LECTURE 05 - PN JUNCTIONS AND CMOS TRANSISTORS

LECTURE ORGANIZATION

Outline

- *pn* junctions
- MOS transistors
- Layout of MOS transistors
- Parasitic bipolar transistors in CMOS technology
- High voltage CMOS transistors
- Summary

***CMOS Analog Circuit Design*, 3rd Edition Reference**

Pages 33-46 and 644-652

PN JUNCTIONS

How are *PN* Junctions used in CMOS?

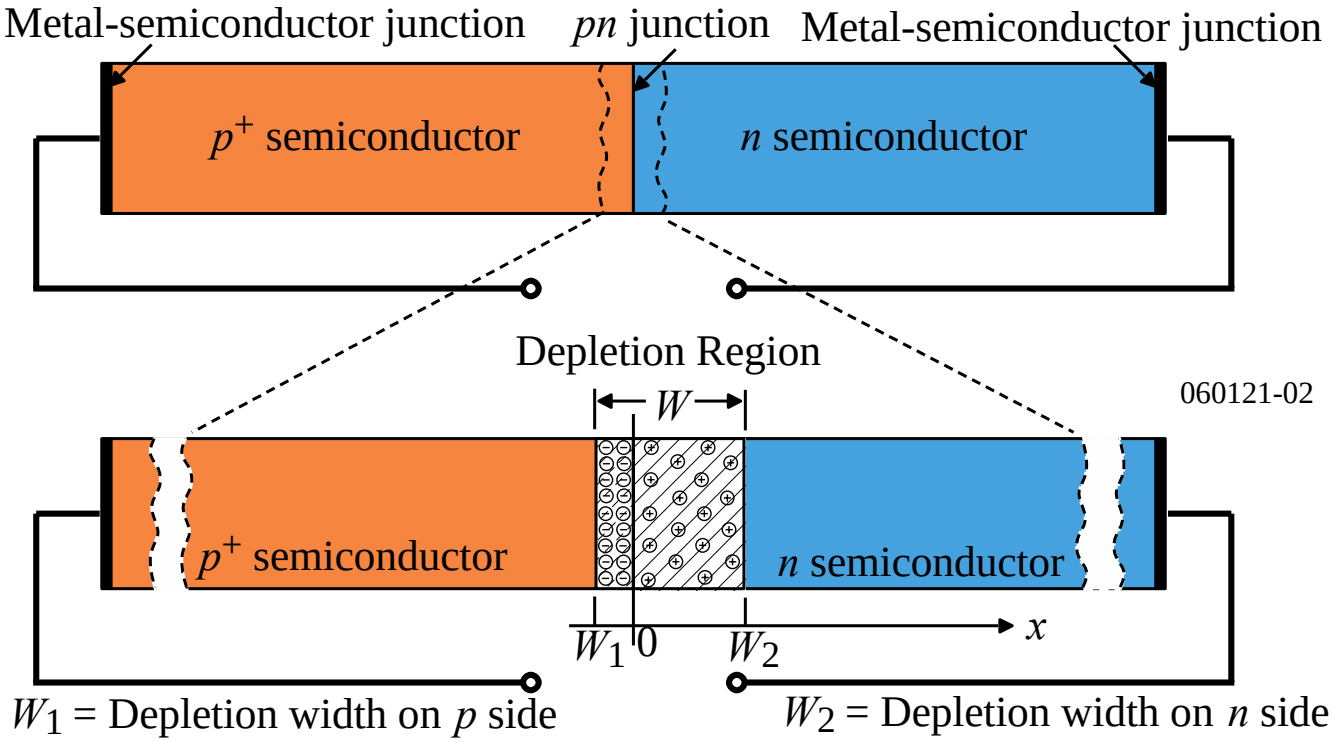
- *PN* junctions are used to electrically isolate one semiconductor region from another
- *PN* diodes
- ESD protection
- Creation of the thermal voltage for bandgap purposes
- Depletion capacitors – voltage variable capacitors (varactors)

Components of a *pn* junction:

1.) *p*-doped semiconductor – a semiconductor having atoms containing a lack of electrons (acceptors). The concentration of acceptors is N_A in atoms per cubic centimeter.

2.) *n*-doped semiconductor – a semiconductor having atoms containing an excess of electrons (donors). The concentration of these atoms is N_D in atoms per cubic centimeter.

Abrupt PN Junction



1. Doped atoms near the metallurgical junction lose their free carriers by diffusion.
2. As these fixed atoms lose their free carriers, they build up an electric field, which opposes the diffusion mechanism.
3. Equilibrium conditions are reached when:

   Current due to diffusion = Current due to electric field

Influence of Doping Level on the Depletion Regions

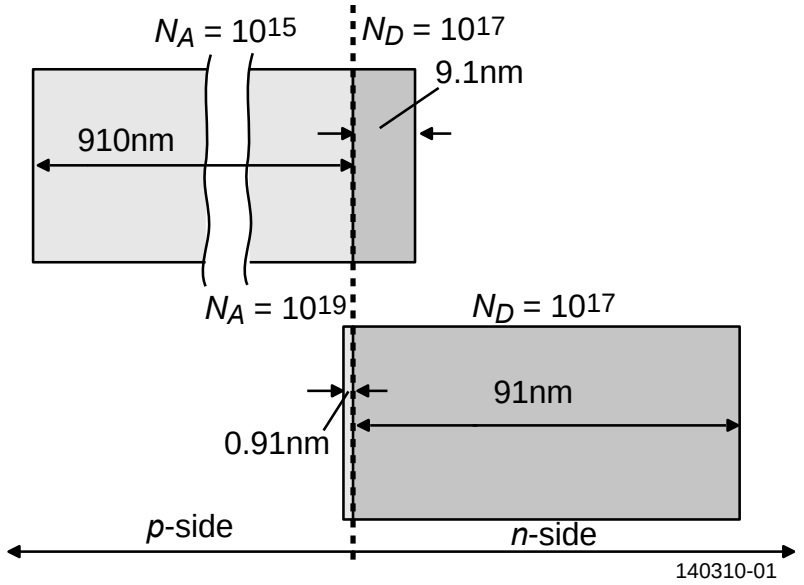
Intuitively, one can see that the depletion regions are inversely proportional to the doping level. To achieve equilibrium, equal and opposite fixed charge on both sides of the junction are required. Therefore, the larger the doping the smaller the depletion region on that side of the junction.

The equations that result are:

$$W_1 = \sqrt{\frac{2\varepsilon(\psi_o - v_D)}{qN_A\left(1+\frac{N_A}{N_D}\right)}} \propto \sqrt{\frac{1}{N_A}}$$

and

$$W_2 = \sqrt{\frac{2\varepsilon(\psi_o - v_D)}{qN_D\left(1+\frac{N_D}{N_A}\right)}} \propto \sqrt{\frac{1}{N_D}}$$



Assume that $v_D = 0$, $\psi_o = 0.637$V and $N_D = 10^{17}$ atoms/cm^3. Find the *p*-side depletion region width if $N_A = 10^{15}$ atoms/cm^3 and if $N_A = 10^{19}$ atoms/cm^3:

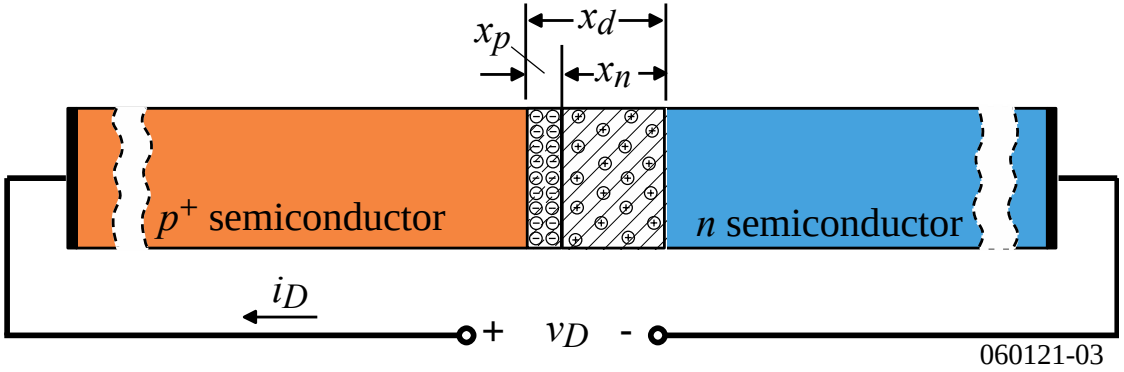
For $N_A = 10^{15}$ atoms/cm^3 the *p*-side depletion width is 0.90 µm.

For $N_A = 10^{19}$ atoms/cm^3 the *p*-side depletion width is 0.9 nm.

Graphical Characterization of the Abrupt PN Junction

Assume the pn junction is open-circuited.

Cross-section of an ideal pn junction:



Symbol for the *pn* junction:

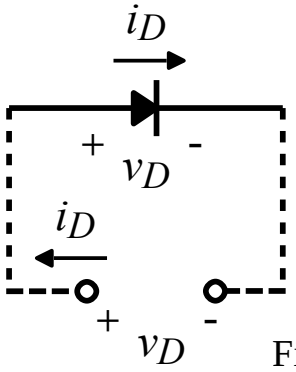


Fig. 06-03

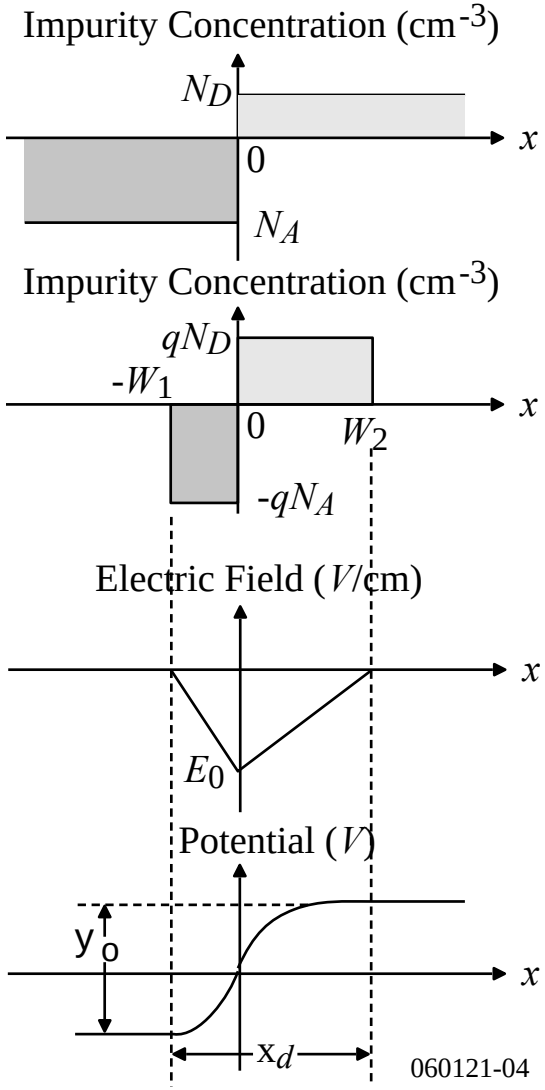
Built-in potential, ψ_o:

$$\psi_o = V_t \ln\left(\frac{N_A N_D}{n_i^2}\right),$$

where

$$V_t = \frac{kT}{q}$$

n_i is the intrinsic concentration of silicon.



Reverse-Biased *PN* Junctions

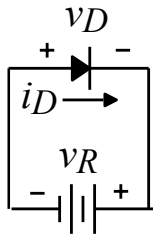
Depletion region:

$$x_d = x_p + x_n = W_1 + W_2$$

$$x_p = W_1 \propto \sqrt{v_R}$$

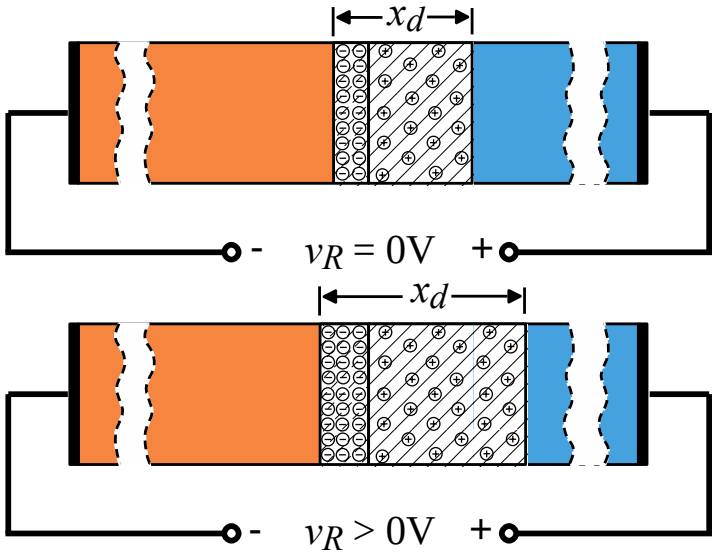
and

$$x_n = W_2 \propto \sqrt{v_R}$$



060121-05

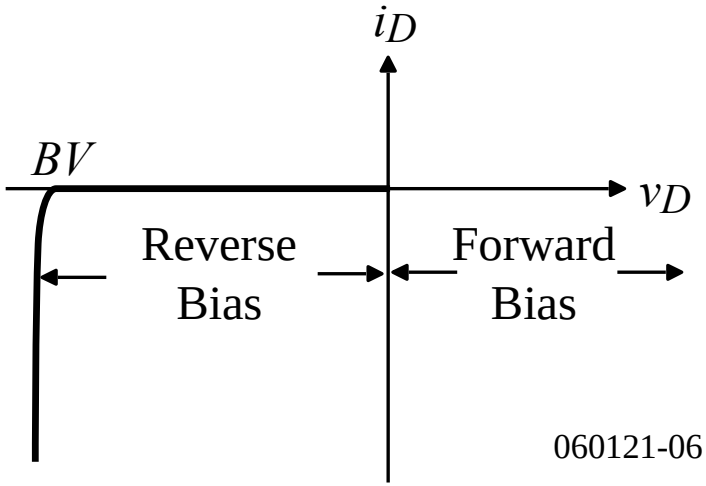
Influence of v_R on depletion region width



Breakdown voltage (*BV*):

In the reverse direction the current can be written as,

$$i_D = \frac{-I_R}{1 - \left(\frac{v_R}{BV}\right)^n}$$



060121-06

Breakdown Voltage as a Function of Doping

It can be shown that†:

$$BV \approx \frac{\varepsilon_{si}(N_A + N_D)}{2qN_AN_D} E_{max}^2$$

where $E_{max} = 3\text{x}10^5$ V/cm for silicon.

An example:

Assume that $N_D = 10^{17}$ atoms/cm^3.

Find BV if $N_A = 10^{15}$ atoms/cm^3 and if $N_A = 10^{19}$ atoms/cm^3:

$N_A = 10^{15}$ atoms/cm^3:

If $N_A << N_D$, then $BV \approx \frac{\varepsilon_{si}}{2qN_A} E_{max}^2 = \frac{1.04\text{x}10^{-12} \cdot 9\text{x}10^{10}}{2 \cdot 1.6\text{x}10^{-19} \cdot 10^{15}} = 291\text{V}$

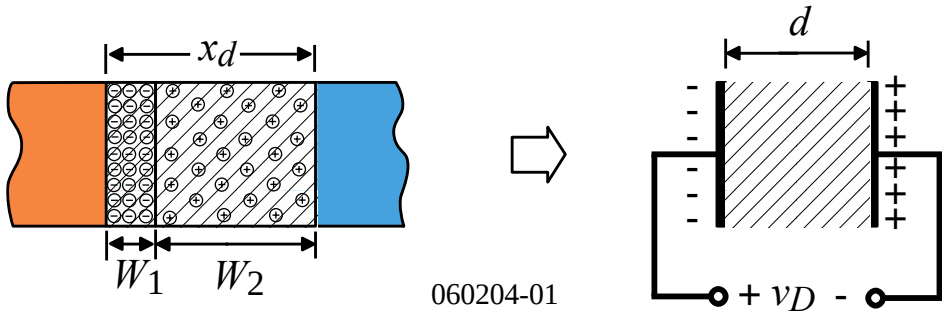
$N_A = 10^{19}$ atoms/cm^3:

If $N_A >> N_D$, then $BV \approx \frac{\varepsilon_{si}}{2qN_D} E_{max}^2 = \frac{1.04\text{x}10^{-12} \cdot 9\text{x}10^{10}}{2 \cdot 1.6\text{x}10^{-19} \cdot 10^{17}} = 2.91\text{V}$

† P. Allen and D. Holberg, *CMOS Analog Circuit Design, 3rd ed.*, Oxford University Press, 2012

Depletion Capacitance

Physical viewpoint of the depletion capacitance:



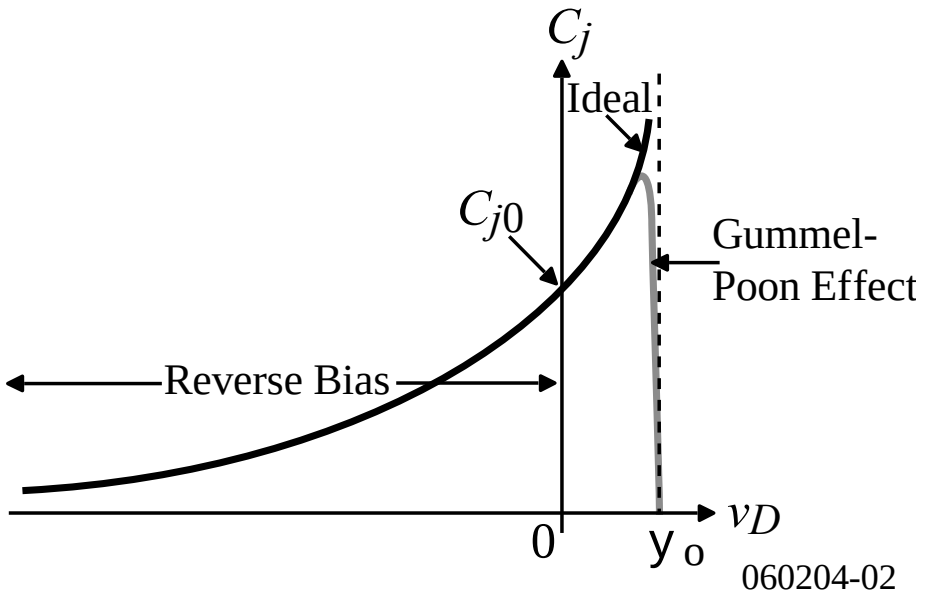
060204-01

$$C_j = \frac{\varepsilon_{si} A}{d} = \frac{\varepsilon_{si} A}{W_1 + W_2}$$

$$= \frac{\varepsilon_{si} A}{\sqrt{\frac{2\varepsilon_{si}(\psi_o - v_D)}{q(N_D + N_A)}}\left[\sqrt{\frac{N_D}{N_A}} + \sqrt{\frac{N_A}{N_D}}\right]}$$

$$= A\sqrt{\frac{\varepsilon_{si} q N_A N_D}{2(N_A + N_D)}}\,\frac{1}{\sqrt{\psi_o - v_D}}$$

$$= \frac{C_{j0}}{\sqrt{1 - \frac{v_D}{\psi_o}}}$$



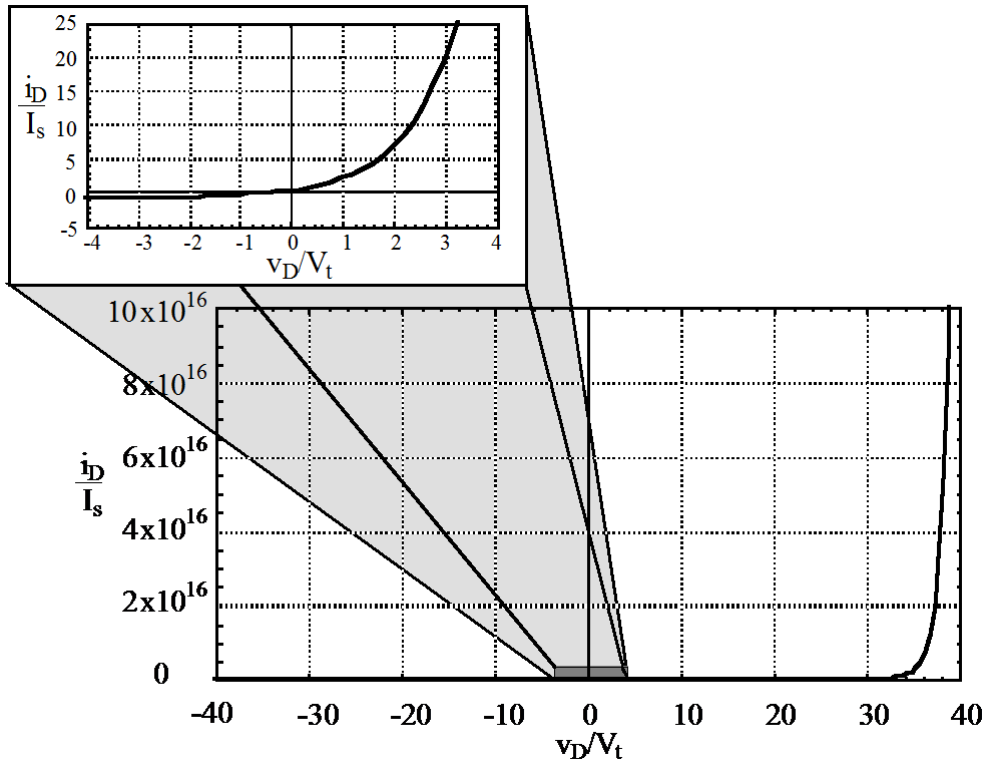
060204-02

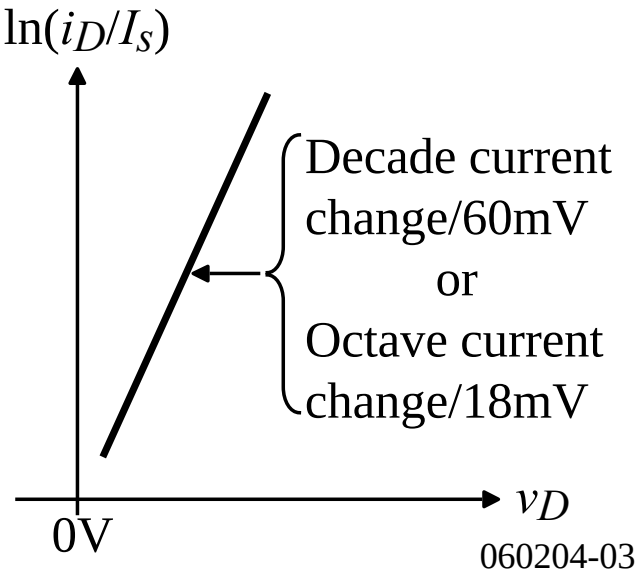
Forward-Biased *PN* Junctions

When the *pn* junction is forward-biased, the potential barrier is reduced and significant current begins to flow across the junction. This current is given by:

$$i_D = I_S\left[\exp\left(\frac{v_D}{V_t}\right) - 1\right] \qquad \text{where } I_S = qA\left[\frac{D_p p_{no}}{L_p} + \frac{D_n n_{po}}{L_n}\right] \approx \frac{qAD}{L}\frac{n_i^2}{N} = KT^3\exp\left(\frac{-V_{GO}}{V_t}\right)$$

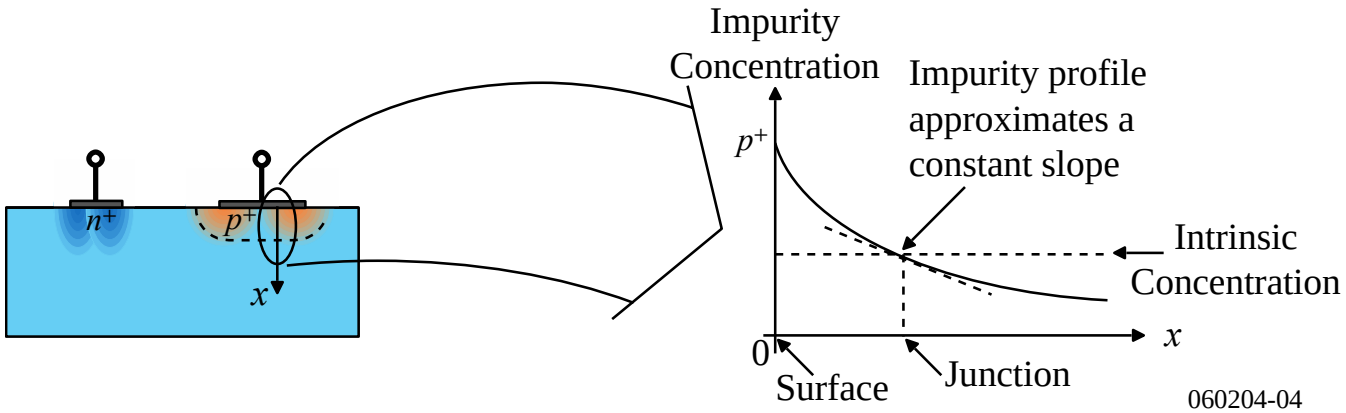
Graphically, the i_D versus v_D characteristics are given as:





Graded *PN* Junctions

In practice, the *pn* junction is graded rather than abrupt.



The previous expressions become:

Depletion region widths-

$$W_1 = \left(\frac{2\varepsilon_{si}(\psi_o - v_D)N_D}{qN_A(N_A+N_D)}\right)^m$$

$$W_2 = \left(\frac{2\varepsilon_{si}(\psi_o - v_D)N_A}{qN_D(N_A+N_D)}\right)^m$$

$$W \propto \left(\frac{1}{N}\right)^m$$

Depletion capacitance-

$$C_j = A\left(\frac{\varepsilon_{si} q N_A N_D}{2(N_A+N_D)}\right)^m \frac{1}{(\psi_o - v_D)^m}$$

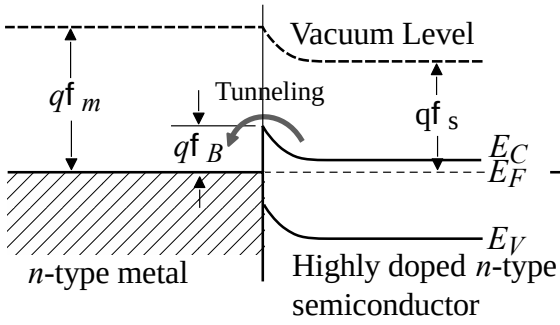
$$= \frac{C_{j0}}{\left(1 - \frac{v_D}{\psi_o}\right)^m}$$

where $0.33 \le m \le 0.5$.

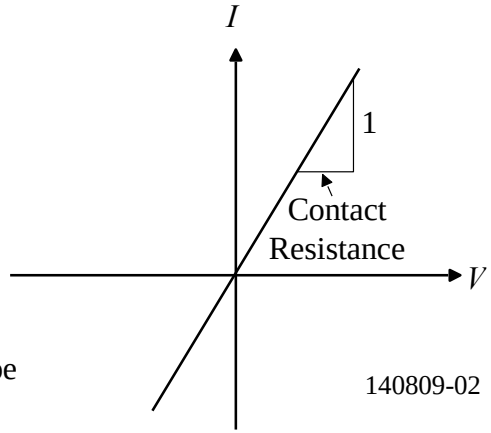
Metal-Semiconductor Junctions

Ohmic Junctions: A metal-semiconductor junction formed by a highly doped semiconductor and metal.

Energy band diagram

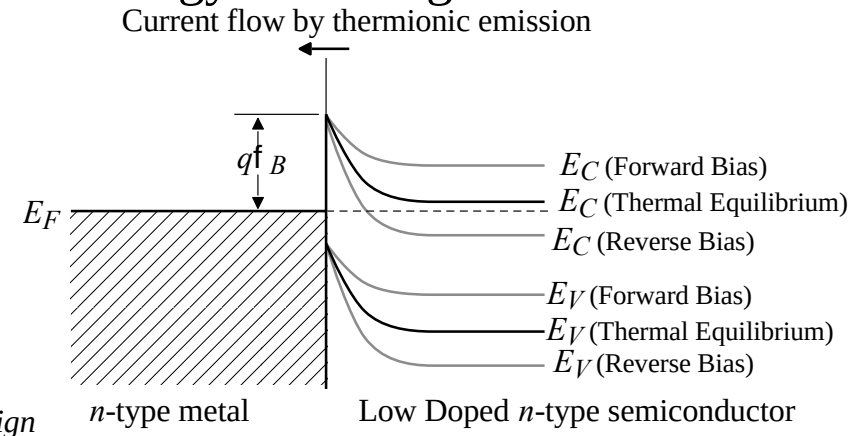


IV Characteristics

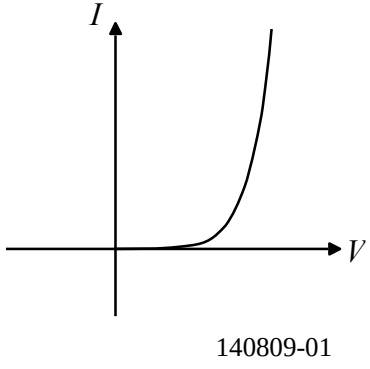


Schottky Junctions: A metal-semiconductor junction formed by a lightly doped semiconductor and metal.

Energy band diagram

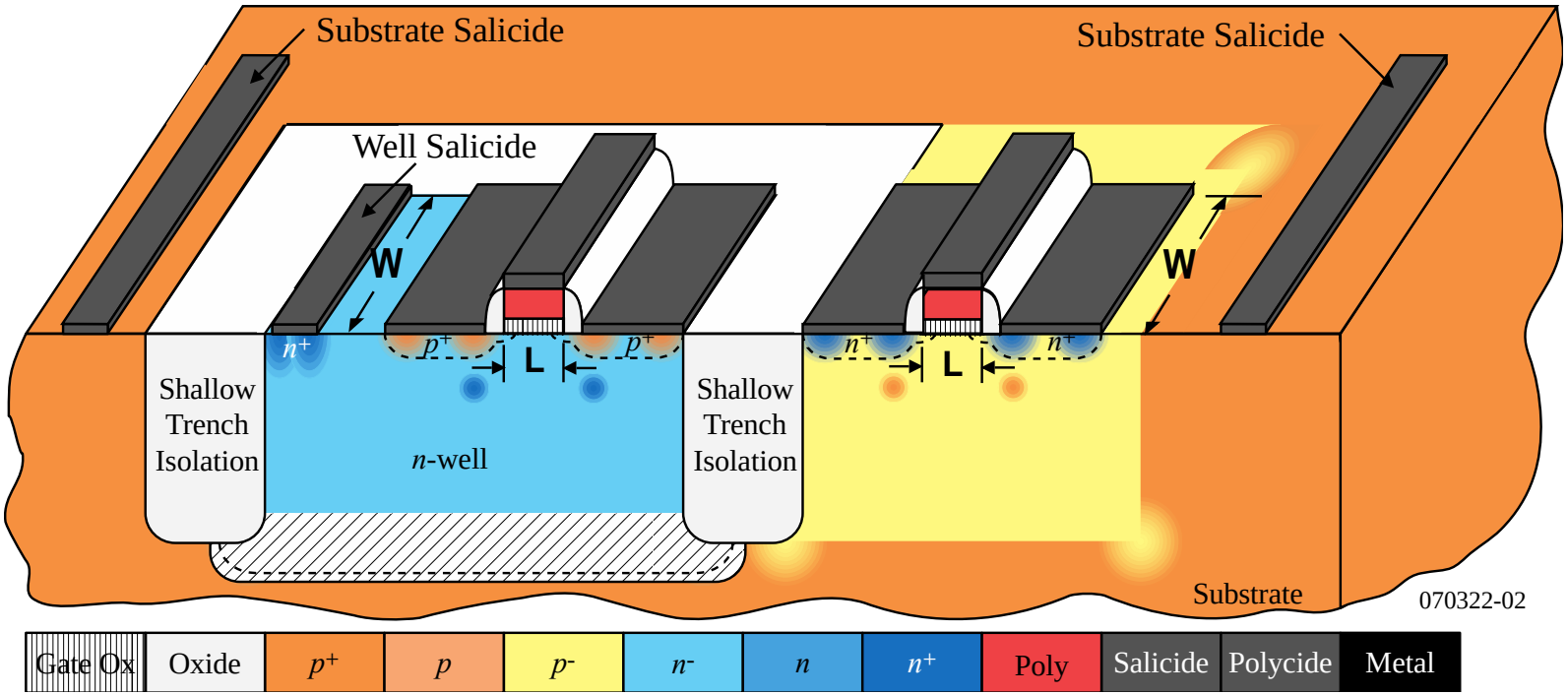


IV Characteristics



MOS TRANSISTORS

Physical Structure of MOS Transistors in an *n*-well Technology



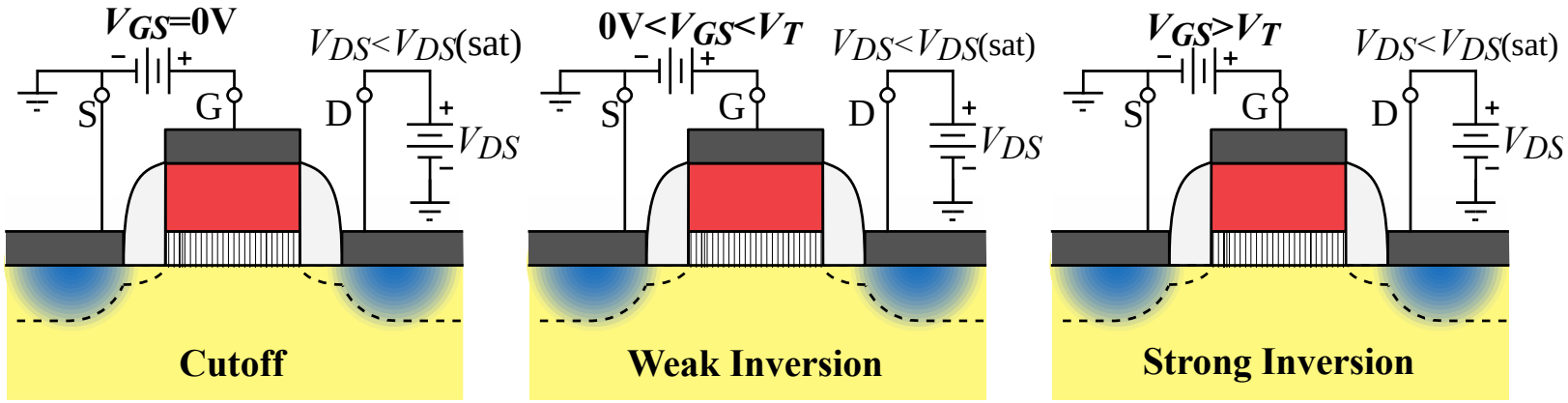
Width (*W*) of the MOSFET = Width of the source/drain diffusion

Length (*L*) of the MOSFET = Width of the polysilicon gate between the S/D diffusions

Note that the MOSFET is isolated from the well/substrate by reverse biasing the resulting *pn* junction

Enhancement MOSFETs

The channel of an enhancement MOSFET is formed when the proper potential is applied to the gate of the MOSFET. This potential inverts the material immediately below the gate to the same type of impurity as the source and drain forming the channel.



060205-06

V_T = Gate-bulk work function (ϕ_{MS}) + voltage to change the surface potential ($-2\phi_F$)
+ voltage to offset the channel-bulk depletion charge ($-Q_b/C_{ox}$)
+ voltage to compensate the undesired interface charge ($-Q_{ss}/C_{ox}$)

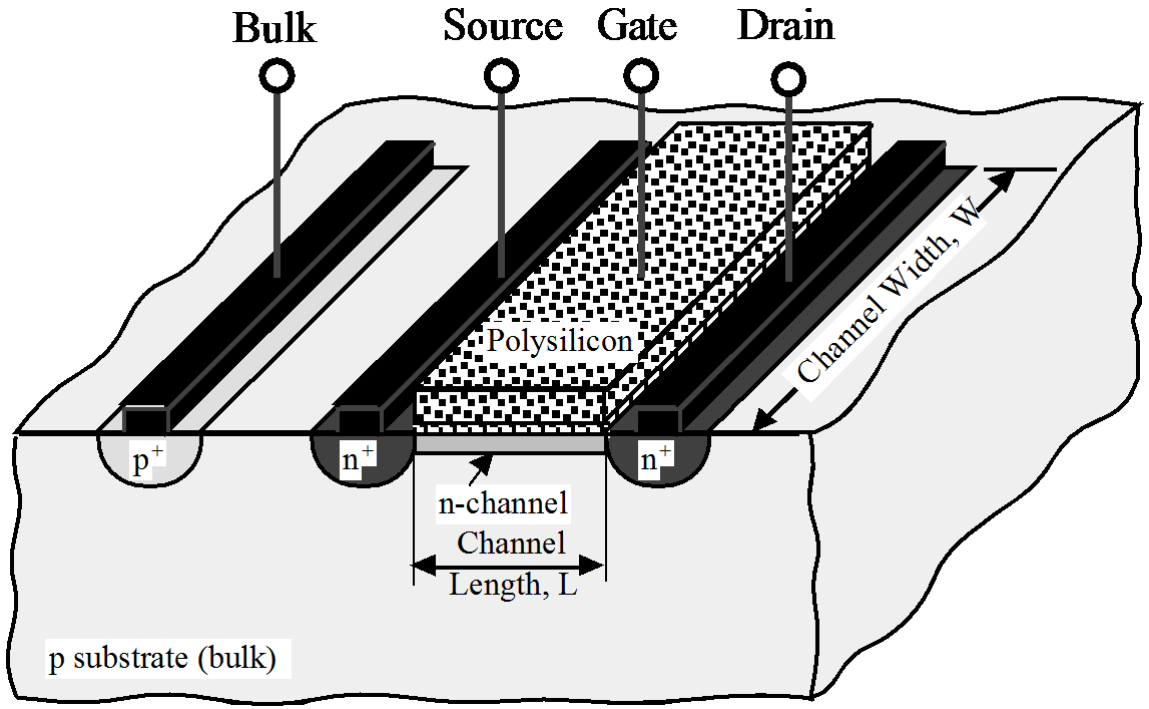
$$V_T = \phi_{MS} - 2\phi_F - \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} - \frac{Q_b - Q_{b0}}{C_{ox}} = V_{T0} + \gamma\left(\sqrt{|-2\phi_F + v_{SB}|} - \sqrt{|-2\phi_F|}\right)$$

where

$$V_{T0} = \phi_{MS} - 2\phi_F - \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}, \quad \gamma = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C_{ox}} \text{ and } \quad Q_b \approx \sqrt{2qN_A\varepsilon_{si}(|-2\phi_F + v_{SB}|)}$$

Depletion Mode MOSFET

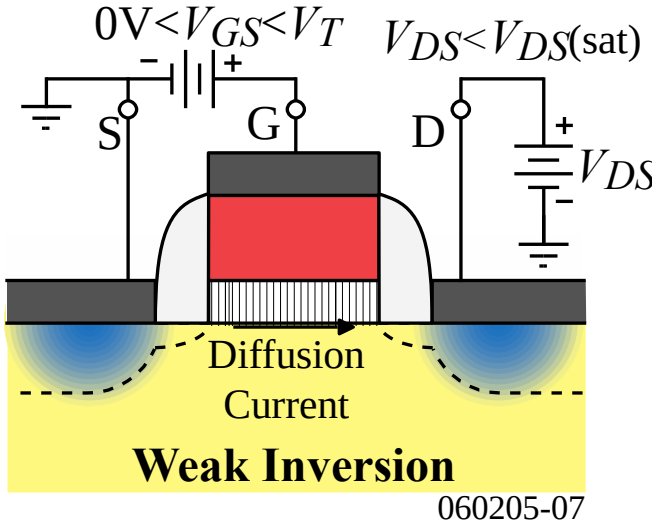
The channel is diffused into the substrate so that a channel exists between the source and drain with no external gate potential.



The threshold voltage for a depletion mode NMOS transistor will be negative (a negative gate potential is necessary to attract enough holes underneath the gate to cause this region to invert to p-type material).

Weak Inversion Operation

Weak inversion operation occurs when the applied gate voltage is below V_T and occurs when the surface of the substrate beneath the gate is weakly inverted.



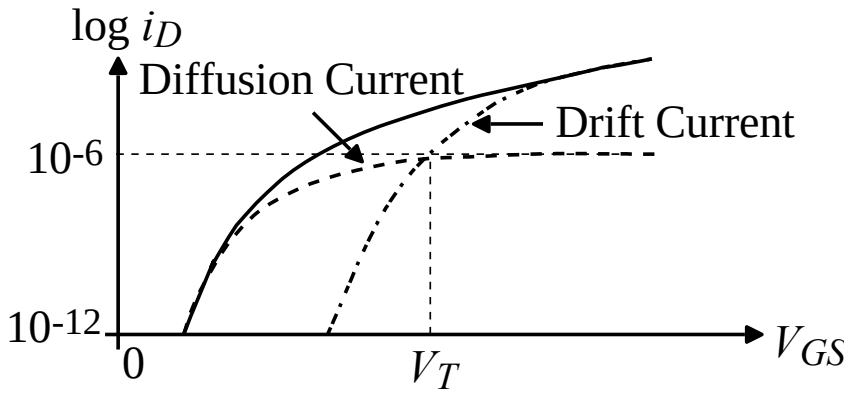
Regions of operation according to the surface potential, ϕ_S.

$\phi_S < \phi_F$: Substrate not inverted

$\phi_F < \phi_S < 2\phi_F$: Channel is weakly inverted (diffusion current)

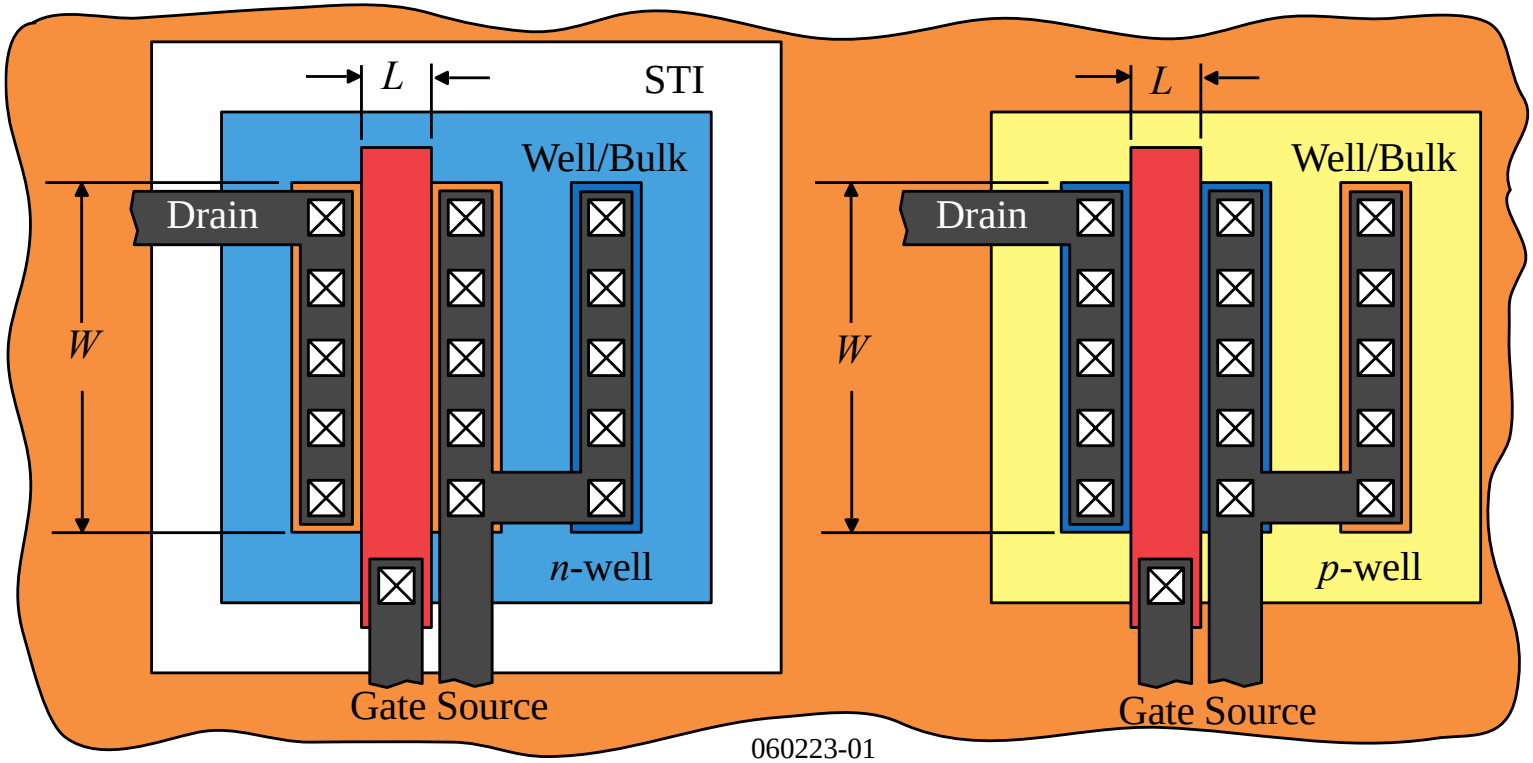
$2\phi_F < \phi_S$: Strong inversion (drift current)

Drift current versus diffusion current in a MOSFET:



LAYOUT OF MOS TRANSISTORS

Layout of a Single MOS transistor:

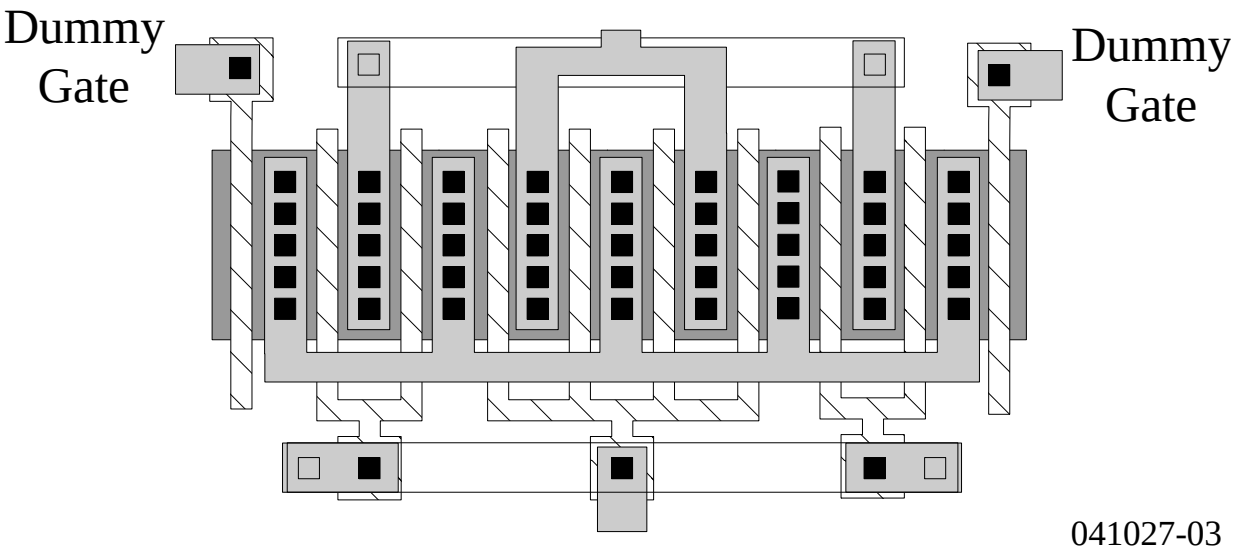


Comments:

- Make sure to contact the source and drain with multiple contacts to evenly distribute the current flow under the gate.
- Minimize the area of the source and drain to reduce bulk-source/drain capacitance.

Diffusion and Etch Effects

• Poly etch rate variation – use dummy elements to prevent etch rate differences.



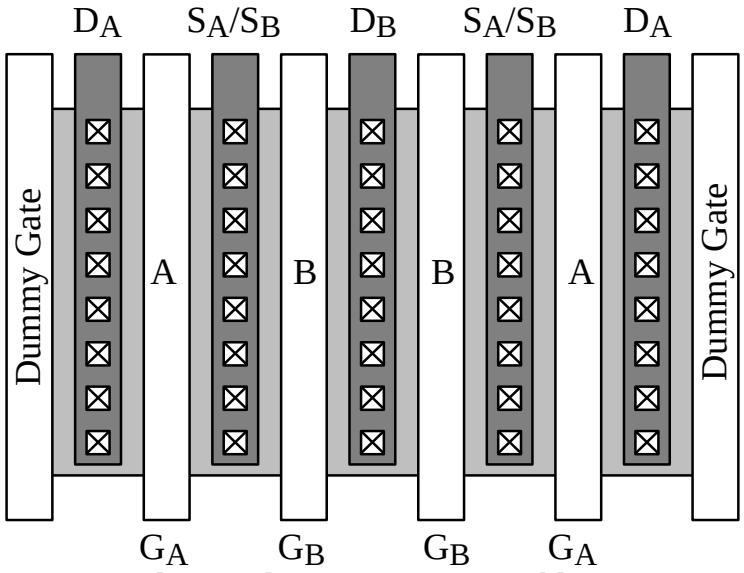
• Do not put contacts on top of the gate for matched transistors.

• Be careful of diffusion interactions for diffusions near the channel of the MOSFET

Thermal and Stress Effects

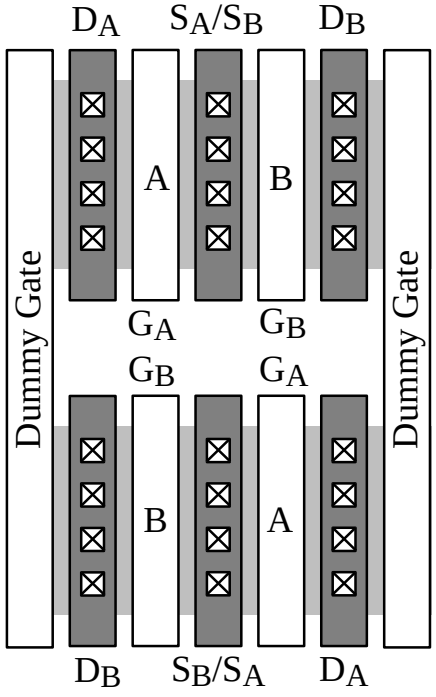
- Oxide gradients – use common centroid geometry layout
- Stress gradients – use proper location and common centroid geometry layout
- Thermal gradients – keep transistors well away from power devices and use common centroid geometry layout with interdigitated transistors

Examples of Common Centroid Interdigitated transistor layout:



Interdigitated, common centroid layout

041027-04

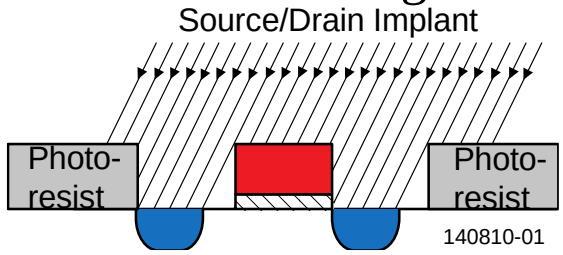


Cross-Coupled Transistors

MOS Transistor Layout

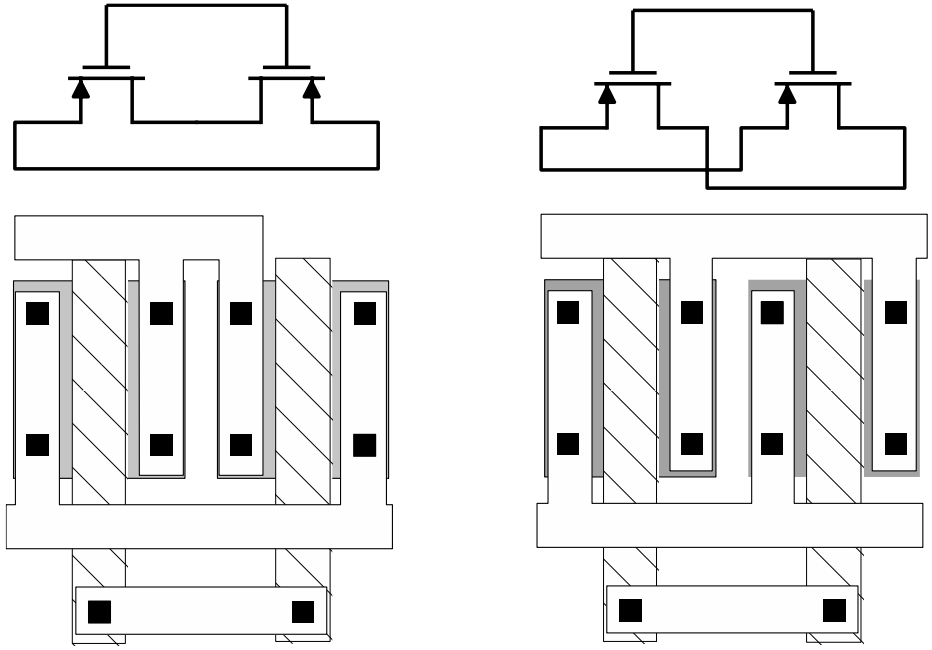
Photolithographic invariance (PLI) are transistors that exhibit identical orientation. PLI comes from optical interactions between the UV light and the masks.

Simple illustration of PLI:



Examples of the layout of matched MOS transistors:

1.) Examples of mirror symmetry and photolithographic invariance.

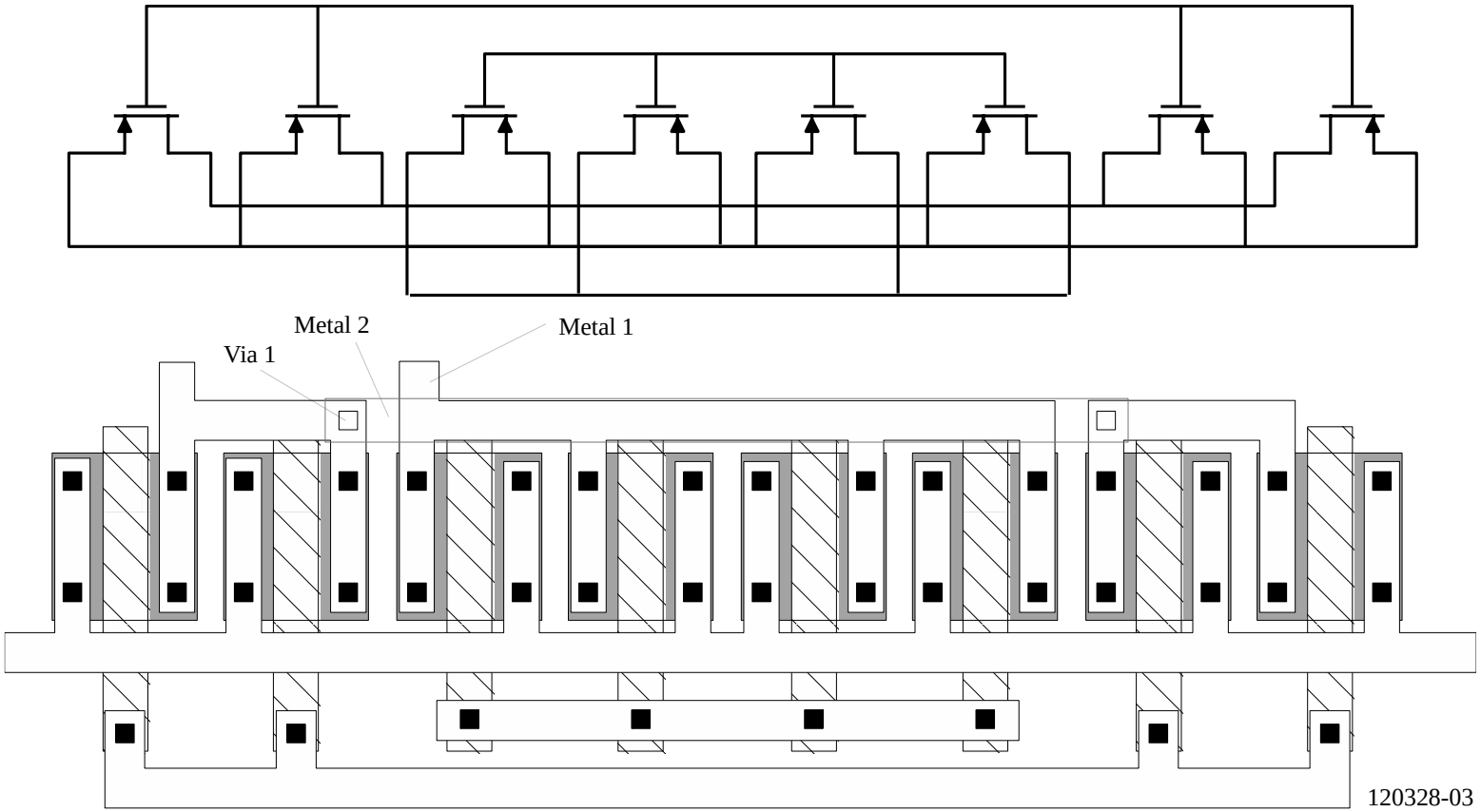


Mirror Symmetry Photolithographic Invariance

120328-02

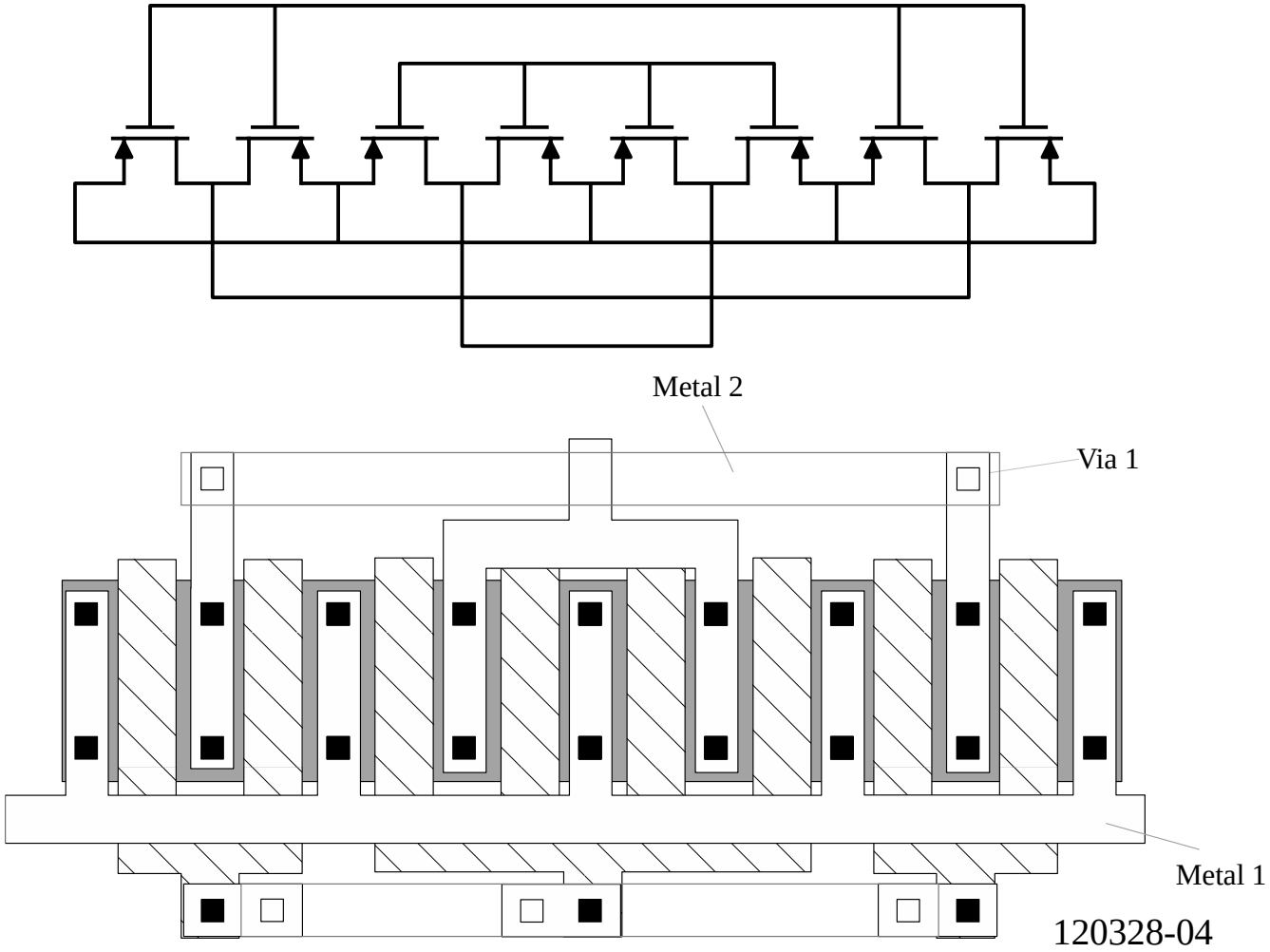
MOS Transistor Layout - Continued

2.) Two transistors sharing a common source and laid out to achieve both photolithographic invariance and common centroid.



MOS Transistor Layout - Continued

3.) Compact layout of the previous example.

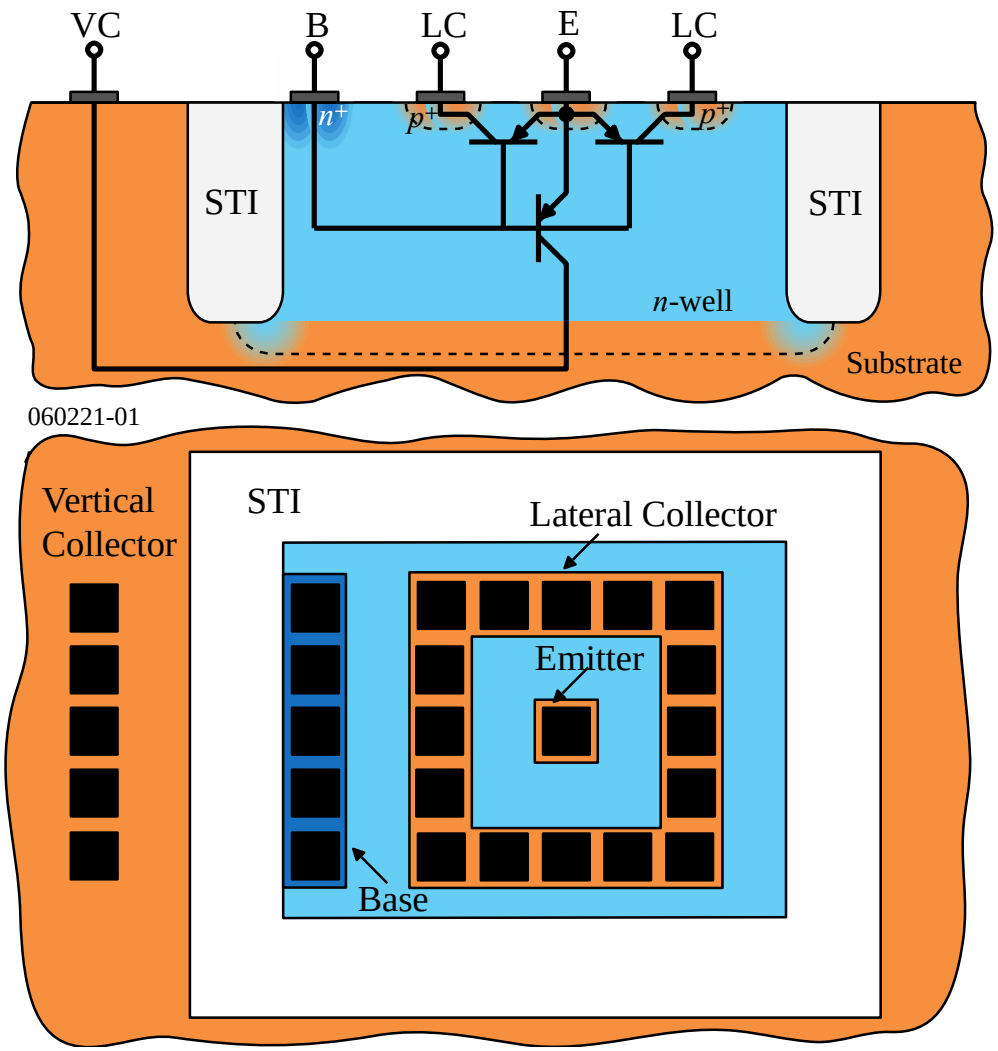


PARASITIC BIPOLAR TRANSISTORS IN CMOS TECHNOLOGY

A Lateral Bipolar Transistor

n-well CMOS technology:

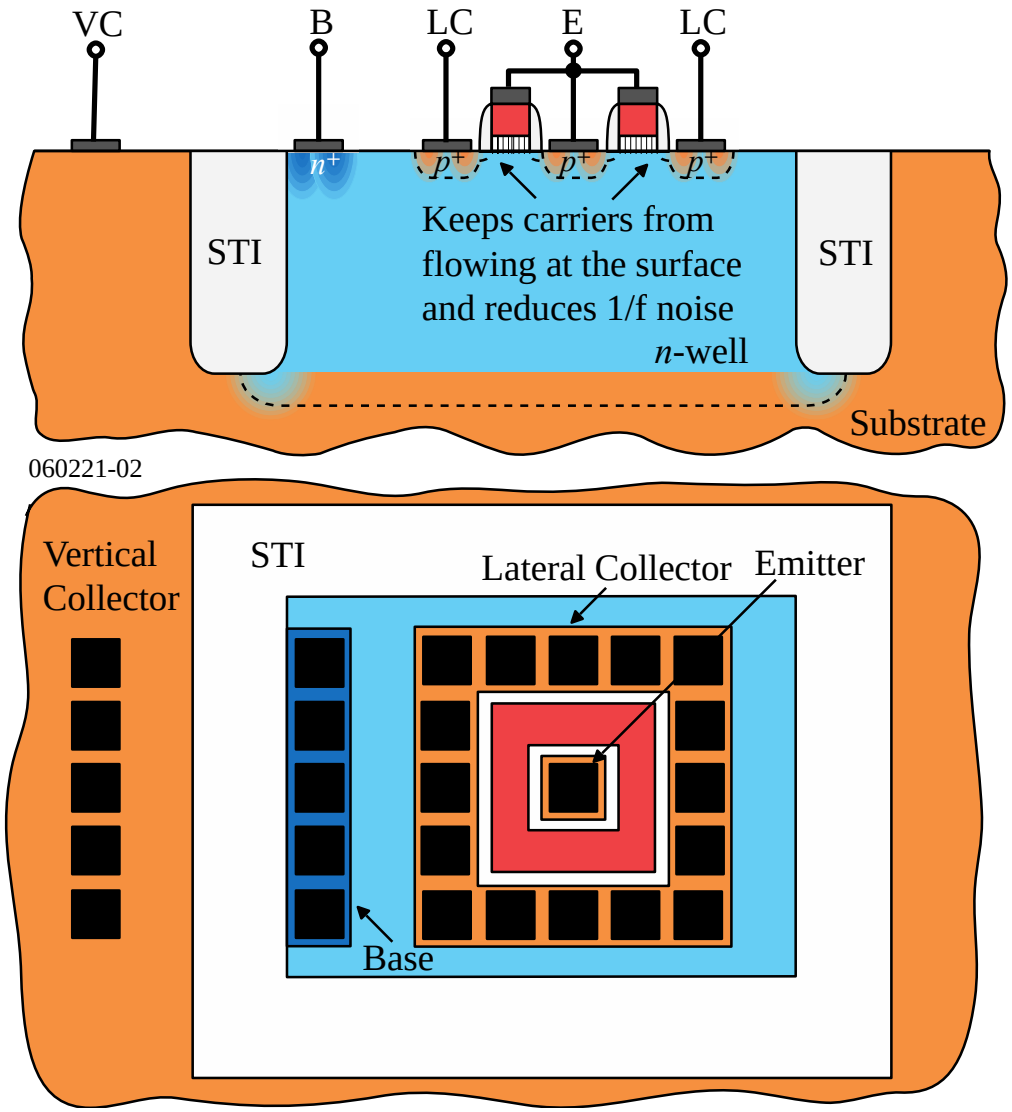
- It is desirable to have the lateral collector current much larger than the vertical collector current.
- Lateral BJT generally has good matching.
- The lateral BJT can be used as a photodetector with reasonably good efficiency.
- Triple well technology allows the current of the vertical collector to avoid the substrate.



A Field-Aided Lateral BJT

Use minimum channel length to enhance beta:

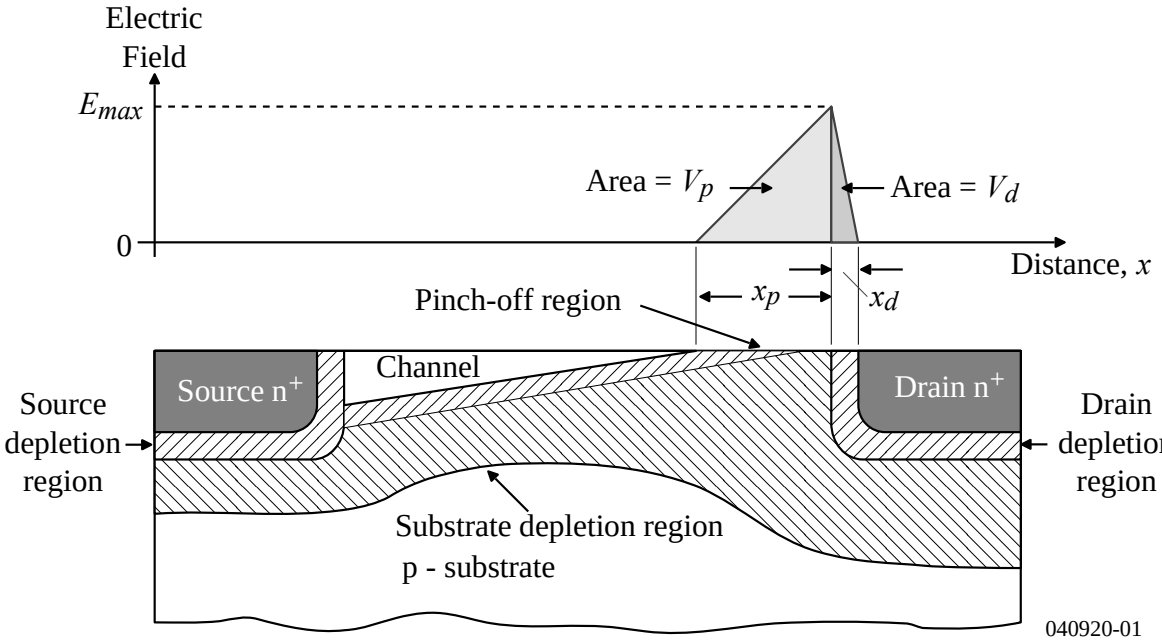
$ß_F \approx 50$ to 100 depending on the process



HIGH VOLTAGE CMOS TRANSISTORS

Extended Voltage MOSFETS

The electric field from the source to drain in the channel is shown below.



The voltage drop from drain to source is,

$$V_{DS} = V_p + V_d = 0.5(E_{max}x_p + E_{max}x_d) = 0.5E_{max}(x_p + x_d)$$

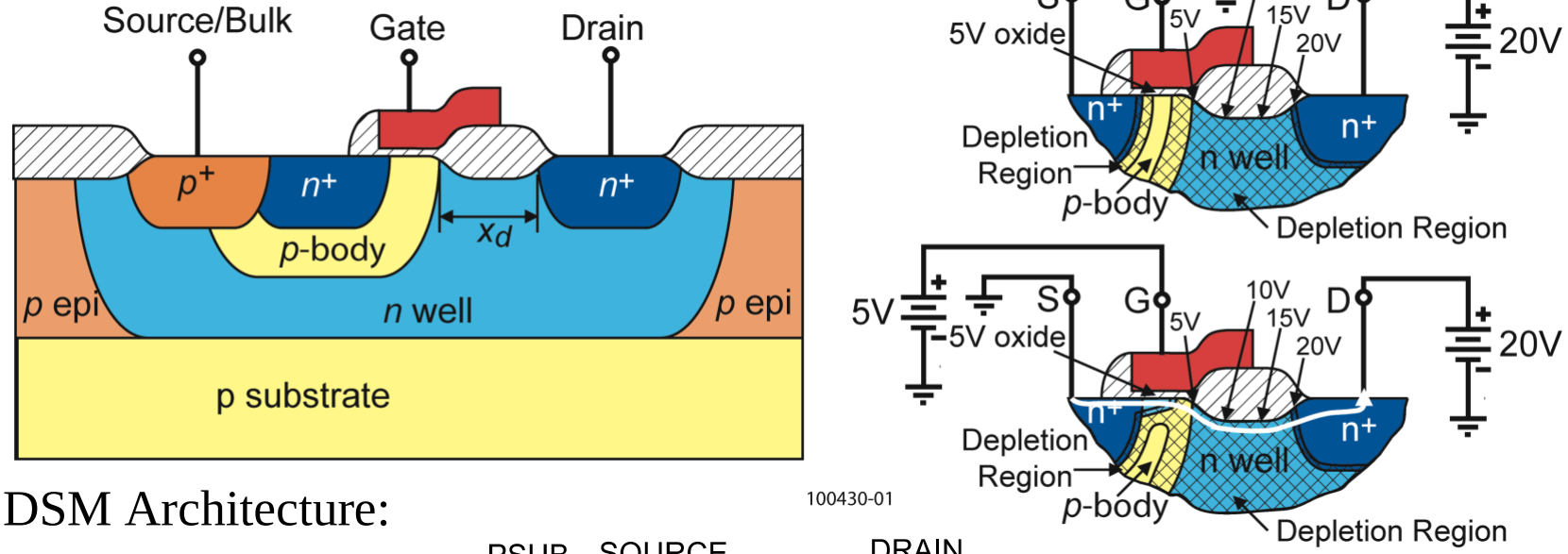
E_{max} and x_p are limited by hot carrier generation and channel length modulation requirements whereas these limitations do not exist for x_d.

Therefore, to get extended voltage transistors, make x_d larger.

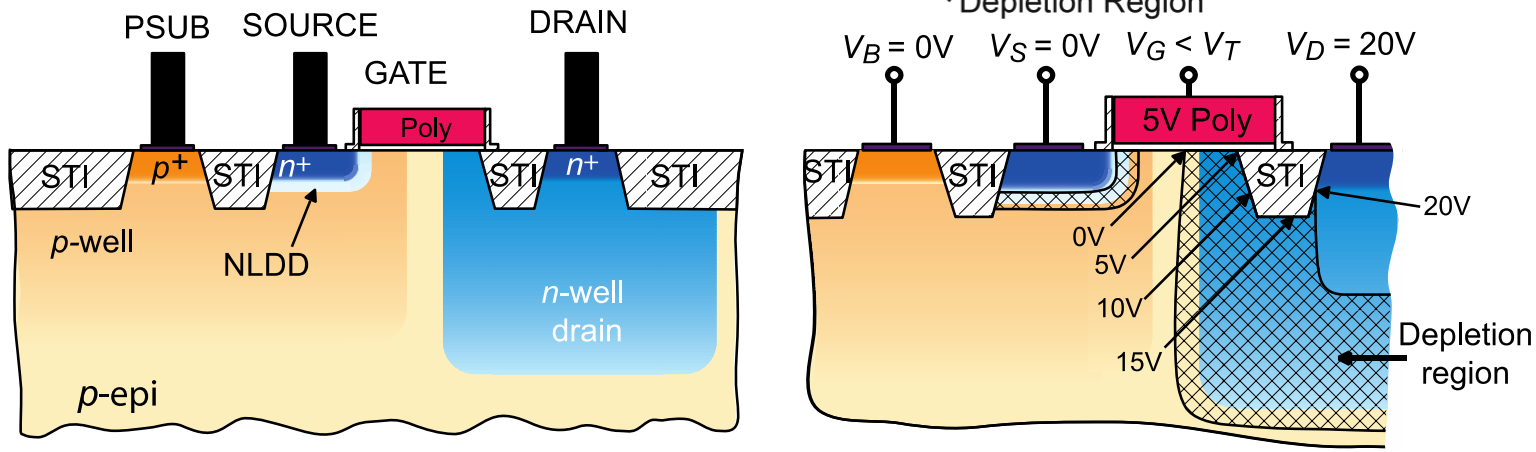
High Voltage Architectures

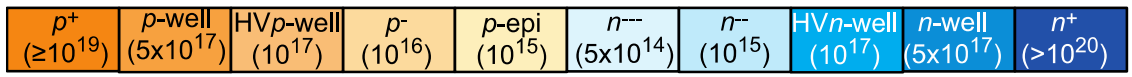
The objective is to create a lightly doped, extended drain region where the high voltage of the drain can drop down to a level that will not cause the gate oxide to breakdown.

LOCOS Architecture:



DSM Architecture:

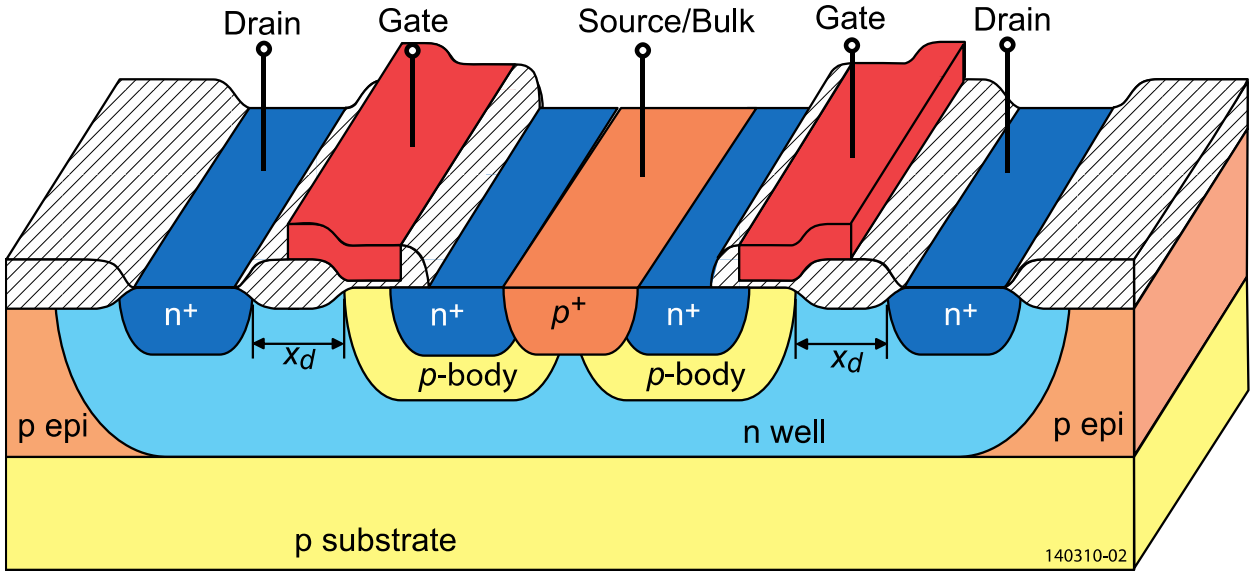




Lateral DMOS (LDMOS) Using LOCOS CMOS Technology

The LDMOS structure is designed to provide sufficient lateral dimension and to prevent oxide breakdown by the higher drain voltages.

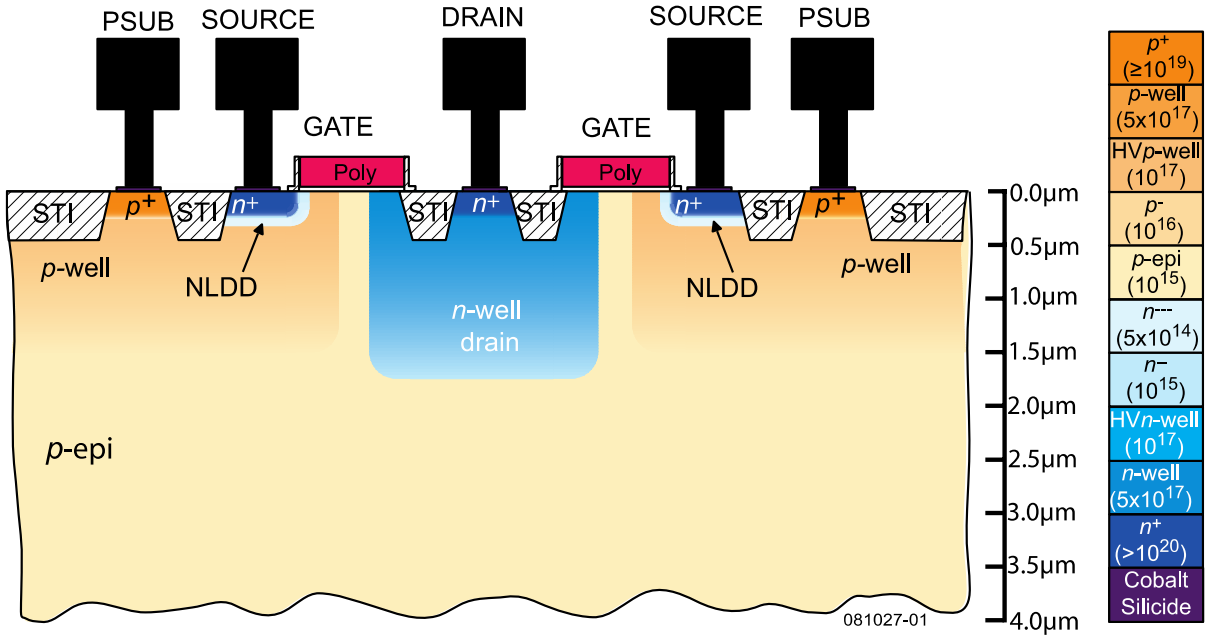
One possible implementation using LOCOS technology:



- Structure is symmetrical about the source/bulk contact
- Channel is formed in the p region under the gates
- The lightly doped *n* region between the drain side of the channel and the n^+ drain contact (x_d) increases the depletion region width on the drain side of the channel/drain *pn* junction resulting in larger values of v_{DS}.
- Drain voltage can be 20-100V depending on the spacing and doping.

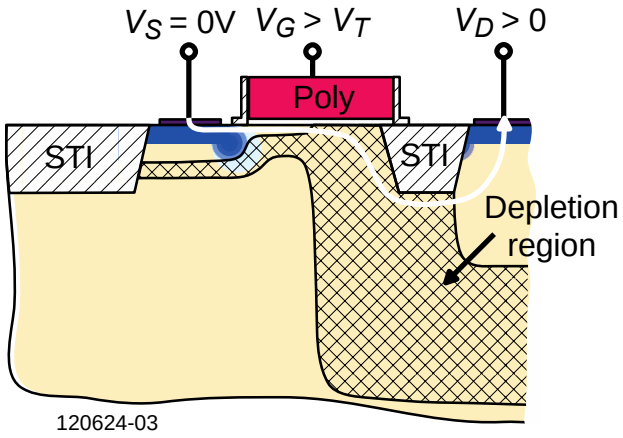
Lateral DMOS (LDMOS) Using DSM CMOS Technology

Cross-section of an NLDMOS using DSM technology:



Differences between an NLDMOS and NMOS:

- Asymmetry
- Non-uniform channel
- Current flow (not all at the surface)
- No self-alignment (larger drain-gate overlap capacitance)
- Note the extended drift region on the drain side of the channel



SUMMARY

- *pn* junction usage in CMOS include:
  - Electrical isolation, *pn* diodes, ESD protection, depletion capacitors
- Depletion region widths are inversely proportional to the doping
- Depletion region widths are proportional to the reverse bias voltage
- Ohmic metal-semiconductor junctions require a highly doped semiconductor
- MOSFETs can be:
  - Enhancement – the applied gate voltage forms the channel
  - Depletion – the channel is physically constructed in fabrication
- The threshold voltage of MOSFETs consists of the following components:
  - Gate bulk work function (ϕ_{MS})
  - Voltage to change the surface potential ($-2\phi_F$)
  - Voltage to offset the channel-bulk depletion charge ($-Q_b/C_{ox}$)
  - Voltage to compensate the undesired interface charge ($-Q_{ss}/C_{ox}$)
- Weak inversion is MOSFET operation with the gate-source voltage less than the threshold voltage
- Layout of the MOSFET is important to its performance and matching capabilities
- Extended drain regions lead to higher voltage capability MOSFETs