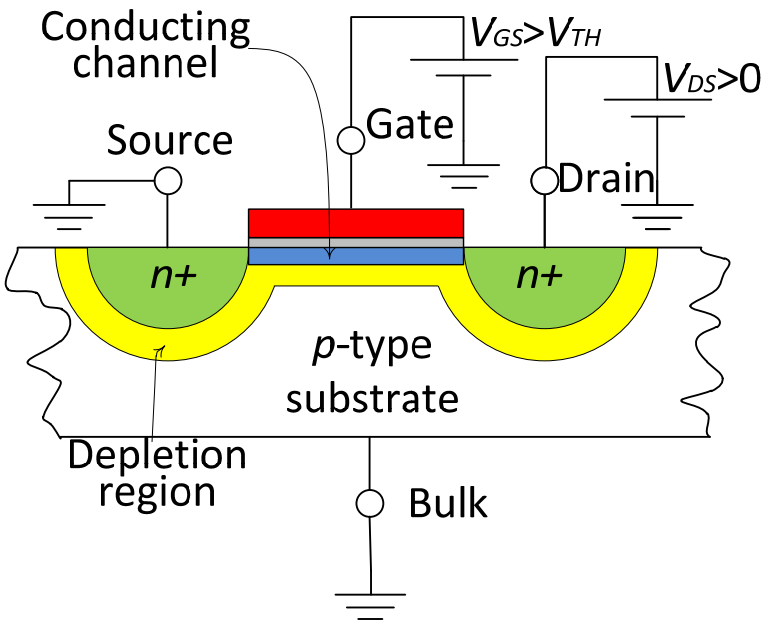

Fig. 4.8 NMOS transistor with a $V_{GS}>V_{TH}$, and $V_{DS}>0$.

The positive V_{DS} produces a horizontal electric field that makes the channel charge, Q_{ch}, flowing between Source and Drain for drift effect. The resulting I_{DS} current can be calculated as follows:

Eq. 4.1
$$I_{DS} = \frac{N \cdot q}{t_d} = \frac{Q_{ch}}{t_d}$$

where N is the total number of electron composing the channel charge, q is the electron charge, and t_d is the drift time required to cross the channel by an electron. Q_{ch} is modulated by the V_{GS} voltage that exceeds the threshold voltage, V_{TH}, i.e. V_{GS}-V_{TH} which is called the overdrive voltage, V_{ov}. V_{ov} produces a vertical electric field which generates electrons accumulation in the conducting channel. Q_{ch} can be calculated as follows:

Eq. 4.2
$$Q_{ch} = C_{ox} \cdot L \cdot W \cdot (V_{GS} - V_{TH}) = C_{ox} \cdot L \cdot W \cdot V_{ov}$$

where L and W are the channel length and width, respectively, and C_{ox} is the Gate capacitance per unit area. C_{ox} is given by the ratio between the permectivity, ε_{ox}, and thickness, t_{ox}, of the silicon dioxide:

Eq. 4.3
$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

In standard CMOS technologies of the last twenty years t_{ox} is about fifty times lower than the minimum channel length L_{min}:

Eq. 4.4
$$t_{ox} = \frac{L_{min}}{50}$$

Drift time, t_d, is directly proportional to the channel length L, and inversely proportional to the drift velocity of electrons, v_d, as equation 4.3 shows:

Eq. 4.5
$$t_d = \frac{L}{v_d}$$

Drift velocity, v_d, is proportional to the horizontal electric field, ε_y, i.e.:

Eq. 4.6 $$v_d = \mu_n \cdot \varepsilon_y$$

where μ_n is the electron mobility in the channel. The value of the horizontal electric field, ε_y, is, approximately, calculated as follows:

Eq. 4.7 $$\varepsilon_y = \frac{V_{DS}}{L}$$

Substituting equations 4.2, 4.3, 4.4, and 4.5, into equation 4.1 the following expression for I_{DS} is get:

Eq. 4.8 $$I_{DS} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot V_{DS} = 2 \cdot k_n \cdot V_{ov} \cdot V_{DS}$$

where k_n is called conductivity factor for NMOS transistors. It is given by:

Eq. 4.9 $$k_n = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L}$$

Equation 4.8 describes the transfer characteristics of a NMOS transistor, as V_{DS} is kept low, i.e. $V_{DS}<V_{GS}-V_{TH}$. In this bias condition, the NMOS transistor operates in the linear, or triode, region. Fig. 4.9 shows the I_{DS}-V_{DS} curves of the NMOS transistor operating in linear region, with V_{GS} as parameter.

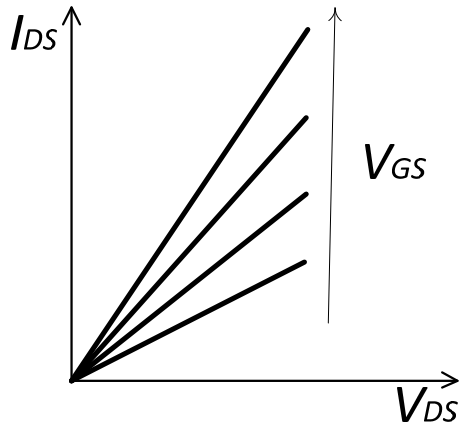

Fig. 4.9 I_{DS}-V_{DS} curves of the NMOS transistor in linear region with V_{GS} as parameter.

Equation 4.6 predicts that I_{DS} is proportional to V_{DS}. Practically, in linear region, the NMOS transistor acts like a variable resistance, R_{on}, whose value depends on V_{ov}. R_{on} is calculated as follows:

Eq. 4.10 $$R_{on} = \frac{1}{2 \cdot k_n \cdot (V_{GS} - V_{TH})} = \frac{1}{2 \cdot k_n \cdot V_{ov}}$$

As V_{DS} is increased, the charge channel narrows at the drain end. In fact, V_{DS} modifies the voltage, and, then, the charge along the channel. The voltage component due to V_{DS} is maximum at the drain end, while it is zero at the Source end. Supposing a linear distribution of the voltage component due to V_{DS} along the channel, its average value can be approximately estimated to be $V_{DS}/2$. Therefore, a more accurate calculation of the channel charge is get by adding $V_{DS}/2$ in equation 4.2:

Eq. 4.11 $$Q_{ch} = C_{ox} \cdot L \cdot W \cdot \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2}\right)$$

Considering equation 4.11, and repeating the same steps than before, a more accurate calculation of the I_{DS} current in linear region is also possible:

Eq. 4.12 $$I_{DS} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2}\right) \cdot V_{DS} = 2 \cdot k_n \cdot \left(V_{ov} - \frac{V_{DS}}{2}\right) \cdot V_{DS}$$

I_{DS} values obtained by adopting equations 4.9 and 4.12, are very close when V_{DS} is much less than V_{ov}.

3.3 Saturation region

As V_{DS} approaches V_{GS}-V_{TH}, the channel charge approaches zero at the Drain end. In facts, the channel charge is sustained by a Gate-Drain voltage, V_{GD}, at the drain side, which is less than V_{GS} at the Source side. When V_{DS} compensates for the overdrive voltage V_{GS}-V_{TH}, V_{GD} results to be equal to the threshold voltage V_{TH}, as equation 4.13 predicts:

Eq. 4.13 $$V_{GD} = V_{GS} - V_{DS} = V_{GS} - (V_{GS} - V_{TH}) = V_{TH}$$

In this bias condition, no channel charge is available at the Drain end. The conducting channel pinches-off, disconnecting the Drain. Fig. 4.10 shows the NMOS transistor situation. The V_{DS} value that produces the channel pinch-off is called saturation voltage ($V_{DS,sat}$). It corresponds to V_{GS}-V_{TH}, i.e the overdrive voltage.

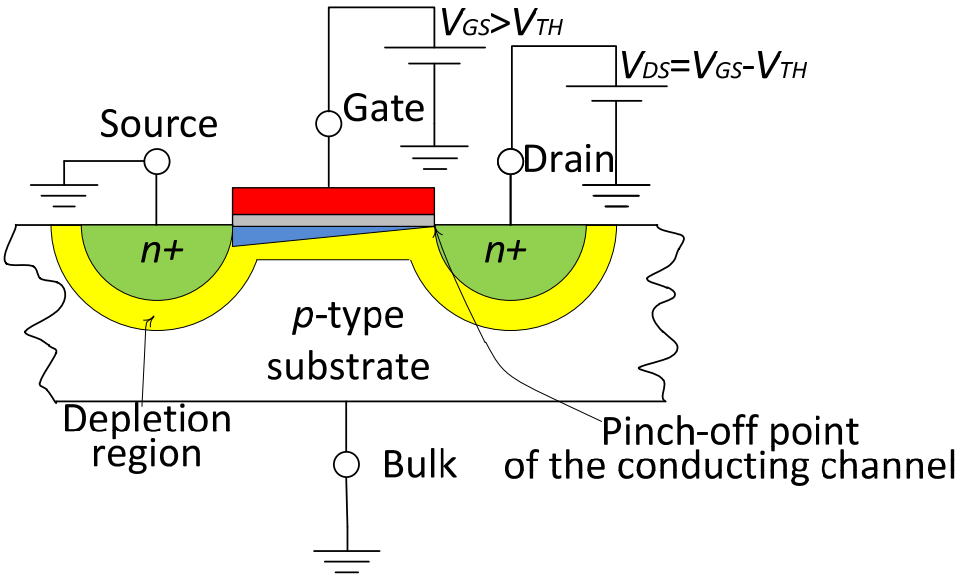

Fig. 4.10 NMOS transistor with a $V_{GS}>V_{TH}$, and $V_{DS}=V_{GS}-V_{TH}$.

Substituting the V_{DS} value in equation 4.10, the expression of the I_{DS} at the pinch-off is obtained:

Eq. 4.14 $$I_{DS} = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot {V_{ov}}^2$$

As V_{DS} increases over $V_{GS} - V_{TH}$, a depletion region is formed between the pinch-off point and the Drain. Since the voltage between the Gate and the pinch-off point is V_{TH} by definition, the V_{DS} part that exceeds $V_{GS} - V_{TH}$, stands across this depletion region. Therefore, when V_{DS} increases, this depletion region enlarges, and the pinch-off point moves toward the Source. Fig. 4.11 shows the NMOS transistor status.

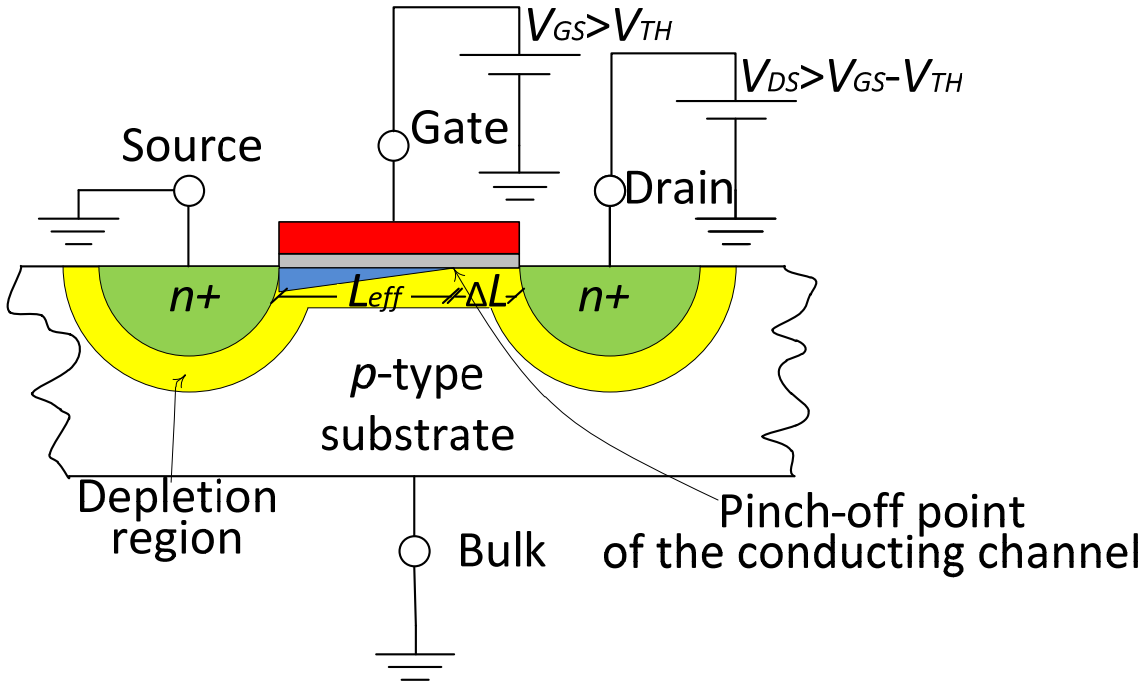

Fig. 4.11 NMOS transistor with a $V_{GS}>V_{TH}$, and $V_{DS}>V_{GS}-V_{TH}$.

Further increments of V_{DS} over $V_{GS}-V_{TH}$ do not produce modifications on the voltage across channel region, therefore the horizontal electric field is kept constant. Thus, the I_{DS} current, which is due to the drift effect of the channel charge by the horizontal electric field, remains equal to that one expressed by equation 4.12. Equation 4.12 predicts that I_{DS} depends only on V_{ov}, not on V_{DS}, when the conducting channel pinches-off. This region of operation is called saturation region.

Fig. 4.12 shows the I_{DS}-V_{GS} input characteristic of the NMOS transistor in saturation region. The I_{DS}-V_{GS} curve fit the square law expressed in equation 4.12, and it meets the V_{GS} axis at V_{TH}.

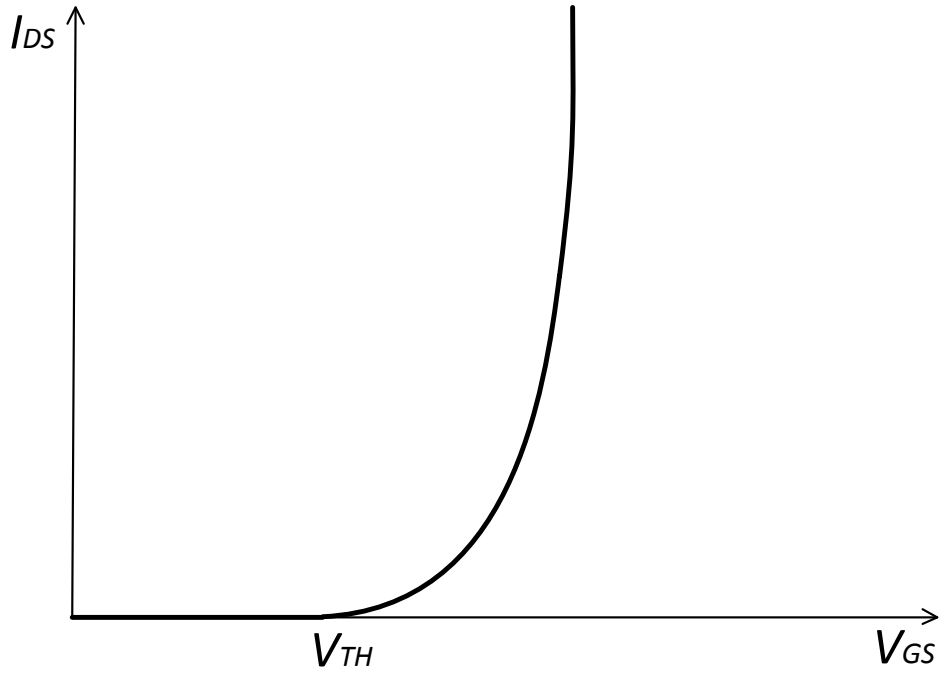

Fig. 4.12 I_{GS}-V_{GS} input characteristic of the NMOS transistor in saturation region.

Fig. 4.13 shows the ideal I_{DS}-V_{DS} output characteristics with V_{GS} as parameter. These curves do not includes the channel modulation effect.

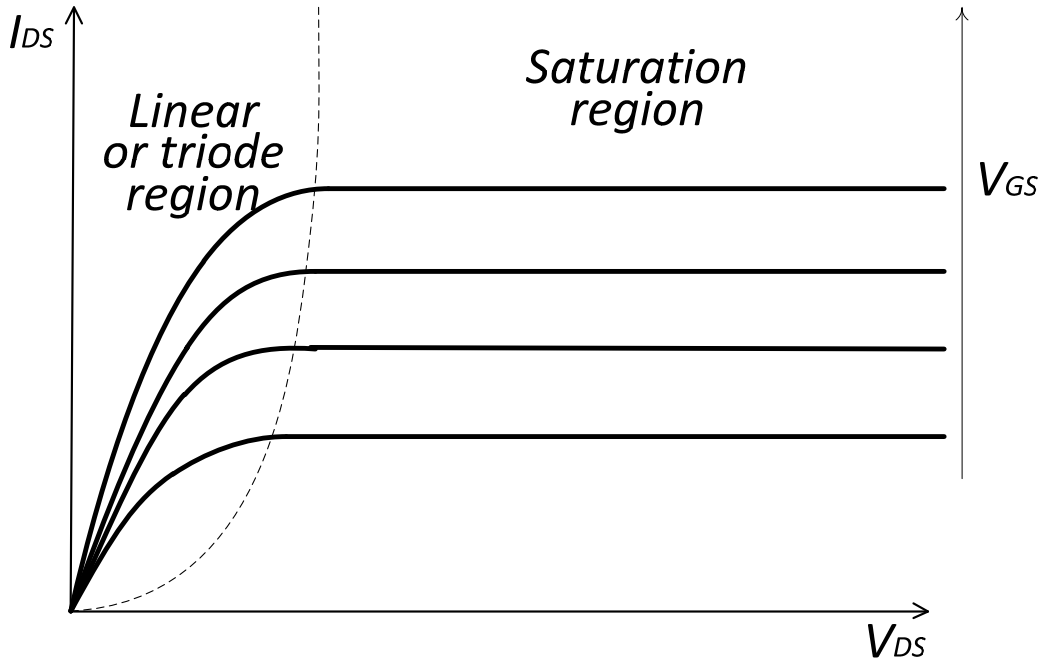

Fig. 4.13 Ideal I_{DS}-V_{DS} curves of the NMOS transistor with V_{GS} as parameter.

Since the depletion region length ΔL augments as V_{DS} increases, the length of the region containing the channel charge, i.e. the effective channel length L_{eff}, is reduced. In facts:

Eq. 4.15 $$L_{eff} = L - \Delta L$$

A more accurate expression of I_{DS}, includes L_{eff}:

Eq. 4. 16 $$I_{DS} = \frac{1}{2}\cdot\mu_n\cdot C_{ox}\cdot\frac{W}{L_{eff}}\cdot(V_{GS} - V_{TH})^2$$

Because L_{eff} depends on V_{DS}, also I_{DS} changes with V_{DS} in the saturation region. This effect is known as channel length modulation. I_{DS} expression can be rearranged as follows:

Eq. 4.17 $$I_{DS} = \frac{1}{2}\cdot\mu_n\cdot C_{ox}\cdot\frac{W}{L_{eff}}\cdot(V_{GS} - V_{TH})^2 = \frac{1}{2}\cdot\mu_n\cdot C_{ox}\cdot\frac{W}{L}\cdot(V_{GS} - V_{TH})^2\cdot\frac{L}{L-\Delta L}$$

According to equation xxx the depletion region length ΔL is calculated as follows:

Eq. 4.18 $$\Delta L = \sqrt{\frac{2}{\varepsilon_{si}\cdot q\cdot N_A}\cdot\left(V_{DS} - V_{DS,sat}\right)}$$

Therefore:

Eq. 4.19 $$\frac{L}{L-\Delta L} = \frac{1}{1-\sqrt{\frac{2}{\varepsilon_{si}\cdot q\cdot N_A\cdot L^2}\cdot\left(V_{DS}-V_{DS,sat}\right)}} \cong 1 + \frac{1}{\varepsilon_{si}\cdot q\cdot N_A\cdot L^2}\cdot\left(V_{DS} - V_{DS,sat}\right)$$

Substituting equation 4.19 in 4.16, and neglecting $V_{DS,sat}$ with respect to V_{DS}, the following expression for I_{DS} is obtained:

Eq. 4.20 $$I_{DS} = \frac{1}{2}\cdot\mu_n\cdot C_{ox}\cdot\frac{W}{L}\cdot(V_{GS} - V_{TH})^2\cdot(1 + \lambda\cdot V_{DS}) = k_n\cdot V_{ov}^2\cdot(1 + \lambda\cdot V_{DS})$$

where λ, which is called channel modulation parameter, is equal to:

Eq. 4.21 $$\lambda = \frac{1}{\varepsilon_{si}\cdot q\cdot N_A\cdot L^2}$$

Fig. 4.14 shows the I_{DS}-V_{DS} curves of the NMOS transistor with channel modulation effect and V_{GS} as parameter.

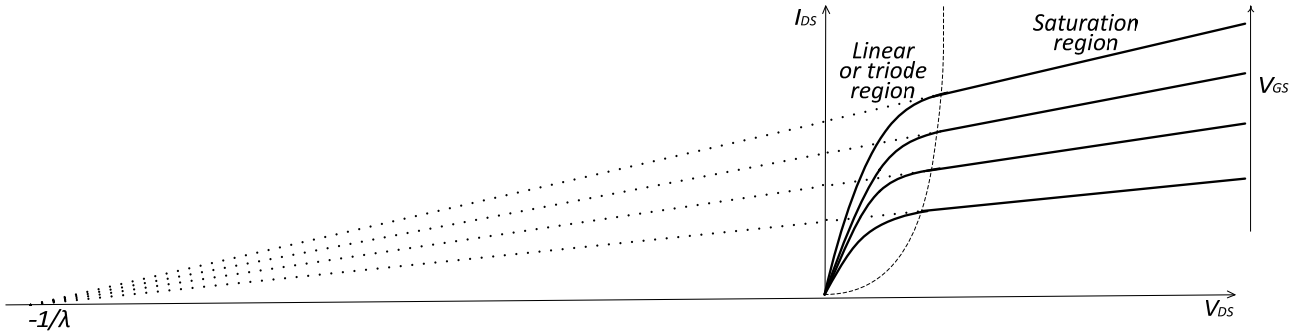

Fig. 4.14 I_{DS}-V_{DS} curves of the NMOS transistor with channel modulation effect and V_{GS} as parameter.

By extrapolating the curves in saturation region, all of them meet the V_{DS} axis in the same point equal to -1/λ. The inverse of λ , 1/λ, can be compared to the Early voltage, V_A, in bipolar transistors.

3.4 Weak inversion

The passage from the off to the on state of the NMOS transistor is not drastic. Supposing the NMOS transistor in off state (i.e. $V_{GS}<V_{TH}$), as V_{GS} approaches V_{TH}, if V_{DS} is non-null, a small I_{DS} current starts flowing.

When V_{GS} is around V_{TH}, the transistor operates in a region called weak inversion.

Fig. 4.15 shows the NMOS transistor situation.

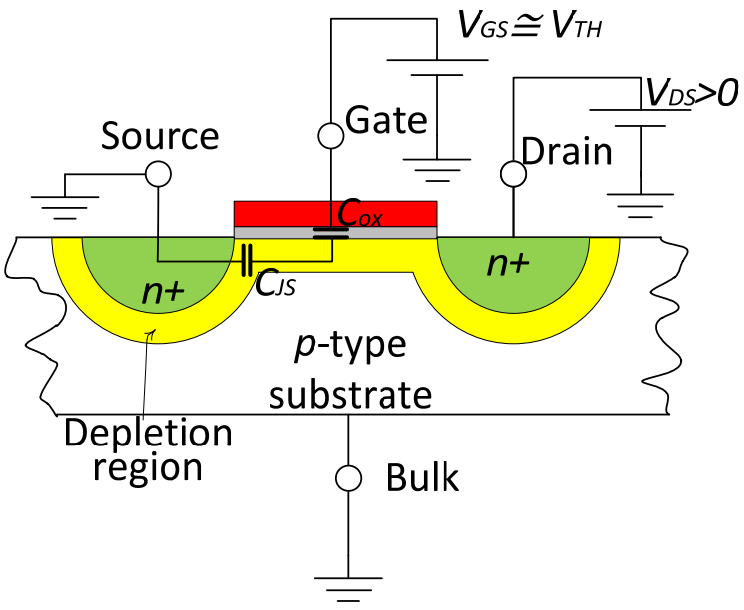

Fig. 4.15 NMOS transistor with a $V_{GS}\cong V_{TH}$, $V_{DS}>0$.

In this condition the conducting channel is not completely formed. Since there is not much charge, current cannot flows for drift but diffusion. Practically, the NMOS transistor behaves like a NPN bipolar transistor, with the Source and the Drain regions corresponding to the Emitter and the Collector, respectively, while the p-type substrate under the silicon dioxide corresponds to the Base. Substituting I_C and V_{BE} with I_{DS} and V_{sur} respectively, in equation 3.xxx, I_{DS} can be calculated as follows:

Eq. 4.22 $$I_{DS} = I_S \cdot e^{\frac{Vsur}{Vt}}$$

where I_S is the process dependent inverse saturation current, V_{sur} is the voltage at the surface between the silicon dioxide and the p-type substrate, and V_t is the thermic voltage.

The oxide capacitance C_{ox} stands between the Gate and the p-type substrate under the silicon dioxide, and the p-type substrate and Source junction forms a capacitance C_{JS}. The surface voltage V_{sur}, can be calculated by considering the capacitive divider made up of C_{ox} and C_{JS}, as follows:

Eq. 4.23 $$V_{sur} = V_{GS} \cdot \frac{C_{ox}}{C_{ox}+C_{JS}} = \frac{V_{GS}}{n}$$

where n is called slope factor. It ranges between 1 and 2. It is calculated as follows:

Eq. 4.23_1 $$n = 1 + \frac{C_{JS}}{C_{ox}}$$

Substituting equation 4.23 in equation 4.22 the following expression of the I_{DS} current in weak inversion is obtaining:

Eq. 4.24 $$I_{DS} = I_0 \cdot e^{\frac{V_{GS}}{n \cdot Vt}}$$

Fig. 4.16 shows the I_{DS}-V_{DS} output characteristics including also the weak inversion region.

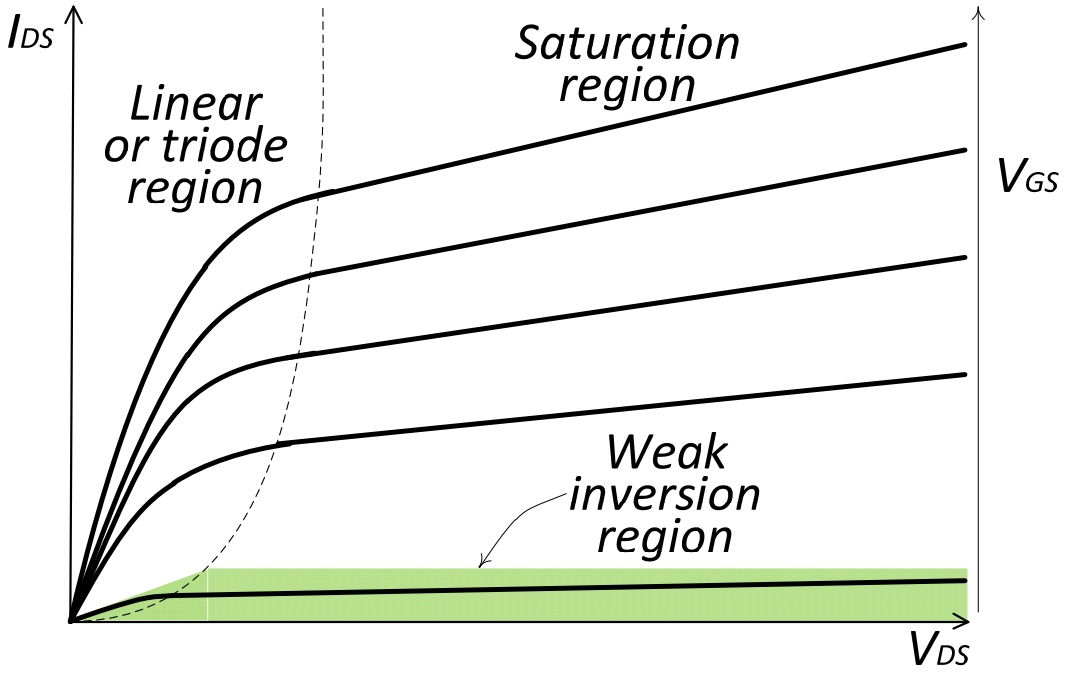

Fig. 4.16 I_{DS}-V_{DS} curves of the NMOS transistor including the weak inversion region.

Table 4.1 summarizes the NMOS transistor characteristic equations for different regions of operation, and the bias conditions required to have them.

Region of operation	Characteristic equation	V_{GS}	V_{DS}
Cut-off	$I_{DS}=0$	$<V_{TH}$	-
Linear or triode	$I_{DS} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2}\right)$	$>V_{TH}$	$<V_{GS}- V_{TH}$
Saturation	$I_{DS} = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2 \cdot (1 + \lambda \cdot V_{DS})$	$>V_{TH}$	$>V_{GS}- V_{TH}$
Weak inversion	$I_{DS} = I_0 \cdot e^{\frac{V_{GS}}{n \cdot Vt}}$	$\cong V_{TH}$	>0 V

Table 4.1 Summary of the NMOS transistor characteristic equations for different regions of operation

- **Example 4.1**

Problem. Consider the NMOS in Fig. 4.17. Assume μ_n=600$\frac{cm}{V^2 \cdot s}$, C_{ox}=$\frac{4fF}{\mu m^2}$, V_{TH}=0.5V, W=10um, L=1um, V_G=1V, V_{DD}=5V. Find the I_{DS} current.

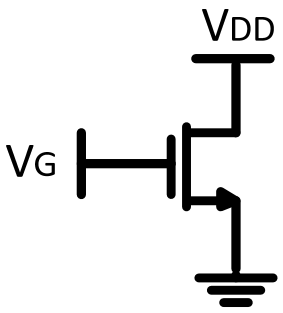

Fig. 4.17 NMOS transistor of the example 4.1.

Solution. Evaluate V_{GS} and compare to V_{TH}:

Eq. 4.25 $$V_{GS} = V_G = 1V > 1V = V_{TH}$$

As V_{GS} is more than V_{TH}, the NMOS transistor is on. Now calculate V_{DS} and compare it to V_{DSsat}:

Eq. 4.26 $$V_{DS} = V_{DD} = 5V > 0.5V = V_{GS} - V_{TH} = V_{ov} = V_{DSsat}$$

As V_{DS} is more than V_{DSsat}, the NMOS transistor operates in saturation region. Note that V_{DSsat} is equal to the transistor overdrive V_{ov}. Now evaluate the transistor conductivity factor k_n:

Eq. 4.27 $$k_n = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} = 1.2 \frac{mA}{V^2}$$

Now, according to equation 4.14, calculate the I_{DS} current:

Eq. 4.28 $$I_{DS} = k_n \cdot V_{ov}^2 = 300 \mu A$$

-

3.5 The threshold voltage

The threshold voltage, V_{TH}, is the voltage required to form the conducting channel under the silicon dioxide. In order to examine V_{TH}, let's consider Source and Drain grounded, and a negative V_{BS} and a

positive V_{GS} voltages applied to the Bulk and the Gate, respectively. Fig. 4.17 shows the situation of the NMOS transistor.

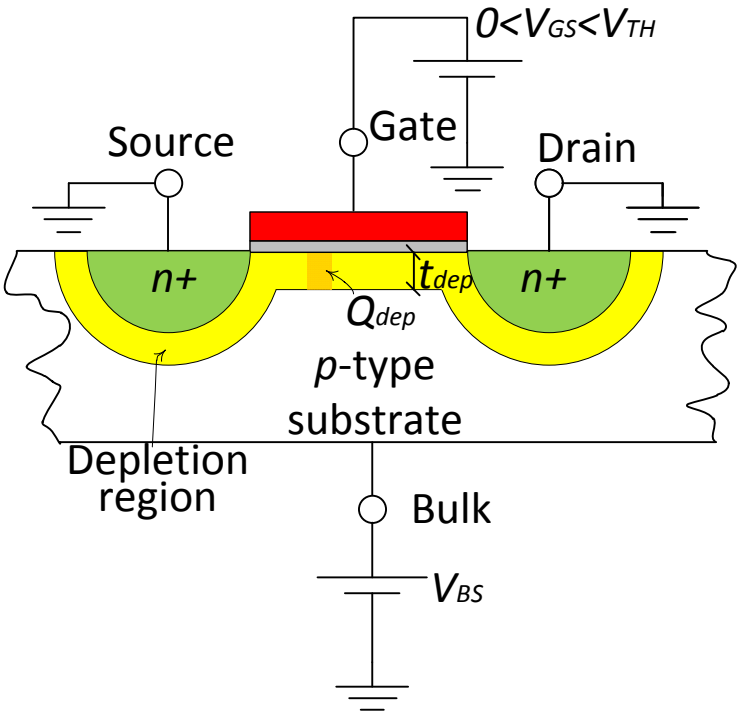

Fig. 4.18 NMOS transistor with a $0<V_{GS}<V_{TH}$, $V_{BS}<0V$, and Source and Drain grounded.

V_{TH} consists of different contributions. Firstly, V_{TH} is required to sustain depletion layer charge. The depletion layer charge per unit area $Q_{dep,}$ is proportional to the charge per unit volume, $q \cdot N_A$, and the depletion region depth, t_{dep}, under the silicon dioxide:

Eq. 4.29
$$Q_{dep} = q \cdot N_A \cdot t_{dep}$$

As the inversion layer channel charge starts forming, the voltage at the surface silicon dioxide-substrate is about twice the Fermi level ϕ_f . According to eq. 2.xxx, t_{dep} is calculated as follows:

Eq. 4.30
$$t_{dep} = \sqrt{\frac{2 \cdot \varepsilon_{si} \cdot (2 \cdot \phi_f - V_{BS})}{q \cdot N_A}}$$

The minimum t_{dep} and, then the minimum depletion layer charge per unit area, Q_{dep0}, is obtained for V_{BS} null. In this bias condition, also the minimum threshold voltage, V_{TH0}, is get:

Eq. 4.31
$$V_{TH0} = \phi_{MS} + 2 \cdot \phi_f + \frac{Q_{dep0}}{C_{ox}} - \frac{Q_{SS}}{C_{ox}} = \phi_{MS} + 2 \cdot \phi_f + \frac{\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{si} \cdot 2 \cdot \phi_f}}{Cox} - \frac{Q_{SS}}{C_{ox}}$$

where ϕ_{MS} is the work-functions difference between the Gate polysilicon and the silicon substrate, and Q_{SS} is a positive charge density due to imperfections of the silicon crystal at the surface between silicon dioxide and p-type substrate.

The calculation of the threshold voltage for a non-null V_{BS} is reported in the following:

Eq.4.32
$$V_{TH} = \phi_{MS} + 2 \cdot \phi_f + \frac{Q_{dep}}{C_{ox}} - \frac{Q_{SS}}{C_{ox}} = \phi_{MS} + 2 \cdot \phi_f + \frac{Q_{dep0}}{C_{ox}} - \frac{Q_{SS}}{C_{ox}} + \frac{Q_{dep} - Q_{dep0}}{C_{ox}} = V_{TH0} + \gamma \cdot \left(\sqrt{2 \cdot \phi_f - V_{BS}} - \sqrt{2 \cdot \phi_f}\right)$$

where:

Eq. 4.33 $$\gamma = \frac{\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{si}}}{Cox}$$

Equation 4.28 predicts the dependency of V_{TH} on V_{BS}. This phenomena is known as **Body effect**, and γ is called Body effect coefficient.

Fig. 4.18 shows the $\sqrt{I_{DS}} - V_{GS}$ curves of the NMOS transistor operating in saturation, with V_{BS} as parameter. As I_{DS} depends on the square of V_{GS} in saturation region, $\sqrt{I_{DS}}$ changes linearly with V_{GS}. $\sqrt{I_{DS}} - V_{GS}$ line intercepts the V_{GS} axes in a point equal to V_{TH}. This point is shifted to the right, i.e. the threshold voltage moves toward higher values, as V_{BS} decreases.

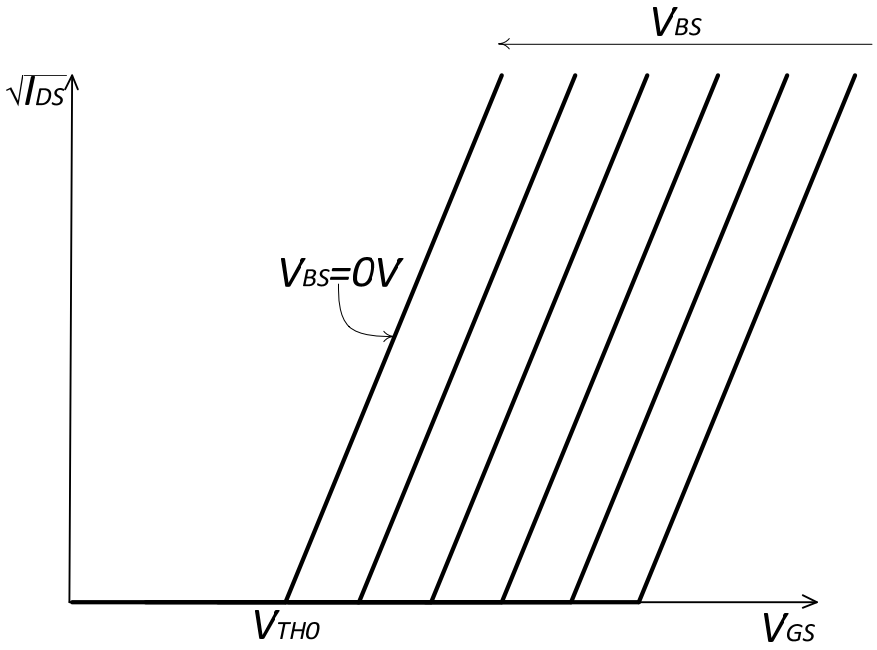

Fig. 4.19 $\sqrt{I_{DS}} - V_{GS}$ curves of the NMOS transistor with V_{BS} as parameter.

4. Large signal model of the NMOS transistor

In the last paragraph, transfer characteristics of the NMOS transistor have been get. However, due fabrication limitations, a number of passive elements must be taken under consideration to get a complete large signal model of the NMOS transistor.

Fig. 4.20 shows the cross section of a typical NMOS transistor with I_{DS} current source and parasitics. I_{DS} current source represents the I_{DS} current model.

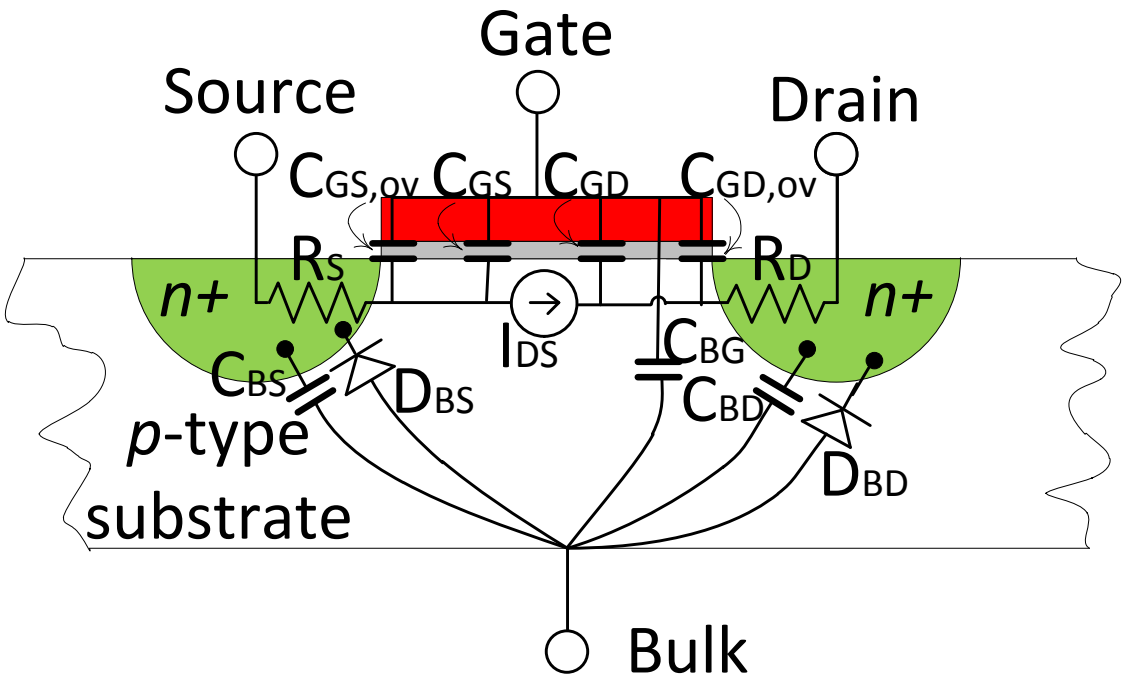

Fig. 4.20 Cross section of a typical NMOS transistor with I_{DS} current source and parasitics.

The two n+ islands of Source and Drain have a certain resistivity. Therefore, they give rise to contact resistances, represented by R_S and R_D. They values are limited to few ohms.

Zooming in between Source and Gate, as done in Fig. 4.21, a small region of overlap of the Gate over the Source is observed. The same happens at the Drain side.

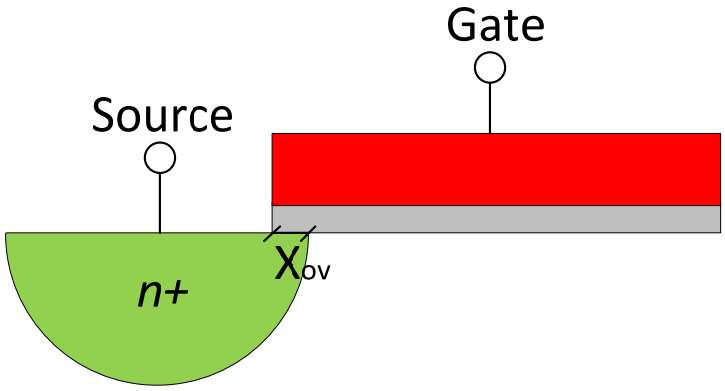

Fig. 4.21 Cross section of a typical NMOS transistor with I_{DS} current source and parasitics.

These overlap regions produce parasitic capacitances modeled by $C_{GS,ov}$ and $C_{GD,\ ov}$.. The values of these capacitances are calculated as follows:

Eq. 4.34 $$C_{GS,ov} = C_{GD,ov} = W \cdot X_{ov} \cdot C_{ox}$$

where X_{ov} is the length of the overlap region.

The contact resistances and the overlap capacitances have a linear behavior, i.e. their values do not depend on the bias condition of the NMOS transistor.

The NMOS transistor structure also includes two pn-junctions formed by the n+ Source and Drain islands and the p-type substrate. These pn-junctions are represented by D_{BS} and D_{BD} diodes. As the p-type substrate is biased at the minimum voltage available for the circuits, D_{BS} and D_{BD} are inversely biased. Therefore they are passed by an inverse current I_{GR}. According to equation 2.xxx I_{GR} value is calculated as follows:

Eq. 4.35 $$I_{GR} = A \cdot \frac{q \cdot n_i \cdot x_j}{2 \cdot \tau_0}$$

where A is the area of the pn-junction, x_j is the depletion region depth, n_i is the intrinsic carrier concentration, τ_0 is the minority carrier lifetime.

As described in 2.xxx, parasitic capacitances are associated with the depletion regions. Three depletion regions can be distinguished in the NMOS transistor structure: two depletion regions around the Source and the Drain island, and a third depletion region under the Gate. The parasitic capacitances associated with these depletion regions are represented by C_{BS}, C_{BG} and C_{BD}. These capacitances have a non -linear behavior, since their values depend on the voltage at their ends, as equation 2.xxx predicts.

The Gate, the silicon dioxide and the p-type substrate form the Gate capacitance, C_G, which is intrinsic to the operation of the NMOS transistor, since it is used to control the channel charge. As C_{ox} is the silicon dioxide capacitance per unit area, then $C_{ox} \cdot W \cdot L$ is the maximum value of the total Gate capacitance. Fig. 4.22 plots C_G over $C_{ox} \cdot W \cdot L$ versus V_{GS}.

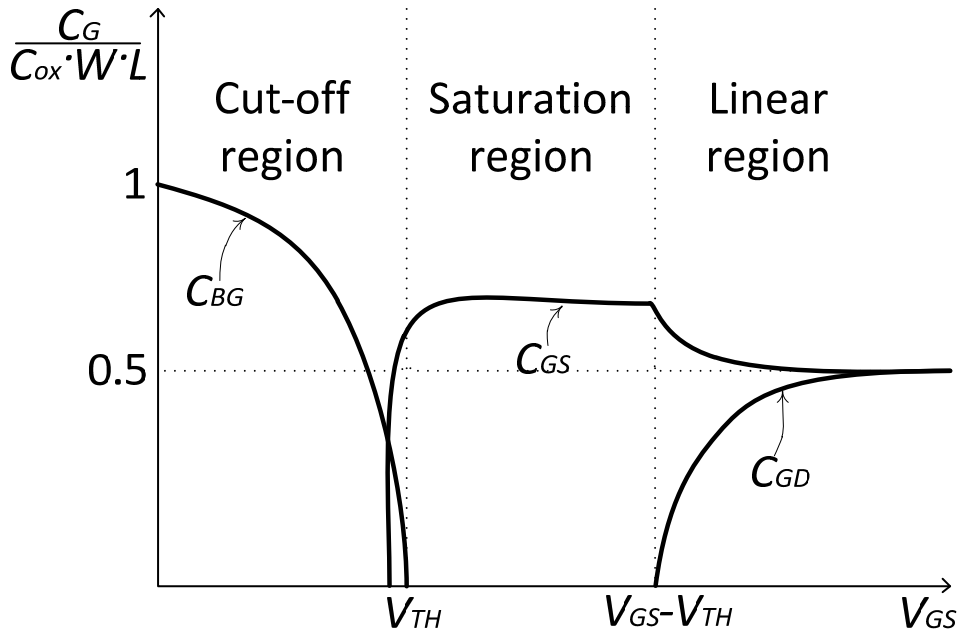

Fig. 4.22 C_G over $C_{ox} \cdot W \cdot L$ versus V_{GS}.

When the NMOS transistor is off, the conducting channel does not exist, and Source and Drain result disconnected. In this condition, the total Gate capacitance stands between Gate and Bulk, adding its contribution to C_{BG} which includes also the capacitance due to the depletion region under the Gate.

When the transistor is on, the conducting channel forms the second plate of the Gate capacitance. In saturation region, only the Source is connected with the conducting channel. Therefore most of the Gate capacitance stands between Gate and Source, forming the C_{GS} capacitance which is about $\frac{2}{3} \cdot C_{ox} \cdot W \cdot L$. The

Drain has a small influence on the channel charge, therefore the Gate-Drain capacitance C_{GD} can be neglected. In linear region, a continuous conducting channel is extended between Source and Drain. The Gate capacitance is divided between Source and Drain in similar parts. In fact, in linear region, C_{GD} and C_{GS} values are around $\frac{1}{2} \cdot C_{ox} \cdot W \cdot L$.

D_{BS} and D_{BD} diodes, C_{GS} and C_{GD}, and capacitances associated with the depletion regions (C_{BG}, C_{BD}, and C_{BS}) have a non-linear behavior, as well as the I_{DS} current.

5. The PMOS transistor

In CMOS technology, the PMOS transistor is fabricated in a N-well obtained in a p-type substrate. Fig. 4.23 shows a cross section of a typical PMOS transistor.

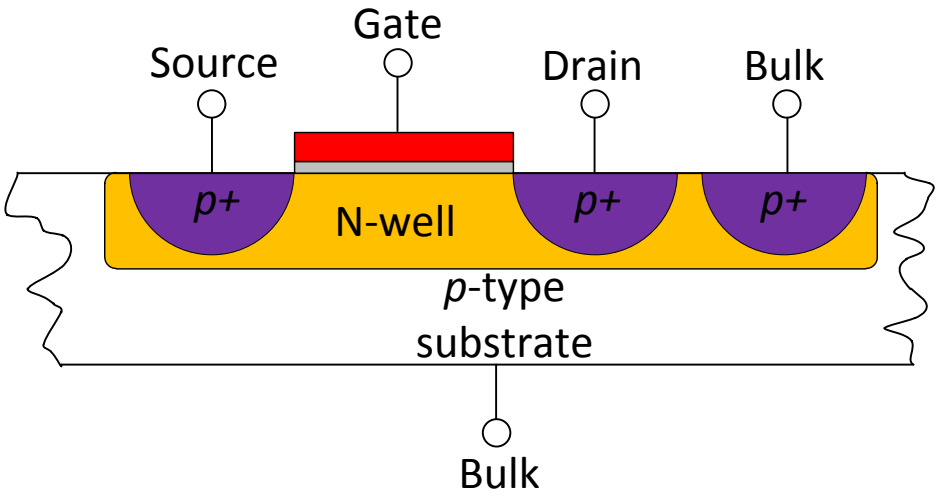

Fig. 4.23 Cross section of a typical P-MOS transistor.

The operation mode of PMOS and NMOS transistors are similar. In both devices, the creation of a conducting channel is required to have a current flowing between Source and Drain. However, in the PMOS transistor, the channel charge is made of holes, while, in the NMOS transistor, the channel charge is made of electrons. In PMOS as well as NMOS transistors, current is due to the drift of the channel charge from Source toward the Drain by the horizontal electric field. In the NMOS transistor, the channel charge (electrons) is negative, generating a negative Source Drain current, I_{SD} (i.e a positive I_{DS}), while, in the PMOS transistor, the channel charge is positive (holes), generating a positive I_{SD} current.

Moreover, in PMOS transistors, Drain and Gate have to be biased to a voltage less than the Source voltage in order to attract holes. In facts, V_{TH}, V_{GS} and V_{DS} are negative in the PMOS transistor, while they are positive in the NMOS transistor.

As for NMOS transistor, also PMOS transistor has four main regions of operations:

- cut-off;
- linear or triode;
- saturation;
- weak inversion.

To derive characteristic equations of PMOS transistors, it is not necessary repeating the analytic procedure used for the NMOS transistor. It is sufficient to perform substitutions reported in table 4.2, in characteristic equations of the NMOS transistor:

NMOS	PMOS
I_{DS}	I_{SD}
V_{GS}	V_{SG}
V_{DS}	V_{SD}
V_{TH}	$-V_{TH}$
μ_n	μ_p

Table 4.2 Substitutions required to derive the characteristic equations of the PMOS transistor.

Table 4.3 summarizes The characteristic equations of the PMOS transistor and the bias conditions required to have them.

Region of operation	Characteristic equation	V_{SG}	V_{SD}
Cut-off	$I_{SD}=0$	$>V_{TH}$	-
Linear or triode	$I_{SD} = \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(V_{SG} + V_{TH} - \frac{V_{SD}}{2}\right) \cdot V_{SD}$	$<V_{TH}$	$>V_{SG}+ V_{TH}$
Saturation	$I_{SD} = \frac{1}{2} \cdot \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{SG} + V_{TH})^2 \cdot (1 + \lambda \cdot V_{SD})$	$<V_{TH}$	$<V_{SG}+ V_{TH}$
Weak inversion	$I_{SD} = I_0 \cdot e^{\frac{V_{SG}}{n \cdot Vt}}$	$\cong V_{TH}$	<0 V

Table 4.3 Summary of the PMOS transistor characteristic equations for different regions of operation

For the PMOS transistor it is also possible to use the same characteristic equations and bias conditions disequations for the NMOS transistor, substituting currents and voltages with their absolute values, as table 4.4 reports:

Region of operation	Characteristic equation	$\lvert V_{GS}\rvert$	$\lvert V_{DS}\rvert$
Cut-off	$\lvert I_{DS}\rvert=0$	$<\lvert V_{TH}\rvert$	-
Linear or triode	$\lvert I_{DS}\rvert = \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(\lvert V_{GS}\rvert - \lvert V_{TH}\rvert - \frac{\lvert V_{DS}\rvert}{2}\right) \cdot \lvert V_{DS}\rvert$	$>\lvert V_{TH}\rvert$	$<\lvert V_{GS}\rvert - \lvert V_{TH}\rvert$
Saturation	$\lvert I_{DS}\rvert = \frac{1}{2} \cdot \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot (\lvert V_{GS}\rvert - \lvert V_{TH}\rvert)^2 \cdot (1 + \lambda \cdot \lvert V_{DS}\rvert)$	$>\lvert V_{TH}\rvert$	$>\lvert V_{GS}\rvert - \lvert V_{TH}\rvert$
Weak inversion	$\lvert I_{DS}\rvert = I_0 \cdot e^{\frac{\lvert V_{GS}\rvert}{n \cdot Vt}}$	$\cong \lvert V_{TH}\rvert$	>0 V

Table 4.4 Summary of the PMOS characteristic equations and bias conditions for different regions of operation with absolute values of current and voltages

As in the PMOS transistor, I_{SD} is due to the drift of holes, the hole mobility, μ_p, has to be considered. The mobility of holes is quite less than electrons, as equation 4.35 predicts:

Eq. 4.36
$$\frac{\mu_n}{\mu_p} \cong 2.5$$

Therefore, a PMOS transistor has to be larger than a NMOS transistor in the same bias conditions in order to provide the same current.

However, in PMOS transistors it is possible to mitigate the threshold voltage shift due to the Body effect. In facts, since p-type substrate is bounded to ground, i.e. to the minimum voltage available for the circuit, the pn-junction formed by the N-well and the p-type substrate is always inversely biased. As the N-well results always isolated, it is possible to connect its Bulk terminal to the Source, in order to null the V_{BS}

voltage, reducing the Body effect. While this effect can be always mitigated in PMOS transistors, it cannot be avoided in NMOS transistors when the Source is connected to a voltage different from the ground, where its Bulk is bounded.

6. Analysis of the bias conditions of a MOS transistor

To solve a circuit including a MOS transistor by hand, it is required to determine its region of operation. As explained in previous paragraphs, the region of operation of a MOS transistor depends on voltages at its terminals. Assuming that voltages at the MOS transistor terminals are known, in order to determine the region of operation of a MOS transistor, at firstly it is required to verify if transistor is on or off by comparing its V_{GS} to V_{TH}. If the MOS transistor is off, then it is operating in cut-off region. Otherwise, if the MOS transistor is on, then its V_{DS} should be compared to V_{DSsat} in order to verify if it is operating in saturation or triode region.

Fig. 4.24 shows the algorithm to follow in order to determine the region of operation of a NMOS transistor as V_{GS} and V_{DS} are known.

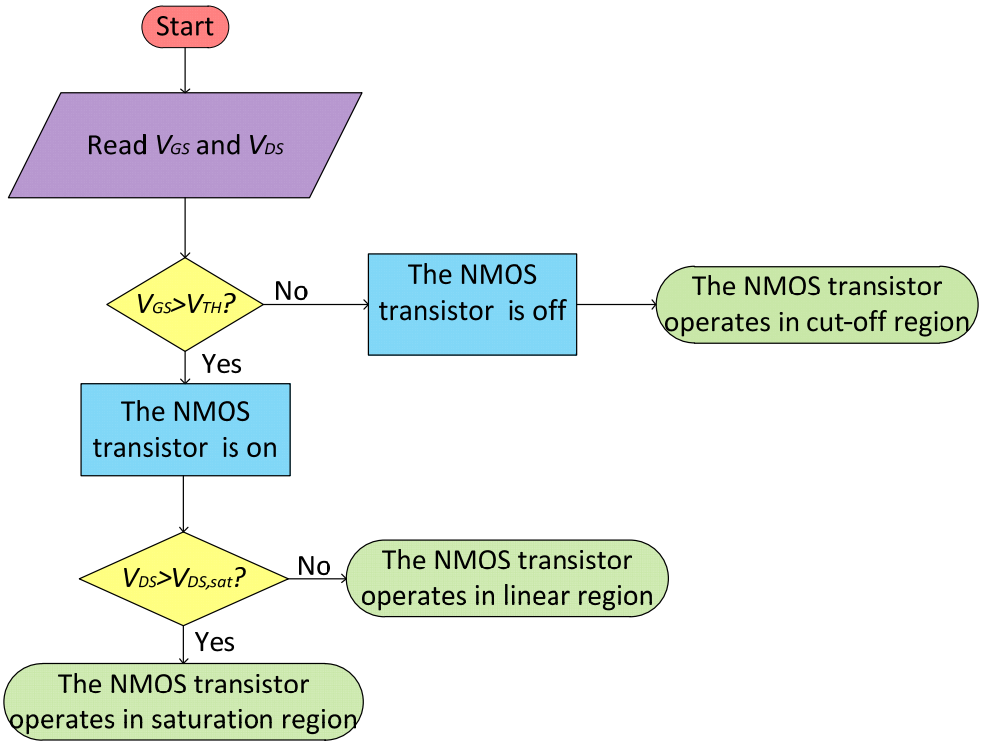

Fig. 4.24 Algorithm for determining the bias conditions of a NMOS transistor by hand.

Considering the absolute values of V_{DS}, V_{GS} and V_{TH} voltages, the same algorithm can be applied to a PMOS transistor.

- **Example 4.2**

Problem. Consider the NMOS transistor reported in Fig. 4.25. Assume V_{TH}=0.5V V_G=0.8V, V_D=2V. Find the operation region of the NMOS transistor.

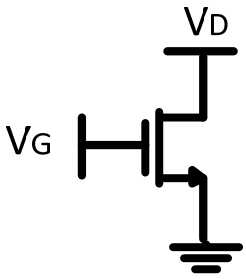

Fig. 4.25 NMOS transistor biased by V_G and V_D voltages.

Solution. Evaluate V_{GS} and compare to V_{TH}:

Eq. 4.37 $$V_{GS} = V_G = 0.8V > 0.5V = V_{TH}$$

As V_{GS} is more than V_{TH}, the NMOS transistor is on. Now calculate V_{DS} and compare it to V_{DSsat}:

Eq. 4.38 $$V_{DS} = V_D = 2V > 0.3V = V_{GS} - V_{TH} = V_{ov} = V_{DSsat}$$

As V_{DS} is more than V_{DSsat}, the NMOS transistor operates in saturation region.

- **Example 4.3**

Problem. Consider the PMOS transistor reported in Fig. 4.26. Assume V_{TH}=-0.5V V_G=3V, V_S=5V and V_D=4V. Find the operation region of the PMOS transistor.

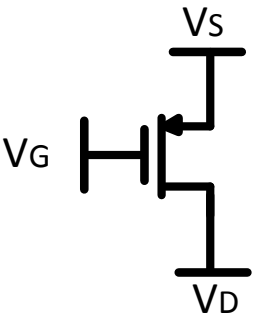

Fig. 4.26 PMOS transistor biased by V_G, V_s, and V_D voltages.

Solution. Evaluate $|V_{GS}|$ and compare to $|V_{TH}|$:

Eq. 4.37 $$|V_{GS}| = |V_G - V_S| = 2V > 0.5V = |V_{TH}|$$

As $|V_{GS}|$ is more than $|V_{TH}|$, the PMOS transistor is on. Now calculate $|V_{DS}|$ and compare it to $|V_{DSsat}|$:

Eq. 4.38 $$|V_{DS}| = |V_D - V_S| = 1V < 1.5V = |V_{GS}| - |V_{TH}| = |V_{ov}| = |V_{DSsat}|$$

As $|V_{DS}|$ is less than $|V_{DSsat}|$, the PMOS transistor operates in triode region.

-

However, as a circuit including a MOS transistor is to be solved by hand, generally, the bias voltages at the MOS transistor terminals are not a priori known. Therefore, it is not possible to a priori determined the MOS transistor region of operation. Hence, it is needed to make a starting hypothesis on its region of operation. Then, the MOS transistor characteristic equation coherent with the starting hypothesis is defined. As the MOS transistor characteristic equation is known, it is possible to solve the circuit. Once the circuit is solved, it is required to verify the starting hypothesis on its region of operation. If the starting hypothesis is not verified, then the procedure for solving the circuit is to be restarted with a different starting hypothesis, until the starting hypothesis is verified. Fig. 4.27 shows the algorithm for solving the a circuit including a MOS transistor by hand.

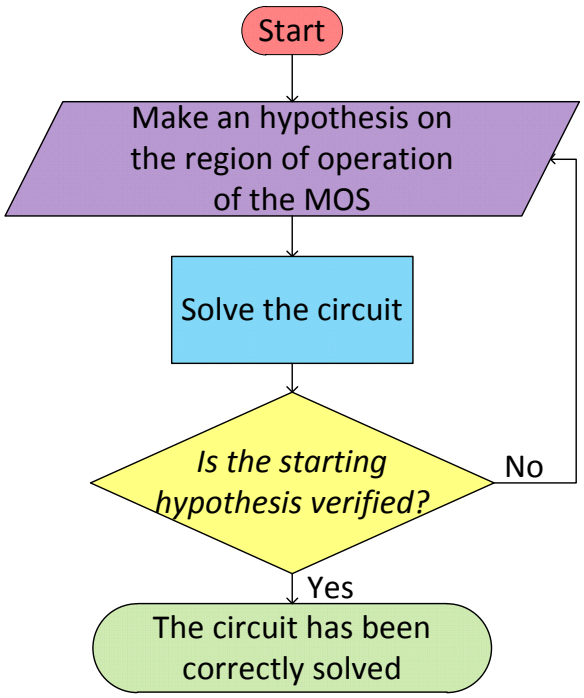

Fig. 4.27 Algorithm for solving a circuit including a MOS transistor by hand.

- **Example 4.4**

Problem. Consider the circuit reported in fig. 4.28. Assume V_{TH}=0.5V, K_n=1mA/V^2, V_G=2V, V_{DD}=5V, R_S=1kΩ, R_D=2 kΩ. Solve the circuit.

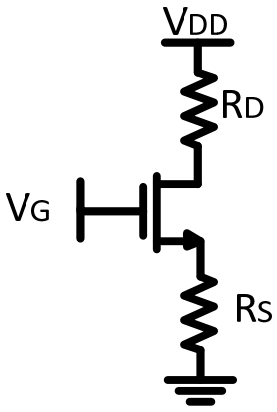

Fig. 4.28 Circuit to be solved in the example 4.4.

Solution. Establish a direction and set a label for each current in the circuit, as done in fig. 4.29.

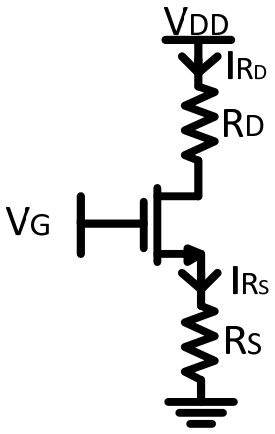

Fig. 4.28 Circuit of the example 4.4 with currents labels and directions.

Applying the Kirchoff's current law to the Source of the NMOS transistor, it is get:

Eq. 4.39 $$I_{DS} = I_{R_S}$$

where I_{DS} is the Drain-Source current flowing into the NMOS transistor. Suppose that the NMOS transistor works in saturation region. Therefore, equation 4.14 can be used to determine I_{DS}. Replacing the expressions of currents of the NMOS transistor and the resistor R_S into equation 4.39, it is get:

Eq. 4.40 $$k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_G - V_S - V_{TH})^2 = \frac{V_S}{R_S}$$

Solving equation 4.40, two solutions for V_S are obtained: 0.677 V and 3.646 V. However, the second one has to be waived as it means that the NMOS transistor has to be on despite its Source voltage is higher than the Gate voltage, which is not physically possible. Therefore, V_S is equal to 0.677 V. Hence, the I_{DS} current is equal to 677µA. V_D is calculate as follows:

Eq. 4.41 $$V_D = V_{DD} - R_D \cdot I_{DS} = 3.646 \text{ V}$$

Once all currents and voltages in the circuit have been calculated, then the hypothesis made about the region of operation of the NMOS transistor has to be verified. Therefore, V_{DS} voltage has to be calculated and compared to V_{DSsat}, i.e.:

Eq. 4.42 $$V_{DS} = V_D - V_S = 2.969 \text{ V} > V_{DSsat} = V_{GS} - V_{TH} = 0.823\ V$$

As V_{DS} is more than V_{DSsat}, the hypothesis of the NMOS transistor operating in saturation region is verified.

-

- **Example 4.5**

Problem. Consider the circuit reported in fig. 4.30. Assume V_{TH}=-0.5V, K_p=1mA/V^2, V_G=3.5V, V_{DD}=5V, R_D=2kΩ. Solve the circuit and find the minimum value of R_D, R_{Dmin}, that brings in triode region the PMOS transistor.

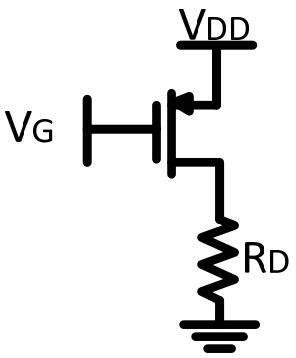

Fig. 4.30 Circuit to be solved in the example 4.5.

Solution. Establish a direction and set a label for the current in the circuit, as done in fig. 4.31.

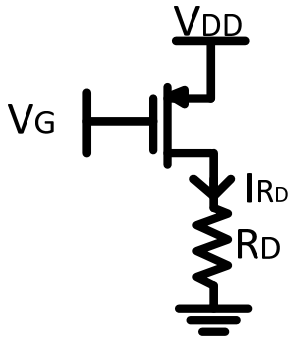

Fig. 4.31 Circuit of the example 4.5 with current label and direction.

Suppose that the PMOS transistor works in saturation region. Therefore, equation for PMOS transistor operating in saturation region reported in table 4.3 can be used to determine I_{SD}, i.e.:

Eq. 4.43 $$I_{SD} = k_p \cdot (V_{SG} + V_{TH})^2 \cdot (1 + \lambda \cdot V_{SD}) \cong k_p \cdot (V_{DD} - V_G + V_{TH})^2 = 1\text{mA}$$

Applying the Kirchoff's current law to the Drain of the PMOS transistor, it is get:

Eq. 4.44 $$I_{SD} = I_{R_D}$$

Hence, the Drain voltage V_D is calculate as follows:

Eq. 4.45 $$V_D = R_D \cdot I_{SD} = 2\ \text{V}$$

Once all currents and voltages in the circuit have been calculated, then the hypothesis made about the region of operation of the PMOS transistor has to be verified. Therefore, V_{SD} voltage has to be calculated and compared to V_{SDsat}, i.e.:

Eq. 4.42 $$V_{SD} = V_{DD} - V_D = 3\ \text{V} > V_{SD,sat} = V_{SG} + V_{TH} = 1\ V$$

As V_{SD} is more than V_{SDsat}, the hypothesis of the PMOS transistor operating in saturation region is verified.

The minimum value of R_D, R_{Dmin}, that brings in triode region the PMOS transistor, makes V_{SD} equal to $V_{SD,sat}$, i.e.:

Eq. 4.43 $$V_{SD} = V_{DD} - V_D = V_{DD} - R_{Dmin} \cdot I_{SD} > V_{SD,sat} = 1\,V$$

From equation 4.43 it is get:

Eq. 4.44 $$R_{Dmin} = \frac{V_{DD} - V_{SD,sat}}{I_{SD}} = 4k\Omega$$

▪

7. Small signal model of the MOS transistor

The small signal model is used in order to simplify the calculation of the gain and the input and output impedances in analog circuits including MOS transistors. The complexity of the model is increased according to the analysis to perform.

7.1 Small signal circuit

Consider the NMOS transistor in fig. 4.32. with bias voltages V_G and V_{DD} applied as shown.

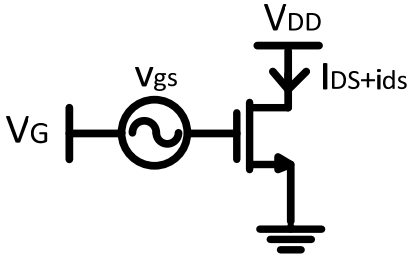

Fig. 4.32 NMOS transistor with biasing and a small signal v_{gs} applied to the Gate.

Assuming that the NMOS transistor is on (i.e. $V_G > V_{th}$), then, a Drain-Source bias current, I_{DS}, flows across the device. As a voltage signal v_{gs} is applied in series with V_G, a small drain current variation i_{ds} is generated. For small values of v_{gs}, i_{ds} is directly related to v_{gs} by the transconductance parameter, g_m, which is defined as follows:

Eq. 4.45 $$g_m = \frac{\partial I_{ds}}{\partial V_{gs}}$$

Therefore i_{ds} is obtained as follows:

Eq. 4.46 $$i_{ds} = g_m \cdot v_{gs}$$

This relation can be explained in an alternative way, by recurring to the Taylor series expansion of the I_{ds} current expression, i.e.:

Eq. 4.47 $$I_{ds} = I_{DS} + i_{ds} = I_{DS} + \frac{\partial I_{ds}}{\partial V_{gs}} \cdot \partial V_{gs} + \frac{1}{2} \cdot \frac{\partial^2 I_{ds}}{\partial^2 V_{gs}} \cdot \partial V_{gs} + \cdots$$

For small Gate-Source voltage variation, ∂V_{gs}, (i.e. for small values of v_{gs}) is it possible to have a good approximation of I_{ds} current by stopping to the first term of the Taylor expansion, i.e.:

Eq. 4.48 $$I_{ds} = I_{DS} + i_{ds} = I_{DS} + \frac{\partial I_{ds}}{\partial V_{gs}} \cdot \partial V_{gs} = I_{DS} + g_m \cdot v_{gs}$$

Consider the NMOS transistor in fig. 4.33, with a small signal, v_{ds}, applied to the Drain. Assuming that the NMOS transistor is on, then the Drain-Source current, I_{ds}, changes with v_{ds}.

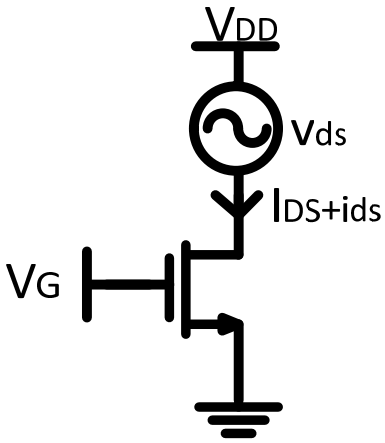

Fig. 4.33 NMOS transistor with biasing and a small signal v_{ds} applied to the Drain.

It is possible to calculate I_{ds} variation, i_{ds}, due to v_{ds}, as follows:

Eq. 4.49 $$i_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} \cdot \partial V_{ds} = g_{ds} \cdot v_{ds}$$

where g_{ds} is the small signal output conductance, calculated as follows:

Eq. 4.50 $$g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}}$$

The Gate of the MOS is isolated from the channel by the silicon oxide. Therefore, at low frequency, the Gate current is about zero, while the input impedance is about infinite, as a consequence.

Combining the preceding small signal elements yields the small model of the MOS transistor shown in Fig. 4.34.

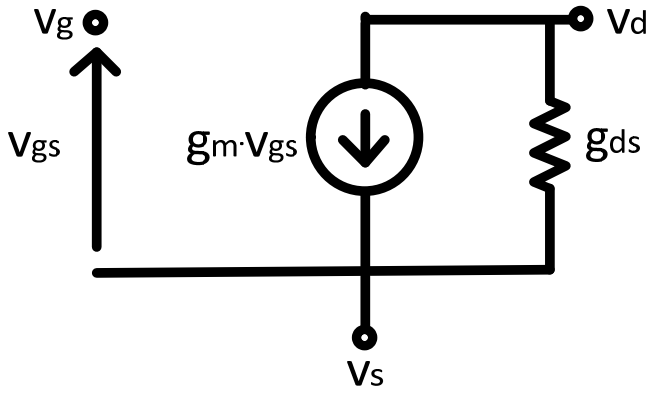

Fig. 4.34 Basic model of the small signal circuit.

This is the basic model of the small signal circuit. It does not includes the dependence of the Drain-Source current on the Bulk voltage due to the Body effect, and the parasitic capacitances. In fact, the Bulk-

Source voltage changes the threshold voltage, which changes the Drain-Source current when the Gate-Source voltage is fixed. Therefore, a further transconductance term, the Bulk transconductance, is required to the MOS model. The Bulk transconductance is defined as follows:

Eq. 4.51 $$g_{mb} = \frac{\partial I_{ds}}{\partial V_{bs}}$$

At high frequency, the effects of the parasitic capacitances cannot be neglected.

Fig. 4.35 shows the small signal circuit including C_{GS} and C_{GD} capacitances and the Bulk transconductance.

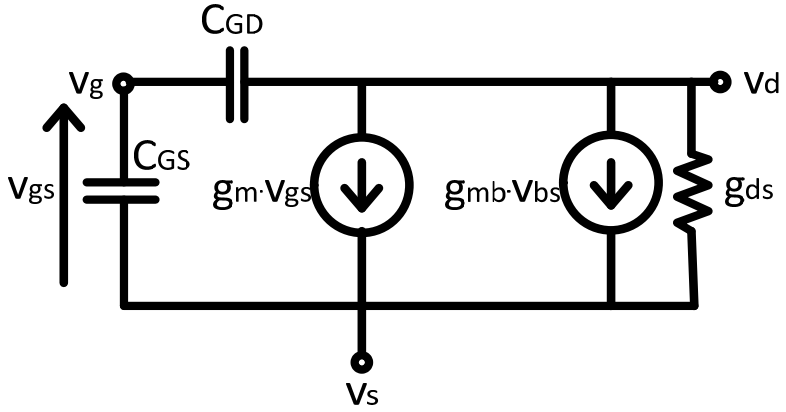

Fig. 4.35 Complete model of the small signal circuit.

This is the complete model of the small signal circuit.

7.2 Small signal circuit parameters calculation

The calculation of the small signal circuit parameters of a MOS transistor depends on its region of operation.

In saturation region, the transconductance of a NMOS transistor is calculated from equation 4.20 by differentiating with respect to Gate-Source voltage, V_{gs}, i.e.:

Eq. 4.52 $$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot (1 + \lambda \cdot V_{DS}) \approx 2 \cdot k_n \cdot V_{ov} = 2 \cdot \frac{I_{ds}}{V_{ov}} = 2 \cdot \sqrt{k_n \cdot I_{DS}}$$

According to the last expression of g_m found in equation 4.52, the transcoductance of a NMOS transistor is proportional to the square root of the bias current, I_{DS}. This is a key difference with respect to the bipolar transistor which has a transconductance proportional to the bias current. Therefore, the transconductance for given current, i.e. the current efficiency, is much higher in bipolar than MOS transistors.

However, in weak inversion, the transconductance of a NMOS transistor is proportional to the bias current I_{DS}, as the transistor current has an exponential dependency on the Gate-Source voltage, V_{gs} . In this case, the current efficiency of MOS transistors is closer to that one of the bipolar transistors. In fact, in weak inversion, the transconductance is calculated from equation 4.24 by differentiating:

Eq. 4.53 $$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} = \frac{I_{DS}}{n \cdot V_t}$$

To push a transistor in weak inversion, reducing the transistor overdrive is needed. This has important impact on the circuit performance like offset, noise, etc., therefore it is not always possible. Moreover, reducing the transistor overdrive by keeping a constant bias current, implies that the transistors sizes must be increased. But larger transistors introduce bigger parasitic capacitances, limiting the operation frequency.

Fig. 4.36 shows the real behavior of the MOS transistor transconductance, g_m, when its Gate-Source voltage, V_{GS}, is increased. It is assumed that the Drain-Source voltage, V_{DS}, is high enough not to make the transistor going in linear region.

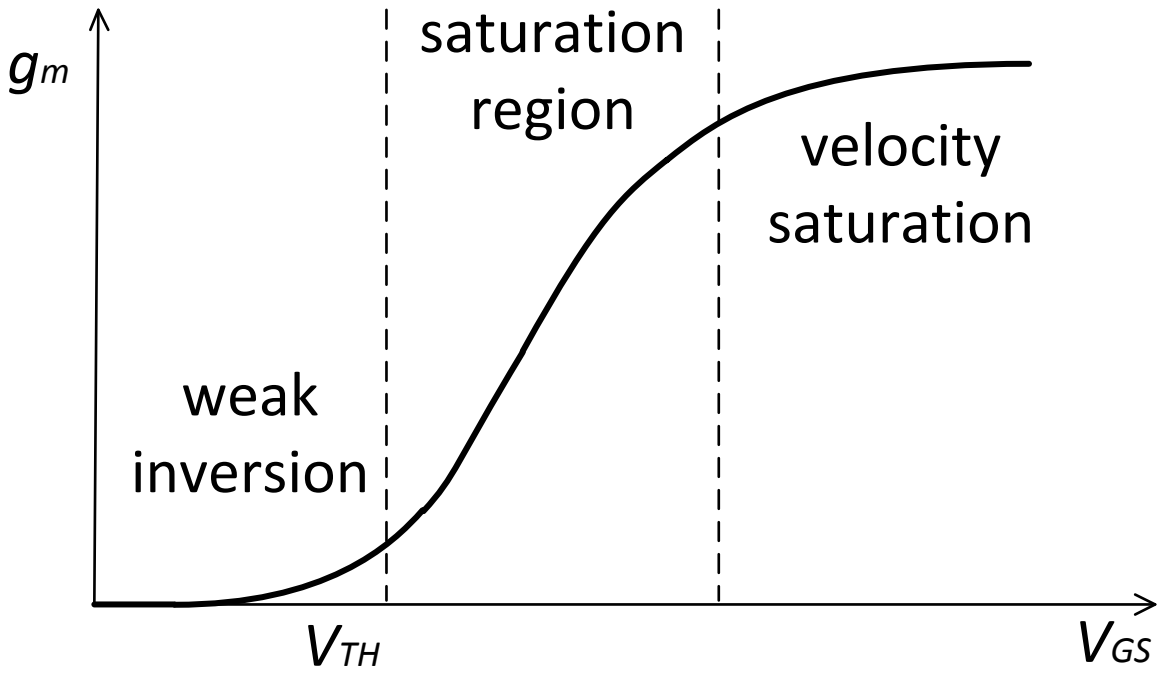

Fig. 4.36 MOS transistor transconductance, g_m, versus V_{GS}.

For high value of V_{GS} the MOS transistor g_m saturates. In fact, increasing V_{GS}, the electric field inside the transistor augments too. For low electric field intensities, the drift velocity of electrons is proportional to the electric field, according to equation 4.6. As the electric field intensity approaches the critical value, $\mathcal{E}_c$, the drift velocity of the electrons reaches its upper limit (i.e. the scattering limited velocity) due to the increasing scattering resistance. This phenomena can be described by introducing a more complex model for the electron mobility. Equation 4.54 shows a first order approximation model for the mobility that consider its reduction at the increase of the Gate-Source voltage, V_{GS} [xxx]:

Eq. 4.54
$$\mu_n = \frac{\mu_{n0}}{1+\Theta\cdot(V_{GS}-V_{TH})}$$

where μ_{n0} is the electron mobility for a null electric field, while the parameter Θ depends on the critical electric filed, $\mathcal{E}_c$, and the transistor length L:

Eq. 4.55
$$\Theta = \frac{1}{L\cdot\mathcal{E}_c}$$

Therefore, the phenomena of velocity saturation is more evident for short channel devices.

By replacing the expression of the electron mobility reported by equation 4.54 in equation 4.20, the following model for the Drain-Source current is found:

Eq. 4.55 $$I_{DS} = \frac{1}{2} \cdot \frac{\mu_{n0}}{1+\theta \cdot (V_{GS}-V_{TH})} \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2 \cdot (1 + \lambda \cdot V_{DS})$$

For large values of V_{GS} (i.e. $V_{GS} \gg \frac{1}{\theta} + V_{TH}$), an substantially linear dependency of the Drain-Source current, I_{DS}, versus the gate-source voltage, V_{GS}, occurs:

Eq. 4.56 $$I_{DS} \approx \frac{1}{2} \cdot \frac{\mu_{n0}}{\theta} \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot (1 + \lambda \cdot V_{DS})$$

In this condition, the transconductance reaches its maximum, $g_{m,sat}$, that does not depends on the Gate-Source, V_{GS}:

Eq. 4.57 $$g_{m,sat} = \frac{1}{2} \cdot \frac{\mu_{n0}}{\theta} \cdot C_{ox} \cdot \frac{W}{L} \cdot (1 + \lambda \cdot V_{DS}) \approx \frac{k_n}{\theta}$$

In linear region, the transconductance of a NMOS transistor is calculated from equation 4.8 by differentiating with respect to Gate-Source voltage, V_{gs}, i.e.:

Eq. 4.58 $$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot V_{DS} = 2 \cdot k_n \cdot V_{DS}$$

In saturation region, the output conductance is calculated from equation 4.20 by differentiating with respect to the Drain-Source voltage, V_{ds}, as follows:

Eq. 4.59 $$g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} = \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2 \cdot \lambda \approx \lambda \cdot I_{DS}$$

Similarly, in linear region, the output conductance is calculated from equation 4.8 by differentiating with respect to the Drain-Source voltage, V_{ds}, as follows:

Eq. 4.60 $$g_{ds} = \frac{\partial I_{ds}}{\partial V_{ds}} = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot \lambda = 2 \cdot k_n \cdot (V_{GS} - V_{TH})$$

Using equation 4.20 the Bulk transconductance in saturation region is calculated as follows:

Eq. 4.61 $$g_{mb} = \frac{\partial I_{ds}}{\partial V_{bs}} = \frac{\partial I_{ds}}{\partial V_t} \cdot \frac{\partial Vt}{\partial V_{bs}} = -g_m \cdot \chi$$

where χ is the rate of change of the threshold voltage with the Bulk-Source voltage due to the Body effect. From equations 4.32 and 2.xxx, χ is calculated as follows:

Eq. 4.62 $$\chi = -\frac{\gamma}{2 \cdot \sqrt{2 \cdot \phi_f - V_{BS}}} = \frac{C_{JS}}{C_{ox}}$$

where C_{JS} is the capacitance per unit area formed by substrate Source junction.

The ratio between g_{mb} and g_m is an important parameter in practice. From equation 4.61, it is get:

Eq. 4.59 $$\frac{g_{mb}}{g_m} = -\chi$$

This is a powerful relationship, but it does not provide an accuratvalue as χ depends on the bias conditions. The χ factor typically ranges from 0.1 to 0.3, therefore the Bulk transconductance is 3-10 times smaller than the MOS transconductance.

To derive the expressions of small signal circuit parameters of PMOS transistors, it is not necessary repeating the analytic procedure used for the NMOS transistor. It is sufficient to perform substitutions reported in table 4.2, in the equations obtained to calculate the small signal circuit parameters of a NMOS transistor.

8. Gain stages with a single MOS transistor

The small signal equivalent circuits of Bipolar and MOS transistors are quite similar. The two devices differs mainly in the values of some parameters. MOS transistors have a substantially infinite Gate resistance, in contrast with the finite Base resistance, r_π, of bipolar transistors. On the other hand, the Bipolar transistor has a transconductance one order of magnitude larger than that of MOS transistors, with the same current. According to the situations, MOS of Bipolar transistors are preferable. For example, if high input impedance amplifiers are required, they can more easily implemented by using a MOS technology. If a larger gain is required, then a Bipolar transistor is more suited. Designers must appreciate the similarities and the differences between Bipolar and MOS technologies, in order to make an appropriate technology choice.

As for the Bipolar transistors, MOS transistors are able to provide a useful gain in three configurations: common Source, common Drain and common Gate configurations.

8.1 Common Source configuration

The common source amplifier is shown in figure 4.37.

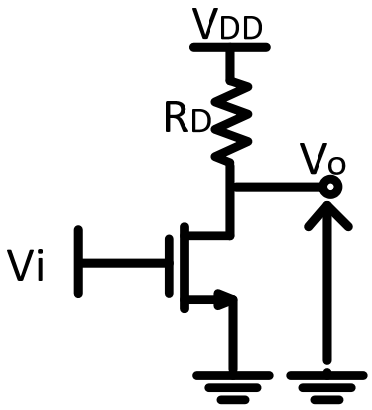

Fig. 4.37 Common source amplifier.

As Figure 4.37 shows, the Source is connected to the ground, i.e. it is common to the ground network. The input signal is applied to the Gate, while the output signal is taken from the Drain.

The output voltage, V_O, is calculated as follows:

Eq. 4.60 $$V_o = V_{DD} - R_D \cdot I_{DS}$$

As $V_i < V_{TH}$, no current flows through the NMOS transistor, therefore:

Eq. 4.61 $$I_{DS} = 0 \text{ and } V_o = V_{DD}$$

As V_i is increased beyond the threshold voltage, V_{TH}, the NMOS transistor starts conducting. The NMOS transistor operates in saturation region until $V_o > V_i - V_{TH}$. According to equation 4.14, as the NMOS transistor is in saturation region, the Drain-Source current is given by:

Eq. 4.62 $$I_{DS} = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_i - V_{TH})^2$$

Combining equations 4.60 e 4.62 the following input-output relationship is get:

Eq. 4.63 $$V_o = V_{DD} - R_D \cdot I_{DS} = V_{DD} - R_D \cdot k_n \cdot (V_i - V_{TH})^2$$

By keeping on increasing the V_i voltage, the NMOS transistor goes in linear region as $V_o < V_i - V_{TH}$. According to equation 4.8, the following input-output relationship is get:

Eq. 4.64 $$V_o = V_{DD} - R_D \cdot I_{DS} = V_{DD} - R_D \cdot 2 \cdot k_n \cdot (V_i - V_{TH}) \cdot V_o$$

As the output conductance augments when the NMOS transistor enter the linear region, the voltage gain drops dramatically.

The resulting input-output characteristic is reported in figure 4.38.

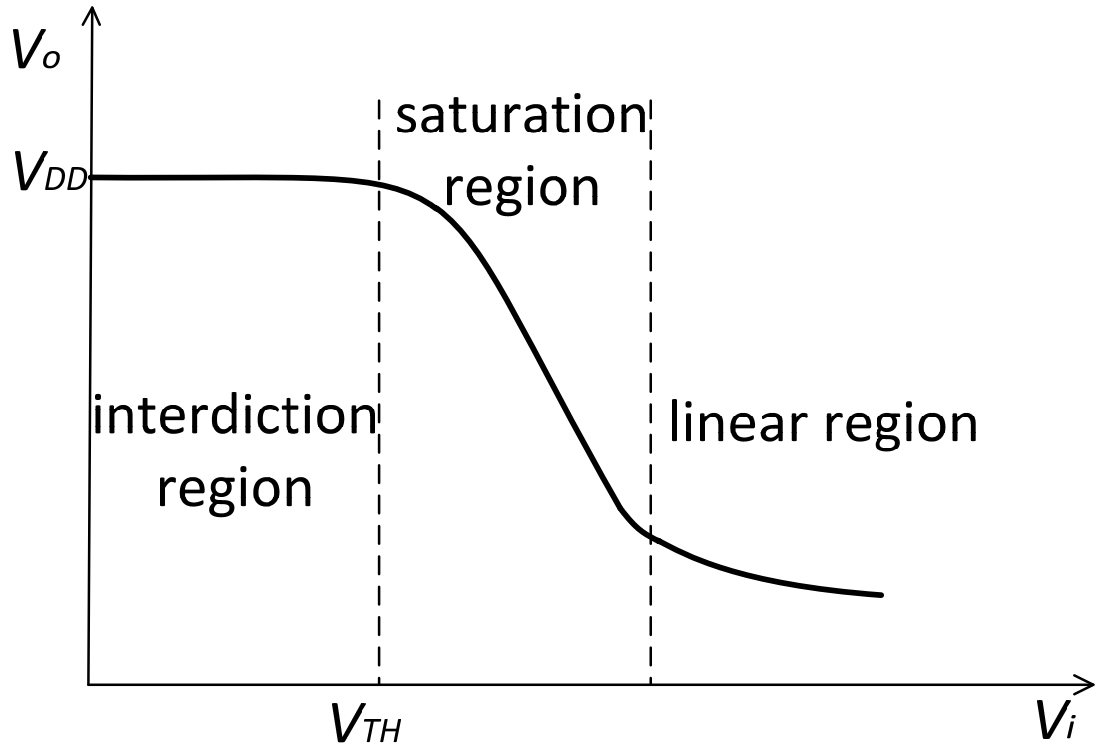

Fig. 4.38 Input-output characteristic of a common Source amplifier.

The slope of this characteristic at any operating point, corresponds to the small signal gain at that point. In linear region the slope is reduced, confirming that the voltage gain drops.

Assuming that a small voltage signal is applied to the Gate, figure 4.39 shows the corresponding small signal equivalent circuit:

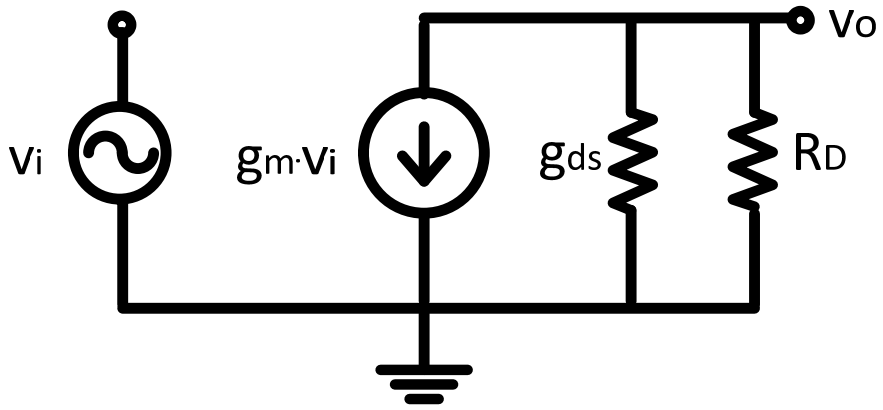

Fig. 4.39 Small signal equivalent circuit of common Source amplifier.

The output voltage, v_o, is given by the voltage drop on the resistive load, $\frac{1}{g_{ds}} \| R_D$. As the current flowing in the resistive load is that provided by the current generator, the output voltage is calculated:

Eq. 4.65 $$v_o = -g_m \cdot v_i \cdot \frac{1}{g_{ds}} \| R_D$$

Therefore, the voltage gain, A_V, is calculated as follows:

Eq. 4.66 $$A_V = \frac{v_o}{v_i} = -g_m \cdot \frac{1}{g_{ds}} \| R_D$$

As $R_D \ll \frac{1}{g_{ds}}$, then the voltage gain becomes:

Eq. 4.67 $$A_V = \frac{v_o}{v_i} \approx -g_m \cdot R_D$$

The voltage gain increases by augmenting R_D. However, for large R_D values the output conductance of the NMOS, g_{ds}, is no more negligible. A least, as $R_D \gg \frac{1}{g_{ds}}$, the following voltage gain is get:

Eq. 4.68 $$A_V = \frac{v_o}{v_i} \approx -\frac{g_m}{g_{ds}}$$

Assuming that the NMOS transistor works in saturation region, by replacing equations 4.52 and 4.59 in equation 4.68, the following dependency of the gain on the Drain-Source current, I_{DS}, is get:

Eq. 4.69 $$A_V = -\frac{g_m}{g_{ds}} = -\frac{2 \cdot \sqrt{k_n \cdot I_{DS}}}{\lambda \cdot I_{DS}} = -\frac{2 \cdot \sqrt{k_n}}{\lambda \cdot \sqrt{I_{DS}}}$$

By reducing the Drain-Source current, I_{DS}, the voltage gain increases. However, according to equation 4.14, reducing the Drain-Source current of an MOS transistor with a constant geometry, produces a decrease of the overdrive voltage. In practical, the MOS transistor is pushed to work in weak inversion. In weak inversion, the transconductance of the NMOS is proportional to the Drain-Source current, I_{DS}. Then, by replacing equations 4.53 and 4.59 in equation 4.68, the expression of the voltage gain changes:

Eq. 4.70 $$A_V = -\frac{g_m}{g_{ds}} = -\frac{\frac{I_{DS}}{n \cdot Vt}}{\lambda \cdot I_{DS}} = -\frac{1}{\lambda \cdot n \cdot Vt}$$

This is the maximum value of the voltage gain. It does not depend on the Drain-Source current, I_{DS}, nor on the MOS transistor aspect ratio, $\frac{W}{L}$. It only depends on the channel modulation parameter, λ, which is reversely proportional to the MOS transistor length, *L*, and the slope factor, *n*, which is mainly correlated to the fabrication process.

As the input signal is applied to the Gate terminal, the input resistance of the amplifier is about infinite. The output resistance is calculated by considering the small signal equivalent circuit reported in Fig. 4.40.

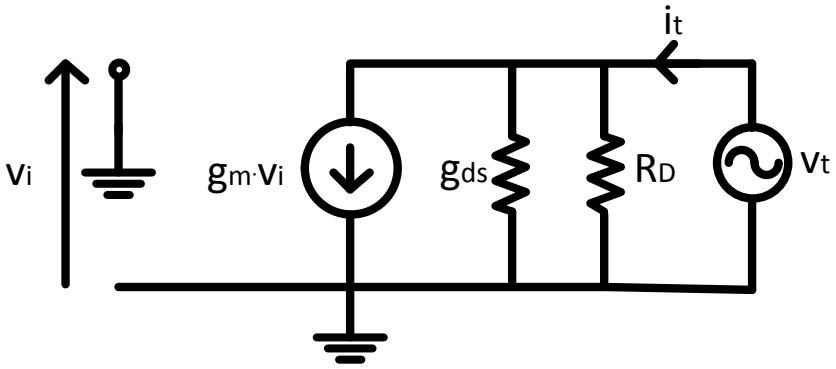

Fig. 4.40 Small signal equivalent circuit for the calculation of the output resistance of a common Source amplifier.

For the calculation of the output resistance, R_o, the input signal, v_i, is nulled by definition. Therefore, the current of the current generator (i.e. $g_m \cdot v_i$) is zero too. In practical, it is like the current generator does not exist, while the parallel of R_D and $1/g_{ds}$ remains. The output resistance, R_o, is calculated as follows:

Eq. 4.70 $$R_o = \frac{v_t}{i_t} = R_D \| \frac{1}{g_{ds}}$$

- **Example 4.6**

Problem. Consider the circuit reported in fig. 4.41. Assume V_{TH}=1V, K_n=5mA/V^2, V_{DD}=10V, R_L=4kΩ, R_1=8.5kΩ, R_2=1.5kΩ, C_L=10nF, $C_i \rightarrow \infty$. Calculate the bias point and the voltage gain versus the frequency, $\frac{v_o}{v_i}(s)$. Trace the Bode diagram of the voltage gain. Evaluate the maximum input voltage amplitude before the NMOS transistor, M_1, goes to the linear region.

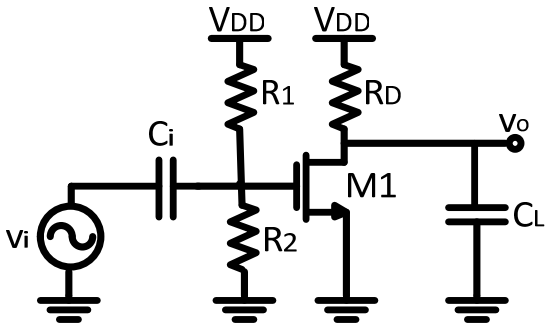

Fig. 4.41 Circuit to be solved in the example 4.6.

Solution. To calculate the bias point it is possible to simplify the circuit by nulling the capacitances and the input signal, as it is done in figure 4.42.

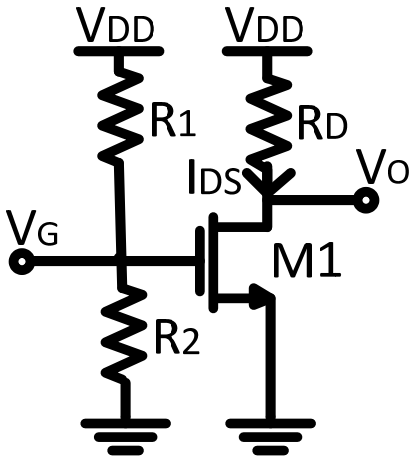

Fig. 4.42 Circuit of the example 4.6 without capacitances and with a null input signal.

As the Gate does not adsorb any current, the Gate voltage V_G, is calculated as a partition of V_{DD} on R_2 resistor:

Eq. 4.71 $$V_G = V_{GS} = V_{DD} \cdot \frac{R_2}{R_2+R_1} = 1.5V$$

As V_G=V_{GS}>V_{TH} the transistor is switched on. Suppose that the NMOS transistor works in saturation region. Therefore, equation for NMOS transistor operating in saturation region reported in equation 4.14 can be used to determine I_{DS}, i.e.:

Eq. 4.72 $$I_{DS} = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_G - V_{TH})^2 = 1.25\text{mA}$$

The output bias voltage, V_O, is calculated as follows:

Eq. 4.73 $$V_o = V_{DD} - R_D \cdot I_{DS} = 5V$$

Once that the circuit has been solved, it is required to verify the initial hypothesis, i.e. the NMOS transistor working in saturation region. It means that the following condition on the drain source voltage V_{DS}, reported in table 4.1, is to be verified:

Eq. 4.74 $$V_{DS} > V_{GS} - V_{TH} \rightarrow V_O > V_G - V_{TH} \rightarrow 5V > 0.5V$$

As this condition is satisfied, the initial hypothesis of NMOS transistor working in saturation region is verified.

In order to calculate the voltage gain, the small signal equivalent circuit is reported in figure 4.43.

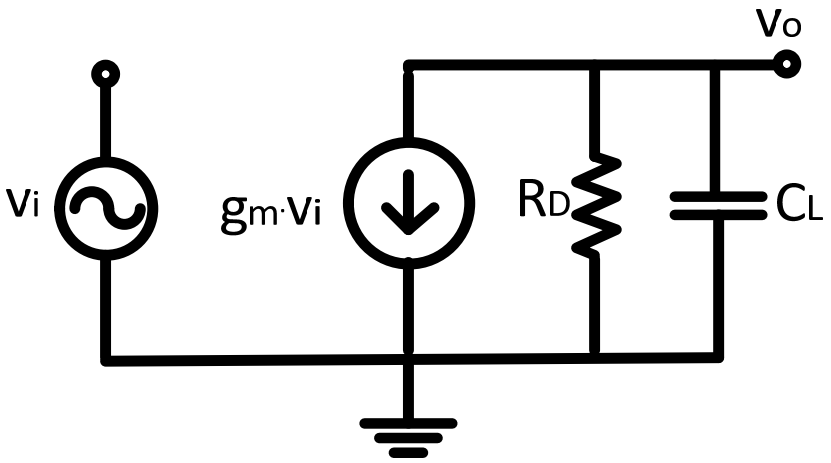

Fig. 4.43 Small signal equivalent circuit of the circuit of the example 4.6.

If the channel modulation parameter is not specified, assume that $\frac{1}{g_{ds}} >> R_D$. Therefore the output transistor conductance, g_{ds}, can be neglected for the drawing of the small signal equivalent circuit, and, then, for the calculation of the voltage gain. Applying the Kirkoff's current law to the output node, it results that the current of the current generator passes through the output load consisting of the parallel of R_D and C_L, and producing a voltage drop equal to v_o. Therefore, v_o is calculated as follows:

Eq. 4.75 $$v_o = -g_m \cdot v_i \cdot R_D \| \frac{1}{s \cdot C_L} = -g_m \cdot v_i \cdot \frac{R_D}{1 + s \cdot R_D \cdot C_L}$$

The voltage gain is, then, calculated:

Eq. 4.76 $$\frac{v_o}{v_i}(s) = -g_m \cdot \frac{R_D}{1 + s \cdot R_D \cdot C_L}$$

As equation 4.76 shows, the voltage gain has a pole, p_1:

Eq. 4.77 $$p_1 = -\frac{1}{R_D \cdot C_L} = -25\text{krad/s}$$

Calculate the absolute value of the voltage dc-gain, $\left|\frac{v_o}{v_i}(0)\right|$, by combining equations 4.76 and 4.52:

Eq. 4.78 $$\left|\frac{v_o}{v_i}(0)\right| = g_m \cdot R_D = 2 \cdot \frac{I_{DS}}{V_{ov}} \cdot R_D = 2 \cdot \frac{I_{DS}}{V_{GS} - V_{TH}} \cdot R_D = 2 \cdot \frac{I_{DS}}{V_G - V_{TH}} \cdot R_D = \ 20 = 26dB$$

Figure 4.44 shows the asymptotic Bode diagram of the voltage gain. At low frequency, the magnitude has a constant value equal to the voltage dc-gain. Starting from the pole frequency, $|p_1|$, the voltage gain decreases with a slope of -20dB/decade. At low frequency the phase is -180° due the negative sign of the voltage gain. The pole p_1 produces a phase shift of -90°, leading the phase to -270° at higher frequencies.

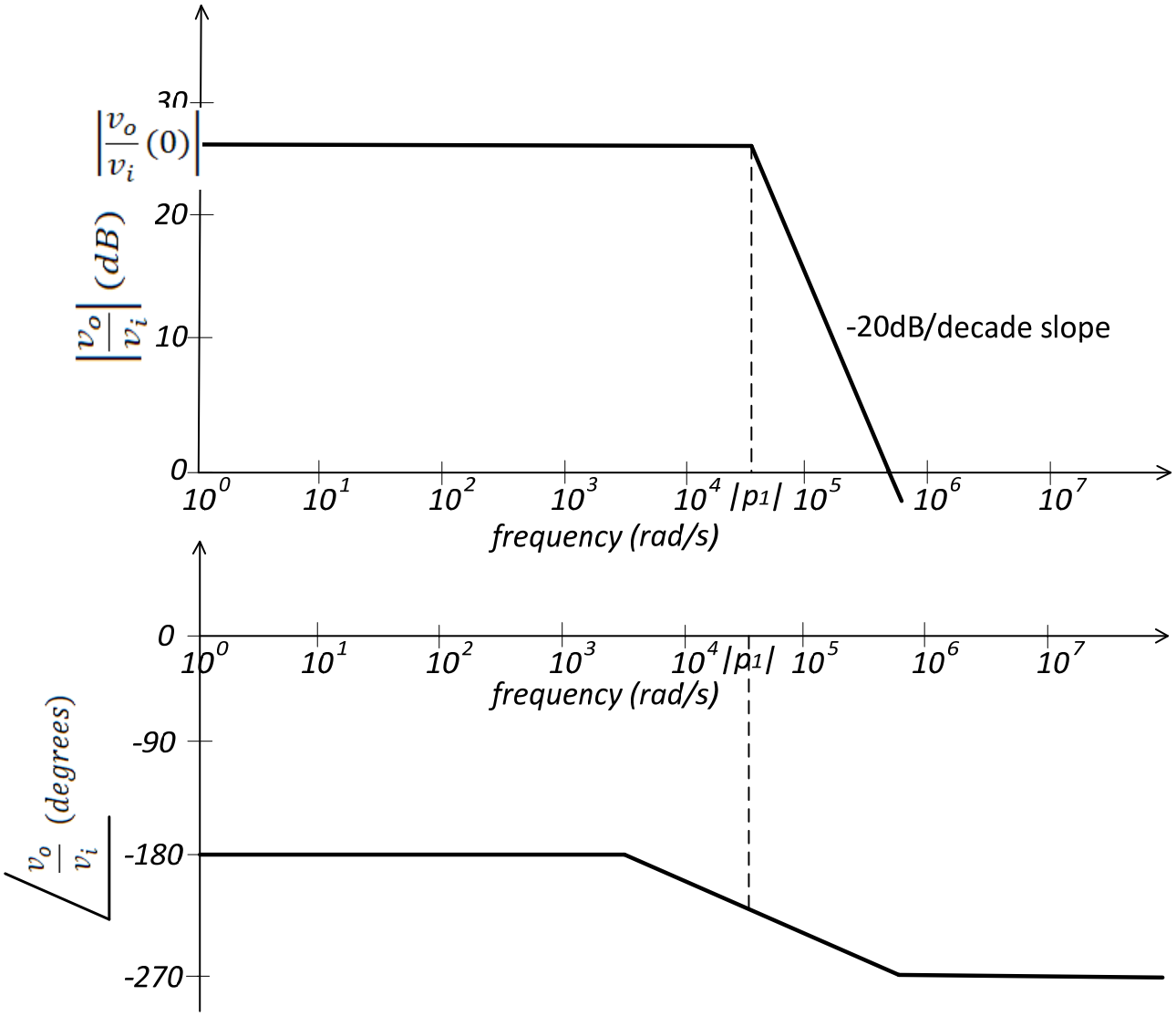

Fig. 4.44 Asymptotic Bode diagram of the voltage gain for the example 4.6.

In order to solve the last point of the example, consider the condition on V_{ds} reported in table 4.1 in order to keep on the NMOS transistor, M_1, operating in saturation region, i.e.:

Eq. 4.79 $$V_{ds} > V_{gs} - V_{TH} \rightarrow V_O + v_o > V_G + v_i - V_{TH}$$

But at dc:

Eq. 4.80 $$v_o = \frac{v_o}{v_i}(0) \cdot v_i = -g_m \cdot R_D \cdot v_i$$

Therefore, by combining equations 4.79 and 4.80 it is get:

Eq. 4.81 $$V_O - g_m \cdot R_D \cdot v_i > V_G + v_i - V_{TH}$$

Disequation 4.81 is satisfied until v_i is kept lower than the edge value $v_{i,max}$ which is calculated by imposing:

Eq. 4.82 $$V_O - g_m \cdot R_D \cdot v_{i,max} = V_G + v_{i,max} - V_{TH} \rightarrow v_{i,max} = \frac{V_O - V_G}{1 + g_m \cdot R_D}$$

By replacing the values of V_O, V_G and $g_m \cdot R_D$ found in equations 4.71, 4.73, and 4.78, in equation 4.82, it is get:

Eq. 4.83 $$v_{i,max} = 167mV$$

8.2 Common Drain configuration

The common drain amplifier is shown in figure 4.45.

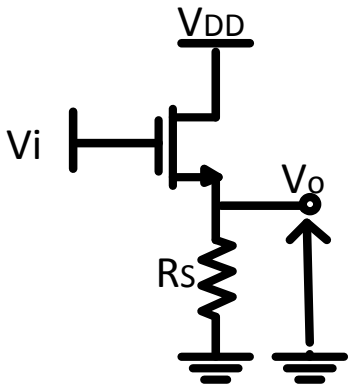

Fig. 4.45 Common drain amplifier.

As Figure 4.45 shows, the Drain is connected to V_{DD} which is ground for the signal, i.e. the Drain node is common to the ground network of the small signal equivalent circuit. The input signal is applied to the Gate, while the output signal is taken from the Source.

The output voltage, V_O, is calculated as follows:

Eq. 4.84 $$V_o = R_S \cdot I_{DS}$$

As $V_i < V_{TH}$, no current flows through the NMOS transistor, therefore:

Eq. 4.85 $$I_{DS} = 0 \text{ and } V_o = 0$$

As V_i is increased beyond the threshold voltage, V_{TH}, the NMOS transistor starts conducting. The NMOS transistor operates in saturation region until $V_i > V_{DD} + V_{TH}$. According to equation 4.14, as the NMOS transistor is in saturation region, the Drain-Source current is given by:

Eq. 4.86 $$I_{DS} = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_i - V_o - V_{TH})^2$$

Combining equations 4.84 e 4.86 the following input-output relationship is get:

Eq. 4.87 $$V_o = R_S \cdot I_{DS} = R_S \cdot k_n \cdot (V_i - V_o - V_{TH})^2$$

From equation 4.87 it is get:

Eq. 4.88 $$V_o = V_i - V_{TH} + \frac{1}{2 \cdot R_S \cdot k_n} - \sqrt{V_i \cdot \left(\frac{1}{R_S \cdot k_n} - V_{TH}\right) + \frac{1}{4 \cdot k_n^2 \cdot R_S^2} - \frac{V_{TH}}{R_S \cdot k_n}}$$

The resulting input-output characteristic is reported in figure 4.46.

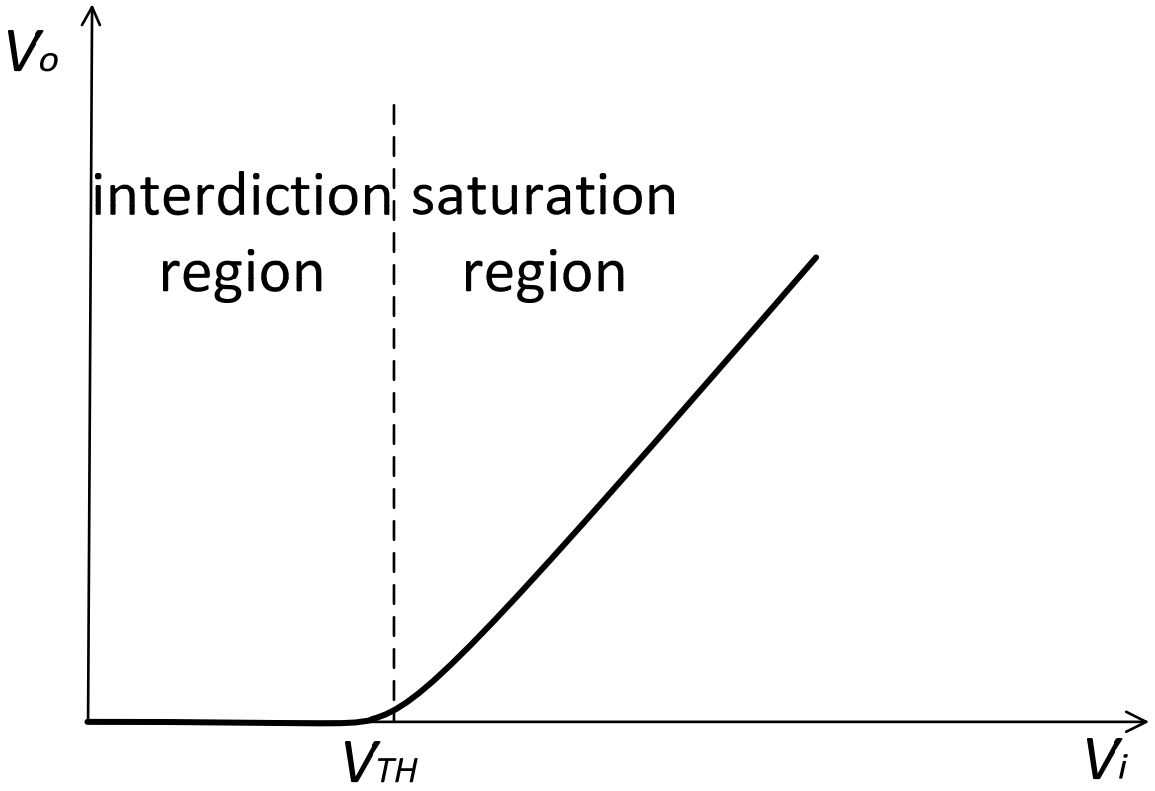

Fig. 4.46 Input-output characteristic of a common Dain amplifier.

Assuming that a small voltage signal is applied to the Gate, figure 4.47 shows the corresponding small signal equivalent circuit:

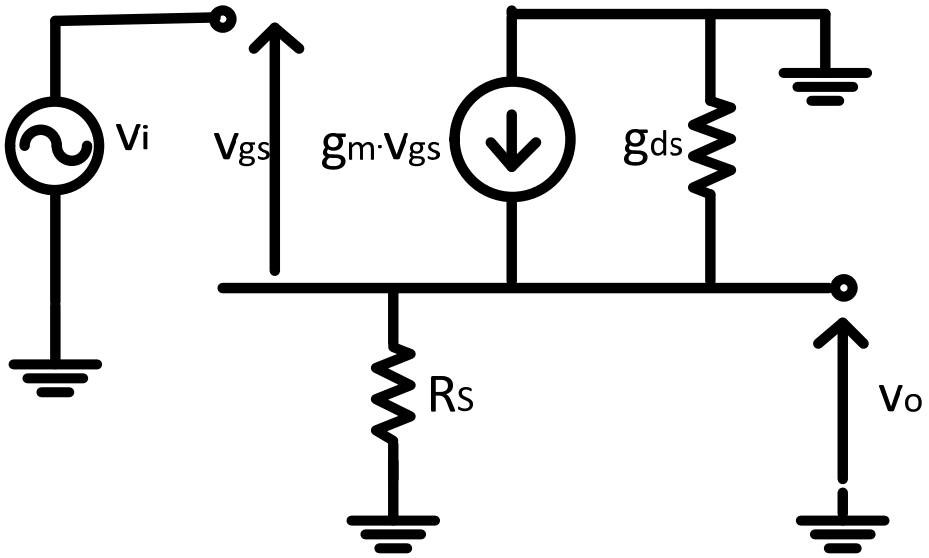

Fig. 4.47 Small signal equivalent circuit of common Drain amplifier.

The output voltage, v_o, is given by the voltage drop on the resistive load, $\frac{1}{g_{ds}} \| R_S$. As the current flowing in the resistive load is that provided by the current generator, the output voltage is calculated:

Eq. 4.89 $$v_o = g_m \cdot v_{gs} \cdot \frac{1}{g_{ds}} \Big\| R_D$$

But $v_{gs} = v_i - v_o$, therefore, the voltage gain, A_V, is calculated as follows:

Eq. 4.90 $$A_V = \frac{v_o}{v_i} = \frac{g_m \cdot \frac{1}{g_{ds}} \Big\| R_S}{1 + g_m \cdot \frac{1}{g_{ds}} \Big\| R_S}$$

As $R_S \ll \frac{1}{g_{ds}}$, then the voltage gain becomes:

Eq. 4.91 $$A_V = \frac{v_o}{v_i} \approx \frac{g_m \cdot R_S}{1 + g_m \cdot R_S}$$

The voltage gain increases by augmenting R_S. A least, as $g_m \cdot R_S \gg 1$, the following voltage gain is get:

Eq. 4.92 $$A_V = \frac{v_o}{v_i} \approx 1$$

As the gain is about unitary, this gain stage is also called Source follower (it is like the Source voltage "follows" the Gate voltage). The input signal is applied to the Gate terminal, therefore the input resistance of the amplifier is about infinite. The output resistance is calculated by considering the small signal equivalent circuit reported in Fig. 4.48.

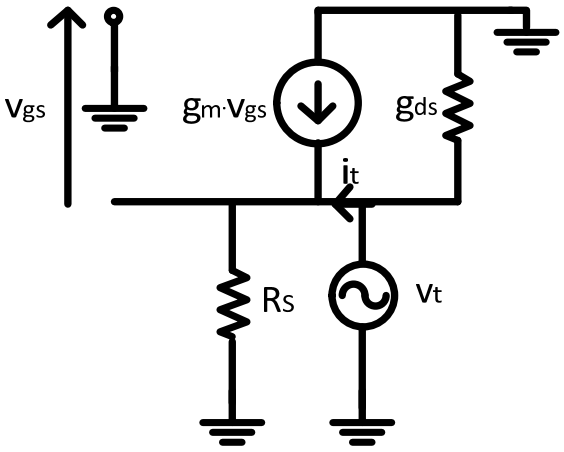

Fig. 4.48 Small signal equivalent circuit for the calculation of the output resistance of a common Drain amplifier.

For the calculation of the output resistance, R_o, the input signal, v_i, is nulled by definition. Therefore, the current of the current generator is equal to $g_m \cdot v_t$. In practical, the current generator produces a current proportional to the voltage drop on its ends (v_t), therefore it behaves like a resistor with $1/g_m$ value. The output resistance, R_o, is calculated as the parallel of $1/g_m$, R_S and $1/g_{ds}$, i.e. :

Eq. 4.93 $$R_o = \frac{v_t}{i_t} = \frac{1}{g_m} \Big\| R_S \Big\| \frac{1}{g_{ds}}$$

Assuming $1/g_m \ll R_S \Big\| \frac{1}{g_{ds}}$, the output resistance, R_o, is as small as $1/g_m$. As this gain stage has an high input impedance, a low output impedance and an about unitary gain, it is often used as voltage buffer to separate two cascaded stages in order to limit the load effects of the second stage on the first one.

- **Example 4.7**

Problem. Consider the circuit reported in fig. 4.49. Assume V_{TH}=1V, K_n=5mA/V^2, V_{DD}=10V, R_S=4kΩ, R_L=4kΩ, R_2=8.5kΩ, R_1=1.5kΩ, C_o=10nF, C_i=4μF. Calculate the bias point and the voltage gain versus the frequency, $\frac{v_o}{v_i}(s)$. Trace the Bode diagram of the voltage gain.

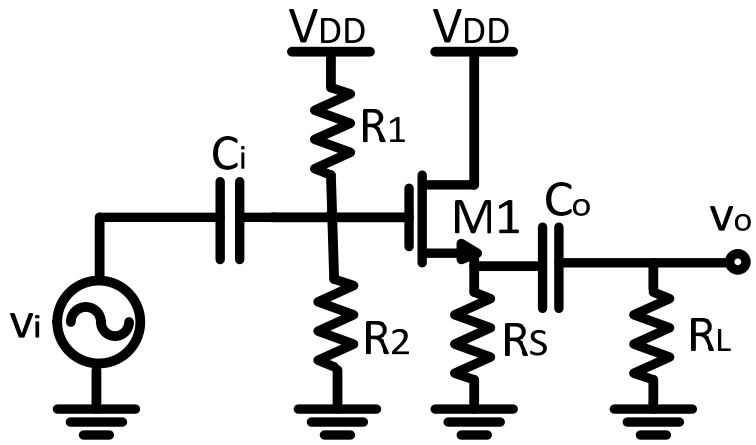

Fig. 4.49 Circuit to be solved in the example 4.7.

Solution. To calculate the bias point it is possible to simplify the circuit by nulling the capacitances and the input signal, as it is done in figure 4.50.

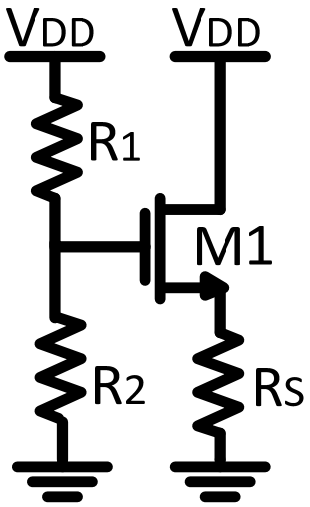

Fig. 4.50 Circuit of the example 4.7 without capacitances and with a null input signal.

As the Gate does not adsorb any current, the Gate voltage V_G, is calculated as a partition of V_{DD} on R_2 resistor:

Eq. 4.94 $$V_G = V_{GS} = V_{DD} \cdot \frac{R_2}{R_2+R_1} = 8.5V$$

As V_G=V_{GS}>V_{TH} the transistor is switched on. Suppose that the NMOS transistor works in saturation region. Therefore, equation for NMOS transistor operating in saturation region reported in equation 4.14 can be used to determine I_{DS}, i.e.:

Eq. 4.95 $$I_{DS} = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_G - V_S - V_{TH})^2$$

But V_S is determined by the voltage drop on R_S resistor, i.e.:

Eq. 4.96 $$V_S = R_S \cdot I_{DS}$$

By replacing V_S in equation 4.95 by the expression found in equation 4.96, a second order equation is get where the only variable is I_{DS}:

Eq. 4.97 $$I_{DS} = k_n \cdot (V_G - R_S \cdot I_{DS} - V_{TH})^2$$

Two solutions are possible for equation 4.97,:

Eq. 4.98 $$I_{DS,1} = 1.73mA \text{ and } I_{DS,2} = 2.03mA$$

But the only $I_{DS,1}$ has a physical significant. $I_{DS,2}$ is not a feasible solution. In fact, if I_{DS}=$I_{DS,2}$, V_{GS} results to be less than a threshold voltage V_{TH}, which means that the NMOS transistor should be switched off, and, at the same time, it should be passed by a non-null current.

Once that the circuit has been solved, it is required to verify the initial hypothesis, i.e. the NMOS transistor working in saturation region. It means that the following condition on the drain source voltage V_{DS}, reported in table 4.1, is to be verified:

Eq. 4.99 $$V_{DS} > V_{GS} - V_{TH} \rightarrow V_{DD} > V_G - V_{TH} \rightarrow 10V > 6.5V$$

As this condition is satisfied, the initial hypothesis of NMOS transistor working in saturation region is verified.

In order to calculate the voltage gain, the small signal equivalent circuit is reported in figure 4.51.

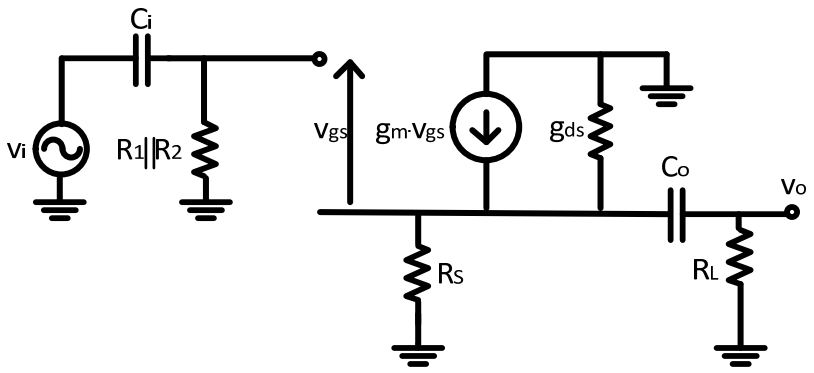

Fig. 4.51 Small signal equivalent circuit of the circuit of the example 4.7.

If the channel modulation parameter is not specified, assume that $\frac{1}{g_{ds}} >> R_S$. Therefore the output transistor conductance, g_{ds}, can be neglected. In order to simplify the gain calculation it is possible to express the voltage gain, A_v, as follows:

Eq. 4.100 $$A_v = \frac{v_o}{v_i}(s) = \frac{v_o}{v_s}(s) \cdot \frac{v_s}{v_g}(s) \cdot \frac{v_g}{v_i}(s)$$

where:

Eq. 4.101 $$\frac{v_g}{v_i} = \frac{s \cdot C_i \cdot R_1 \| R_2}{1 + s \cdot C_i \cdot R_1 \| R_2}.$$

The overall impedance, Z_s, connected to the source of the NMOS transistor is calculated as follows:

Eq. 4.102 $$Z_S = R_S \| \left(R_L + \frac{1}{s \cdot C_o} \right)$$

By considering equation 4.91, and replacing R_S with Z_S the following $\frac{v_s}{v_g}$ gain is get:

Eq. 4.103 $$\frac{v_s}{v_g} = \frac{g_m \cdot Z_S}{1 + g_m \cdot Z_S}.$$

The $\frac{v_s}{v_g}$ gain is get as partition of V_S voltage on R_L, i.e.:

Eq. 4.103 $$\frac{v_o}{v_s} = \frac{s \cdot C_o \cdot R_L}{1 + s \cdot C_o \cdot R_L}$$

Therefore, by replacing equations 4.101, 4.102, 4.103 in equation 4.100, the overall voltage gain, A_V, is get as follows:

Eq. 4.104 $$A_v = \frac{s \cdot C_i \cdot R_1 \| R_2}{1 + s \cdot C_i \cdot R_1 \| R_2} \cdot \frac{g_m \cdot R_S \cdot R_L \cdot s \cdot C_o}{1 + g_m \cdot R_S + s \cdot C_o \cdot (R_L + R_S + g_m \cdot R_L \cdot R_S)}$$

As equation 4.104 shows, the voltage gain, A_v, has two zero at the null frequency and two poles, p_1 and p_2 :

Eq. 4.105 $$p_1 = -\frac{1 + g_m \cdot R_S}{C_o \cdot (R_L + R_S + g_m \cdot R_L \cdot R_S)} = -25\text{krad/s}; \qquad p_2 = -\frac{1}{R_1 \| R_2 \cdot C_i} = -196\text{rad/s}.$$

From equation 4.104, calculate the absolute value of the voltage gain at high frequencies ($s \to \infty$):

Eq. 4.106 $$A_v(s \to \infty) = \frac{g_m \cdot R_S \cdot R_L}{R_L + R_S + g_m \cdot R_L \cdot R_S} \cong 1 = 0dB$$

Figure 4.52 shows the asymptotic Bode diagram of the voltage gain, A_v. The two zeros at the null frequency introduce a 40dB/decade slope of the gain curve at low frequency. The two poles p_1 and p_2 determine a gain slope variations of -20dB/decade each. At high frequencies, the magnitude has a constant value about equal to 0dB. At low frequency the phase is 180° due the two zeros. The two poles introduce a phase shift of -90° each, leading the phase to 0° at higher frequencies.

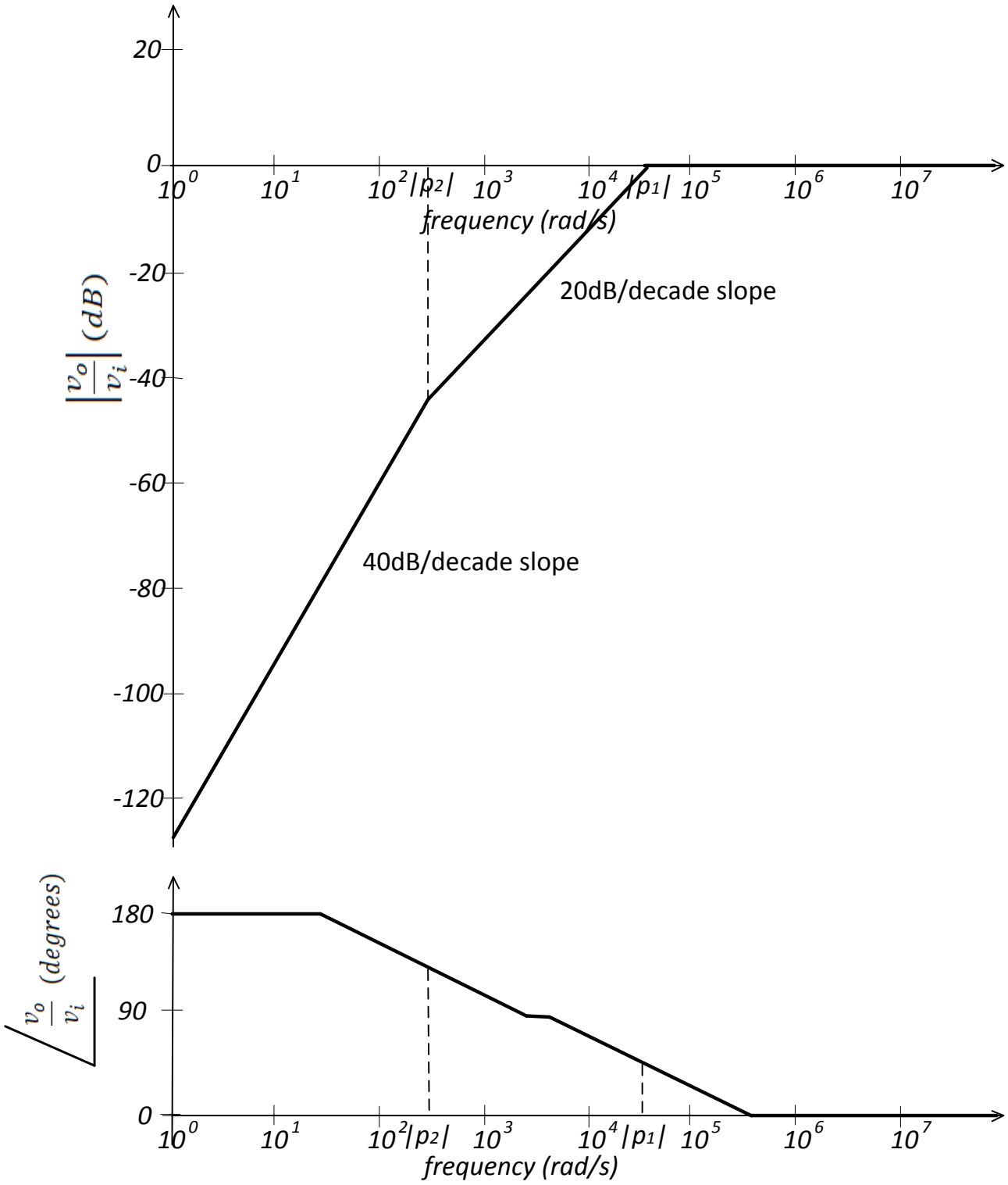

Fig. 4.52 Asymptotic Bode diagram of the voltage gain for the example 4.7.

8.3 Common Gate configuration

The common Gate amplifier is shown in figure 4.53.

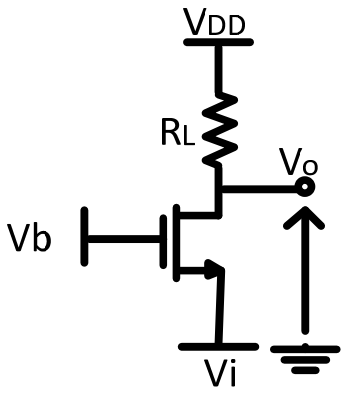

Fig. 4.53 Common Gate amplifier.

As Figure 4.53 shows, the Gate is connected to a fixed voltage V_b which is ground for the signal, i.e. the Gate node is common to the ground network of the small signal equivalent circuit. The input signal is applied to the Source, while the output signal is taken from the Drain.

The output voltage, V_O, is calculated as follows:

Eq. 4.107 $$V_o = V_{DD} - R_D \cdot I_{DS}$$

As $V_i > V_b - V_{TH}$, no current flows through the NMOS transistor, therefore:

Eq. 4.108 $$I_{DS} = 0 \text{ and } V_o = V_{DD}$$

As V_i is decreased beyond $V_b - V_{TH}$, the NMOS transistor starts operating in saturation region. The Drain-Source current , I_{DS}, is given by:

Eq. 4.109 $$I_{DS} = k_n \cdot (V_{GS} - V_{TH})^2 = k_n \cdot (V_b - V_i - V_{TH})^2$$

According to equation 4.108 the following input-output relationship is get:

Eq. 4.110 $$Vo = V_{DD} - R_D \cdot k_n \cdot (V_b - V_i - V_{TH})^2$$

Equation 4.111 is valid until the NMOS transistor keep on operating in saturation region, i.e. until the following condition is satisfied:

Eq. 4.111 $$V_{DS} > V_{GS} - V_{TH} \rightarrow Vo > V_b - V_{TH} \rightarrow V_{DD} - R_D \cdot k_n \cdot (V_b - V_i - V_{TH})^2 > V_b - V_{TH}$$

The previous disequation is satisfied as V_i is higher than the limit value $V_{i,min}$:

Eq. 4.112 $$V_{i,min} = V_b - V_{TH} - \sqrt{\frac{V_{DD} - V_b + V_{TH}}{k_n \cdot R_D}}$$

Fo $V_i < V_{i,min}$ the NMOS transistor operates in linear region.

The resulting input-output characteristic is reported in figure 4.54.

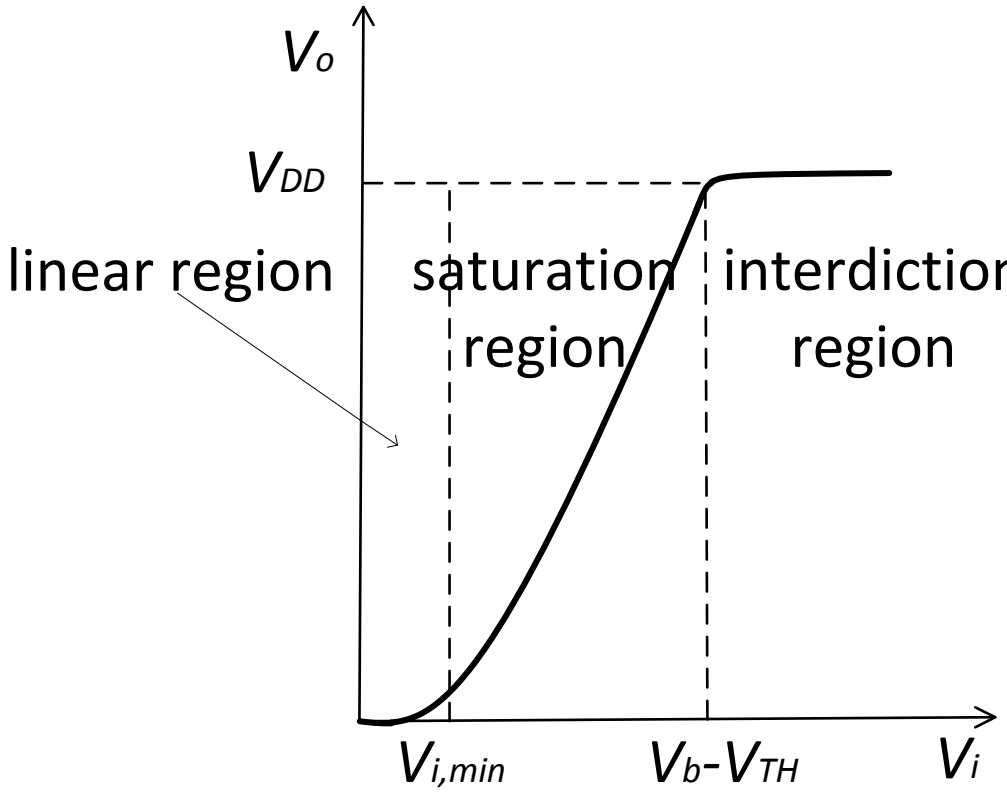

Fig. 4.54 Input-output characteristic of a common Gate amplifier.

Assuming that a small voltage signal is applied to the Source, figure 4.55 shows the corresponding small signal equivalent circuit:

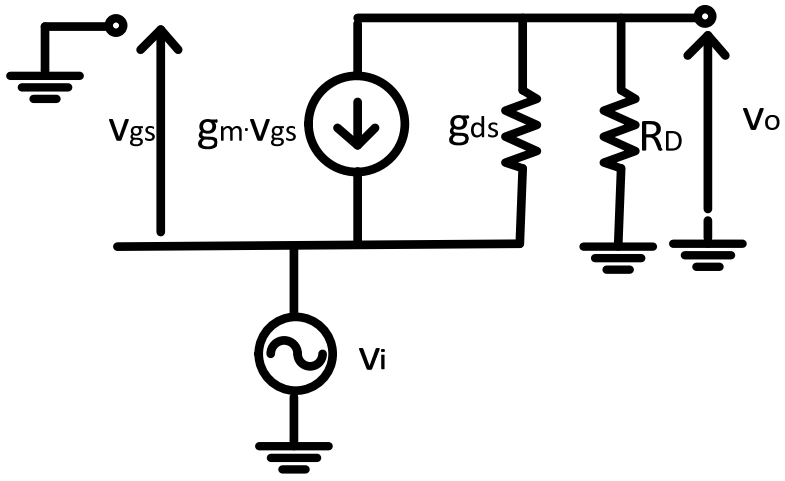

Fig. 4.55 Small signal equivalent circuit of common Gate amplifier.

Assuming $\frac{1}{g_{ds}}$ negligible, the output voltage, v_o, is given by the voltage drop on the resistive load, R_D. As the current flowing in the resistive load is provided by the current generator, the output voltage is calculated as follows:

Eq. 4.113 $$v_o = -g_m \cdot v_{gs} \cdot R_D$$

But $v_{gs} = -v_i$, therefore, the voltage gain, A_V, is calculated as follows:

Eq. 4.114 $$A_V = \frac{v_o}{v_i} = g_m \cdot R_D$$

The input resistance is calculated by considering the small signal equivalent circuit reported in Fig. 4.56.

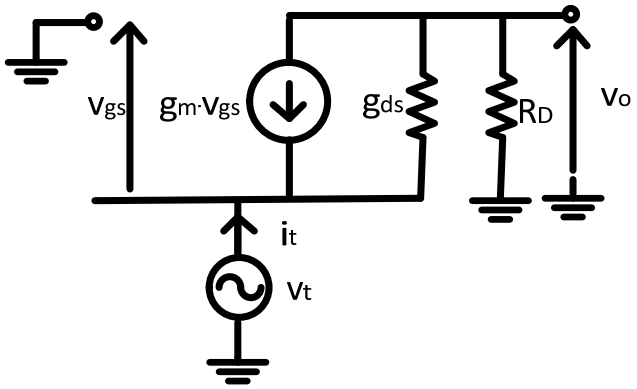

Fig. 4.56 Small signal equivalent circuit for the calculation of the input resistance of a common Gate amplifier.

By applying the Kirkoff's current law to the Source node, and neglecting the current flowing in g_{ds} as it is assumed to be very small, the following equation is get:

Eq. 4.115 $$i_t = -g_m \cdot v_{gs} = g_m \cdot v_t$$

From equation 4.115 the input resistance, R_i, is get:

Eq. 4.116 $$R_i = \frac{v_t}{i_t} = \frac{1}{g_m}.$$

The output resistance, R_o, is calculated by considering the small signal equivalent circuit reported in Fig. 4.57.

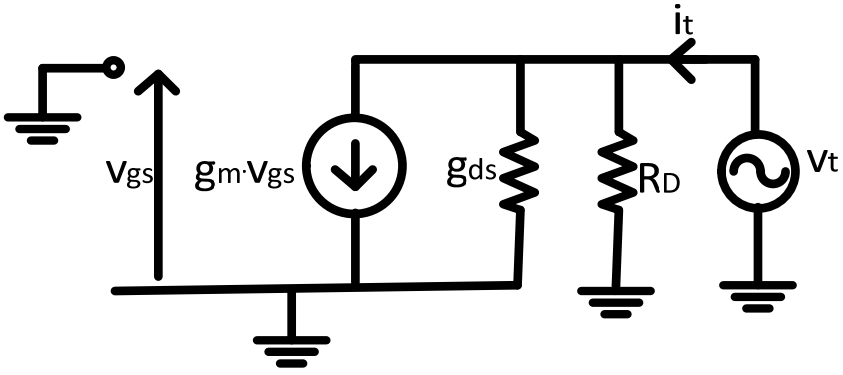

Fig. 4.57 Small signal equivalent circuit for the calculation of the output resistance of a common Gate amplifier.

For the calculation of the output resistance, R_o, the input signal, v_i, is nulled by definition. Therefore, the current of the current generator is null. Therefore, from the output node, it is seen the parallel of R_D and $1/g_{ds}$, i.e.:

Eq. 4.117 $$R_o = \frac{v_t}{i_t} = \frac{1}{g_m} \Big\| R_D$$

Assuming $1/g_{ds}$<<R_D, the output resistance, R_o, is about equal to R_D.

- **Example 4.8**

Problem. Consider the circuit reported in fig. 4.58. Assume V_{TH}=1V, K_n=5mA/V^2, V_{DD}=10V, R_D=4kΩ, V_b=5V, V_{DD}=10V. Design the amplifier in order to have a voltage gain, $A_v = \frac{v_o}{v_i} = 20dB$ in the frequency range between 20Hz and 20kHz.

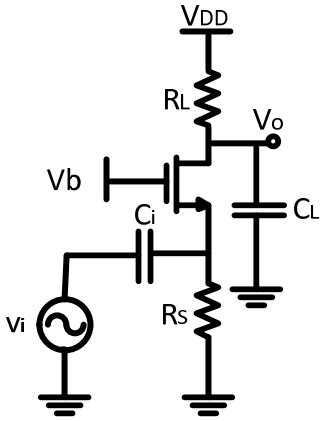

Fig. 4.58 Circuit to be solved in the example 4.8.

Solution. Assume that the NMOS transistor operates in saturation region. Consider the small signal equivalent circuit reported in figure 4.59, in order to calculate the voltage gain A_v.

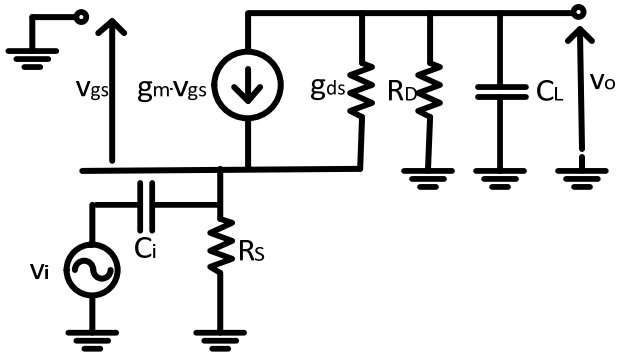

Fig. 4.59 Small signal equivalent circuit of the circuit of the example 4.8.

To simplify the calculation it is possible to decompose the voltage gain as follows:

Eq. 4.118 $$A_v = \frac{v_o}{v_i}(s) = \frac{v_o}{v_s}(s) \cdot \frac{v_s}{v_i}(s)$$

From equation 4.114 it is possible to derive $\frac{v_o}{v_s}(s)$ by replacing R_D with $Z_L = R_L \| C_L$, i.e.:

Eq. 4.119 $$\frac{v_o}{v_s}(s) = g_m \cdot Z_L = g_m \cdot R_L \| C_L = \frac{g_m \cdot R_L}{1 + s \cdot R_L \cdot C_L}$$

To calculate $\frac{v_s}{v_i}(s)$ consider the circuit reported in figure 4.60, where the Thevenin's equivalent circuit seen from the Source of the NMOS transistor has been considered instead of the overall small signal equivalent circuit. The Thevenin's equivalent circuit corresponds to the input resistance ($= 1/g_m$) of the common Gate amplifier.

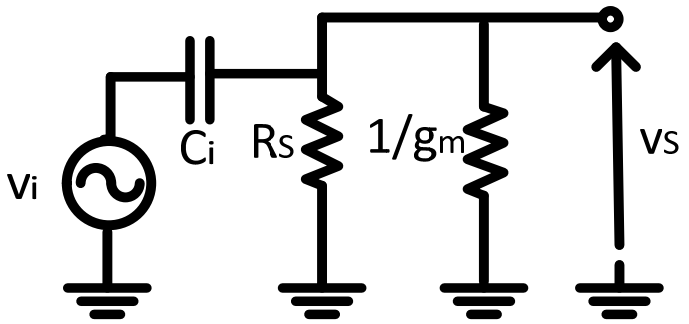

Fig. 4.60 Small signal equivalent circuit of the example 4.8 with the Thevenin's equivalent circuit seen from the Source of the NMOS transistors.

The v_s voltage is obtained as partition of the input voltage, v_i, on $R_S \| \frac{1}{g_m}$, therefore:

Eq. 4.120
$$\frac{v_s}{v_i}(s) = \frac{R_S \| \frac{1}{g_m}}{\frac{1}{s \cdot C_i} + R_S \| \frac{1}{g_m}} = \frac{s \cdot C_i \cdot R_S \| \frac{1}{g_m}}{1 + s \cdot C_i \cdot R_S \| \frac{1}{g_m}}$$

By combining equations 4.118, 4.119, and 4.120 it is get:

Eq. 4.121
$$A_v = \frac{v_o}{v_i}(s) = \frac{v_o}{v_s}(s) \cdot \frac{v_s}{v_i}(s) = \frac{g_m \cdot R_L}{1 + s \cdot R_L \cdot C_L} \cdot \frac{s \cdot C_i \cdot R_S \| \frac{1}{g_m}}{1 + s \cdot C_i \cdot R_S \| \frac{1}{g_m}}$$

The decoupling capacitance C_i generates a zero at the null and a low frequency pole, p_1, while the C_L capacitance generates an high frequency pole p_2:

Eq. 4.122
$$p_1 = -\frac{1}{C_i \cdot R_S \| \frac{1}{g_m}}; \qquad p_2 = -\frac{1}{R_L \cdot C_L}.$$

In the frequency range between the two poles, i.e. $|p_1| < \omega < |p_2|$, the voltage gain is calculated as follows:

Eq. 4.123
$$A_v(|p_1| < \omega < |p_2|) \cong g_m \cdot R_L$$

According to the specifications, $A_v(|p_1| < \omega < |p_2|)$ has to be equal to 20dB, therefore:

Eq. 4.124
$$g_m = \frac{A_v(|p_1| < \omega < |p_2|)}{R_L} = \frac{10}{R_L} = 2.5 \frac{mA}{V}$$

According to eq. 4.52, from the transconductane g_m and the conducibility factor K_n of the NMOS transistor, it is get the value of the Drain-Source current, I_{DS}:

Eq. 4.125
$$I_{DS} = \frac{g_m^2}{4 \cdot k_n} = 312.5 \mu A$$

According to eq. 4.14, the NMOS transistor overdrive, V_{ov}, is given by:

Eq. 4.126
$$V_{ov} = \sqrt{\frac{I_{DS}}{k_n}} = 250 mV$$

With reference to fig. 4.59, the Source voltage at dc, i.e. considering the capacitor C_i as an open circuit, is calculated by applying the Kirkoff's law for voltages, i.e.:

Eq. 4.127 $$V_S = V_b - V_{GS} = V_b - V_{TH} - V_{ov} = 3.750V$$

At dc, the R_S resistor is passed by the I_{DS} current, while the V_S voltage drops at its ends. Therefore its value is get as follows:

Eq. 4.128 $$R_S = \frac{V_S}{I_{DS}} = 12k\Omega$$

By specifications the lower frequency pole, p_1, has to be set at 20Hz, therefore:

Eq. 4.129 $$|p_1| = \frac{1}{C_i \cdot R_S \| \frac{1}{gm}} \rightarrow C_i = \frac{1}{|p_1| \cdot R_S \| \frac{1}{gm}} = 10\mu F$$

By specifications the lower frequency pole, p_2, has to be set at 20kHz, therefore:

Eq. 4.130 $$|p_2| = \frac{1}{R_L \cdot C_L} \rightarrow C_L = \frac{1}{|p_2| \cdot R_L} = 2nF$$

The last check to do is on the operation region of the NMOS transistor. It has been assumed that the NMOS transistor operates in saturation region. This happens if the following condition is satisfied:

Eq. 4.131 $$V_{DS} > V_{GS} - V_{TH} \rightarrow V_D > V_b - V_{TH} = 4V$$

As $V_D = V_{DD} - R_L \cdot I_{DS} = 8.75V$, the previous disequation is satisfied.