MOSFET transistor

- Physical structure

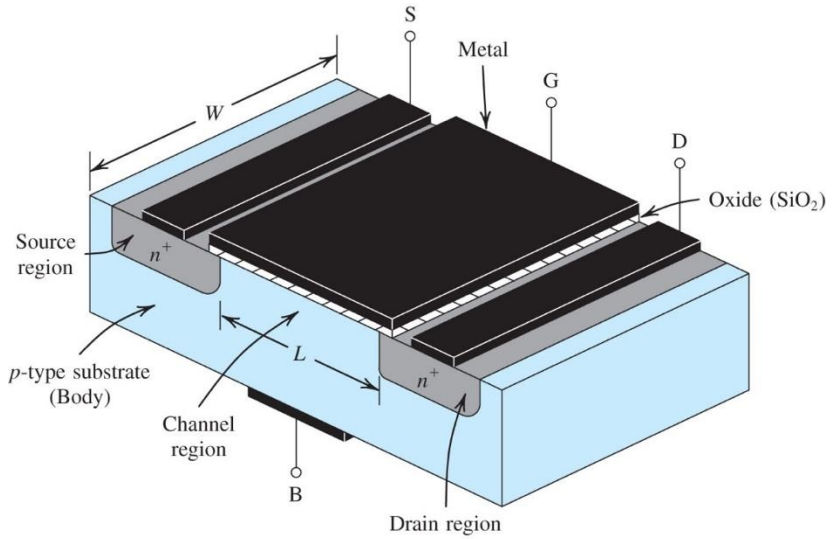
The MOSFET transistor

- The MOSFET transistor is the most widely used electronic component
- Most analogue and digital functionalities can be implemented using MOSFET transistors only, no extra components needed
- The use of MOSFET transistors is ubiquitous in the design of IC (i.e. entire circuits fabricated on a single silicon chip)
  - MOSFET transistors are used to design analogue, digital and mixed-mode IC
  - CMOS (Complementary MOS) is the main IC technology
- Compared to BJT
  - Smaller feature size and easier fabrication process → Very Large Scale Integration (VLSI) → powerful memories and microprocessors
  - Higher input impedance, and lower power consumption

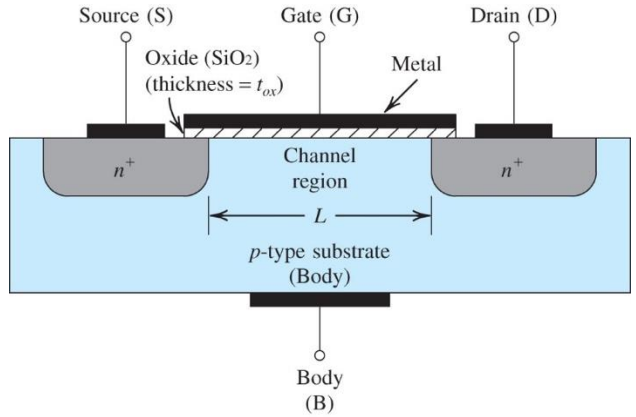
Physical structure: Enhancement-type NMOS

- There are two types of MOSFET: enhancement-type and depletion-type
  - We will consider only enhancement-type MOSFET that is the most widely used

- MOSFET of each type can be n-channel or p-channel, shortly NMOS or PMOS



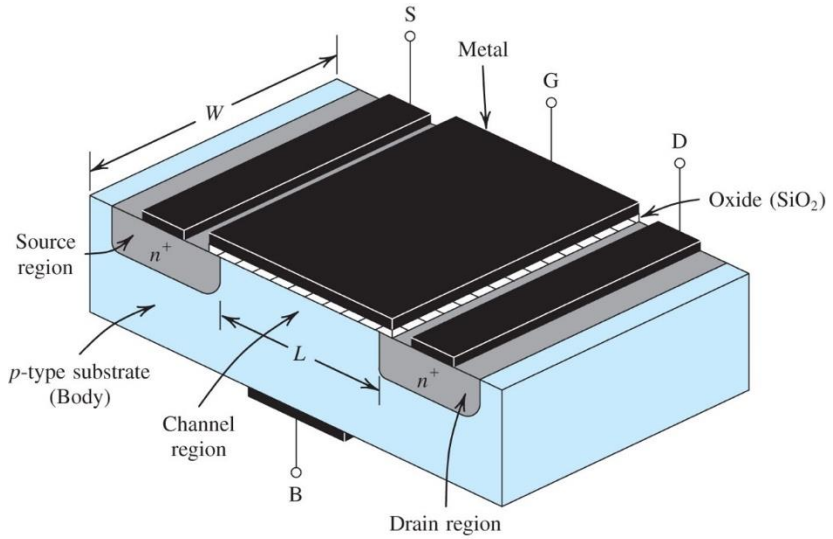
(a)



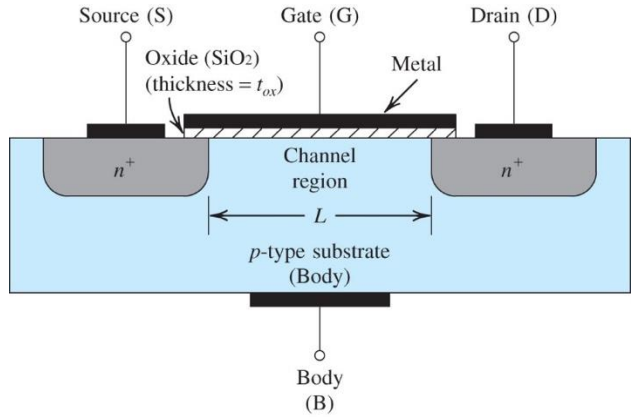
Physical structure: Enhancement-type NMOS

- Enhancement-type, n-channel MOSFET
  - p-type substrate (bulk contact)
  - Two highly doped n-type regions (source and drain contacts)
  - Oxide layer (typ. SiO_2) of thickness t_{OX} of a few nm is grown on top of the region between source and drain
  - Metal layer on top of the oxide layer (gate contact)

- The MOSFET gets its name from its physical structure

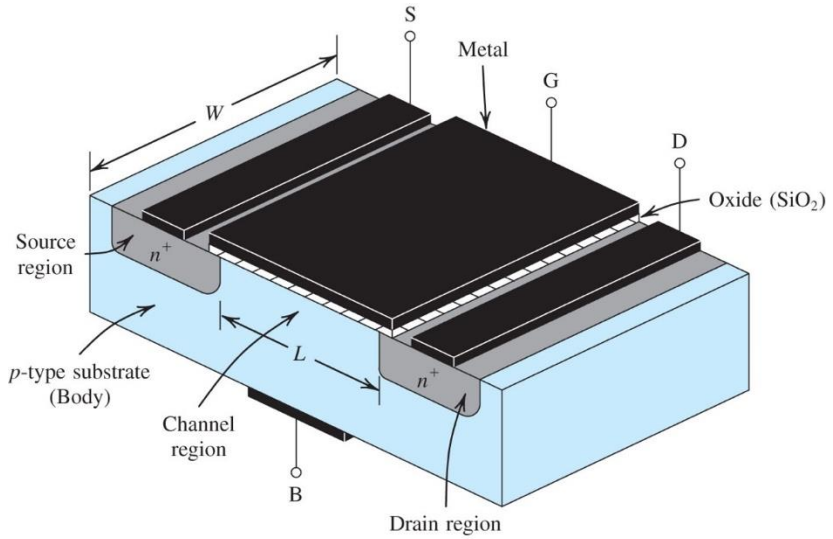


(a)

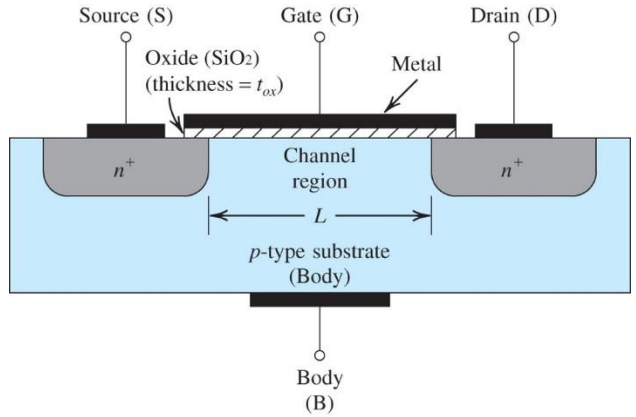


Physical structure: Enhancement-type NMOS

- The MOSFET is a symmetrical device
  - Source and drain have the same doping level and geometry and can be interchanged
- Source and drain are defined by the applied voltages
  - The drain is always more positive than the source
  - All voltages are referred to the source
- In standard operation the two pn-junctions (S-B and D-B) are always reverse biased

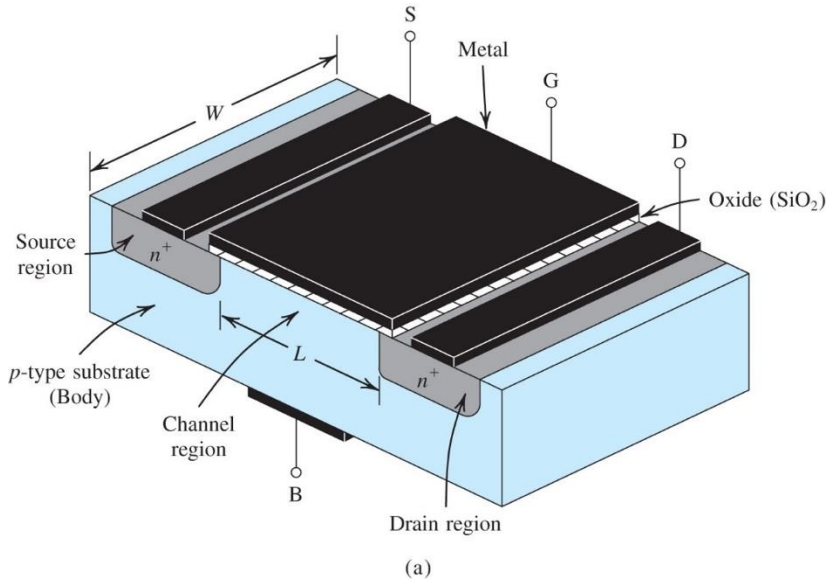


(a)

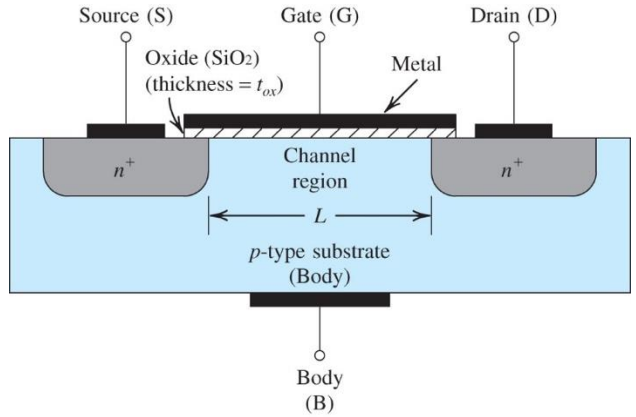


Physical structure: Enhancement-type NMOS

- The gate is electrically isolated from the bulk
  - There is no current flowing in the gate!!!
- The bulk is usually considered connected to the source, so it has no influence on the device operation
  - We will always consider $v_{BS} = 0$ and thus ignore the bulk effect
- The voltage applied to the gate controls the current flowing between source and drain
  - The current flows in a channel region of width W and length L



(a)

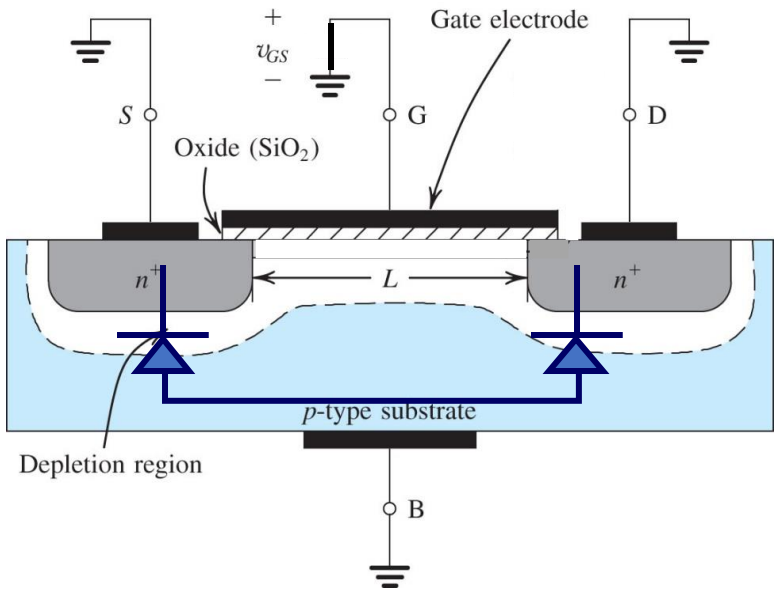


MOSFET transistor

- Channel formation

Channel formation

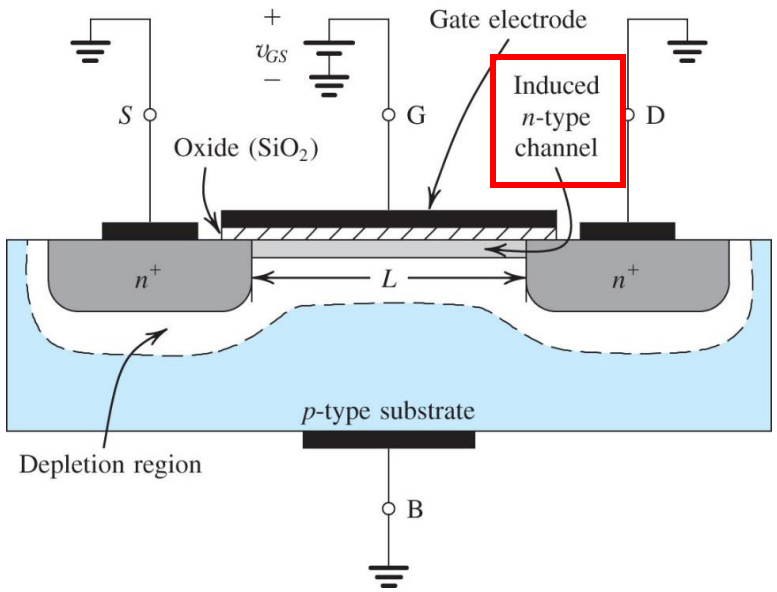
- Let's consider $v_{DS} = 0$, $v_{BS} = 0$
- For $v_{GS} = 0$
  - Two diodes back-to-back in series between source and drain
  - No current can flow even if $v_{DS} > 0$



Channel formation

For $v_{GS} > 0$

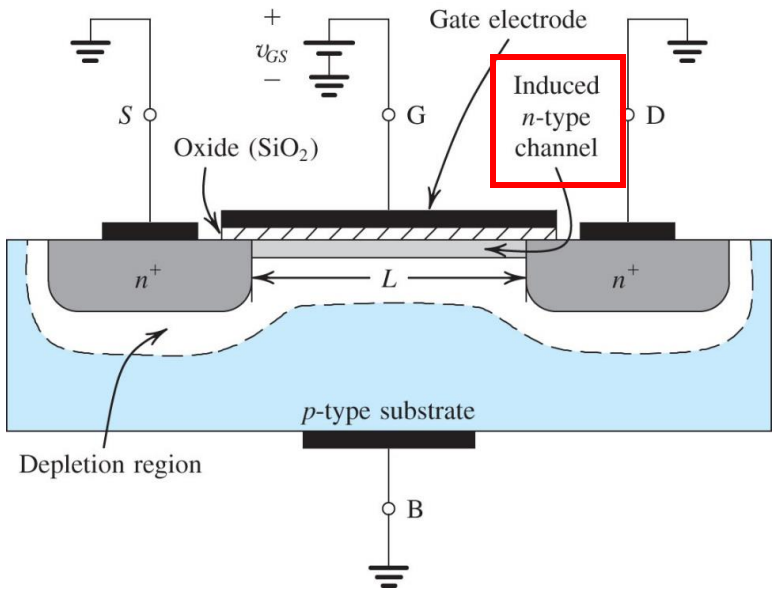
- Holes are repelled from the p-substrate region below the gate, leaving behind a carrier-depleted region
- Electrons from the n+ source and drain regions are attracted below the gate
- A conductive channel, made of free charge carriers, forms between source and drain → current can flow by applying $v_{DS} > 0$



Channel formation

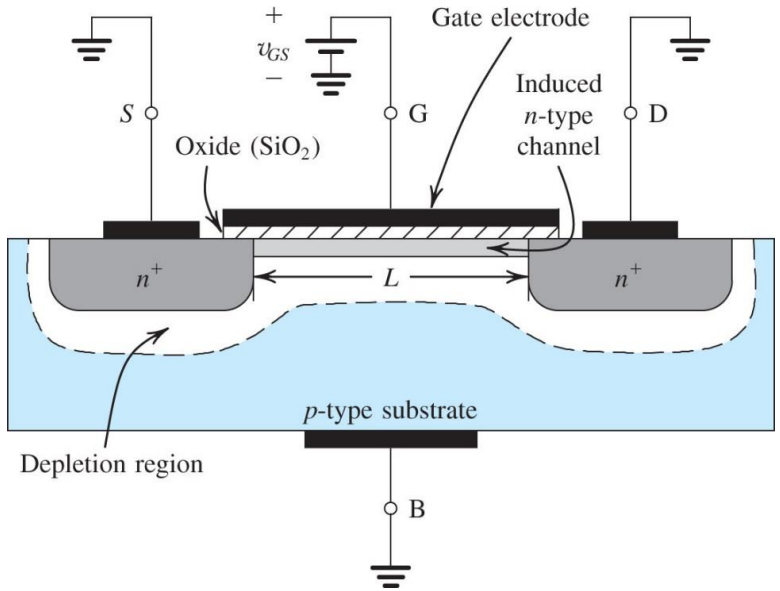
For $v_{GS} > 0$

- In this device the free charge carriers are electrons, hence the name n-channel MOSFET or NMOS
- Note: the substrate in an NMOS is p-type
- The value of v_{GS} for which the channel forms is called threshold voltage V_t
  - The value of V_t is defined by the manufacturing process



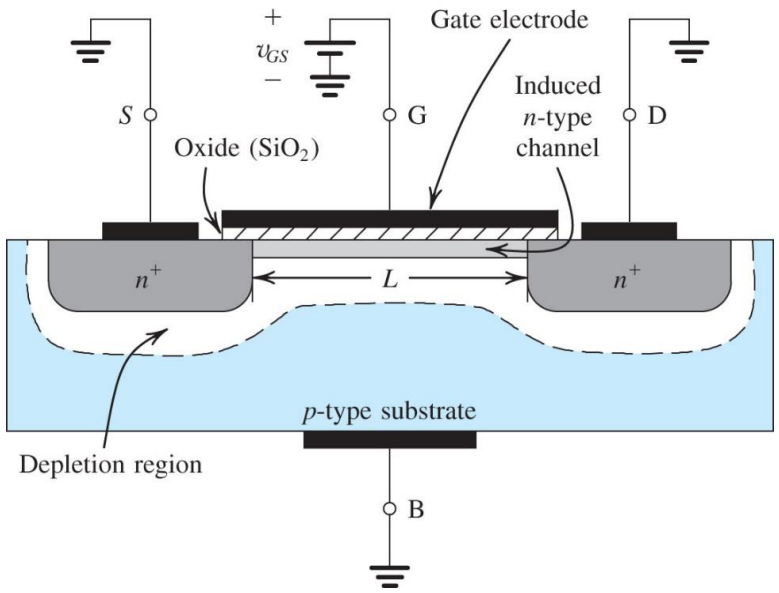
Channel formation

- The gate, SiO_2, channel stack forms a parallel plate capacitor
- Positive charge accumulates at the gate electrode to mirror the negative charge in the channel
- An electrical field exists in the vertical direction
- This electrical field controls the amount of charge in the channel and thus the conductivity of the device → field-effect-transistor



Channel formation

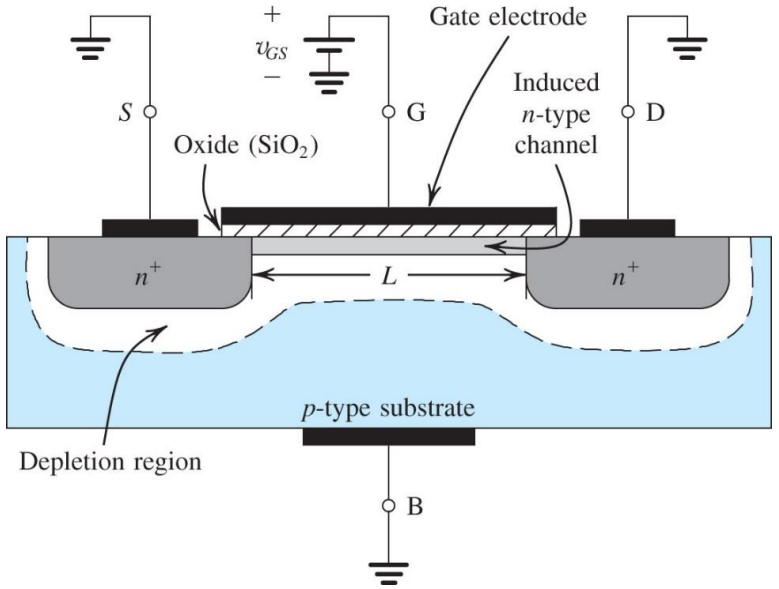
- The voltage across the parallel plate capacitor must exceed V_t for the channel to form
- Assuming $v_{DS} = 0$, the voltage along the channel is 0
- The voltage across the oxide is uniform and equal to v_{GS}



Channel formation

- The excess of v_{GS} over V_t is called overdrive voltage: $V_{OV} \equiv v_{GS} - V_t$
  - The overdrive voltage determines the amount of charge in the channel

- Charge accumulated in the channel: $|Q| = C_{OX}(WL)v_{OV}$
  - Assuming $v_{DS} = 0$, v_{GS} is uniform along the channel
  - $C_{OX} = \frac{\varepsilon_{OX}}{t_{OX}}$: oxide capacitance per unit gate area
    - $\varepsilon_{OX} = 3.45 \times 10^{-11} F/m$: permittivity of SiO2



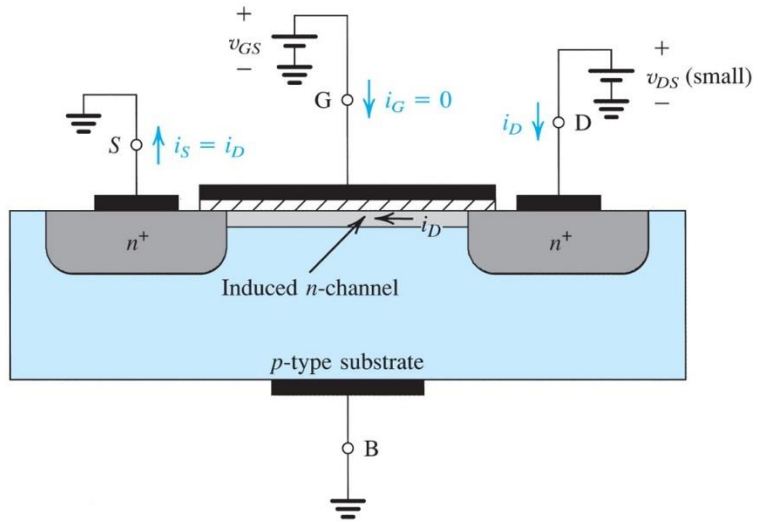
MOSFET transistor

- Current flow

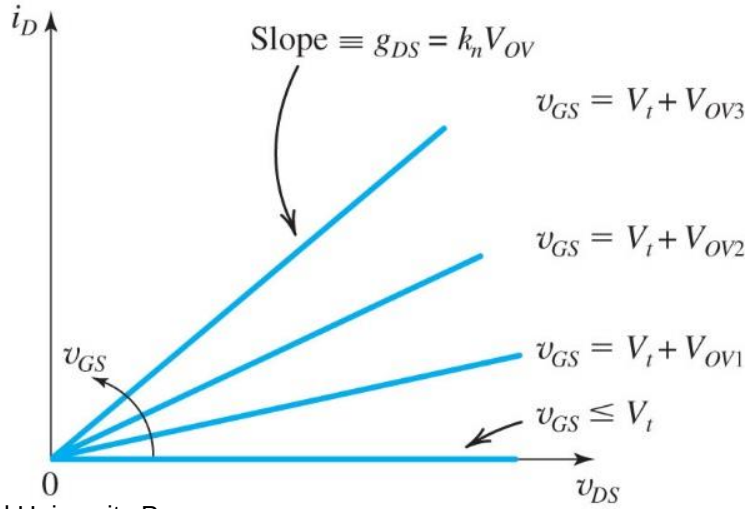
Current flow: small v_{DS}

- For small v_{DS}, the channel behaves as a linear resistance whose value is controlled by the overdrive voltage v_{OV}, which in turn is determined by v_{GS}

$$i_D = [(\mu_n C_{OX})\left(\frac{W}{L}\right)v_{OV}]v_{DS}$$

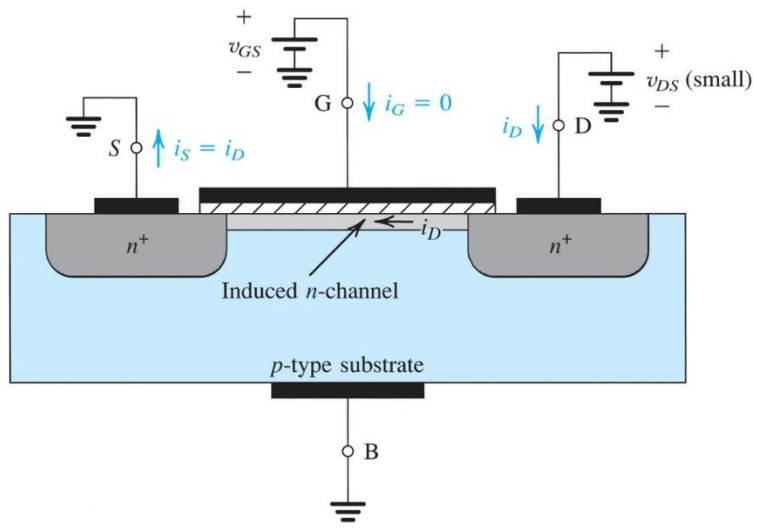


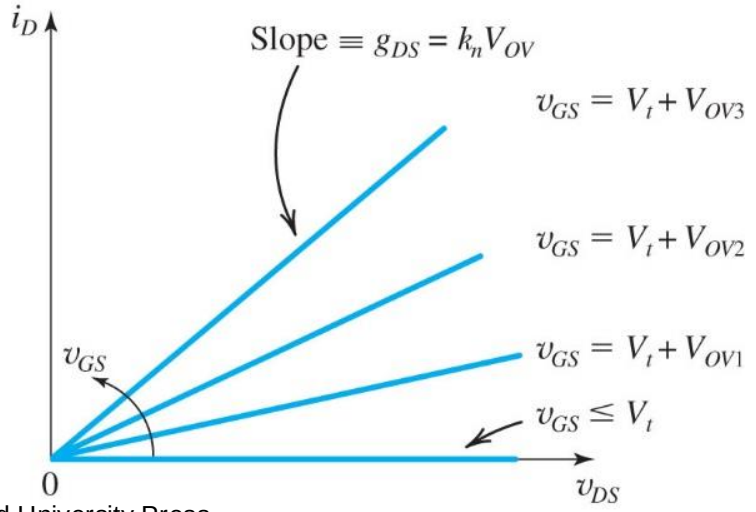
- Conductance $g_{DS} = \frac{i_D}{v_{DS}} = (\mu_n C_{OX})\left(\frac{W}{L}\right)v_{OV} = \frac{1}{r_{DS}}$



Current flow: small v_{DS}

- Process transconductance parameter $k'_n = (\mu_n C_{OX})$
  - Determined by the process technology
- Transistor aspect ratio $\frac{W}{L}$
  - Determined by the designer
  - Technology set a limit on min L
- MOSFET transconductance parameter $k_n = (\mu_n C_{OX})\frac{W}{L}$



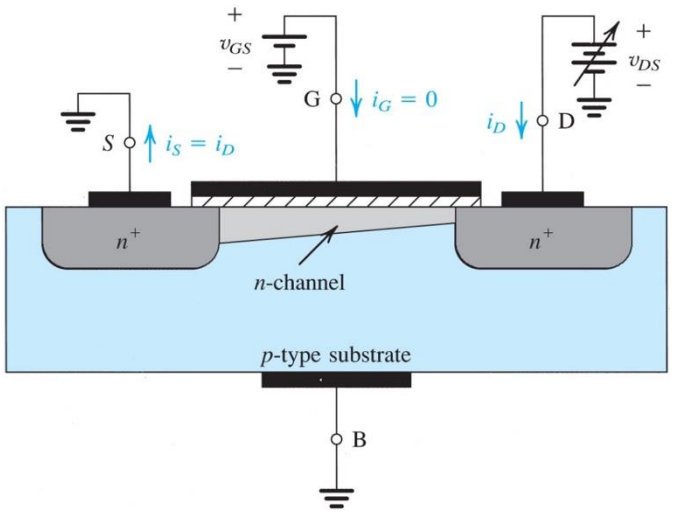


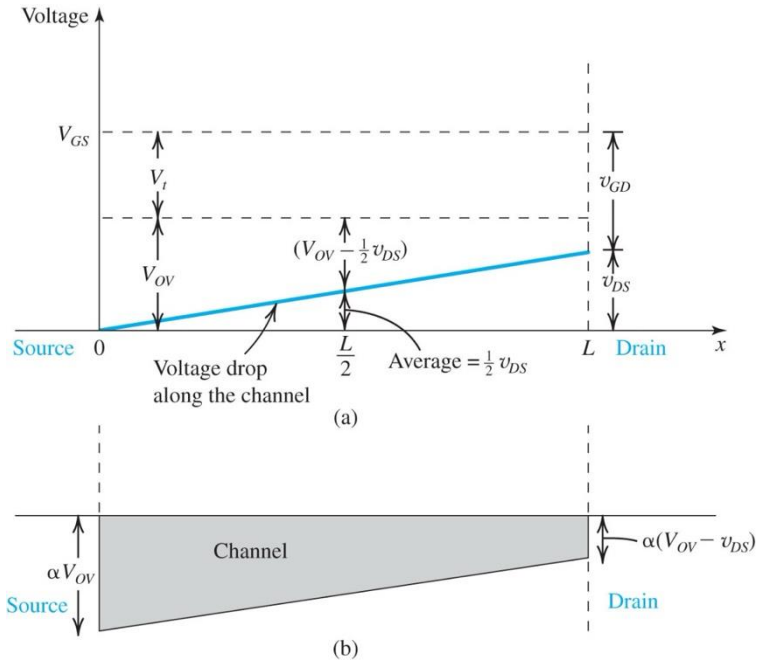
Current flow: v_{DS} increases

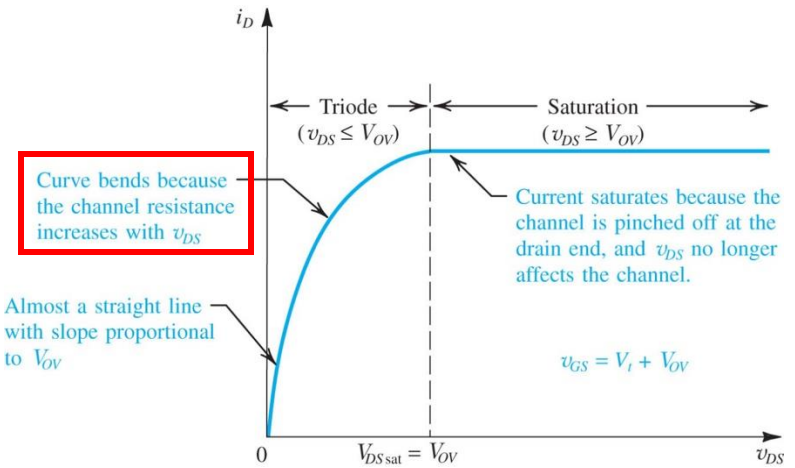
- As v_{DS} increases (with constant v_{GS}) the channel becomes more tapered, and its resistance increases accordingly

$$i_D = (\mu_n C_{OX})\left(\frac{W}{L}\right)(V_{OV} - \frac{1}{2}v_{DS})v_{DS}$$

  - This equation is valid from $v_{DS} = 0$ to $v_{DS} \leq V_{OV}$





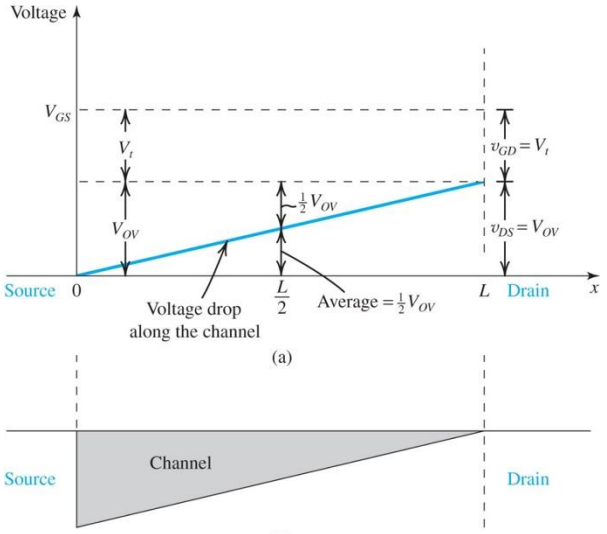


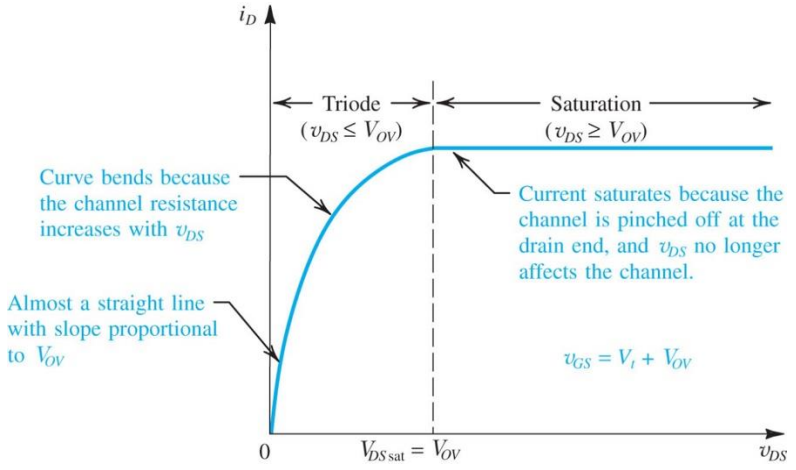
Current flow: $v_{DS} \geq V_{OV}$

- For $v_{DS} = V_{OV}$ → $v_{GS} = V_t$ → the channel depth at the drain is 0 → channel pinch-off

- For $v_{DS} > V_{OV}$ there in no change in the shape of the channel, the charge stays the same → the current stay constant at the value reached for $v_{DS} = V_{OV}$

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





Current flow: $v_{DS} \geq V_{OV}$

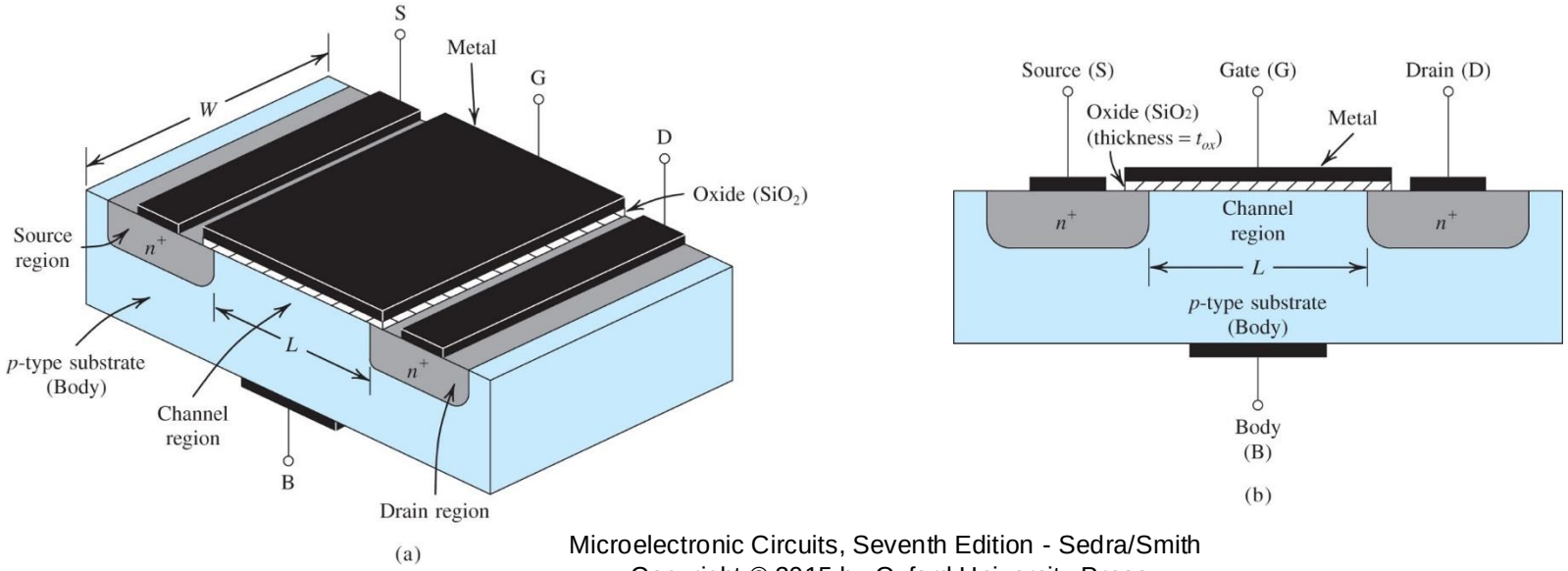
- After this point the MOSFET is in saturation
  - i_D is independent of v_{DS}
  - $V_{Dsat} = V_{OV} = V_{GS} - V_t$

- Any increase of v_{DS} over V_{Dsat} appears as a voltage drop across the depletion region at the drain-bulk junction
  - Voltage drop and current along the channel stay constant

- Current continues to flow between source and drain!!!
  - Electrons that reach the drain are swept by the electric field of the drain-substrate junction into the drain terminal

Recap

- Enhancement-type n-channel MOS transistor, i.e. NMOS



- Channel formation
  - $v_{OV} \equiv v_{GS} - V_t$
  - $|Q| = C_{OX}(WL)v_{OV}$

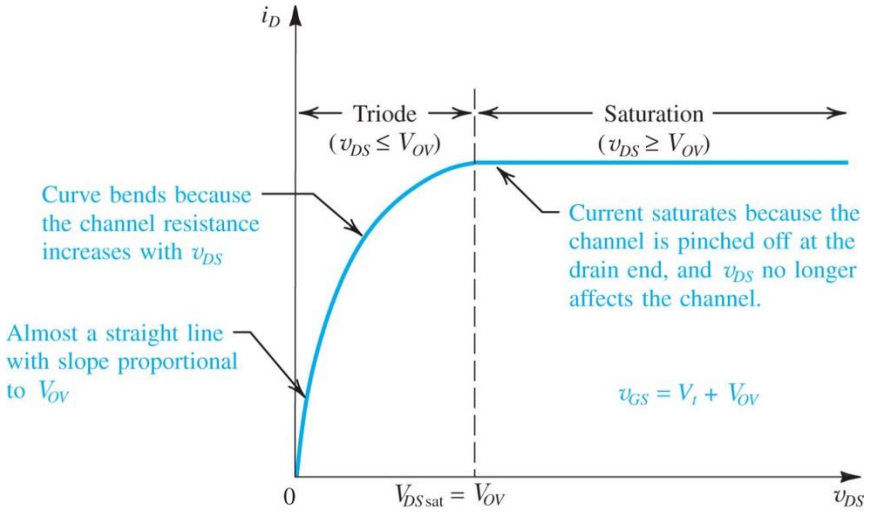
Recap

- From $v_{DS} = 0$ to $v_{DS} < v_{OV}$

$$i_D = (\mu_n C_{OX})\left(\frac{W}{L}\right)(v_{OV} - \frac{1}{2}v_{DS})v_{DS}$$

- For $v_{DS} \geq v_{OV}$

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



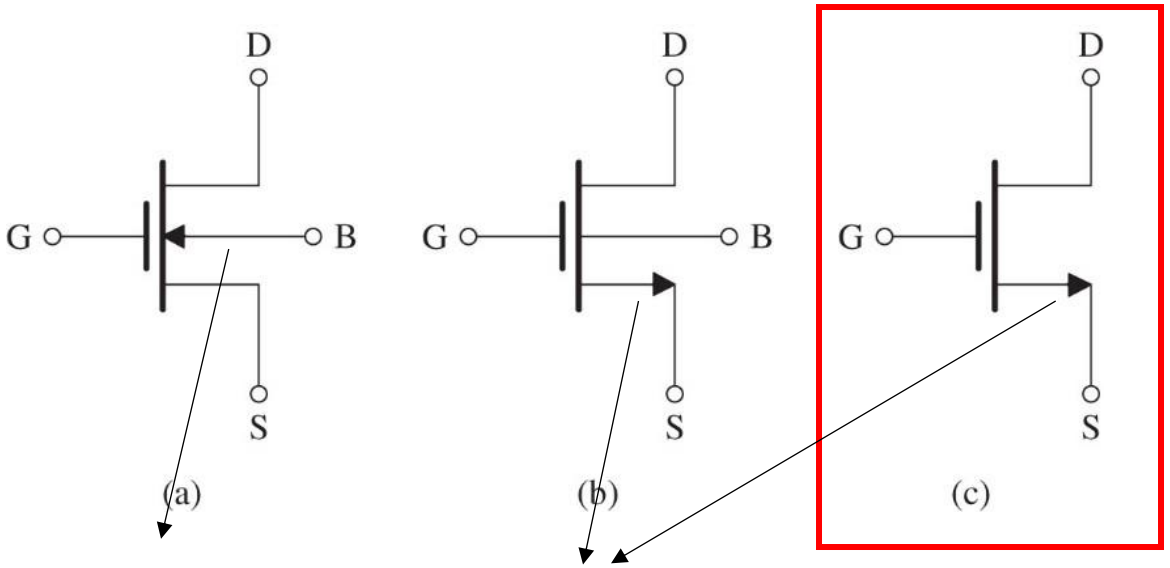
Recap

- Conductance $g_{DS} = \frac{i_D}{v_{DS}} = (\mu_n C_{OX})\left(\frac{W}{L}\right) v_{OV} = \frac{1}{r_{DS}}$
- Process transconductance parameter $k'_n = (\mu_n C_{OX})$
  - Determined by the process technology
- Transistor aspect ratio $\frac{W}{L}$
  - Determined by the designer
  - Technology set a limit on min L
- MOSFET transconductance parameter $k_n = (\mu_n C_{OX})\frac{W}{L}$

MOSFET transistor

- Circuit symbols
- I-V characteristics
- Finite output resistance in saturation

Circuit symbol: NMOS



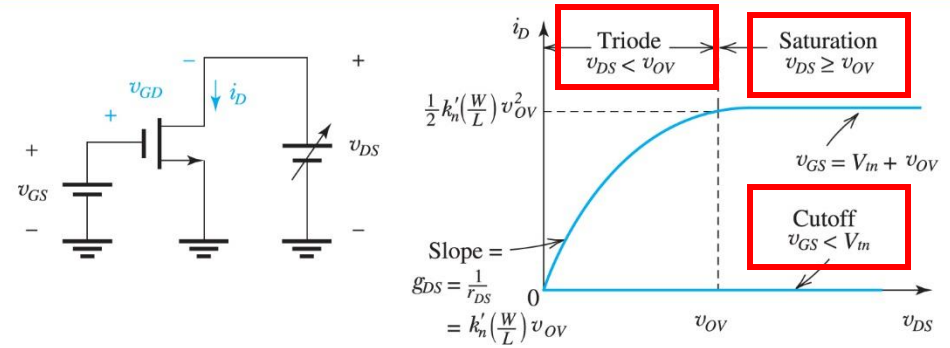
This arrow indicates the polarity of the substrate (p-type)

This arrow indicates the current flow

This is the symbol we use. It shows the direction of current flow. The B contact is omitted because we consider B connected to S

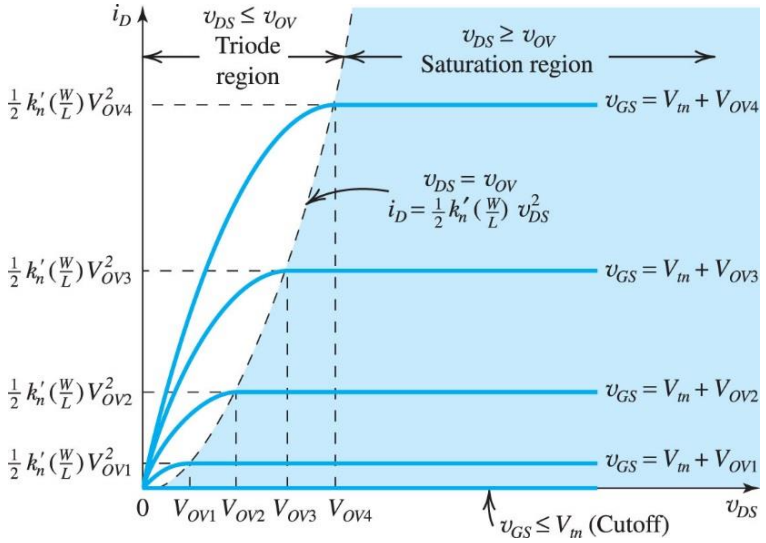
IV-characteristics: NMOS

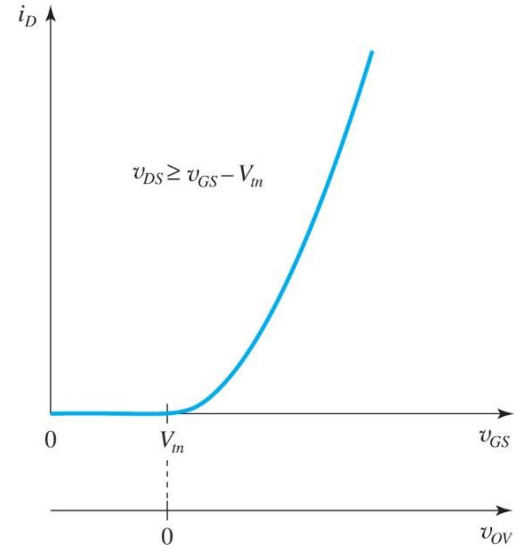
Table 5.1 Regions of Operation of the Enhancement NMOS Transistor



- $v_{GS} < V_{tn}$: no channel; transistor in cutoff; $i_D = 0$
- $v_{GS} = V_{tn} + v_{OV}$: a channel is induced; transistor operates in the triode region or the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region	*Saturation Region*
Continuous channel, obtained by:	Pinched-off channel, obtained by:
$v_{GD} > V_{tn}$	$v_{GD} \leq V_{tn}$
or equivalently:	or equivalently:
$v_{DS} < v_{OV}$	$v_{DS} \geq v_{OV}$
Then,	Then
$i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_{tn})v_{DS} - \frac{1}{2}v_{DS}^2\right]$	$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2$
or equivalently,	or equivalently,
$i_D = k_n'\left(\frac{W}{L}\right)\left(v_{OV} - \frac{1}{2}v_{DS}\right)v_{DS}$	$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2$



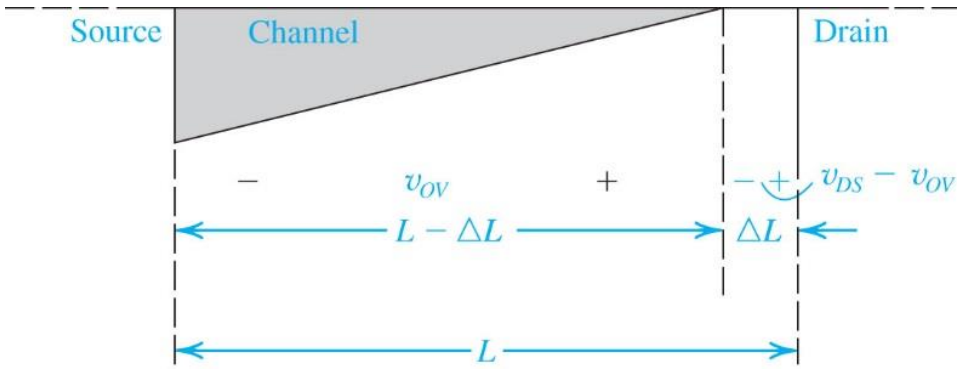


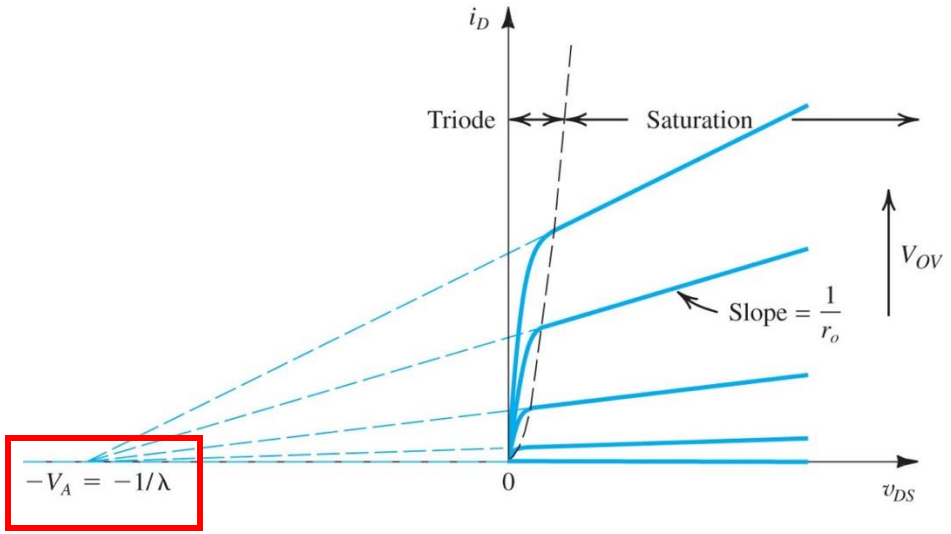
- Cut-off and ohmic regions are used when the transistor is operated as switch
- Saturation is used for amplification

Finite output resistance in saturation

- As v_{DS} is increased beyond v_{OV}, the pinch-off point moves towards the source → the channel length decrease
- This effect is called channel-length modulation
- i_D increases as it is inversely proportional to L

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



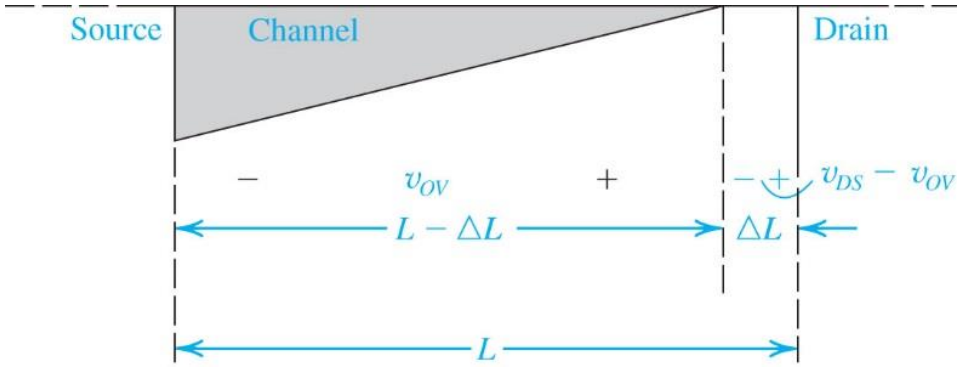


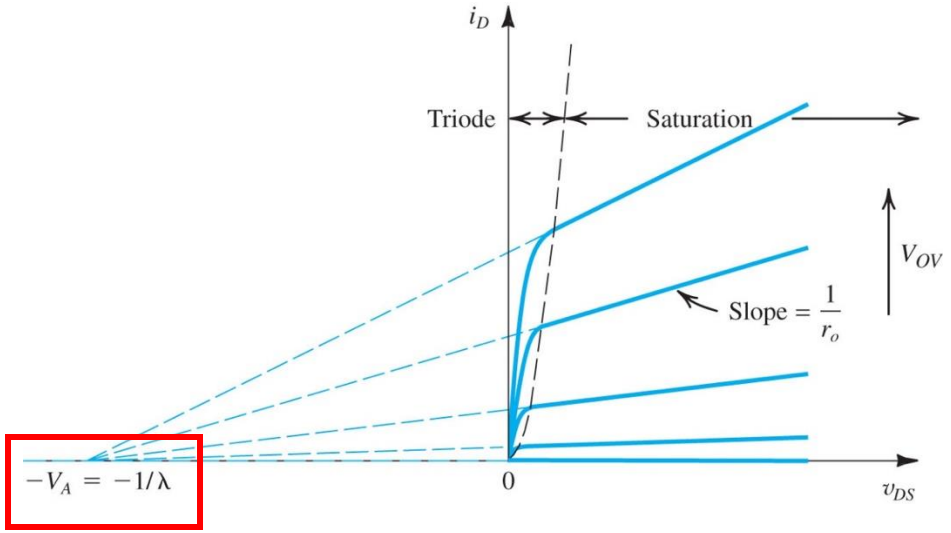
Finite output resistance in saturation

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

- The value of λ depends on process technology and channel length and it is the reciprocal of the Early voltage V_A

- Output resistance: $r_O = \dfrac{\partial i_D}{\partial v_{DS}} = \dfrac{1}{\lambda i_D}$



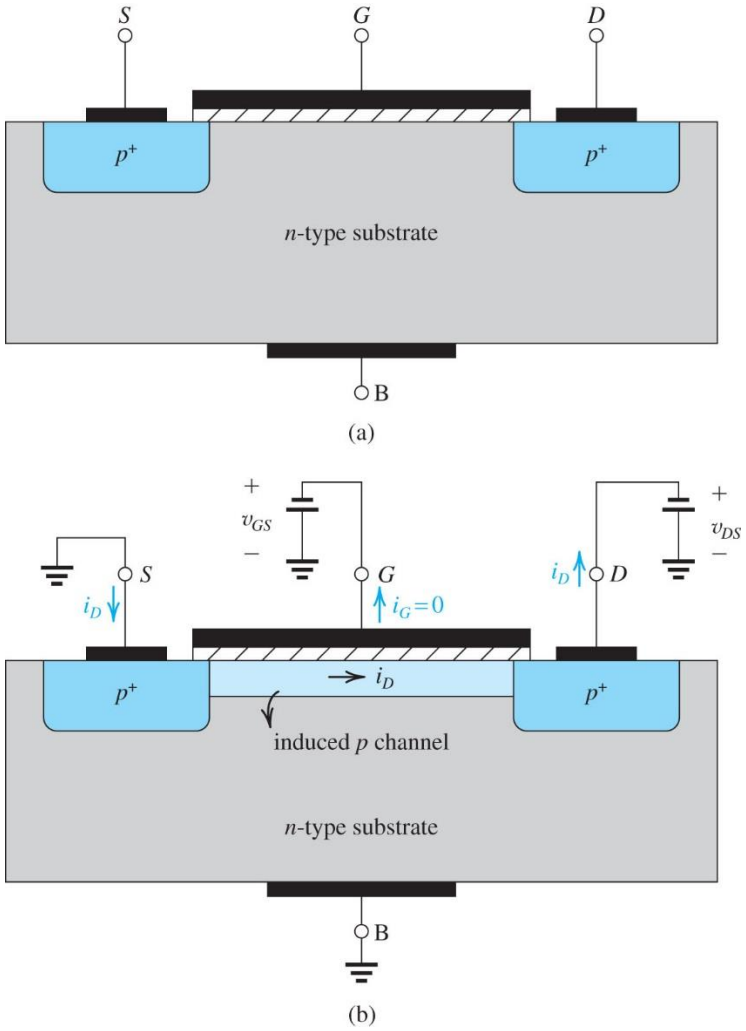


MOSFET transistor

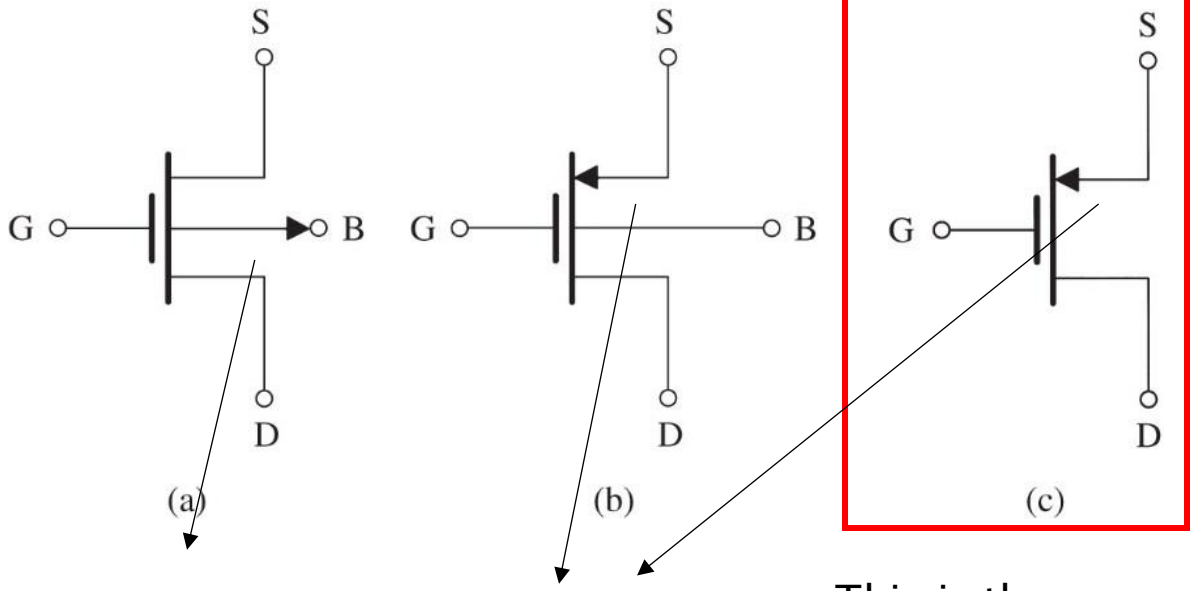
- PMOS

Enhancement-type PMOS

- n-type substrate
- p-doped source and drain regions
- Current is given by holes flowing between source and drain
- Reverse voltage polarity with respect to NMOS
  - $V_{GS} \leq V_t$ to create the channel
  - A negative V_{DS} is needed to have current flow



Circuit symbol: PMOS

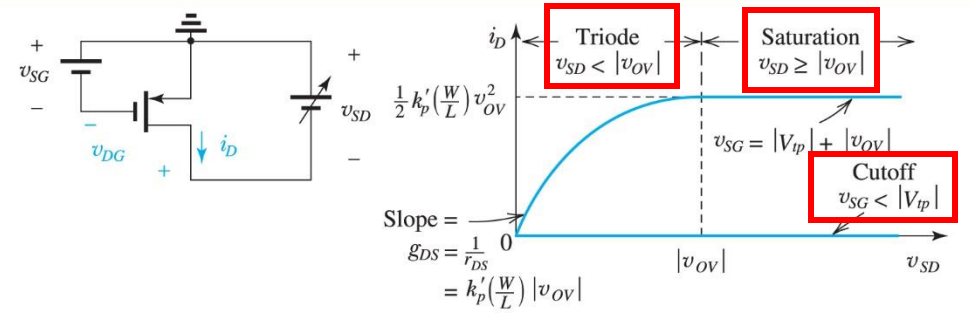


This arrow indicates the polarity of the substrate (n-type)

This arrow indicates the current flow

This is the symbol we use. It shows the direction of current flow. The B contact is omitted because we consider B connected to S

Table 5.2 Regions of Operation of the Enhancement PMOS Transistor



- $v_{SG} < |V_{tp}|$: no channel; transistor in cutoff; $i_D = 0$
- $v_{SG} = |V_{tp}| + |v_{OV}|$: a channel is induced; transistor operates in the triode region or in the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{DG} > |V_{tp}|$$

or equivalently

$$v_{SD} < |v_{OV}|$$

Then

$$i_D = k_p'\left(\frac{W}{L}\right)\left[(v_{SG} - |V_{tp}|)v_{SD} - \frac{1}{2}v_{SD}^2\right]$$

or equivalently

$$i_D = k_p'\left(\frac{W}{L}\right)\left(|v_{OV}| - \frac{1}{2}v_{SD}\right)v_{SD}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{DG} \leq |V_{tp}|$$

or equivalently

$$v_{SD} \geq |v_{OV}|$$

Then

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)(v_{SG} - |V_{tp}|)^2$$

or equivalently

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)v_{OV}^2$$

MOSFET transistor

- CMOS technology

CMOS technology

- Most widely used IC technology
- IC manufacturing was initially dominated by PMOS due to technical difficulties in the fabrication of NMOS transistors
- Once NMOS transistors could be fabricated they became the transistor of choice as electron mobility is higher than holes mobility
  - Greater gain and speed
- Now both devices can be fabricated in the same substrate → CMOS technology

CMOS technology

- Complementary MOS
  - Both PMOS and NMOS transistors are used
  - Allows design of many circuit configurations
  - Both analogue and digital circuits can be fabricated

- Characterized by feature size
  - Example: 65-nm CMOS technology means that the smallest transistor length L=65nm