

Physics of Semiconductor Devices

Module – III.3

Metal Oxide Semiconductor Field Effect Transistor



FET & MOSFET

- Transistor is a device that presents a high input resistance to the signal source, drawing little input power, and a low resistance to the output circuit, capable of supplying a large current to drive the circuit load.
- **Field-effect transistor or FET** :- The gate turns the transistor (inversion layer) on and off with an **electric field** through the oxide.
- FETs have three main parts ; i.e., Source, Channel & Drain. It come in different form;
 - ❖ JFET: Junction FET:- the gate voltage varies the depletion layer width of a reverse biased PN junction(channel)
 - ❖ MESFET: Metal-Semiconductor FET:- Channel is a Schottky barrier
 - ❖ MISFET: Metal-Insulator-Semiconductor FET;- Channel is Insulator; If it is an Oxide, it is called MOSFET

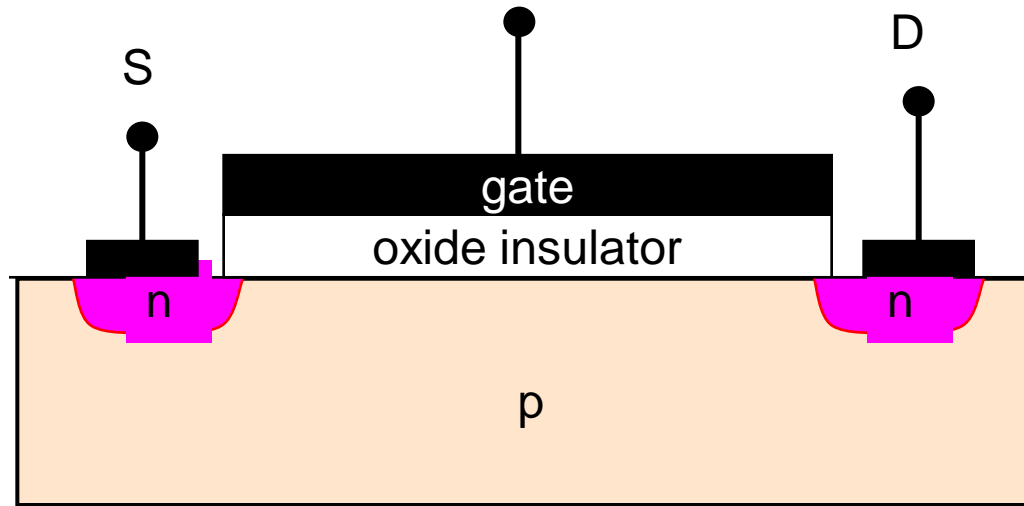
- Two important features of a PN junction are:
- (i) **Injection of minority carriers with forward bias**:- used in BJT (bipolar), controlled by current & has low input impedance.
- (ii) **Variation of depletion layer width with reverse bias**:- used in FET(unipolar), controlled by voltage & has high input resistance.
- Both are three terminal devices; Source, Channel & Drain
- FETs are suitable for controlled switching between a conducting state & non-conducting state, useful for digital circuits.
- FETs are useful for integration of many devices on a single chip; used in semiconductor memory devices & microprocessors.
- In n-channel; electrons flow from source-to-drain—opposite to the current flow
- In p-channel; holes flow from source-to-drain in the direction of current

MOSFET

- MOSFET Field effect transistor is a unipolar transistor, which acts as a voltage-controlled current device and is a device in which current at two electrodes; drain and source is controlled by the action of an electric field at another electrode gate.

Usefulness of MOSFET:-

- MOSFET is by far the most prevalent semiconductor device in ICs.
- It is the basic building block of digital, analog, and memory circuits.
- Its small size allows the making of inexpensive and dense circuits such as giga-bit (Gb) memory chips.
- Its low power and high speed make possible chips for gigahertz (GHz) computer processors and radio-frequency (RF) cellular phones.

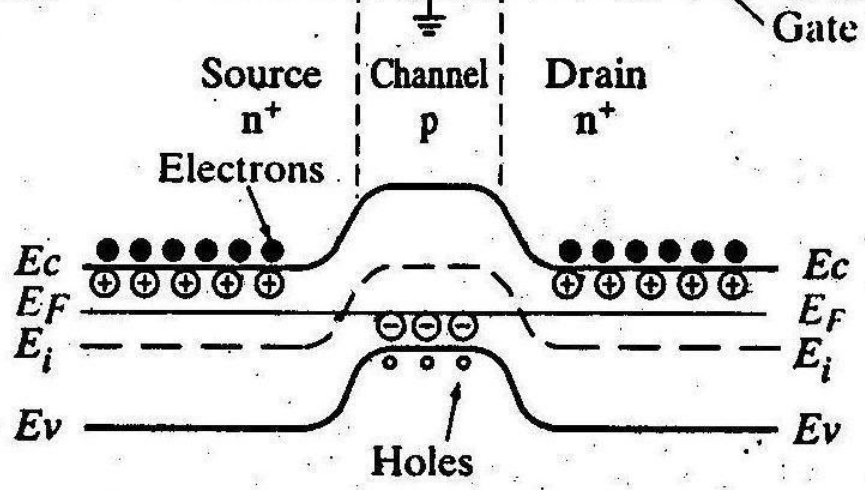
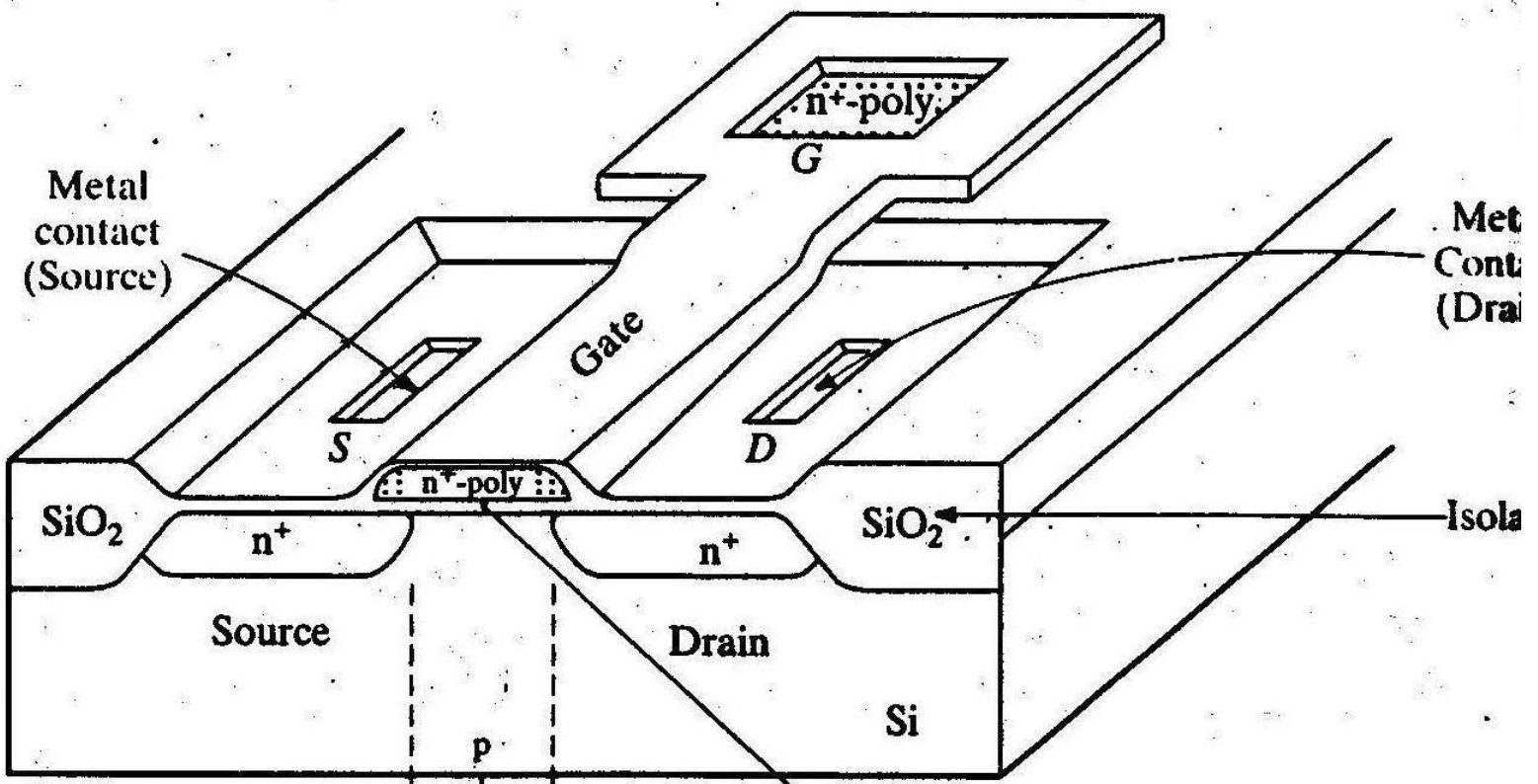


An electric field is applied normal to the surface of the semiconductor (by applying a voltage to an overlying electrode), to modulate the conductance of the semiconductor

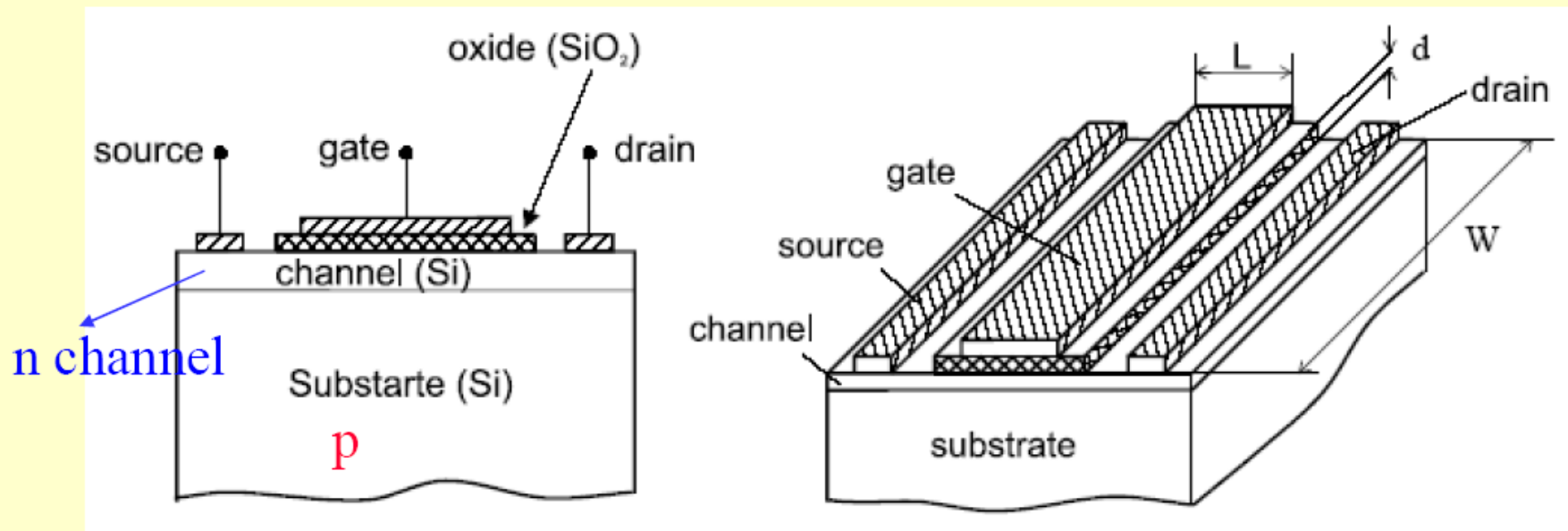
- Modulate drift current flowing between 2 contacts (“source” and “drain”) by varying the voltage on the “gate” electrode.
- Without a gate-to-source voltage (V_{GS}) applied, no current can flow between the source and drain regions.
- Above a certain gate-to-source voltage (***threshold voltage V_t***), a conducting layer of mobile electrons (n-channel) is formed at the Si surface beneath the oxide. These electrons flow between source and drain, resulting current from drain-to-source (I_D)

MOSFET Structure

- Two pn junctions placed back-to back
- The two n^+ end layers “Source” and “Drain” are heavily doped to approximately the same level.
- The p type middle layer is termed the body (or substrate) and has moderate doping level (2 to 3 orders of magnitude lower than n^+ regions on both sides).
- The n^- drain drift region has the lowest doping density. Thickness of this region determines the breakdown voltage of the device.
- The gate terminal is placed over the n^- and p type regions of the cell structure and is insulated from the semiconductor body by a thin layer of silicon dioxide (also called the gate oxide).

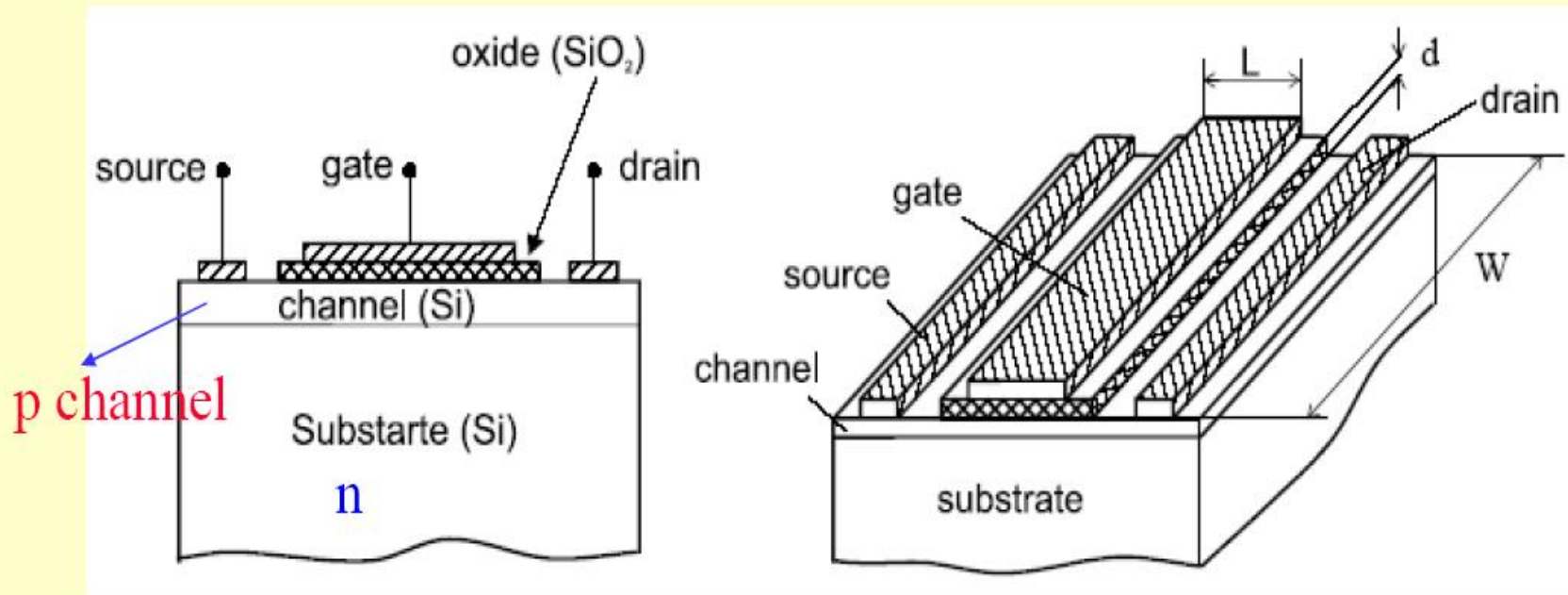


n-channel depletion Metal-Oxide-Semiconductor FET (MOSFET)



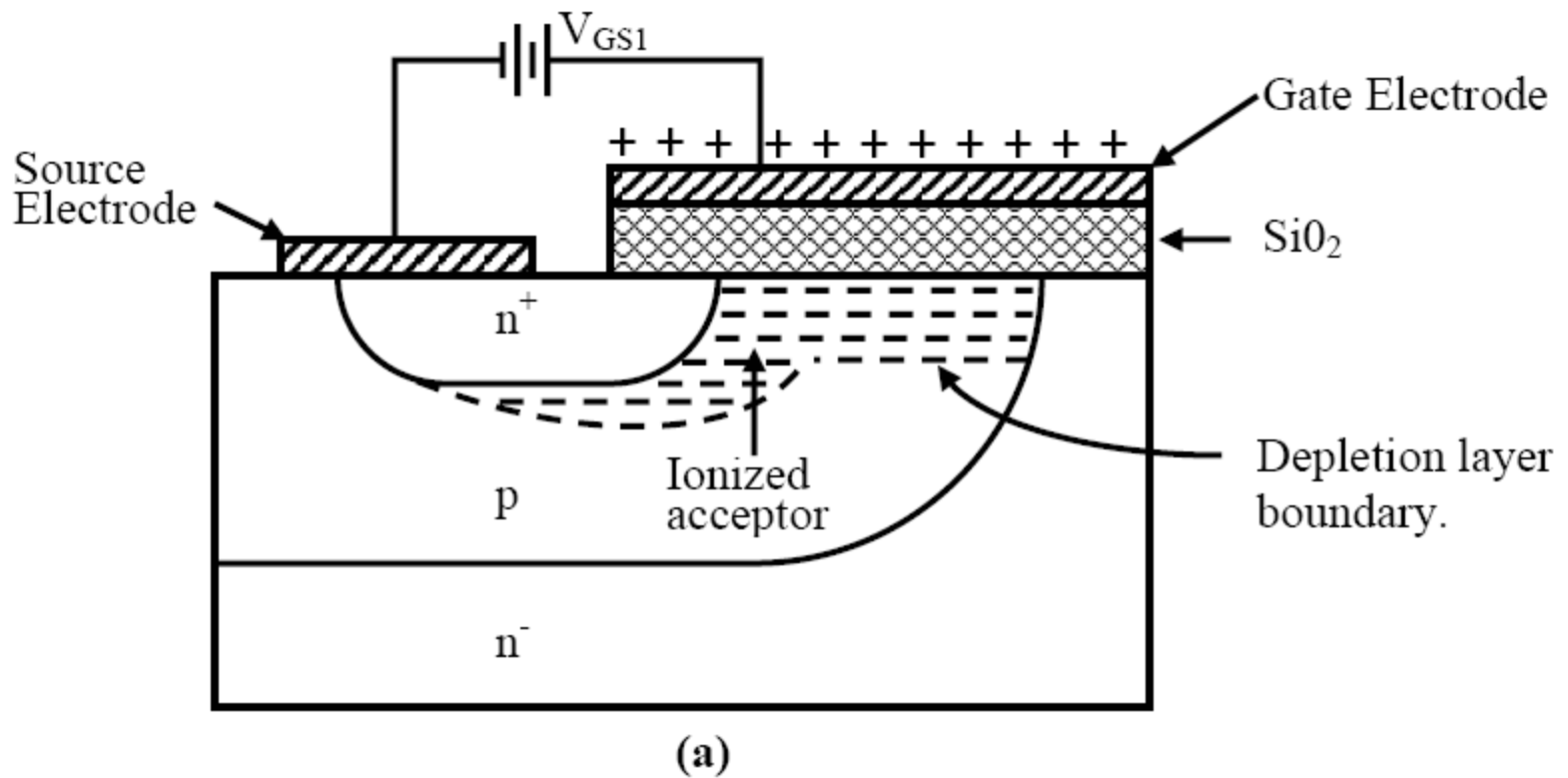
The gate-channel insulator is made out of dielectric (SiO₂), $\epsilon = 3.9$

p-channel depletion Metal-Oxide-Semiconductor FET (MOSFET)

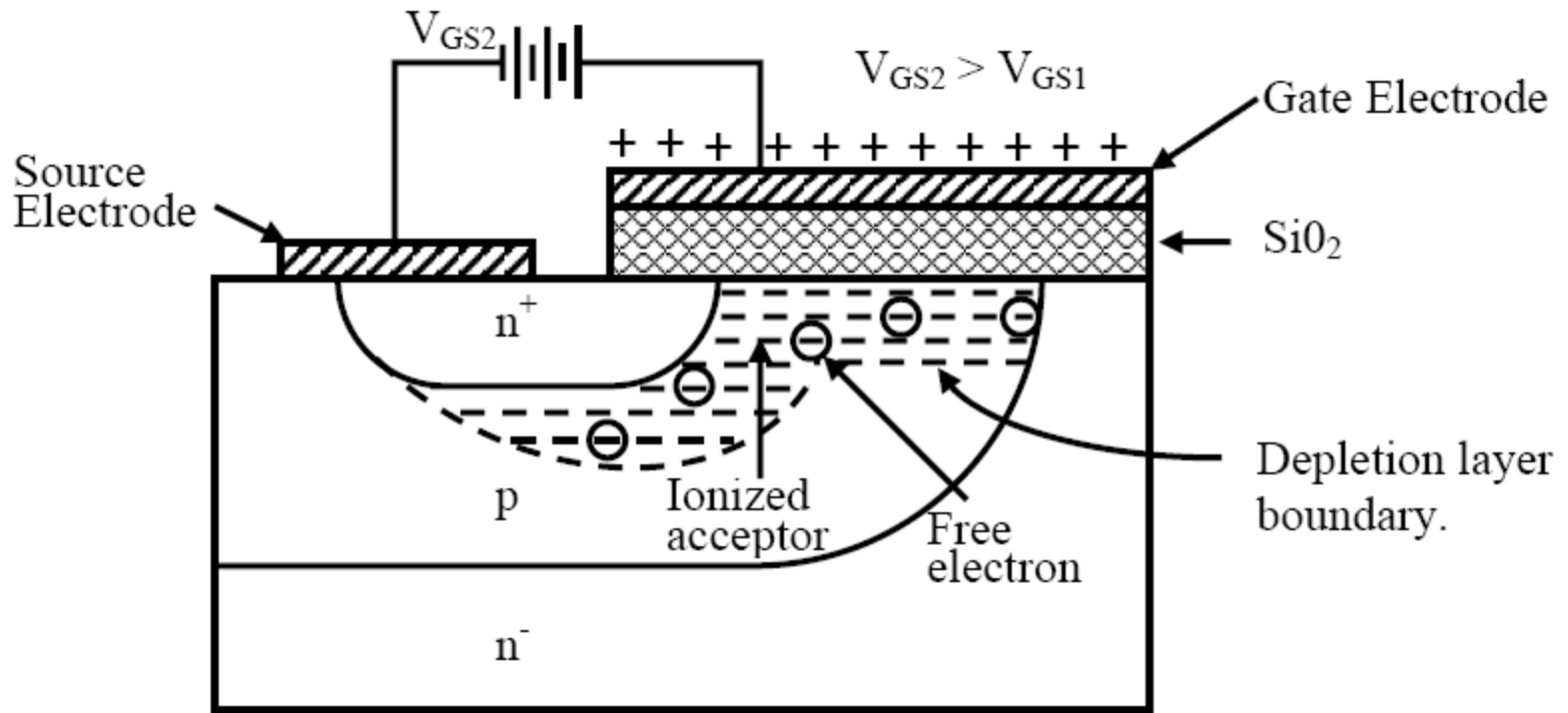


The gate-channel insulator is made out of dielectric (SiO₂), $\epsilon = 3.9$

A small voltage applied to the capacitor with gate terminal positive with respect to the source a depletion region forms at the interface between the oxide and substrate.

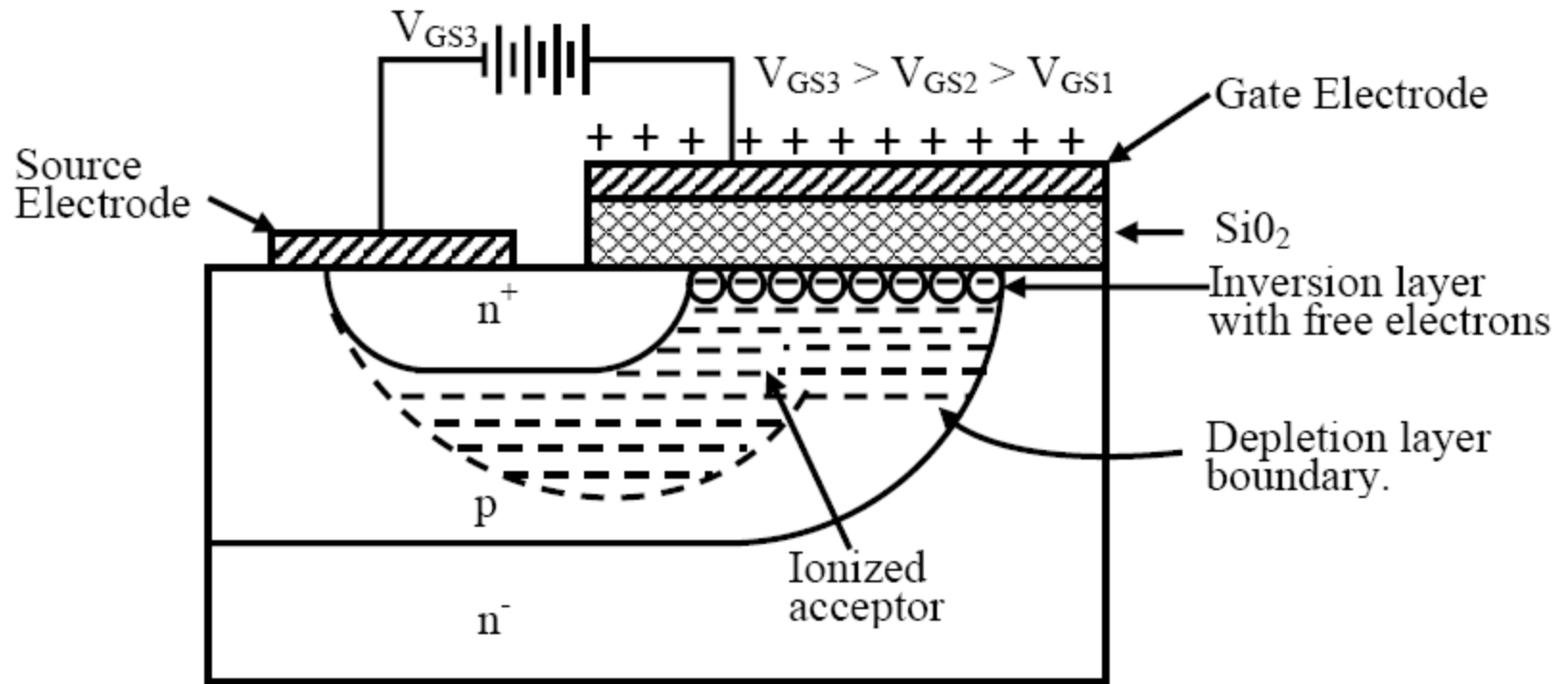


Further increase in V_{GS} causes the depletion layer to grow in thickness. At the same time the electric field at the oxide-silicon interface gets larger and begins to attract free electrons



(b)

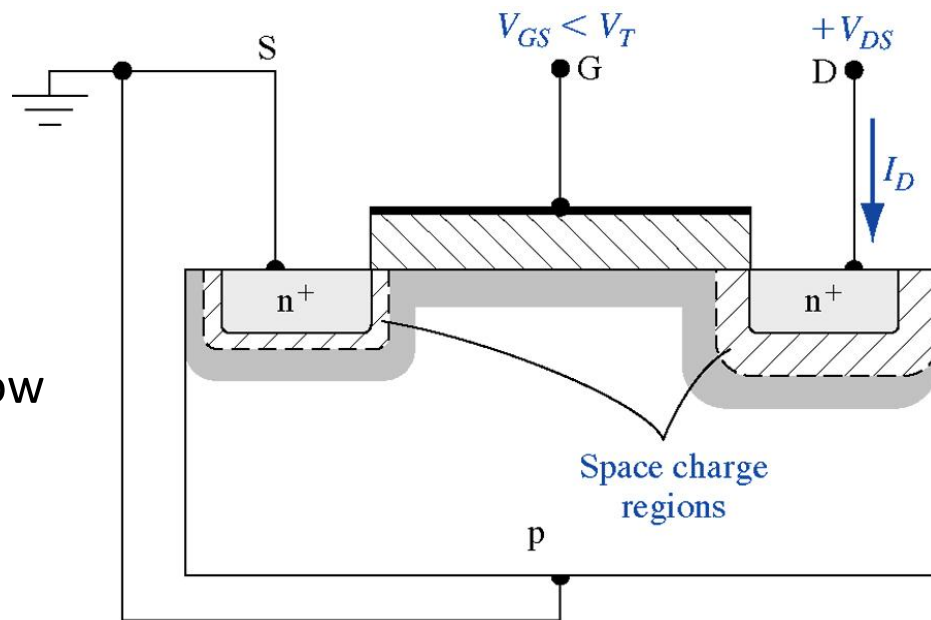
As V_{GS} increases further the density of free electrons at the interface becomes equal to the free hole density in the bulk of the body region beyond the depletion layer. The layer of free electrons at the interface is called the inversion layer :- is a conductive path or “channel” between the drain and the source which permits flow of current between the drain and the source.



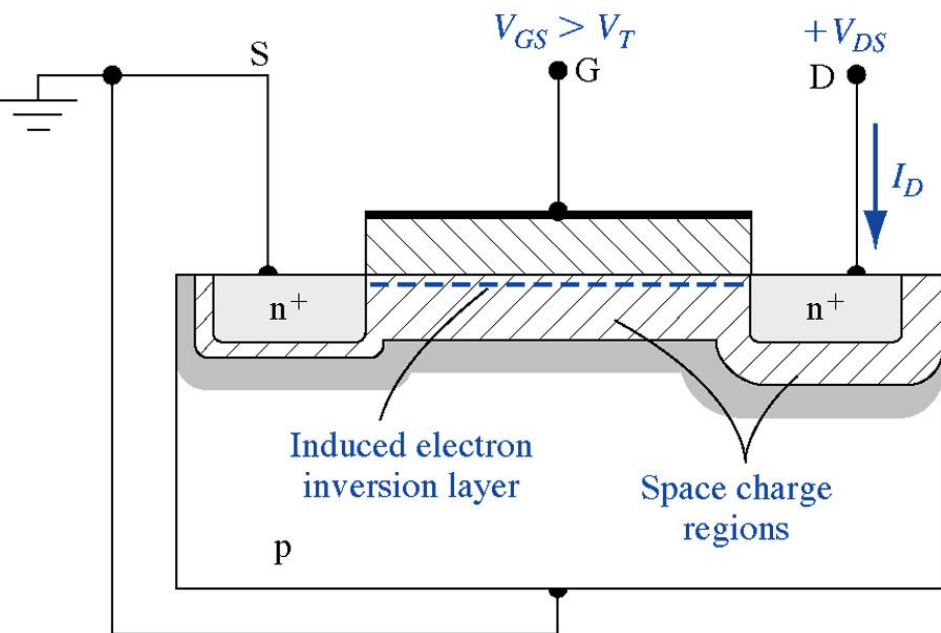
(c)

MOSFET operation

- Transistor is ON when $V_{GS} = 0V$
- $V_{DS} > 0$, electrons flow from source to drain under EF.
- $V_{GS} > V_t$ -- Inversion mode--current flow
- Current is drift current.
- Channel behaves as resistance.
- I_D varies linearly with V_{DS}



(a)

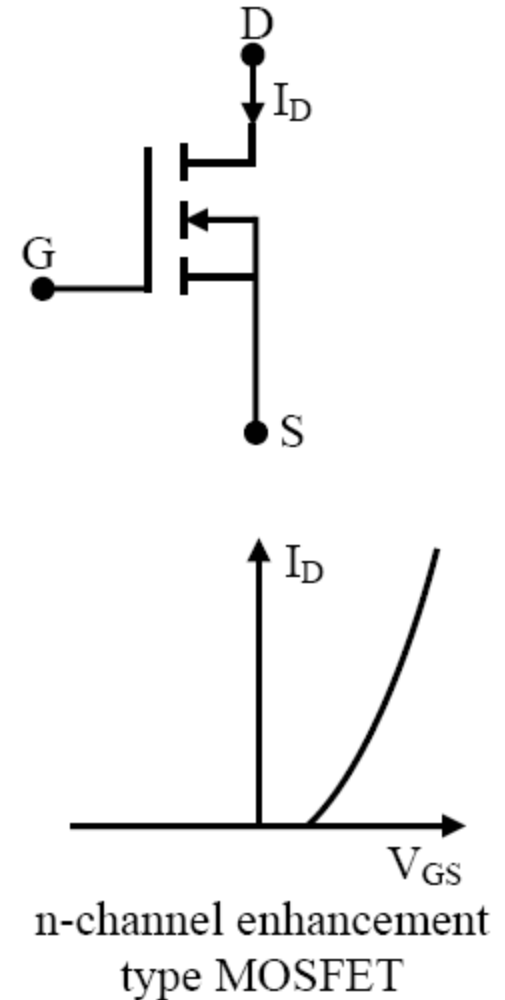
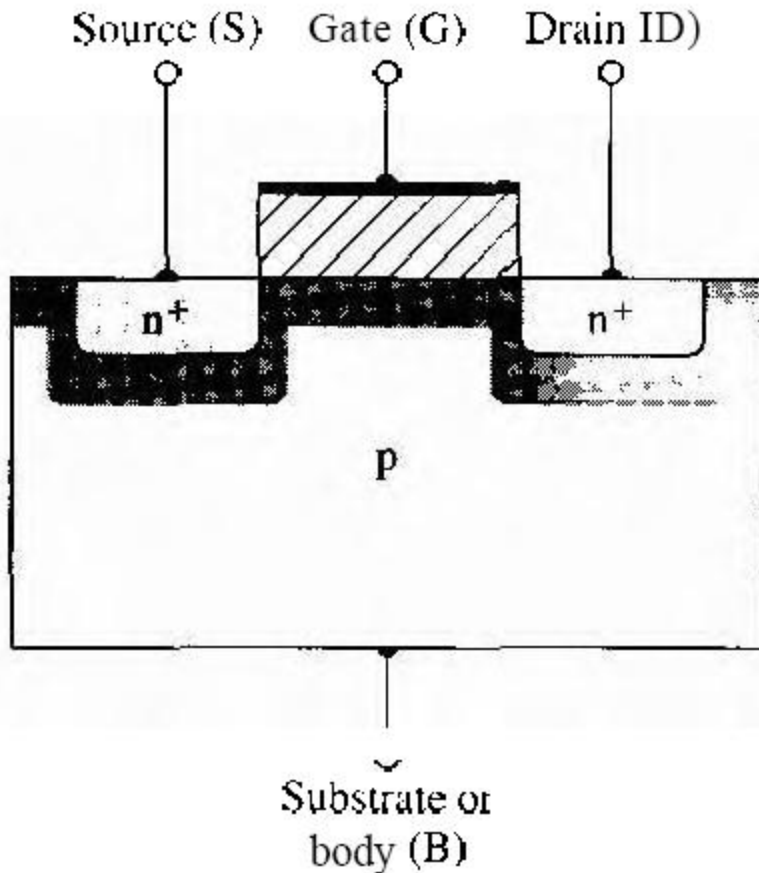


(b)

- V_{GS} fixed, V_{DS} increases, Voltage across oxide & electron conc. reduces from source-to-drain.
- $V_{DS} = V_{DSAT}$ — Pinch-off — no Inversion layer at drain end.
- $V_{DS} < V_{DSAT}$ — linear region.
- $V_{DS} > V_{DSAT}$ — saturation region (I_D is const.)

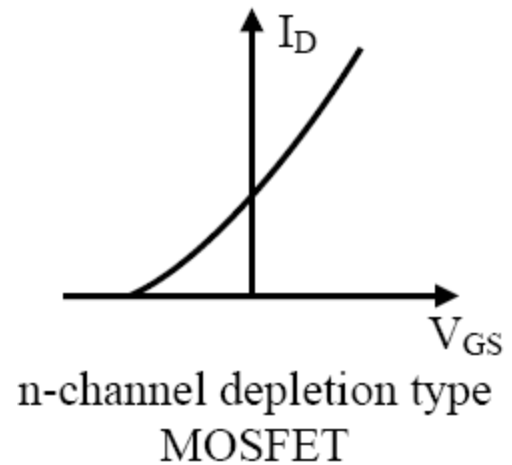
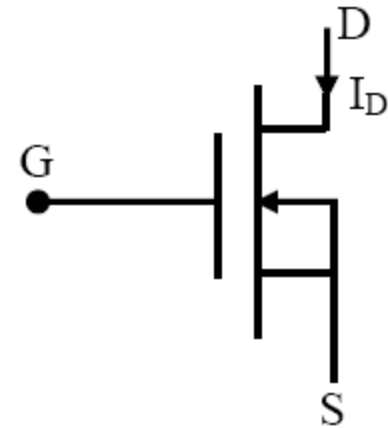
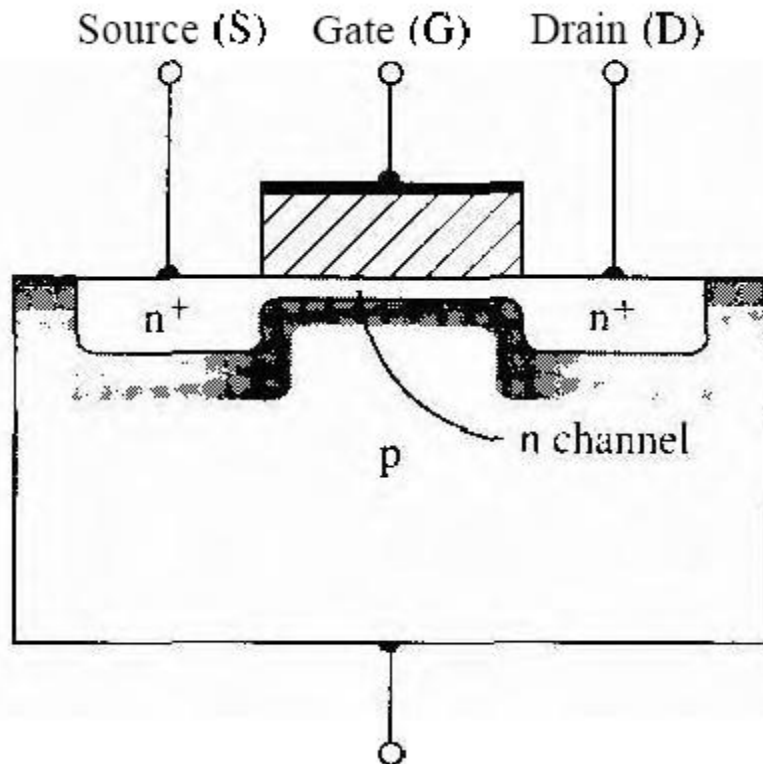
Basic MOSFETs

Positive gate voltage induces the n-channel, electrons flow from the source to the drain; conventional current will enter the drain and leave the source.

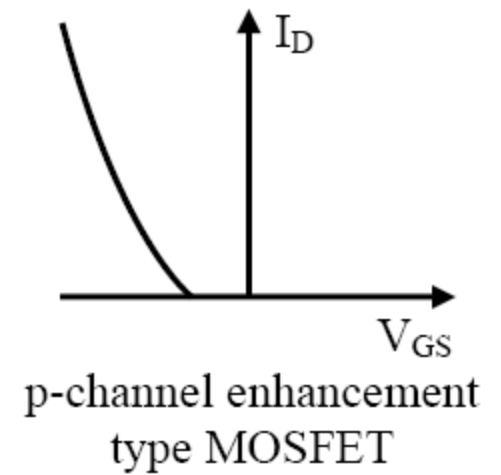
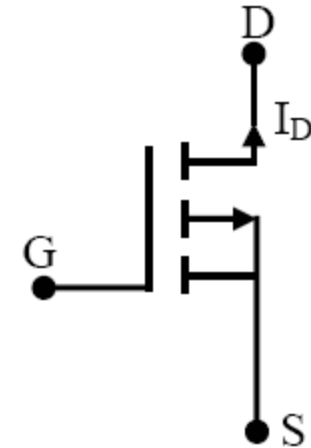
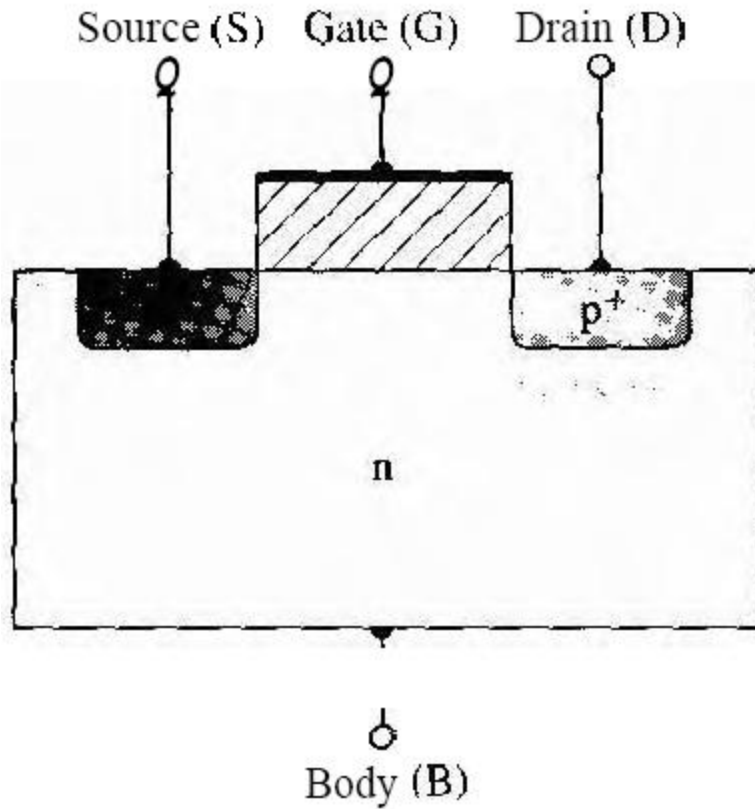


If the semiconductor is lightly doped and ϕ_{ms} is sufficient to induce an inversion layer even with zero applied gate voltage. Also V_{th} may be negative for a p-substrate.

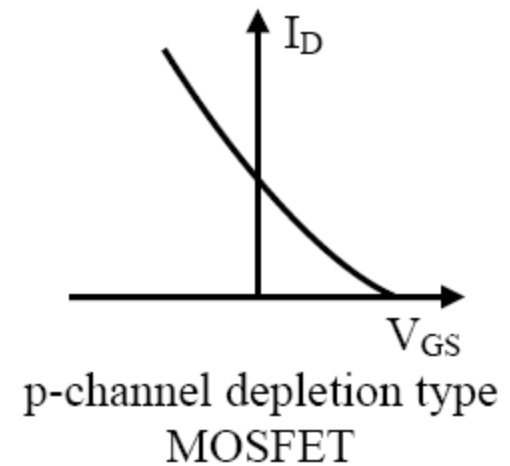
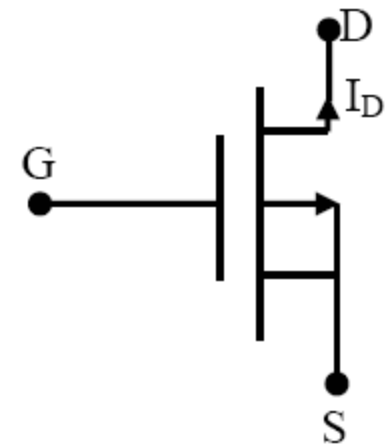
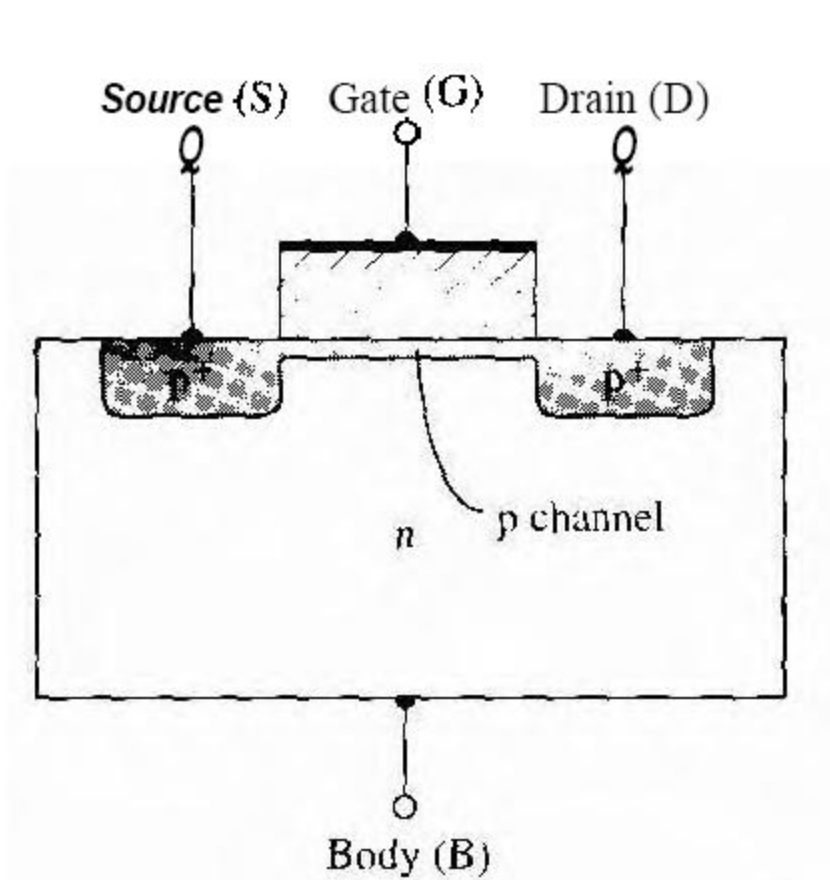
So, an n-channel region exists under the oxide with zero volts applied to the gate



p-channel enhancement mode MOSFET



p-channel depletion mode MOSFET



Surface Mobility

To have large transistor current in the MOSFET is ensured by electron or hole mobility in the surface inversion layer.

Mobility of carriers near surface of silicon (μ_{ns} or μ_{ps}) (i.e., inside the inversion layer) can be different from the mobility inside bulk silicon (μ_n or μ_p).

When a small V_{ds} is applied, the drain to source current, I_{ds} ,

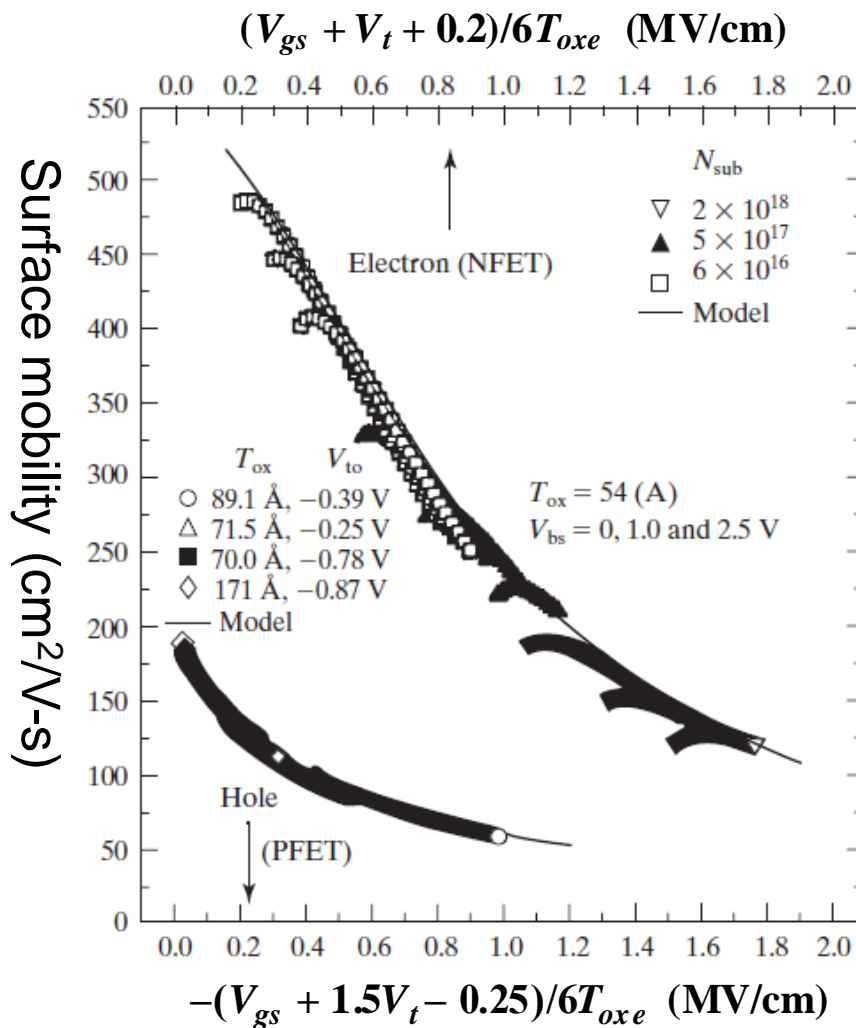
$$\begin{aligned} I_{ds} &= W \cdot Q_{inv} \cdot v = W Q_{inv} \mu_{ns} \mathcal{E} = W Q_{inv} \mu_{ns} V_{ds} / L \\ &= W C_{oxe} (V_{gs} - V_t) \mu_{ns} V_{ds} / L \end{aligned}$$

W is the **channel width**, the vertical dimension of the channel in.

Q_{inv} (C/cm^2) is the inversion charge density.

L is the **channel length**.

μ_{ns} is the **electron surface mobility, or the effective mobility**, smaller than the bulk mobilities



μ_{ns} follow the $T^{-3/2}$ temperature dependence characteristic of phonon scattering.

The surface mobility
At $V_g \approx V_t$, especially in the heavily doped semiconductor, is lower than the universal mobility due to dopant ion scattering.

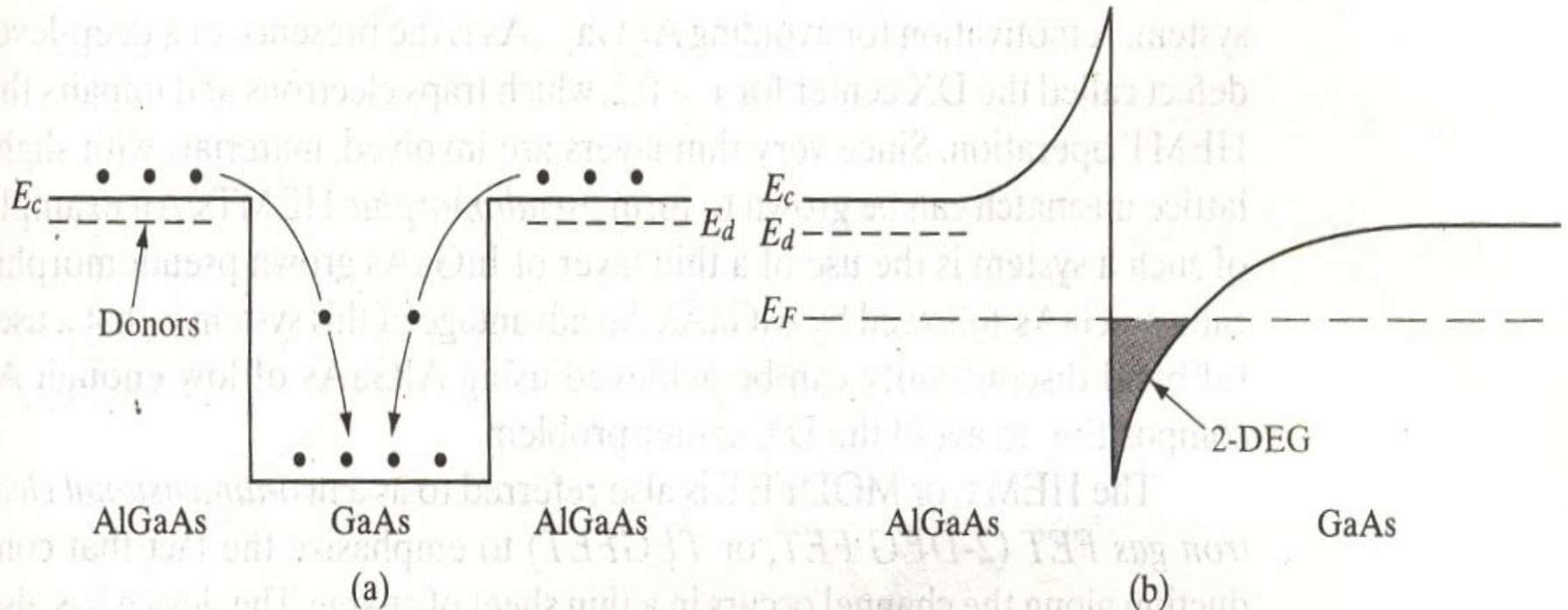
At higher V_g , dopant ion scattering effect is screened out by the inversion layer carriers

Universal Mobility Degradation Curve

HEMT:- High Electron Mobility Transistor

- Higher carrier mobility allows the carriers to travel faster and the transistors to operate at higher speeds.
- High-speed devices improve the throughput of electronic equipment & used for inexpensive microwave communication applications.
- The dopants in the channel may increase the conductivity, but significantly reduce the electron mobility through impurity scattering. If the channel is undoped, the mobility can be much higher.
- GaAs is used in (**MESFET**).
- To increase the carrier conc. for high conductance, but simultaneously improve the mobility by avoiding impurity scattering by doping.
- MOS-like structure is made by growing a thin epitaxial layer of GaAlAs over the undoped GaAs substrate.

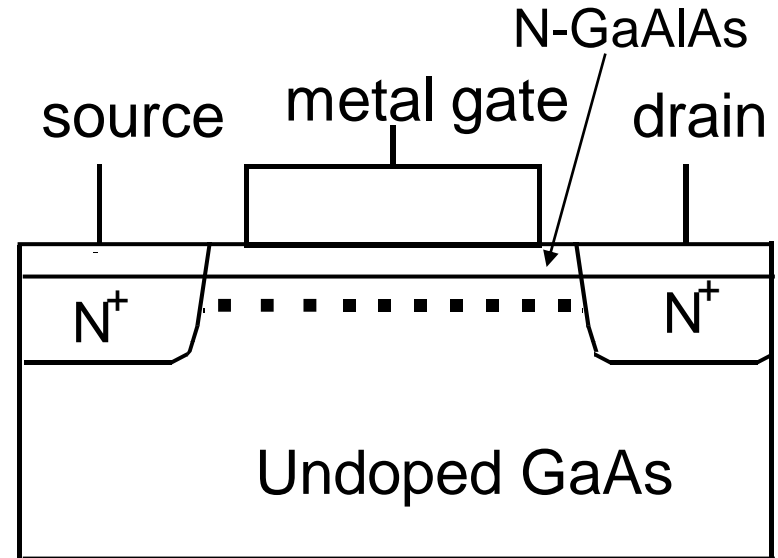
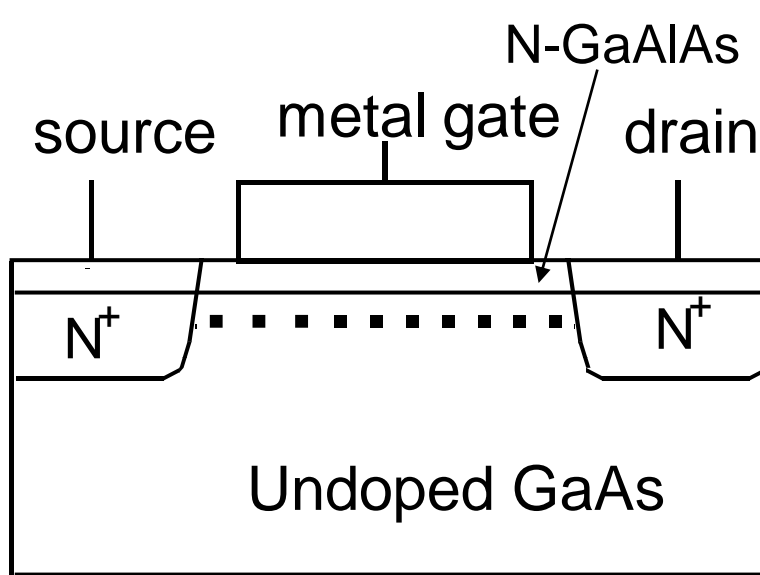
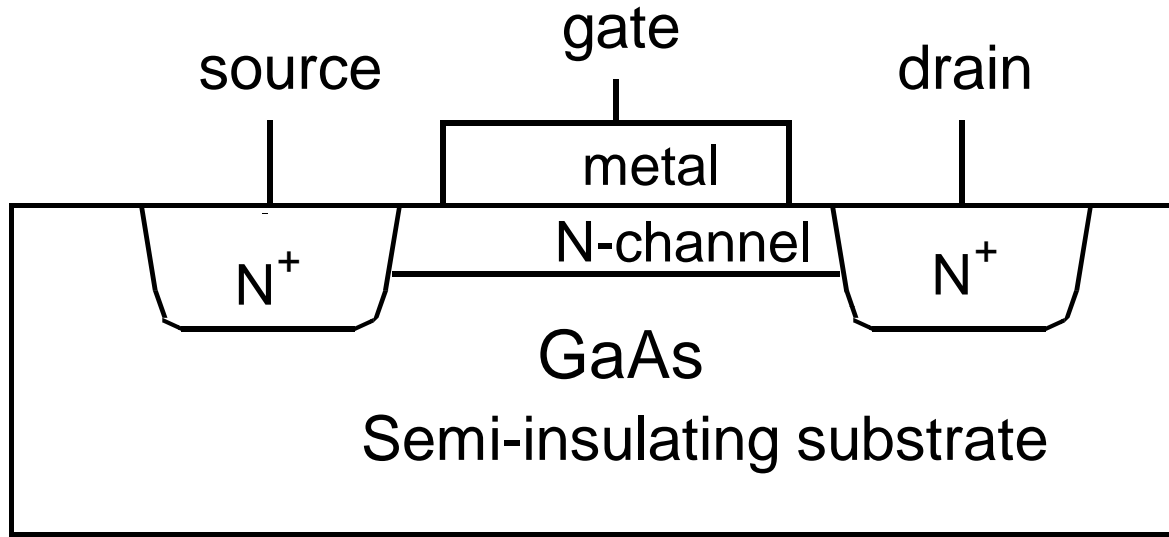
- GaAs is bounded by GaAlAs, which has a larger band gap than GaAs.
- It creates an energy well and a thin layer of electrons at the GaAs–GaAlAs interface. The curvature in the GaAlAs band diagram is due to the presence of the dopant ions as in the depletion layer of a PN junction.
- Electron from the donors in GaAlAs falls into the well & are trapped there. Carriers in this device does not suffer from mobility degradation by impurity scattering in undoped GaAs.
- The trapped electrons is called 2D-electron-gas, the equivalent of the inversion or accumulation layer, when connected with channel in series provides high carrier conc. with high mobility.
- This configuration is called modulation doping.



(a) Simplified view of modulation doping, showing only the conduction band. Electrons in the donor-doped AlGaAs fall into the GaAs potential well and become trapped. As a result, the undoped GaAs becomes n-type, without the scattering by ionized donors which is typical of bulk n-type material. (b) Use of a single AlGaAs/GaAs heterojunction to trap electrons in the undoped GaAs. The thin sheet of charge due to free electrons at the interface forms a two-dimensional electron gas (2-DEG), which can be exploited in HEMT devices.

- This device is called **HEMT or high electron-mobility transistor or modulation doped field effect transistor (MODFET)**
- It is used in microwave communication, satellite TV receivers, etc.

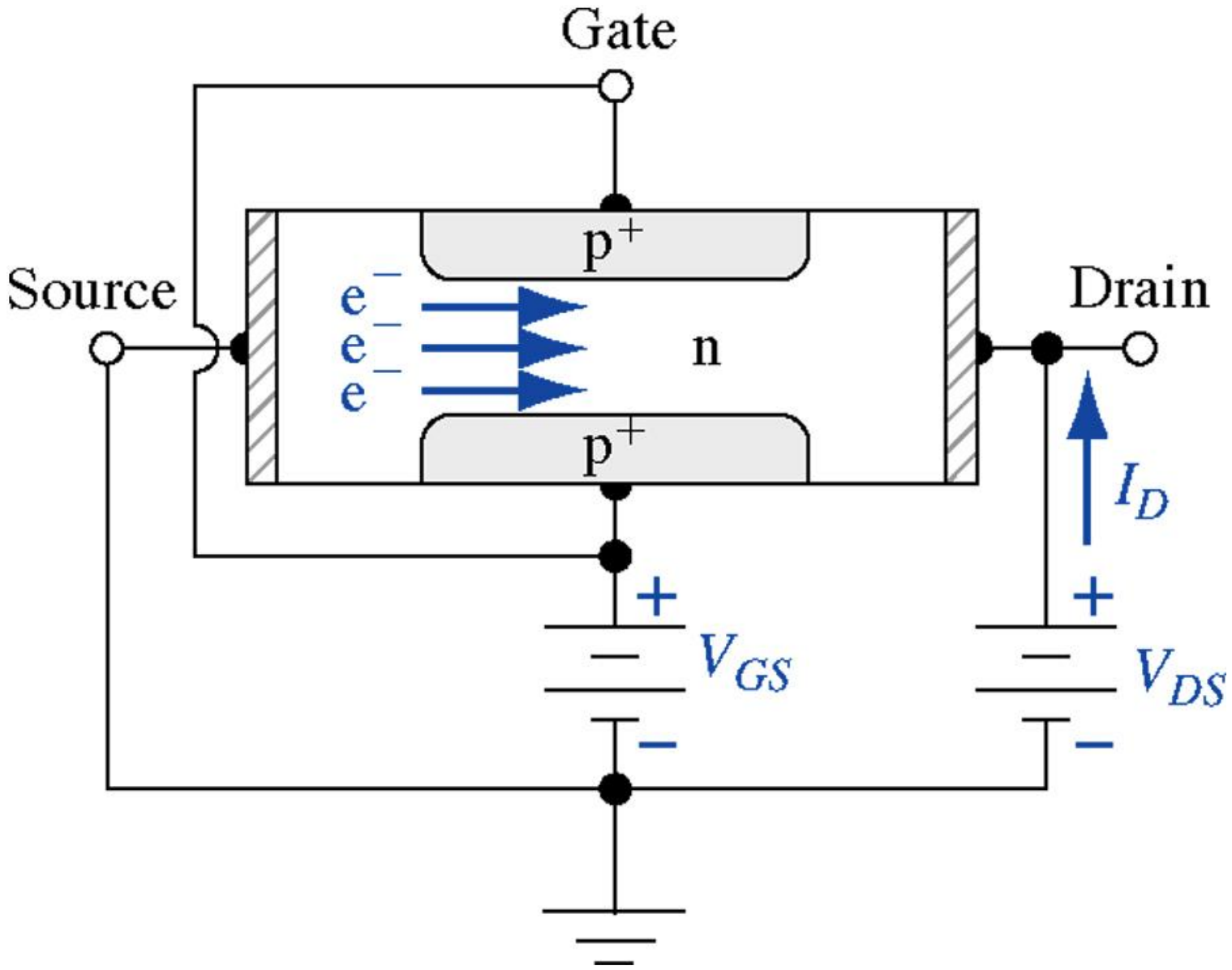
HEMT, High Electron Mobility Transistor



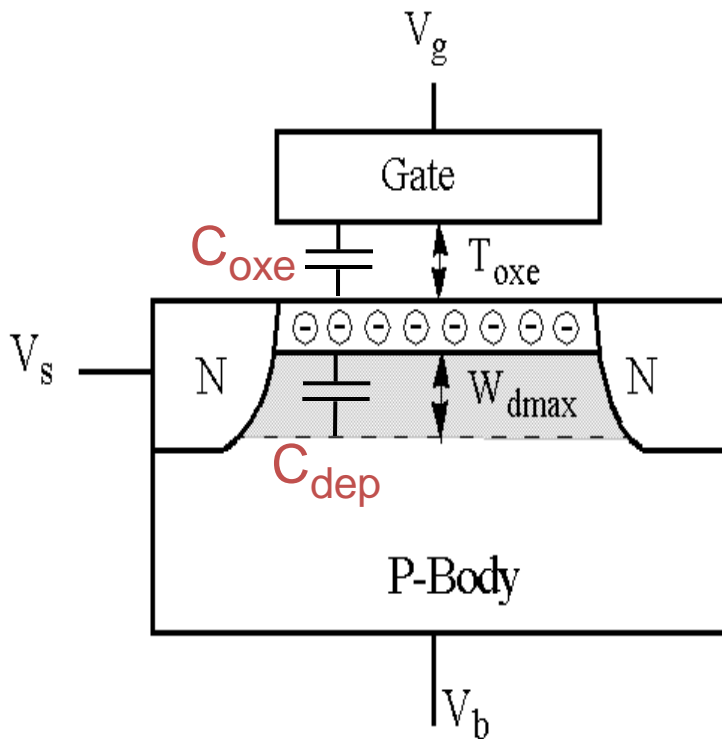
JFET

- If the Schottky junction is replaced with a P⁺N junction, the new structure is called a **JFET or junction field-effect transistor**.
- As in a MESFET, a reverse bias would expand the depletion layer and constrict the conduction channel.
- JFET current is controlled with the gate voltage.
- JFET provided a low input current and capacitance device because its input is a reverse-biased diode.
- JFET can be fabricated with bipolar transistors and coexist in the same IC chip.

JFET



MOSFET V_t and Body Effect



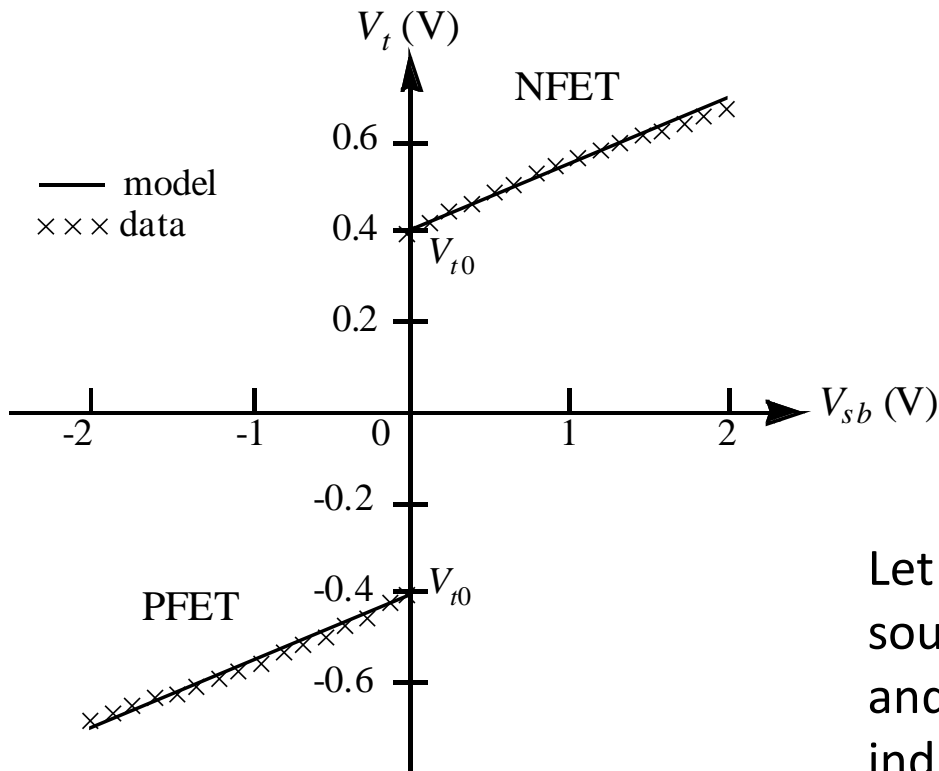
The inversion layer of a MOSFET can be thought of as a resistive N-type film (1–2 nm thin) that connects the source and the drain.

This film, at potential V_s , forms a capacitor with the gate, the oxide being the capacitor dielectric.

It also forms a second capacitor with the body and the capacitor dielectric is the depletion layer.

At strong inversion:- $V_G > V_T$

$$V_G = V_{FB} + \phi_s + V_{ox} = V_{FB} + 2\phi_{fp} + [-Q_{dep} - Q_{inv}] / C_{ox}$$



V_T is an approximately linear function of the body to source bias voltage.

The polarity of the body bias is normally that which would reverse bias the body-source junction.

$$V_G = V_T - Q_{inv} / C_{ox}$$

$$Q_{inv} = -C_{oxe}(V_{gs} - V_t)$$

Let there is also a voltage between the source and the body, V_{sb} . Since the body and the channel are coupled by C_{dep} , V_{sb} induces a charge in the inversion layer, $C_{dep}V_{sb}$.

Therefore,

$$\begin{aligned} Q_{inv} &= -C_{oxe}(V_{gs} - V_t) + C_{dep}V_{sb} \\ &= -C_{oxe}\left(V_{gs} - \left(V_t + \frac{C_{dep}}{C_{oxe}}V_{sb}\right)\right) \end{aligned}$$

$$Q_{inv} = -C_{oxe}(V_{gs} - V_t(V_{sb}))$$

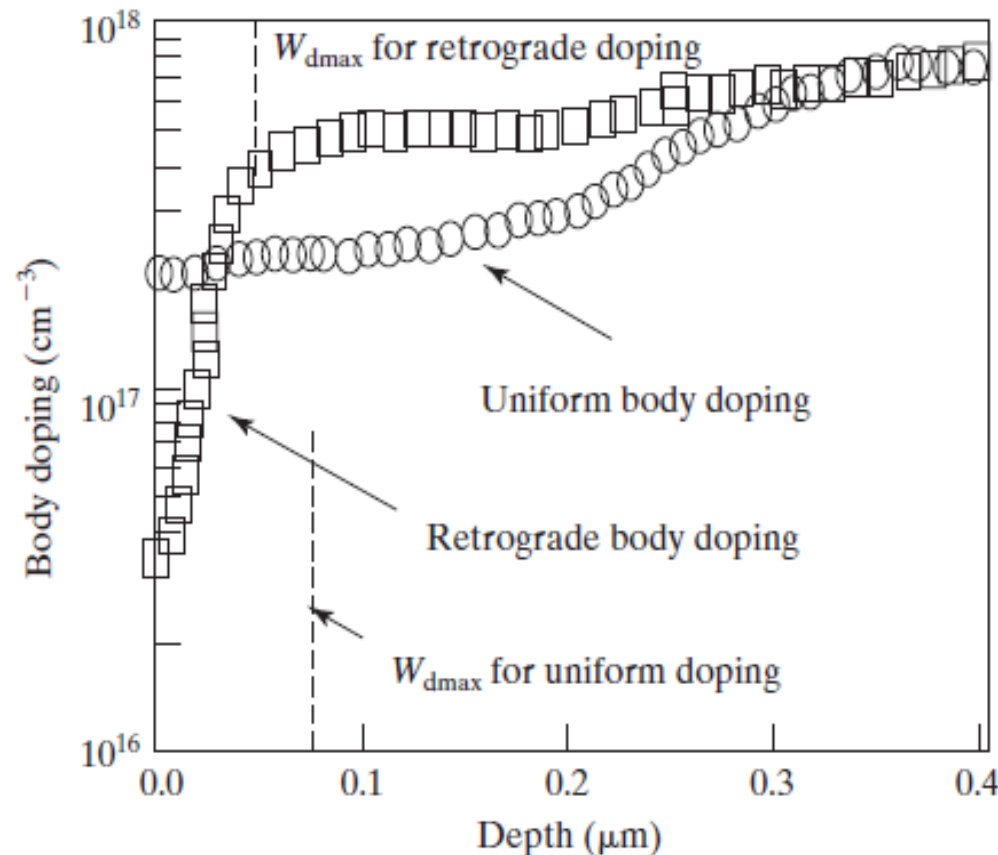
$$V_t(V_{sb}) = V_{t0} + \frac{C_{dep}}{C_{oxe}} V_{sb} = V_{t0} + \alpha V_{sb}$$

So, V_T is a function of V_{sb} . When the source-body junction is reverse-biased, the NFET V_T becomes more positive and the PFET V_T becomes more negative.

V_T is a function of the body bias is called the **body effect**.

- When the source-body junction is reverse-biased, V_T increases.
- Body effect coefficient: $\alpha = C_{dep}/C_{ox} = 3T_{ox} / W_{dep}$

- Some transistors' source–body junctions are reversed biased.
- So it raises their V_T and reduces I_{ds} and the circuit speed.
- Circuits therefore perform best when V_T is as insensitive to V_{sb} as possible, i.e., the body effect should be minimized.
- This can be accomplished by minimizing the T_{ox}/X_{dT} ratio.
- “ α ” can be extracted from the slope of the curve



$$I_{ds} = W \cdot Q_{inv} \cdot v = W Q_{inv} \mu_{ns} \mathcal{E} = W Q_{inv} \mu_{ns} V_{ds} / L$$

$$= W C_{oxe} (V_{gs} - V_t) \mu_{ns} V_{ds} / L$$

RETROGRADE DOPING

- Modern transistors employ steep **retrograde body doping profiles** (**light** doping in a thin surface layer and very heavy doping underneath).
- Steep retrograde doping allows transistor shrinking to smaller sizes for cost reduction and reduces impurity scattering.
- **The depletion-layer thickness is basically the thickness of the lightly doped region.**
- As V_{sb} increases, the depletion layer does not change significantly.
- Therefore C_{dep} and α are basically constants. As a result, modern transistors **exhibit a more or less linear relationship between V_T and V_{sb} .**
- A linear relationship means that X_{dT} and therefore the C_{dep}/C_{ox} ratio are independent of the body bias.

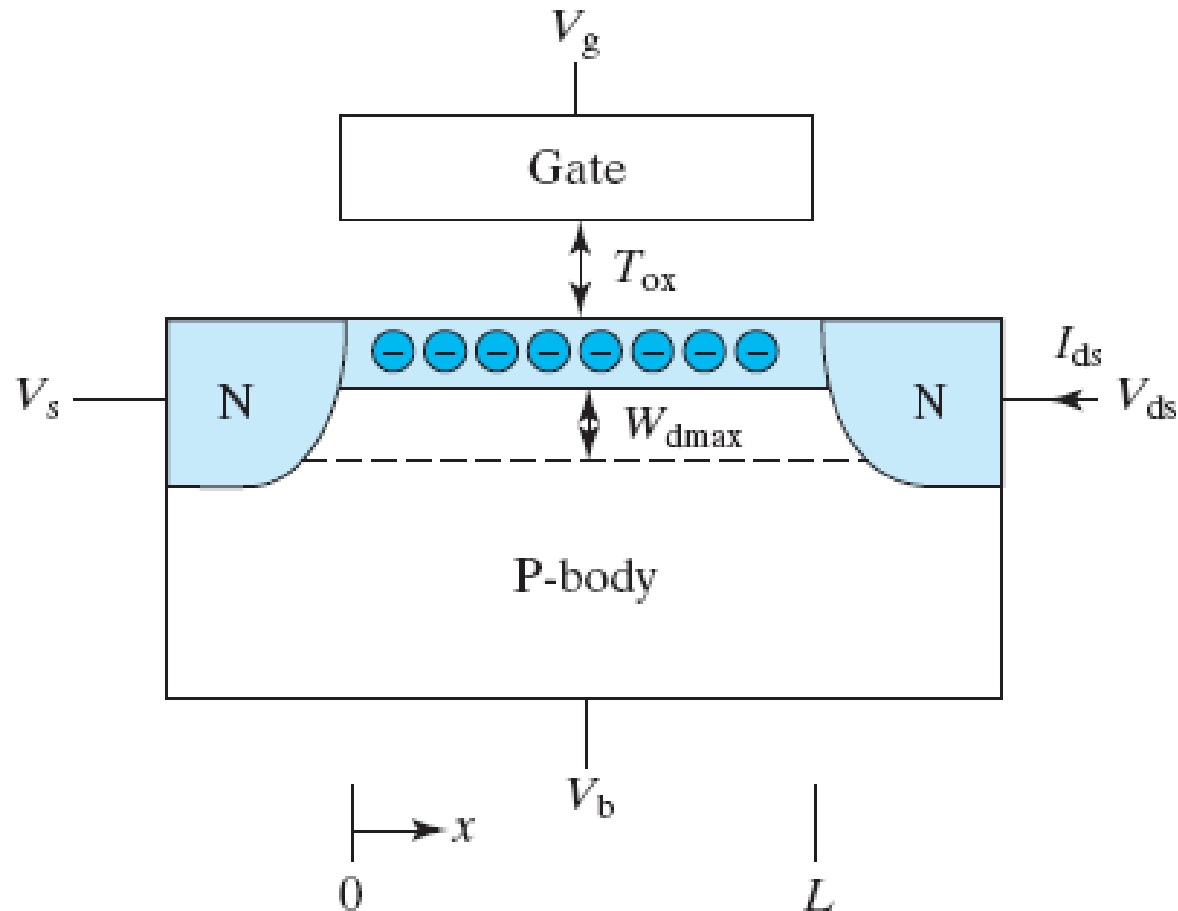
When the source/body junction is reverse-biased, there are two quasi-Fermi levels (E_{fn} and E_{fp}) which are separated by eV_{sb} . An NMOSFET reaches threshold of inversion when E_c is close to E_{fn} , not E_{fp} . This requires the band-bending to be $2\phi_{fp} + V_{sb}$,

V_T can be obtained by replacing the $2\phi_{fp}$ term (band bending in the body) with $2\phi_{fp} + V_{sb}$.

$$\begin{aligned} V_t &= V_{t0} + \frac{\sqrt{qN_a 2\epsilon_s}}{C_{oxe}} (\sqrt{2\phi_B + V_{sb}} - \sqrt{2\phi_B}) \\ &\equiv V_{t0} + \gamma (\sqrt{2\phi_B + V_{sb}} - \sqrt{2\phi_B}) \end{aligned}$$

' γ ' is the body-effect parameter.

Q_{INV} IN MOSFET



When $V_{ds} \neq 0$, the channel voltage V_c is a function of x .

The channel voltage, V_c , is a function of x . $V_c = V_s$ at $x = 0$ and $V_c = V_d$ at $x = L$.

voltage in the middle of the channel is higher than at source end, so there is less voltage across the oxide capacitor and across the depletion layer capacitor. Therefore, there will be fewer electrons on the capacitor electrode (the inversion layer).

So, V_{gs} term should be replaced by $V_{gs} - V_{cs}(x)$ and V_{sb} by $V_{sb} + V_{cs}(x)$.

- $$Q_{inv}(x) = -C_{ox}(V_{gs} - V_{cs} - V_{t0} - \alpha(V_{sb} + V_{cs}))$$

$$= -C_{ox}(V_{gs} - V_{cs} - (V_{t0} + aV_{sb}) - aV_{cs})$$

$$= -C_{ox}(V_{gs} - mV_{cs} - V_t)$$

- $$m = 1 + \alpha = 1 + 3T_{ox}/X_{dT}$$

“m” is called the body-effect factor or bulk-charge factor

MOSFET IV characteristics

$$Q_{\text{inv}} = -C_{\text{ox}}(V_{\text{Gs}} - mV_{\text{cs}} - V_{\text{T}}) \text{ -----[1]}$$

$$m = 1 + \alpha = 1 + 3T_{\text{oxe}}/X_{\text{dT}}$$

$$\begin{aligned} I_{\text{ds}} &= W \cdot Q_{\text{inv}}(x) \cdot v = W \cdot Q_{\text{inv}} \mu_{\text{ns}} \mathcal{E} \\ &= WC_{\text{oxe}}(V_{\text{gs}} - mV_{\text{cs}} - V_{\text{t}}) \mu_{\text{ns}} dV_{\text{cs}}/dx \end{aligned}$$

$$\int_0^L I_{\text{ds}} dx = WC_{\text{oxe}} \mu_{\text{ns}} \int_0^{V_{\text{ds}}} (V_{\text{gs}} - mV_{\text{cs}} - V_{\text{t}}) dV_{\text{cs}}$$

$$I_{\text{ds}} L = WC_{\text{oxe}} \mu_{\text{ns}} \left(V_{\text{gs}} - V_{\text{t}} - \frac{m}{2} V_{\text{ds}} \right) V_{\text{ds}}$$

$$I_{\text{ds}} = WC_{\text{oxe}} \mu_{\text{s}} \left(V_{\text{gs}} - V_{\text{t}} - \frac{m}{2} V_{\text{ds}} \right) \frac{V_{\text{ds}}}{L}$$

Drain-to-source current depends on channel width, E.F. in the channel and Average inversion layer charge density.

When V_{ds} is very small, the $mV_{ds}/2$ term is negligible

$I_{ds} \propto V_{ds}$, i.e., the **transistor behaves as a resistor**.

As V_{ds} increases, the average $(Q_{inv})_{av}$ **decreases** and dI_{ds}/dV_{ds} decreases.

at a certain V_{ds} , called V_{dsat} , dI_{ds}/dV_{ds} becomes zero.

$$\frac{dI_{ds}}{dV_{ds}} = 0 = \frac{W}{L} C_{ox} \mu_{ns} (V_{gs} - V_t - mV_{ds}) \quad \text{at} \quad V_{ds} = V_{dsat}$$

$$V_{dsat} = \frac{V_{gs} - V_t}{m}$$

V_{dsat} is called the **drain saturation voltage**, beyond which the drain current is saturated. For each V_G , there is a different V_{dsat} .

The part of the I-V curves with $V_{ds} \ll V_{dsat}$ is the linear region (Ohmic region),
The part with $V_{ds} > V_{dsat}$ is the saturation region (active region).

substituting V_{dsat} for V_{ds}

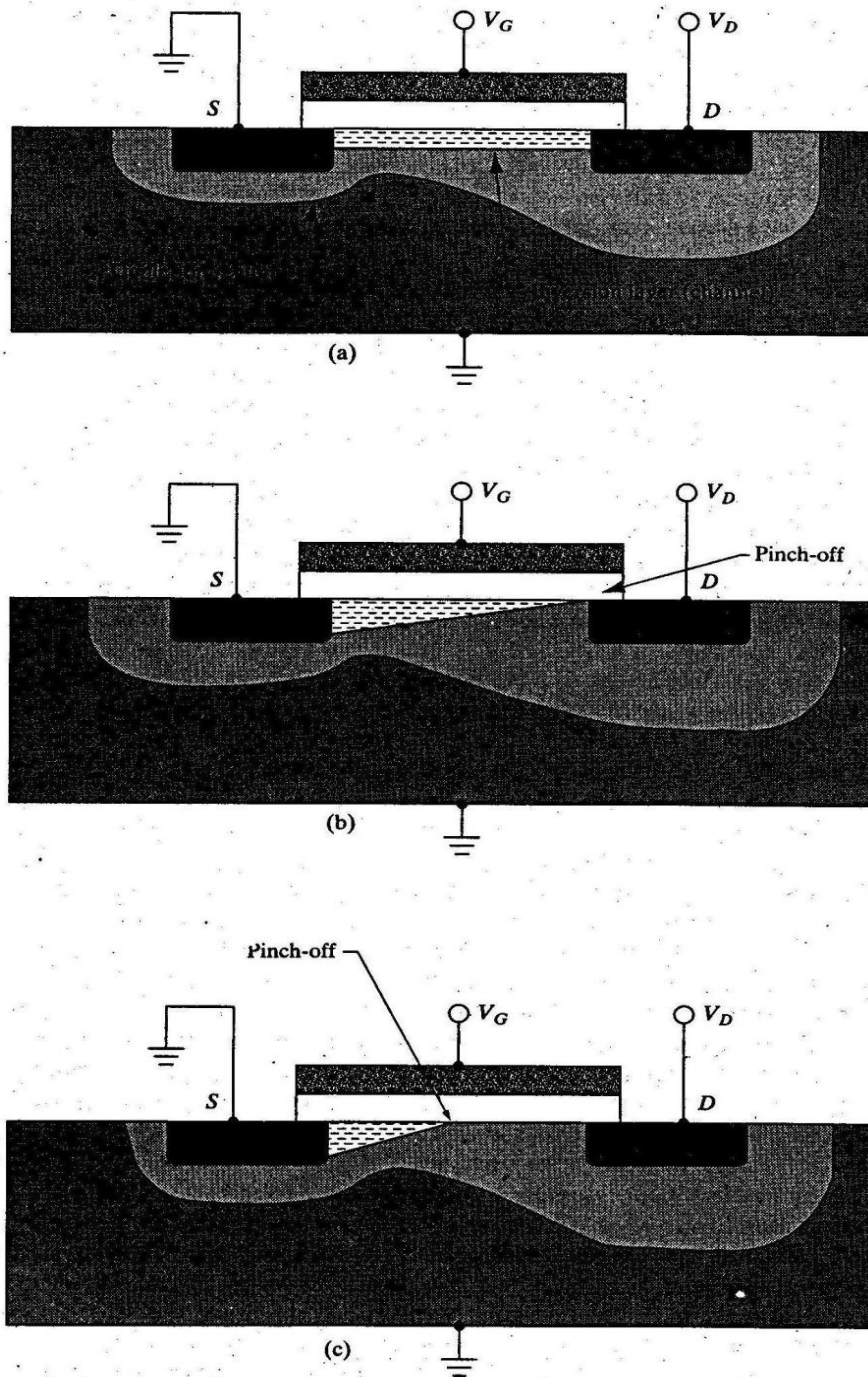
$$I_{dsat} = \frac{W}{2 \text{ mL}} C_{oxe} \mu_{ns} (V_{gs} - V_t)^2$$

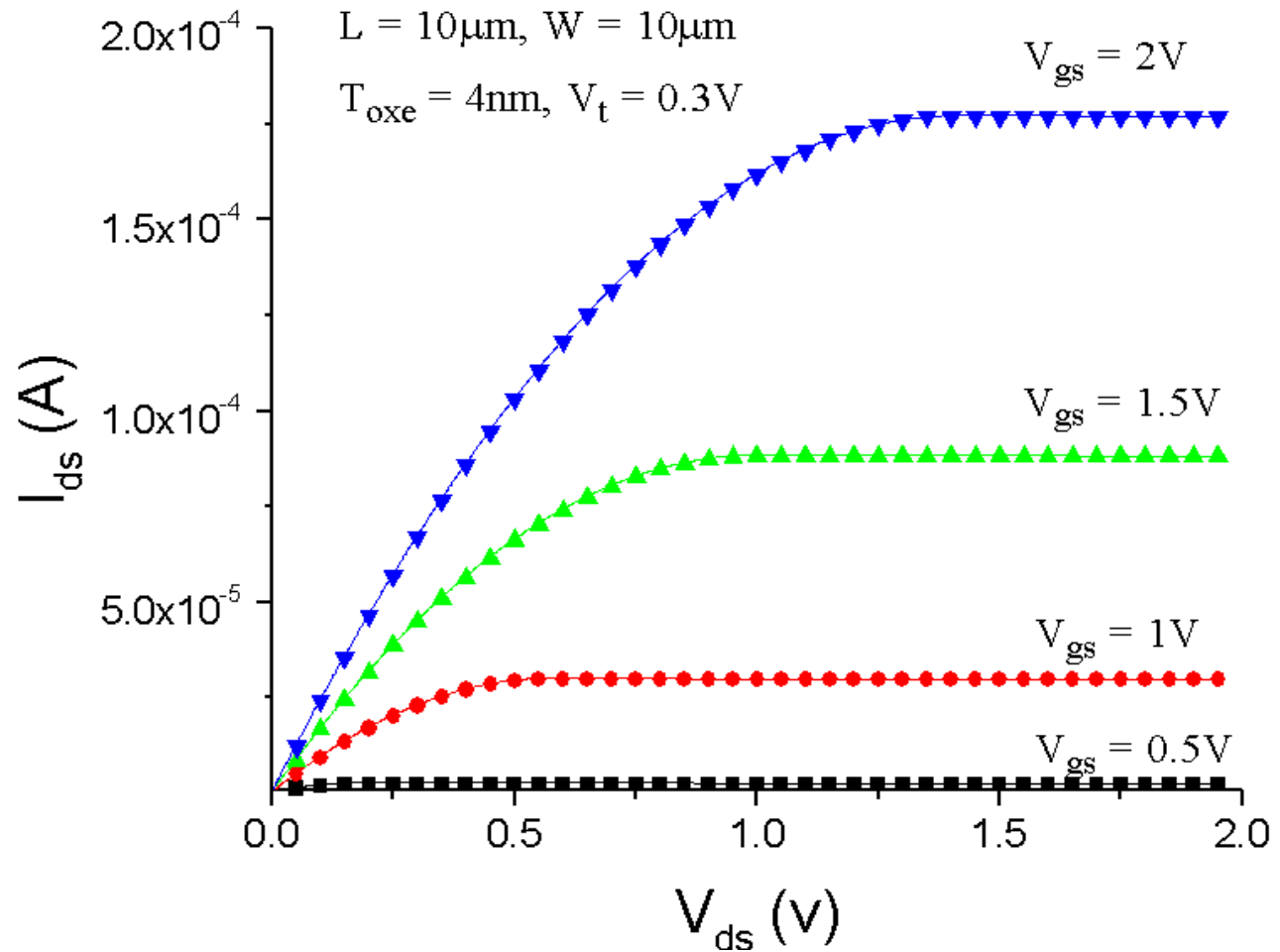
When $V_{ds} = V_{dsat}$, Q_{inv} at the drain end of the channel, is zero.

This disappearance of the inversion layer is called channel **pinch-off** (Inversion charge density is pinched off).

- A depletion region does not stop current flow as long as there is a supply of the right carriers.
- when the electrons reach the pinch-off region of a MOSFET, they are swept down the steep potential drop.
- Therefore, the pinch-off region does not present a barrier to current flow.

Figure 6-11
 n-channel
 MOSFET cross-
 sections under dif-
 ferent operating
 conditions: (a) lin-
 ear region for
 $V_G > V_T$ and
 $V_D < (V_G - V_T)$;
 (b) onset of
 saturation
 at pinch-off,
 $V_G > V_T$ and
 $V_D = (V_G - V_T)$;
 (c) strong saturation,
 $V_G > V_T$ and
 $V_D > (V_G - V_T)$.





- $I_{\text{ds}} - V_{\text{ds}}$ curves also change with gate voltage (V_G).
- When V_G increases, slope of the curves increases.
- $V_{\text{ds}}(\text{Sat})$ is a function of V_G .

Transconductance:

$$g_m = [di_{ds}/dV_{gs}] \text{ at } V_{ds}$$

- It is a measure of a transistor's sensitivity to the input voltage.
- In general, a large g_m is desirable

$$I_{dsat} = \frac{W}{2mL} C_{oxe} \mu_{ns} (V_{gs} - V_t)^2$$

$$g_{msat} = \frac{W}{mL} C_{oxe} \mu_{ns} (V_{gs} - V_t)$$

MOSFET I-V Characteristics

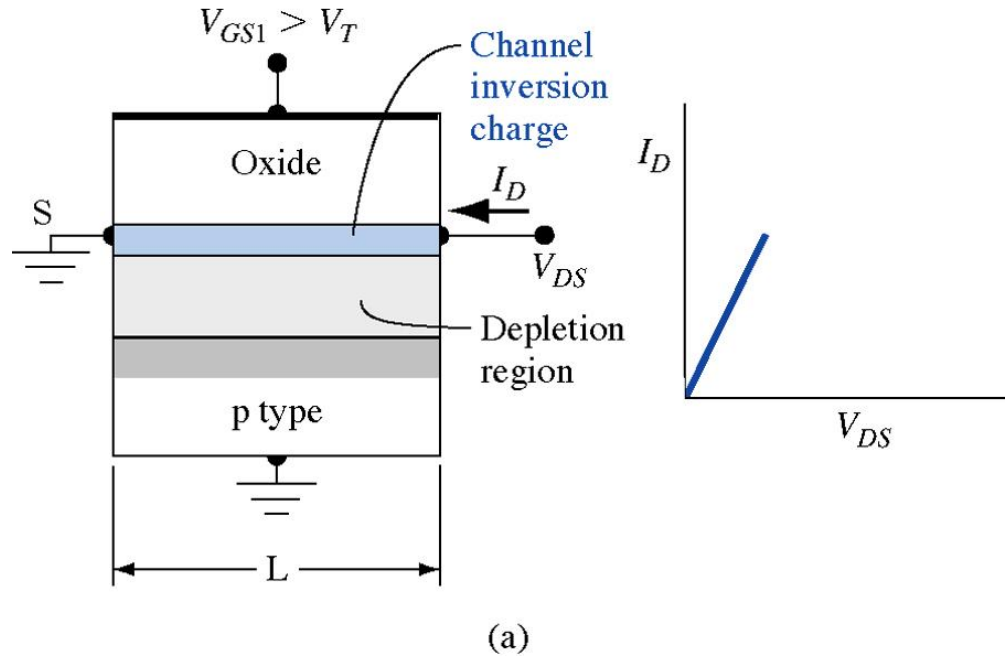
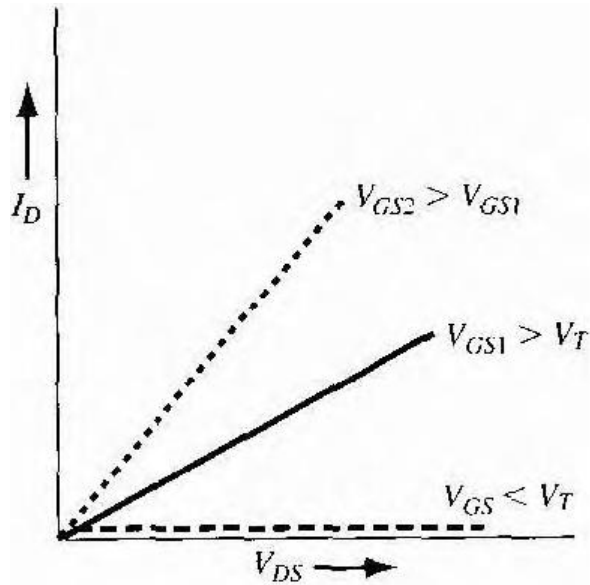
$V_G < V_T$ & V_{DS} is small

- Source & body terminals are grounded
- Drain-to-source reverse biased
- No inversion layer
- Drain current is zero

$V_G > V_T$ & V_{DS} small

- Inversion layer is formed
- Drain current flows for small V_{DS} value from drain-to-source
- No current through the oxide to gate channel
- Channel acts as a resistor
- $I_{DS} = g_m V_{DS}$
- Q_{inv} is a function of gate voltage V_G
- MOSFET basis action is modulation of channel conductance by gate voltage, which determines the drain current.

“n-channel ENHANCEMENT MODE MOSFET”

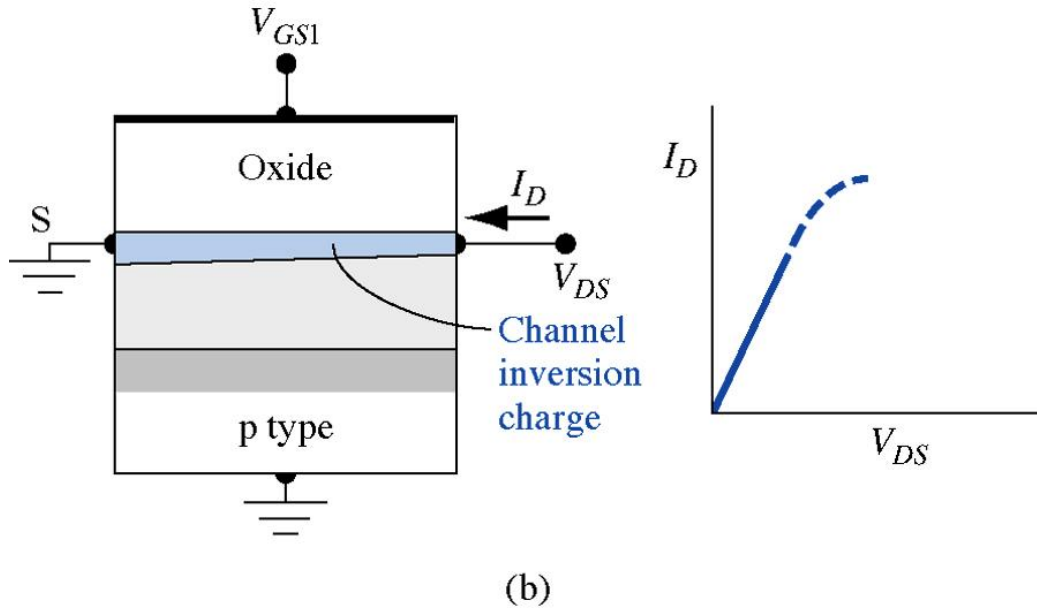


When $V_G < V_T$:-
drain current is zero

When $V_G > V_T$:-

- Drain current increases.
- g_m slope of the I-V curve is large.
- Thickness of the channel is const. along the entire length of the channel.

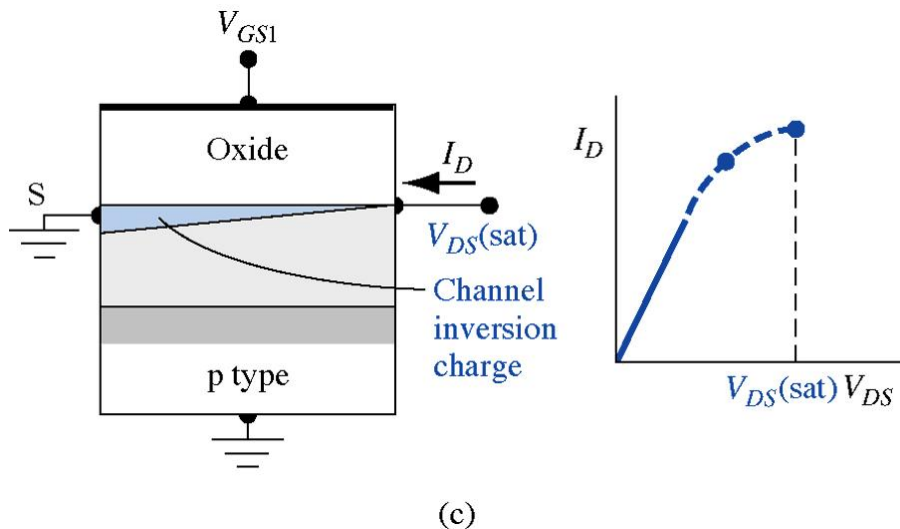
“n-channel ENHANCEMENT MODE MOSFET”



When V_{DS} increases:-

- The voltage drop across oxide layer near the drain increases.
- Inversion charge density near drain decreases.
- Conductance decreases;- slope of $I_D - V_{DS}$ decreases

“n-channel ENHANCEMENT MODE MOSFET”



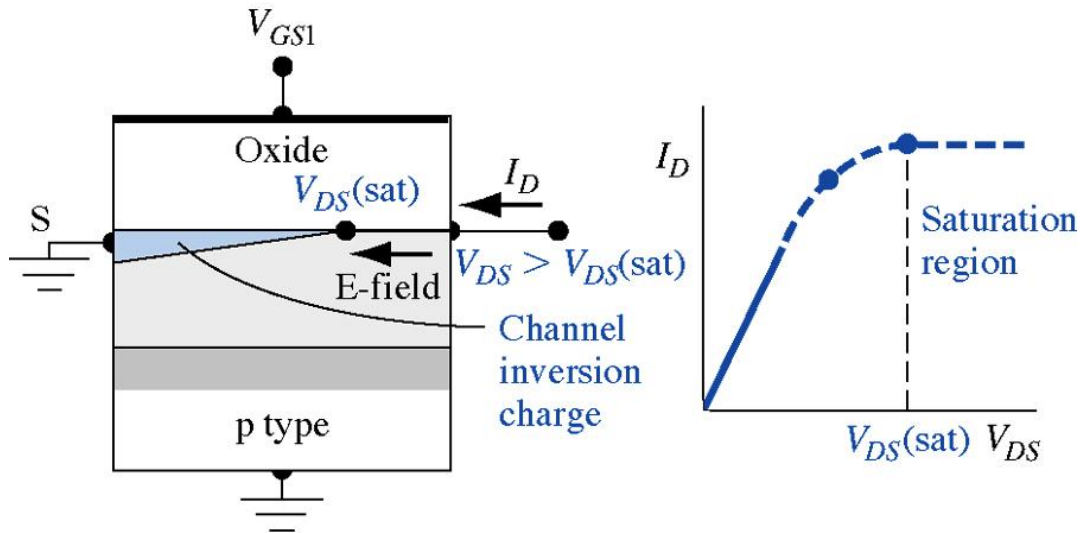
When $V_{DS} = V_{DS}(\text{Sat})$:-

- Potential drop across the oxide near drain = V_T
- Inversion charge density $Q_{inv} = 0$
(Pinch-Off Condition)
- Conductance is zero.
- $V_{DS}(\text{Sat}) = V_G - V_T$

➤ $V_{DS}(\text{Sat})$ is the value of drain-to-source voltage for which inversion charge density near drain is zero;- **Pinch-off condition.**

➤ I_{ds} is a function of gate-to-source voltage only.

“n-channel ENHANCEMENT MODE MOSFET”

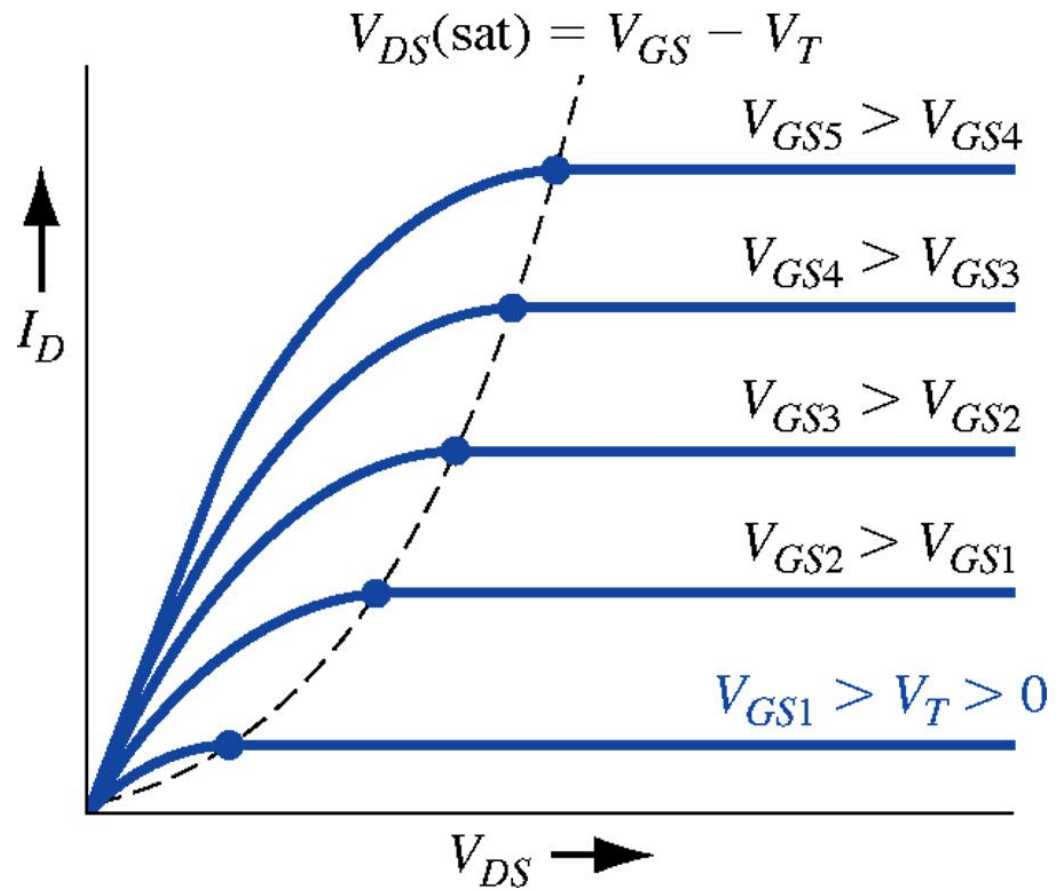
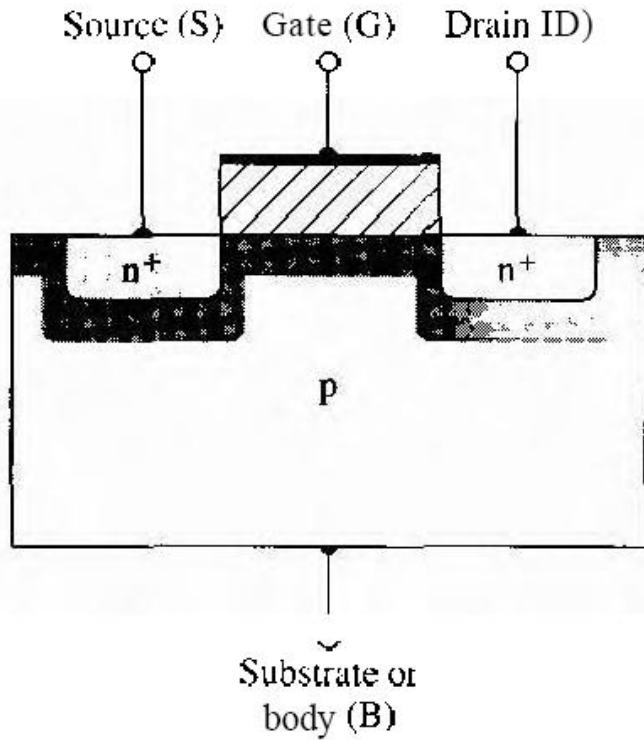


(d)

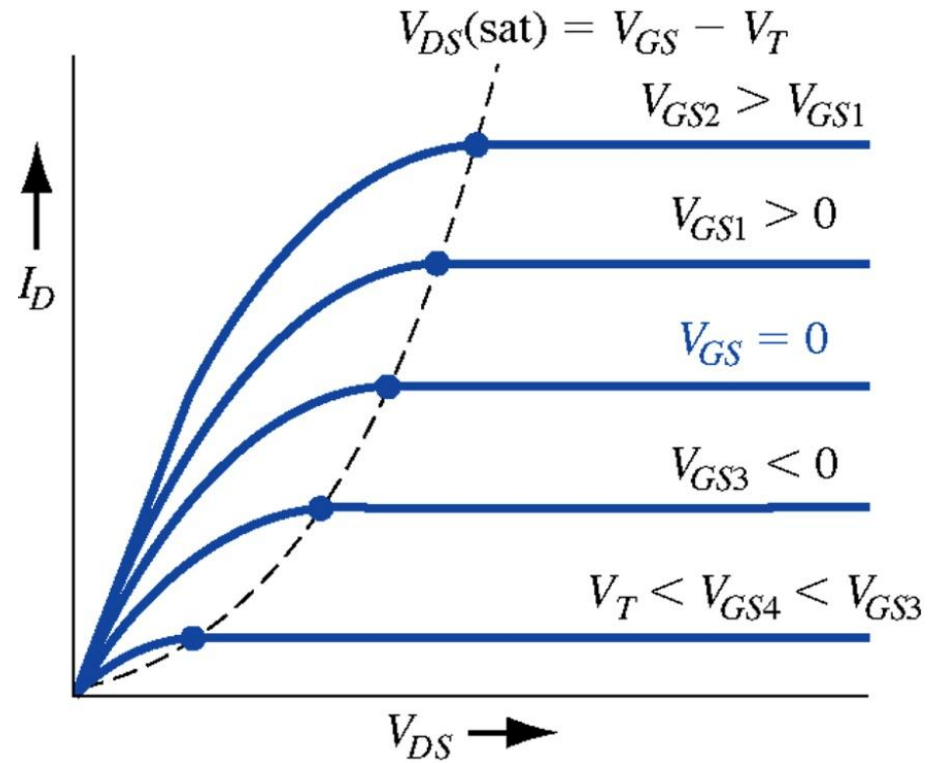
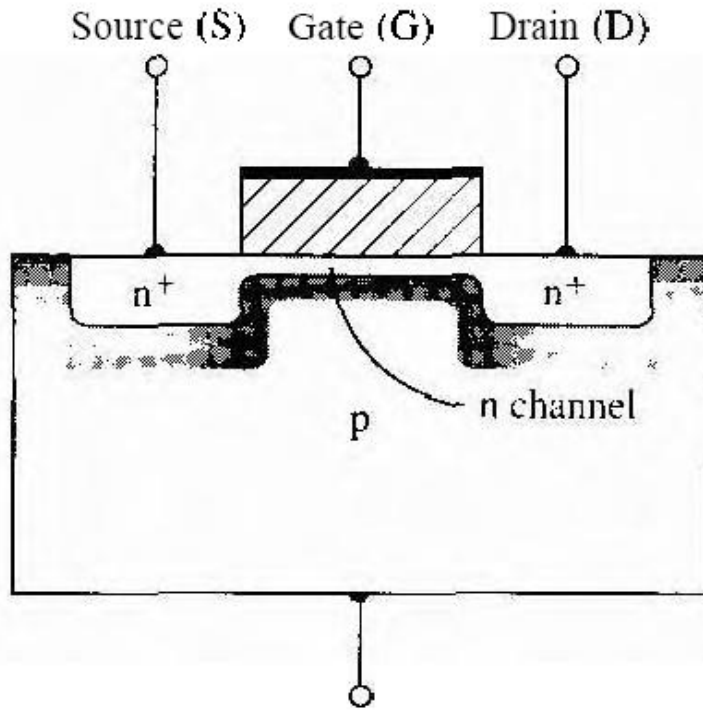
$V_{DS} > V_{DS(sat)}$:-

- Pinch-off point moves into the channel
- At this point, electrons injected into the depletion region & are swept away by the EF at drain contact.
- Channel length decreases. $V_{cs}(x)$, $Q_{inv}(x)$ and therefore I_{ds} are the same.
- For small decrease in channel length in this case, I_D is const.:- **Saturation Point**

“n-channel ENHANCEMENT MODE MOSFET”

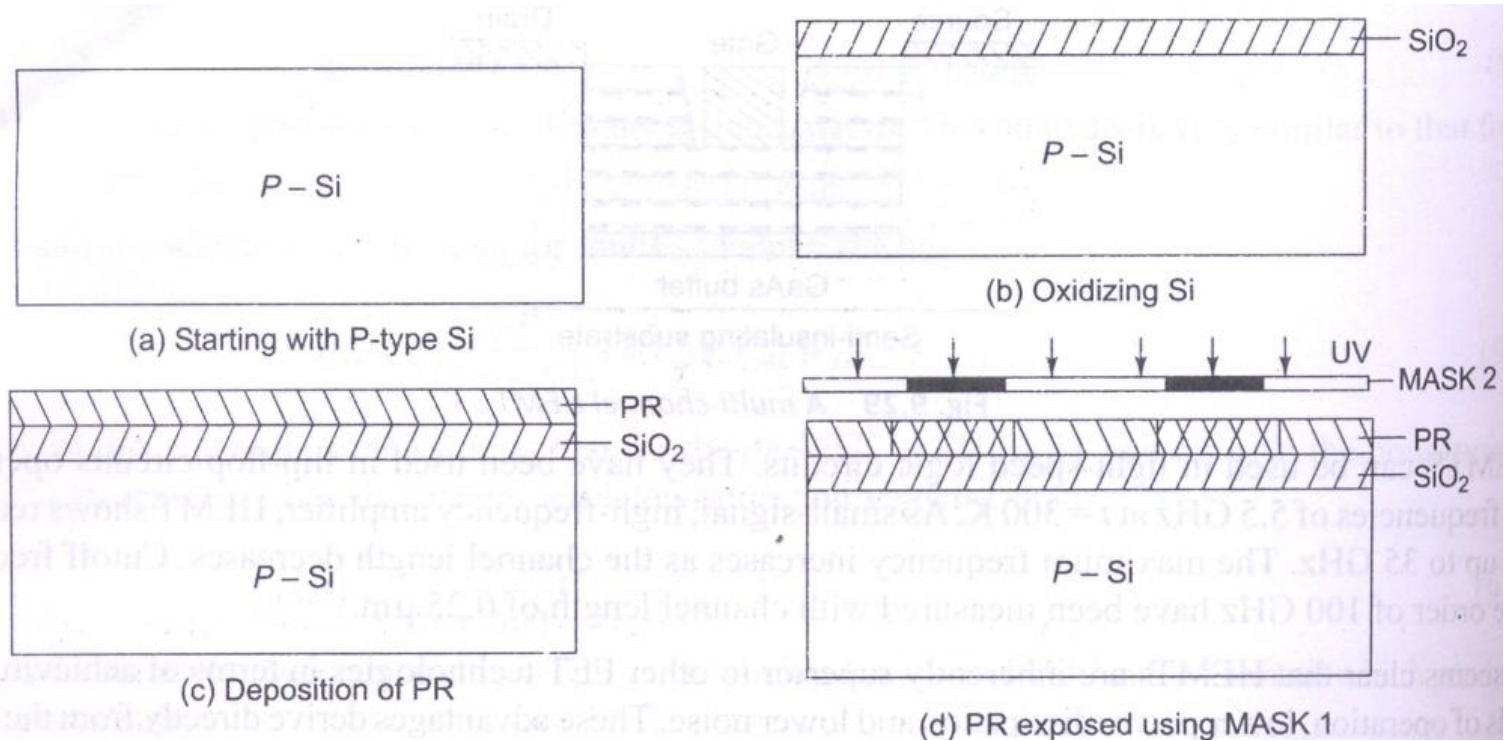


“n-channel DEPLETION MODE MOSFET”



Device is “normally on” & is turned “off” by applying negative gate voltage

MOSFET Fabrication

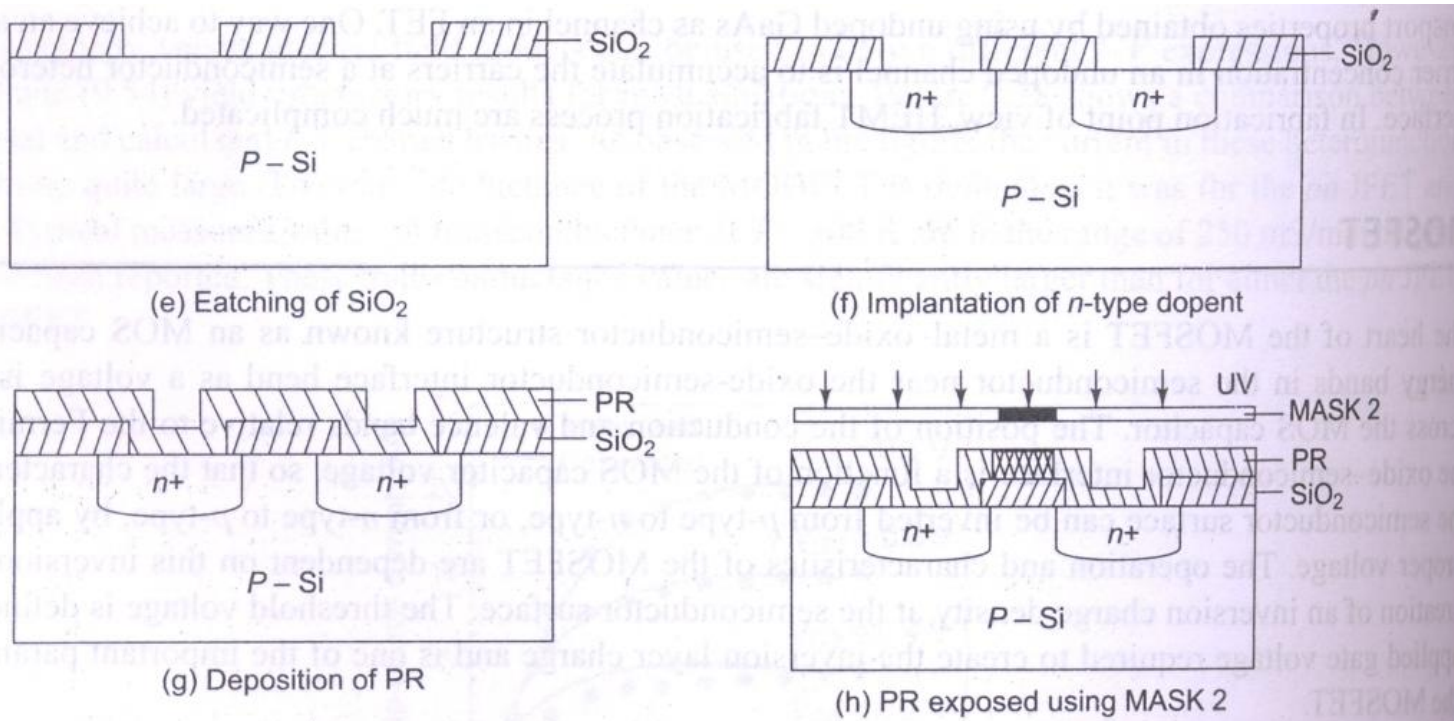


Step (a):- A thin wafer (0.4mm & 75.1nm dia) cut from a single crystal of high pure 'Si' & doped with p-type impurity.

(b) A layer of SiO_2 of $1\mu\text{m}$ is grown on the surface by wet or dry oxidation.

(c) Surface is covered with Photo-Resist uniformly.

(d) A mask is placed on PR, through which UV radiation is incident, s.t. the exposed portion get polymerized and hardened.



(e) Unpolymerised portion is etched away with SiO₂ layer using suitable etchant Solution. Then polymerised PR is also removed creating windows on SiO₂ coated Si, for diffusion of desired impurity.

(f) 0.1 μm thickness SiO₂ layer (thinox) is grown over entire chip surface by oxidation.

(g) PR coating, masking & etching is used to deposit polysilicon on thin oxide layer by CVD at the window.

(h) The thinox layer is removed by suitable etching procedure.

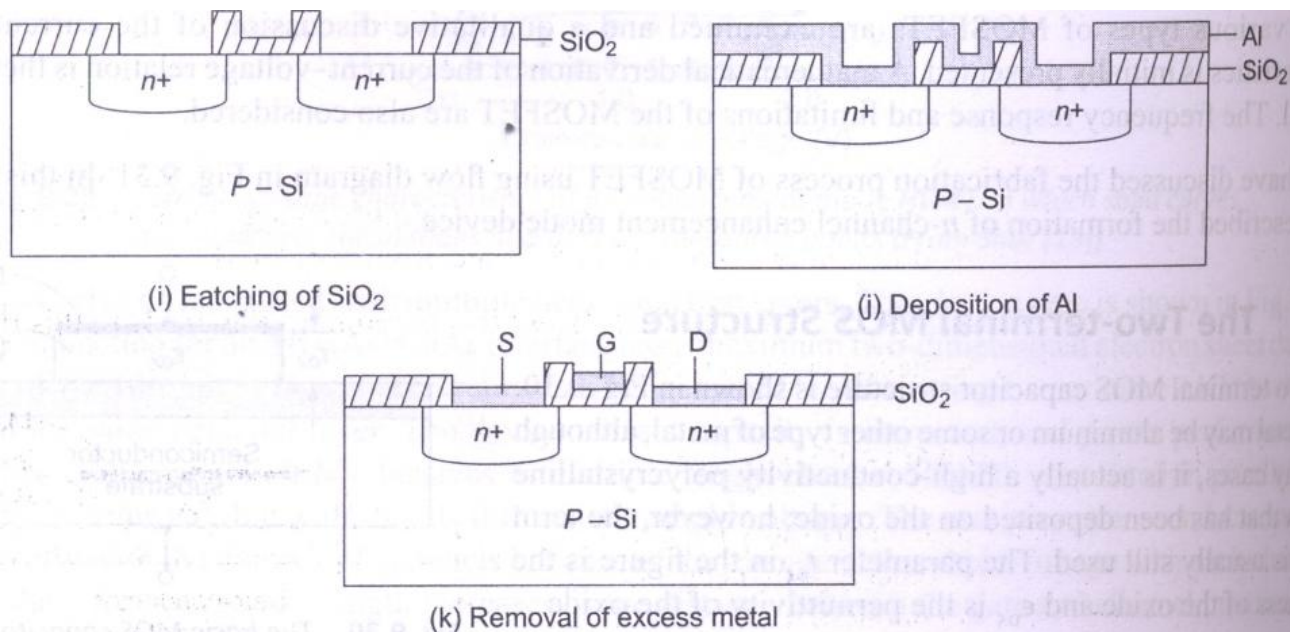


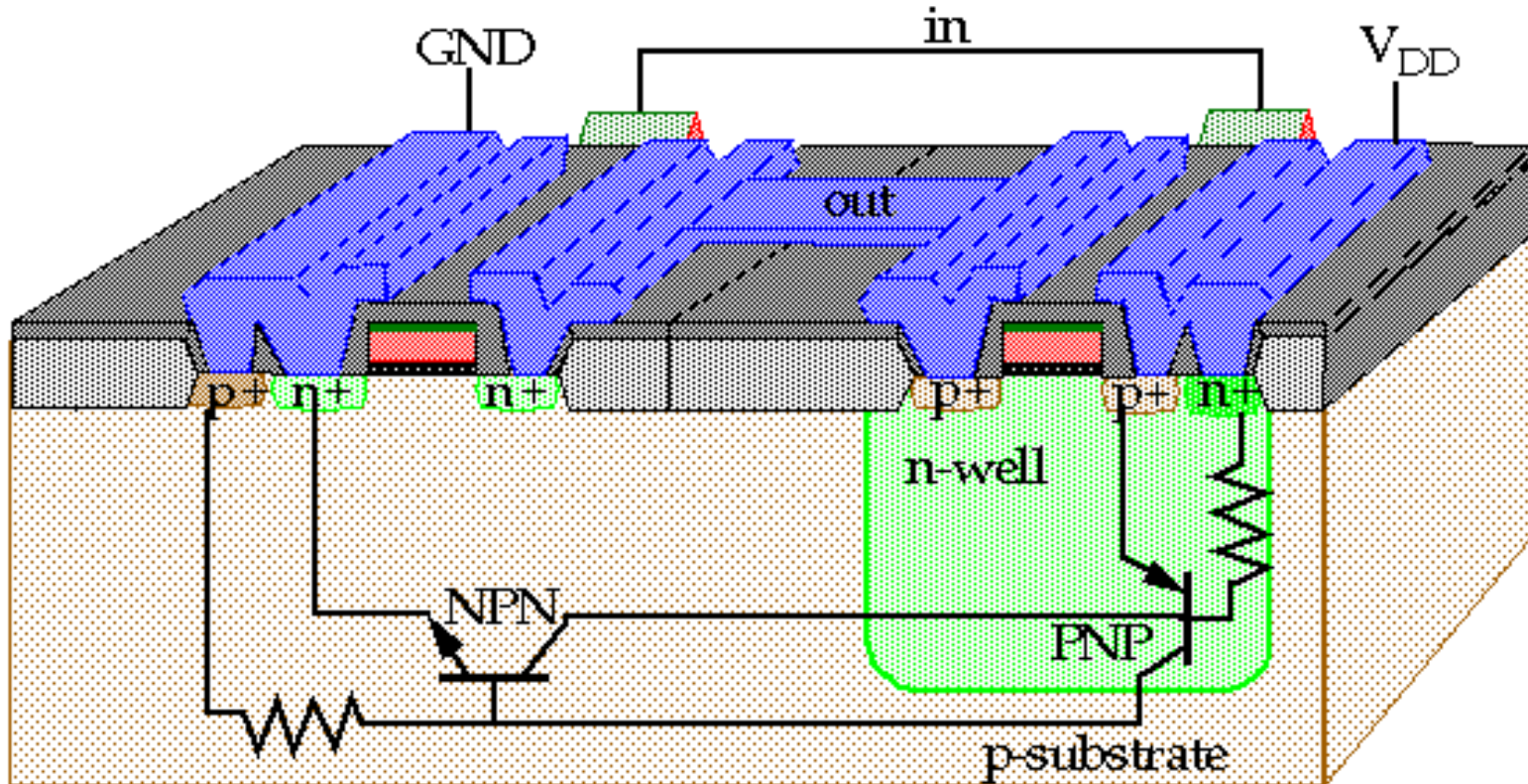
Fig. 9.31 Fabrication of MOSFET

(i) The n-type impurities are diffused into exposed area of p-type Si, by heating to high temp. and passing a gas containing the impurity for source and for drain.

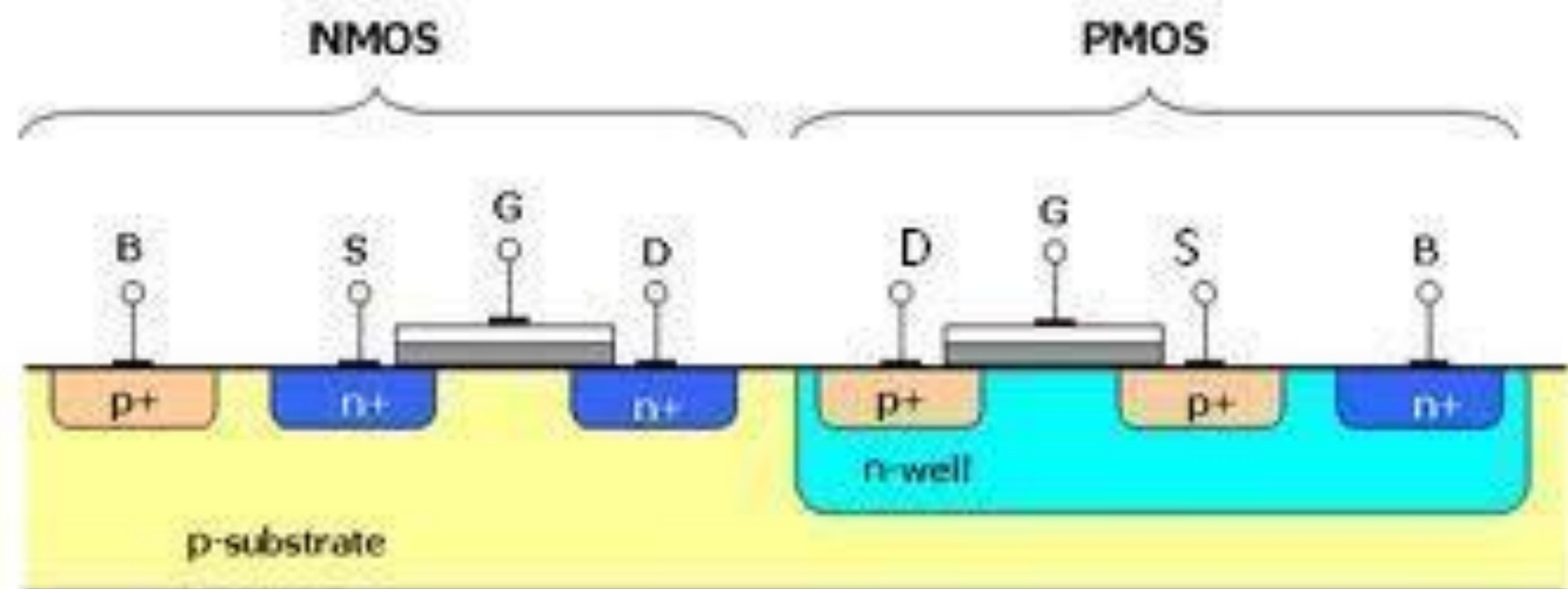
(j) Thick oxide layer is formed on whole surface, coated with PR and mask to select areas for polycrystalline gate, source and drain connections. Contact points are determined by exposing to UV radiation and etching as before.

(k) A thin 'Al' film (1 μ m) is coated over the surface, masked and etched to fix contact points for source, drain and gate.

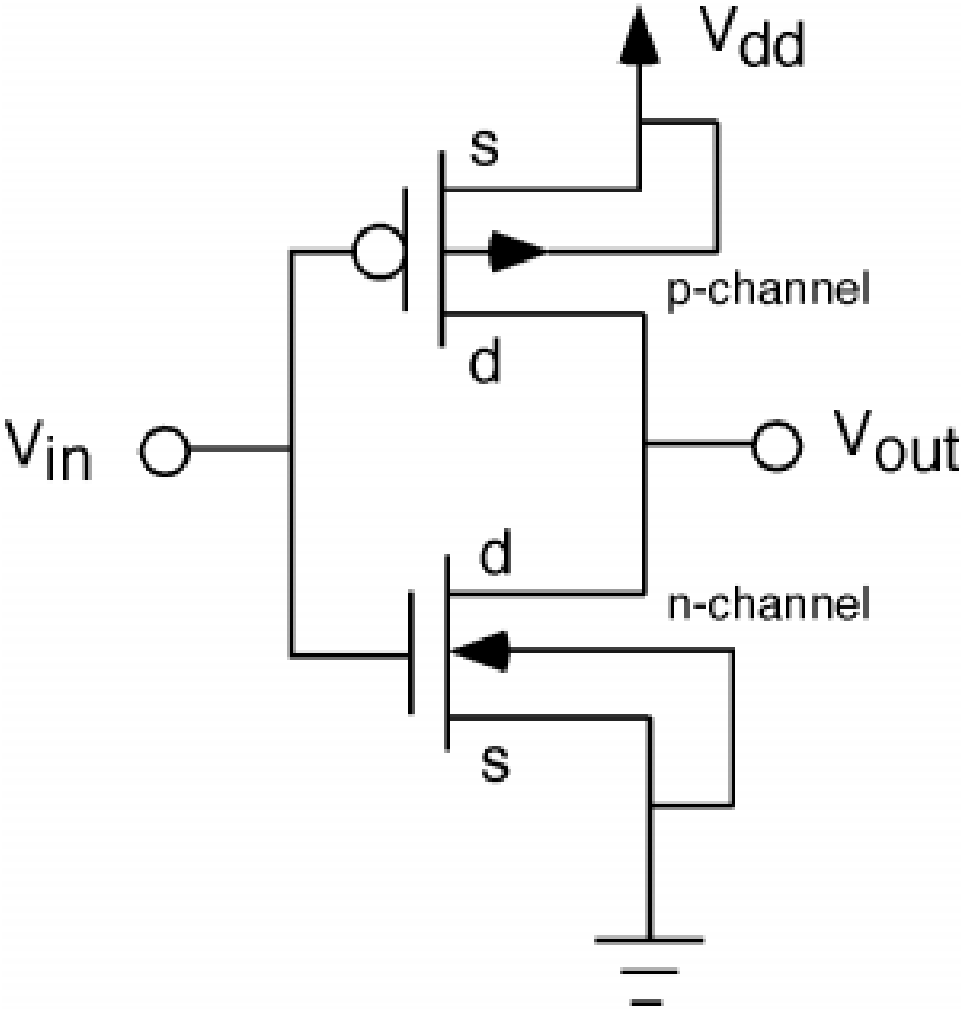
Complementary MOSFETs(CMOS) Technology



- Nfet and Pfet can be fabricated on the same chip. Potential problem in CMOS is **latch-up**, the high current, low voltage condition that may occur in a four layer *pnpn* junction.



CMOS (Complementary MOS) Inverter



CMOS inverter is made of a PFET pull-up device and a NFET pull-down device.

Latch-Up in CMOS

- Latch-up is a term used in integrated circuits (ICs) to describe a particular type of short circuit which can occur in an improperly designed circuit.
- More specifically it is the undesirable creation of a low-impedance path between the power supply rails of a MOSFET circuit, triggering a parasitic structure which disrupts proper functioning of the part, leading to its destruction due to over current.
- The parasitic structure is usually equivalent to a PNPN structure which acts as a PNP and an NPN transistor stacked next to each other.
- During a latch-up when one of the transistors is conducting, the other one begins conducting too. They both keep each other in saturation for as long as the structure is forward-biased and some current flows through it.
- The parasitic structure is formed as a part of the PMOS and NMOS transistor pair on the output drivers of the gates.
- A power cycle is required to correct this situation.

Latch-Up

