Lecture 12:

MOSFET Devices

Gu-Yeon Wei
Division of Engineering and Applied Sciences
Harvard University
guyeon@eecs.harvard.edu

Overview

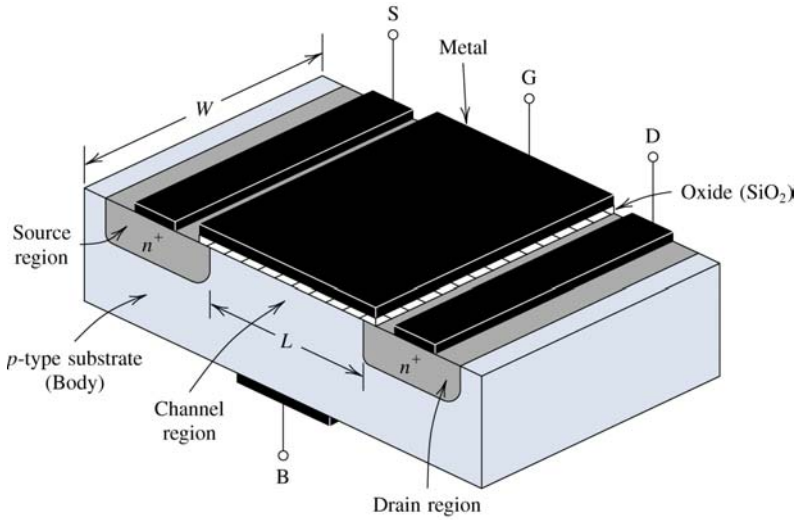
- **Reading**
  - S&S: Chapter 5.1~5.4
- **Supplemental Reading**

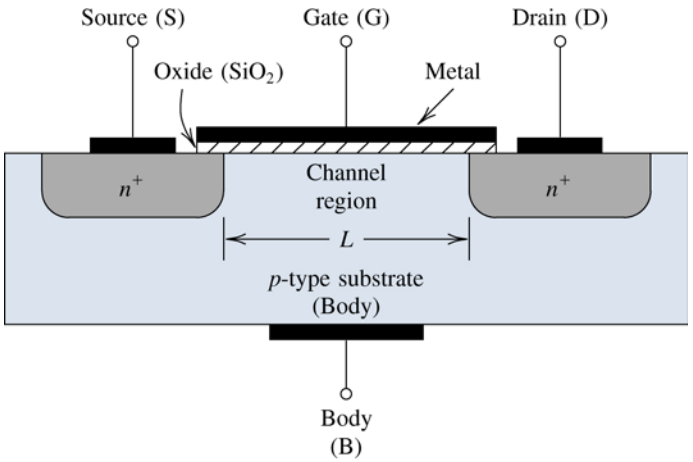
- **Background**
  - We are now going to switch gears and look at a different of transistor device called a MOSFET. Most modern ICs are built using these transistors. While they are commonly used to implement digital circuits, we will look at their analog characteristics and talk about how to build amplifiers equivalent to ones built with BJTs.

    We begin with the physical structure and a qualitative understanding of how MOSFETs operate. We will derive some current-voltage equations for the transistor. We will also use band diagrams to provide some theoretical rigor to our initial qualitative understanding. Then, we will look at some non-ideal characteristics of the transistor. Lastly, we will analyze the DC operation of MOSFETs.

Enhancement-Type MOSFET

- Most widely used field effect transistor (enhancement type)
- Let's look at its structure and physical operation

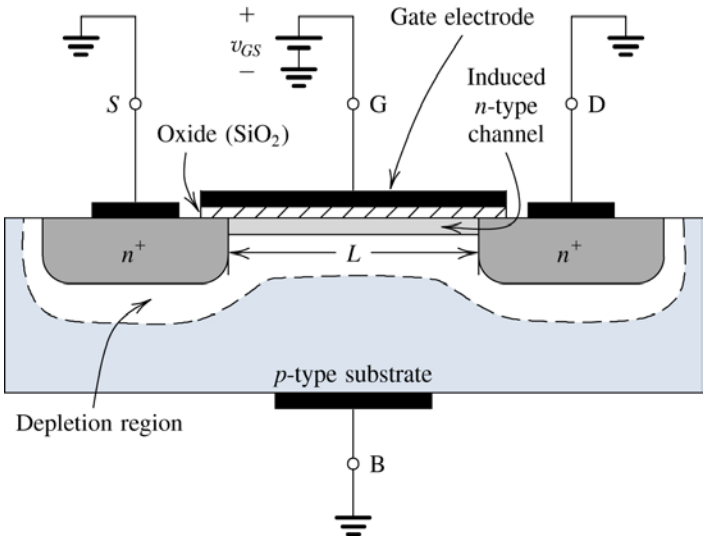




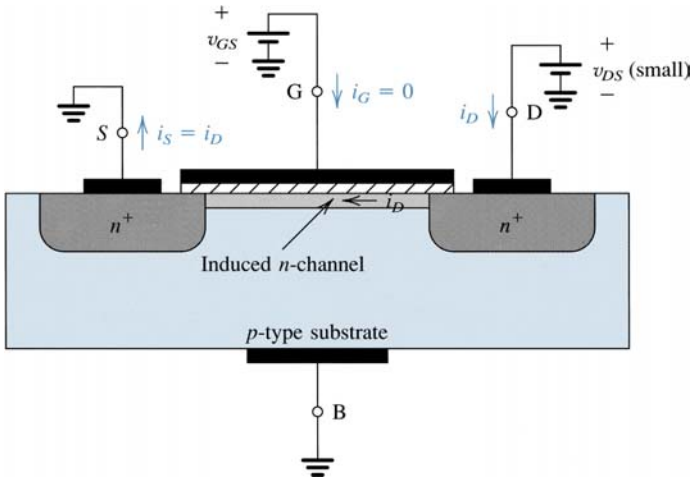
  - 3 terminal device like the BJT but different names
  - Additional body (or bulk) terminal (generally at DC and not used for signals)
  - No connection between the gate and drain/source (separated by oxide)
  - Voltage on gate controls current flow between source and drain

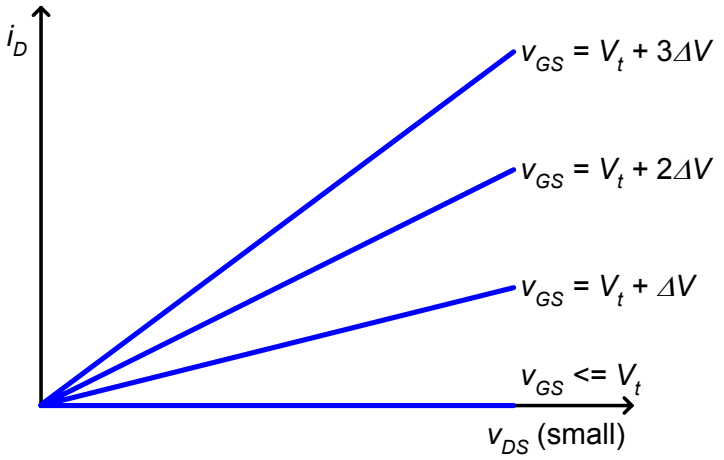
Device Operation

- No gate voltage (v_{GS} = 0)
  - Two back to back diodes both in reverse bias
  - no current flow between source and drain when voltage between source and drain is applied (v_{DS} >0)
  - There is a depletion region between the p (substrate) and n⁺ source and drain regions
- Apply a voltage on $v_{GS} > 0$
  - Positive potential on gate node pushes free holes away from the region underneath the gate and leave behind a negatively charged carrier depletion region
    - transistor in depletion mode
  - As v_{GS} increases, electrons start to gather at the surface underneath the gate (onset of inversion)
  - When v_{GS} is high enough, a n-type channel is induced underneath the gate oxide with electrons supplied by the n⁺ source and drain regions
    - This induced region is called an inversion layer (or channel) and forms when v_{GS} > some **threshold voltage** V_t and current can flow between S & D
    - Transistor is in inversion mode
    - When V_{DS} = 0, no current flows between source and drain



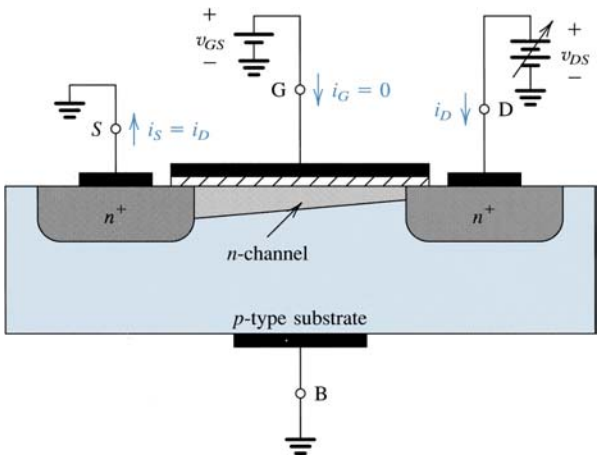
Linear Operation

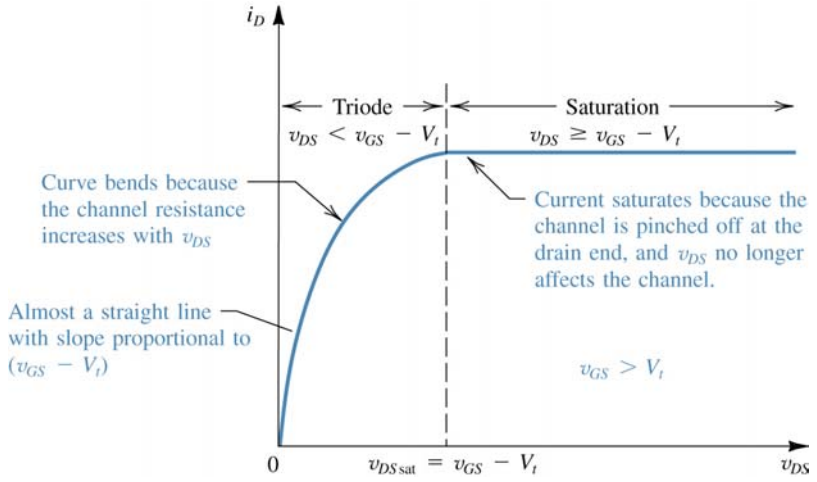




- With v_{GS} large enough to induce a channel, apply a small potential v_{DS}
  - Causes current to flow between source and drain (electrons flow from source to drain)
  - Magnitude of i_D depends on density of electrons in channel which depends on v_{GS} (larger v_{GS} = higher density of electrons)
  - Conductance of channel is proportional to v_{GS}-V_t (called excess gate voltage or effective voltage or gate overdrive)
  - Current is proportional to v_{GS}-V_t and v_{DS} that causes current to flow
  - i-v curve shows the transistor operates like a voltage-controlled linear resistor
- Notice $i_D = i_S$ and $i_G = 0$ due to the gate oxide

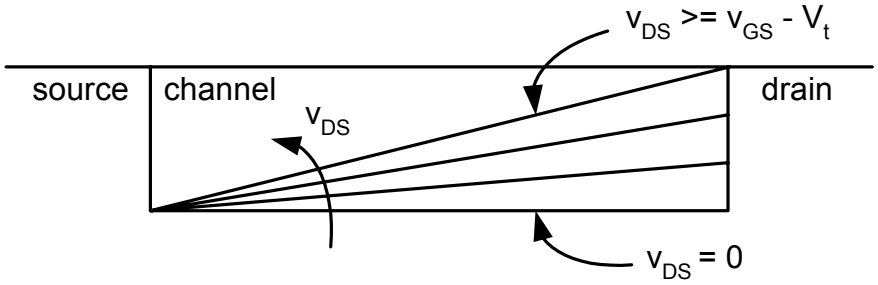
Triode to Saturation Region

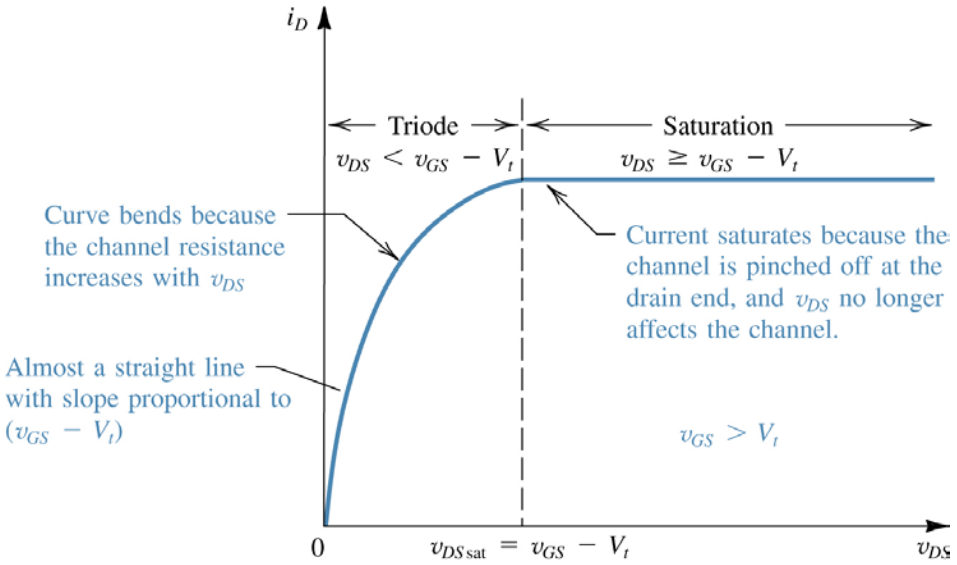




- Assume v_{GS} is at a constant value > V_t and increase v_{DS}
  - v_{DS} appears as a voltage drop across the channel and at different points along the channel, the voltage is different
  - Voltages between the gate and points along the channel are also different ranging from v_{GS} at the source to v_{GS}-v_{DS} at the drain
    - Induced channel is a function of voltage across the oxide at the different points and so channel depth varies across the length of the transistor
  - i-v curve bends over as v_{DS} increases due to the smaller channel depth
  - At v_{DS} = v_{GS}-V_t channel depth is almost zero at the drain side
    - Current stays flat for higher voltages $v_{DS} > v_{GS}$-V_t
    - The transistor is said to now operate in the saturation region (not to be confused with the saturation region in BJTs)

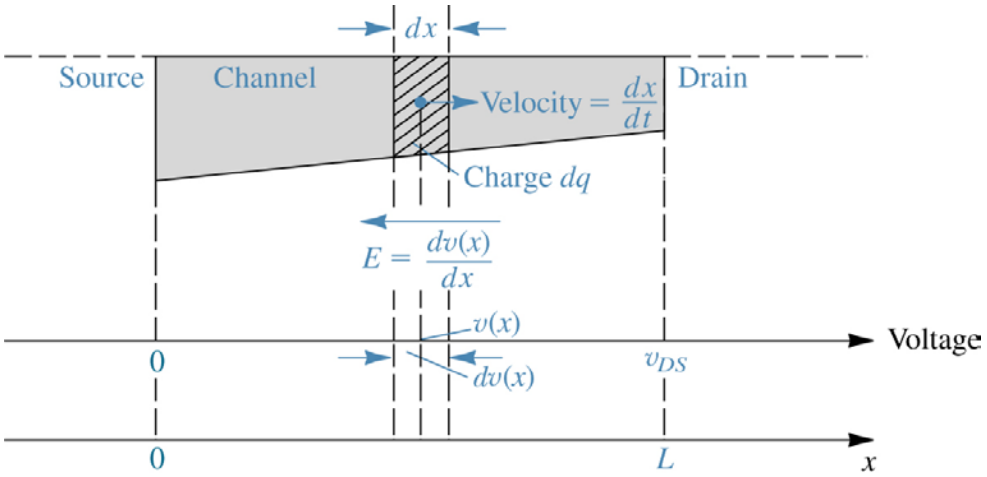
Saturation Region





- As v_{DS} increases, the channel gets smaller and smaller on the drain side until $v_{DS} = v_{GS} - V_t$ at which point the channel is said to be **pinched off**
  - Increasing v_{DS} beyond this point as little (ideally no) effect on the channel shape
  - Current remains constant and said to saturate
  - Transistor enters saturation at $v_{DSsat} = v_{GS} - V_t$

Deriving the i_D-v_{DS} Relationship



- First consider the linear (triode) region of operation $v_{DS} < v_{GS} - V_t$ ($v_{GS} > V_t$ is assumed)
  - Consider a point along the channel of infinitesimal width dx at x and voltage $v(x)$
  - The electron charge at this point is:

$$dq(x) = -C_{ox} W dx \left[v_{GS} - v(x) - V_t \right]$$

where C_{ox} is the parallel-plate cap formed by the gate electrode and channel

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

  - v_{DS} produces as electric field along the channel (in the negative x direction)

$$E(x) = -\frac{dv(x)}{dx}$$

- The electric field causes electron charge $dq(x)$ to drift with a velocity dx/dt

$$\frac{dx}{dt} = -\mu_n E(x) = \mu_n \frac{dv(x)}{dx}$$

  - Where μ_n is the electron mobility in the channel
  - Current is the movement of charge and so…

$$i_D = -dq(x)\frac{dx}{dt} = \mu_n C_{ox} W dx\left[v_{GS} - v(x) - V_t\right]\frac{dv(x)}{dx}$$

- Rearrange the equation and integrate along the length of the channel

$$i_D dx = \mu_n C_{ox} W dx\left[v_{GS} - v(x) - V_t\right]dv(x) \quad \Rightarrow \quad \int_0^L i_D dx = \int_0^{v_{DS}} \mu_n C_{ox} W dx\left[v_{GS} - v(x) - V_t\right]dv(x)$$

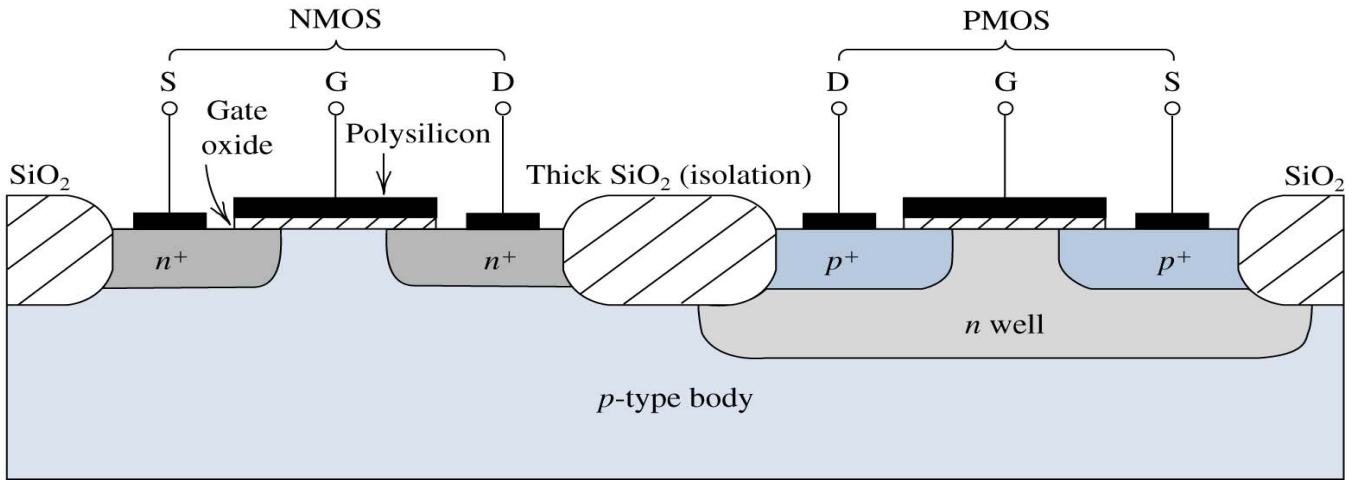
  - Gives the current in the linear (triode) region:

$$i_D = \mu_n C_{ox} \frac{W}{L}\left[\left(v_{GS} - V_t\right)v_{DS} - \frac{1}{2}{v_{DS}}^2\right]$$

  - When $v_{DS}=v_{GS}-V_t$, we get the saturation current equation

$$i_D = \frac{1}{2}\mu_n C_{ox} \frac{W}{L}\left(v_{GS} - V_t\right)^2$$

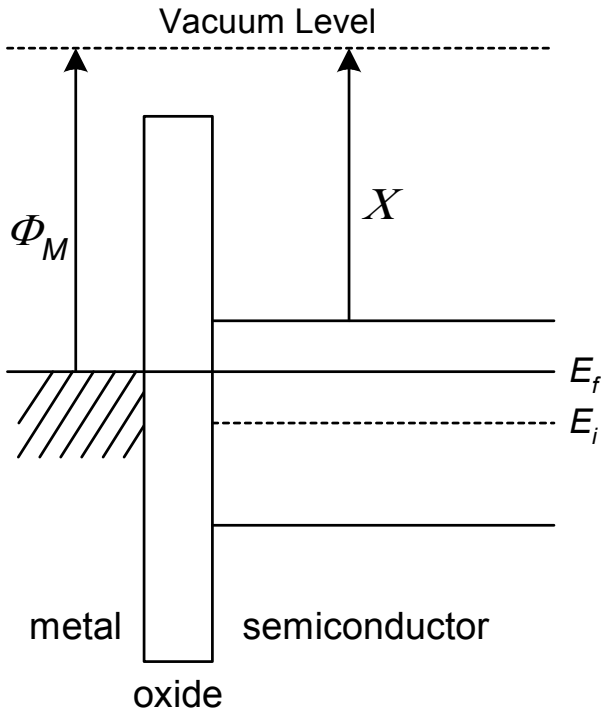
nMOS and pMOS



- We've just seen how current flows in nMOS devices. A complementary version of the nMOS device is a pMOS shown above
  - pMOS operation and current equations are the same except current is due to drift of holes
  - The mobility of holes (μ_p) is lower than the mobility of electrons (μ_n)
  - Current is lower in pMOS devices given the same dimension and voltages.

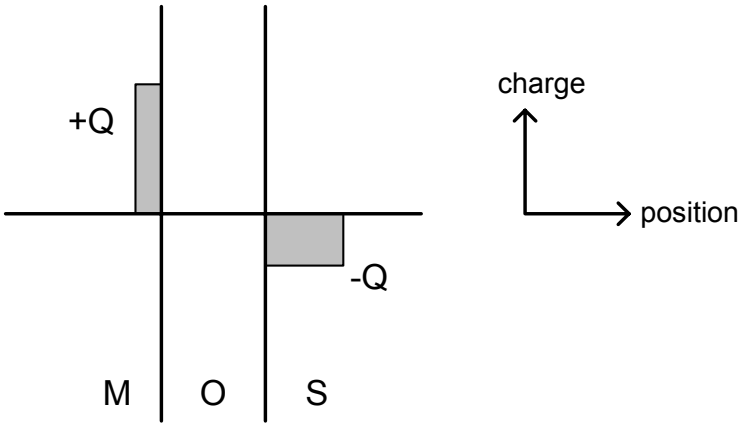
MOSFET Band Diagrams

- A more rigorous look at MOSFETs requires us to again use band diagrams
  - energy diagram drawn relative to the vacuum level at equilibrium (no voltage applied)
  - metal work function, Φ_M, energy required to completely free an electron from the metal
  - electron affinity, X, is the energy between the conduction band and vacuum level



Block Charge diagram

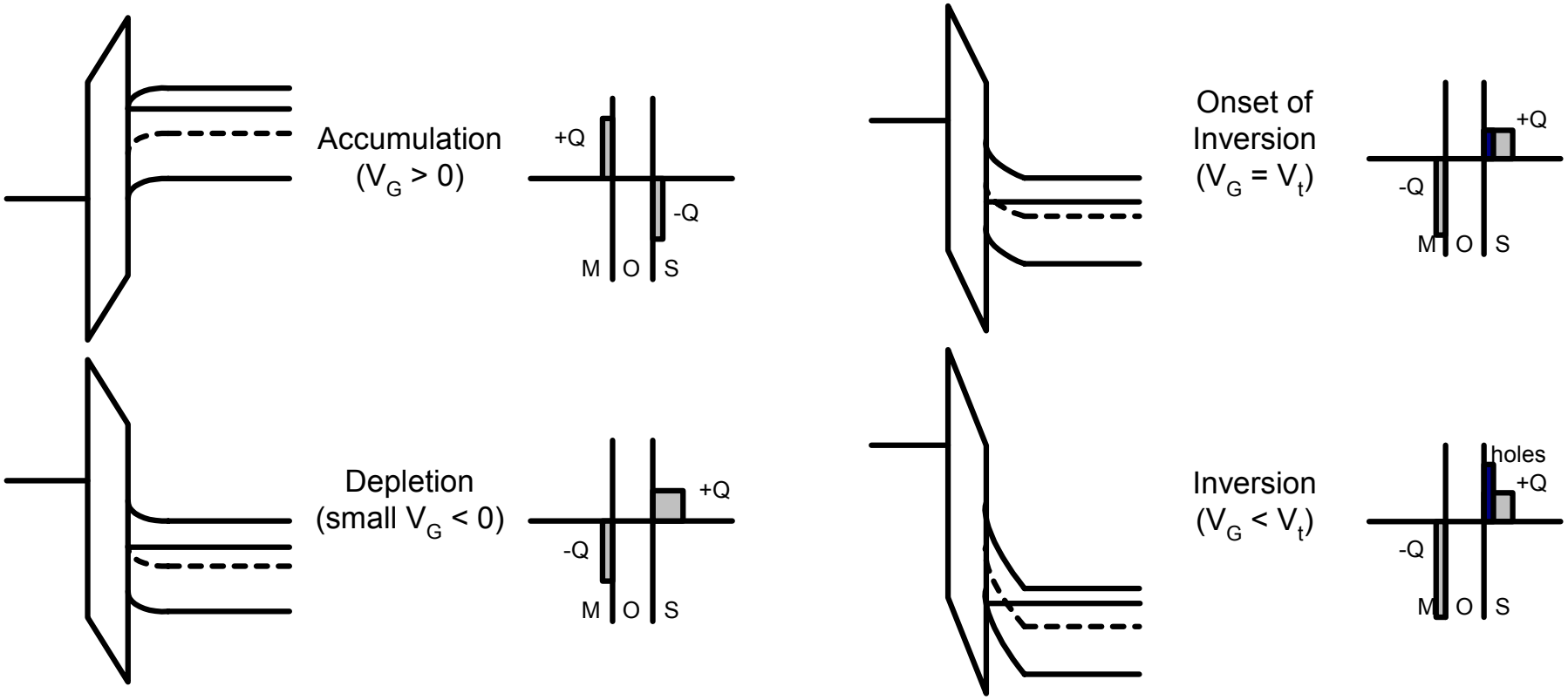
- Provides information about the charge distribution inside a MOS structure
  - no charges at equilibrium
  - when bias is applied, charge appears within the metal and semiconductor at the interfaces to the oxide
    - voltage drop across the oxide and there is an electric field due to the +Q and –Q charge separated by the oxide



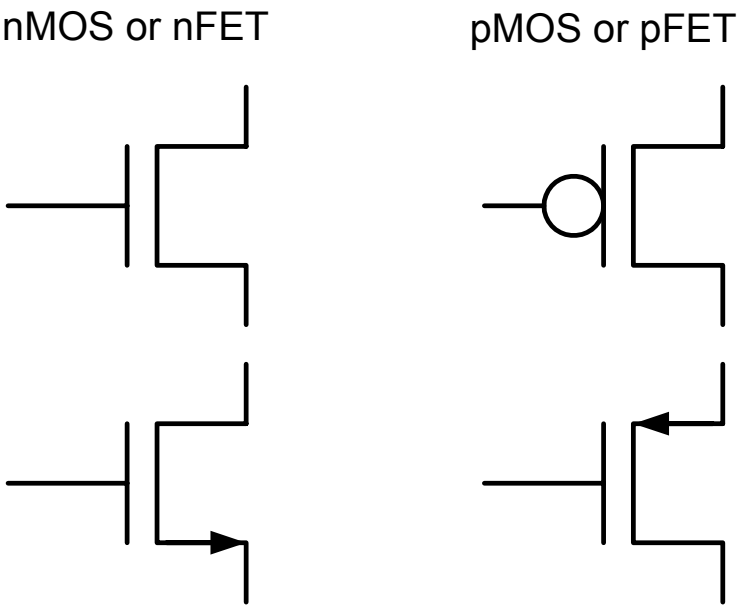
- We will use band diagrams and block charge diagrams to better understand how MOS devices work

Applying Bias

- Look at a pMOS device and see how applying a bias on the gate affects the band and block charge diagrams

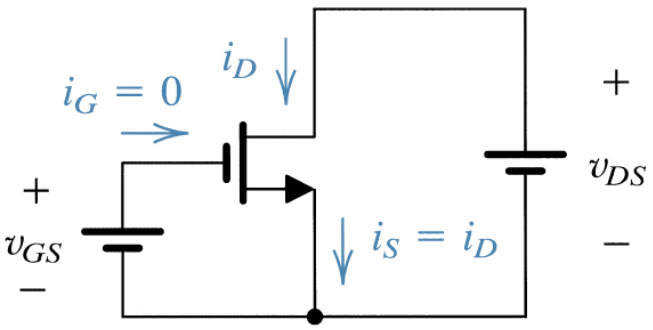


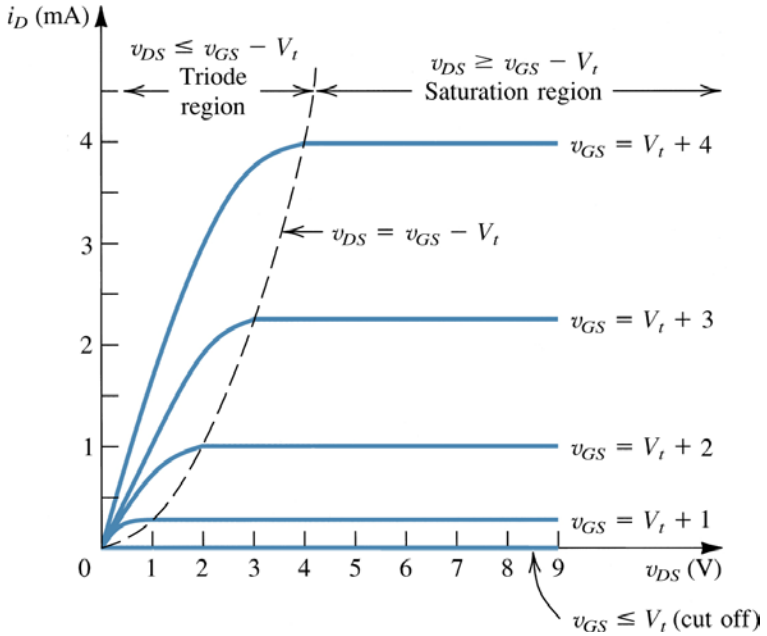
Circuit Symbols



- We represent MOSFETs with the following symbols
  - The book specifies nMOS vs. pMOS with arrows
  - I will use bubbles b/c they are easier to distinguish quickly
    - a digital circuit designers way of drawing symbols
- These are symmetric devices and so drain and source can be used interchangeably

i-v Characteristics





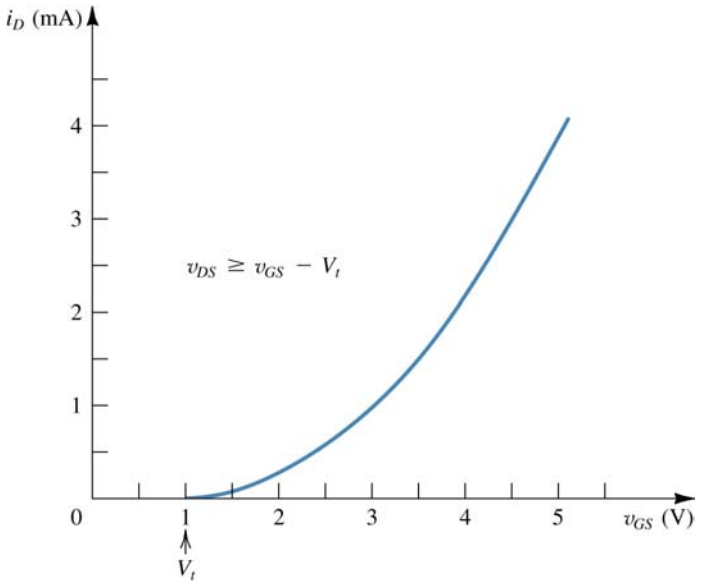
- For small values of v_{DS}, v_{DS}^2 is small and so near the origin, we can approximate the transistor as a linear resistor

$$i_D = \mu_n C_{ox} \frac{W}{L}\left[(v_{GS} - V_t)v_{DS} - \frac{1}{2}v_{DS}^{\,2}\right] \Rightarrow i_D \cong \mu_n C_{ox} \frac{W}{L}(v_{GS} - V_t)v_{DS} \text{ when } v_{DS} \text{ is small}$$

$$\Rightarrow r_{DS} \equiv \frac{vDS}{iD} = \left[\mu_n C_{ox} \frac{W}{L}(v_{GS} - V_t)\right]^{-1}$$

- We can get a relationship between i_D and v_{GS} when the transistor is in saturation
  - Let v_{GS}-V_t = V_{DS}

$$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS} - V_t)^2 \quad \Rightarrow \quad i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}v_{DS}{}^2$$

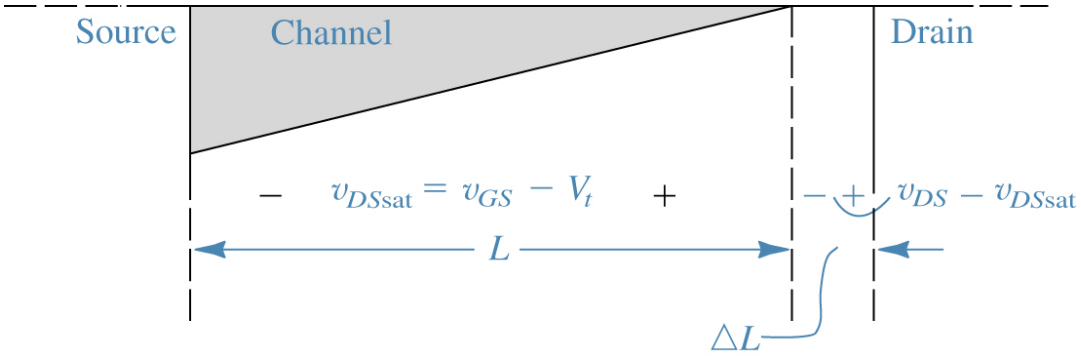


- MOS vs. BJT
  - Current is quadratic with voltage in MOS vs. exponential relationship in BJT

Some Non-Ideal Characteristics

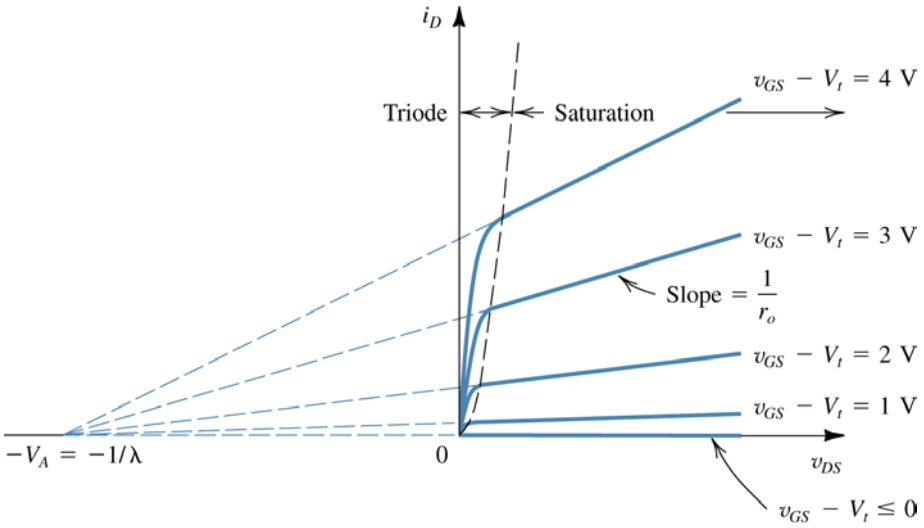
- Channel-length modulation
- Body effect
- Velocity saturation

Channel-Length Modulation



- Like the Early effect in BJTs, there is an effect in MOSFETs that causes drain current to vary with v_{DS} in saturation (finite output resistance)
- As v_{DS} increases beyond v_{DSsat}, the pinch off point moves away from the drain by ΔL and has the effect of changing the effective channel length in the transistor
  - Account for this effect with a $(1+\lambda v_{DS})$ term in the saturation current equation

$$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS} - V_t)^2(1 + \lambda v_{DS})$$



- Channel-length modulation makes the output resistance in saturation finite

$$r_o \equiv \left[\frac{\partial iD}{\partial vDS}\right]^{-1}_{v_{GS}=\text{constant}} \Rightarrow r_o = \left[\lambda \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS}-V_t)^2\right]^{-1} \cong [\lambda I_D]^{-1}$$

$$r_o = \frac{V_A}{I_D}$$

Body Effect

- So far, we have been ignoring the substrate (or bulk or body) of the transistor and assumed that is it tied to the source. However, we cannot always make that assumption.
  - In integrated circuits, body is common to many MOS transistors and channel is connected to most negative (positive) supply for nMOS (pMOS) transistors.
- The resulting reverse-bias voltage between the source and substrate affects device operation.
  - Reverse bias will widen the depletion region and reduces channel depth – which can be modeled as changing the threshold voltage

$$V_t = V_{t0} + \gamma\left[\sqrt{2\phi_f + V_{SB}} + \sqrt{2\phi_f}\right]$$

where V_{t0} is the threshold voltage when V_{SB}=0, ff is a physical parameter, γ is a fabrication-process parameter…

$$\gamma = \frac{\sqrt{2qN_A\varepsilon_s}}{C_{ox}}$$

γ is typically 0.5-$V^{1/2}$

- As V_{SB} increases, V_t increases which affects the transistor's i-v characteristics

Temperature Effects

- V_t and mobility $\mu_{n,p}$ are sensitive to temperature
  - V_t decreases by 2-mV for every 1ºC rise in temperature
  - mobility $\mu_{n,p}$ decreases with temperature
- Overall, increase in temperature results in lower drain currents

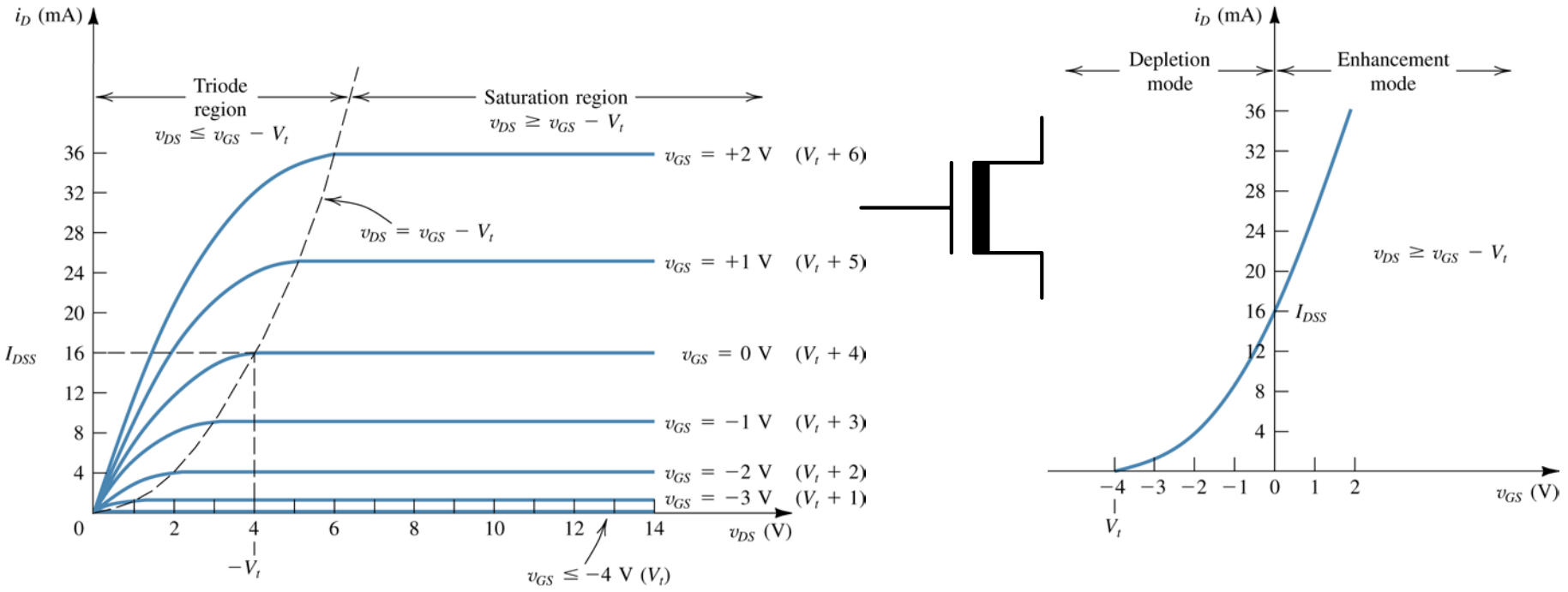
Velocity Saturation

- So far, the saturation current equation is quadratic with overdrive voltage (v_{GS}-V_t) and said to obey the "square law" which is valid for long channel length (>1-μm) devices
- As transistor dimensions decrease, gate oxide gets thinner and there is a higher vertical electrical field that the electrons moving through the channel experience
  - Causes electrons to bounce up to the oxide (more scattering) and saturates the velocity at which current flows across the channel
  - Can **approximate** the effect of velocity saturation with the following power-law equation for saturation current

$$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(v_{GS} - V_t\right)^{\alpha}$$

    α ranges from 1 to 2 depending on process technology (transistor length)
  - This approximation is not rigorous, but convenient to use. More accurate models of the velocity saturation equation can be found in the literature

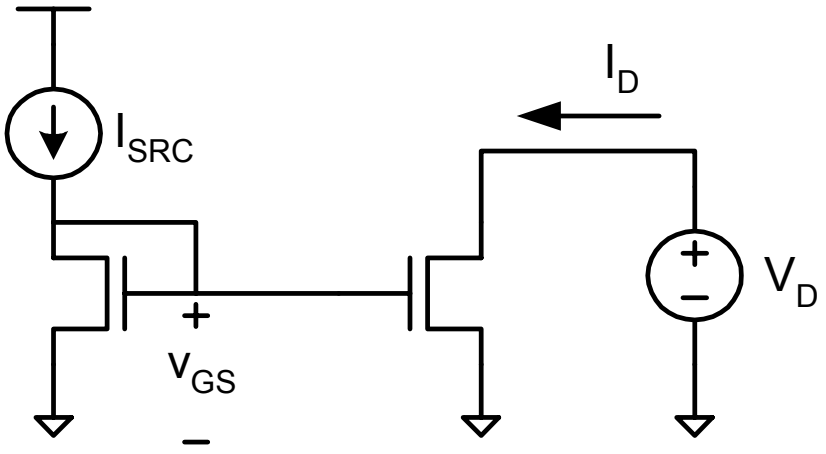
Depletion-type MOSFETs



- Depletion-type MOSFETs have a channel implant and has a channel with zero v_{GS} (symbol is drawn with channel)
  - must apply negative v_{GS} to "turn off" device

MOSFETs at DC Examples

- Example 5.1~3
- Current Mirror Example
  - What is v_{GS}?
  - How is I_D related to I_{SRC}?
  - What is I_D vs. V_D?



- Example 5.2

- Example 5.3

- Current Mirror Example