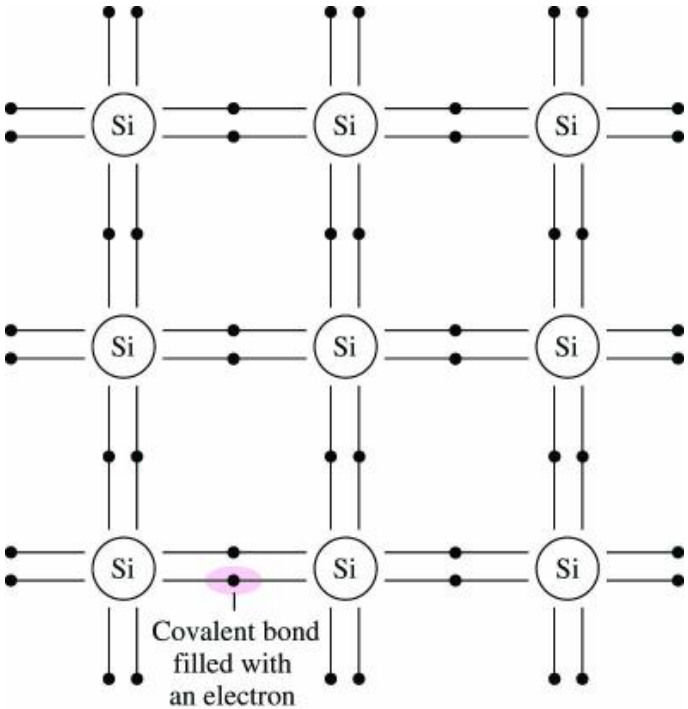
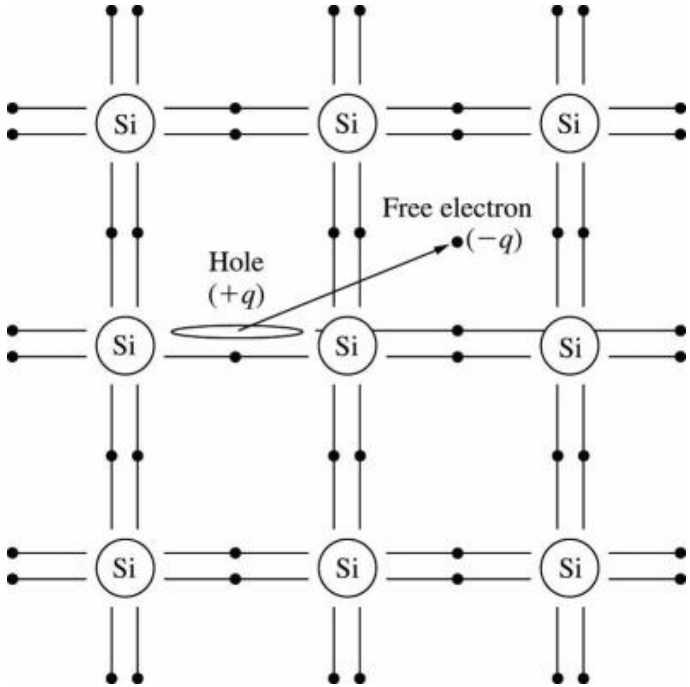
Portion of the Periodic Table including Typical Semiconductor Materials

	IIIA	IVA	VA	VIA
	5 10.811 **B** Boron	6 12.01115 **C** Carbon	7 14.0067 **N** Nitrogen	8 15.9994 **O** Oxygen
	13 26.9815 **Al** Aluminum	14 28.086 **Si** Silicon	15 30.9738 **P** Phosphorus	16 32.064 **S** Sulfur
IIB				
30 65.37 **Zn** Zinc	31 69.72 **Ga** Gallium	32 72.59 **Ge** Germanium	33 74.922 **As** Arsenic	34 78.96 **Se** Selenium
48 112.40 **Cd** Cadmium	49 114.82 **In** Indium	50 118.69 **Sn** Tin	51 121.75 **Sb** Antimony	52 127.60 **Te** Tellurium
80 200.59 **Hg** Mercury	81 204.37 **Tl** Thallium	82 207.19 **Pb** Lead	83 208.980 **Bi** Bismuth	84 (210) **Po** Polonium

Silicon Covalent Bond Model



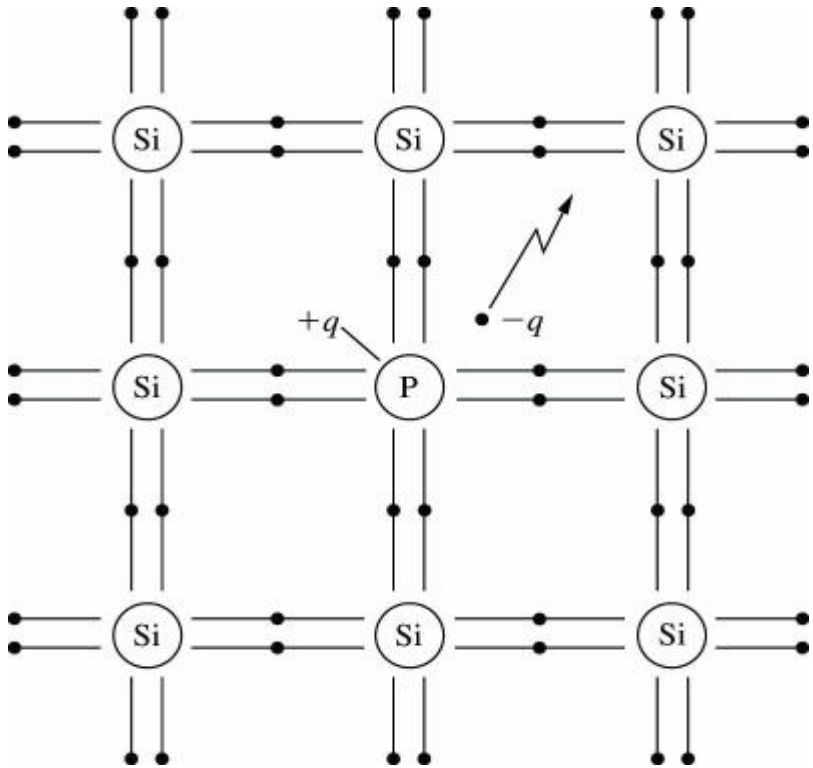
Near absolute zero, all bonds are complete. Each Si atom contributes one electron to each of the four bond pairs.



Increasing temperature adds energy to the system and breaks bonds in the lattice, generating electron-hole pairs.

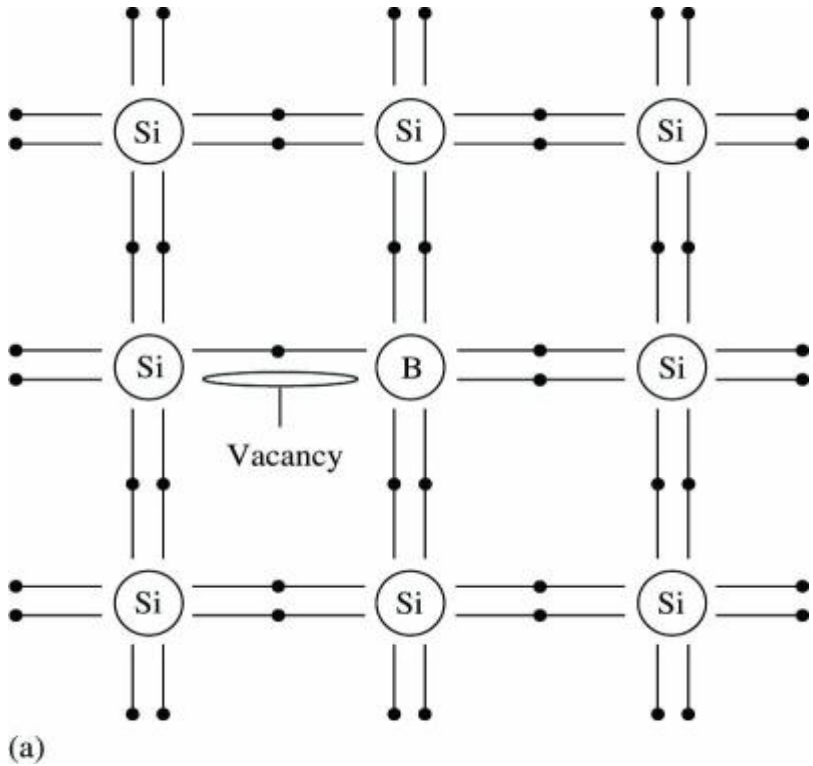
Donor Impurities in Silicon

- Phosphorous (or other column V element) atom replaces silicon atom in crystal lattice.
- Since phosphorous has five outer shell electrons, there is now an ‘extra’ electron in the structure.
- Material is still charge neutral, but very little energy is required to free the electron for conduction since it is not participating in a bond.

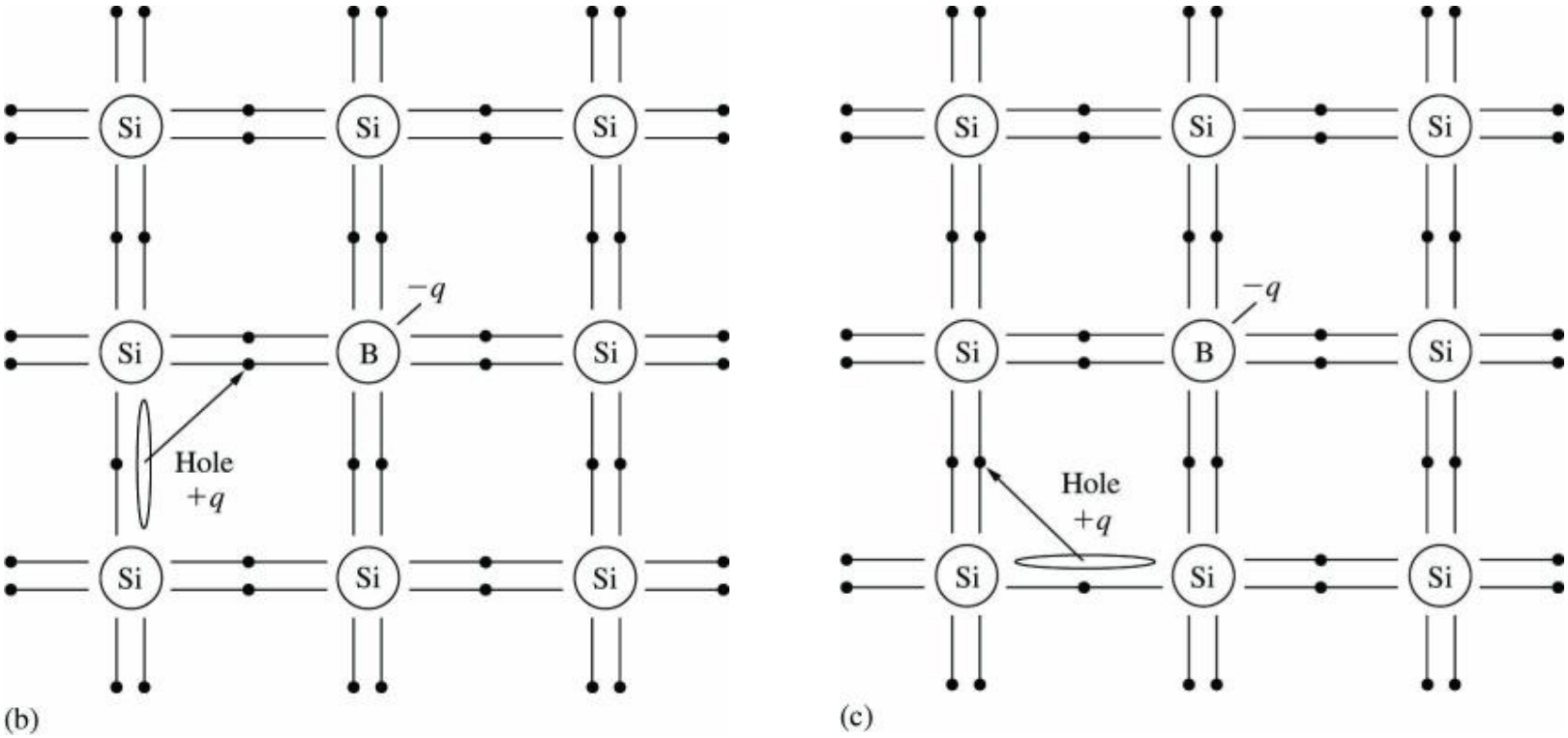


Acceptor Impurities in Silicon

- Boron (column III element) has been added to silicon.
- There is now an incomplete bond pair, creating a vacancy for an electron.
- Little energy is required to move a nearby electron into the vacancy.
- As the ‘hole’ propagates, charge is moved across the silicon.

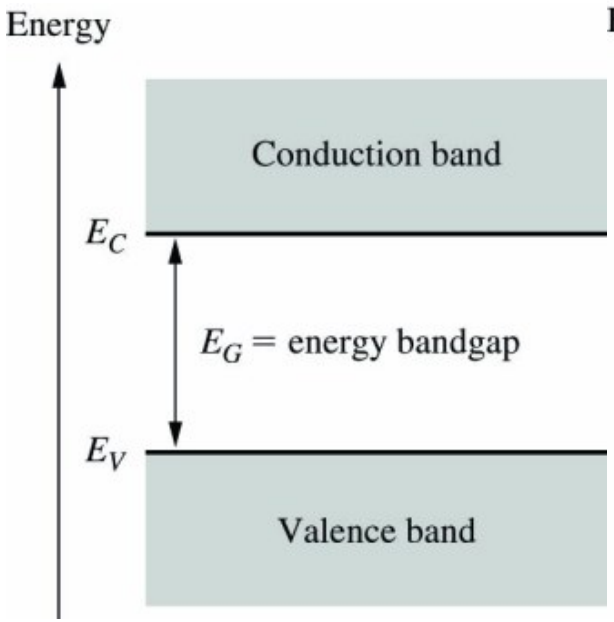


Acceptor Impurities in Silicon

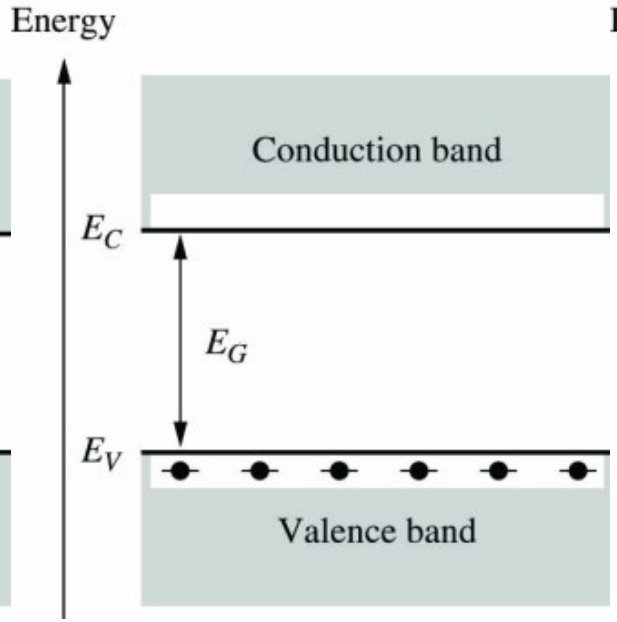


Hole is propagating through the silicon.

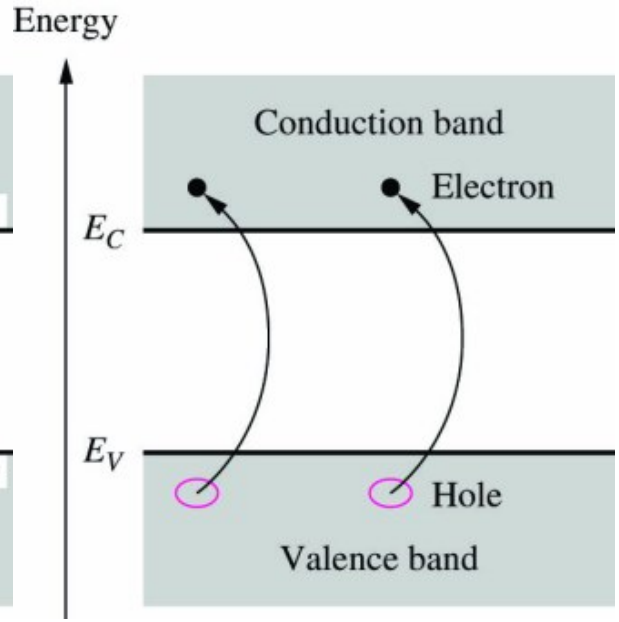
Semiconductor Energy Band Model



Semiconductor energy band model. E_C and E_V are energy levels at the edge of the conduction and valence bands.

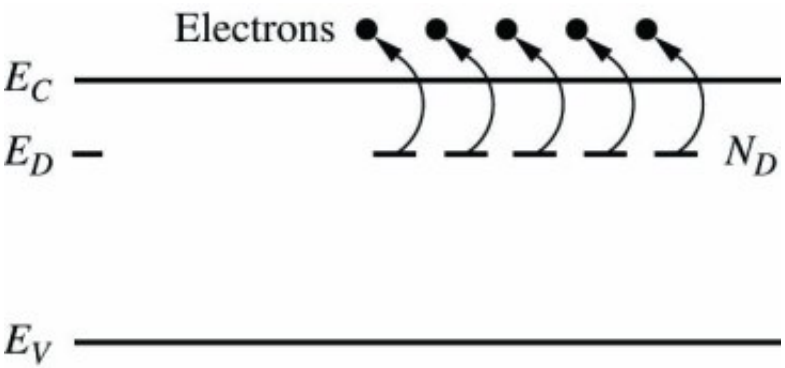


Electron participating in a covalent bond is in a lower energy state in the valence band. This diagram represents 0 K.

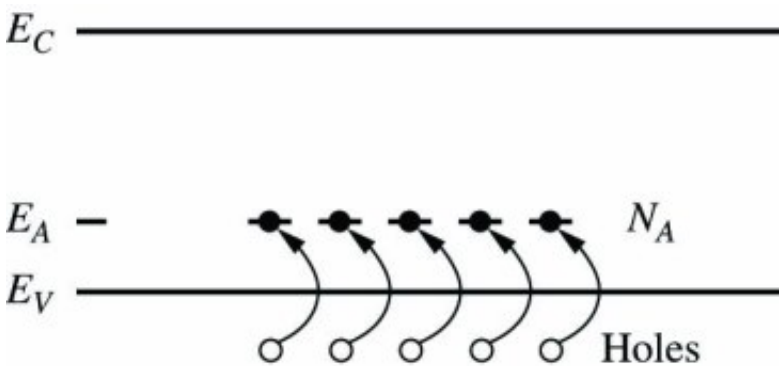


Thermal energy breaks covalent bonds and moves the electrons up into the conduction band.

Energy Band Model for a Doped Semiconductor

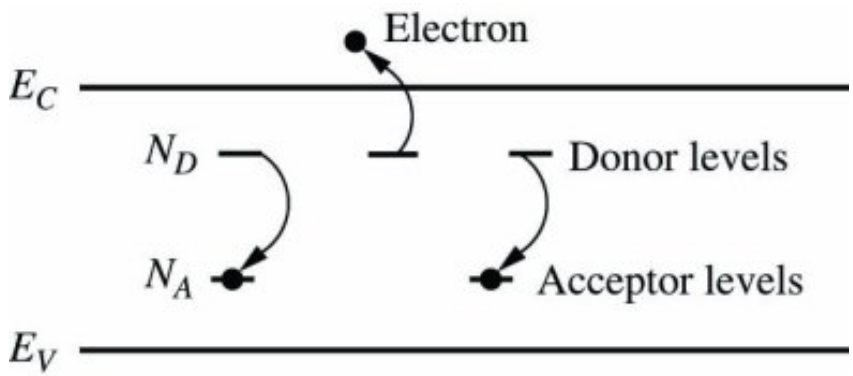


Semiconductor with donor or n-type dopants. The donor atoms have free electrons with energy E_D. Since E_D is close to E_C, (about 0.045 eV for phosphorous), it is easy for electrons in an n-type material to move up into the conduction band.



Semiconductor with acceptor or p-type dopants. The donor atoms have unfilled covalent bonds with energy state E_A. Since E_A is close to E_V, (about 0.044 eV for boron), it is easy for electrons in the valence band to move up into the acceptor sites and complete covalent bond pairs.

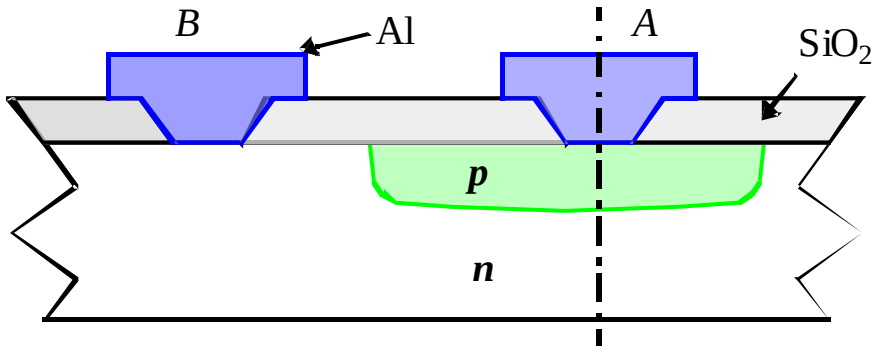
Energy Band Model for Compensated Semiconductor



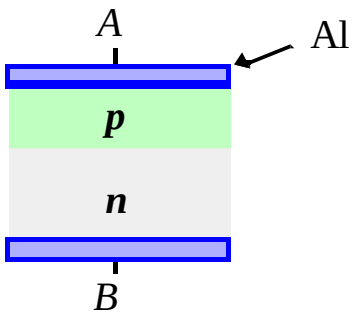
A compensated semiconductor has both n-type and p-type dopants. If $N_D > N_A$, there are more N_D donor levels. The donor electrons fill the acceptor sites. The remaining N_D-N_A electrons are available for promotion to the conduction band.

The combination of the covalent bond model and the energy band models are complementary and help us visualize the hole and electron conduction processes.

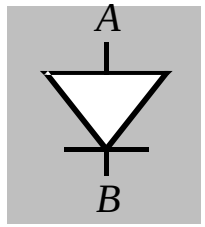
The Diode



Cross-section of *pn* junction in an IC process

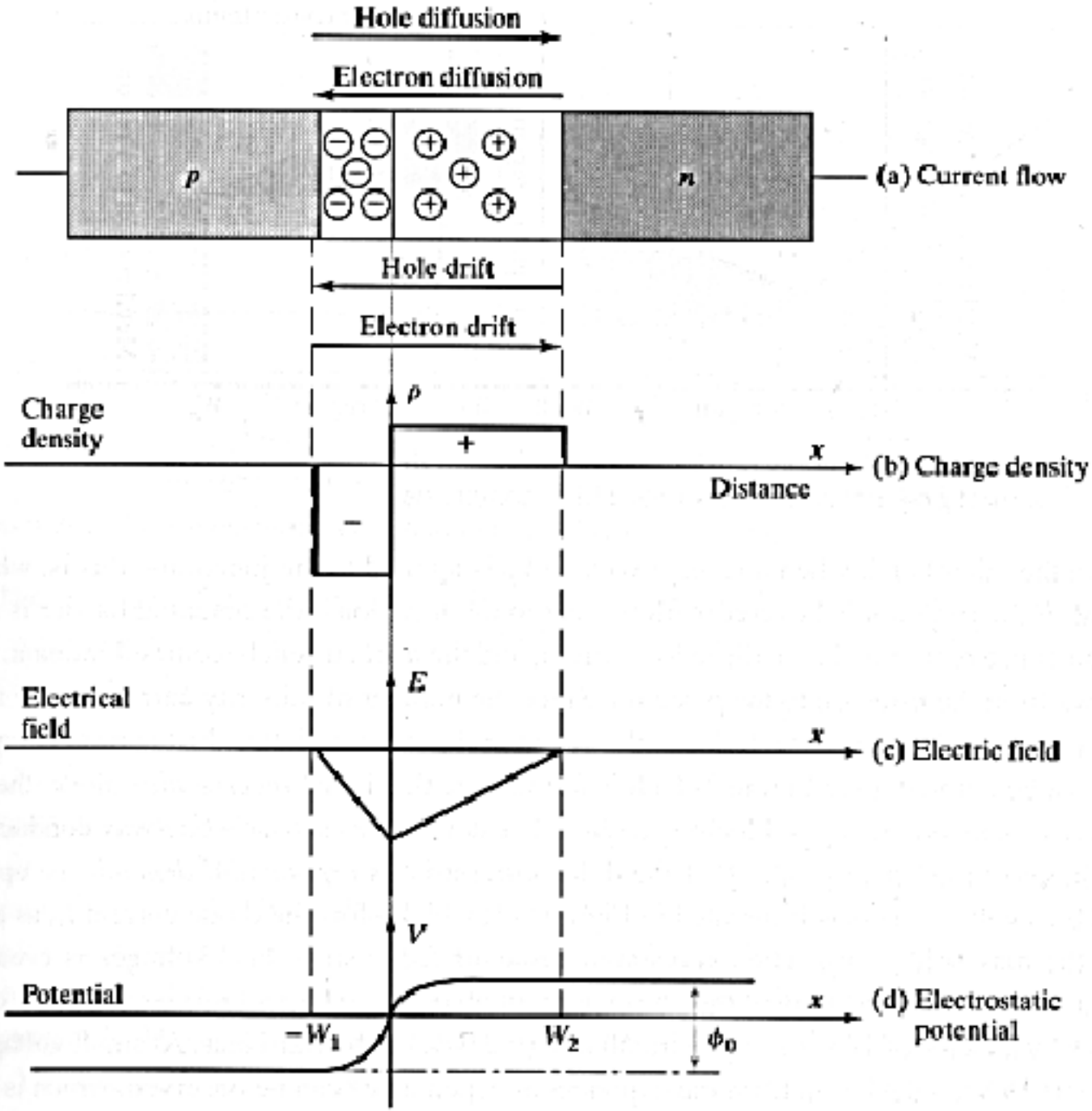


One-dimensional representation

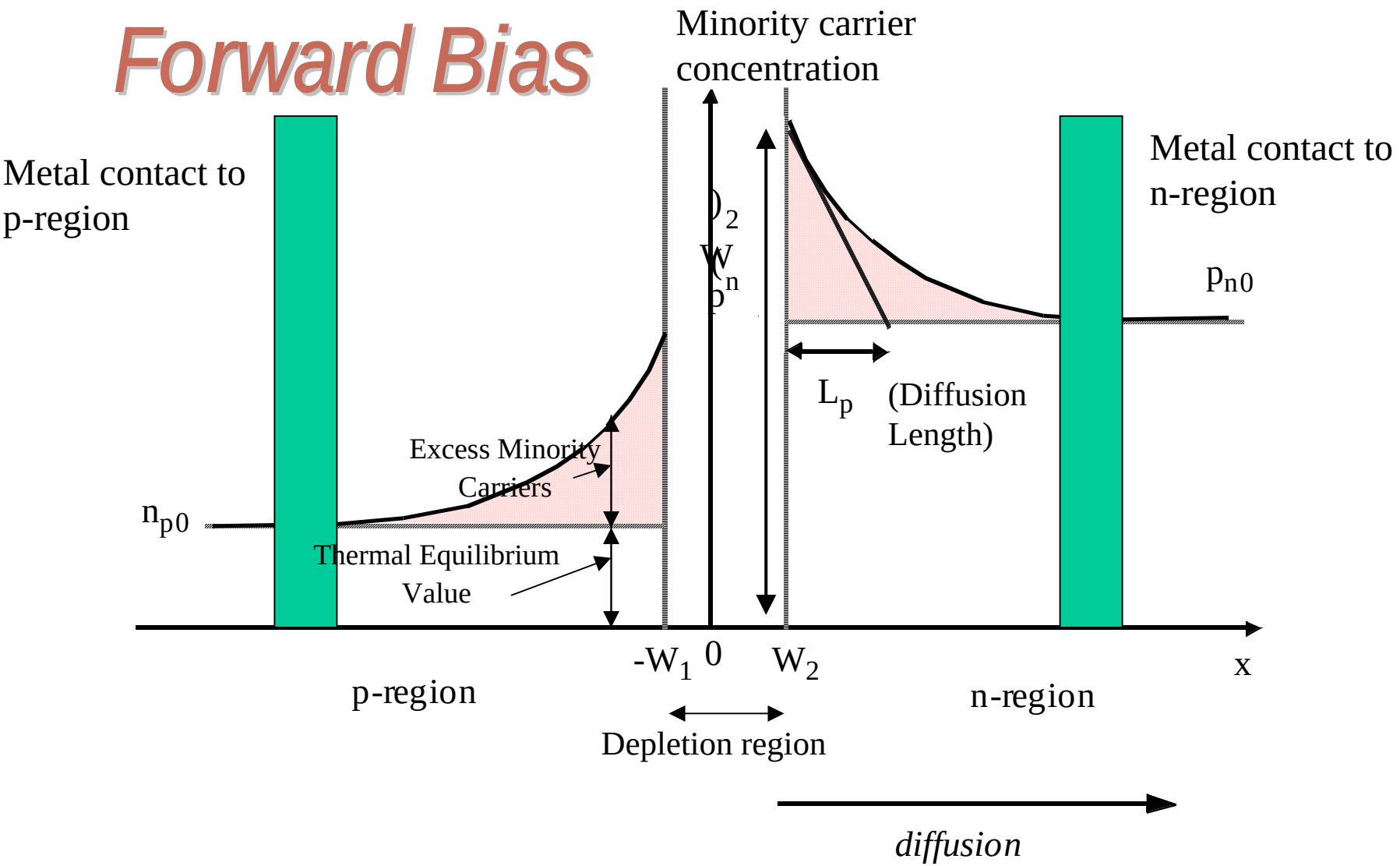


diode symbol

Mostly occurring as parasitic element in Digital ICs

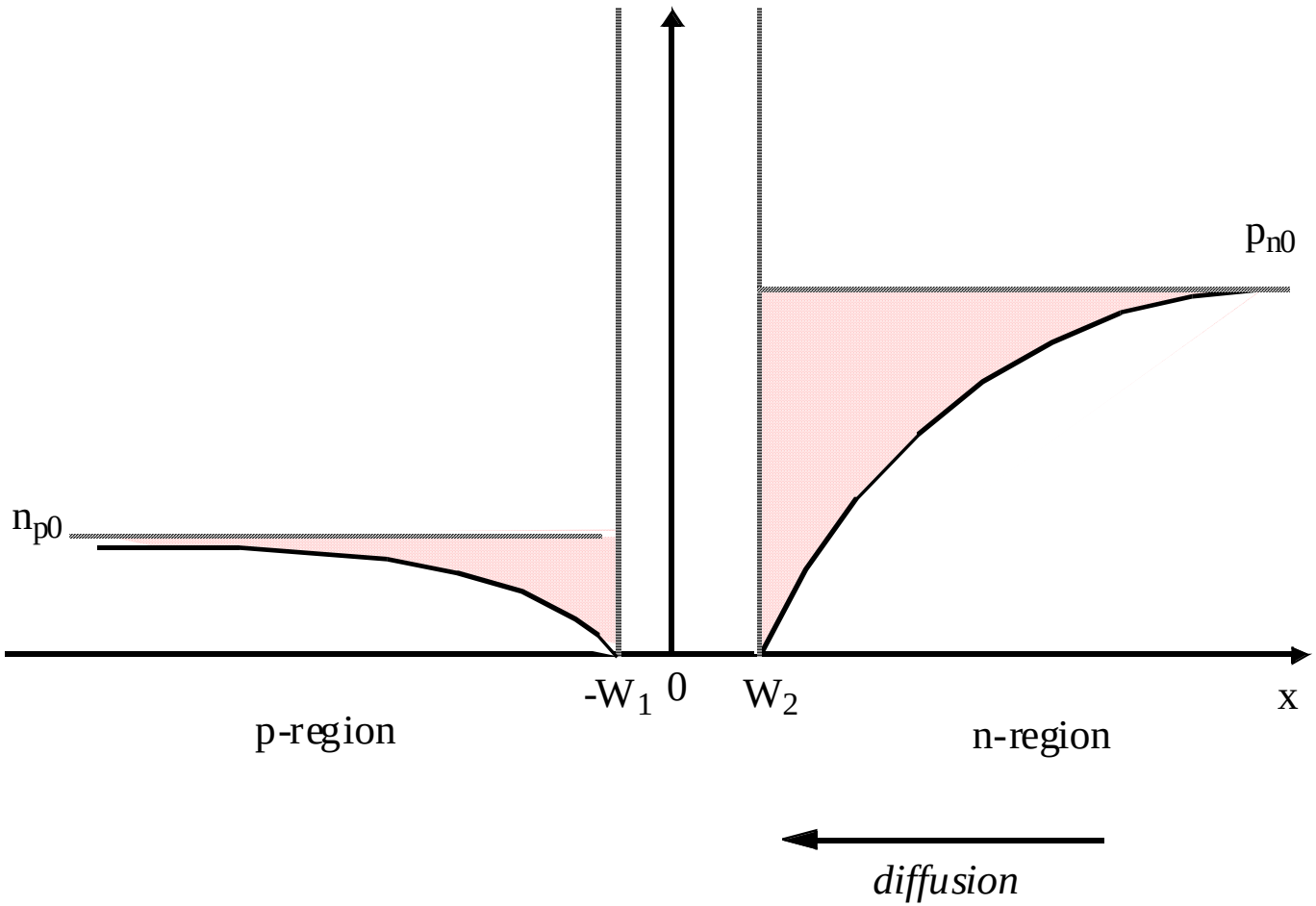

Hole diffusion
Electron diffusion
p
n
(a) Current flow
Hole drift
Electron drift
Charge density
ρ
+
−
x
Distance
(b) Charge density
Electrical field
E
x
(c) Electric field
Potential
V
x
$-W_1$
W_2
ϕ_0
(d) Electrostatic potential

Forward Bias



Typically avoided in Digital ICs

Reverse Bias



The Dominant Operation Mode

$$I_D = I_S(e^{V_D/\Phi_T} - 1)$$

$$\text{Build-in potential} = \Phi_0 = \Phi_T \ln\left[\frac{N_A N_D}{n_i^2}\right]$$

$$\text{Thermal Voltage} = \Phi_T = \frac{kT}{q} = 26mV \quad at\ 300\text{K}$$

N_A = Acceptor concentration

N_D = Donor concentration

n_i = intrinsic carrier concentration $\cong 1.5 \times 10^{10}\ cm^{-3}$

for silicon, at 300K

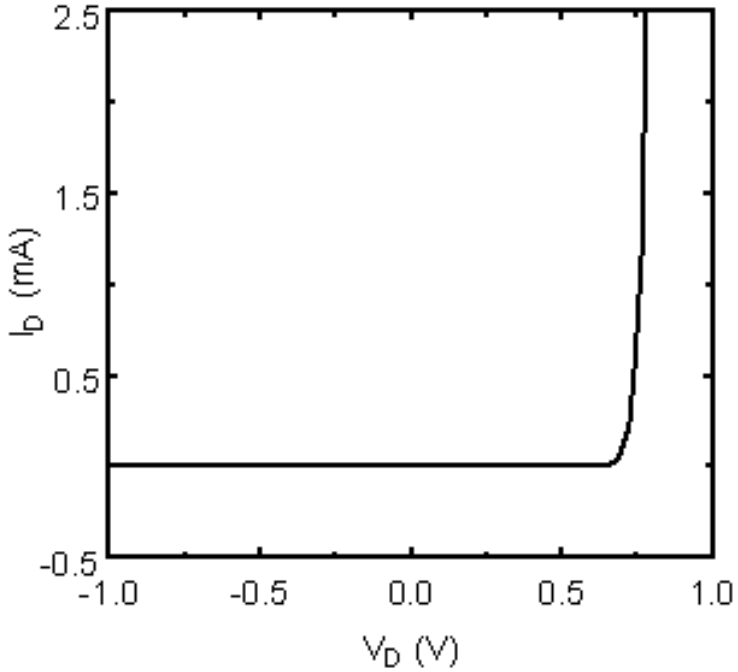
k = Boltzmann's constant = 1.38×10^{-23}joules/kelvin

T = Absolute temperature in kelvins = 273 + C^o

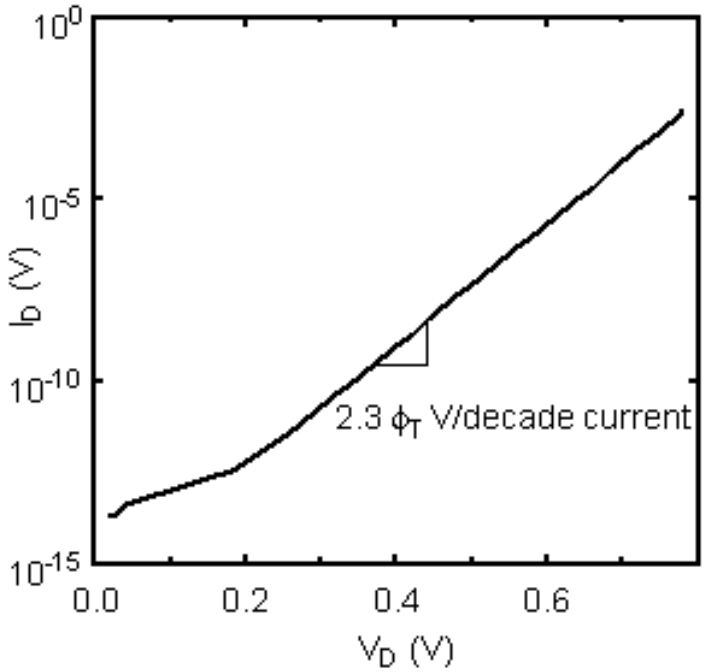
q = Electric charge = 1.6×10^{-19} coulomb

- I_S is usually called the saturation current and it is directly proportional to the cross-sectional area of the diode.
- The value of I_S is a very strong function of the temperature and it generally doubles for every 5^oC rise in temperature.
- In actual devices, the reverse current is substantially (about 3 times) larger than I_S. That is due to the thermal generation of hole and electron pairs in the depletion region, which are swept out by the electric field, generating additional current component. Reverse current doubles for every 10^oC in temperature.
- For typical silicon junctions, I_S is nominally in the range of 10^{-17} $A/\mu m^2$
- ϕ_0 is also called the cut-in voltage (0.6 – 0.8V)

Diode Current



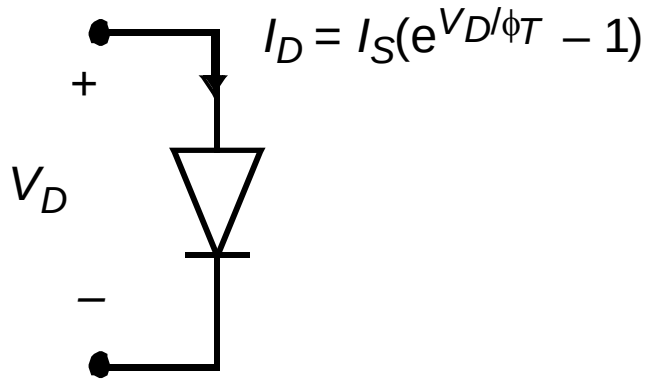
(a) On a linear scale.



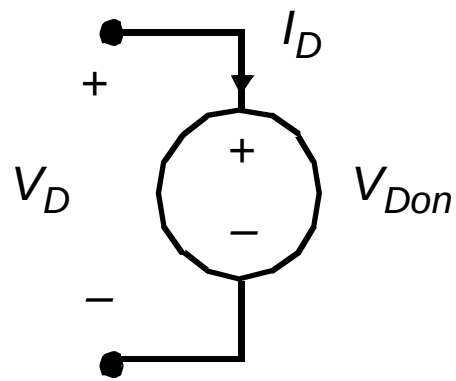
(b) On a logarithmic scale (forward bias).

$$I_D = I_S\left(e^{V_D/\phi_T} - 1\right)$$

Models for Manual Analysis



(a) Ideal diode model



(b) First-order diode model

Consider the simple network of Figure 3-7 and assume that $V_S = 3$ V, $R_S = 10$ kΩ, and $I_S = 0.5 \times 10^{-16}$ A. The diode current and voltage are related by the following network equation:

$$V_S - R_S I_D = V_D$$

Inserting the ideal diode equation and (painfully) solving the nonlinear equation using numerical or iterative techniques yields the following solution: $I_D = 0.224$ mA, and $V_D = 0.757$ V. The simplified model with $V_{Don} = 0.7$ V produces similar results ($V_D = 0.7$ V, $I_D = 0.23$ A) with far less effort. Hence, it makes considerable sense to use this model when determining a first-order solution of a diode network.

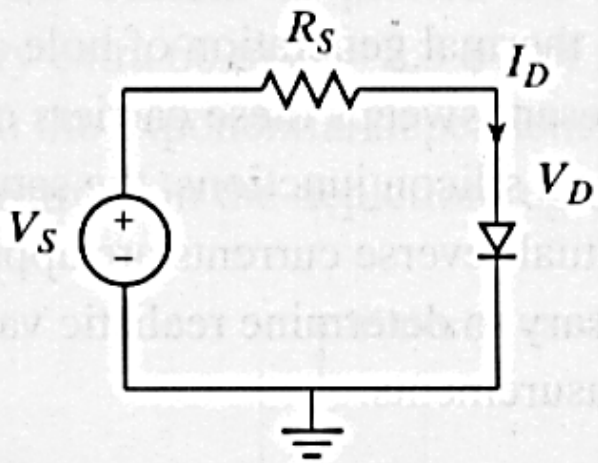


Figure 3-7 A simple diode circuit.

Depletion region charge:

Space charge is equal at both sides (p&n) of depletion region

$$-qW_1N_AA_D = qW_2N_DA_D \rightarrow -\frac{W_2}{W_1} = \frac{N_A}{N_D} \rightarrow W_1 = -W_2\frac{N_D}{N_A} \text{ ..... (1)}$$

Depletion region width: $W_j = W_2 - W_1 \rightarrow W_2 = W_j + W_1$ ..... (2)

Inserting (1) into (2) $\rightarrow W_2 = W_j - W_2\frac{N_D}{N_A} \rightarrow W_2\left(1+\frac{N_D}{N_A}\right) = W_j$

$$W_2 = \frac{W_j}{1+N_D/N_A} \text{ ..... (3)}$$

Note that W_1 is negative

Depletion region charge: Q_j $\quad Q_j = Q_N = qW_2N_DA_D$ ..... (4)

Inserting (3) into (4) $\rightarrow Q_j = qN_D\frac{W_jA_D}{1+N_D/N_A} = \frac{qN_DW_jA_D}{(N_A+N_D)/N_A}$

$$Q_j = q\frac{N_A N_D}{N_A + N_D} A_D W_j \quad \text{..... (5)}$$

Device physics state that: $W_j = \sqrt{\frac{2\varepsilon_{si}}{q}\frac{N_A + N_D}{N_A N_D}(\Phi_0 - V_D)}$ ..... (6)

Inserting (6) into (5) $\rightarrow$ $Q_j = q\frac{N_A N_D}{N_A + N_D} A_D \sqrt{\frac{2\varepsilon_{si}}{q}\frac{N_A + N_D}{N_A N_D}(\Phi_0 - V_D)}$

$$Q_j = A_D \sqrt{\frac{2\varepsilon_{si}\cancel{q^2}}{\cancel{q}}\frac{(N_A N_D)^{\cancel{2}}}{(N_A + N_D)^{\cancel{2}}}\frac{\cancel{N_A + N_D}}{\cancel{N_A N_D}}(\Phi_0 - V_D)}$$

$$\boxed{Q_j = A_D \sqrt{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D}\right)(\Phi_0 - V_D)}}$$

Note that, V_D is negative for reverse bias

Maximum electric field:

Gaussian Law: $\varepsilon_0 \Phi_E = \varepsilon_0 EA = q \;\rightarrow\; E = \dfrac{q}{\varepsilon_0 A}$

$$E_j = \frac{Q_j}{\varepsilon_{si} A_D} = \frac{\cancel{A_D}\sqrt{2\varepsilon_{si} q \dfrac{N_A N_D}{N_A + N_D}(\Phi_0 - V_D)}}{\varepsilon_{si}\cancel{A_D}}$$

$$E_j = \sqrt{\frac{2\cancel{\varepsilon_{si}} q \dfrac{N_A N_D}{N_A + N_D}(\Phi_0 - V_D)}{\varepsilon_{si}^{\cancel{2}}}} = \boxed{\sqrt{\frac{2q}{\varepsilon_{si}} \frac{N_A N_D}{N_A + N_D}(\Phi_0 - V_D)}}$$

Junction Capacitance:

Because the space charge region contains few mobile carriers,

it acts as an insulator with a dielectric constant ε_{si}

n and p regions act as capacitor plates.

$$Q_j = A_D \sqrt{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)(\Phi_0 - V_D)} = A_D \left[\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)(\Phi_0 - V_D) \right]^{1/2}$$

$$C_j = \frac{dQ_j}{dV_D} = A_D \frac{1}{2} \left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right) \left[\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)(\Phi_0 - V_D) \right]^{-1/2}$$

$$= \frac{A_D}{2} \frac{2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D}}{\sqrt{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)(\Phi_0 - V_D)}} = \frac{A_D}{2} \sqrt{\frac{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)^2}{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right)(\Phi_0 - V_D)}}$$

$$= A_D \sqrt{\frac{1}{\cancel{4}\cancel{2}} \frac{\cancel{2}\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D}}{\Phi_0 - V_D}} = A_D \sqrt{\frac{1}{2} \varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} (\Phi_0 - V_D)^{-1}}$$

let $C_{j0} = C_j$ when $V_D = 0$ (zero bias) $\rightarrow$ $C_{j0} = A_D \sqrt{\frac{\varepsilon_{si} q}{2} \frac{N_A N_D}{N_A + N_D} \Phi_0^{-1}}$

$$C_j = A_D \sqrt{\frac{\frac{\varepsilon_{si} q}{2} \frac{N_A N_D}{N_A + N_D}}{\Phi_0 - V_D}} = A_D \sqrt{\frac{\frac{\varepsilon_{si} q}{2} \frac{N_A N_D}{N_A + N_D}}{\Phi_0}} \frac{\Phi_0}{\Phi_0 - V_D}$$

$$C_j = \frac{C_{j0}}{\sqrt{\frac{\Phi_0 - V_D}{\Phi_0}}} \qquad \frac{\Phi_0 - V_D}{\Phi_0} = 1 - \frac{V_D}{\Phi_0}$$

$$\boxed{C_j = \frac{C_{j0}}{\sqrt{1 - \frac{V_D}{\Phi_0}}}}$$

Diffusion Capacitance

The dynamic behavior of the diode is modeled by the nonlinear capacitance C_D, which combines the charge storage effects of the space charge and the excess minority carriers.

$Q =$ excess minority carrier stored charge $= \tau_T I_D = \tau_T I_S (e^{V_D / n\Phi_T} - 1)$

$Q \cong \tau_T I_S e^{V_D / n\Phi_T}$ for $I_D >> I_S$

where $n =$ emission coefficient $= 1$ for most diodes, and

τ_T is the mean transit time. Since $C_d = \dfrac{dQ}{dV_D} =$ diffusion capacitance,

$$C_D = C_j + C_d = \frac{C_{j0}}{(1 - V_D / \Phi_0)^m} + \frac{\tau_T I_S}{\Phi_T} e^{V_D / n\Phi_T}$$

C_d is significant under forward bias conditions, which is not a case in digital integrated circuits. Here, m is called the grading coefficient and it is equal to 1/2 for abrupt junctions.

1. Depletion region charge (V_D is negative for reverse bias):

$$Q_j = A_D \sqrt{\left(2\varepsilon_{si} q \frac{N_A N_D}{N_A + N_D} \right) \left(\Phi_0 - V_D \right)}$$

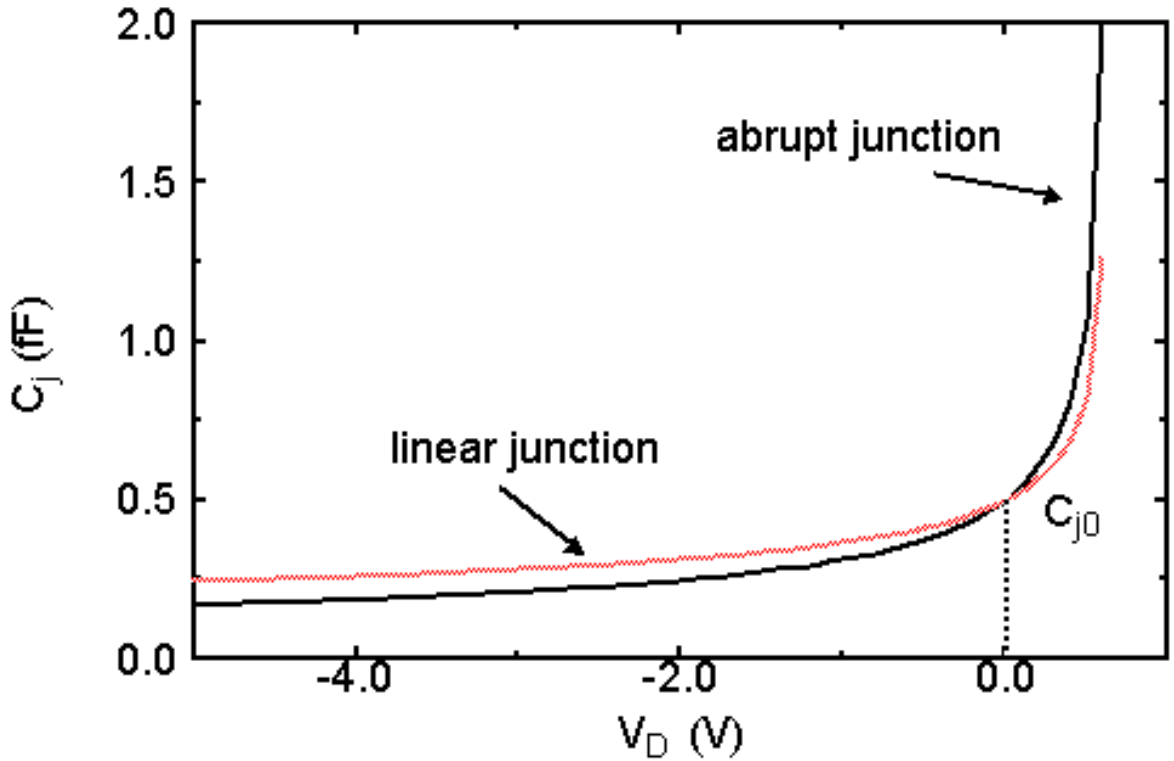
2. Depletion region width:

$$W_j = W_2 - W_1 = \sqrt{\frac{2\varepsilon_{si}}{q} \frac{N_A + N_D}{N_A N_D} \left(\Phi_0 - V_D \right)}$$

3. Maximum electric field:

$$E_j = \sqrt{\left(\frac{2q}{\varepsilon_{si}} \frac{N_A N_D}{N_A + N_D} \right) \left(\Phi_0 - V_D \right)}$$

Junction Capacitance



$$C_j = \frac{C_{j0}}{(1 - V_D/\phi_0)^m}$$

m = 0.5: abrupt junction
m = 0.33: linear junction

$$C_{j0} = A_D \sqrt{\left(\frac{\varepsilon_{si} q}{2} \frac{N_A N_D}{N_A + N_D} \right) \Phi_0^{-1}}$$

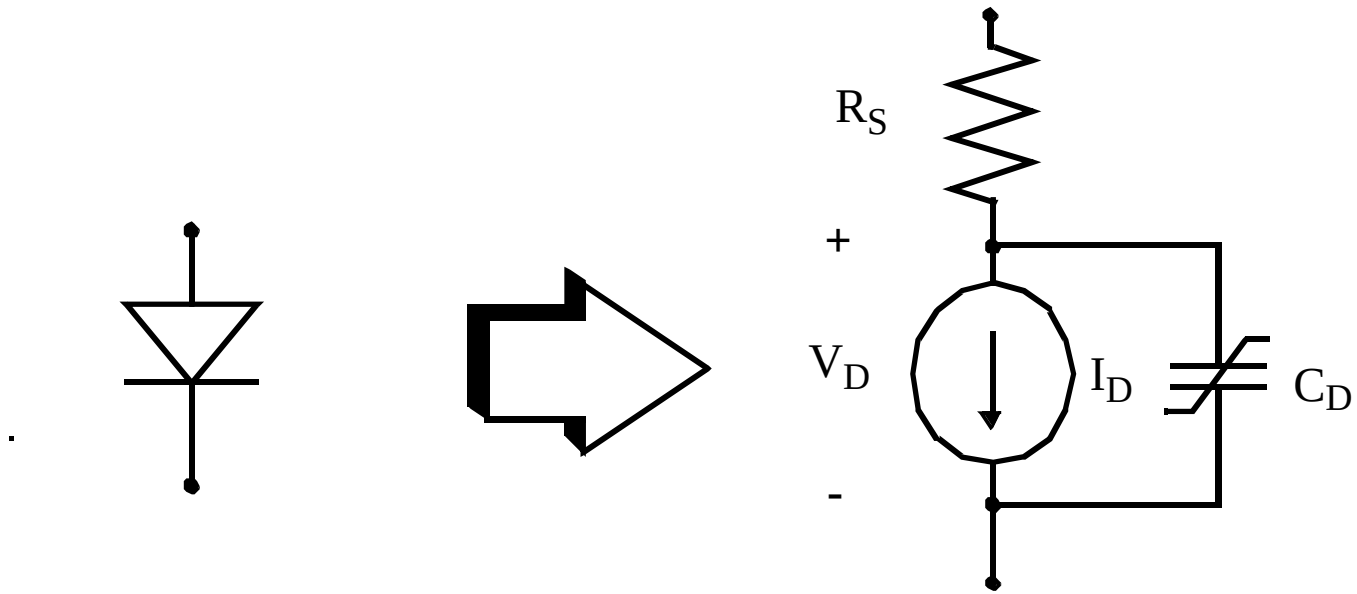
Large-Signal Depletion Region Capacitance

Junction capacitance is a voltage-dependent parameter. In digital circuits, operating voltages tend to move rapidly over a wide range V_{high}-V_{low}. It is more appropriate to replace the voltage dependent nonlinear capacitance C_j by an equivalent linear capacitance C_{eq}

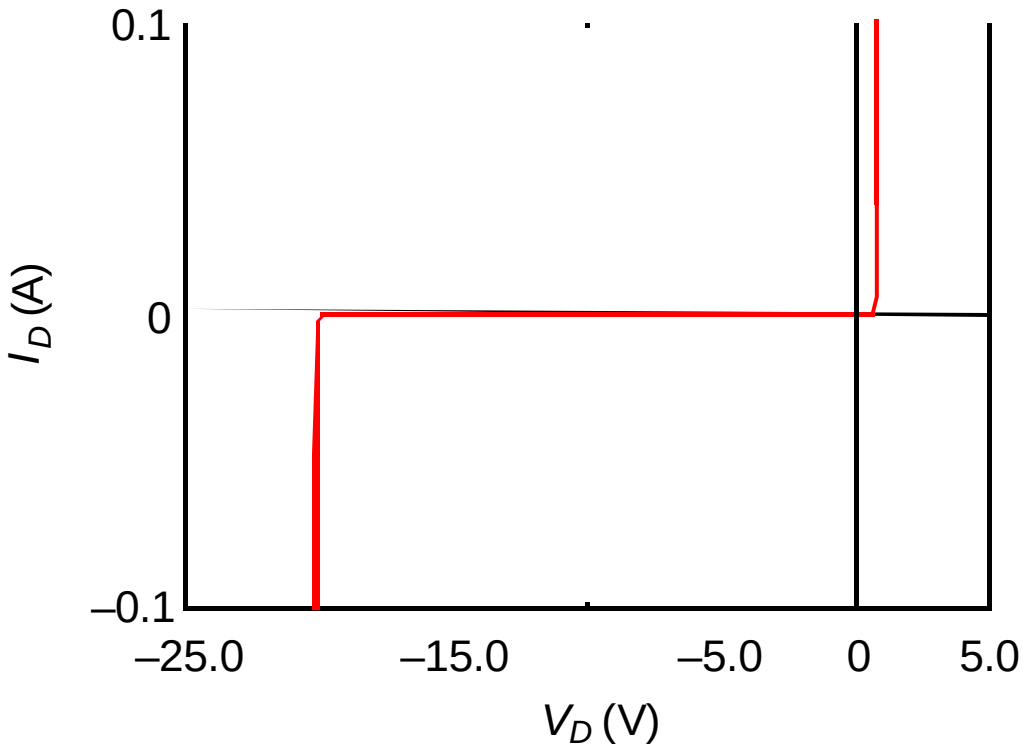
$$C_{eq} = \frac{\Delta Q_j}{\Delta V_D} = \frac{Q_j(high) - Q_j(low)}{V_{high} - V_{low}} = K_{eq} C_{j0}$$

$$K_{eq} = \frac{-\Phi_0^{\,m}}{(V_{high} - V_{low})(1-m)} \left[(\Phi_0 - V_{high})^{1-m} - (\Phi_0 - V_{low})^{1-m} \right]$$

Diode Model



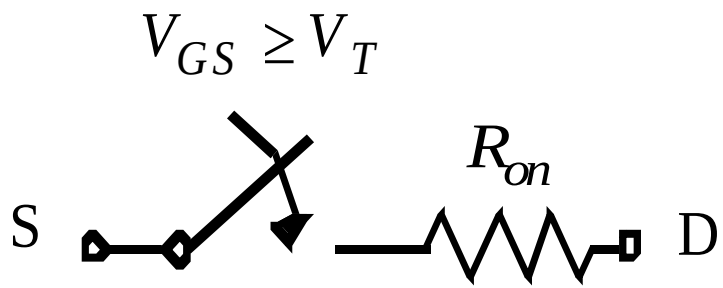
Secondary Effects

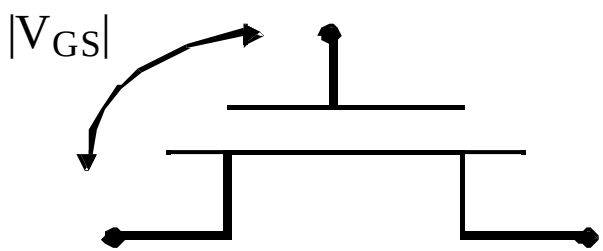


Avalanche Breakdown

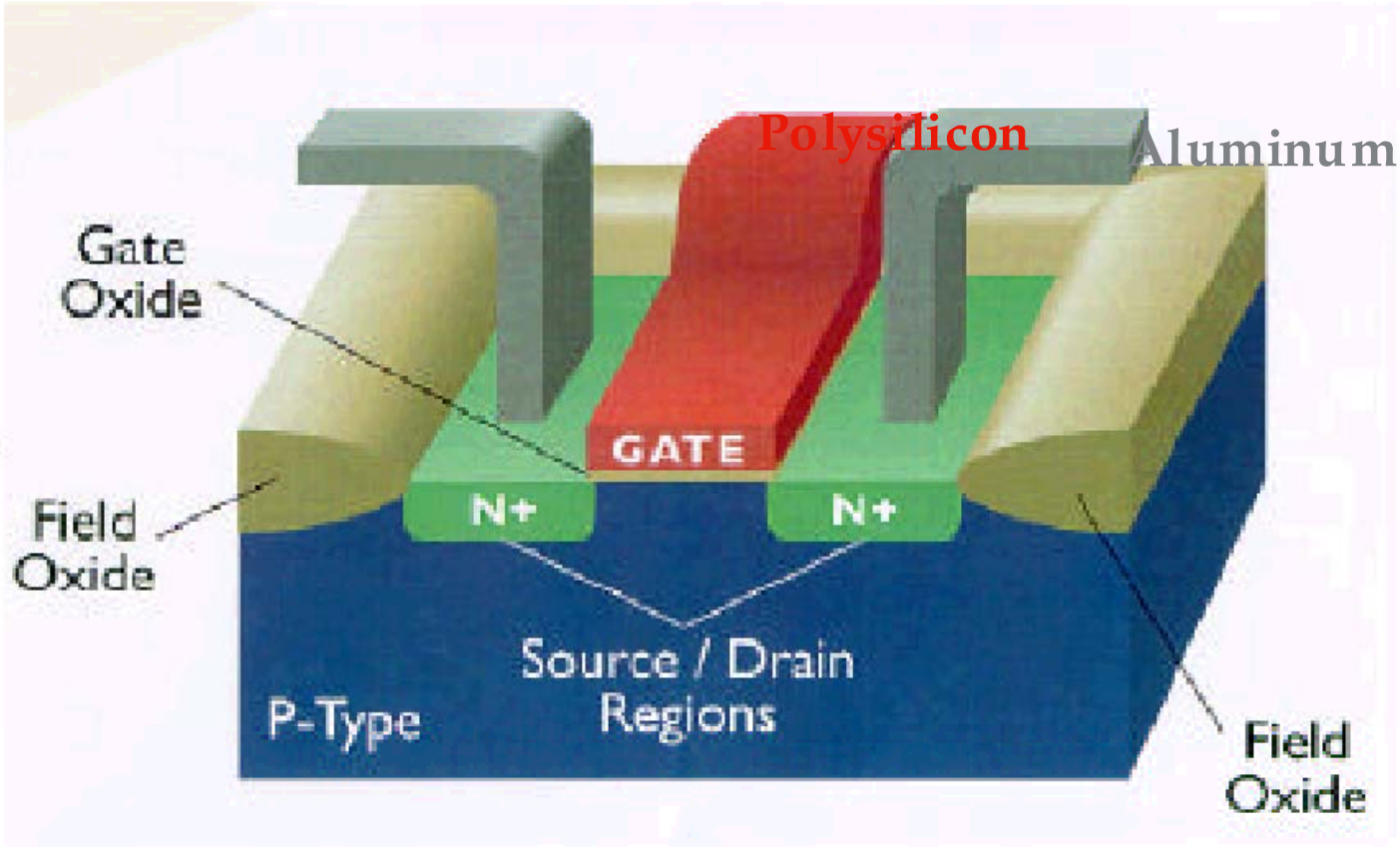
What is a Transistor?

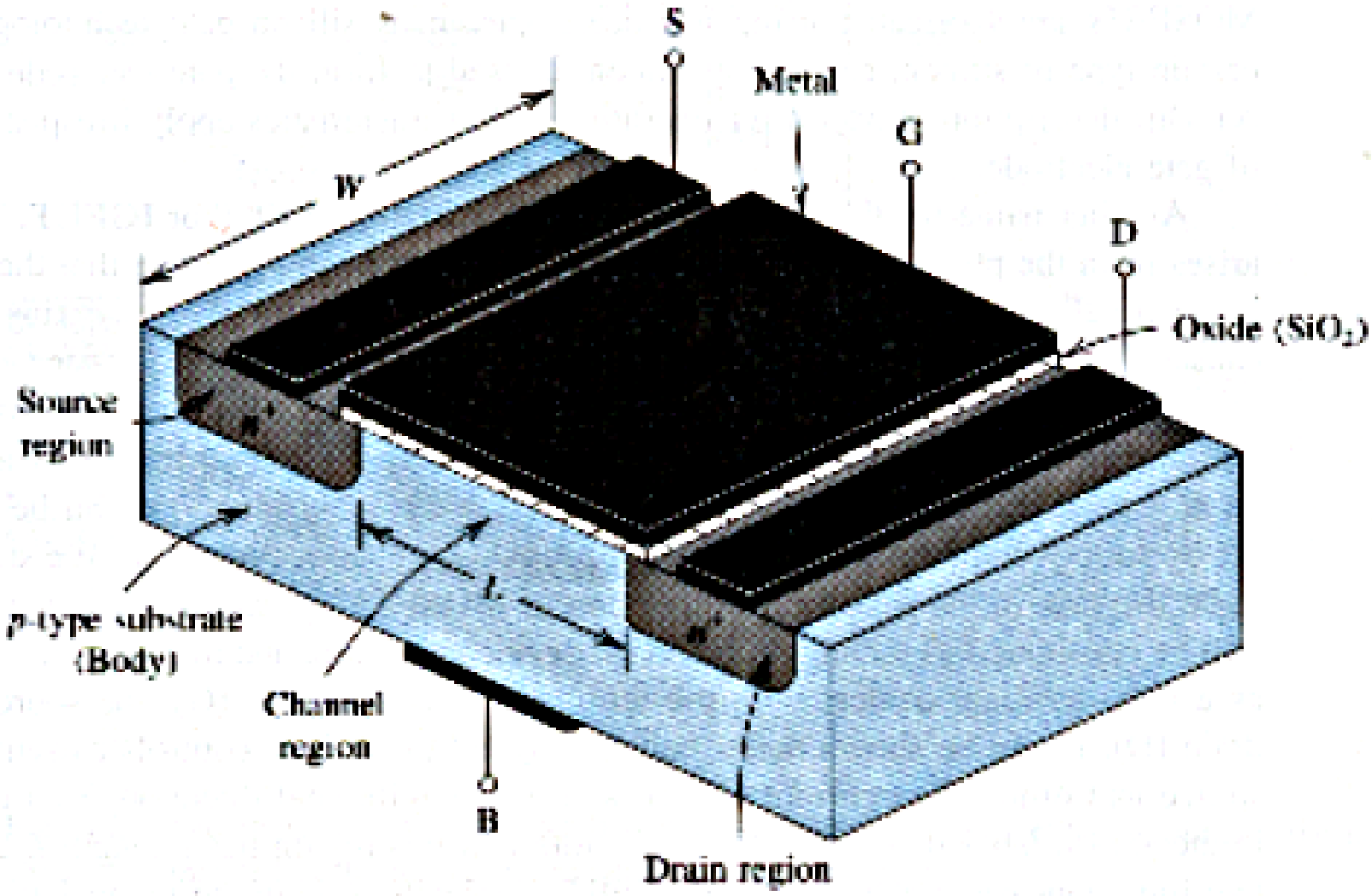
A Switch! ⟷ A MOS Transistor

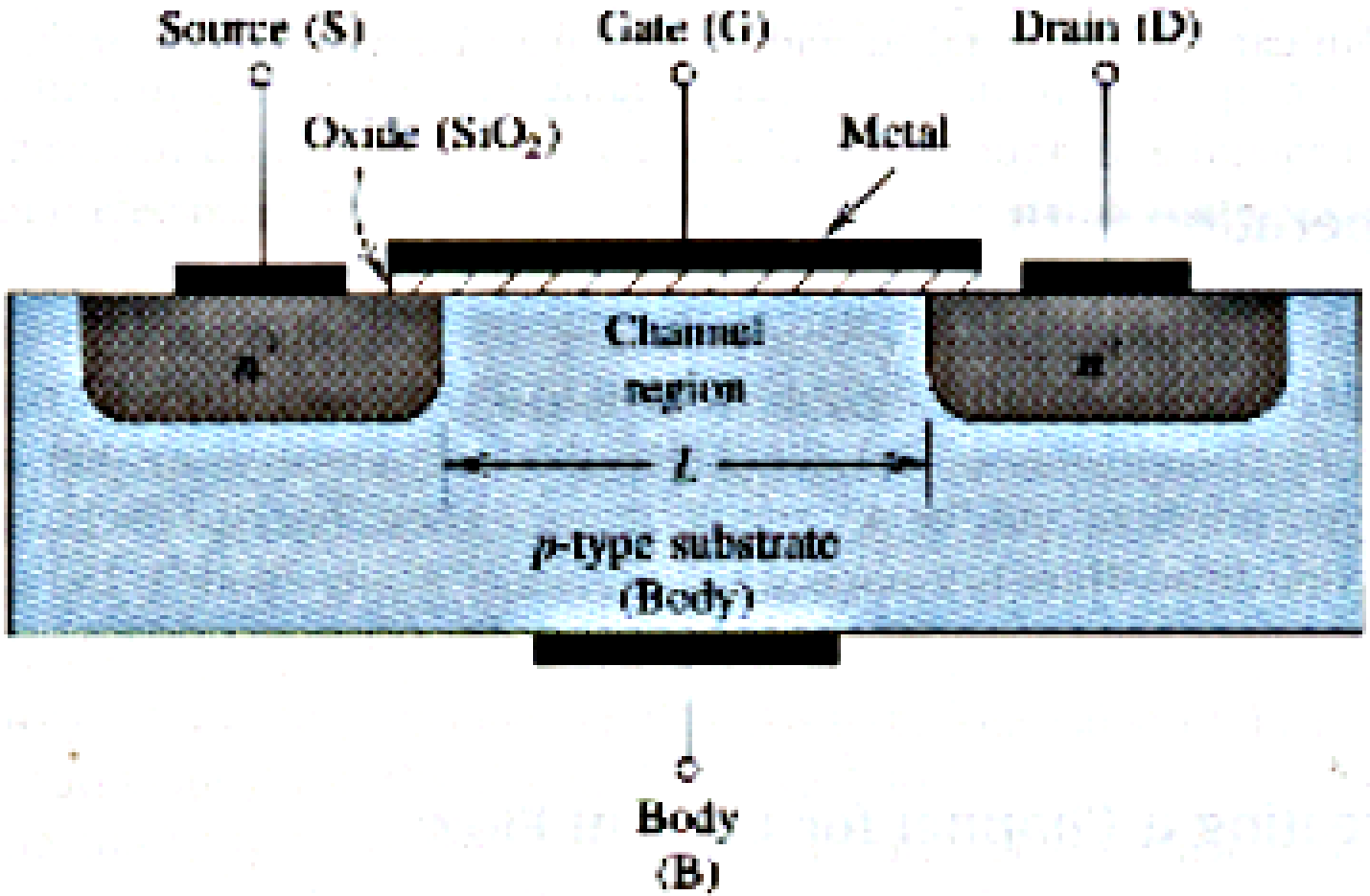




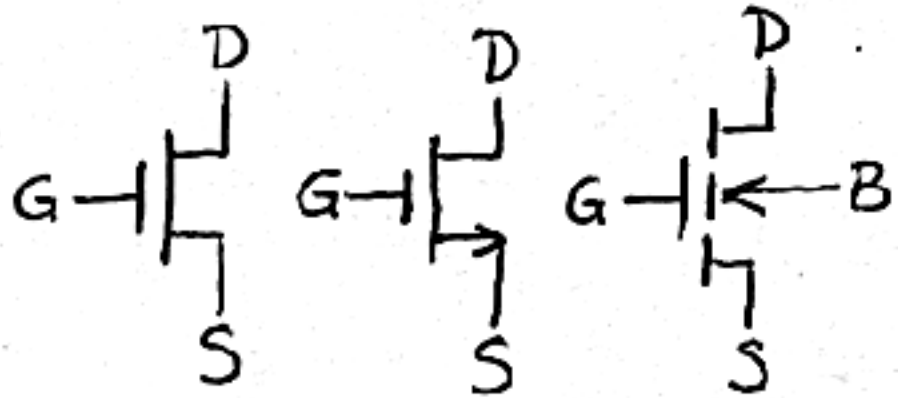
The MOS Transistor

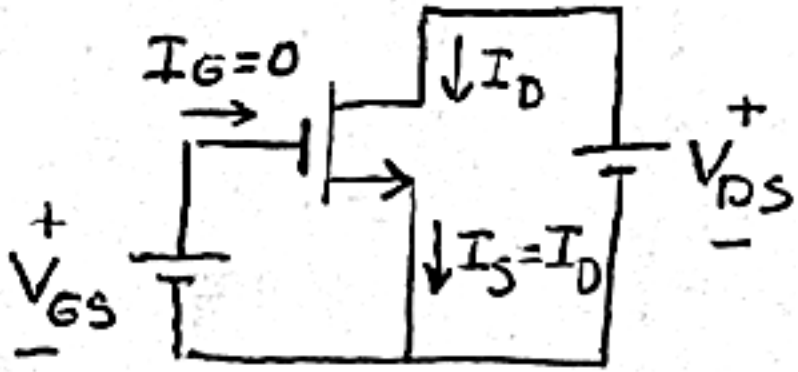


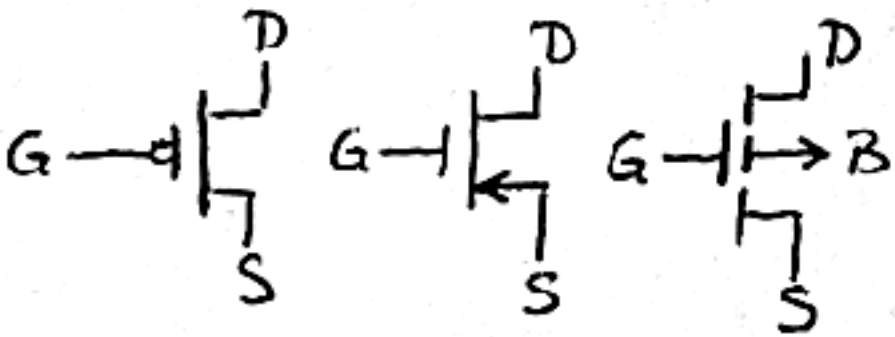

S
Metal
G
D
W
Oxide (SiO_2)
Source region
p-type substrate (Body)
Channel region
L
B
Drain region

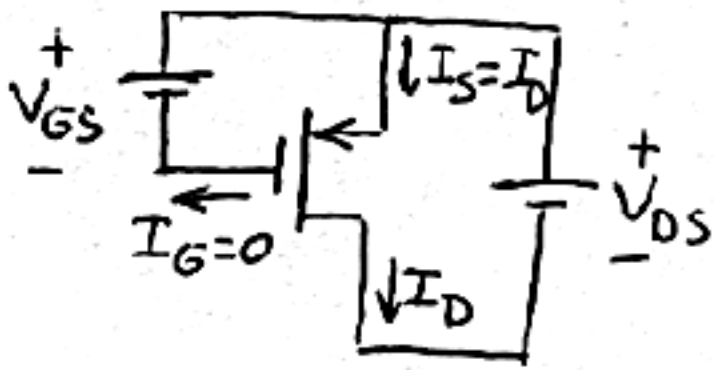

Source (S)
Gate (G)
Drain (D)
Oxide (SiO_2)
Metal
n+
Channel region
n+
L
p-type substrate (Body)
Body (B)

NMOS Enhancement

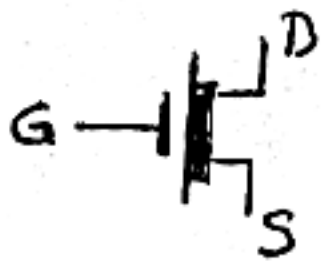






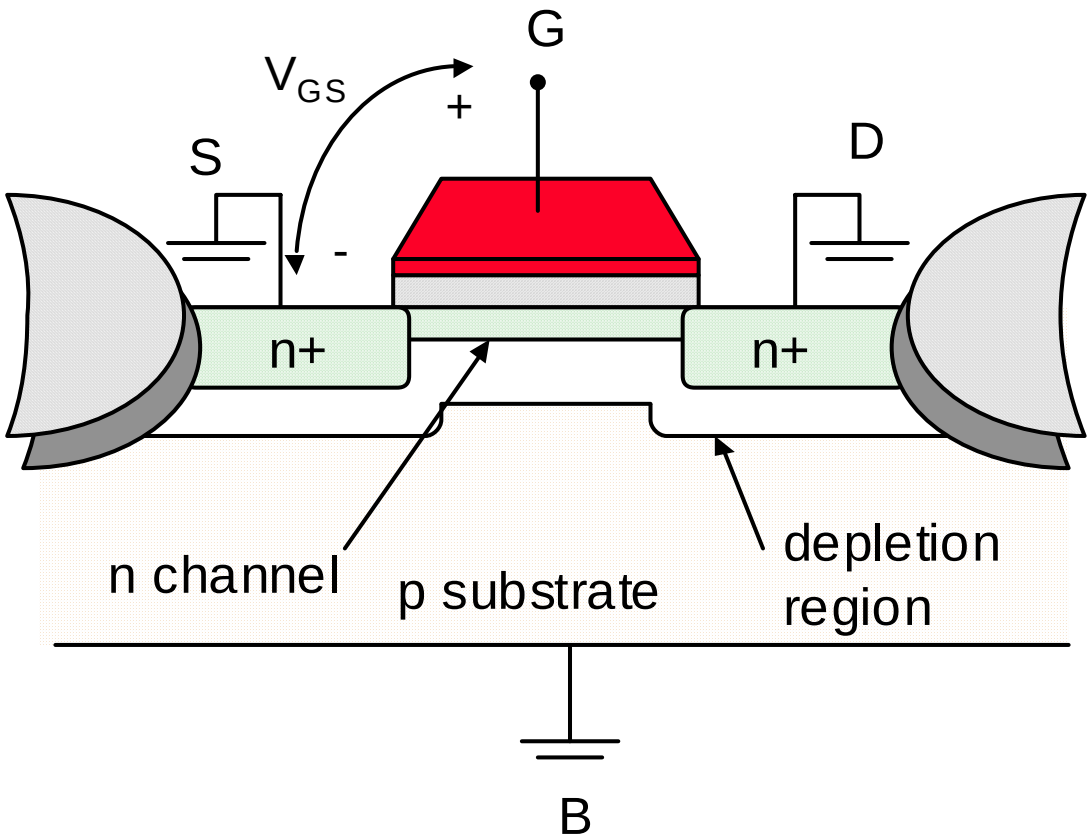


PMOS Enhancement



NMOS Depletion

Threshold Voltage Concept



The value of V_{GS} where strong inversion occurs is called the threshold voltage, V_T

$$W_d = \sqrt{\frac{2\varepsilon_{si}\,\Phi}{q\,N_A}} \qquad Q_d = \sqrt{2q\,N_A\,\varepsilon_{si}\,\Phi}$$

Where ϕ is the voltage across the depletion layer (potential at the oxide-silicon boundary, N_A is the substrate doping.

In the presence of an inversion layer, the charge stored in the depletion region is fixed and equal to:

$$Q_{B0} = \sqrt{2q\,N_A\,\varepsilon_{si}\left|2\Phi_F\right|}$$

When a substrate bias voltage V_{SB} is applied, the surface potential required for strong inversion increases and the charge stored can be expressed as:

$$Q_B = \sqrt{2q\,N_A\,\varepsilon_{si}\left(\left|(-2)\,\Phi_F + V_{SB}\right|\right)}$$

The Threshold Voltage

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

V_{T0} is the threshold voltage at $V_{SB} = 0$ and is mostly a function of the manufacturing process

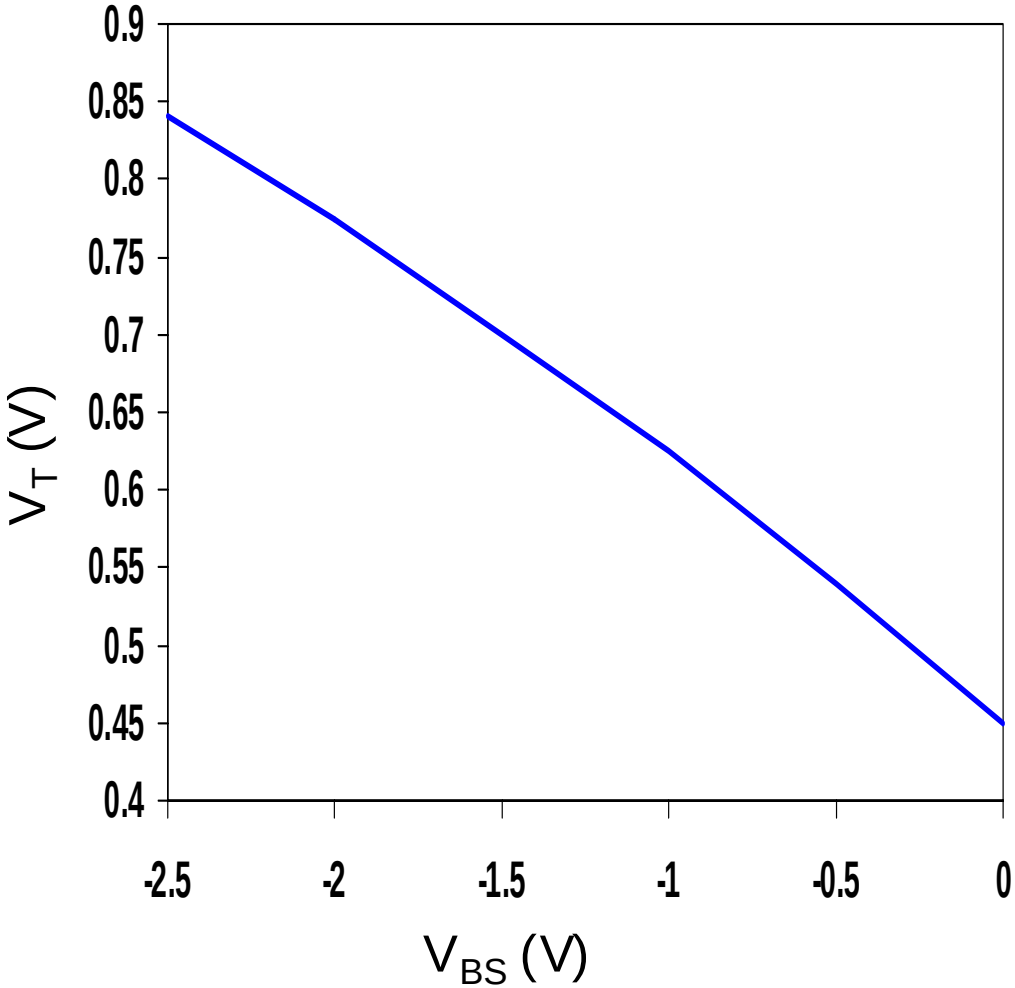
- Difference in work-function between gate and substrate material, oxide thickness, Fermi voltage, charge of impurities trapped at the surface, dosage of implanted ions, etc.

V_{SB} is the source-bulk voltage

$\phi_F = -\phi_T \ln(N_A/n_i)$ is the Fermi potential (~ -0.3V) ($\phi_T = kT/q$ = 26mV at 300K is the thermal voltage; N_A is the acceptor ion concentration; $n_i \approx 1.5 \times 10^{10}$ cm^{-3} at 300K is the intrinsic carrier concentration in pure silicon)

$\gamma = \sqrt{(2q\varepsilon_{si}N_A)}/C_{ox}$ is the body-effect coefficient (impact of changes in V_{SB}) ($\varepsilon_{si}=1.053 \times 10^{-10}$F/m is the permittivity of silicon; $C_{ox} = \varepsilon_{ox}/t_{ox}$ is the gate oxide capacitance with $\varepsilon_{ox}=3.5 \times 10^{-11}$F/m)

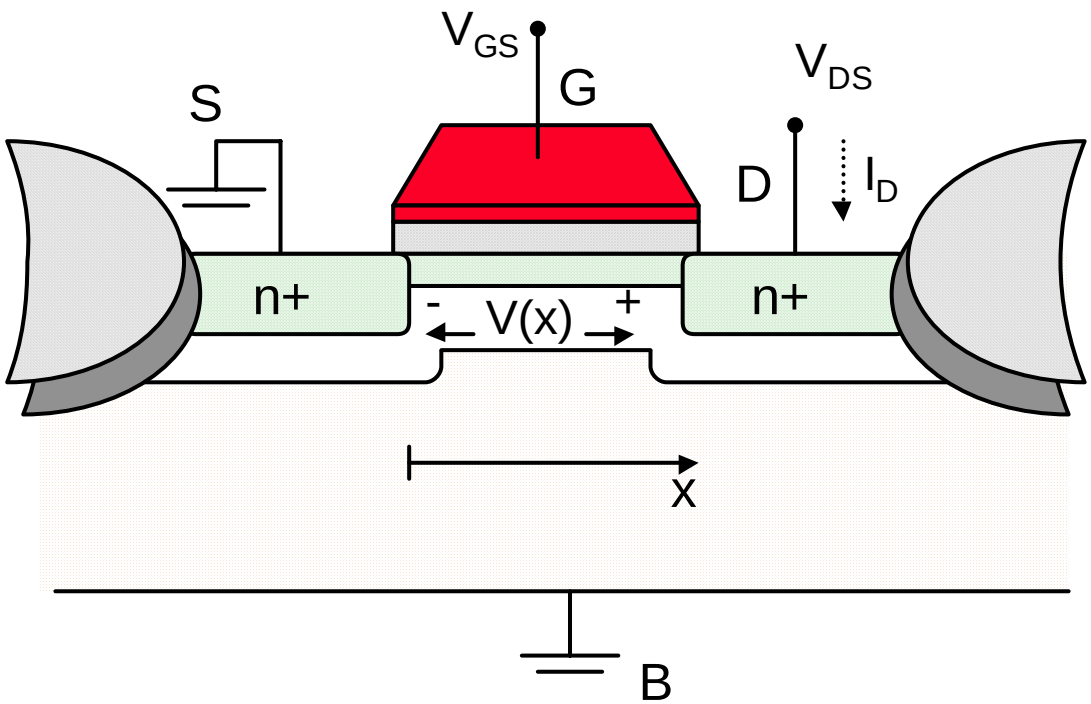
The Body Effect



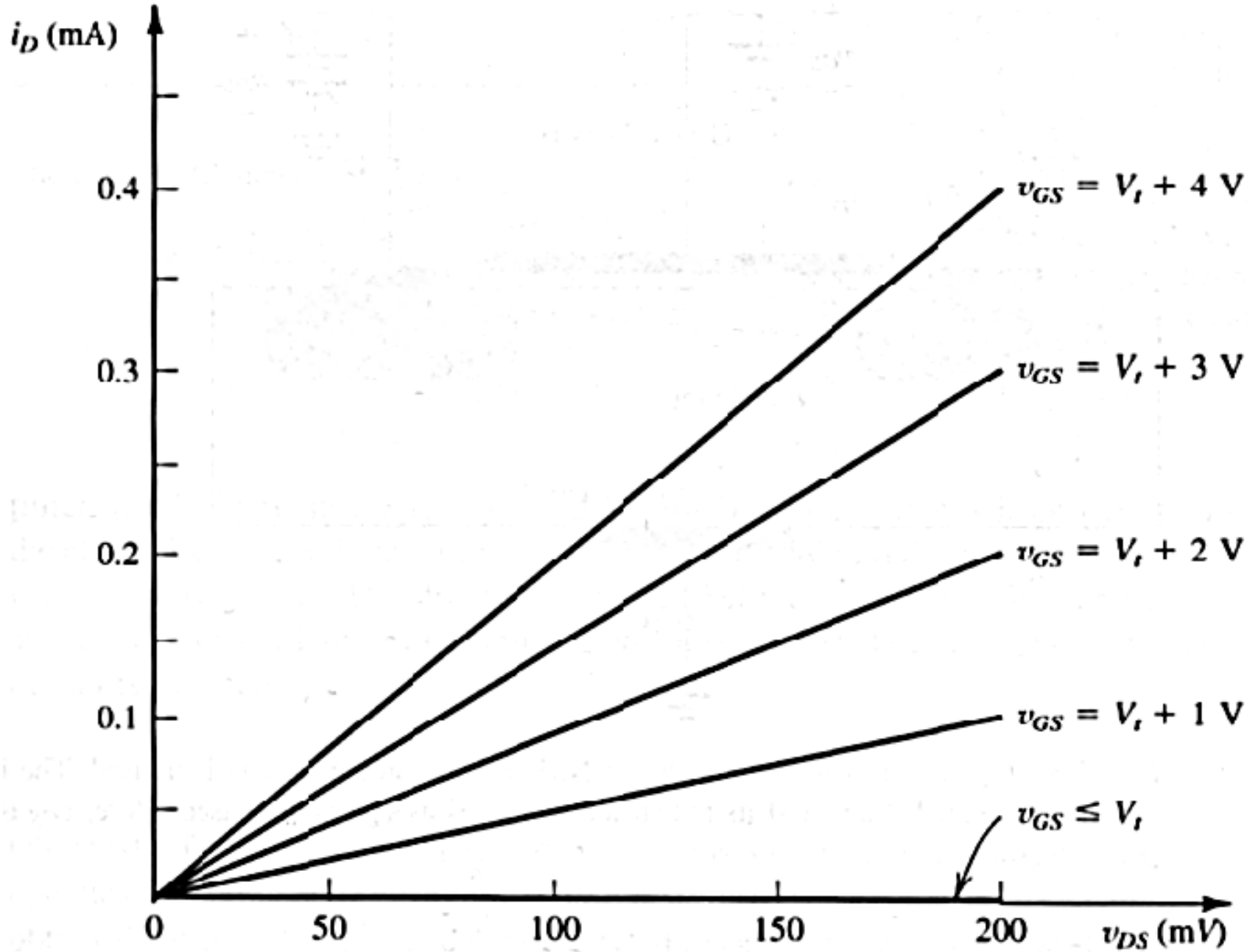
- V_{SB} is the substrate bias voltage (normally positive for n-channel devices with the body tied to ground)
- A negative bias causes V_T to increase from 0.45V to 0.85V

Transistor in Linear Mode

Assuming $V_{GS} > V_T$

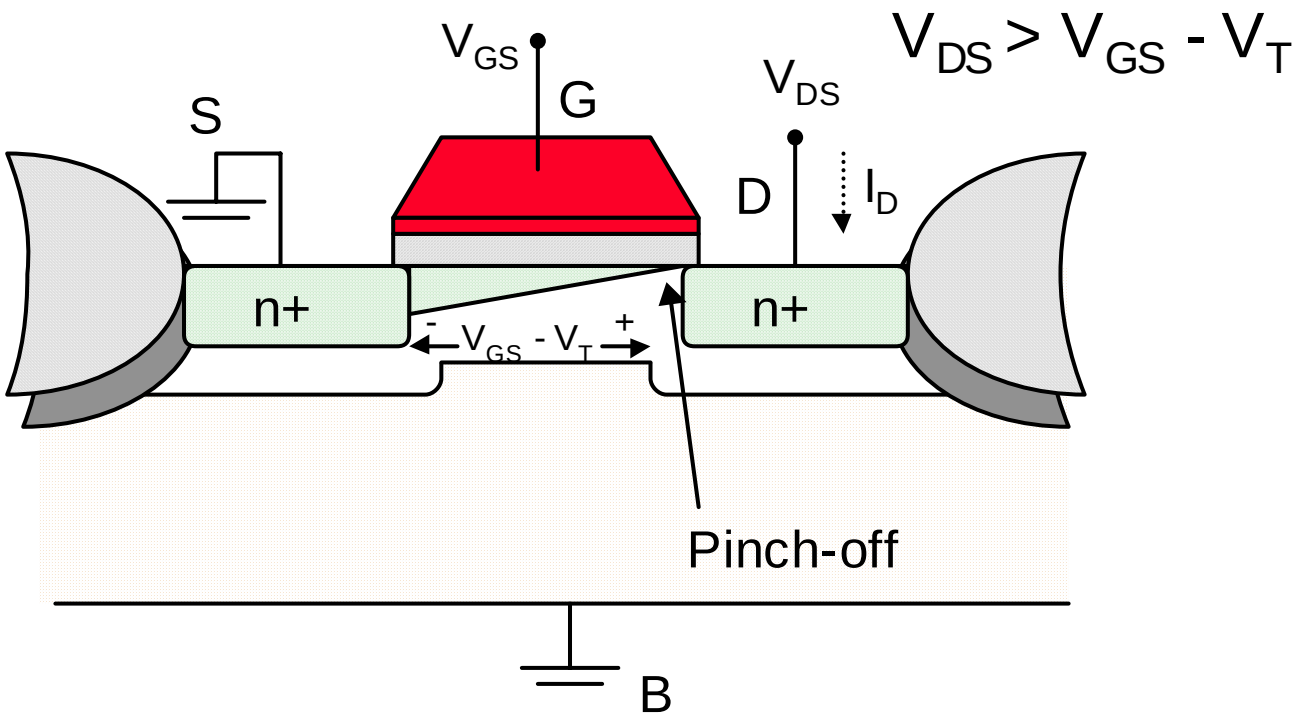


The current is a linear function of both V_{GS} and V_{DS}

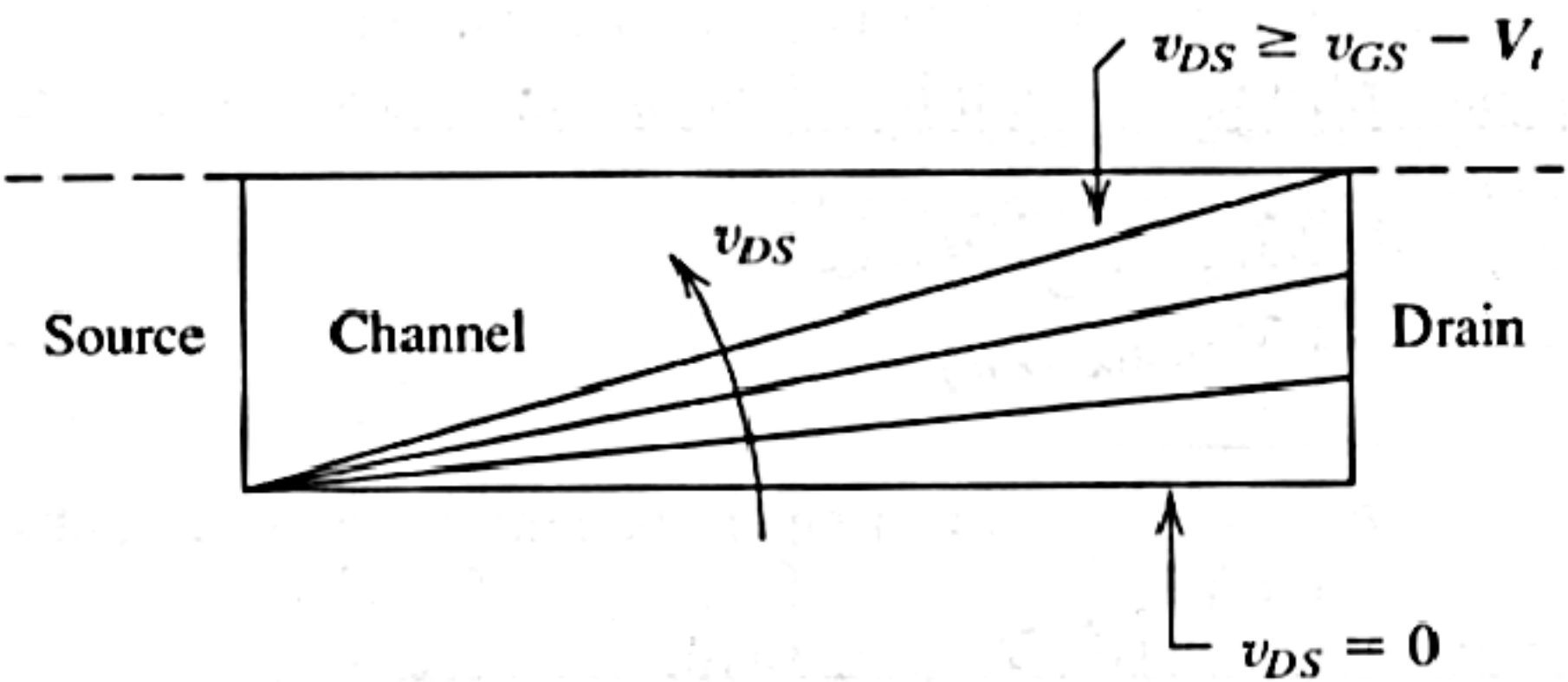

i_D (mA)
0.4
0.3
0.2
0.1
0
50
100
150
200
v_{DS} (mV)
$v_{GS} = V_t + 4$ V
$v_{GS} = V_t + 3$ V
$v_{GS} = V_t + 2$ V
$v_{GS} = V_t + 1$ V
$v_{GS} \leq V_t$

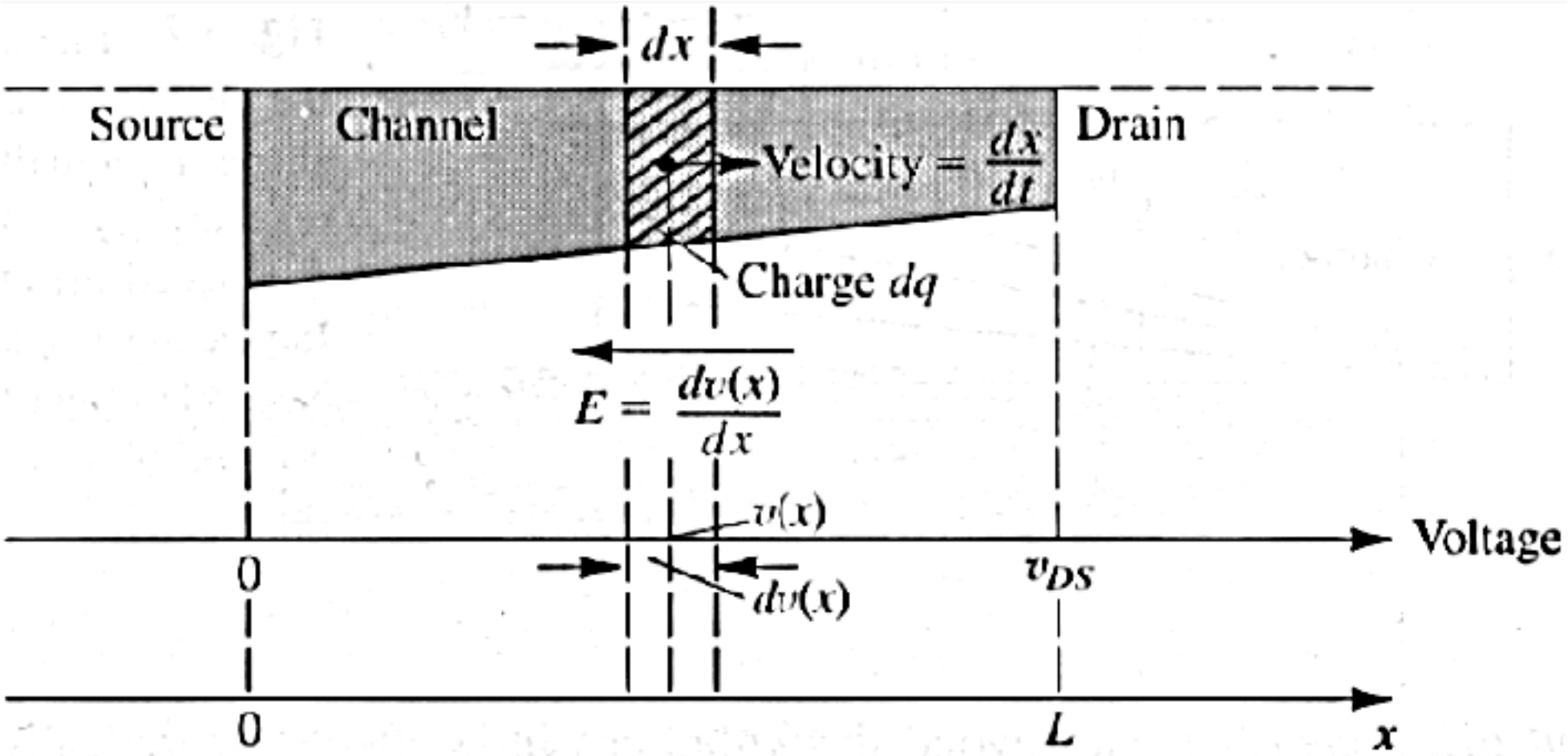
Transistor in Saturation Mode

Assuming $V_{GS} > V_T$



The current remains constant (saturates).


$v_{DS} \geq v_{GS} - V_t$
v_{DS}
Source
Channel
Drain
$v_{DS} = 0$



Assuming $V_{GS}>V_T$ and $V_{DS}<V_{GS}-V_T$ (resistive operation)

Resistive Operation:

$dq(x) = -C_{ox} W dx \left[V_{GS} - v(x) - V_T \right]$ ..... (1)

where $dq(x)$ is the charge in an infinitesimal portion of the channel.

Charge per unit area: $Q_i(x) = \dfrac{dq(x)}{dx W} = -C_{ox} \left[V_{GS} - v(x) - V_T \right]$

Capacitance per unit area: $C_{ox} = \dfrac{\varepsilon_{ox}}{t_{ox}}$ where, ε_{ox}: oxide permittivity,

t_{ox}: oxide thickness. The electric field at point x: $E(x) = -\dfrac{dv(x)}{dx}$

$E(x)$ causes the electron charge $dq(x)$ to drift toward the drain

with a velocity $\dfrac{dx}{dt} = v_n = -\mu_n E(x) = \mu_n \dfrac{dv(x)}{dx}$

since, $i = \dfrac{dq}{dt} = \dfrac{dx}{dt}\dfrac{dq(x)}{dx} = \mu_n \dfrac{dv(x)}{dx}\dfrac{dq(x)}{dx}$ ..... (2)

Using (1) $\rightarrow$ $\dfrac{dq(x)}{dx} = -C_{ox}W\left[V_{GS} - v(x) - V_T\right]$

inserting into (2) $\rightarrow$ $i = -\mu_n C_{ox} W\left[V_{GS} - v(x) - V_T\right]\dfrac{dv(x)}{dx}$

$I_D = -i = \mu_n C_{ox} W\left[V_{GS} - v(x) - V_T\right]\dfrac{dv(x)}{dx}$ rearranging & integrating both sides:

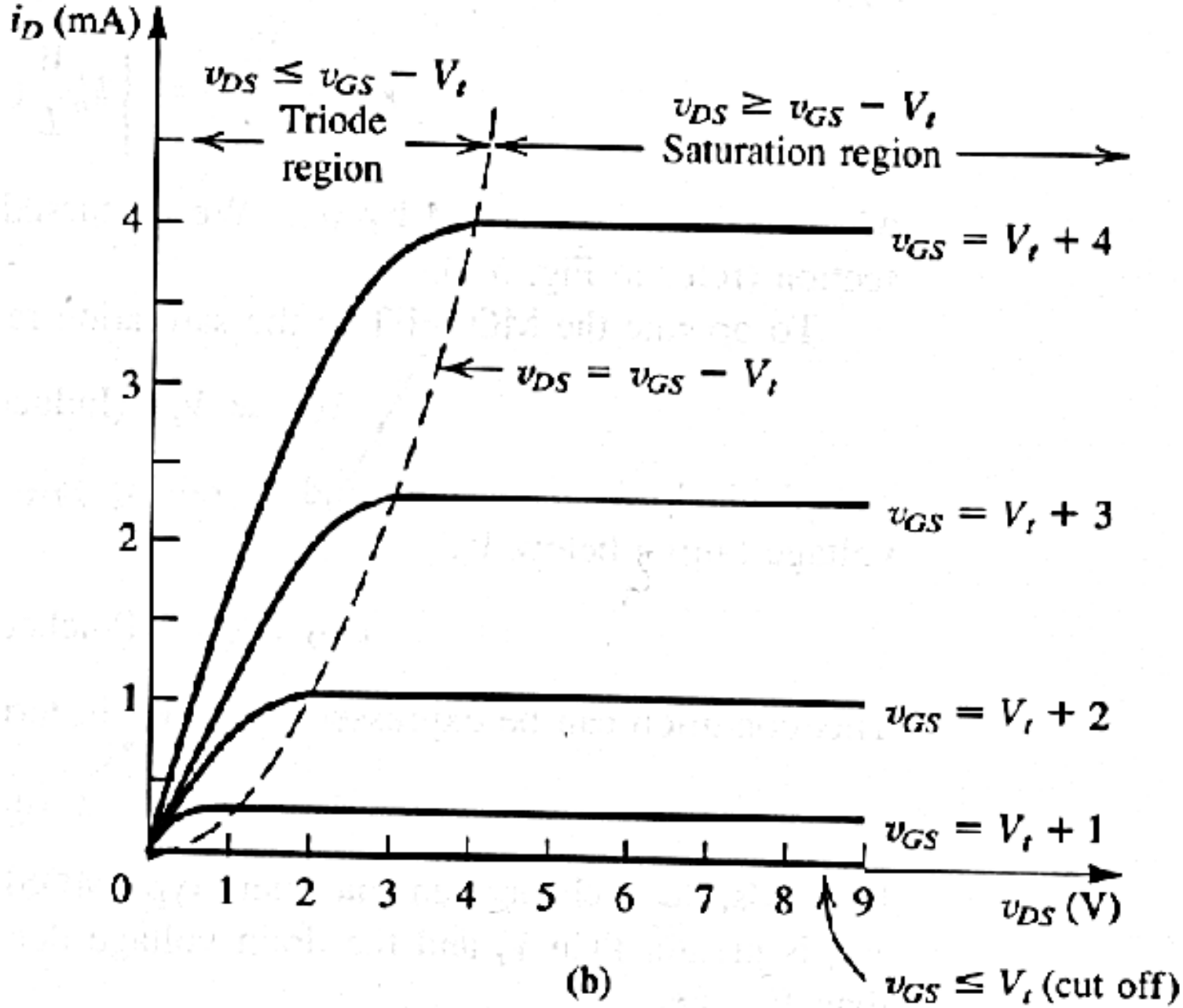
$$\int_0^L I_D\,dx = \int_0^{V_{DS}} \mu_n C_{ox} W\left[V_{GS} - V_T - v(x)\right]\,dv(x)$$

(V_{DS} ← v(L), 0 ← v(0))

$$I_D = \left(\mu_n C_{ox}\right)\left(\frac{W}{L}\right)\left[\left(V_{GS} - V_T\right)V_{DS} - \frac{1}{2}V_{DS}^{\,2}\right]$$

where: $\mu_n C_{ox} = K_n'$: process transconductance parameter,

$K_n = K_n'\dfrac{W}{L}$: gain factor, $\dfrac{W}{L}$: aspect ratio


i_D (mA)
$v_{DS} \le v_{GS} - V_t$
Triode region
$v_{DS} \ge v_{GS} - V_t$
Saturation region
$v_{GS} = V_t + 4$
$v_{DS} = v_{GS} - V_t$
$v_{GS} = V_t + 3$
$v_{GS} = V_t + 2$
$v_{GS} = V_t + 1$
4
3
2
1
0
1
2
3
4
5
6
7
8
9
v_{DS} (V)
(b)
$v_{GS} \le V_t$ (cut off)