CMOS

Digital Integrated Circuits

Analysis and Design

Chapter 3

MOS Transistor

The Metal Oxide Semiconductor (MOS) structure

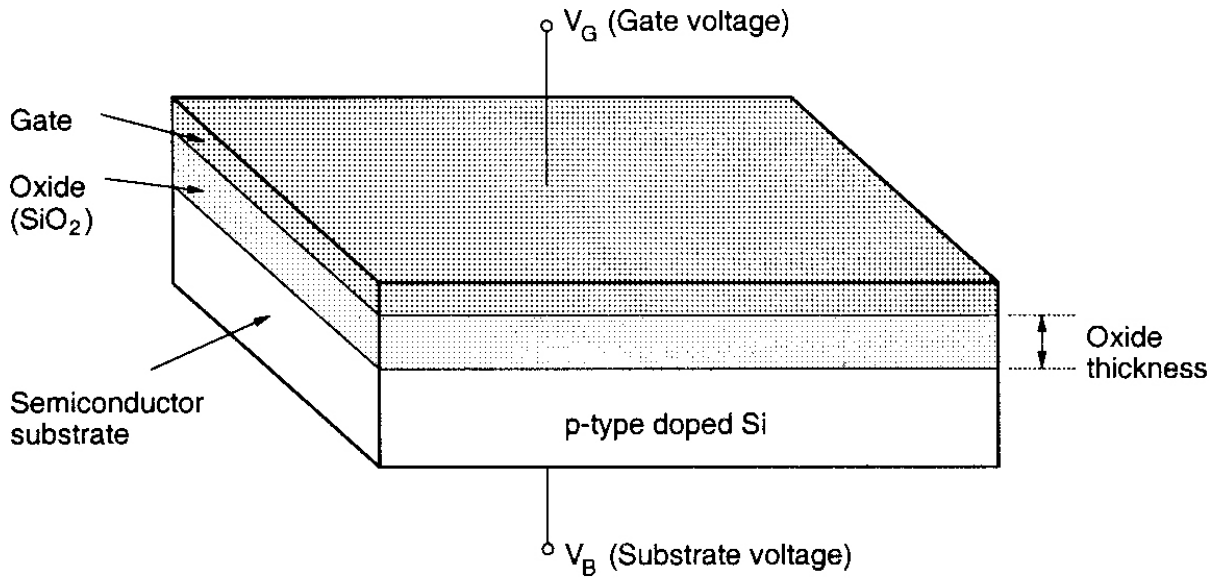


Figure 3.1 Two-terminal MOS structure.

- The structure consists of three layer
  - The metal gate electrode
  - The insulating oxide (SiO2) layer
  - The p-type bulk semiconductor

- The basic properties of the semiconductor

$$\textit{The mass action law: } n \cdot p = n_i^2$$

$$\textit{Assume the substrate doping concentration } N_A$$

$$\textit{then } p_{n0} \cong \frac{n_i^2}{N_A}, \quad p_{p0} \cong N_A$$

Energy band diagram of a p-type silicon substrate

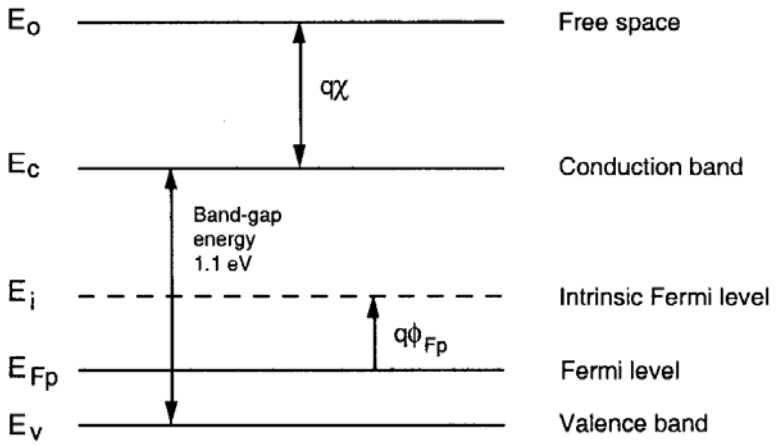


Figure 3.2 Energy band diagram of a p-type silicon substrate.

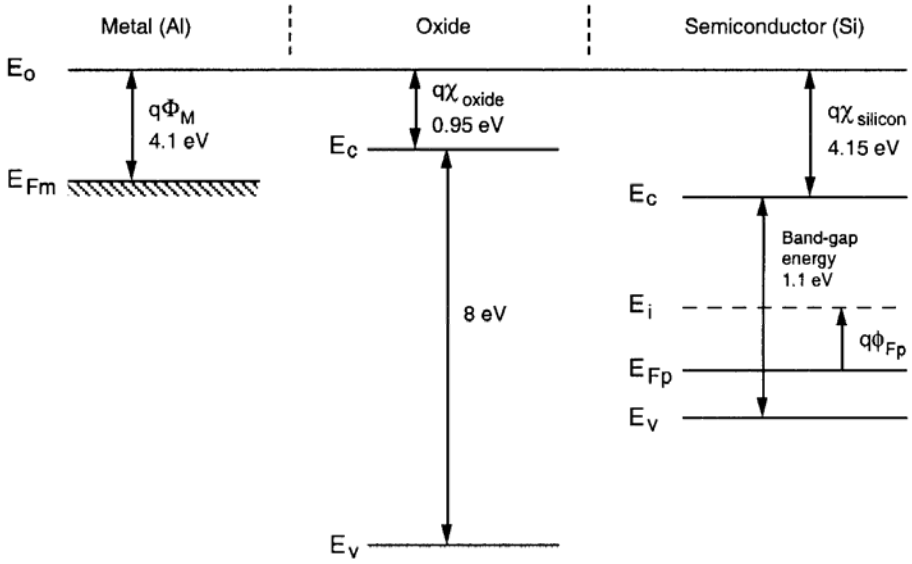


Figure 3.3 Energy band diagrams of the components that make up the MOS system.

The Fermi potential $\varphi_F = \dfrac{E_F\text{-}E_i}{q}$

For a p-type semiconductor, $\varphi_{Fp} = \dfrac{kT}{q}\ln\dfrac{n_i}{N_A}$

For a n-type semiconductor, $\varphi_{Fn} = \dfrac{kT}{q}\ln\dfrac{N_D}{n_i}$

The energy required for an electron to move from the Fermi level int*o free space is called the work function*

$q\varphi_s = q\chi + (E_c\text{-}E_F)$

Energy diagram of the combined MOS system

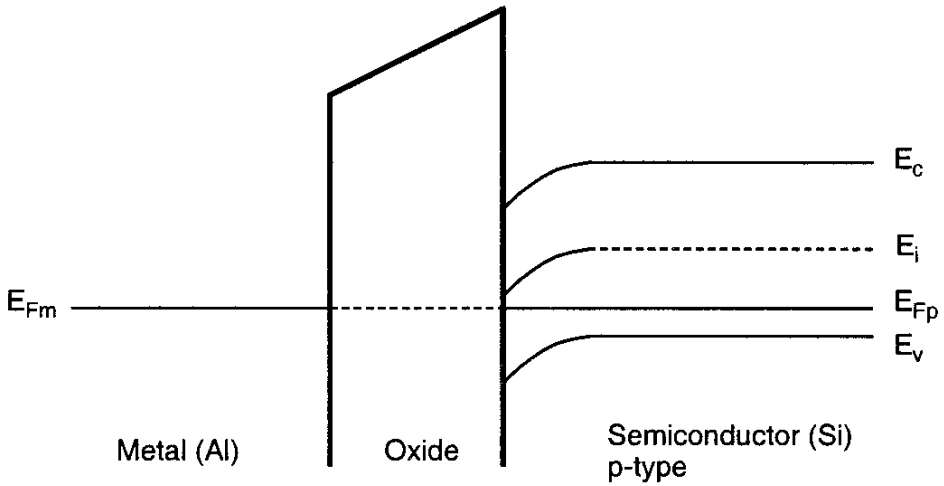


Figure 3.4 Energy band diagram of the combined MOS system.

- The equilibrium Fermi levels of the semiconductor (Si) substrate and the metal gate are at the same potential
- The bulk Fermi level is not significantly affected by the bending
- The surface Fermi level moves closer to the intrinsic Fermi level

Example 1

Consider the MOS structure that consists of a p-type doped silicon substrate, a silicon dioxide layer, and a metal (aluminum) gate. The equilibrium Fermi potential of the doped silicon substrate is given as $q\phi_{F_p} = 0.2$ eV. Using the electron affinity for silicon and the work function for aluminum given in Fig. 3.3, calculate the built-in potential difference across the MOS system. Assume that the MOS system contains no other charges in the oxide or on the silicon-oxide interface.

First, we have to calculate the work function for the doped silicon, which is given by (3.6). Since the electron affinity of silicon is 4.15 eV, the work function $q\Phi_S$ is found as

$$q\Phi_S = 4.15 \text{ eV} + 0.75 \text{ eV} = 4.9 \text{ eV}$$

Now calculate the work function difference between the silicon substrate and the aluminum gate. Note that the work function of aluminum is given as 4.1 eV in Fig. 3.3. Thus, the built-in potential difference across this MOS system is

$$q\Phi_M - q\Phi_S = 4.1 \text{ eV} - 4.9 \text{ eV} = -0.8 \text{ eV}$$

If a voltage corresponding to this potential difference is applied externally between the gate and the substrate, the bending of the energy bands near the surface can be compensated; i.e., the energy bands become "flat." Thus, the voltage defined by

$$V_{FB} = \Phi_M - \Phi_S$$

is called the *flat-band* voltage.

The MOS System under External Bias - accumulation

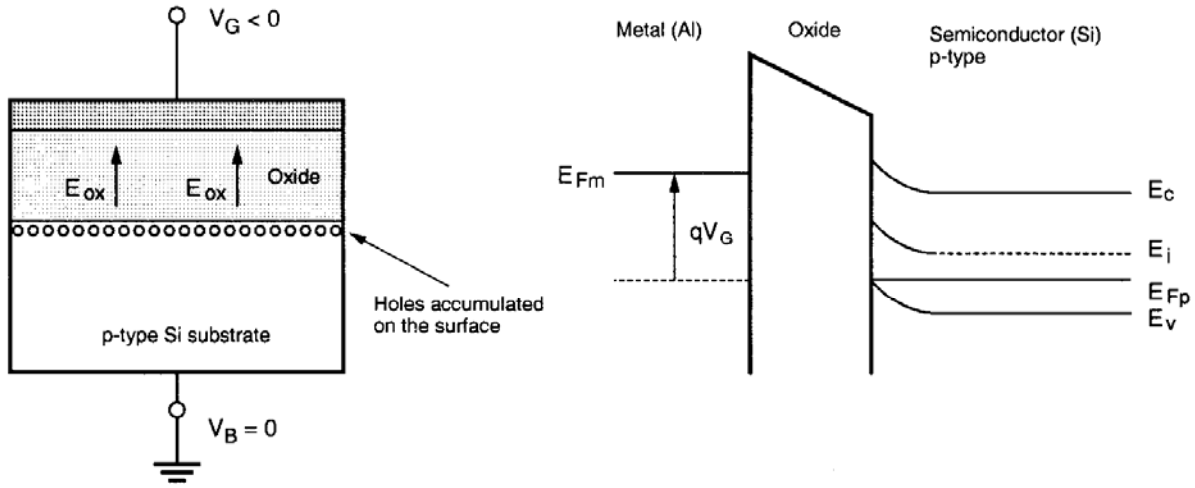


Figure 3.5 The cross-sectional view and the energy band diagram of the MOS structure operating in accumulation region.

- A negative voltage V_G is applied to the gate electrode.
  - The holes in the p-type substrate are attracted to the semiconductor-oxide surface
  - The majority carrier concentration > the equilibrium hole concentration
    - The electron concentration (minority carrier) decreases as the negatively charged electron are pushed deeper into the substrate
  - The oxide electric field is directed towards the gate electrode
  - Causing the energy bands bend up-ward near the surface

The MOS System under External Bias – depletion

- A small positive gate bias V_G is applied to the gate electrode
  - The oxide electric field will be directed towards the substrate
  - Causing the energy bands to bend downward near the surface
  - The majority carrier (hole) will be repelled backed into the substrate
    - Leaving negatively charged fixed acceptor ions behind (depletion region)

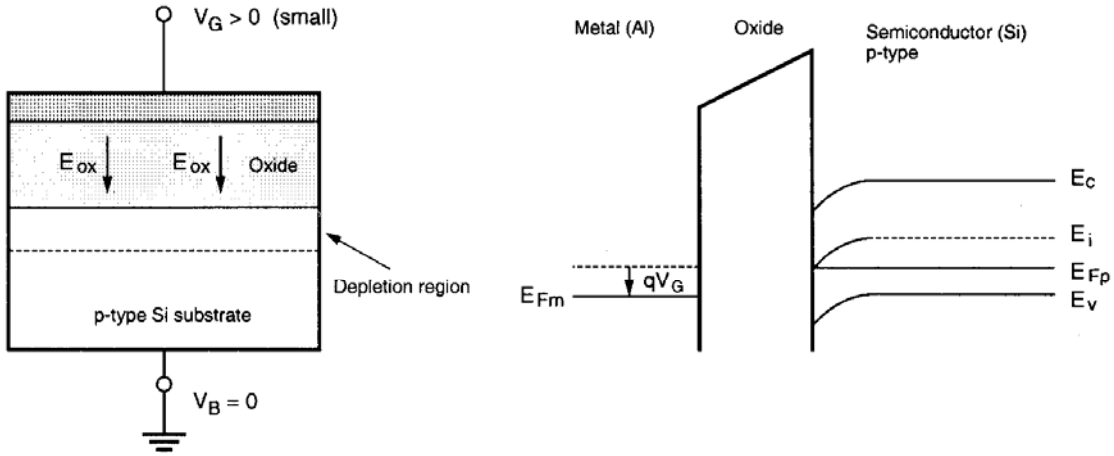


Figure 3.6 The cross-sectional view and the energy band diagram of the MOS structure operating in depletion mode, under small gate bias.

$$dQ = -q \cdot N_A \cdot dx$$

$$d\phi_s = -x \cdot \frac{dQ}{dx} = \frac{q \cdot N_A \cdot x}{\varepsilon_{Si}} dx$$

$$\int_{\phi_F}^{\phi_s} d\phi_s = \int_0^{x_d} \frac{q \cdot N_A \cdot x}{\varepsilon_{Si}} dx$$

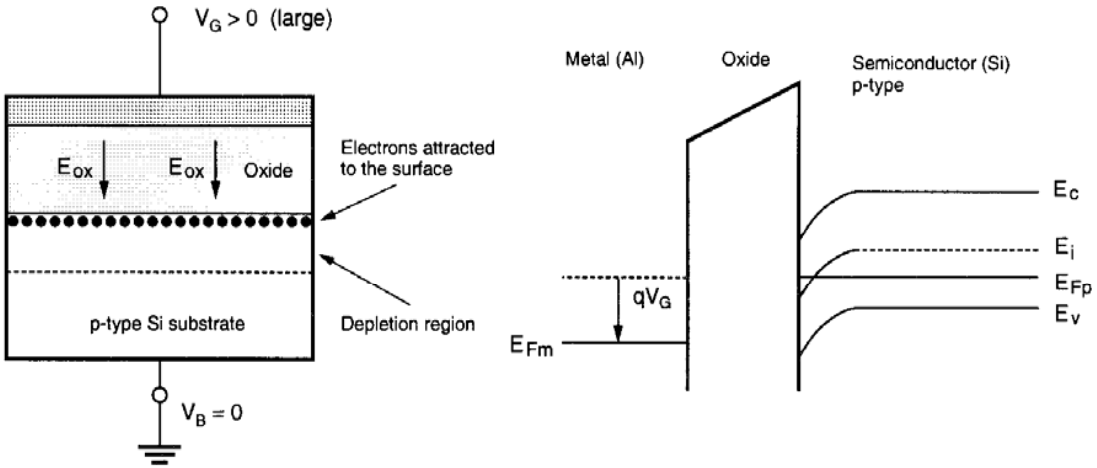
$$\phi_s - \phi_F = \frac{q \cdot N_A \cdot x_d^2}{\varepsilon_{Si}}$$

$$x_d = \sqrt{\frac{2\varepsilon_{Si} \cdot |\phi_s - \phi_F|}{q \cdot N_A}}$$

$$Q = -q \cdot N_A \cdot x_d = -\sqrt{2q \cdot N_A \cdot \varepsilon_{Si} \cdot |\phi_s - \phi_F|}$$

The MOS System under External Bias – inversion

- A further increase in the positive gate bias
  - Increasing surface potential ⇨the downward bending of the energy bands will increase
  - The mid-gap energy level E_i becomes smaller than the Fermi level E_{Fp} on the surface
    - The substrate semiconductor in this region become n-type
    - The electron density is larger than the majority hole density
    - Inversion layer, surface inversion
    - Can be utilized for conducting current between two terminal of the MOS transistor
  - The surface is said to be *inverted*
    - *The density of mobile electrons on the surface becomes equal to the density of holes in the bulk substrate*
    - Requiring the surface potential has the same magnitude, but the reverse polarity, as the bulk Fermi potential ϕ_F
    - Further increase gate voltage ⇨ electron concentration↑ ⇨ but not to an increase of the depletion depth



$$x_{dm} = \sqrt{\frac{2 \cdot \varepsilon_{Si} \cdot |2\phi_F|}{q \cdot N_A}}$$

Figure 3.7 The cross-sectional view and the energy band diagram of the MOS structure in surface inversion, under larger gate bias voltage.

The physical structure of a n-channel enhancement-type MOSFET

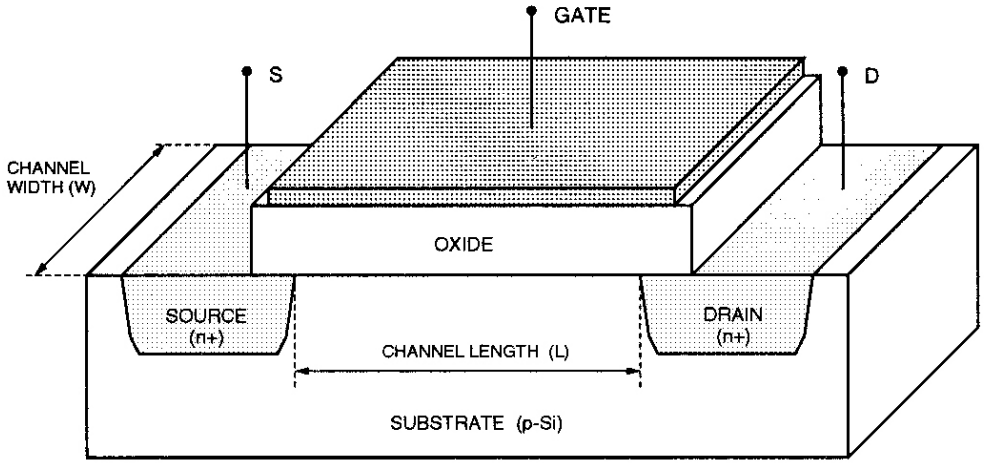


Figure 3.8 The physical structure of an n-channel enhancement-type MOSFET.

- MOS structure
  - polysilicon gate, thin oxide layer, semiconductor
- Source, drain n^+-region
  - The current conducting terminals of the device
    - Conducting channel, channel length L, channel width W
  - The device structure is completely symmetrical with respect to the drain and source
- The simple operation of this device
  - *Controlling the current conduction between the source and the drain, using the electric field generated by the gate voltage as a control variable*

Circuit symbols for enhancement-type MOSFET

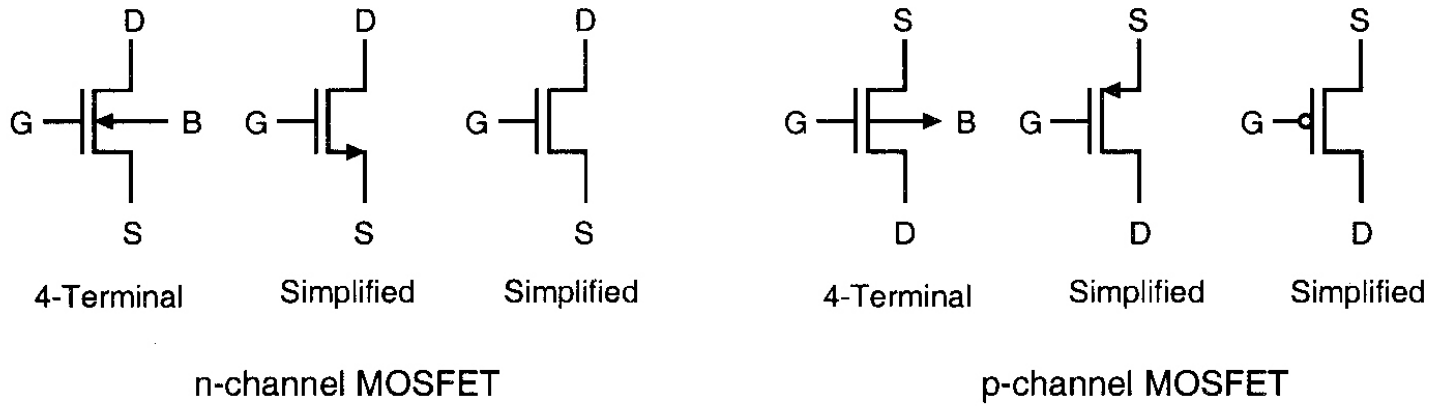


Figure 3.9 Circuit symbols for n-channel and p-channel enhancement-type MOSFETs.

- Enhancement-mode MOSFET
  - No conducting region at zero gate bias
- Depletion-mode MOSFET
  - A conducting channel already exists at zero gate bias
- The abbreviations used for device terminals are
  - G for the gate, D for the drain, S for the source, and B for the substrate
- The small arrow always marks the source terminal

Formation of a depletion region

- For small gate voltage level
  - The majority carriers (holes) are repelled back into the substrate
  - The surface of the p-type substrate is depleted
  - Current conduction between S and D is not possible

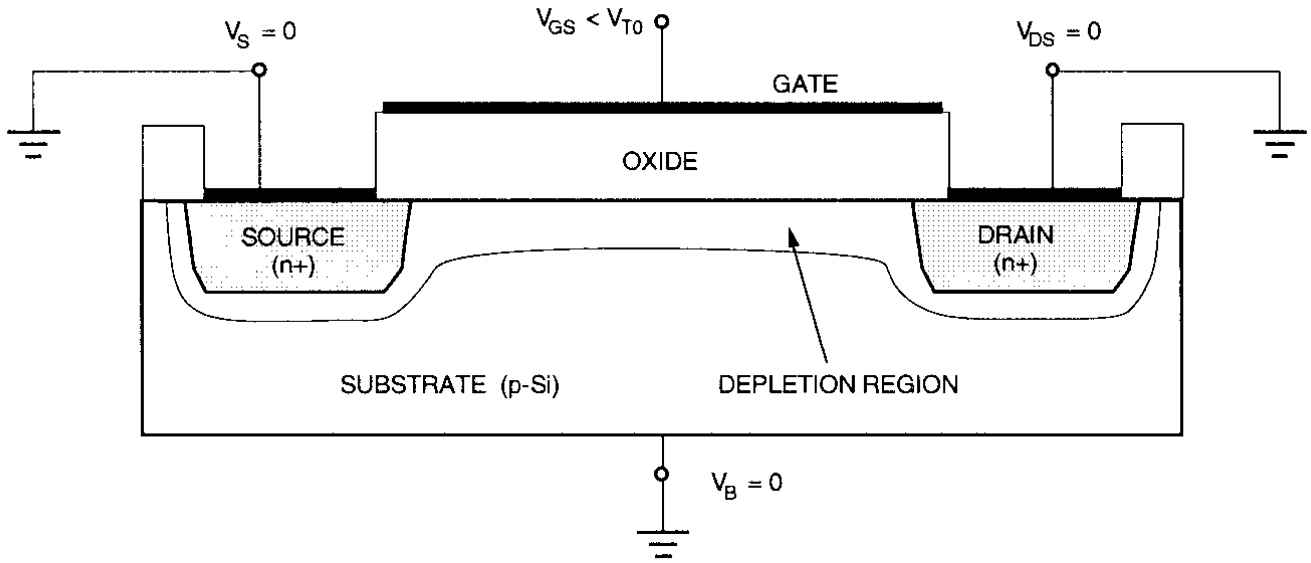


Figure 3.10 Formation of a depletion region in an n-channel enhancement-type MOSFET.

Formation of an inversion layer

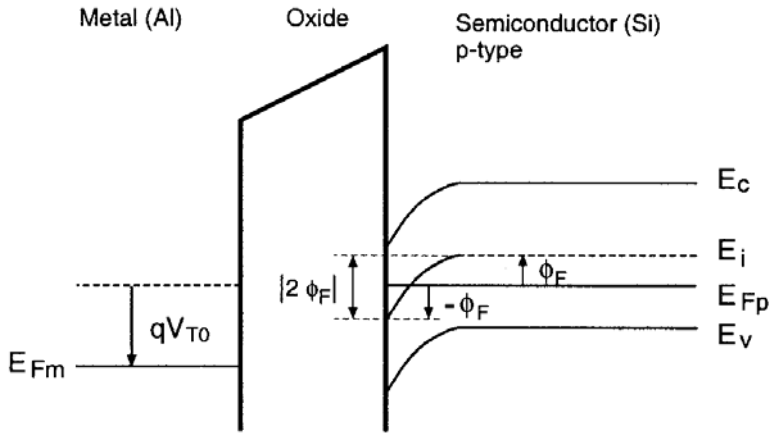


Figure 3.11 Band diagram of the MOS structure underneath the gate, at surface inversion. Notice the band bending by $|2\phi_F|$ at the surface.

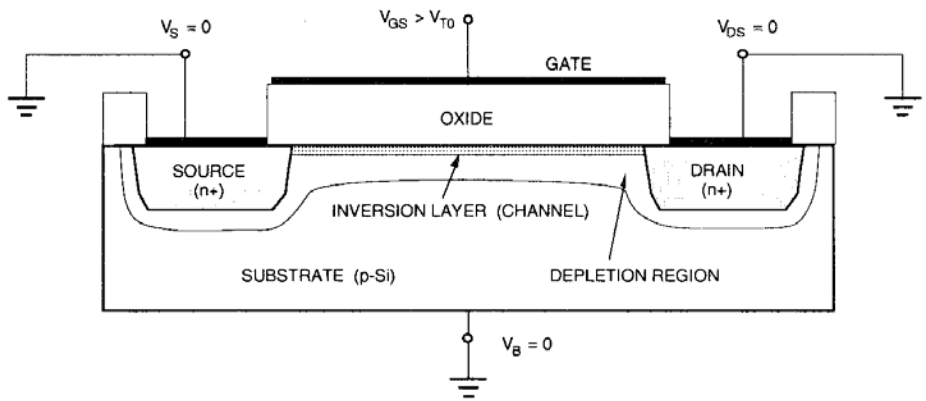


Figure 3.12 Formation of an inversion layer (channel) in an n-channel enhancement-type MOSFET.

- As the gate-to-source voltage is further increased
  - The surface potential reaches $-\phi_{Fp}$ ⇨ surface inversion will be established ⇨ conducting channel between S and D
  - Allowing current flow, as log as there is a potential difference between S and D
  - $V_{GS}<V_{T0}$ (threshold voltage)
    - Not sufficient to establish an inversion layer
    - No current between S and D
  - $V_{GS}>V_{T0}$ (threshold voltage)
    - Electrons are attracted to the surface
      - Contributing to channel current conduction
  - Further increase gate voltage
    - Not affect the surface potential and the depletion region depth

The threshold voltage

- Four physical components of V_{T0}
  - The work function difference between gate and the channel
    - ϕ_{GC}= ϕ_F(substrate)- ϕ_M for metal gate
    - ϕ_{GC}= ϕ_F(substrate)- ϕ_F(gate) for polysilicon gate
  - The gate voltage component to change the surface potential
    - To change the surface potential by -2ϕ_F
  - The gate voltage component to offset the depletion region charge
    - $-Q_B/C_{ox}$
    - $$Q_B = -\sqrt{2q \cdot N_A \cdot \varepsilon_{Si} \cdot \left|-2\phi_F + V_{SB}\right|}$$
      $$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$
  - The voltage component to offset the fixed charge in the gate oxide and in the silicon-oxide interface
    - $-Q_{ox}/C_{ox}$
    - $$V_{T0} = \Phi_{GC} - 2\phi_F - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \text{ (no body effect)}$$
      $$V_T = V_{T0} + \gamma \cdot \left(\sqrt{\left|-2\phi_F + V_{SB}\right|} - \sqrt{\left|2\phi_F\right|}\right) \text{(with body effect)}$$
      $$\text{where } \gamma = \frac{\sqrt{2q \cdot N_A \cdot \varepsilon_{Si}}}{C_{ox}}$$

- Compared with the p-MOSFET
  - The substrate Fermi potential ϕ_F is negative in NMOS, positive in pMOS
  - The depletion region charge densities Q_{B0} and Q_B are negative in nMOS, positive in pMOS
  - The substrate bias coefficient γ is positive in nMOS, negative in pMOS
  - The substrate bias voltage V_{SB} is positive in nMOS, negative in pMOS

- Threshold voltage adjustment
  - Implanting p-type impurity ⇨ V_T increased
  - Implanting n-type impurity ⇨ V_T decreased
  - The amount of change in the threshold voltage
    - Shift qN_I/C_{ox}

Example 2

Calculate the threshold voltage V_{T0} at $V_{SB} = 0$, for a polysilicon gate n-channel MOS transistor, with the following parameters: substrate doping density $N_A = 10^{16}$ cm^{-3}, polysilicon gate doping density $N_D = 2 \times 10^{20}$ cm^{-3}, gate oxide thickness $t_{ox} = 500$ Å, and oxide-interface fixed charge density $N_{ox} = 4 \times 10^{10}$ cm^{-2}.

First, calculate the Fermi potentials for the p-type substrate and for the n-type polysilicon gate:

$$\phi_F(substrate) = \frac{kT}{q} \ln\left(\frac{n_i}{N_A}\right) = 0.026 \text{ V} \cdot \ln\left(\frac{1.45 \cdot 10^{10}}{10^{16}}\right) = -0.35 \text{ V}$$

Since the doping density of the polysilicon gate is very high, the heavily doped n-type gate material is expected to be degenerate. Thus, we may assume that the Fermi potential of the polysilicon gate is approximately equal to the conduction band potential, i.e., $\phi_F(gate) = 0.55$ V. Now, calculate the work function difference between the gate and the channel:

$$\Phi_{GC} = \phi_F(substrate) - \phi_F(gate) = -0.35 \text{ V} - 0.55 \text{ V} = -0.90 \text{ V}$$

The depletion region charge density at $V_{SB} = 0$ is found as follows:

$$\begin{aligned} Q_{B0} &= -\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{Si} \cdot |-2\phi_F(substrate)|} \\ &= -\sqrt{2 \cdot 1.6 \cdot 10^{-19} \cdot 10^{16} \cdot 11.7 \cdot 8.85 \cdot 10^{-14} |-2 \cdot 0.35|} \\ &= -4.82 \cdot 10^{-8} \text{ C/cm}^2 \end{aligned}$$

The oxide-interface charge is:

$$Q_{ox} = q \cdot N_{ox} = 1.6 \cdot 10^{-19} \text{ C} \times 4 \cdot 10^{10} \text{ cm}^{-2} = 6.4 \cdot 10^{-9} \text{ C/cm}^2$$

The gate oxide capacitance per unit area is calculated using the dielectric constant of silicon dioxide and the oxide thickness t_{ox}.

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \frac{3.97 \cdot 8.85 \cdot 10^{-14} \text{ F/cm}}{500 \cdot 10^{-8} \text{ cm}} = 7.03 \cdot 10^{-8} \text{ F/cm}^2$$

Now, we can combine all components and calculate the threshold voltage.

$$\begin{aligned} V_{T0} &= \Phi_{GC} - 2\phi_F(substrate) - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \\ &= -0.90 - (-0.70) - (-0.69) - 0.09 = 0.40 \text{ V} \end{aligned}$$

In this simplified analysis, the doping concentrations of the source and the drain diffusion regions and the geometry (physical dimensions) of the channel region have no influence upon the threshold voltage V_{T0}.

Circuit symbols for n-channel depletion-type MOSFETs

- Using selective ion implantation into the channel
  - The threshold voltage for nMOSFET can be made negative
  - Having a conducting channel at $V_{GS}=0$

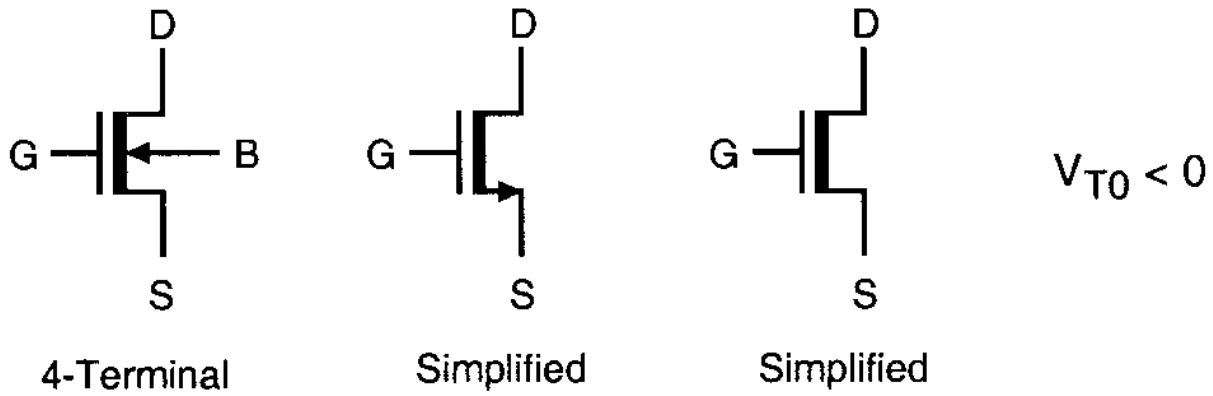


Figure 3.13 Circuit symbols for n-channel depletion-type MOSFETs.

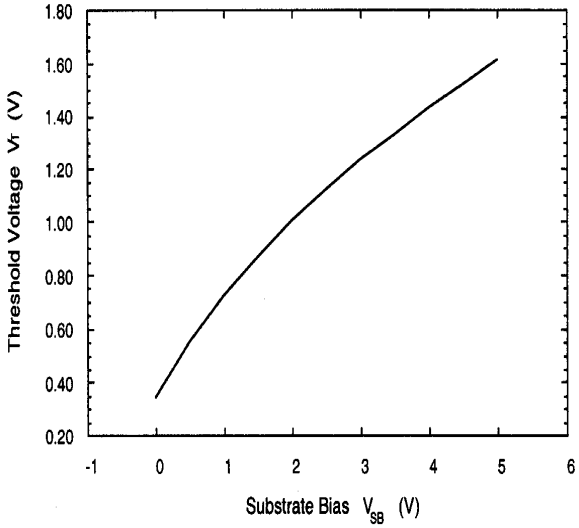
Example 3

Consider the n-channel MOSFET process given in Example 3.2. In several digital circuit applications, the condition $V_{SB} = 0$ cannot be guaranteed for all transistors. We will examine in this example how a nonzero source-to-substrate voltage V_{SB} affects the threshold voltage of the MOS transistor.

First, we must calculate the substrate-bias coefficient γ using the process parameters given in Example 3.2.

$$\gamma = \frac{\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{Si}}}{C_{ox}} = \frac{\sqrt{2 \cdot 1.6 \cdot 10^{-19} \cdot 10^{16} \cdot 11.7 \cdot 8.85 \cdot 10^{-14}}}{7.03 \cdot 10^{-8}}$$

$$= 0.82\ \mathrm{V}^{\frac{1}{2}}$$



Variation of the threshold voltage as a function of the source-to-substrate voltage.

Now we compute and plot the threshold voltage V_T as a function of the source-to-substrate voltage V_{SB}. The voltage V_{SB} will be assumed to vary between zero and 5 V.

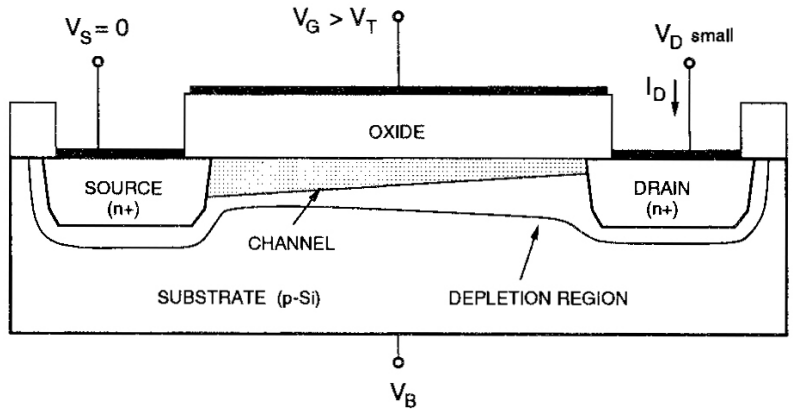
$$V_T = V_{T0} + \gamma \cdot \left(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right)$$

$$= 0.40 + 0.82 \cdot \left(\sqrt{0.7 + V_{SB}} - \sqrt{0.7}\right)$$

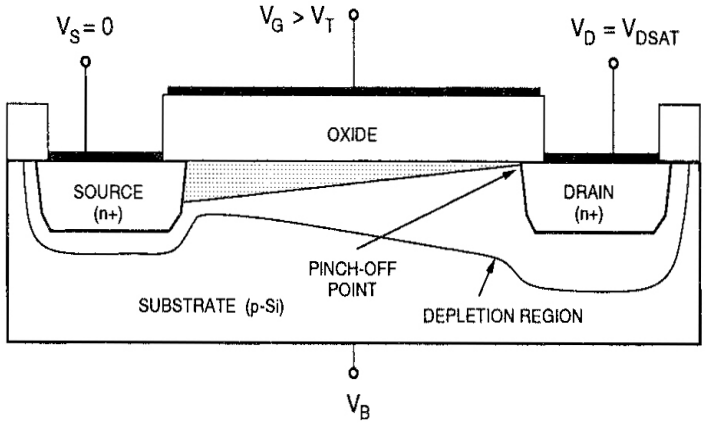
It is seen that the threshold voltage variation is about 1.3 V over this range, which could present serious design problems if neglected. We will see in the following chapters that the substrate-bias effect is unavoidable in most digital circuits and that the circuit designer usually must take appropriate measures to account for and/or to compensate for the threshold voltage variations.

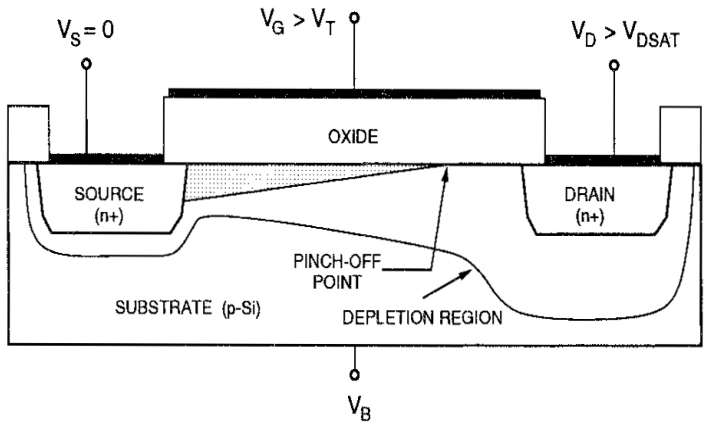
MOSFET operation: linear region

- The MOSFET consists
  - A MOS capacitor, two pn junction adjacent to the channel
  - The channel is controlled to the MOS gate
- The carrier (electron in nMOSFET)
  - Entering through source, controlling by gate, leaving through drain
- To ensure that both p-n junctions are reverse-biased initially
  - The substrate potential is kept lower than the other three terminal potentials
- When $0<V_{GS}<V_{T0}$
  - G-S region depleted, G-D region depleted
  - No current flow
- When $V_{GS}>V_{T0}$
  - Conduction channel formed
  - Capable of carrying the drain current
  - As $V_{DS}=0$
    - $I_D=0$
  - As $V_{DS}>0$ and small
    - I_D proportional to V_{DS}
    - Flowing from S to D through the conducting channel
    - The channel act as a voltage controlled resistor
    - The electron velocity much lower than the drift velocity limit
    - As $V_{DS}\uparrow\Rightarrow$the inversion layer charge and the channel depth at the drain end start to decrease



MOSFET operation: saturation region





- For $V_{DS}=V_{DSAT}$
  - The inversion charge at the drain is reduced to zero
  - *Pitch off* point
- For $V_{DS}>V_{DSAT}$
  - A depleted surface region forms adjacent to the drain
  - As further increases V_{DS} ⇨ this depletion region grows toward the source
  - The channel-end remains essentially constant and equal to V_{DSAT}
  - The pitch-off (depleted) section
    - Absorbing most of the excess voltage drop, V_{DS}-V_{DSAT}
    - A high-field forms between the channel-end of the drain boundary
      - Accelerating electrons, usually reaching the drift velocity limit

MOSFET current-voltage characteristics-gradual channel approximation (GCA)(1)

- Considering linear mode operation
  - $V_S=V_B=0$, the V_{GS} and V_{DS} are the external parameters controlling the drain current I_D
  - $V_{GS} > V_{T0}$ (assume constant through the channel) to create a conducting inversion layer
  - Defining
    - X-direction: perpendicular to the surface, pointing down into the substrate
    - Y-direction: parallel to the surface
      - The y=0 is at the source end of the channel
      - Channel voltage with respect to the source, $V_c(y)$
  - Assume the electric field E_y is *dominant* compared with E_x
    - This assumption reduced ⇨the current flow in the channel to the y-direction only
  - Let $Q_I(y)$ be the total mobile electron charge in the surface inversion layer
    - $Q_I(y)=-C_{ox}[V_{GS}-Vc(y)-V_{T0}]$

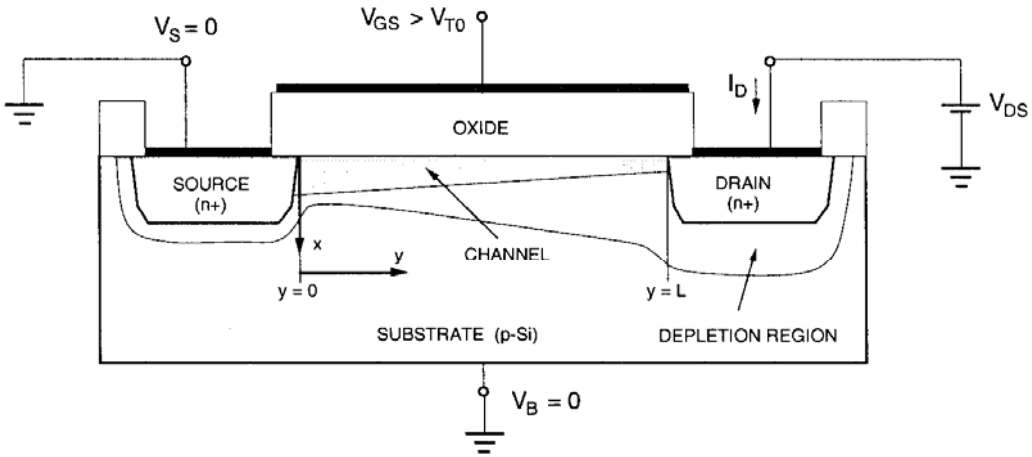


Figure 3.15 Cross-sectional view of an n-channel transistor, operating in linear region.

MOSFET current-voltage characteristics-gradual channel approximation (GCA)(2)

Assumeing that all mobile electrons in the inversion layer has a constant surfacr mobility μ_n

$dR = -\dfrac{dy}{W \cdot \mu_n \cdot Q_I(y)}$ (mimus sign is due to the negative polarity of the inversion layer charge Q_I)

The electron surface mobility μ_n dependents on the doping concentration of the channel region, and its magnitude is typically about one - half of that of the bulk electron mobility

$$dV_C = I_D \cdot dR = -\frac{I_D}{W \cdot \mu_n \cdot Q_I(y)} \cdot dy$$

$$\int_0^L I_D \cdot dy = -W \cdot \mu_n \int_0^{V_{DS}} Q_I(y) \cdot dV_C$$

$$I_D \cdot L = W \cdot \mu_n \cdot C_{ox} \int_0^{V_{DS}} \left(V_{GS} - V_C - V_{T0}\right) \cdot dV_C$$

$$I_D = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot \left(V_{GS} - V_{T0}\right) V_{DS} - V_{DS}^2\right]$$

$$I_D = \frac{k'}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot \left(V_{GS} - V_{T0}\right) V_{DS} - V_{DS}^2\right] \quad where\ k' = \mu_n C_{ox}$$

$$I_D = \frac{k}{2} \cdot \left[2 \cdot \left(V_{GS} - V_{T0}\right) V_{DS} - V_{DS}^2\right] \quad where\ k = k' \cdot \frac{W}{L}$$

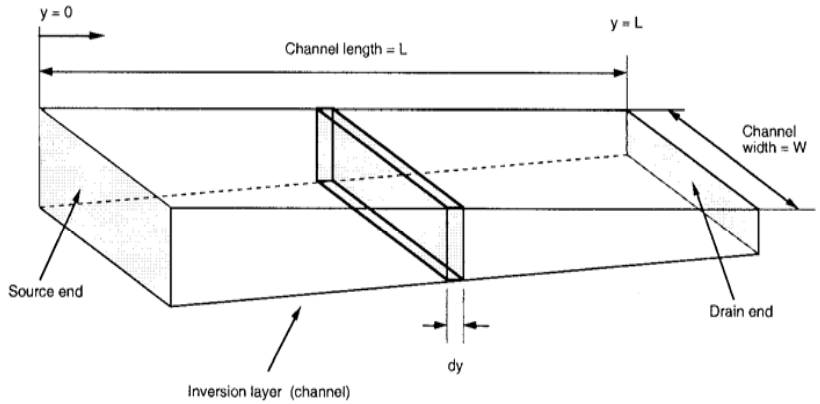


Figure 3.16 Simplified geometry of the surface inversion layer (channel region).

Example 4

For an n-channel MOS transistor with $\mu_n = 600\,\text{cm}^2/\text{V}\cdot\text{s}$, $C_{ox} = 7 \cdot 10^{-8}$ F/cm^2, $W = 20\ \mu$m, $L = 2\ \mu$m and $V_{T0} = 1.0$ V, examine the relationship between the drain current and the terminal voltages.

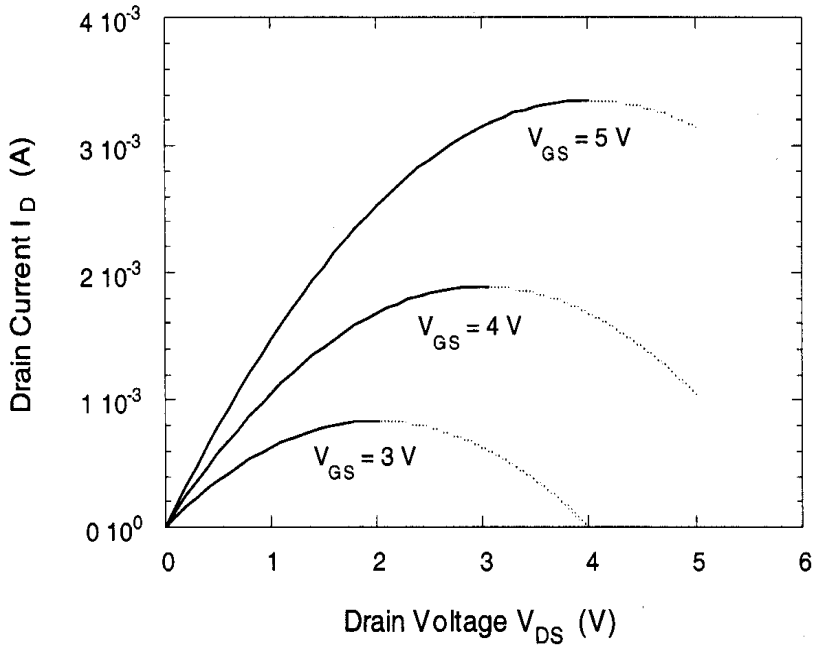
First, calculate the parameter k:

$$k = \mu_n \cdot C_{ox} \cdot \frac{W}{L} = 600\ \text{cm}^2/\text{V}\cdot\text{s} \times 7 \cdot 10^{-8}\ \text{F/cm}^2 \times \frac{20\ \mu\text{m}}{20\ \mu\text{m}} = 0.42\ \text{mA/V}^2$$

Now, the current-voltage equation (3.34) can be written as follows.

$$I_D = 0.21\ \text{mA/V}^2 \left[2 \cdot (V_{GS} - 1.0) \cdot V_{DS} - V_{DS}^2\right]$$

To examine the effect of the gate-to-source voltage and the drain-to-source voltage upon the drain current, we will plot I_D as a function of V_{DS}, for different (constant) values of V_{GS}. It can easily be seen that the second-order current-voltage equation given above produces a set of inverted parabolas for each constant V_{GS} value.



The drain current-drain voltage curves shown on page 104 reach their peak value for $V_{DS} = V_{GS} - V_{T0}$. Beyond this maximum, each curve exhibits a *negative* differential conductance, which is not observed in actual MOSFET current-voltage measurements (section shown by the dashed lines). We must remember now that the drain current equation (3.32) has been derived under the following voltage assumptions,

$$V_{GS} \geq V_{T0}$$

$$V_{GD} = V_{GS} - V_{DS} \geq V_{T0}$$

which guarantee that the entire channel region between the source and the drain is inverted. This condition corresponds to the *linear* operating mode for the MOSFET, which was examined qualitatively in Section 3.4. Hence, the current equation (3.32) is valid only for the linear mode operation. Beyond the linear region boundary, i.e., for V_{DS} values *larger* than $V_{GS} - V_{T0}$, the MOS transistor will be assumed to be in *saturation*. A different current-voltage expression will be necessary for the MOSFET operating in this region.

MOSFET current-voltage characteristics-gradual channel approximation (GCA)-saturation region

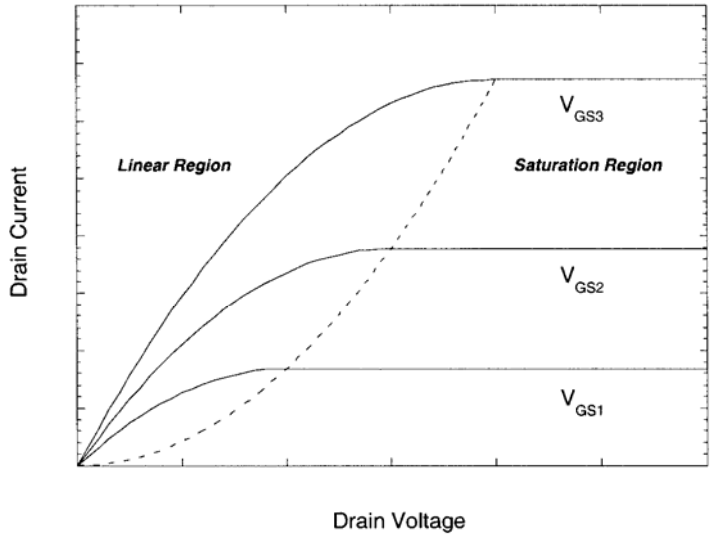


Figure 3.17 Basic current-voltage characteristics of an n-channel MOS transistor.

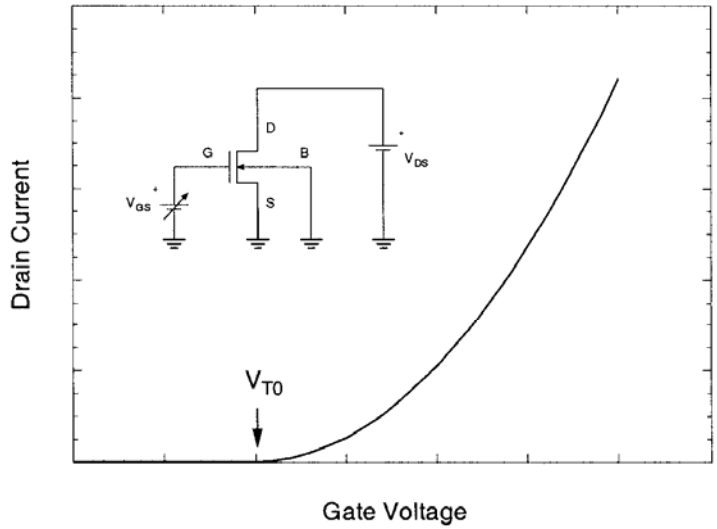


Figure 3.18 Drain current of the n-channel MOS transistor as a function of the gate-to-source voltage V_{GS}, with $V_{DS} > V_{DSAT}$ (transistor in saturation).

- For $V_{DS} \geq V_{DSAT} = V_{GS} - V_{T0}$
  - $$I_{D(sat)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot (V_{GS} - V_{T0}) \cdot (V_{GS} - V_{T0}) - (V_{GS} - V_{T0})^2\right]$$
    $$= \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_{T0})^2$$
  - The drain current becomes a function only of V_{GS}, beyond the saturation boundary

Channel length modulation

The inversion layer charge at the source end of the channel is

$$Q_I(y=0) = -C_{ox} \cdot (V_{GS} - V_{T0})$$

and the inversion layer charge at the drain end of the channel is

$$Q_I(y=L) = -C_{ox} \cdot (V_{GS} - V_{T0} - V_{DS})$$

Note that at the edge of saturation, $V_{DS} = V_{DSAT} = V_{GS} - V_{T0}$

The inversion layer charge at the drain end become very small

$$Q_I(y=L) \approx 0$$

The effective channel length $L' = L - \Delta -$

where $\Delta\Delta$ is the length of the channel segment with $Q_I = 0$

$$I_{D(sat)} = \frac{\mu_n C_{ox}}{2} \cdot \frac{W}{L'} \cdot (V_{GS} - V_{T0})^2$$

$$I_{D(sat)} = \left(\frac{1}{1 - \frac{\Delta L}{L}} \right) \frac{\mu_n C_{ox}}{2} \cdot \frac{W}{L'} \cdot (V_{GS} - V_{T0})^2$$

$$\Delta L \propto \sqrt{V_{DS} - V_{DSAT}}$$

We use $1 - \frac{\Delta L}{L} \approx 1 - \lambda \cdot V_{DS}$, λ channel length modulation coeffic

Assuming that $\lambda\lambda_{DS} << 1$

$$I_{D(sat)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_{T0})^2 \cdot (1 + \lambda V_{DS})$$

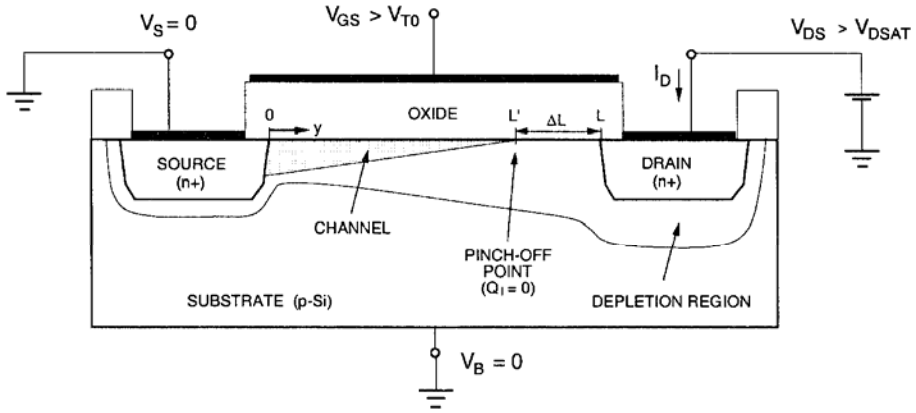


Figure 3.19 Channel length modulation in an n-channel MOSFET operation in saturation mode.

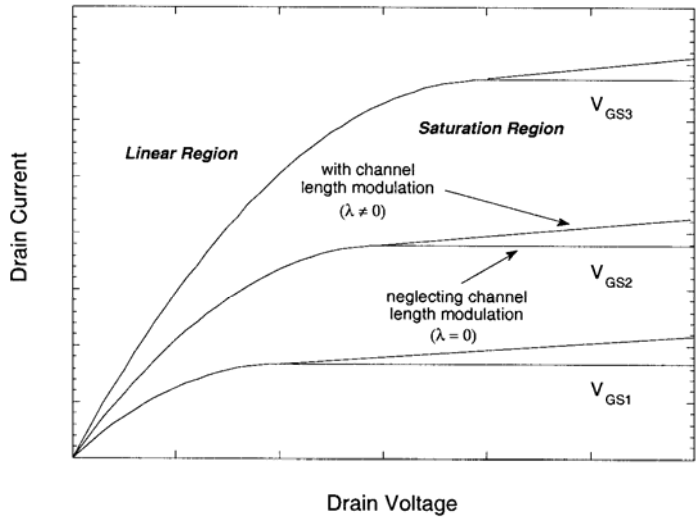


Figure 3.20 Current-voltage characteristics of an n-channel MOS transistor, including the channel length modulation effect.

Substrate bias effect

- The discussion in the previous has been done under the assumption
  - The substrate potential is equal to the source potential, i.e. $V_{SB}=0$
- On the other hand
  - the source potential of an nMOS transistor can be larger than the substrate potential, i.e. $V_{SB}>0$
  - $V_T(V_{SB}) = V_{T0} + \gamma \cdot \left(\sqrt{|2\phi_F| + V_{SB}} - \sqrt{|2\phi_F|} \right)$

$$I_{D(lin)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot \left(V_{GS} - V_T(V_{SB}) \right) V_{DS} - V_{DS}^2 \right]$$

$$I_{D(sat)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left(V_{GS} - V_T(V_{SB}) \right)^2 \cdot \left(1 + \lambda \cdot V_{DS} \right)$$

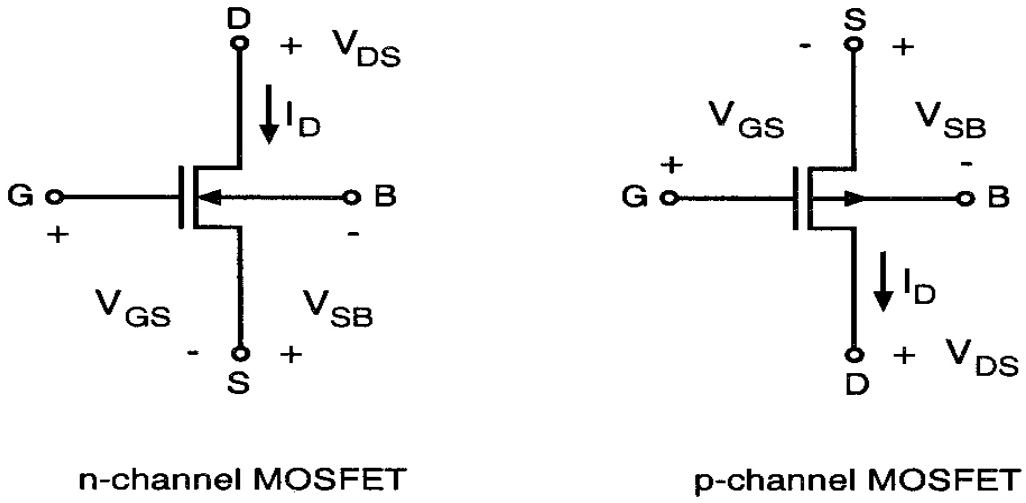


Figure 3.21 Terminal voltages and currents of the nMOS and the pMOS transistor.

Current-voltage equation of n-, p-channel MOSFET

For n - channel MOSFET

$$I_D = 0, \quad \text{for } V_{GS} < V_T$$

$$I_{D(lin)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot (V_{GS} - V_T) V_{DS} - V_{DS}^2\right] \text{ for } V_{GS} \geq V_T$$

$$\text{and } V_{DS} < V_{GS} - V_T$$

$$I_{D(sat)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2 \cdot (1 + \lambda \cdot V_{DS}) \text{ for } V_{GS} \geq V_T$$

$$\text{and } V_{DS} \geq V_{GS} - V_T$$

For p - channel MOSFET

$$I_D = 0, \quad \text{for } V_{GS} > V_T$$

$$I_{D(lin)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left[2 \cdot (V_{GS} - V_T) V_{DS} - V_{DS}^2\right] \text{ for } V_{GS} \leq V_T$$

$$\text{and } V_{DS} > V_{GS} - V_T$$

$$I_{D(sat)} = \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2 \cdot (1 + \lambda \cdot V_{DS}) \text{ for } V_{GS} \leq V_T$$

$$\text{and } V_{DS} \leq V_{GS} - V_T$$

Measurement of parameters- k_n, V_{T0}, and γ

- The V_{SB} is set at a constant value
  - The drain current is measured for different values of V_{GS}
  - V_{DG}=0
    - V_{DS}>V_{GS}-V_T is always satisfied ⇨ saturation mode
    - Neglecting the channel length modulation effect
      - $I_{D(sat)} = \frac{k_n}{2} \cdot (V_{GS} - V_{T0})^2, \sqrt{I_D} = \sqrt{\frac{k_n}{2}} \cdot (V_{GS} - V_{T0})$
      - Obtaining the parameters k_n, V_{T0}, and γ
      - $$\gamma = \frac{V_T(V_{SB}) - V_{T0}}{\sqrt{|2\phi_F| + V_{SB}} - \sqrt{|2\phi_F|}}$$

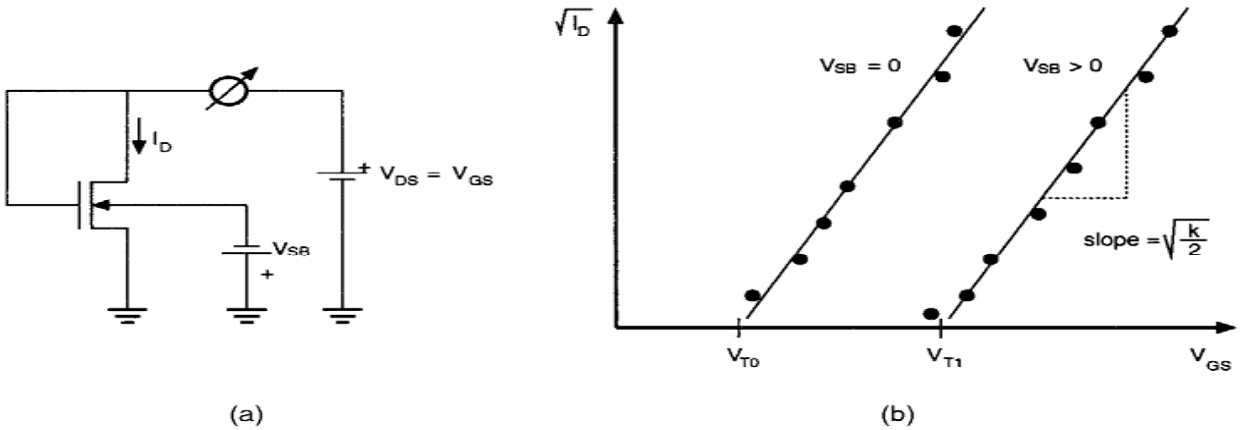


Figure 3.22 (a) Test circuit arrangement and (b) measured data for experimental determination of the parameters k_n, V_{T0}, and γ.

Measurement of parameters- λ

- The voltage V_{GS} is set to $V_{T0}+1$
- The voltage V_{DS} is chosen sufficiently large ($V_{DS}>V_{GS}-V_{T0}$) that the transistor operates in the saturation mode, V_{DS1}, V_{DS2}
  - $I_{D(sat)}-(k_n/2)(V_{GS}-V_{T0})^2(1+\lambda V_{DS})$
    - Since $V_{GS}=V_{T0}+1 \Rightarrow I_{D2}/I_{D1}=(1+\lambda V_{DS2})/(1+\lambda V_{DS1})$
    - Which can be used to calculate the channel length modulation coefficient λ
    - This is in fact equivalent to calculating the slope of the drain current versus drain voltage curve in the saturation region
      - The slope is $\lambda k_n/2$

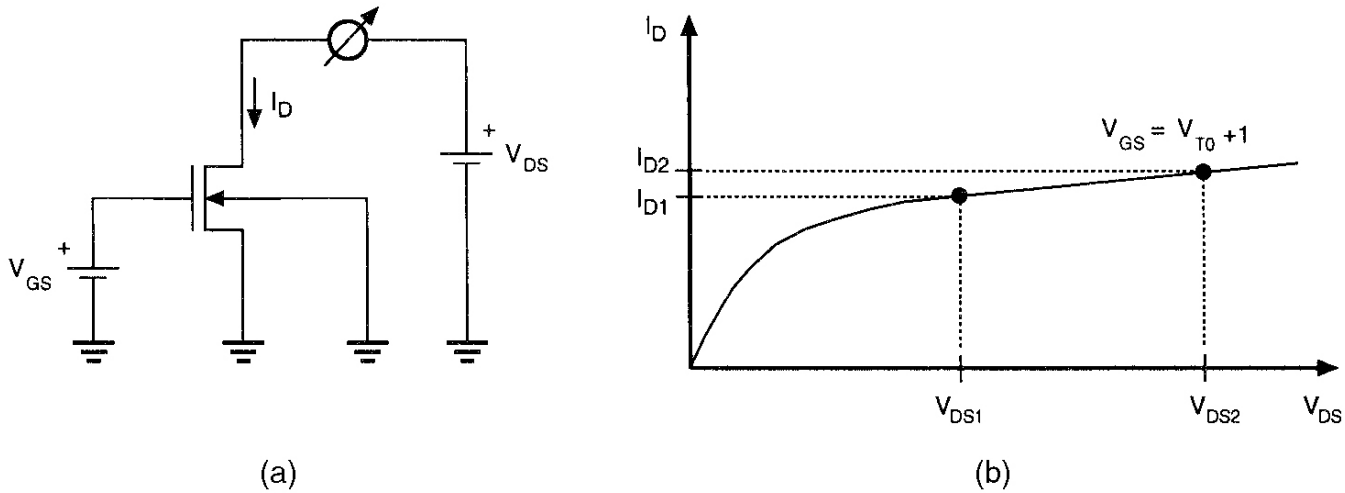


Figure 3.23 (a) Test circuit arrangement and (b) measured data for experimental determination of the channel length modulation coefficient λ.

Example 5

Measured voltage and current data for a MOSFET are given below. Determine the type of the device, and calculate the parameters k_n, V_{T0}, and γ. Assume $\phi_F = -0.3$ V.

V_{GS} (V)	V_{DS} (V)	V_{SB} (V)	I_D (μA)
3	3	0	97
4	4	0	235
5	5	0	433
3	3	3	59
4	4	3	173
5	5	3	347

First, the MOS transistor is on ($I_D > 0$) for $V_{GS} > 0$ and $V_{DS} > 0$. Thus, the transistor must be an n-channel MOSFET. Assume that the transistor is enhancement-type and, therefore, operating in saturation mode for $V_{GS} = V_{DS}$. Neglecting the channel length modulation effect, the saturation mode current is written as

$$I_D = \frac{k_n}{2} \cdot (V_{GS} - V_T)^2 \qquad \Leftrightarrow \qquad \sqrt{I_D} = \sqrt{\frac{k_n}{2}} \cdot (V_{GS} - V_T)$$

Let (V_{GS1}, I_{D1}) and (V_{GS2}, I_{D2}) be any two current-voltage pairs obtained from the table. Then, the square-root of the transconductance parameter k_n can be calculated.

$$\sqrt{\frac{k_n}{2}} = \frac{\sqrt{I_{D1}} - \sqrt{I_{D2}}}{V_{GS1} - V_{GS2}} = \frac{\sqrt{433\ \mu\text{A}} - \sqrt{97\ \mu\text{A}}}{5\ \text{V} - 3\ \text{V}} = 5.48 \times 10^{-3}\ \text{A}^{1/2}/\text{V}$$

Thus, the transconductance parameter of this n-channel MOSFET is:

$$k_n = 2 \cdot (5.48 \times 10^{-3})^2 = 60 \times 10^{-6}\ \text{A/V}^2 = 60\ \mu\text{A/V}^2$$

The *extrapolated threshold voltage* V_{T0} at zero substrate bias can be found by calculating the x-axis intercept of the *square-root of* (I_D) versus V_{GS} curve.

$$V_{T0} = V_{GS} - \sqrt{\frac{2 \cdot I_D}{k_n}} = 1.2\ \text{V}$$

To find the substrate bias coefficient γ, we must first determine the threshold voltage V_T at the source-to-substrate voltage of 3 V. Using one of the current-voltage data pairs corresponding to $V_{SB} = 3$ V, V_T can be calculated as follows:

$$V_T(V_{SB} = 3\ \text{V}) = V_{GS} - \sqrt{\frac{2 \cdot I_D}{k_n}} = 4\ \text{V} - \sqrt{\frac{2 \cdot 173\ \mu\text{A}}{60\ \mu\text{A/V}^2}} = 1.6\ \text{V}$$

Finally, the substrate bias coefficient is found as:

$$\gamma = \frac{V_T(V_{SB} = 3\ \text{V}) - V_{T0}}{\sqrt{|2\phi_F| + V_{SB}} - \sqrt{|2\phi_F|}} = \frac{1.6\ \text{V} - 1.2\ \text{V}}{\sqrt{0.6\ \text{V} + 3\ \text{V}} - \sqrt{0.6\ \text{V}}} = 0.36\ \text{V}^{1/2}$$

MOSFET scaling and small-geometry effects

- High density chip
  - The sizes of the transistors are as small as possible
  - The operational characteristics of MOS transistor will change with the reduction of iys dimensions
- There are two basic types of size-reduction strategies
  - Full scaling (constant-field scaling)
  - Constant-voltage scaling
- A new generation of manufacturing technology replaces the previous one about
  - every two or three years
  - The down-scaling factor S about 1.2 to1.5
- The scaling of all dimensions by a factor of S>1 leads to the reduction of the area occupied by the transistor by a factor of S^2

Table 3.1 Reduction of the minimum feature size (minimum dimensions that can be defined and manufactured on chip) over the years, for a typical CMOS gate-array process

Year	1985	1987	1989	1991	1993	1995	1997	1999
Feature size (μm)	2.5	1.7	1.2	1.0	0.8	0.5	0.35	0.25

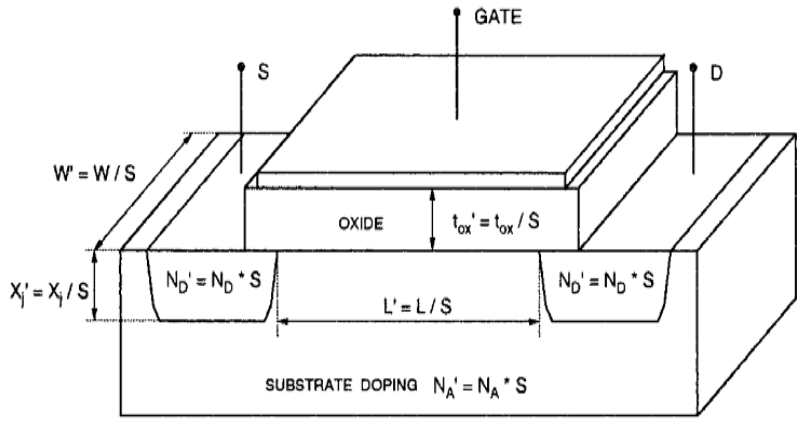


Figure 3.24 Scaling of a typical MOSFET by a scaling factor of *S*.

Full scaling (constant-field scaling)

To achieve this goal, all potentials must be scaled down proportionally, by the same scaling factor

Assuming the surface mobility μ_n is not significantly affected by the scaled doping density

The gate oxide capacitance per unit area

$$C_{ox}^{'} = \frac{\varepsilon_{ox}}{t_{ox}^{'}} = S \cdot \frac{\varepsilon_{ox}}{t_{ox}} = S \cdot C_{ox}$$

The aspect ratio *W/L* unchanged $\Rightarrow$ the k_n will also scaled by a factor of S

The linear mode drain current

$$I_D^{'}(lin) = \frac{k_n^{'}}{2} \cdot \left[2 \cdot \left(V_{GS}^{'} - V_T^{'}\right) \cdot V_{DS}^{'} - V_{DS}^{'2}\right]$$

$$= \frac{S \cdot k_n}{2} \cdot \frac{1}{S^2} \cdot \left[2 \cdot \left(V_{GS} - V_T\right) \cdot V_{DS} - V_{DS}^2\right] = \frac{I_{D(lin)}}{S}$$

The saturation mode drain current

$$I_D^{'}(sat) = \frac{k_n^{'}}{2} \cdot \left(V_{GS}^{'} - V_T^{'}\right)^2 = \frac{S \cdot k_n}{2} \cdot \frac{1}{S^2} \cdot \left(V_{GS} - V_T\right)^2 = \frac{I_{D(sat)}}{S}$$

The power dissipation

$$P^{'} = I_D^{'} \cdot V_{DS}^{'} = \frac{1}{S^2} \cdot I_D \cdot V_{DS} = \frac{P}{S^2}$$

Table 3.2 Full scaling of MOSFET dimensions, potentials, and doping densities

Quantity	Before scaling	After scaling
Channel length	L	$L' = L/S$
Channel width	W	$W' = W/S$
Gate oxide thickness	t_{ox}	$t'_{ox} = t_{ox}/S$
Junction depth	x_j	$x'_j = x_j/S$
Power supply voltage	V_{DD}	$V'_{DD} = V_{DD}/S$
Threshold voltage	V_{T0}	$V'_{T0} = V_{T0}/S$
Doping densities	N_A	$N'_A = S \cdot N_A$
	N_D	$N'_D = S \cdot N_D$

Table 3.3 Effects of full scaling upon key device characteristics

Quantity	Before scaling	After scaling
Oxide capacitance	C_{ox}	$C'_{ox} = S \cdot C_{ox}$
Drain current	I_D	$I'_D = I_D/S$
Power dissipation	P	$P' = P/S^2$
Power density	$P/Area$	$P'/Area' = P/Area$

The significant reduction of the power dissipation is one of the most attractive features of full scaling

The *power density* per unit area remaining virtually unchanged

C_g is scaled down by a factor of $S \Rightarrow$ the charge - up, and charge - down time improved

A reduction of various parasitic capacitances abd resistances

Constant-voltage scaling

All dimensions of the MOSFETare reduced by a factor of S.

The power supply voltage and the terminal voltages remained unchanged.

The doping densities must be increased by a factor of S^2 in order to preserve the charge - field relations

The gate oxide capacitance per unit area C_{ox} is increased by a factor of S

$\Rightarrow$ The transconductance parameter is also increased by S

The linear mode drain current

$$I_D'(lin) = \frac{k_n'}{2} \cdot \left[2 \cdot \left(V_{GS}' - V_T'\right) \cdot V_{DS}' - V_{DS}'^2\right]$$

$$= \frac{S \cdot k_n}{2} \cdot \left[2 \cdot \left(V_{GS} - V_T\right) \cdot V_{DS} - V_{DS}^2\right] = S \cdot I_D(lin)$$

The saturation mode drain current

$$I_D'(sat) = \frac{k_n'}{2}\left(V_{GS}' - V_T'\right)^2 = \frac{S \cdot k_n}{2} \cdot \left(V_{GS} - V_T\right)^2 = S \cdot I_D(sat)$$

The drain current density increased by a factor of S^3

The power dissipation

$$P' = I_D' \cdot V_{DS}' = (S \cdot I_D) \cdot V_{DS} = S \cdot P$$

The power density incresaed by a factor of S^3

Table 3.5 Effects of constant-voltage scaling upon key device characteristics

Quantity	Before scaling	After scaling
Oxide capacitance	C_{ox}	$C_{ox}' = S \cdot C_{ox}$
Drain current	I_D	$I_D' = S \cdot I_D$
Power dissipation	P	$P' = S \cdot P$
Power density	$P/Area$	$P'/Area' = S^3 \cdot (P/Area)$

To summarized, constant - voltage scaling may be preferred over full scaling in mamy practical cases because of the external voltage - level constraints.

Disadv. $\Rightarrow$ increasing current density, power density

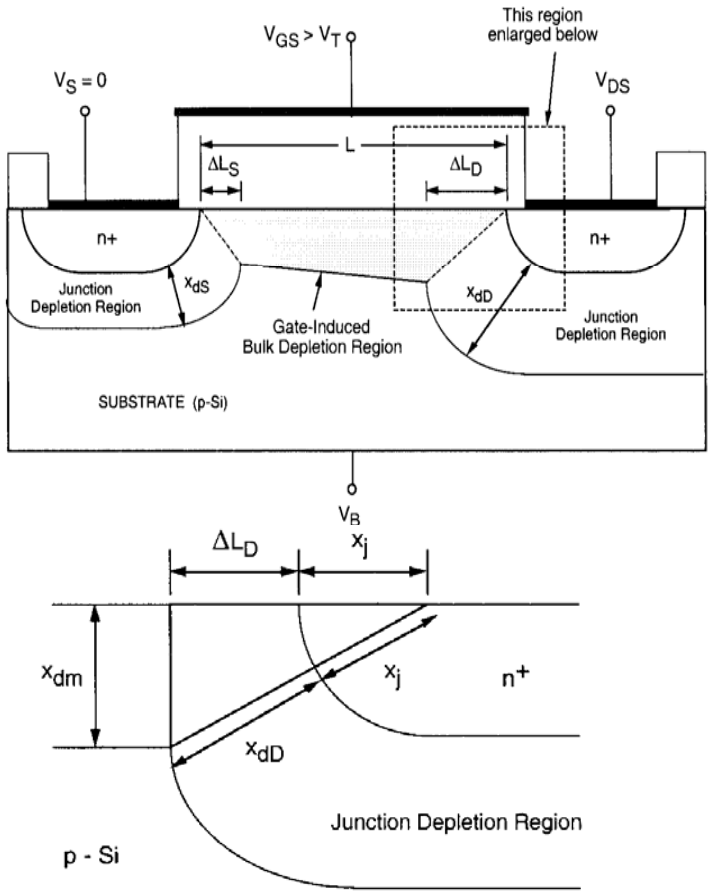
$\Rightarrow$ electromigration, hot carrier degradation, oxide breakdown, and electrical over - stress

Short-channel effects

- A MOS transistor is called a short-channel device
  - If its channel length is on the same order of magnitude as the depletion region thickness of the *S* and *D* junction
  - The effective channel length $L_{eff} \approx$ *S, D* junction depth x_j
  - Two physical phenomena arise from short-channel effects
    - The limitations imposed on electron drift characteristics in the channel
      - The lateral electric field E_y increased, v_d reached saturation velocity
      - $$I_{D(sat)} = W \cdot v_{d(sat)} \cdot \int_0^{L_{eff}} q \cdot n(x) \cdot dx = W \cdot v_{d(sat)} \cdot |Q_I| = W \cdot v_{d(sat)} \cdot C_{ox} \cdot V_{DSAT}$$
        - » No longer a quadratic function of V_{GS}, virtually independent of the channel length
      - The carrier velocity in the channel also a function of E_x
        - » Influence the scattering of carriers in the surface
        - » $$\mu_n(eff) = \frac{\mu_{no}}{1 + \Theta \cdot Ex} = \frac{\mu_{no}}{1 + \frac{\Theta \varepsilon_{ox}}{t_{ox} \varepsilon_{Si}} \cdot (V_{GS} - V_c(y))} = \frac{\mu_{no}}{1 + \eta \cdot (V_{GS} - V_T)}$$
    - The modification of the threshold voltage due to the shortening channel length

Short-channel effects-modification of V_T

- The n^+ drain and source diffusion regions in p-type substrate induce a significant amount of depletion charge
  - The long channel VT, overetimates the depletion charge support by the gate voltage
  - The bulk depletion region ⇨ asymmetric trapezoidal shape
    - A significant portion of the total depletion region charge is due the S and D junction depletion



$$V_{T0}(short\ channel) = V_{T0} - \Delta V_{T0}$$

$$Q_{B0} = -\left(1 - \frac{\Delta L_S + \Delta L_D}{2L}\right) \cdot \sqrt{2 \cdot q \cdot \varepsilon_{Si} \cdot N_A \cdot |2\varphi_F|}$$

$$x_{dS} = \sqrt{\frac{2 \cdot \varepsilon_{Si}}{q \cdot N_A} \cdot \varphi_0}, \quad x_{dD} = \sqrt{\frac{2 \cdot \varepsilon_{Si}}{q \cdot N_A} \cdot (\varphi_0 + V_{DS})}, \quad \varphi_0 = \frac{kT}{q} \cdot \ln\left(\frac{N_D \cdot N_A}{n_i^2}\right)$$

$$(x_j + x_{dD})^2 = x_{dm}^2 + (x_j + \Delta L_D)^2$$

$$\Delta L_D^2 + 2 \cdot x_j \cdot \Delta L_D + x_{dm}^2 - x_{dD}^2 - 2 \cdot x_j \cdot x_{dD} = 0$$

$$\Delta L_D = -x_j + \sqrt{x_j^2 - (x_{dm}^2 - x_{dD}^2) + 2 x_j x_{dD}} \cong x_j \cdot \left(\sqrt{1 + \frac{2x_{dD}}{x_j}} - 1\right)$$

$$\Delta L_S \cong x_j \cdot \left(\sqrt{1 + \frac{2x_{dS}}{x_j}} - 1\right)$$

$$\Delta V_{T0} = \frac{1}{C_{ox}} \cdot \sqrt{2 \cdot q \cdot \varepsilon_{Si} \cdot N_A \cdot |2\varphi_F|} \cdot \frac{x_j}{2L} \cdot \left[\left(\sqrt{1 + \frac{2x_{dD}}{x_j}} - 1\right) + \left(\sqrt{1 + \frac{2x_{dS}}{x_j}} - 1\right)\right]$$

Example 6 (1)

Consider an n-channel MOS process with the following parameters: substrate doping density $N_A = 10^{16}$ cm^{-3}, polysilicon gate doping density N_D (gate) $= 2 \times 10^{20}$ cm^{-3}, gate oxide thickness $t_{ox} = 50$ nm, oxide-interface fixed charge density $N_{ox} = 4 \times 10^{10}$ cm^{-2}, and source and drain diffusion doping density $N_D = 10^{17}$ cm^{-3}. In addition, the channel region is implanted with p-type impurities (impurity concentration $N_I = 2 \times 10^{11}$ cm^{-2}) to adjust the threshold voltage. The junction depth of the source and drain diffusion regions is $x_j = 1.0$ μm.

Plot the variation of the zero-bias threshold voltage V_{T0} as a function of the channel length (assume that $V_{DS} = V_{SB} = 0$). Also find V_{T0} for $L = 0.7$ μm, $V_{DS} = 5$ V, and $V_{SB} = 0$.

First, we have to find the zero-bias threshold voltage using the conventional formula (3.23). The threshold voltage *without* the channel implant was already calculated for the same process parameters in Example 3.2, and was found to be $V_{T0} = 0.40$ V. The additional p-type channel implant will increase the threshold voltage by an amount of qN_I/C_{ox}. Thus, we find the long-channel zero-bias threshold voltage for the process described above as

$$V_{T0} = 0.40 \text{ V} + \frac{q \cdot N_I}{C_{ox}} = 0.40 \text{ V} + \frac{1.6 \times 10^{-19} \cdot 2 \times 10^{11}}{7.03 \times 10^{-8}} = 0.855 \text{ V}$$

Next, the amount of threshold voltage reduction due to short-channel effects must be calculated using (3.88). The source and drain junction built-in voltage is

$$\phi_0 = \frac{kT}{q} \cdot \ln\left(\frac{N_D \cdot N_A}{n_i^2}\right) = 0.026 \text{ V} \cdot \ln\left(\frac{10^{17} \cdot 10^{16}}{2.1 \times 10^{20}}\right) = 0.76 \text{ V}$$

Example 6 (2)

For zero drain bias, the depth of source and drain junction depletion regions is found as

$$x_{dS} = x_{dD} = \sqrt{\frac{2 \cdot \varepsilon_{Si}}{q \cdot N_A} \cdot \phi_0} = \sqrt{\frac{2 \cdot 11.7 \cdot 8.85 \times 10^{-14}}{1.6 \times 10^{-19} \cdot 10^{16}} \cdot 0.76}$$

$$= 31.4 \times 10^{-6} \text{ cm} = 0.314\ \mu\text{m}$$

Now, the threshold voltage shift ΔV_{T0} due to short-channel effects can be calculated as a function of the gate (channel) length L.

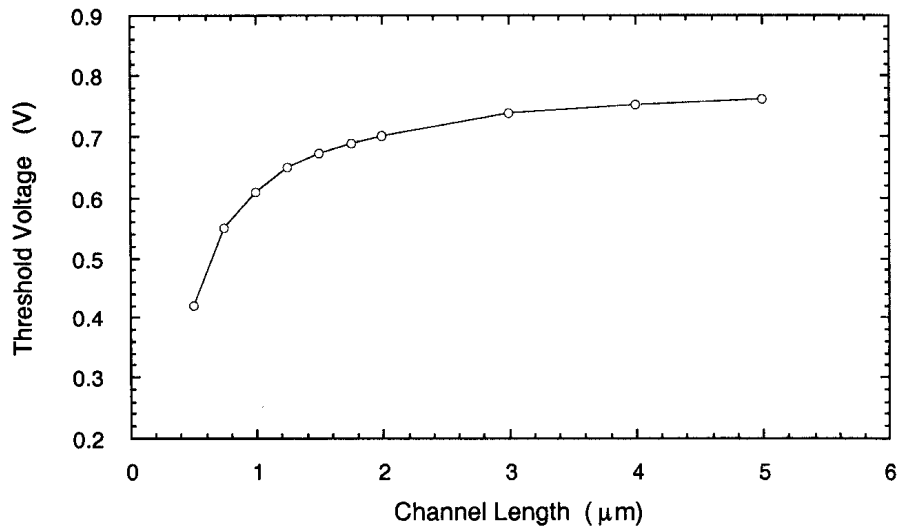
$$\Delta V_{T0} = \frac{1}{C_{ox}} \cdot \sqrt{2q\varepsilon_{Si} N_A |2\phi_F|} \cdot \frac{x_j}{2L} \cdot \left[\left(\sqrt{1 + \frac{2x_{dS}}{x_j}} - 1 \right) + \left(\sqrt{1 + \frac{2x_{dD}}{x_j}} - 1 \right) \right]$$

$$= \frac{4.82 \times 10^{-8} \text{ C/cm}^2}{7.03 \times 10^{-8} \text{ F/cm}^2} \cdot \frac{1.0\ \mu\text{m}}{L} \cdot \left(\sqrt{1 + \frac{2 \cdot 0.314\ \mu\text{m}}{1.0\ \mu\text{m}}} - 1 \right)$$

Finally, the zero-bias threshold voltage is found as

$$V_{T0}(\textit{short channel}) = 0.855 \text{ V} - 0.19 \text{ V} \cdot \frac{1}{L[\mu\text{m}]}$$

The following plot shows the variation of the threshold voltage with the channel length. The threshold voltage decreases by as much as 50% for channel lengths in the submicron range, while it approaches the value of 0.8 V for larger channel lengths.



Example 6 (3)

Since the conventional threshold voltage expression (3.23) is not capable of accounting for this drastic reduction of V_{T0} at smaller channel lengths, its application for short-channel MOSFETs must be carefully restricted.

Now, consider the variation of the threshold voltage with the applied drain-to-source voltage. Equation (3.82) shows that the depth of the drain junction depletion region increases with the voltage V_{DS}. For a drain-to-source voltage of $V_{DS} = 5$ V, the drain depletion depth is found as:

$$x_{dD} = \sqrt{\frac{2 \cdot \varepsilon_{Si}}{q \cdot N_A} \cdot (\phi_0 + V_{DS})}$$

$$= \sqrt{\frac{2 \cdot 11.7 \cdot 8.85 \times 10^{-14}}{1.6 \times 10^{-19} \cdot 10^{16}} \cdot (0.76 + 5.0)} = 0.863\ \mu\text{m}$$

The resulting threshold voltage shift can be calculated by substituting x_{dD} found above in (3.88).

$$\Delta V_{T0} = \frac{1}{C_{ox}} \cdot \sqrt{2q\varepsilon_{Si} N_A |2\phi_F|} \cdot \frac{x_j}{2L} \cdot \left[\left(\sqrt{1 + \frac{2x_{dS}}{x_j}} - 1 \right) + \left(\sqrt{1 + \frac{2x_{dD}}{x_j}} - 1 \right) \right]$$

$$= \frac{4.82 \times 10^{-8}}{7.03 \times 10^{-8}} \cdot \frac{1.0}{2 \cdot 0.7} \cdot \left[\left(\sqrt{1 + \frac{2 \cdot 0.314}{1.0}} - 1 \right) + \left(\sqrt{1 + \frac{2 \cdot 0.863}{1.0}} - 1 \right) \right]$$

$$= 0.45\ V$$

The threshold voltage of this short-channel MOS transistor is calculated as

$$V_{T0} = 0.855\ \text{V} - 0.45\ \text{V} = 0.405\ \text{V}$$

which is significantly lower than the threshold voltage predicted by the conventional long-channel formula (3.23).

Narrow-channel effect

- Channel width W on the same order of magnitude as the maximum depletion region thickness x_{dm}
- The actual threshold voltage of such device is larger than that predicted by the conventional threshold voltage
- Fringe depletion region under field oxide

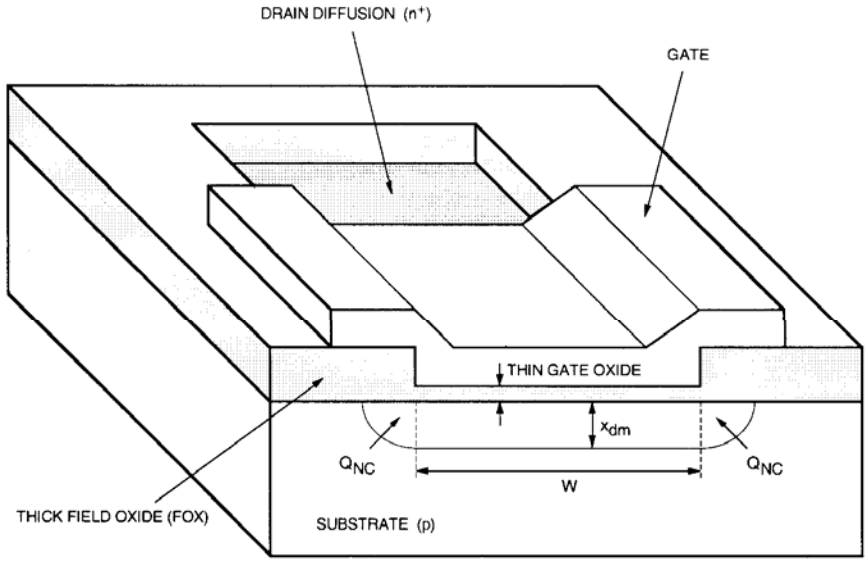


Figure 3.26 Cross-sectional view (across the channel) of a narrow-channel MOSFET. Note that Q_{NC} indicates the extra depletion charge due to narrow-channel effects.

- $V_{T0}(\text{narrow channel}) = \mathrm{V_{T0}} + \Delta \mathrm{V_{T0}}$

$$\Delta \mathrm{V_{T0}} = \frac{1}{C_{ox}} \cdot \sqrt{2q\varepsilon_{Si} N_A \left|2\phi_F\right|} \cdot \frac{\kappa \cdot x_{dm}}{W}$$

$$\kappa = \frac{\pi}{2} \text{ for depletion region modeled by quarter - circular arcs}$$

Other limitations imposed by small-device geometries

- The current flow in the channel are controlled by two dimensional electric field vector
- Subthreshold conduction
  - Drain-induced barrier lowering (DIBL)
  - A nonzero drain current I_D for $V_{GS}<V_{T0}$
  - $$I_D(subthreshold) \cong \frac{qD_n W x_c n_0}{L_B} \cdot e^{\frac{q\phi_r}{kT}} \cdot e^{\frac{q}{kT}(A \cdot V_{GS} + B \cdot V_{DS})}$$
- Punch-through
  - The gate voltage loses its control upon the drain current, and the current rises sharply
- Gate oxide thickness t_{ox} scaled to t_{ox}/S, is restricted by processing difficulties
  - Pinholes, oxide breakdown
- Hot-carrier effect

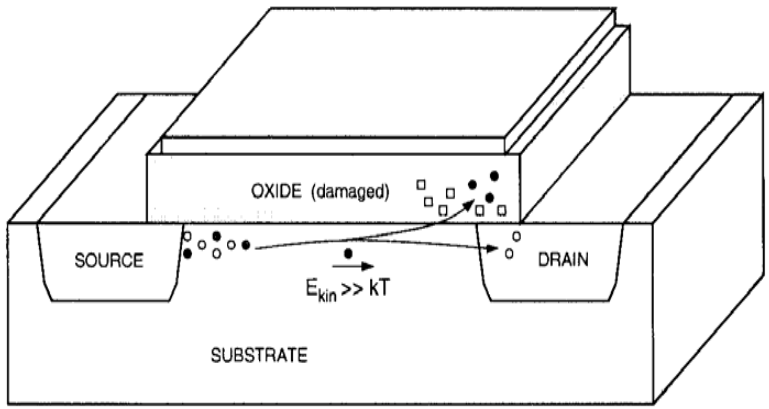


Figure 3.27 Hot-carrier injection into the gate oxide and resulting oxide damage.

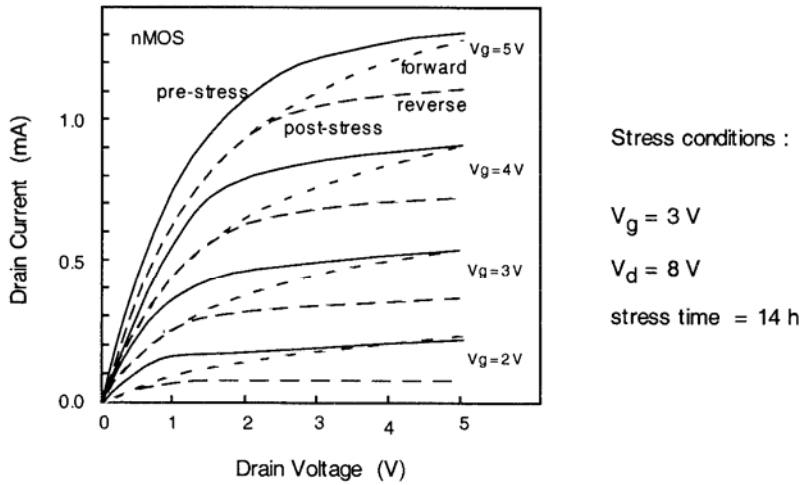


Figure 3.28 Typical drain current vs. drain voltage characteristics of an n-channel MOS transistor before and after hot-carrier induced oxide damage.

MOSFET capacitances

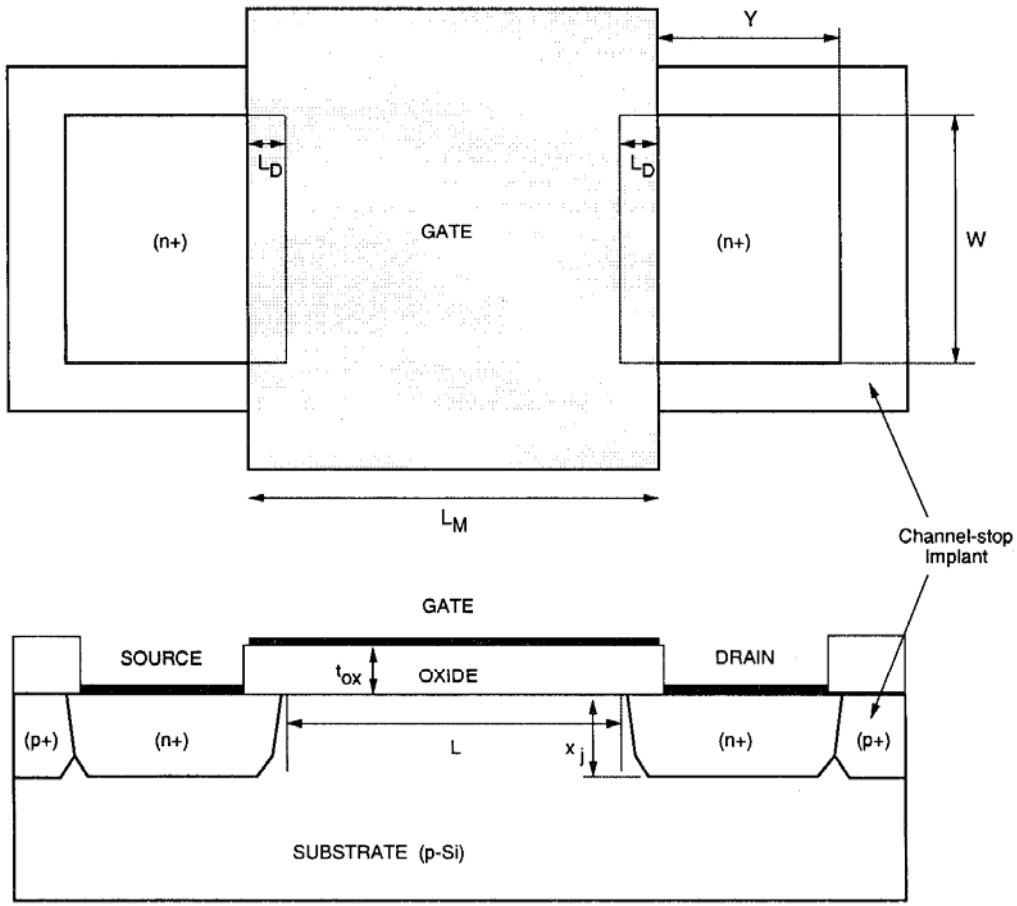


Figure 3.29 Cross-sectional view and the top view (mask view) of a typical n-channel MOSFET.

- $L=L_M-2L_D$
  - L: the actual channel length
  - L_M: the mask length of the gate
  - L_D: the gate-drain, the gate-source overlap
    - On the order of 0.1μm

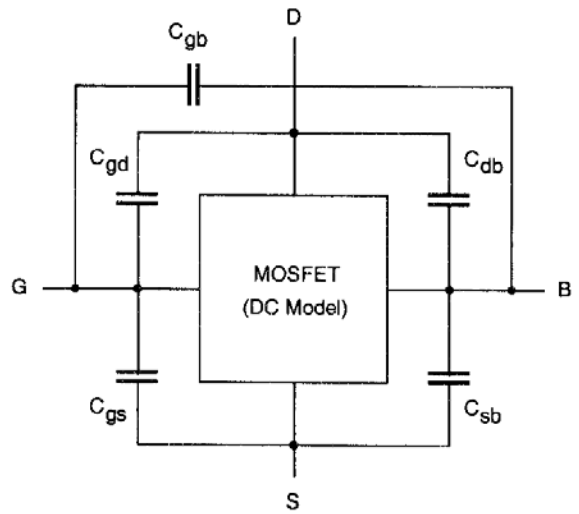


Figure 3.30 Lumped representation of the parasitic MOSFET capacitances.

Oxide related capacitance(1)

- The gate electrode overlap capacitance
  - $C_{GD(overlap)}=C_{ox}WL_D$
  - $C_{GS(overlap)}=C_{ox}WL_D$
    - *With* $C_{ox}=\varepsilon_{ox}/t_{ox}$
  - Both capacitance do not depend on the bias condition, they are voltage-independent
- The capacitances result from the interaction between the gate voltage and the channel charge
  - Cut-off mode
    - $C_{gs}=C_{gd}=0$
    - $C_{gb}=C_{ox}WL$
  - Linear mode
    - $C_{gb}=0$
    - $C_{gs}\cong C_{gd}\cong(1/2)\ C_{ox}WL$
  - Saturation mode
    - $C_{gb}= C_{gd} =0$
    - $C_{gs}\cong(2/3)\ C_{ox}WL$

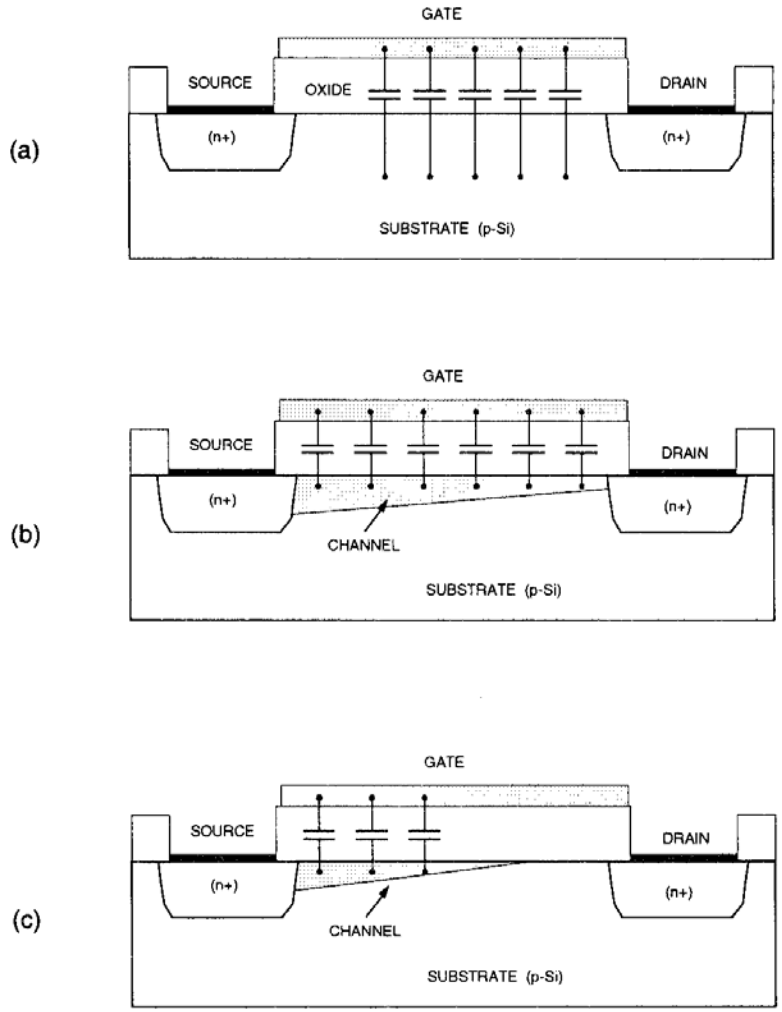


Figure 3.31 Schematic representation of MOSFET oxide capacitances during (a) cut-off, (b) linear, and (c) saturation modes.

Oxide related capacitance(2)

- The sum of all three voltage-dependent (distributed) gate oxide capacitances ($C_{gb}+C_{gs}+C_{gd}$)
  - A minimum value of $0.66C_{ox}WL$, in saturation mode
  - A maximum value of $C_{ox}WL$, in cut off and linear modes
  - For simple hand calculation
    - The three capacitances can be considered to be in parallel
    - A constant worst-case value of $C_{ox}W(L+2L_D)$ can be used for the sum of MOSFET gate oxide capacitances

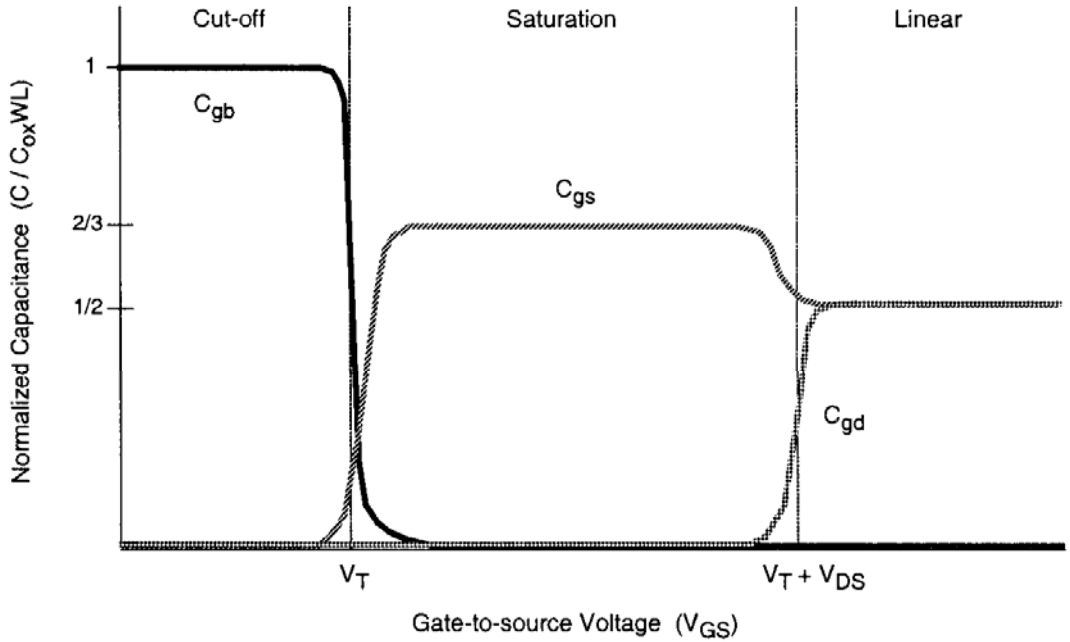


Figure 3.32 Variation of the distributed (gate-to-channel) oxide capacitances as functions of gate-to-source voltage V_{GS}.

Junction capacitance(1)

The depletion region thickness $x_d = \sqrt{\frac{2\cdot\varepsilon_{Si}}{q}\cdot\frac{N_A+N_D}{N_A\cdot N_D}(\varphi_0 - V)}$

The built - in potential $\varphi_0 = \frac{kT}{q}\cdot\ln\left(\frac{N_A\cdot N_D}{n_i^2}\right)$

The depletion region charge $Q_j = A\cdot q\cdot\left(\frac{N_A\cdot N_D}{N_A+N_D}\right)\cdot x_d = A\sqrt{2\cdot\varepsilon_{Si}\cdot q\cdot\frac{N_A\cdot N_D}{N_A+N_D}(\varphi_0 - V)}$

The junction capacitance $C_j = \left|\frac{dQ_j}{dV}\right| = A\cdot\sqrt{\frac{\varepsilon_{Si}\cdot q}{2}\cdot\left(\frac{N_A\cdot N_D}{N_A+N_D}\right)}\cdot\frac{1}{\sqrt{\varphi_0 - V}}$

$C_j(V) = \frac{AC_{j0}}{\left(1-\frac{V}{\varphi_0}\right)^m}$, the parameter m is grading coefficient

The zero bias junction capacitance per unit area $C_{j0} = \sqrt{\frac{\varepsilon_{Si}\cdot q}{2}\cdot\left(\frac{N_A\cdot N_D}{N_A+N_D}\right)\cdot\frac{1}{\varphi_0}}$

The equivalent large - signal capacitance can be defined as

$$C_{eq} = \frac{\Delta Q}{\Delta V} = \frac{Q_j(V_2) - Q_j(V_1)}{V_2 - V_1} = \frac{1}{V_2 - V_1}\int_{V_1}^{V_2} C_j(V)dV$$

$$= -\frac{A\cdot C_{j0}\cdot\phi_0}{(V_2 - V)\cdot(1-m)}\cdot\left[\left(1-\frac{V_2}{\phi_0}\right)^{1-m} - \left(1-\frac{V_1}{\phi_0}\right)^{1-m}\right]$$

For the special case of abrupt pn - junctions

$$C_{eq} = -\frac{2\cdot A\cdot C_{j0}\cdot\varphi_0}{(V_2 - V)}\cdot\left[\sqrt{1-\frac{V_2}{\phi_0}} - \sqrt{1-\frac{V_1}{\phi_0}}\right]$$

$$C_{eq} = A\cdot C_{j0}\cdot K_{eq}$$

$$K_{eq} = -\frac{2\sqrt{\phi_0}}{V_2 - V_1}\cdot\left(\sqrt{\phi_0 - V_2} - \sqrt{\phi_0 - V_1}\right)$$

Table 3.7 Types and areas of the pn-junctions shown in Fig. 3.33

Junction	Area	Type
1	$W\cdot x_j$	n^+/p
2	$Y\cdot x_j$	n^+/p^+
3	$W\cdot x_j$	n^+/p^+
4	$Y\cdot x_j$	n^+/p^+
5	$W\cdot Y$	n^+/p

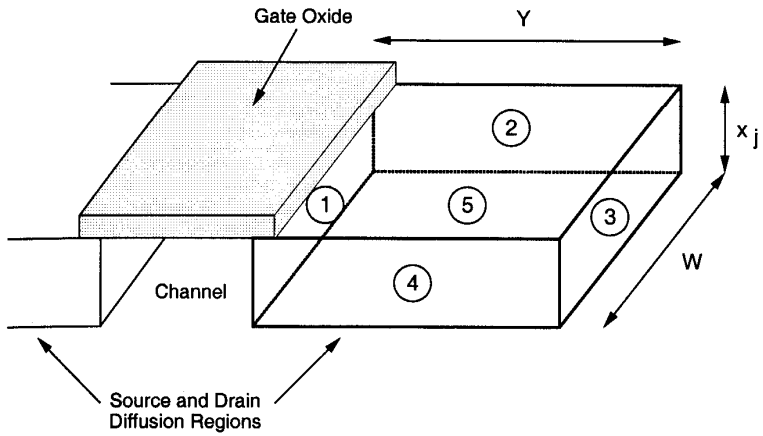
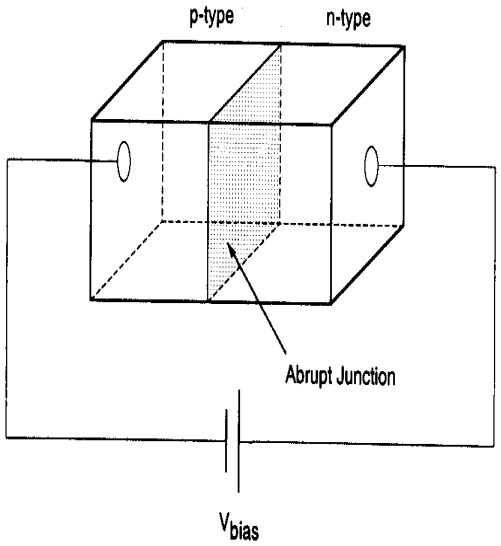


Figure 3.33 Three-dimensional view of the n^+ diffusion region within the p-type substrate.

Example 7

Consider a simple abrupt pn-junction, which is reverse-biased with a voltage V_{bias}. The doping density of the n-type region is $N_D = 10^{19}$ cm^{-3}, and the doping density of the p-type region is given as $N_A = 10^{16}$ cm^{-3}. The junction area is $A = 20\,\mu$m $\times$ $20\,\mu$m.



First, we will calculate the zero-bias junction capacitance per unit area, C_{j0}, for this structure. The built-in junction potential is found as

$$\phi_0 = \frac{kT}{q} \cdot \ln\left(\frac{N_A \cdot N_D}{n_i^2}\right) = 0.026\ \text{V} \cdot \ln\left(\frac{10^{16} \cdot 10^{19}}{2.1 \times 10^{20}}\right) = 0.88\ \text{V}$$

Using (3.105), we can calculate the zero-bias junction capacitance:

$$C_{j0} = \sqrt{\frac{\varepsilon_{Si} \cdot q}{2} \cdot \left(\frac{N_A \cdot N_D}{N_A + N_D}\right) \cdot \frac{1}{\phi_0}}$$

$$= \sqrt{\frac{11.7 \cdot 8.85 \times 10^{-14}\ \text{F/cm} \cdot 1.6 \times 10^{-19}\ \text{C}}{2} \cdot \left(\frac{10^{16} \cdot 10^{19}}{10^{16} + 10^{19}}\right) \cdot \frac{1}{0.88\ \text{V}}}$$

$$= 3.1 \times 10^{-8}\ \text{F/cm}^2$$

Next, find the equivalent large-signal junction capacitance assuming that the reverse bias voltage changes from $V_1 = 0$ to $V_2 = -5$ V. The voltage equivalence factor for this transition can be found as follows:

$$K_{eq} = -\frac{2\sqrt{\phi_0}}{V_2 - V_1} \cdot \left(\sqrt{\phi_0 - V_2} - \sqrt{\phi_0 - V_1}\right)$$

$$= -\frac{2\sqrt{0.88}}{-5} \cdot \left(\sqrt{0.88 - (-5)} - \sqrt{0.88}\right) = 0.56$$

Then, the average junction capacitance can be found simply by using (3.109).

$$C_{eq} = A \cdot C_{j0} \cdot K_{eq} = 400 \times 10^{-8}\ \text{cm}^2 \cdot 3.1 \times 10^{-8}\ \text{F/cm}^2 \cdot 0.56 = 69\ \text{fF}$$

Junction capacitance(2)

The sidewalls of a typical MOSFET source or drain diffusion region are surrounded by a p^+ channel-stop implant, with a higher doping density than the substrate doping density N_A

Assume the sidewall doping density is given by $N_{A(sw)}$,

the zero-bias capacitance per unit area can be found as

$$C_{j0sw} = \sqrt{\frac{\varepsilon_{Si} \cdot q}{2} \cdot \left(\frac{N_{A(sw)} \cdot N_D}{N_{A(sw)} + N_D} \right) \cdot \frac{1}{\varphi_{0sw}}}$$

$$C_{jsw} = C_{j0sw} \cdot x_j$$

The sidewall voltage equivalence factor

$$K_{eq(sw)} = -\frac{2\sqrt{\phi_{0sw}}}{V_2 - V_1} \cdot \left(\sqrt{\phi_{0sw} - V_2} - \sqrt{\phi_{0sw} - V_1} \right)$$

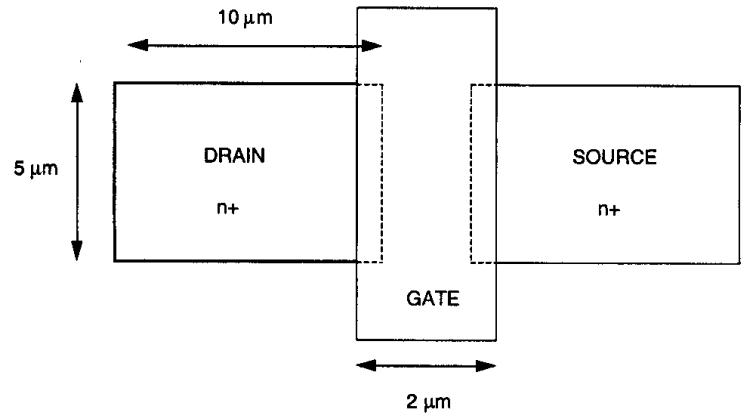
The equivalent large-signal junction capacitance $C_{eq(sw)}$ for a sidewall of length (perimeter) P can be

$$C_{eq(sw)} = P \cdot C_{jsw} \cdot K_{eq(sw)}$$

Example 8 (1)

Consider the n-channel enhancement-type MOSFET shown below. The process parameters are given as follows:

Substrate doping	$N_A = 2 \times 10^{15}\ \text{cm}^{-3}$
Source/drain doping	$N_D = 10^{19}\ \text{cm}^{-3}$
Sidewall (p^+) doping	$N_A(sw) = 4 \times 10^{16}\ \text{cm}^{-3}$
Gate oxide thickness	$t_{ox} = 45\ \text{nm}$
Junction depth	$x_j = 1.0\ \mu\text{m}$



Note that both the source and the drain diffusion regions are surrounded by p^+ channel-stop diffusion. The substrate is biased at 0 V. Assuming that the drain voltage is changing from 0.5 V to 5 V, find the average drain-substrate junction capacitance C_{db}.

First, we recognize that three sidewalls of the rectangular drain diffusion structure form n^+/p^+ junctions with the p^+ channel-stop implant, while the bottom area and the sidewall facing the channel form n^+/p junctions. Start by calculating the built-in potentials for both types of junctions.

$$\phi_0 = \frac{kT}{q} \cdot \ln\left(\frac{N_A \cdot N_D}{n_i^2}\right) = 0.026\ \text{V} \cdot \ln\left(\frac{2 \times 10^{15} \cdot 10^{19}}{2.1 \times 10^{20}}\right) = 0.837\ \text{V}$$

$$\phi_{0sw} = \frac{kT}{q} \cdot \ln\left(\frac{N_A(sw) \cdot N_D}{n_i^2}\right) = 0.026\ \text{V} \cdot \ln\left(\frac{4 \times 10^{16} \cdot 10^{19}}{2.1 \times 10^{20}}\right) = 0.915\ \text{V}$$

Example 8 (2)

Next, we calculate the zero-bias junction capacitances per unit area:

$$C_{j0} = \sqrt{\frac{\varepsilon_{Si} \cdot q}{2} \cdot \left(\frac{N_A \cdot N_D}{N_A + N_D}\right) \cdot \frac{1}{\phi_0}}$$

$$= \sqrt{\frac{11.7 \cdot 8.85 \times 10^{-14}\ \text{F/cm} \cdot 1.6 \times 10^{-19}\ \text{C}}{2} \cdot \left(\frac{2 \times 10^{15} \cdot 10^{19}}{2 \times 10^{15} + 10^{19}}\right) \cdot \frac{1}{0.837\ \text{V}}}$$

$$= 1.41 \times 10^{-8}\ \text{F/cm}^2$$

$$C_{j0sw} = \sqrt{\frac{\varepsilon_{Si} \cdot q}{2} \cdot \left(\frac{N_A(sw) \cdot N_D}{N_A(sw) + N_D}\right) \cdot \frac{1}{\phi_{0sw}}}$$

$$= \sqrt{\frac{11.7 \cdot 8.85 \times 10^{-14}\ \text{F/cm} \cdot 1.6 \times 10^{-19}\ \text{C}}{2} \cdot \left(\frac{4 \times 10^{16} \cdot 10^{19}}{4 \times 10^{16} + 10^{19}}\right) \cdot \frac{1}{0.915\ \text{V}}}$$

$$= 6.01 \times 10^{-8}\ \text{F/cm}^2$$

The zero-bias sidewall junction capacitance per unit length can also be found as follows.

$$C_{jsw} = C_{j0sw} \cdot x_j = 6.01 \times 10^{-8}\ \text{F/cm}^2 \cdot 10^{-4}\ \text{cm} = 6.01\ \text{pF/cm}$$

In order to take the given drain voltage variation into account, we must now calculate the voltage equivalence factors, K_{eq} and $K_{eq}(sw)$, for both types of junctions. This will allow us to find the average large-signal capacitance values.

$$K_{eq} = -\frac{2\sqrt{0.837}}{-5-(-0.5)} \cdot \left(\sqrt{0.837+5} - \sqrt{0.837+0.5}\right) = 0.51$$

$$K_{eq}(sw) = -\frac{2\sqrt{0.915}}{-5-(-0.5)} \cdot \left(\sqrt{0.915+5} - \sqrt{0.915+0.5}\right) = 0.53 \cong K_{eq}$$

The total area of the n^+/p junctions is calculated as the sum of the bottom area and the sidewall area facing the channel region.

$$A = (10 \times 5)\ \mu\text{m}^2 + (5 \times 1)\ \mu\text{m}^2 = 55\ \mu\text{m}^2$$

The total length of the n^+/p^+ junction perimeter, on the other hand, is equal to the sum of three sides of the drain diffusion area. Thus, the combined equivalent (average) drain-substrate junction capacitance can be found as follows:

$$\langle C_{db} \rangle = A \cdot C_{j0} \cdot K_{eq} + P \cdot C_{jsw} \cdot K_{eq}(sw)$$

$$= 55 \times 10^{-8}\ \text{cm}^2 \cdot 1.41 \times 10^{-8}\ \text{F/cm}^2 \cdot 0.51$$

$$+ 25 \times 10^{-4}\ \text{cm} \cdot 6.01 \times 10^{-12}\ \text{F/cm} \cdot 0.53 = 11.9 \times 10^{-15}\ \text{F} = \underline{\underline{11.9\ \text{fF}}}$$