Lecture 16. Electric Field Effect and FETs. Part I

- Electric Field Effect
- MOSFETs: band structure, depletion and inversion
- Long-channel MOSFETs

The Electric Field Effect

\[ en_{2D} = \varepsilon \varepsilon_0 E \]

\[ n_{2D} = \frac{1}{e} \frac{4\pi \varepsilon \varepsilon_0}{4\pi} E \sim (10^8 - 10^9)E \]

Maximum \( E \) is limited by the dielectric breakdown

\( \sim 10^9 V/m \) for SiO\(_2\), up to \( 10^{10} V/m \) for ionic gel gating. For comparison, \( \frac{e}{4\pi \varepsilon_0 a_B^2} \sim 10^{11} V/m \)

\[ n_{2D}^{max} \sim 10^{19} \text{ } m^{-2} \text{ (conventional dielectrics)} \]

\[ 10^{20} \text{ } m^{-2} \text{ (ion gel).} \]

Electron concentration in a 1-nm-thick copper:

\[ 10^{29} m^{-3} \times 10^{-9} m = 10^{20} m^{-2} \]

- the field effect is insignificant
  (the surface layer is “shunted” by the bulk).

The situation is quite different in semi-metals and semiconductors.

Limitation of the ionic gel technique:
- cannot vary \( n_{2D} \) in-situ below the gel freezing \( T \sim 200 K \);
- polarization of the ionic gel is a slow (\( \sim 1s \)) process.
The Field Effect (cont’d)

General Idea:
- to fabricate an artificial 2D system at the interface between the gate dielectric and a semiconductor/insulator with the charge density controlled by the gate;
- to study the charge transport in the system.

New states of matter can be explored with the field effect.


Carrier Concentration in FETs

If the drain-source voltage $V_{DS}$ is small in comparison with the gate-source voltage $V_{GS}$, the charge carrier concentration is uniform along the channel.

Provided all $V_{GS}$ is applied btw the gate and the channel:

$$n_{2D} = \frac{C_{di}(V_{GS} - V_T)}{e}$$

$V_T$ - the threshold voltage

$C_{di}$ is the gate dielectric capacitance per unit area.

More accurately, taking into account the voltage divider $C_{di}, C_{sc}$:

$$n_{2D} = \frac{C_{di}C_{sc}}{C_{di} + C_{sc}} \frac{(V_{GS} - V_T)}{e}$$

$C_{SC}$ - the semiconductor space-charge capacitance

$V_T$ is proportional to the density of charge traps at the Si-$SiO_2$ interface. There are no mobile charges in the channel unless $V_{GS}$ exceeds $V_T$. The early history of MOSFET development was all about the struggle to reduce $V_T$.

Why $Si$? Because the processes of elimination of charge traps in the bulk of thermally grown $SiO_2$ and at the the $Si - SiO_2$ interface have been developed (at Bell Labs, around 1960).
Early History of Si MOSFETs

Julius Lilienfeld proposed FET, but never constructed a working device...

Julius E. Lilienfeld (1882 –1963)

Martin M. (John) Atalla (1924 – 2009)

Dawon Kahng (1931 – 1992)


D. Kahng, M. Atalla (Bell)
S. Hofstein, F. Heiman (RCA Sarnoff)
created practical Si FETs

Shockley and Brattain attempted to make FET, but failed...

Shockley

Bardeen

Fred Heiman

Steven Hofstein

attributed the failure of Shockley’s efforts to the surface states

Michael Riordan & Lillian Hoddeson. “Crystal Fire: The Invention of the Transistor and the Birth of the Information Age”.

Oskar Heil
- Electric Field Effect
- MOSFETs: band structure, depletion and inversion
- Long-channel MOSFETs
We’ll consider just one type of FETs (an \( n \)-type \( Si \)-based MOSFET). Many more types of FETs exist, both \( n \)- and \( p \)-type, with inversion and accumulation channels, “normally-on” and “normally-off”, inorganic and organic, etc.
Let’s first consider the case when the gate and the semiconductor are not connected to an external circuit.

Chemical potentials are not aligned yet.
The gate and the semiconductor are electrically isolated.

The gate and the semiconductor are connected through an external circuit $V_{GS} = 0$.

To line up chemical potentials, charges are redistributed through an external circuit. As the result (a) the bands bend near the $Si – SiO_2$ interface, (b) the region near the interface is **depleted of holes**, primary carriers in the $p$-doped $Si$.

The total built-in voltage $V_{bi} = \frac{1}{e} |\Phi_G - \Phi_S|$ where $\Phi_G$ and $\Phi_S$ are the work functions of the gate and semiconductor, respectively. This voltage drops across the oxide and the depletion region.
nMOSFET band diagram (from Gate to Channel) $V_{GS} \neq 0$

Minority carriers (electrons) form an inversion layer at the $Si - SiO_2$ interface. The layer is isolated from the bulk by the depletion layer.
Figure 1.5: Normalized total semiconductor charge per unit area versus normalized surface potential for p-type Si with $N_a = 10^{16}/\text{cm}^3$, $Q_{th} = (2e q N_a V_{th})^{1/2} \approx 9.3 \times 10^{-5} \text{ C/cm}^2$ and $V_{th} \approx 0.026 \text{ V}$ at $T = 300 \text{ K}$. The arrows indicate flat-band condition and onset of strong inversion.

$C_{dl}$ - the gate dielectric capacitance

$C_{SC}$ - the semiconductor space-charge capacitance

$C_{it}$ - the interface trap capacitance
MOSFET: degenerate or non-degenerate?

Operation at 300K – we need to compare $k_B T$ with $E_F(n_{2D})$.

$$n_{2D} = \frac{\varepsilon \varepsilon_0}{e t_{ox}} V = (\varepsilon = 4, t_{ox} = 2nm, V = 0.8V) = 8 \times 10^{16} m^{-2}$$

$$E_F(n_{2D}) = \frac{\hbar^2}{m^*} \pi n_{2D} \quad m^* = 0.2 m_e$$

$$n_{2D} = 10^{16} m^{-2} \quad E_F(n_{2D}) \approx 100K \quad \text{- non-degenerate Fermi gas}$$

$$n_{2D} = 10^{17} m^{-2} \quad E_F(n_{2D}) \approx 1000K \quad \text{- (almost) degenerate Fermi gas}$$

- the experiments in Lecture 12 which demonstrated a sharp occupancy step at the Fermi energy were performed at low $T \ll E_F$. 
- Electric Field Effect
- MOSFETs: band structure, depletion and inversion
- Long-channel MOSFETs
MOSFET Current-Voltage Characteristics and Modes of Operation

Analog applications
MOSFETS as amplifiers

Digital applications
MOSFETS as switches

\[ V_{DS} = V_{GS} - V_T \]

Subthreshold regime
\[ V_{GS} < V_T \]

“Linear” regime
Saturation regime
The potential barrier for electrons traveling from source to drain is controlled by the gate and drain voltages.
Effect of $T$: Subthreshold Swing

Provided all $V_{GS}$ is applied btw the gate and the channel (which is not 100% true because of the divider $C_{di}/C_{sc}$):

$$n(V_{GS}) = N_C e^{-\frac{E_C-E_{F,\text{equil}}-eV_{GS}}{k_B T}}$$

- concentration of “non-degenerate” electrons in the channel.

**Subthreshold swing $S_{s-th}$**: voltage difference that corresponds to a 10-times drop of $I_{DS}$ (or, the same, $n$) in the subthreshold regime.

$$\frac{n_1(V_{GS,1})}{n_2(V_{GS,2})} = e^{-\frac{e(V_{GS,1}-V_{GS,2})}{k_B T}}$$

$$S_{s-th}(300K) = \ln(10) \frac{k_B T}{e} \approx 60 \text{ mV}$$

**Conclusion**: at least ~0.5V is required at 300K to go from “off” to “on” – imposes the lower limit on energy dissipation!
MOSFET band diagram (from Source to Drain) $V_{DS} \neq 0$

- $E_C$ along the channel
- $V_{GS} = const$
- $E_{FS}$
- $V_{DS}$
- $E_{FD}$
Long-channel MOSFET current-voltage characteristics

Long channel: negligible non-uniformity of fields/concentration near source and drain

\[ L \gg L^* = \sqrt{t_{ch} t_{di} (\epsilon_{ch}/\epsilon_{di})} \]

We need to qualitatively understand the transistor operation in the linear and saturation regimes.
The Linear Regime

- describes the MOSFET behavior for small $V_{DS}$, where the MOSFET acts as a variable resistor. $\mu, n_{2D}, E_{DS}$ are considered constant along the channel.

\[
E_{DS} = \frac{V_{DS}}{L} \quad n_{2D} = \frac{C_{di}(V_{GS} - V_T)}{-e} \quad |V_{DS}| \ll (V_{GS} - V_T)
\]

$C_{di}$ is the gate dielectric capacitance per unit area.

\[
j_{DS} = (-en_{2D}\mu)E_{DS} = -e\frac{C_{di}(V_{GS} - V_T)}{-e} \frac{V_{DS}}{\sigma_{2D}}\mu \frac{V_{DS}}{L}
\]

\[
I_{DS} = Wj_{DS} = \mu C_{di}(V_{GS} - V_T) \frac{W}{L} V_{DS}
\]

To increase $I_{DS}$, we need higher mobility and lower $V_T$. Greater $C_{di}$ (i.e. larger $n_{2D}$) also helps, though the $RC$ time constant increases (not good, lower operation frequencies).
The Drain Current in the Saturation Regime

For larger $V_{DS}$ comparable with $V_{GS} - V_T$ we have to take into account variations of $n_{2D}$ along the channel. The guiding principle – the continuity equation, $j_{DS}$ is the same in any cross section.

$$j_{DS}(x) = (-e n_{2D} \mu) E(x) = \mu C_{di} (V_{GS} - V_T - V(x)) \frac{dV(x)}{dx}$$

$$j_{DS}(x) = j_{DS} = \text{const}$$

$$\int_0^L j_{DS} \, dx = \mu C_{di} \int_0^{V_{DS}} (V_{GS} - V_T - V(x)) \frac{dV(x)}{dx} \, dx$$

$$j_{DS} L = \mu C_{di} \left[ (V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2 \right]$$

$$I_{DS} = \mu C_{di} \frac{W}{L} \left[ (V_{GS} - V_T) V_{DS} - \frac{1}{2} V_{DS}^2 \right]$$

$I_{SD}$ increases linearly for small $V_{DS}$, but then reaches a maximum when $n_{2D}$ becomes zero at the drain end of the channel. Holes are separated from the channel by the depletion layer and cannot contribute to $I_{SD}$. $I_{SD}$ saturates and the depletion layer near the drain end of the channel accommodates the “excessive” $V_{DS}$.

$$I_{DS,\text{sat}} = \mu C_{di} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2} \quad V_{DS} > V_{GS} - V_T$$
In the pinch-off regime, $E$ at the source is independent of $V_{DS}$. $I_{DS}$ also becomes independent of $V_{DS}$ (continuity equation).

$IV$ characteristics of an ideal FET.

**OFF state** - $V_{GS} < V_T$.

**ON state** - $V_{DS} > V_{DS}^{sat}$, saturation regime.

Saturation regime is of primary importance for digital applications, where the output current of a FET should only depend on the input (gate) voltage and not on $V_{DS}$. 
Conclusion: at least $\sim 0.5V$ is required at $300K$ to go from “off” to “on” – imposes the lower limit on energy dissipation!

Transconductance characteristic of MOSFET ($V_{GS} = 2V$)

For non-ideal MOSFETs, the subthreshold swing is greater than $0.5V$. Reduction of $S_{S-th}$ is important for minimization of Joule heating.

Why does fast “on-off” switching generate heat?
Energy Loss in Charging a Capacitor

Energy conservation:

\[ \mathcal{E}i = R i^2 + i V_{bc} \]

Energy dissipated in the resistor:

\[ \int_{0}^{\infty} R i^2(t) dt = \int_{0}^{\infty} \frac{\mathcal{E}^2}{R} e^{-2t/\tau} dt = \frac{\mathcal{E}^2}{R} \left(-\frac{\tau}{2}\right) \left( e^{-\infty/\tau} - e^{-0/\tau} \right) = \frac{\mathcal{E}^2 C}{2} \]

Half of the work done by the battery is wasted as heat (doesn’t depend on \( R \)).

Solution: slowly (in comparison with \( \tau \)) ramp up the emf in the process of charging.
Desirable FET Characteristics

- Small $V_T$ and $V_{ON} - V_{OFF}$ (small subthreshold swing) $\Rightarrow$ small energy dissipation

- ON/OFF current ratio $> 10^4$ $\Rightarrow$ heating minimization in the OFF regime, small energy dissipation

- Well-defined saturation regime
FET Characteristics (cont’d)

- **Fast switching**

\[ \tau \sim R_{tot} C_{eff} = (R_{cont} + R_{ch}) \frac{C_{di} C_{sc}}{C_{di} + C_{sc}} \]

- **Small physical dimensions**

The energy required to switch the transistor between “on” and “off” regimes is \( \frac{1}{2} C_{GS} V_{GS}^2 \).

When a transistor is “recharged”, half of this energy is dissipated as heat!

- **Absence of gate leakage** \( \Rightarrow \) reduction of static power consumption

On the one hand, reduction of all dimensions requires reduction of the gate dielectric thickness (to reduce parasitic capacitances). On the other hand, the gate dielectric cannot be too thin, otherwise the current flows between the gate and the channel due to quantum tunneling \( \Rightarrow \) excessive energy dissipation. Use of high-\( k \) dielectrics (e.g., \( HfO_2 \)) helps to increase the oxide thickness (without decreasing \( n_{2D} \)) and to block tunneling.
Summary

- Electric Field Effect – powerful tool to explore new phases of matter
- Long-channel MOSFETs: linear and saturation regimes

Next lecture:
- Nanoscale MOSFETs: short-channel effects and contacts
- Future progress of electronics, Moore’s Law – quantum limitations

Samsung Exynos 7420 Application Processor