VII Field Effect Transistors

7.1 A brief History

The age of semiconductor technology essentially started in 1948 with the invention of the point-contact transistor and the Bipolar Junction Transistor (BJT) approximately a year later. Interestingly enough, the Field Effect Transistor (FET) was conceived decades earlier (in the late 1920's), but its inventor (Lilienfeld) never got it to work since he did not fully appreciate the crucial role of surface defects or surface states.

BJT technology reigned supreme during the 1950's and 1960's.

The first functional FET was demonstrated in 1960 after the problem of surface states was resolved by growing an oxide insulator on Silicon. This new FET, named Metal-Oxide Semiconductor (MOS) FET, gradually replaced the BJT and is the supreme ruler in today's semiconductor industry.

However, some interesting alternatives to the MOSFET emerged over the last few decades that deserve mentioning.

Today's FET family includes:

FET
- MOSFET
- JFET (Junction FET)
- MESFET (Metal-Semic. FET)
- MISFET (Metal-Insulator-Semic. FET)

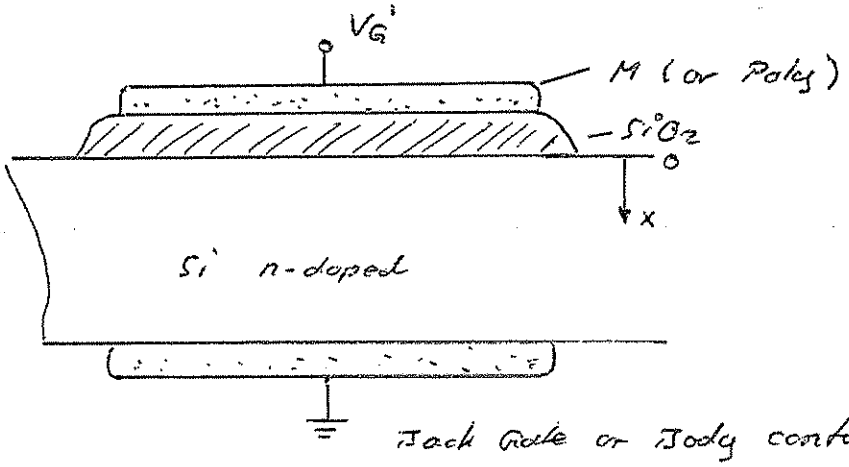
The common feature of all FETs is the control of the channel charge by means of an electric field (imposed by a "Gate" voltage).

Due to its comparatively simple physical implementation, the MOSFET's market dominance is unquestioned. However, MESFETs and MISFETs represent interesting "High-Speed" alternatives.

JFETs and MESFETs control the channel charge via a reverse biased junction (a p-n junction for the JFET and a Shottky barrier for the MESFET, respectively). The MOSFET controls the charge in the channel through a dielectric layer (the gate oxide). Hence, the Metal (or Poly-Silicon)-Oxide-Semiconductor Capacitor plays a key role in its functionality. We therefore begin our investigation with the MOS capacitor.

7.2 The MOS Capacitor

Basic Configuration



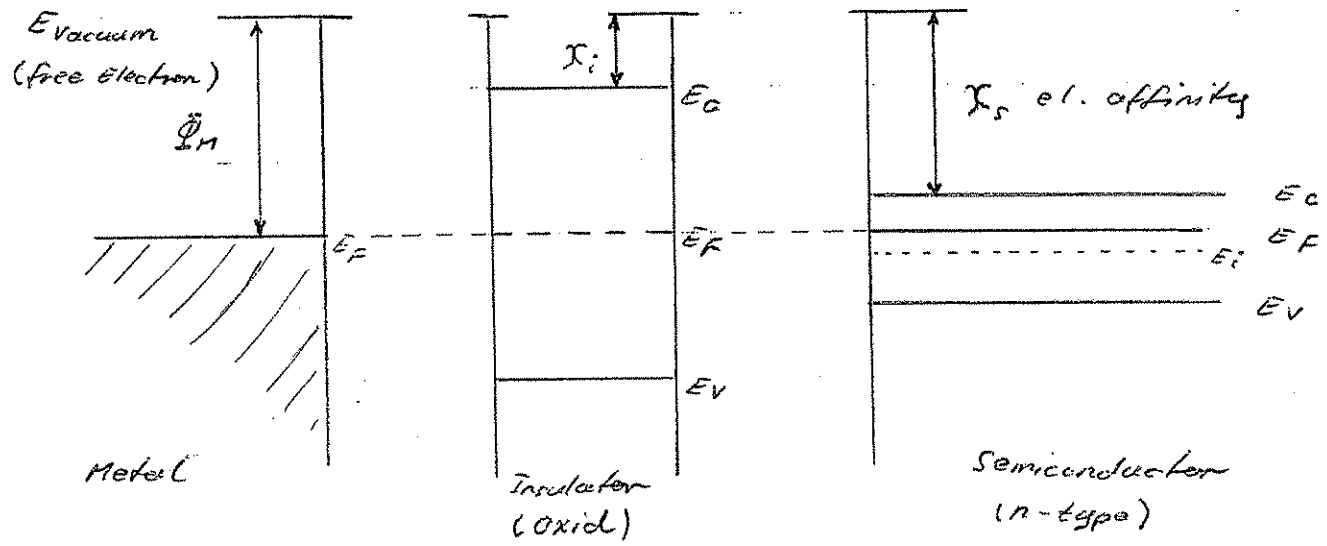
assumptions:

1. Gate is equipotential region
2. Oxide is perfect insulator
3. No charge centers in oxide
4. semiconductor is uniformly doped
5. Semic. is sufficiently thick so that a field-free region is formed before reaching the back contact
6. Back contact is ohmic
7. Capacitor is a one-dimensional structure in x
8. $\Phi_M = \chi_s + (E_c - E_F)$

Work fct. of Metal — el. affinity

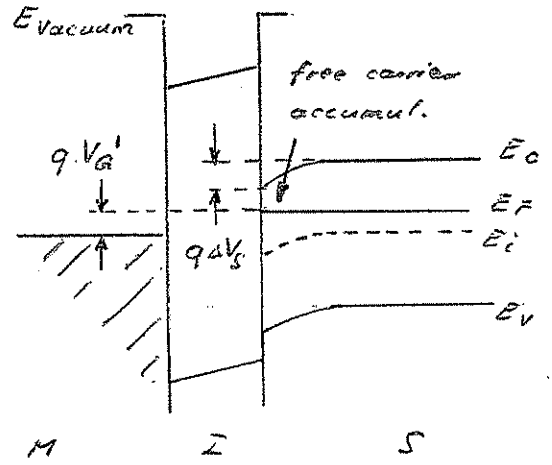
Energy-Band Diagrams

zero Bias ($V_G' = 0$)



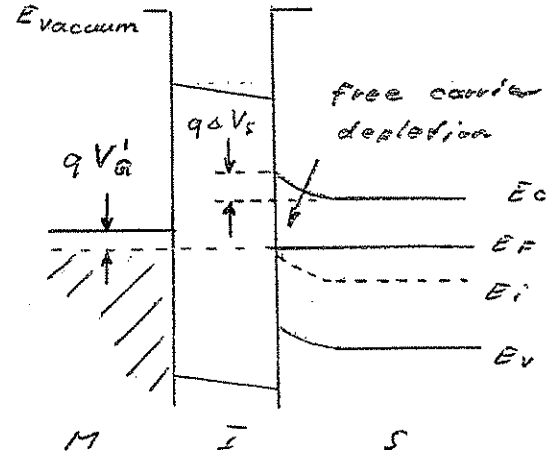
Nonzero Bias

a) $V_G' > 0$



ΔV_s: semiconductor surface potential

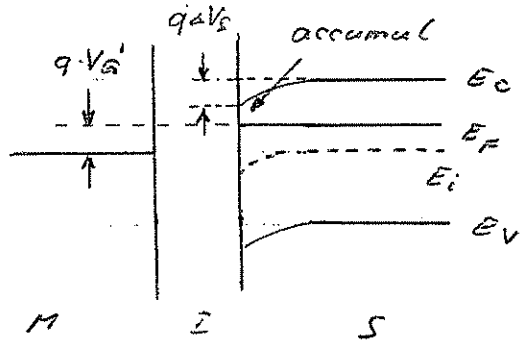
b) $V_G' < 0$ but $q\Delta V_s < (E_F - E_{i\infty})$

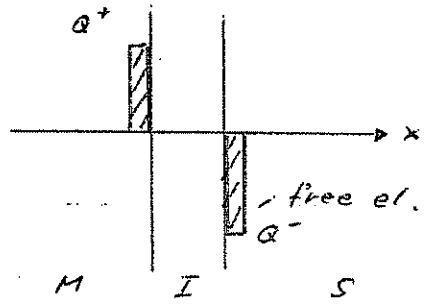


$E_{i\infty}$ denotes intrinsic Fermi level in field-free (bulk) Si

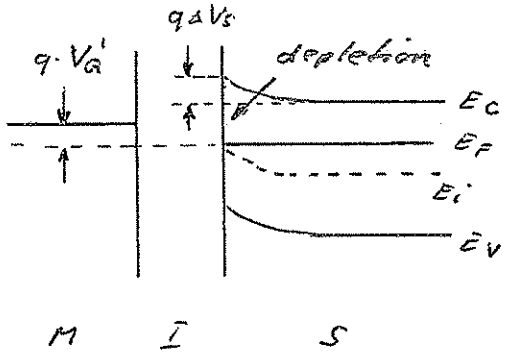
Charge distribution for different bias voltages

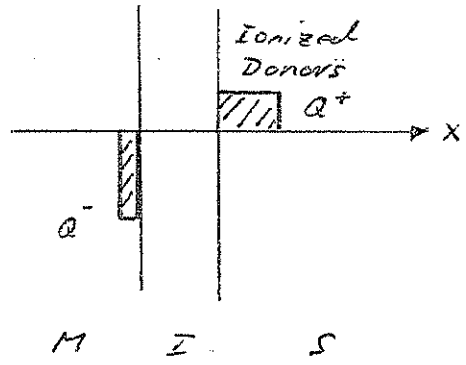
a) $V_G' > 0$



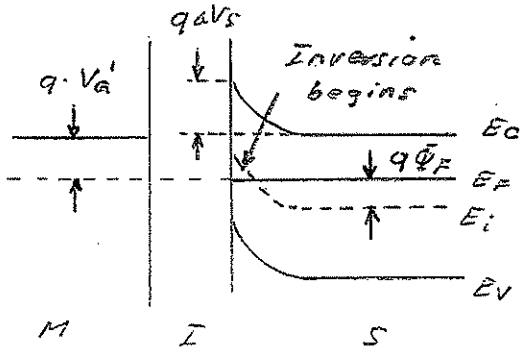


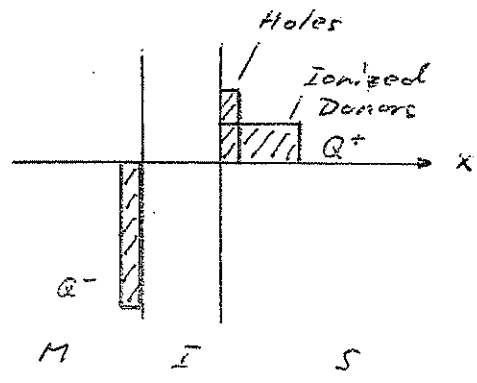
b) $V_G' < 0$ but $q\Delta V_S < (E_F - E_{i\infty})$





c) $V_G' < 0$ but $q\Delta V_S > (E_F - E_{i\infty})$





Inversion is achieved if

$$E_i(\text{surface}) - E_{i,\text{bulk}} \geq 2(E_F - E_{i,\text{bulk}}) \quad \text{Inversion}$$

or $$-q\Delta V_s \geq 2(E_F - E_{i,\text{bulk}}) \quad (n\text{-type } S)$$

The gate voltage V_G' that must be applied to achieve inversion is called Threshold voltage V_T'. Since $V_G' = \Delta V_{ox} + \Delta V_s$, we obtain:

$$V_T' = -2\frac{(E_F - E_{i,\text{bulk}})}{q} + \Delta V_{ox} = +2\Phi_F + \Delta V_{ox} \quad (n\text{-type } S)$$

Under inversion ($V_G' = V_T'$), the hole-concentration at the surface of the semiconductor is equal to the donor concentration N_D (Definition). Hence

$$p_s = N_D = n_i e^{\frac{(E_F - E_{i,\text{bulk}})}{kT}} = n_i e^{-\frac{\Phi_F q}{kT}}$$

$$W_T = \sqrt{\frac{2\varepsilon_s 2\Phi_F}{qN_D}}$$

and $$\Delta V_{ox} = -\frac{Q_{dep}}{C_{ox}} = -\frac{q \cdot N_D \cdot W_T}{C_{ox}} \qquad \left(C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}\right)$$

Solving for Φ_F and V_T' yields:

$$\Phi_F = -\frac{kT}{q} \ln\left[\frac{N_D}{n_i}\right] \quad (n\text{-type } S)$$

$$V_T' = +2\Phi_F - \frac{q \cdot N_D W_T}{C_{ox}} \quad p\text{-channel device}$$

In analogy to this result, we obtain for a p-doped semicond.

$$\Phi_F = +\frac{kT}{q} \ln\left[\frac{N_A}{n_i}\right] \quad (p\text{-type } S)$$

$$V_T' = +2\Phi_F + \frac{q N_A W_T}{C_{ox}} \quad n\text{-channel device}$$

Charge Density Distribution

The exact solution of the charge density distribution inside the semiconductor is obtained by solving Poisson's eq.

$$\left| \frac{d\mathcal{E}}{dx} = \frac{\rho}{\varepsilon} \right|$$

where $\rho = q\,(p(x) - n(x) + N_D - N_A)$

$$\left| \begin{aligned} p(x) &= n_i\, e^{\frac{(E_i(x) - E_F)}{kT}} \\ n(x) &= n_i\, e^{\frac{(E_F - E_i(x))}{kT}} \\ N_D - N_A &= n_i \left[e^{-\frac{\phi_F}{kT}} - e^{+\frac{\phi_F}{kT}} \right] \end{aligned} \right|$$

Potential

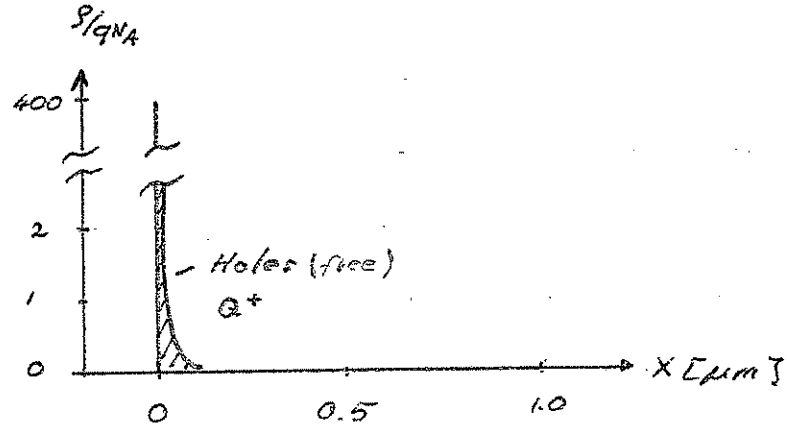
since $\rho = 0$ and $\phi = 0$ in semiconductor bulk

Example: Solution of Poisson's eq. for p-type Si with

$N_A = 10^{15}\,cm^{-3}$ ($\phi_F = 0.3\,V$)

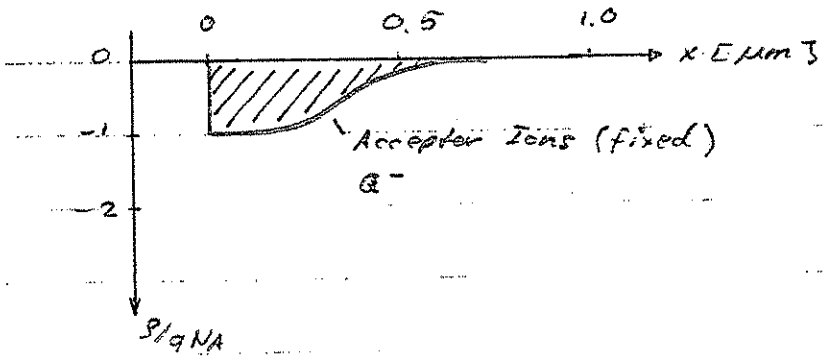
a) Accumulation

$\Delta V_s = -\frac{1}{2}\phi_F$



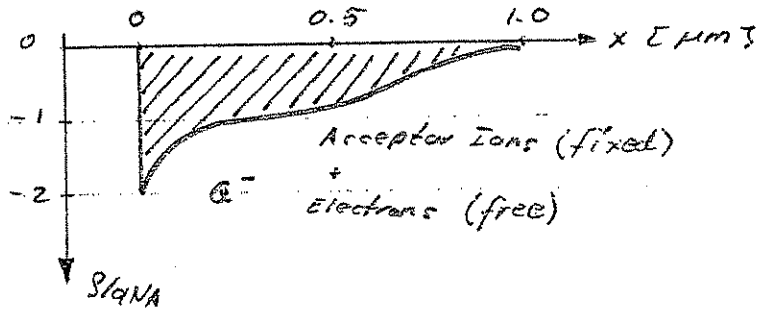
b) Depletion

$\Delta V_S = \phi_F$



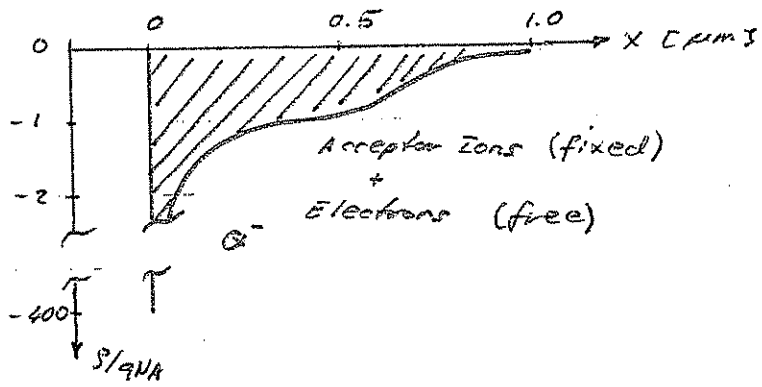
c) Onset of Inversion

$\Delta V_S = 2\,\phi_F$



d) Deep Inversion

$\Delta V_S = 2.5\,\phi_F$



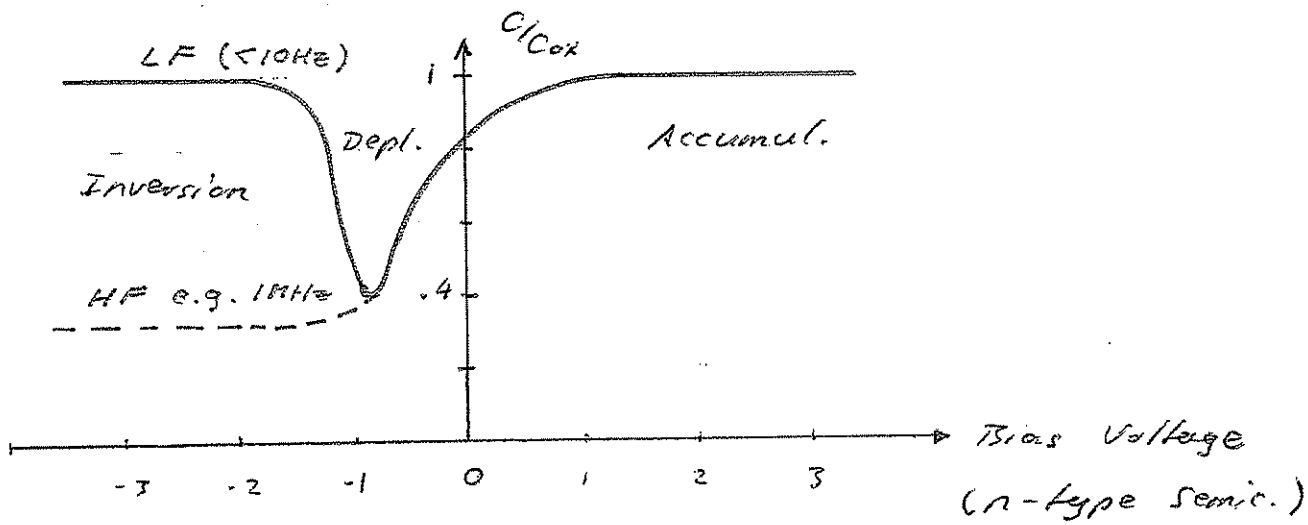
Note: The free charge carriers created through accumulation (holes in p-Si) or inversion (el in p-Si) reside in an extremely narrow portion of the semiconductor immediately adjacent to the oxide layer. By comparison, the depletion portion (acceptor ions in p-Si) extends much deeper into the semiconductor.

The depletion width at the onset of inversion is approximately given by:

$$\left| W_T = \sqrt{\frac{2\varepsilon_s 2\phi_F}{q N_A}} \right| \qquad (p\text{-type})$$

Note that the depletion width approx. remains constant if the bias voltage is increased above the threshold voltage.

Capacitance - Voltage Characteristics



Accumulation

$$C_{acc} \approx C_{ox} A_G = A_G \frac{\varepsilon_{ox}}{t_{ox}}$$

majority carrier

$\tau \approx 1ps$

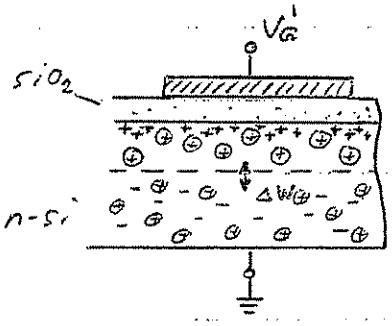
Note: $C_{ox} = \left[\frac{F}{m^2}\right]$

Depletion

$$C_{dep} \approx \frac{C_{ox} C_s}{C_{ox} + C_s}$$

ionized acceptors in depletion region

where $C_s = \varepsilon_s \frac{A_a}{W_T}$



$$C_{depl} = \frac{C_{ox}}{1 + \frac{\varepsilon_{ox}}{\varepsilon_S}\frac{W_T}{t_{ox}}}$$

depl. width is changed by changing el. concentration at the boundary between depl. region and the bulk semiconductor
→ involves majority carriers → short time const.

Inversion

In this case, charge variations in the semiconductor can be caused by small changes in the depletion width W (majority carriers → short time constant) or by changes in the inversion layer at the semiconductor-oxid interface (minority carriers must be generated → large time const.)
thus, the effective capacitance under inversion depends on the frequency of the applied field

a) $\omega \to 0$

$$C_{inv} \approx C_{ox}$$

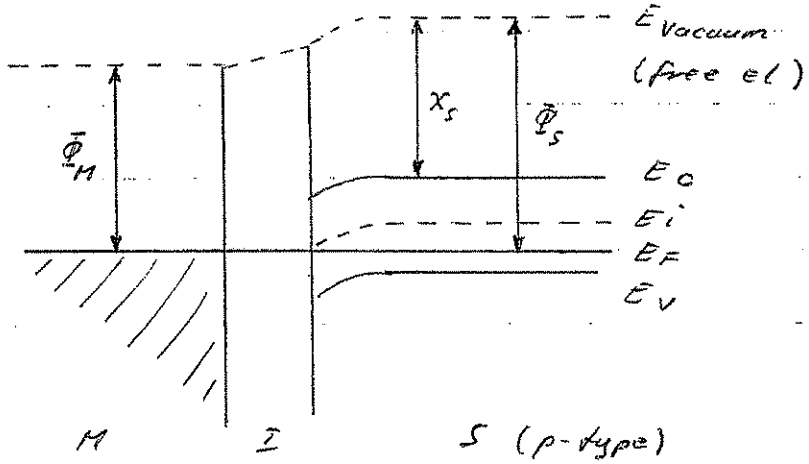
b) $\omega \to \infty$

$$C_{inv} \approx \frac{C_{ox}}{1 + \frac{C_{ox}}{C_{depl}}} = \frac{C_{ox}}{1 + \frac{\varepsilon_{ox}}{\varepsilon_S}\frac{W_T}{t_{ox}}}$$

Deviations from the MOS Ideal

1. Metal-Semiconductor Workfunction Difference

Assumption A), i.e. $\Phi_M = \chi_s + (E_c - E_F)$ is generally not true. In other words, the semiconductor does not exhibit a flat-band diagram under zero bias conditions. The exact band structure looks rather like



In order to achieve flat band conditions we must apply a certain gate voltage ΔV_{G1}, where

$$\Delta V_{G1} = V_{FB1} = \frac{1}{q}\left[\Phi_M - \Phi_S\right] = \frac{1}{q}\Phi_{MS} = \phi_{MS}$$

In most cases, V_{FB1} is negative and on the order of a few tenth of a volt (depending on doping conc. and semic. type)

2. Oxide Charges

Mobile oxide charges (mostly Na^+)

Poisson's eq. $\frac{dE_{ox}}{dx} = \frac{\rho_{Mox}}{\varepsilon_{ox}}$ ... mobile oxide charge per area

$$\Delta V_{G2_1} = V_{FB2_1} = -\frac{1}{\varepsilon_{ox}} \int_0^{t_{ox}} x \, \rho_{Mox} \, dx = -\gamma_M \frac{Q_M}{C_{ox}} \qquad (0 < \gamma_M < 1)$$

effect reduced by "neutralization" of semiconductor - oxide interface e.g. by HCl, Cl_2 (Na^+ is neutralized by Cl^-)

Fixed oxide charges

From Poisson's eq. (positive ions)

$$\Delta V_{G2_2} = V_{FB2_2} = -\frac{Q_{Fox}}{C_{ox}}$$

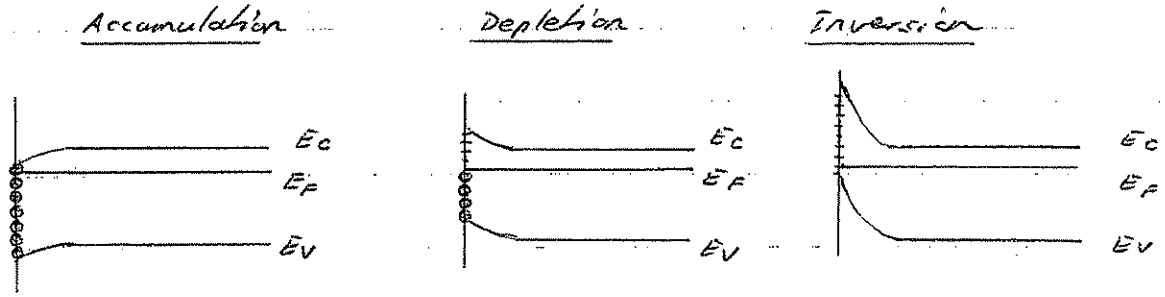
properties of Q_F:

1. Q_F is independent of t_{ox}
2. Q_F varies with Si surface orientation
3. Q_F is strong function of oxidation conditions
4. Annealing of oxidized Si in Ar or N_2 atmosphere reduces Q_F considerably

3. Interfacial Traps (Surface States)

Interfacial traps are allowed energy states in the band-gap in which electrons are "trapped" in the vicinity of the material's surface. They are created by "dangling bonds" at the semiconductor-oxide interface.

Example (n-type Si)



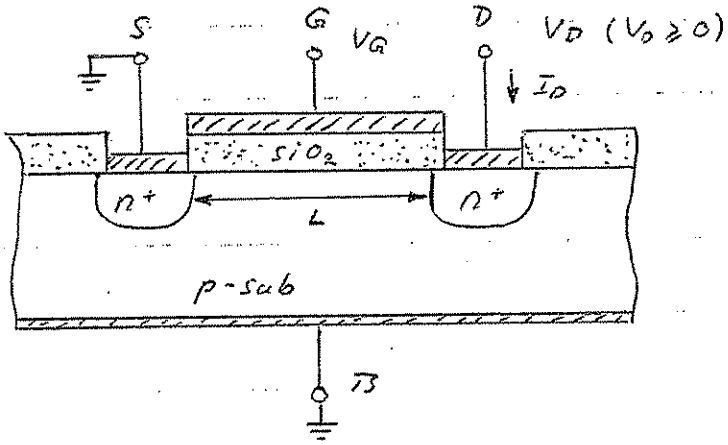
Since the charge Q_{SS} stored in the surface states, like Q_F, is located at the semiconductor-oxide interface, we can write, by direct analogy with the fixed oxide charge analysis,

$$\Delta V_{G3} = V_{FB_3} = -\frac{Q_{SS}}{C_{ox}} \qquad \text{pos or neg}$$

Q_{SS} is very sensitive to even minor fabrication details and also depends on the orientation of the semiconductor surface.
The interfacial trap concentration can be minimized in one of two ways, namely, through post-metallization annealing or hydrogen (H_2) ambient annealing.

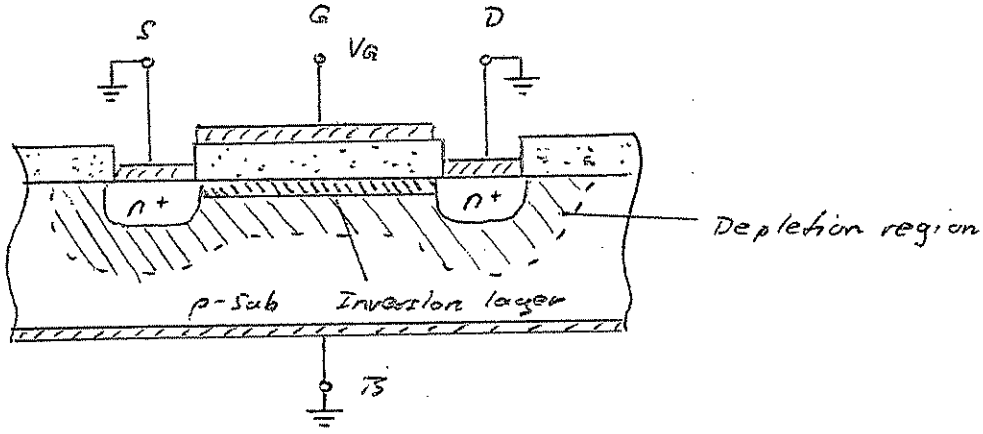
7.3 MOSFET Device Characteristics

Basic Structure (n-channel)



Operation (enhancement-mode)

$V_G > V_T$ ($V_D = 0$)



Threshold Voltage V_t

$$V_{t0} = V_{FB} + 2\phi_F + \frac{Q_{depl\,0}}{C_{ox}}$$ derived from MOS capacitor

where $V_{FB} = \phi_{MS} - \gamma_M \frac{Q_M}{C_{ox}} - \frac{Q_F}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}$ $\left(C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}\right)$

and $Q_{depl\,0} = q N_A W_T = \sqrt{2 \varepsilon_s q N_A 2\phi_F}$ $\left(\phi_F = \frac{kT}{q} \ln\left[\frac{N_A}{n_i}\right]\right)$

$V_{SB} \neq 0$ (Body-Effect)

$$\begin{aligned} V_t &= V_{FB} + 2\phi_F + \frac{Q_{depl\,0}}{C_{ox}} + \frac{Q_{depl} - Q_{depl\,0}}{C_{ox}} \\ &= V_{t0} + \gamma \left[\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F} \right] \\ &= V_{t0} + \gamma \sqrt{2\phi_F} \left[\sqrt{1 + \frac{V_{SB}}{2\phi_F}} - 1 \right] \end{aligned}$$

where $\gamma = \frac{\sqrt{2 \varepsilon_s q N_A}}{C_{ox}}$ Body effect increases with $\sqrt{N_A} \cdot \frac{t_{ox}}{\varepsilon_{ox}}$

and $V_{t0} = V_{FB} + 2\phi_F + \gamma \sqrt{2\phi_F} = V_{FB} + \sqrt{2\phi_F} \left[\sqrt{2\phi_F} + \gamma \right]$

Example $N_A = 5 \cdot 10^{22}\ m^{-3}$

$V_{FB} = -0.6\,V$

$\phi_F = 0.4\,V$

$\varepsilon_s = 10^{-10} \frac{As}{Vm}$

$C_{ox} = 2.5 \cdot 10^{-7} \frac{F}{m^2}$

$\Rightarrow$ $\gamma = 0.50 \sqrt{V}$, $V_{t0} = 0.65\,V$

$\therefore \gamma \cdot \sqrt{2\phi_F} = 0.45\,V$

$\therefore V_t(V_{SB}) = 0.65 + 0.45 \left[\sqrt{1 + \frac{V_{SB}}{2\phi_F}} - 1 \right]$

Threshold Voltage tuning via Substrate Doping

Recall: $$V_{to} = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F}$$

where

$$\gamma = \frac{1}{C_{ox}}\sqrt{2\varepsilon_s q N_{sub}}$$

$$\phi_F = V_T \ln\left(\frac{N_{sub}}{n_i}\right) \qquad V_T = \frac{kT}{q}$$

$$V_{FB} = \frac{1}{q}\left[\Phi_M - (\chi_s + \phi_F q)\right] + \frac{Q_{ox}}{C_{ox}}$$

$$\left|\frac{dV_{to}}{dN_{sub}} = \frac{d\phi_F}{dN_{sub}} + \frac{d\gamma}{dN_{sub}}\sqrt{2\phi_F} + \frac{1}{2}\gamma\frac{1}{\sqrt{2\phi_F}}\frac{d\phi_F}{dN_{sub}}\right|$$

with $\frac{d\phi_F}{dN_{sub}} = \frac{V_T}{N_{sub}}$ and $\frac{d\gamma}{dN_{sub}} = \frac{1}{C_{ox}}\sqrt{\frac{\varepsilon_s q}{N_{sub}}}$

we obtain

$$\left|\frac{dV_{to}}{dN_{sub}} = \frac{V_T}{N_{sub}}\left(1 + \frac{1}{2}\gamma\frac{1}{\sqrt{2\phi_F}}\right) + \gamma\frac{\sqrt{\phi_F}}{N_{sub}}\right|$$

We thus conclude that increasing the substrate doping concentration increases the (magnitude) of the threshold voltage. The change is inversly proportional to N_{sub}.

For small changes of N_{sub}, we can approximate the threshold voltage deviation by

$$\left|\Delta V_{to} \cong \frac{dV_{to}}{dN_{sub}}\Delta N_{sub} = \frac{\Delta N_{sub}}{N_{sub}}\left[V_T\left(1 + \frac{1}{2}\gamma\frac{1}{\sqrt{2\phi_F}}\right) + \gamma\sqrt{\phi_F}\right]\right|$$

dominant

Derivation of I/V Characteristics

$$\left| d\pi = \frac{dV(y)}{I_D} \right| \quad \text{or} \quad \left| I_D = \frac{dV(y)}{d\pi} \right|$$

$$\left| d\pi = \pi_D \cdot \frac{dy}{W} \right|$$

where $\pi_D = \frac{1}{\mu \cdot Q_{ch}}$

$$\therefore \left| I_D = \frac{dV(y)}{dy} \mu W Q_{ch} \right|$$

or

$$\left| \int_0^L I_D \cdot dy = \mu W \int_0^{V_{DS}} Q_{ch} \cdot dV = I_D \cdot L \right|$$

Finally

$$Q_{ch} = C_{ox} \cdot [V_{GS} - V_t - V(y)]$$

$$\therefore \left| I_D = \mu C_{ox} \frac{W}{L} \int_0^{V_{DS}} [V_{GS} - V_t - V] dV \right|$$

$$\left\| I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_t) V_{DS} - \frac{1}{2} V_{DS}^2 \right] \right\| \quad \begin{array}{l} V_{GS} > V_t \\ V_{DS} < V_{GS} - V_t \end{array}$$

Post Pinch-off Characteristics

$$I_D\Big|_{V_D > V_{Dsat}} = I_D\Big|_{V_D = V_{Dsat}} = I_{Dsat} \quad \text{(assumption: current remains constant)}$$

or

$$I_{Dsat} = \mu_n \frac{W}{L} C_{ox} \left[(V_G - V_t) V_{Dsat} - \frac{1}{2} V_{Dsat}^2 \right]$$

Since $Q_N(L) = 0$ when $V_y(L) = V_{Dsat}$, we can write

$$Q_N(L) = -C_{ox} (V_{GS} - V_t - V_{Dsat}) = 0$$

$$\Rightarrow \underline{\underline{V_{Dsat} = [V_{GS} - V_t] = V_{eff}}}$$

$$\Rightarrow \boxed{I_D\Big|_{V_D \geq V_{Dsat}} = I_{Dsat} = \mu_n \frac{W}{L} C_{ox} \frac{1}{2} [V_{GS} - V_t]^2}$$

post-pinch-off characteristics

According to this eq., the saturation current varies as the square of the gate voltage above turn-on, the so called "square-law" dependence.

Channel Length Modulation

If $V_{DS} > V_{eff}$, the drain side of the conducting channel is pinched-off. The effective channel length is then given by

$$\boxed{L_{eff} = L - W_T(V_{DS})}$$

where $$W_T(V_{DS}) = \sqrt{\frac{2\varepsilon_s\varepsilon_0(\phi_0 + V_{DS} - V_{eff})}{q N_{sub}}}$$

$$\frac{dI_D}{dV_{DS}} = \frac{dI_D}{dL_{eff}}\frac{dL_{eff}}{dV_{DS}} = -\frac{1}{2}\frac{W}{L_{eff}^2}\mu C_{ox} V_{eff}^2 \frac{dL_{eff}}{dV_{DS}}$$

$$\frac{dL_{eff}}{dV_{DS}} = -\frac{dW_T}{dV_{DS}} = -\sqrt{\frac{\varepsilon_s \varepsilon_0}{2q N_{sub}(\phi_0 + V_{DS} - V_{eff})}}$$

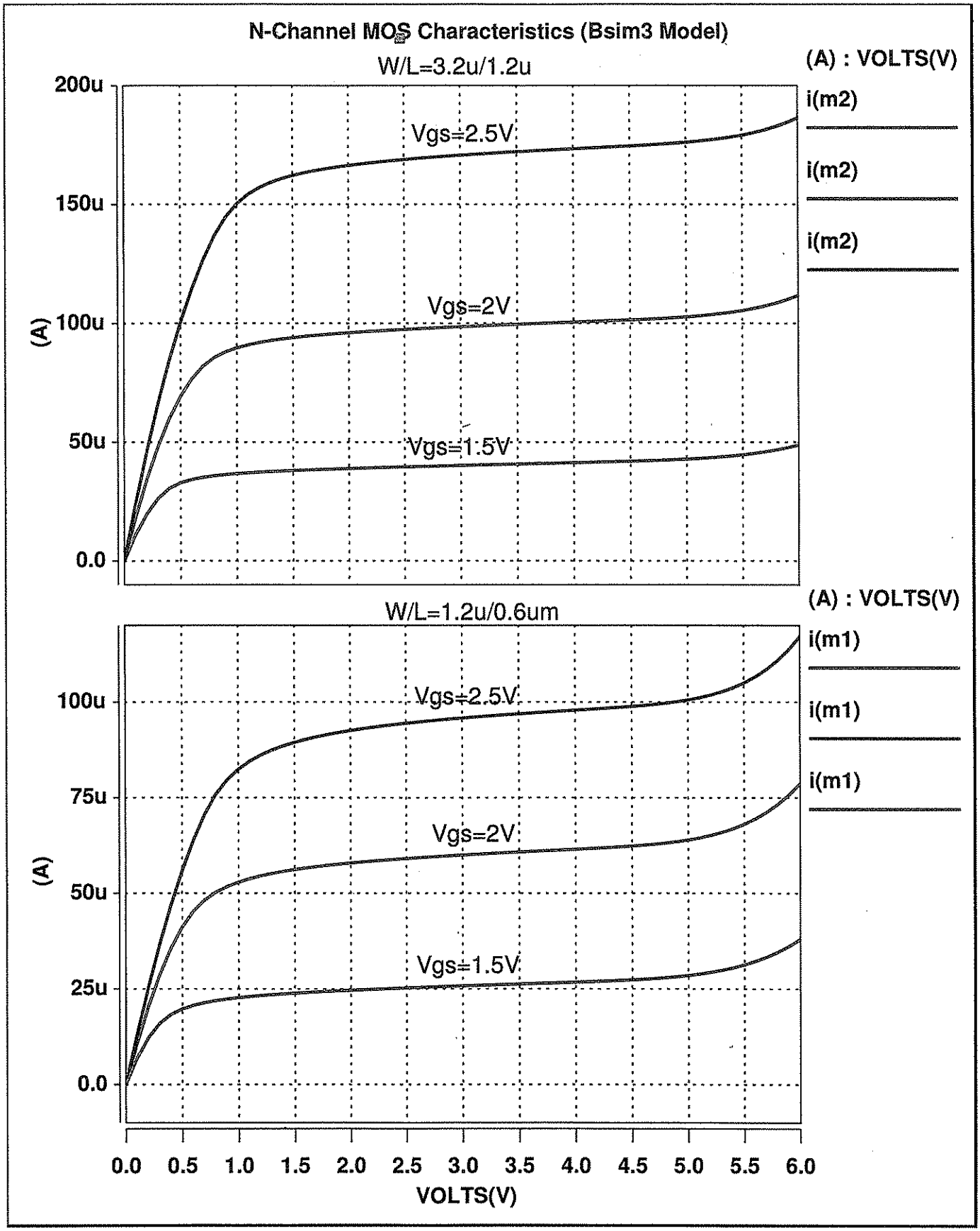
$$\therefore \frac{dI_D}{dV_{DS}} = \underbrace{\frac{1}{2}\frac{W}{L}\mu C_{ox} V_{eff}^2}_{I_{D0}} \frac{1}{L_{eff}}\sqrt{\frac{\varepsilon_s\varepsilon_0}{2qN_{sub}(\phi_0 + V_{DS} - V_{eff})}}$$

$$\therefore \boxed{I_D(V_{DS}) = I_{D0}(1 + \lambda[V_{DS} - V_{eff}])}$$

where $$I_{D0} = \frac{1}{2}\frac{W}{L}\mu C_{ox} V_{eff}^2$$

$$\lambda = \frac{1}{L_{eff}}\sqrt{\frac{\varepsilon_s\varepsilon_0}{2qN_{sub}(\phi_0 + V_{DS} - V_{eff})}}$$

$$\boxed{I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L} V_{eff}^2 (1 + \lambda[V_{DS} - V_{eff}])}$$


N-Channel MOS Characteristics (Bsim3 Model)
W/L=3.2u/1.2u
(A) : VOLTS(V)
i(m2)
i(m2)
i(m2)
200u
150u
100u
50u
0.0
(A)
Vgs=2.5V
Vgs=2V
Vgs=1.5V
W/L=1.2u/0.6um
(A) : VOLTS(V)
i(m1)
i(m1)
i(m1)
100u
75u
50u
25u
0.0
(A)
Vgs=2.5V
Vgs=2V
Vgs=1.5V
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0
VOLTS(V)

Subthreshold Conduction ($V_{GS} < V_t$)

(Berkeley Short Channel Insulated-Gate FET Model)

In weak inversion, the current is due mainly to diffusion between the drain and the source, similar to the current of a BJT.

$$\left| I_{D\,weak} = \frac{I_{exp} \cdot I_{limit}}{I_{exp} + I_{limit}} \right|$$

where

$$\left| I_{exp} \cong G \mu C_{ox} \frac{W}{L} (V_T)^2 e^{\frac{V_{GS} - V_t}{V_T}} \left[1 - e^{-\frac{V_{DS}}{V_T}} \right] \right|$$

and

$$\left| I_{limit} \cong \frac{q}{2} \mu C_{ox} \frac{W}{L} (V_T)^2 \right| \qquad \left(V_T = \frac{kT}{q} \right)$$

If $V_{DS} > 2V_T$ then

$$\left\| I_{exp} \cong I_{no} \frac{W}{L} e^{\frac{V_{GS} - V_t}{V_T}} \right\| \qquad (V_{GS} < V_t)$$

where

$$\left| I_{no} \cong G \mu C_{ox} (V_T)^2 \right|$$

e.g. $T = 300^\circ K$

$C_{ox} = 2.5 \times 10^{-7} \frac{F}{m^2}$ $\quad \therefore \left| I_{no} \cong 400\,nA \right|$

$\mu_n = 4 \times 10^{-2} \frac{m^2}{VS}$

$V_{GS} - V_t = -2V_T = -52\,mV$

$\frac{W}{L} = 2$ $\quad \therefore \left| I_{exp} \cong 108\,nA \right|$

Velocity Saturation (Hot electron Effect)

Recall: Current density in a homogeneous semicond. ($\nabla_r n = 0$)

$$\boxed{\bar{J} = q n \mu \bar{\mathcal{E}}}$$

Energy relaxation eq.

$$\boxed{\frac{\partial \bar{T}_c}{\partial t} = \frac{2}{3}\frac{\bar{J}\bar{\mathcal{E}}}{k_B n} - \frac{(\bar{T}_c - \bar{T}_L)}{\tau_E}}$$

$\bar{T}_c$: El. temp in C.B.
$\bar{T}_L$: Lattice temp.

where $\tau_E = \tau_{E0}\sqrt{\frac{\bar{T}_c}{\bar{T}_L}}$ Si at $T = 300°k$

If we solve for the current density $\bar{J}$ under steady-state conditions, we obtain

$$\boxed{\bar{J} = \frac{3}{2}\frac{k_B n}{\bar{\mathcal{E}}}\frac{(\bar{T}_c - \bar{T}_L)}{\tau_E} \cong \frac{3}{2}\frac{k_B n}{\bar{\mathcal{E}}}\frac{(\bar{T}_c - \bar{T}_L)}{\tau_{E0}}\sqrt{\frac{\bar{T}_L}{\bar{T}_c}}}$$

Equating this current dens. with the drift current yields

$$q n \mu \bar{\mathcal{E}} = \frac{3}{2}\frac{k_B n}{\bar{\mathcal{E}}}\frac{(\bar{T}_c - \bar{T}_L)}{\tau_{E0}}\sqrt{\frac{\bar{T}_L}{\bar{T}_c}}$$

or $$\boxed{q\mu\mathcal{E}^2 = \frac{3}{2}\frac{k_B}{\tau_{E0}}(T_c - T_L)\sqrt{\frac{T_L}{T_c}}}$$

Then the energy balance can be written as

$$\left| \frac{\partial}{\partial t}\left(\frac{3}{2}k_B T_c\right) = \frac{\bar{J}\,\bar{\mathcal{E}}}{n} - \frac{3}{2}\frac{k_B(T_c - T_L)}{\tau_E} \right|$$

or

$$\left| \frac{\partial}{\partial t}(T_c) = \frac{2}{3}\frac{\bar{J}\,\bar{\mathcal{E}}}{k_B n} - (T_c - T_L)\frac{1}{\tau_E} \right|$$

(k_B: Boltzmann const.)

This clearly shows that the temp. of the electron gas is always raised in the presence of an el. field $\bar{\mathcal{E}}$. If $T_c \gg T_L$, one calls the electrons "hot electrons".

Note: The energy relaxation time τ_E depends on the dominant scattering process and thus on the particle's (kinetic) energy → thermal energy

Example: Si at room temp. ($T_L \cong 300°K$)

Optical phonon scattering dominates

$$\tau_E \approx 4 \cdot \sqrt{\frac{T_c}{T_L}}\ [ps] \qquad T_c \geq T_L$$

Steady State: $\left(\frac{\partial}{\partial t} T_c = 0\right)$

$$\Rightarrow \left| (T_c - T_L) = \frac{2}{3}\frac{\bar{J}\,\bar{\mathcal{E}}}{k_B n}\tau_E \right|$$

Furthermore: $\bar{J} = q n \mu \bar{\mathcal{E}}$

Thus $$\left| T_c \cong T_L + \frac{2}{3}\frac{q\mu\mathcal{E}^2}{k_B}\tau_E \right|$$

$$\langle E \rangle \cong \frac{3}{2}k_B T_L + q\mu\mathcal{E}^2\tau_E$$

If we approximate the mobility by

$$\boxed{\mu = \mu_0 \sqrt{\frac{T_L}{T_c}}}$$

we obtain

$$q\mu_0 \mathcal{E}^2 = \frac{3}{2}\frac{k_B}{\tau_{E0}}(\bar{T}_c - T_L)$$

or $$\boxed{\bar{T}_c = T_L + \frac{2}{3}\frac{\tau_{E0}}{k_B} q\mu_0 \mathcal{E}^2}$$

Inserting this value of $\bar{T}_c$ into the equation for the carrier mobility yields

$$\boxed{\mu = \mu_0 \sqrt{\frac{1}{1 + \frac{2}{3}\frac{\tau_{E0}}{k_B T_L} q\mu_0 \mathcal{E}^2}}}$$

The drift current density can now be written as

$$\boxed{\bar{j} = qn\mu_0 \bar{\mathcal{E}} \sqrt{\frac{1}{1 + \frac{2}{3}\frac{\tau_{E0}}{k_B T_L} q\mu_0 \mathcal{E}^2}}}$$

As can be seen, $\bar{j}$ follows Ohm's law for small el. fields $\mathcal{E}$ (small compared to $\mathcal{E}_0 = \sqrt{\frac{3}{2}\frac{k_B T_L}{\tau_{E0} q\mu_0}} \approx \underline{3 \times 10^5 \frac{V}{m}}$).
For large fields $\bar{\mathcal{E}}$, however, $\bar{j}$ becomes a constant

$$\boxed{j_{sat} = \frac{qn\mu_0}{\sqrt{\frac{2}{3}\frac{\tau_{E0}}{k_B T_L} q\mu_0}} = n\sqrt{\frac{3 q\mu_0 k_B T_L}{2\tau_{E0}}}}$$

$$\boxed{j_{sat} = qn\mu_0 \mathcal{E}_0 = qn v_{sat}}$$

Thus, the saturated drift velocity is

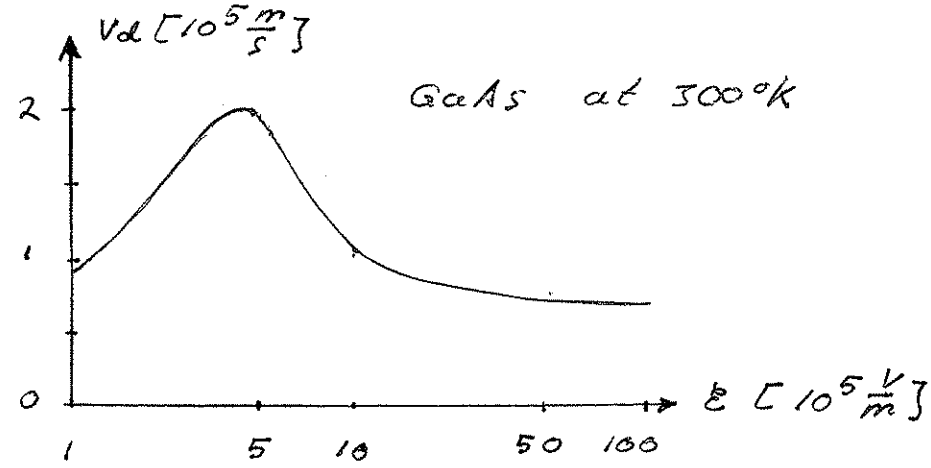
$$\left\| v_{sat} = \mu_0 \mathcal{E}_0 = \sqrt{\frac{3}{2} \mu_0 \frac{k_B T_L}{\tau_{E0} q}} \right\|$$

e.g. n-type Si° around $T = 300°K$

$\mu_0 \approx 0.1 \frac{m^2}{Vs}$

$\tau_{E0} \approx 4\,ps$ $\Rightarrow$ $\underline{\mathcal{E}_0 \approx 3 \times 10^5 \frac{V}{m}}$; $\underline{v_{sat} \approx 3 \times 10^4 \frac{m}{s}}$

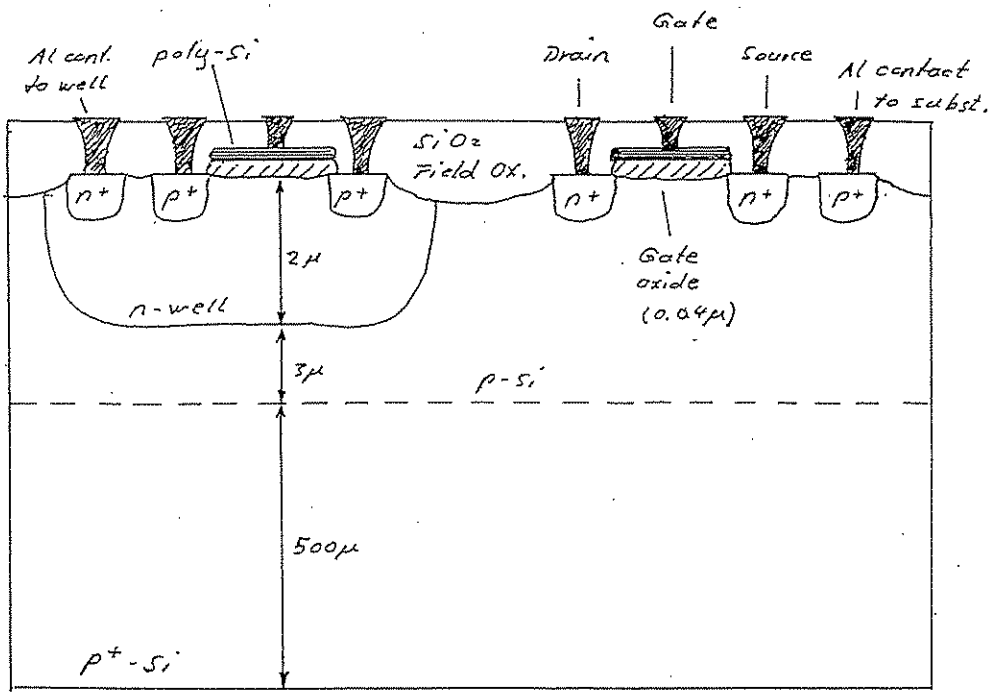
Note The above described velocity saturation does apply to many semiconductors (e.g. Si). However, some III-V compounds such as GaAs exhibit a different drift-velocity vs. el. field characteristic. As the field is increased, more and more electrons from the Γ minimum are excited to the X and L valleys where they have a much larger effective mass. Thus, their mobility is much smaller compared to the Γ mobility and the average carrier velocity drops. This behavior is the basis for the Gunn effect.



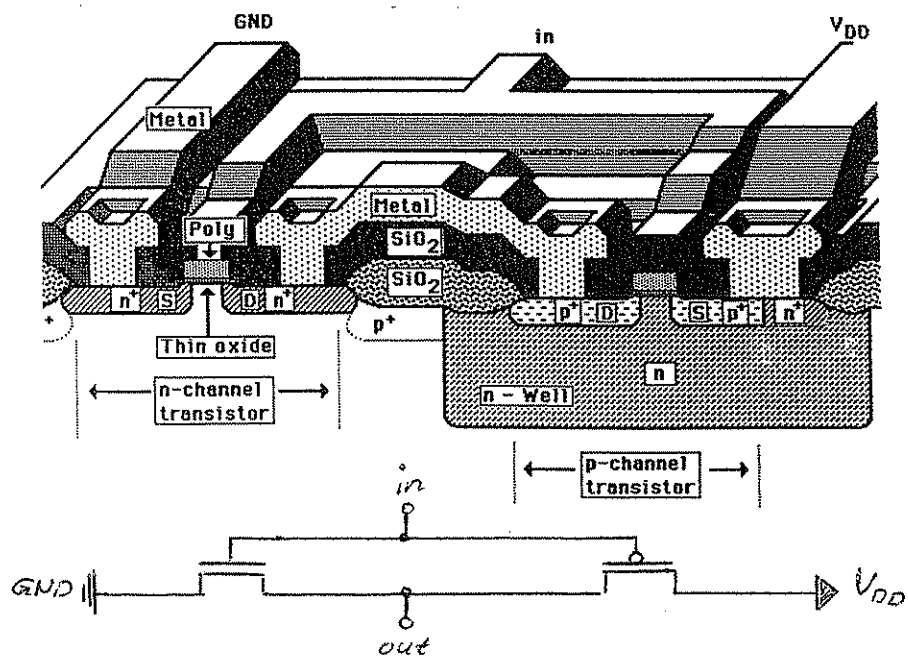
7.4 The CMOS Process

Implementation of a CMOS Inverter

Wafer Cross Section

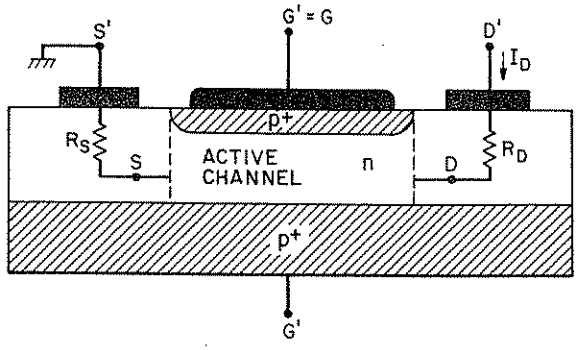


3-D Perspective

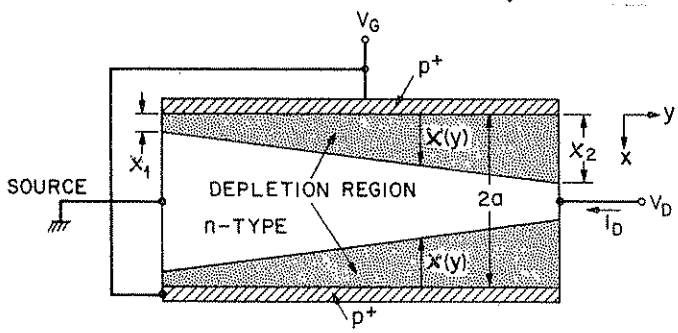


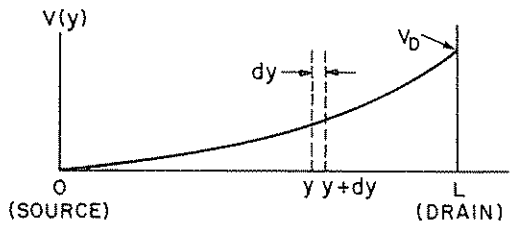
7.5 Junction Field Effect Transistors (JFETs)

In a JFET, the channel charge is controlled by altering the depletion width of a pn junction via a bias voltage (gate voltage).



The current in the above figure flows through an n-type channel between two p^+ regions. A reverse bias between these p^+ regions and the channel causes the depletion region to widen and thus restrict the effective width of the channel.
The figure below illustrates the drain voltage variation along the channel before the onset of pinch-off at the drain end.





Note:
This picture illustrates a "normally-on" or depletion mode device i.e. the channel is on if $V_G = 0$

Current-Voltage Characteristics

The voltage drop across an elemental section dy of the channel (before pinch-off) is given by

$$dV(y) = I_d\, dR = \frac{I_d\, dy}{2q\mu_n N_D \cdot W[a - x(y)]} \qquad (1)$$

where

$$x(y) = \sqrt{\frac{2\varepsilon_s[-V(y) + V_G + V_{bi}]}{qN_D}} \qquad (2)$$

Since I_d is a constant independent of y, we can rewrite eq. (1) as

$$I_d\, dy = 2q\mu_n N_D W[a - x(y)]\, dV \qquad (3)$$

From eq. (2) we obtain

$$\frac{d}{dy}x^2 = 2x\frac{dx}{dy} = \frac{2\varepsilon_s}{qN_D}\frac{dV}{dy} \qquad (4)$$

or

$$dV = x\, dx\, \frac{qN_D}{\varepsilon_s} \qquad (4a)$$

Inserting (4a) into (3) and integrating over the channel length L yields

$$I_d \cdot L = q\mu_n N_D W \frac{qN_D}{\varepsilon_s}\left[a(x_2^2 - x_1^2) - \frac{2}{3}(x_2^3 - x_1^3)\right] \qquad (5)$$

or

$$I_d = \frac{W}{L}\mu_n \frac{q^2 N_D^2}{\varepsilon_s}\left[a(x_2^2 - x_1^2) - \frac{2}{3}(x_2^3 - x_1^3)\right] \qquad (5a)$$

Finally, we can use eq. (2) again to obtain

equivalent expressions for x_1 and x_2, respectively. The current/voltage relationship can therefore be written as

$$\boxed{I_d = I_p \left[\frac{V_d}{V_p} - \frac{2}{3}\left(\frac{V_d + V_a + V_{bi}}{V_p}\right)^{3/2} + \frac{2}{3}\left(\frac{V_a + V_{bi}}{V_p}\right)^{3/2} \right]} \qquad (6)$$

where

$$\boxed{I_p = \frac{W}{L}\mu_n \frac{q^2 N_D^2 a^3}{\varepsilon_s}} \qquad (6a)$$

$$\boxed{V_p = \frac{1}{2}\frac{q N_D a^2}{\varepsilon_s}} \quad \text{pinch-off voltage} \qquad (6b)$$

Note: At pinch-off $x_2 = a$

$$\boxed{I_{d_{sat}} = I_p \left[\frac{1}{3} - \left(\frac{V_a + V_{bi}}{V_p}\right) + \frac{2}{3}\left(\frac{V_a + V_{bi}}{V_p}\right)^{3/2} \right]} \qquad (7)$$

The corresponding saturation voltage is

$$\boxed{V_{dsat} = V_p - V_a - V_{bi}} \qquad (8)$$

In the ideal case, the current I_d remains constant if V_d exceeds V_{dsat}. In practicality, increasing V_d above V_{dsat} widens the pinch-off region and so reduces the ohmic portion of the channel. The result is a small increase of the current I_d in the saturation region, similar to what we have observed for the MOSFET (cf. channel length modulation).

7.6 Metal Semiconductor FETs (MESFETs)

The MESFET was proposed in 1966. Its operation is identical to a JFET. The major difference is that the MESFET uses a rectifying metal-semiconductor contact instead of the pn junction.

Practical MESFETs are fabricated by growing an epitaxial layer on a semi-insulating substrate to minimize parasitic capacitances. Most MESFETs are made of n-type III-V compound semiconductors, such as GaAs. The high electron mobilities of these materials minimize resistances and the high (electron) saturation velocities yield high cutoff frequencies.

To maximize (switching) speed and minimize power dissipation, MESFETs are typically fabricated as "normally-off" or enhancement mode devices. In such a structure, the built-in potential V_{bi} of the gate junction is sufficient to deplete the channel. To achieve this property, a GaAs MESFET employs a very thin epitaxial layer on a semi-insulating substrate.

The minimum (positive) bias voltage causing a current flow through the channel is called the threshold voltage V_t and is given by

$$\boxed{V_t = V_{bi} - V_p} \qquad (9)$$

where V_p denotes the pinch-off voltage defined by eq. (6b). Near the threshold, the drain current in the saturation region, I_{dsat}, can be obtained by substituting V_{bi} in eq. (7) by the equivalent expression derived from (9), i.e., $V_{bi} = V_t + V_p$. Since eq. (7) describes a depletion mode device, we have to replace V_G by $-V_G$. Furthermore, the MESFET only features an upper channel, which halves the current flow. We thus obtain

$$\boxed{I_{dsat} = \frac{1}{2} I_P \left[\frac{1}{3} - \left(1 - \frac{V_G - V_t}{V_P}\right) + \frac{2}{3}\left(1 - \frac{V_G - V_t}{V_P}\right)^{3/2} \right]} \qquad (10)$$

Finally by assuming $(V_G - V_t) \ll V_p$, the saturation current can be approximated by

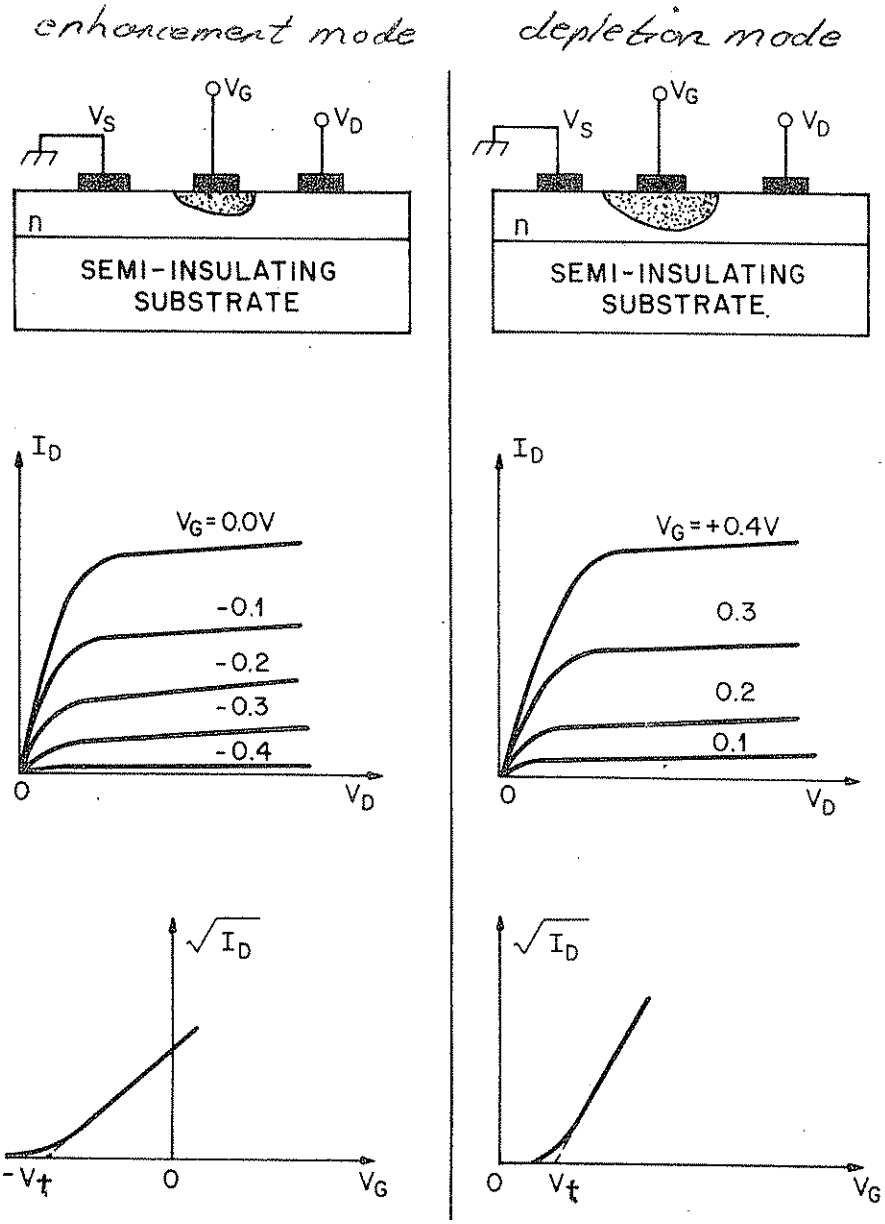
$$\boxed{I_{dsat} \cong \frac{1}{2} \frac{I_P}{4 V_P^2} (V_G - V_t)^2} \qquad (11)$$

with $\frac{I_P}{4 V_P^2} = \frac{W}{L} \mu_n \frac{\varepsilon_s}{a}$, we can rewrite (11) as

$$\boxed{I_{dsat} \cong \frac{1}{2} \frac{W}{L} \mu_n \frac{\varepsilon_s}{a} (V_G - V_t)^2} \qquad (11a)$$

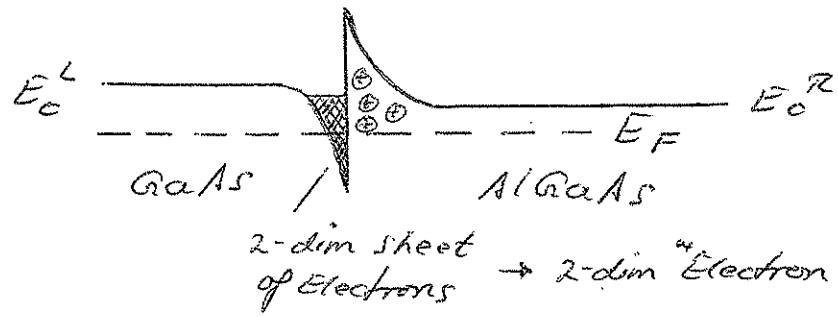
Implementation of MESFETs

The figure below illustrates the basic configuration for an enhancement mode and a depletion mode MESFET, respectively. MESFETs are typically realized on a semi-insulating GaAs substrate.



7.7 High Electron Mobility Transistors (HEMTs)

The channel conductivity (and the switching speed) of an FET can be enhanced by increasing the channel doping and thus the carrier concentration. Unfortunately, higher channel doping increases scattering, which leads to a degradation of carrier mobility.
A clever solution to this dilemma is to grow a thin undoped well (e.g. in GaAs) bounded by a wider band gap, higher doped material (e.g. AlGaAs), as described in the previous chapter (pp. VI-24).



This process, called modulation doping, separates the "trapped" electrons from their parent donors and allows them to travel relatively unimpeded in the 2-dimensional well formed by the hetero-junction. The gain in mobility is most pronounced at low temperature, where lattice scattering is low (and impurity scattering dominates).