ENGR 305

MOS Field-Effect Transistors

September 23, 2025

Introduction

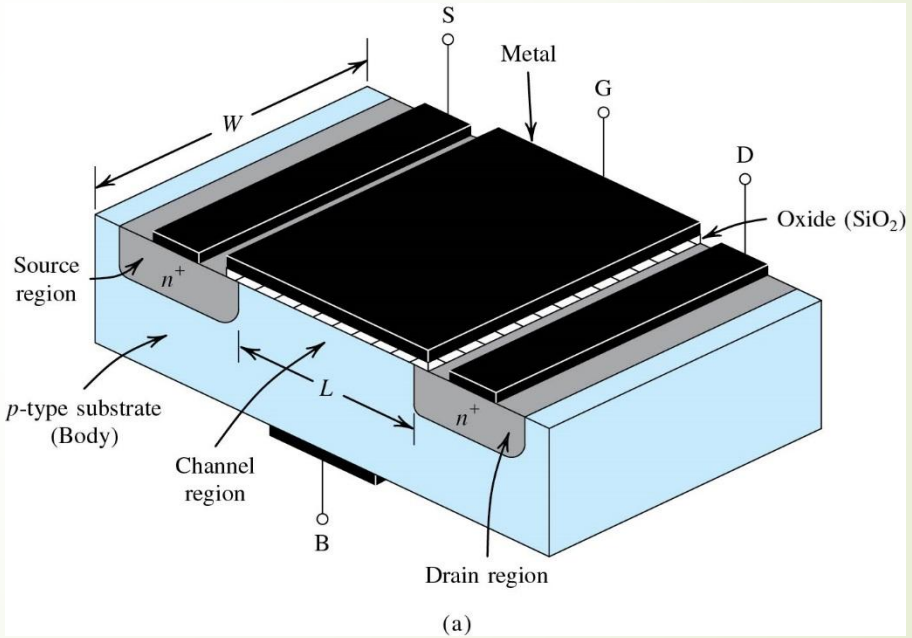
- We studied the junction diode, the basic two-terminal semiconductor device.
- Three-terminal (transistor) devices are more useful than two-terminal devices.
- They operate by using the voltage between two terminals to control the current flowing in a third terminal.
- The three-terminal device can be used as a controlled source, which is the basis for amplifier design.
- There are two types of three-terminal semiconductor devices
  - The metal-oxide-semiconductor field-effect transistor (MOSFET)
  - The bipolar junction transistor (BJT)
- We cover the first one in this chapter.

Device structure and physical operation

- Here we study the *enhancement-type MOSFET*.

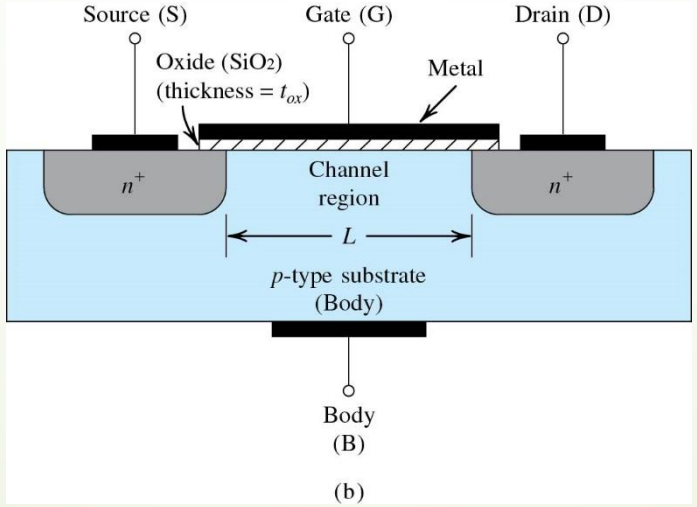
Shown is the physical structure of the *n*-channel enhancement-type MOSFET.

- p-type substrate
- n-type source and drain regions, denoted n+ because they are heavily doped
- A thin layer of silicon dioxide of thickness t_{ox} is grown on the surface of the substrate
- Metal is deposited on top of that to form the **gate electrode**



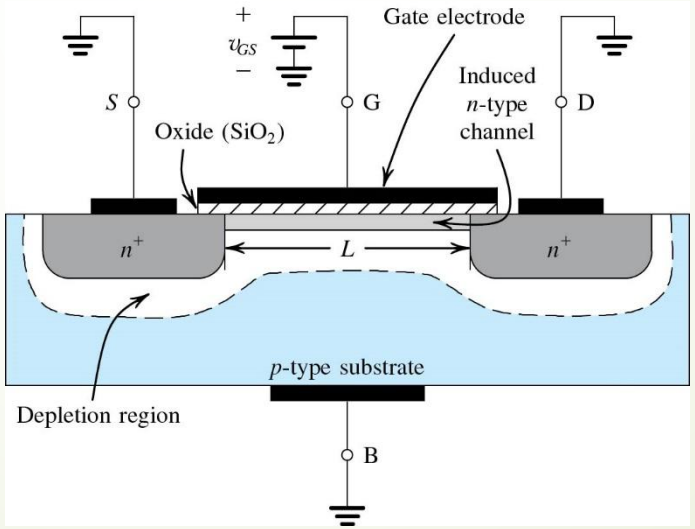
Device structure and physical operation

- Here we see the cross section of the transistor.
  - The channel length L = 20 nm to 1 μm
  - The channel width W = 30 nm to 100 μm
  - The oxide thickness t_{ox} is in the range of 1 to 10 nm



Creating channel for current flow

- Here we have grounded the source and the drain and applied a positive voltage to the gate. The voltage between gate and source is v_{GS}.
- The positive voltage on the gate causes the free holes to be repelled from the area of the substrate just under the gate, leaving behind a carrier-depletion region.



Channel for current flow

- The depletion region is populated by the bound negative charge associated with the acceptor atoms.
  - These charges are "uncovered" because the neutralizing holes have been pushed downward into the substrate.
- The positive gate voltage also attracts electrons from the n+ source and drain regions into the channel region.
- When a sufficient number of electrons accumulate near the surface of the substrate under the gate, an *n* region is effectively created, connecting the source and drain regions to one another.
- The *induced n* region forms a **channel** for current flow from drain to source.
- This MOSFET is called an **n-channel MOSFET** or an **NMOS transistor.**

Creating channel for current flow

- An n-channel MOSFET is formed in a p-type substrate
  - The channel is formed by *inverting* the substrate surface from *p* type to *n* type.
- The value of v_{GS} at which a sufficient number of mobile electrons accumulate under the gate to form a conducting channel is called the **threshold voltage, V_t.**
  - For an n-channel FET the threshold voltage is positive.
- The gate and channel region of the MOSFET form a parallel-plate capacitor with the oxide layer acting as the capacitor dielectric.
  - The positive voltage on the gate causes positive charge to accumulate on the top plate (gate electrode).
  - The corresponding negative charge on the bottom plate is formed by the electrons in the induced channel.
  - An electric field develops in the vertical direction, from positive to negative.

Creating a channel for current flow

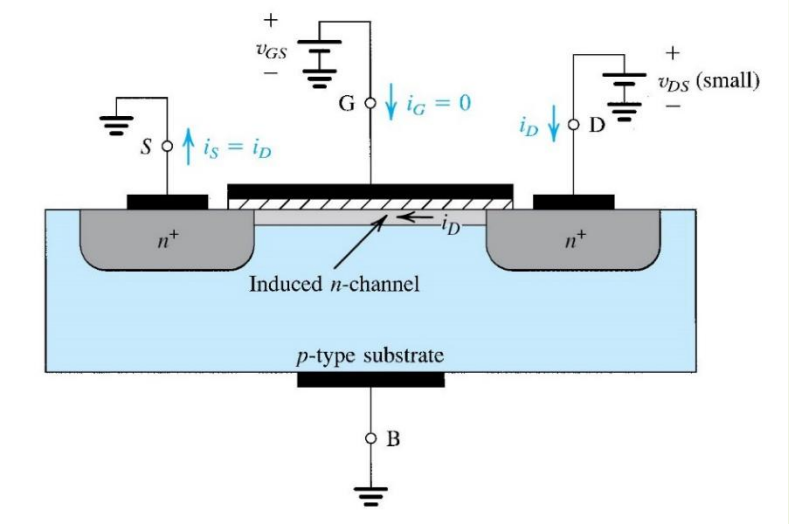
- The voltage across the oxide must exceed V_t in order for a channel to form.
  - When v_{DS} = 0, the voltage at every point along the channel is zero and the voltage across the oxide is uniform and equal to v_{GS}.
  - The excess of v_{GS} over V_t is called the **overdrive voltage** and is the quantity that determines the charge in the channel.
  - $v_{GS} - V_t \equiv v_{OV}$
- We can express the magnitude of the electron charge in the channel by
  - $|Q| = C_{ox}(WL)v_{OV}$
  - C_{ox} is called the **oxide capacitance** and is the capacitance of the parallel-plate capacitor per unit gate area (in units of F/m^2)
  - W is the width of the channel
  - L is the length of the channel

Creating a channel for current flow

- The oxide capacitance C_{ox} is given by
  - $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$
  - ϵ_{ox} is the permittivity of silicon dioxide
  - $\epsilon_{ox} = 3.9\epsilon_0 = 3.9 \times 8.854 \times 10^{-12} = 3.45 \times 10^{-11} F/m$
- The oxide thickness t_{ox} is determined by the process used to fabricate the MOSFET. As an example
  - $C_{ox} = \frac{3.45 \times 10^{-11} F/m}{4 \times 10^{-9}\, m} = 8.6 \times 10^{-3} F/m^2 = 8.6\, fF/\mu m^2$
- For a MOSFET fabricated in this technology with a channel length of $L = 0.18 \mu m$ and $W = 0.72\ \mu m$, the total capacitance between gate and channel is
  - $C = C_{ox} WL = 8.6 \times 0.18 \times 0.72 = 1.1 fF$ (fF denotes femtofarad, 10^{-15}F)

Applying a small v_{DS}

- Having induced a channel, we now apply a positive voltage v_{DS} between drain and source.
- When v_{DS} is small (50 mV or so), the voltage causes a current i_D to flow through the induced channel.
  - Current is carried by free electrons traveling from source to drain (current in opposite direction)



Applying a small v_{DS}

- We wish to calculate the drain current i_D.
- Because v_{DS} is small, we can assume that the voltage between the gate and various points along the channel remains approximately constant and equal to the value at the source end, v_{GS}.
- So, the effective voltage (or overdrive voltage) between the gate and various points along the channel remains equal to v_{OV}, and the channel charge Q is the same as what we had calculated, namely $|Q| = C_{ox}(WL)v_{OV}$.
- To calculate the drain current we first find the charge per unit channel length
  - $\frac{|Q|}{L} = C_{ox}Wv_{OV}$

Applying a small v_{DS}

- The voltage v_{DS} establishes an electric field E across the length of the channel
  - $|E| = \frac{v_{DS}}{L}$
- This electric field then causes the channel electrons to drift toward the drain with a velocity
  - Electron drift velocity = $\mu_n|E| = \mu_n \frac{v_{DS}}{L}$
  - Where μ_n is the mobility of the electrons at the surface of the channel, a physical parameter whose value depends on the fabrication process.
- We find the value of i_D by multiplying the charge per unit channel length by the electron drift velocity
  - $i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right)v_{OV}\right]v_{DS}$

Applying a small v_{DS}

- For small v_{DS}, the channel behaves as a linear resistance whose value is controlled by the overdrive voltage
  - $i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)\right] v_{DS}$
- The channel conductance can be found as
  - $g_{DS} = (\mu_n C_{ox})\left(\frac{W}{L}\right) v_{OV} = (\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)$
- There are three factors that make up the conductance
- The first factor, $(\mu_n C_{ox})$, is determined by the process technology used to fabricate the MOSFET.
  - The factor $(\mu_n C_{ox})$ is called the **process transconductance** parameter and given the symbol k_n', *where* $k_n' = \mu_n C_{ox}$.

Applying a small v_{DS}

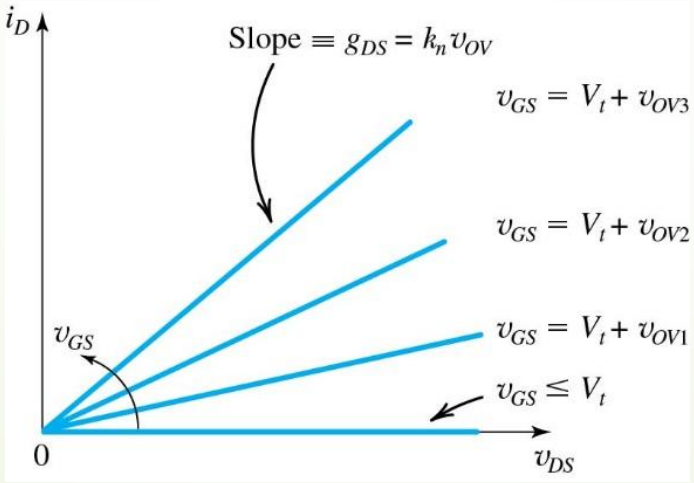
- The second factor in the expression for conductance is the transistor **aspect ratio** (W/L).
  - That the channel conductance is proportional to the channel width W and inversely proportional to the channel length L makes perfect physical sense.
  - The (W/L) ratio is a quantity that can be determined by the device designer to get the desired *i-v* characteristics.
- The product of the process transconductance parameter k'_n and the transistor aspect ratio (W/L) is called the **MOSFET transconductance parameter, k_n**
  - $k_n = k'_n(W/L) = (\mu_n C_{ox})$(W/L)
  - Both k'_n and k_n have the dimensions of $^{A}/_{V^2}$.

Applying a small v_{DS}

- The third term in the expression of the channel conductance is the overdrive voltage v_{OV}. The overdrive voltage directly determines the magnitude of electron charge in the channel.
- We now note that with v_{DS} kept small, the MOSFET behaves as a linear resistance r_{DS} whose value is controlled by the gate voltage
  - $r_{DS} = \frac{1}{g_{DS}}$
  - $r_{DS} = \frac{1}{(\mu_n C_{ox})(W/L)v_{OV}}$
  - $r_{DS} = \frac{1}{(\mu_n C_{ox})(W/L)(v_{GS}-V_t)}$

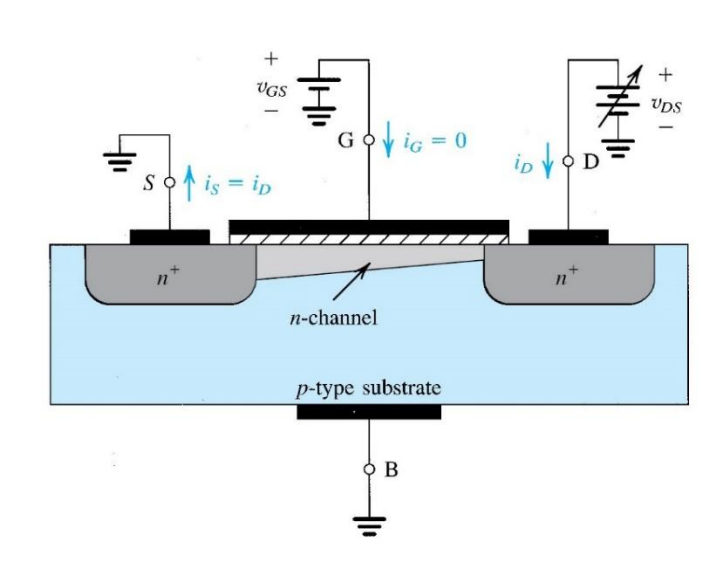
Applying a small v_{DS}

- The operation of the MOSFET as a voltage-controlled resistance is shown.
- This is a graph of $i_D vs. v_{DS}$ for various values of v_{GS}.
- The resistance is infinite for $v_{GS} \leq V_t$ and decreases as v_{GS} increases above V_t.



Operation as v_{DS} is increased

- Let v_{GS} be constant at a value greater than V_t (this means a constant v_{OV})
- As we travel from source to drain along the channel, the voltage (measured relative to the source) increases from zero to v_{DS}.
- The voltage between the gate and points along the channel decreases from $v_{GS} = V_t + v_{OV}$ at the source end to $v_{GD} = v_{GS} - v_{DS} = V_t + v_{OV} - v_{DS}$ at the drain end.



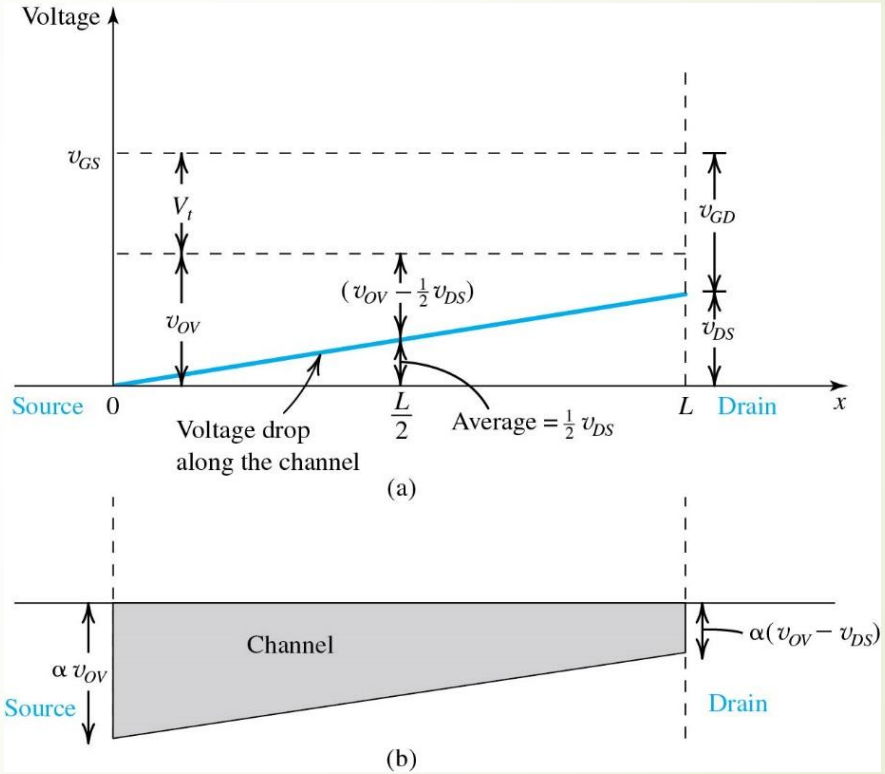
Operation as v_{DS} is increased

- The channel depth depends on the voltage v_{GD} and specifically by the amount that this voltage exceeds V_t, and the channel no longer has uniform depth.
- The channel is deepest at the source end (where the depth is proportional to v_{OV}) and shallowest at the drain end (where depth is proportional to $v_{OV} - v_{DS}$).
- The induced channel acquires a tapered shape and its resistance increases as v_{DS} is increased.

Operation as v_{DS} is increased

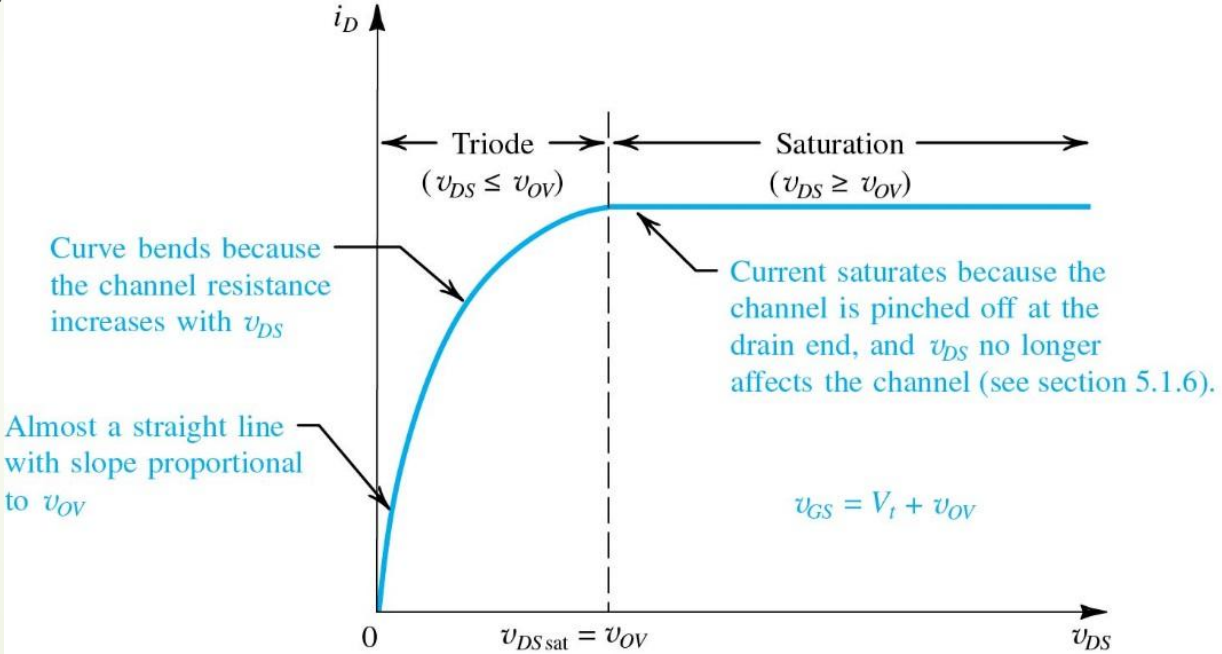
(a) For a MOSFET with $v_{GS} = V_t + v_{OV}$, applying v_{DS} causes the voltage drop along the channel to vary linearly, with an average value of $\frac{1}{2}v_{DS}$ at the midpoint. Since $v_{GD} > V_t$, the channel still exists at the drain end.

(b) The channel shape corresponding to the situation in (a). While the depth of the channel at the source end is still proportional to v_{OV}, that at the drain end is proportional to $v_{OV} - v_{DS}$.



Operation as v_{DS} is increased

- Since the resistance of the channel increases with v_{DS}, the $i_D - v_{DS}$ curve bends.

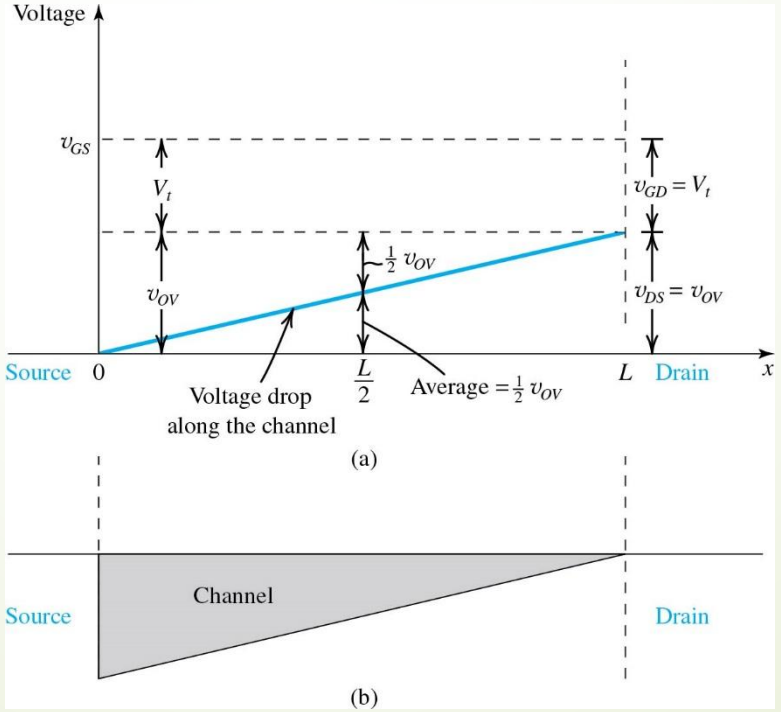


Operation as v_{DS} is increased

- The charge in the tapered channel is proportional to the channel cross-sectional area.
- This area in turn is proportional to $\frac{1}{2}[v_{OV} + (v_{OV} - v_{DS})] = \left(v_{OV} - \frac{1}{2}v_{DS}\right)$
- The relationship between i_D and v_{DS} can be found by replacing v_{OV} by $\left(v_{OV} - \frac{1}{2}v_{DS}\right)$ in the equation we had for i_D.
  - $i_D = k_n'\left(\frac{W}{L}\right)\left(v_{OV} - \frac{1}{2}v_{DS}\right)v_{DS}$
- This relationship describes the semiparabolic portion of the $i_D - v_{DS}$ curve.
- The equation for drain current in this region may be written
  - $i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_t)v_{DS} - \frac{1}{2}v_{DS}^2\right]$

Channel pinch-off and current saturation

- The channel has a finite depth at the drain as long as v_{DS} is sufficiently small that the voltage between gate and drain, v_{GD}, exceeds V_t.
- When $v_{DS} = v_{OV}$, $v_{GD} = V_t$, and the channel depth at the drain end reduces to zero.

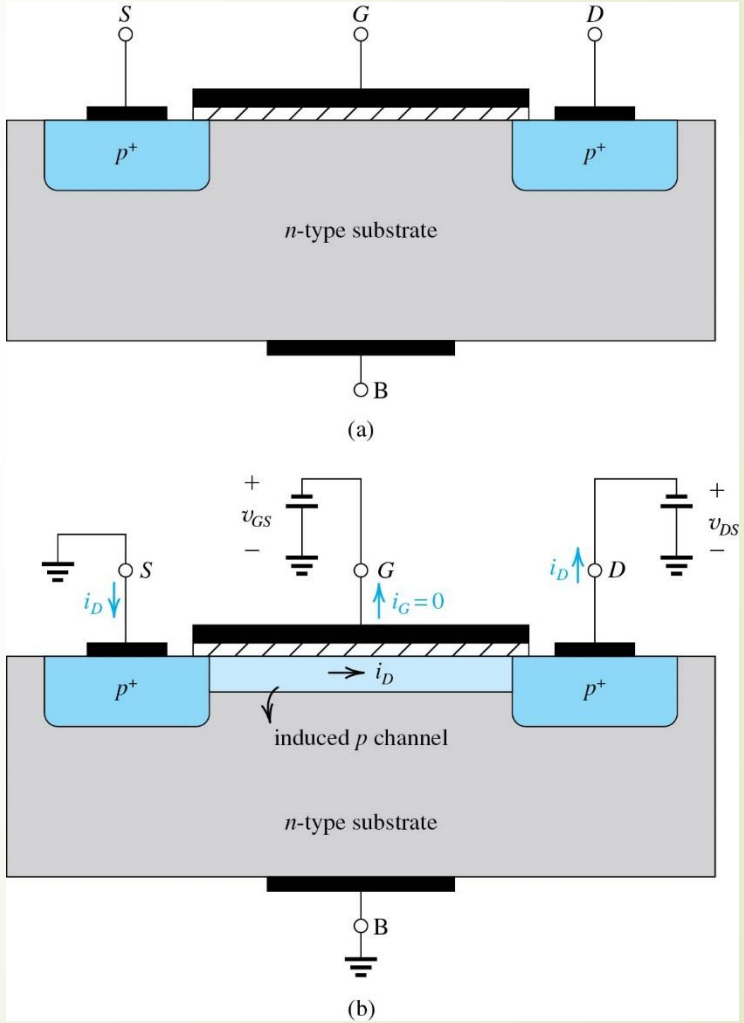


Channel pinch-off and current saturation

- The zero depth of the channel at the drain end is called **channel pinch-off.**
- Increasing v_{DS} such that $v_{DS} > v_{OV}$ has no effect on the channel shape and charge.
- The current through the channel remains constant at the value for $v_{DS} = v_{OV}$.
- The drain current **saturates** at the value found by substituting $v_{DS} = v_{OV}$ into the expression for i_D.
  - $i_D = \frac{1}{2} k'_n \left(\frac{W}{L}\right) v_{OV}^2$
- The MOSFET has then entered the **saturation region**.
- The voltage at which saturation occurs is denoted v_{DSsat}
  - $v_{DSsat} = v_{OV} = v_{GS} - V_t$
- Both the current through the channel and the voltage drop across it remain constant in saturation.

The *p*-channel MOSFET

The structure of the p-channel enhancement type MOSFET is similar to that of the NMOS device except that here the substrate is *n* type and the source and drain regions are *p*+.



The p-channel MOSFET

- To induce a channel for current flow, a negative voltage is applied to the gate.
  - By increasing the magnitude of the negative v_{GS} beyond the magnitude of the threshold voltage V_{tp}, which by convention is negative, a *p* channel is established.
  - This condition is described by $v_{GS} \leq V_{tp}$ or $|v_{SG}| \geq |V_{tp}|$
- To make a current i_D flow in the channel, we apply a negative voltage v_{DS} to the drain.
- We define the process transconductance parameter for the PMOS device as
  - $k'_p = \mu_p C_{ox}$ where μ_p is the mobility of holes in the induced *p* channel.
- The transistor transconductance parameter k_p is defined as
  - $k_p = k'_p(W/L)$

Complementary MOS or CMOS

- Complementary MOS technology uses MOS transistors of both polarities.
- CMOS is now the most widely used of all the IC technologies in both analog and digital circuits.