Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

1 Introduction

Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is a unipolar, three terminal device (Body being the fourth terminal). The conduction of current takes place by only one type of charge carriers (either holes or electrons), depending upon the induced electric field which is controlled by the voltage applied to the gate terminal. There are two types of MOSFETs: *a*) Enhancement type, *b*) Depletion Type. These are further divided into two types: NMOS and PMOS depending upon the charge carrier used for conduction. Enhancement type MOSFETs are widely used. We will discuss the structure, operation and characteristics of the same in this experiment.

Structure of an enhancement type MOSFET:

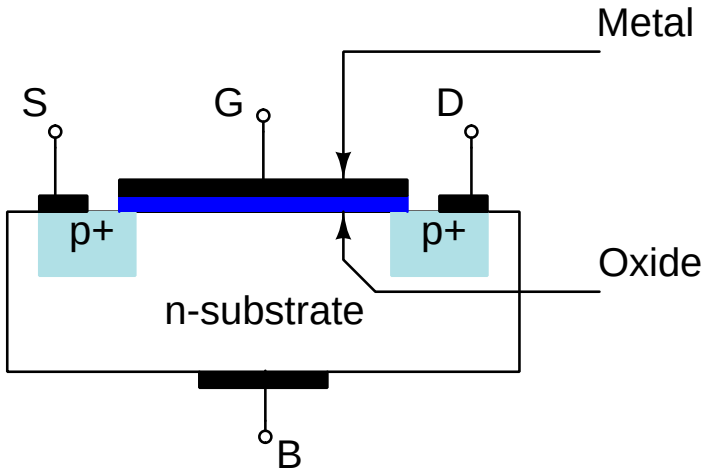

Figure 1: Physical structure of an PMOS enhancement MOSFET

The PMOS transistor has a n-type substrate that provides physical support to the device. Two heavily doped p^+ regions are created in the substrate. They are named as source region and drain region. A thin layer of SiO_2 $(2nm - 50nm)$ is grown on the surface of substrate covering the area between source and drain as shown in Fig 1. This layer is an excellent insulator and acts as Gate electrode. Metal contacts are drawn from source, drain and Gate and substrate (sometimes referred to as body).

Formation of the channel

The two p^+n junctions formed by source (p^+) with substrate (n) and drain (p^+) with substrate (n) are in series opposition and hence no current can flow between drain and source if voltage is applied between drain and source.

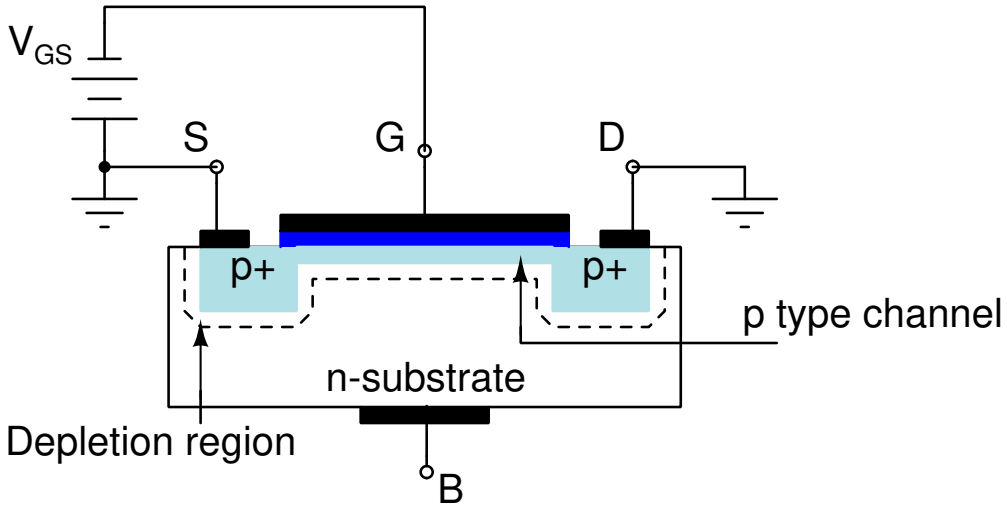

Figure 2: Channel formation

Now assume that no voltage is applied between drain and source and a **small negative** voltage V_{GS} is applied on the gate terminal with respect to source. This negative voltage causes the electrons in the substrate to get repelled by the negative gate voltage and to get pushed downwards into substrate region forming a depletion region. Since the electrons in this region have moved away, the region has a positive charge. The negative gate voltage also causes the holes in source and drain region to get accumulated under SiO_2 layer. If voltage V_{GS} is sufficiently negative, a large number of holes get accumulated under the SiO_2 layer on the surface of substrate so as to form

channel between the source and the drain as shown in Fig. 2. This induced channel is formed due to holes (p-type) hence called p-channel MOSFET.

Operation of a PMOS transistor

Once the channel is formed, application of a voltage between drain and source induces flow of current through the channel. Since the electric field in the channel is due to V_{GS}, by controlling V_{GS} one can control the electric field in the channel which in effect, controls the conductivity of the channel and hence the drain current.

The voltage V_{GS} at which a sufficient number of holes accumulate to form a channel is called V_T (threshold voltage).

At $V_{GS} = V_T$, the channel has just induced and the drain current is extremely small. If V_{GS} goes more negative than V_T, more holes are attracted towards the channel, increasing the conductivity of the channel and also the depth. The drain current is thus proportional to the "excess voltage" (V_{GS} - V_T) that causes the channel to form and V_{DS} causes I_D to flow.

I_D-V_{DS} characteristics

These characteristics are obtained by measuring values of I_D for different values of V_{DS} at constant V_{GS}. The I_D-V_{DS} characteristics are as shown in Fig.3.

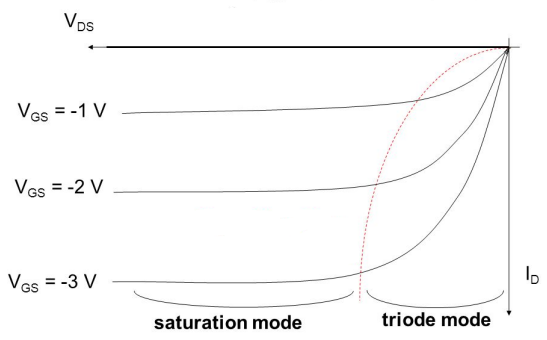

Figure 3: Output Characteristics of PMOS

There are three different regions of operation for a P-MOSFET depending on the relative voltages on its terminals.

V_T for a PMOS is NEGATIVE

- Cut-off region, when $|V_{GS}| \leqslant |V_T|$ i.e. $V_{GS} \geqslant V_T$
- Triode region, when $V_{DS} > V_{GS} - V_T$ and $V_{GS} < V_T$
- Saturation region, when $V_{DS} \leqslant V_{GS} - V_T$ and $V_{GS} < V_T$

Cut-off region

As we have seen earlier that the channel is induced only when $V_{GS} \leqslant V_T$. Hence for voltages $V_{GS} \geqslant V_T$ the device can not conduct and acts as an OFF switch with $I_D = 0$.

Triode or linear region

For $V_{GS} \leqslant V_T$ but $V_{DS} > V_{GS} - V_T$, the part of the characteristic to the left $V_{DS} = V_{GS} - V_T$, is called triode region or linear region.

The drain current I_D of MOSFET depends on both V_{DS} and V_{GS}. The general drain current equation of MOSFET is given by is given by,

$$I_D = \frac{1}{2}K\Big[2(V_{GS} - V_T)V_{DS} - V_{DS}^2\Big] \tag{1}$$

where the constant K is given by,

$$K = \mu_p C_{ox}\Big(\frac{W}{L}\Big) \tag{2}$$

where, μ_p is hole mobility, C_{ox} is oxide layer capacitance, W and L represent width and length of the inversion channel. K is device parameter and has units A/V^2. For very small values of V_{DS} (linear region), the quadratic term in Eq.1 becomes negligibly small and the drain current is then given by,

$$I_D = \frac{1}{2}K\Big[2(V_{GS} - V_T)V_{DS}\Big] \tag{3}$$

This means, for very small values of V_{DS}, the relation between I_D and V_{DS} is linear. The transistor acts like voltage dependent resistor with value,

$$r_{DS} = \frac{V_{DS}}{I_D} = K\Big[(V_{GS} - V_T)\Big]^{-1} \tag{4}$$

The value of the resistance can be controlled by changing V_{GS}.

Saturation region

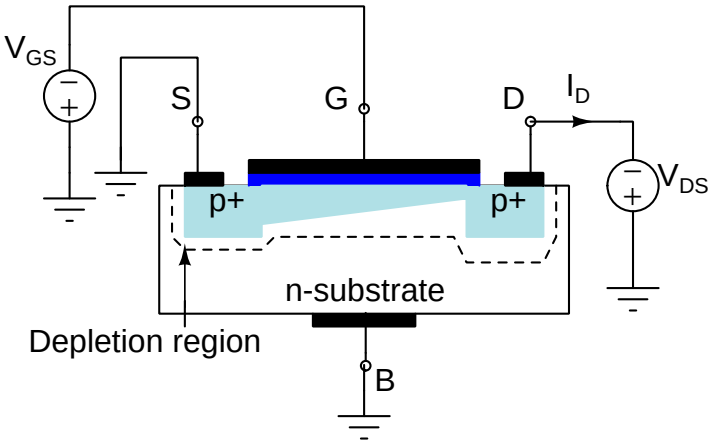

Figure 4: Tapered channel at higher values of V_{DS} with $V_{GS} \leq V_T$).

When V_{DS} is made further negative, ($V_{DS} \leqslant V_{GS} - V_T$) as mentioned earlier, the induced channel starts tapering off towards drain region as shown in Fig. 4. The depth of the channel decreases to almost zero near drain end and the channel is said to be pinched-off. The decrease in V_{DS}, beyond $(V_{GS} - V_T)$ does not result in significant increase in I_D due to pinch-off. The device is said to operate in saturation region. At the boundary i.e. at V_{DS} $= V_{GS} - V_T$ drain current is given by,

$$I_D = \frac{1}{2}K\left[2(V_{GS} - V_T)(V_{GS} - V_T) - (V_{GS} - V_T)^2\right] = \frac{1}{2}K(V_{GS} - V_T)^2 \quad (5)$$

In saturation region, I_D has a square relationship with V_{GS} and ideally independent of V_{DS}, i.e. at constant V_{GS}, if V_{DS} is taken further negative, then I_D should remain constant. The MOSFET can be used as an amplifier in this region.

Channel Length Modulation

In reality, the drain current I_D is not independent of V_{DS}. For larger negative values of V_{DS}, after pinch-off, the effective channel width is found to get reduced as the pinch-off point moves away from drain towards source. The effect of reduced channel length is to increase the value of K causing I_D to increase in it's magnitude as shown in Fig.5. Hence, the I_D/V_{DS} characteristics do not go horizontal beyond $V_{DS(sat)}$, but show some linear dependence

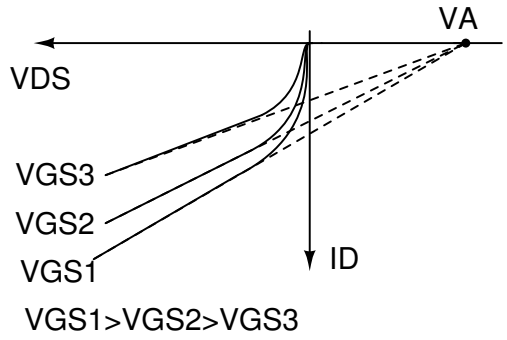

Figure 5: Increase in magnitude of I_D due to the effect of channel length modulation in Saturation region

given by,

$$I_D = \frac{1}{2}K(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \tag{6}$$

where, λ is a MOSFET parameter and is always positive. The intercept of the linear part of saturation region curve ($1+\lambda\ V_{DS}$) part of the characteristics), when extrapolated, on X-axis , is defined as "Early voltage". (Recall BJT output characteristics.) The reciprocal of slope of this characteristics gives output resistance $r_o = 1\ /\lambda$.

To extract the value of V_T

The I_D/V_{GS} characteristics are obtained at constant V_{DS}. These can be obtained by biasing the device in two regions:

1. **Linear region**:- when V_{DS} is very small (around -100mV). In this region, the current I_D is given by Eq.3. For constant value of V_{DS}, we expect a linear relationship between I_D and V_{GS}. However, we get the characteristic as depicted in Fig. 6. For small values of V_{GS}, the $V_{GS} - V_T$ difference is less than V_{DS}, and for this small region, the transistor is actually in saturation, while we expect it to be in linear region! Hence I_D varies non-linearly with V_{GS} for a small range of V_{GS}, after which it changes linearly. So we extrapolate the characteristic to estimate the value of threshold voltage as illustrated in Fig. 6.

2. **Saturation region**:- This is when V_{DS} goes more negative than $V_{DS(sat)}$. In this region the I_D/V_{GS} relation is quadratic. The Eq.5 can be rewritten as

$$\sqrt{I}_D = \sqrt{\frac{K}{2}}(V_{GS} - V_T) \tag{7}$$

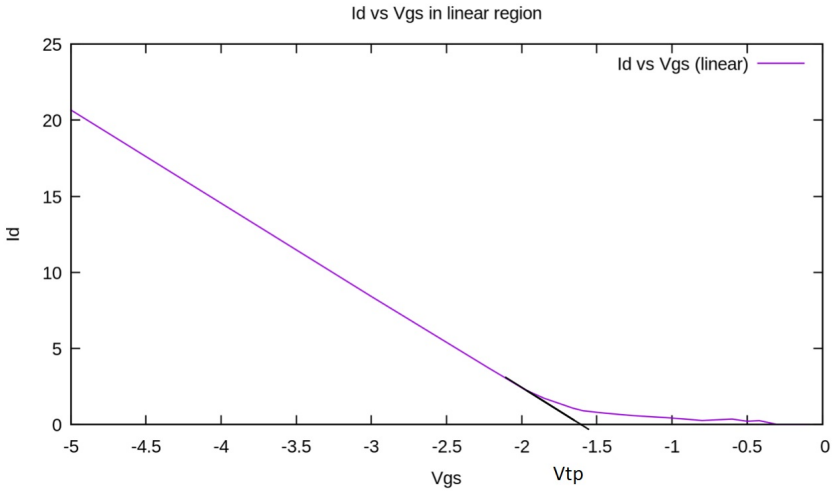

Figure 6: $I_D V_{GS}$ plot of MOSFET in linear region

This is the straight line with slope $\sqrt{\frac{K}{2}}$ and the intercept on X-axis is used to extract V_T.

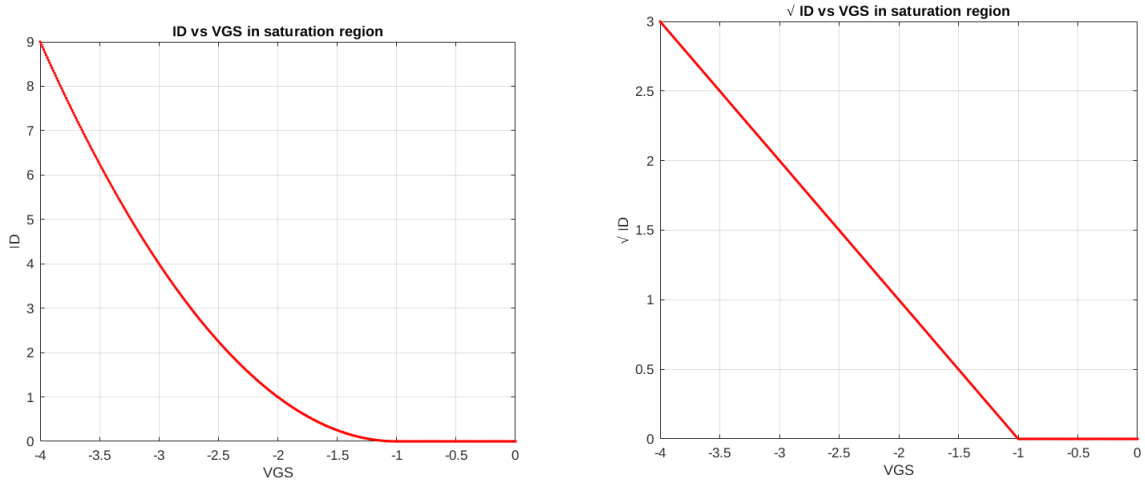

Figure 7: I_D vs. V_{GS} and $\sqrt{I_D}$ vs. V_{GS} plots of MOSFET in saturation region

Substrate (Body) Bias Effect

Source-Body(substrate) and Drain-Body junctions of a MOSFET can be visualised as two PN junctions. Since no current is expected to flow through these junctions, we need to reverse bias them. Connecting the Body of the PMOS to the highest potential and that of the NMOS to the lowest potential serves this purpose. The voltage applied to the Body terminal will influence the flow of minority carriers which are responsible for the channel formation and hence can change the threshold voltage.

To investigate the Body Bias effect, we connect a voltage source (V_{BS}) between Body and Source of the MOSFET as shown in Fig. 8, such that it applies reverse bias to body-substrate junction. The transfer characteristics is observed at different values of V_{BS}. Potential drop between Source and Body changes the depletion layer below the gate and hence the threshold voltage. The threshold voltage of MOSFET is given by,

$$V_T = V_{T0} + \gamma(\sqrt{\phi_s + V_{BS}} - \sqrt{\phi_s}) \tag{8}$$

V_T is affected by V_{BS} according to Eq.8. The effect of body bias on transfer characteristics (I_D-V_{GS}) of the MOSFET is as shown in the Fig. 8.
It is clear that the characteristics for an PMOS transistor shift to left with increase in the body bias resulting in more negative threshold voltage.

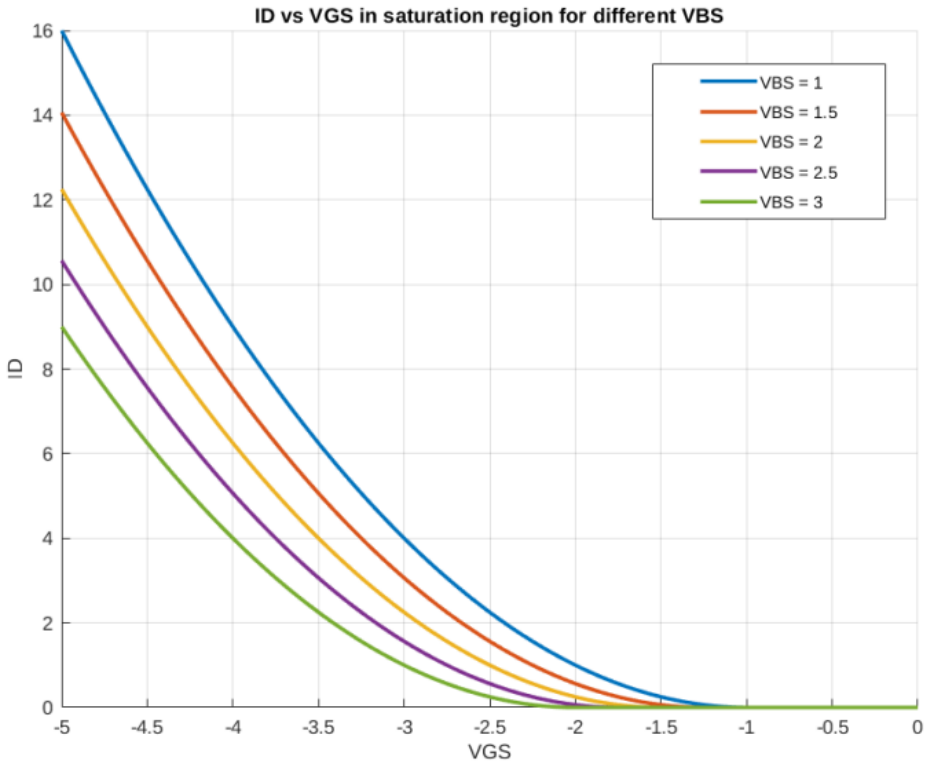

Figure 8: I_D *vs.* V_{GS} characteristics showing body bias effect.