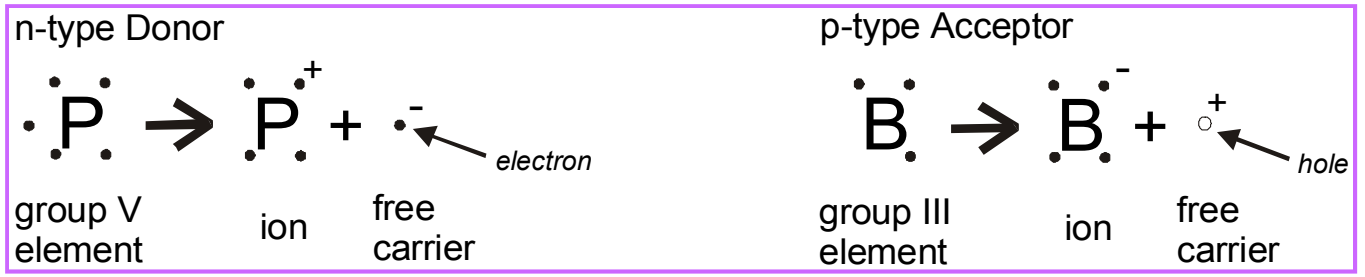
Conduction in Semiconductors -Review

Chap. 3

- Intrinsic (undoped) Semiconductors
  - **intrinsic carrier concentration** ≡ $n_i = 1.45 \times 10^{10}$ cm^{-3}, at room temp.
  - $n = p = n_i$, in intrinsic (undoped) material
    - n ≡ number of electrons, p ≡ number of holes
  - mass-action law, $np = n_i^2 = (1.45E10)^2$, applies to undoped and doped material
- Extrinsic (doped) Semiconductors
  - **dopants** added to modify material/electrical properties



- n-type (n+), add elements with an extra electron
  - N_d ≡ conc. of donor atoms [cm^{-3}]
  - $n_n = N_d$, n_n ≡ conc. of electrons in n-type material
  - $p_n = n_i^2/N_d$, using mass-action law,
  - p_n ≡ conc. of holes in n-type material
  - always a lot more n than p in n-type material
- p-type = p+, add elements with an extra hole
  - N_a ≡ concentration of acceptor atoms [cm^{-3}]
  - $p_p = N_a$, p_p ≡ conc. of holes in p-type material
  - $n_p = ni^2/N_a$, using mass-action law,
  - n_p ≡ conc. of electrons in p-type material
  - always a lot more p than n in p-type material

Conduction in Semiconductors

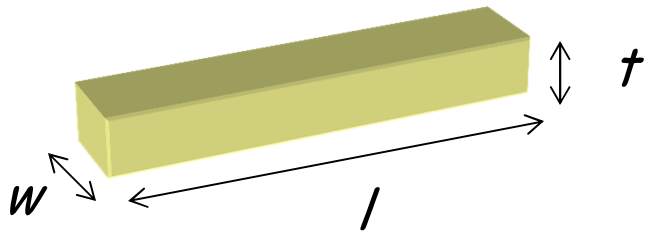
- doping provides free charge carriers, alters conductivity
- conductivity, σ, in semic. w/ carrier densities *n* and *p*
  - $\sigma = q(\mu_n n + \mu_p p)$, q ≡ electron charge, $q = 1.6\times10^{-19}$ [Coulombs]
    - μ ≡ mobility [cm^2/V-sec], $\mu_n \cong 1360$, $\mu_p \cong 480$ (typical values)
- in n-type region, $n_n >> p_n$
  - $\sigma \approx q\mu_n n_n$
- in p-type region, $p_p >> n_p$
  - $\sigma \approx q\mu_p n_p$
- resistivity, $\rho = 1/\sigma$
- resistance of an n+ or p+ region
  - $R = \dfrac{\rho\, l}{A}$, $A = wt$

mobility = average velocity per unit electric field

$\mu_n > \mu_p$

electrons more mobile than holes

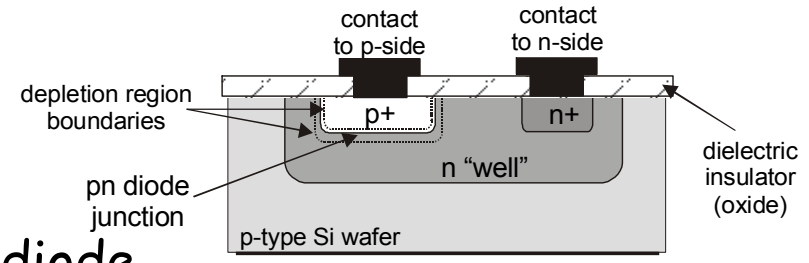
⇒conductivity of n+ > p+



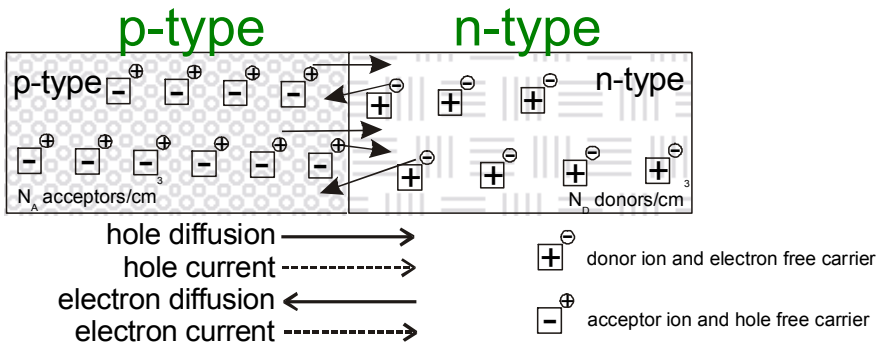
- drift current (flow of charge carriers in presence of an electric field, E_x)
  - n/p drift current density: $J_{xn} = \sigma_n E_x = q\mu_n n_n E_x$, $J_{xp} = \sigma_p E_x = q\mu_p p_p E_x$
  - total drift current density in x direction $J_x = q(\mu_n n + \mu_p p) E_x = \sigma E_x$

pn Junctions: Intro

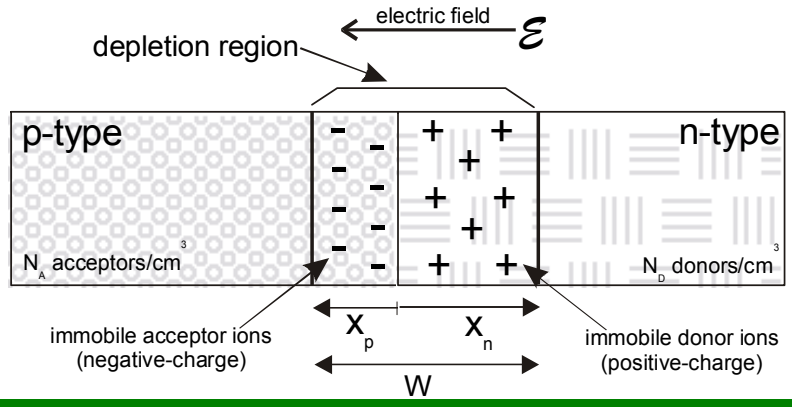
- What is a pn Junction?
  - interface of p-type and n-type semiconductor
  - junction of two materials forms a diode



- In the Beginning…
  - ionization of dopants at material interface

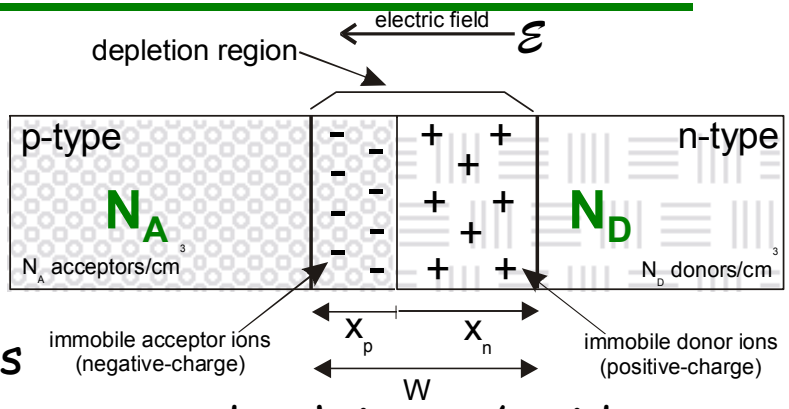


- Diffusion -movement of charge to regions of lower concentration
  - free carriers diffuse out
  - leave behind immobile ions
  - region become depleted of free carriers
  - ions establish an electric field
    - acts against diffusion



pn Junctions: Equilibrium Conditions

- Depletion Region
  - area at pn interface
    void of free charges
  - charge neutrality
    - must have equal charge on both sides
    - $q\,A\,x_p\,N_A = q\,A\,x_n\,N_D$, A=junction area; x_p, x_n depth into p/n side
    - $\Rightarrow x_p N_A = x_n N_D$
    - depletion region will extend further into the more lightly doped side of the junction



- Built-in Potential
  - diffusion of carriers leaves behind immobile charged ions
  - ions create an electric field which generates a built-in potential

$$\Psi_0 = V_T \ln\left(\frac{N_A N_D}{n_i^{\,2}}\right)$$

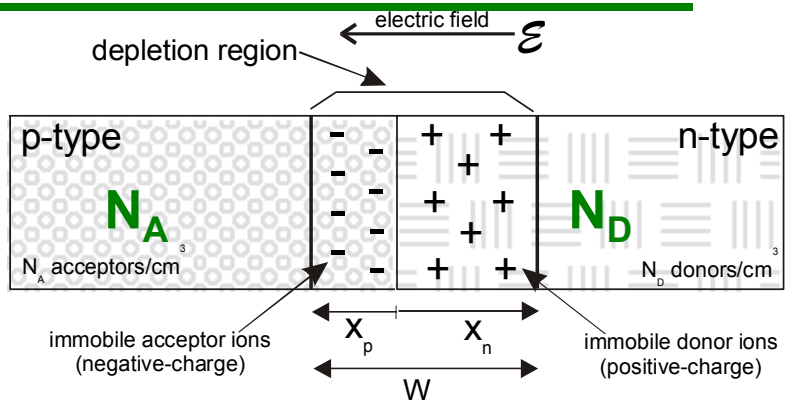
    - where $V_T = kT/q = 26mV$ at room temperature

pn Junctions: Depletion Width

- Depletion Width

  use Poisson's equation & charge neutrality

  - $W = x_p + x_n$



$$x_p = \left[\frac{2\varepsilon(\Psi_0 + V_R)N_D}{qN_A(N_D + N_A)}\right]^{1/2} \qquad x_n = \left[\frac{2\varepsilon(\Psi_0 + V_R)N_A}{qN_D(N_D + N_A)}\right]^{1/2}$$

  - where V_R is applied reverse bias

$$W = \left[\frac{2\varepsilon(\Psi_0 + V_R)}{q}\frac{N_D + N_A}{N_D N_A}\right]^{1/2}$$

$$\Psi_0 = V_T \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

ε is the permittivity of Si

$\varepsilon = 1.04\text{x}10^{-12}$ F/cm

$\varepsilon = K_S\varepsilon_0$, where $\varepsilon_0 = 8.85\text{x}10^{-14}$ F/cm

and $K_S = 11.8$ is the relative permittivity of silicon

- One-sided Step Junction

  - if $N_A \gg N_D$ (p+n diode)
    - most of junction on n-side

$$W \cong x_n = \left[\frac{2\varepsilon(\Psi_0 + V_R)}{qN_D}\right]^{1/2}$$

  - if $N_D \gg N_A$ (n+p diode)
    - most of junction on p-side

$$W \cong x_p = \left[\frac{2\varepsilon(\Psi_0 + V_R)}{qN_A}\right]^{1/2}$$

pn Junctions - Depletion Capacitance

- Free carriers are separated by the depletion layer
- Separation of charge creates junction capacitance
  - Cj = εA/d ⇒ (d = depletion width, W)

$$C_j = A\left[\frac{q\varepsilon N_A N_D}{2(N_A + N_D)}\right]^{1/2}\left(\frac{1}{\sqrt{\Psi_0 + V_R}}\right)$$

ε is the permittivity of Si
ε = 11.8 ε_0 = 1.04x10^{-12} F/cm
V_R = applied reverse bias

  - A is complex to calculate in semiconductor diodes
    - consists of both bottom of the well and side-wall areas
  - Cj is a strong function of biasing
    - must be re-calculated if bias conditions change

$$C_j = \left(\frac{C_{jo}}{\sqrt{1 + \frac{V_R}{\Psi_0}}}\right) \qquad C_{jo} = A\left[\frac{q\varepsilon N_A N_D}{2\Psi_0(N_A + N_D)}\right]^{1/2}$$

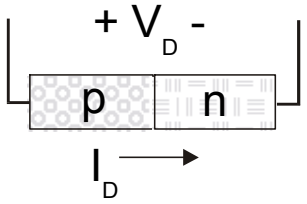
  - CMOS doping is not linear/constant
    - graded junction approximation

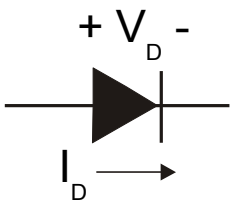
$$C_j = \left(\frac{C_{jo}}{\sqrt[3]{1 + \frac{V_R}{\Psi_0}}}\right)$$

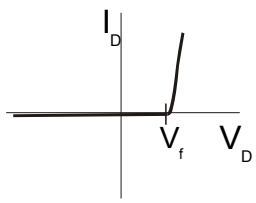
- Junction Breakdown
  - if reverse bias is too high (typically > 30V) can get strong reverse current flow



Diode Biasing and Current Flow



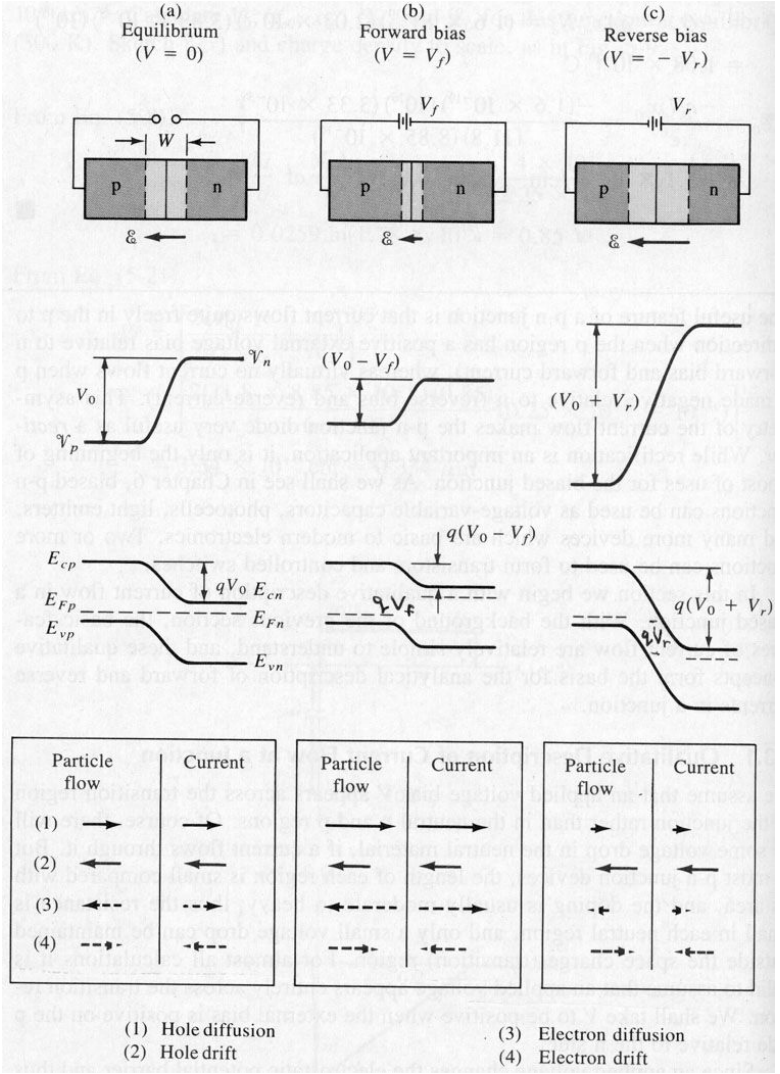




- Forward Bias; $V_D > \Psi_0$
  - acts against built-in potential
  - depletion width reduced
  - diffusion currents increase with V_D
    - minority carrier diffusion

$$I_D = I_S\left(e^{V_D/V_T} - 1\right) \qquad I_S \propto A\left(\frac{1}{N_D} + \frac{1}{N_A}\right)$$

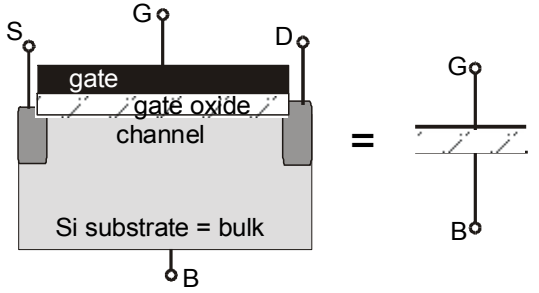
- Reverse Bias; $V_R = -V_D > 0$
  - acts to support built-in potential
  - depletion width increased
  - electric field increased
  - small drift current flows
    - considered leakage
    - small until V_R is too high and breakdown occurs

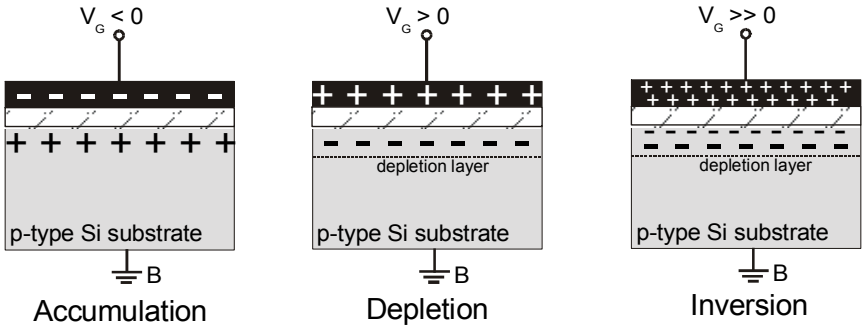


MOSFET Capacitor

- MOSFETs move charge from drain to source underneath the gate, if a conductive channel exists under the gate
- Understanding how and why the conductive channel is produced is important
- MOSFET capacitor models the gate/oxide/substrate region
  - source and drain are ignored
  - substrate changes with applied gate voltage
- Consider an nMOS device
  - Accumulation, $V_G < 0$, (-)ve charge on gate
    - induces (+)ve charge in substrate
    - (+)ve charge accumulate from substrate p+
  - Depletion, $V_G > 0$ but small
    - creates depletion region in subst.
    - (-) charge but no free carriers
  - Inversion, $V_G > 0$ but larger
    - further depletion requires high energy
    - (-) charge pulled from Ground
    - e- free carriers in channel

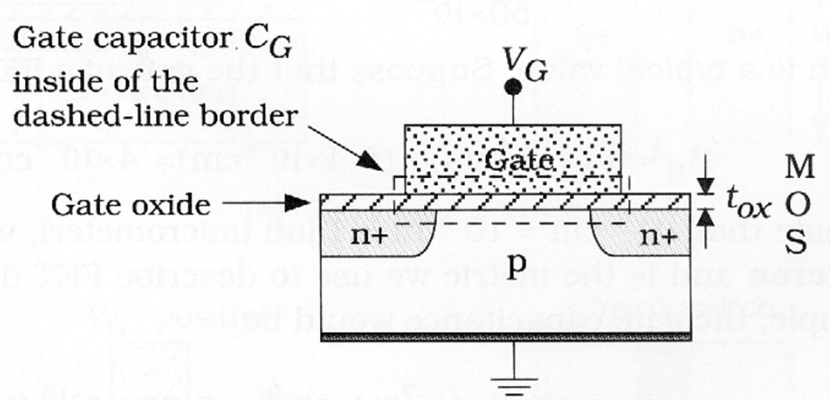
Weak Inversion



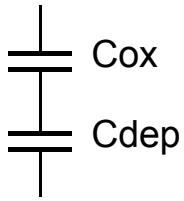


Capacitance in MOSFET Capacitor

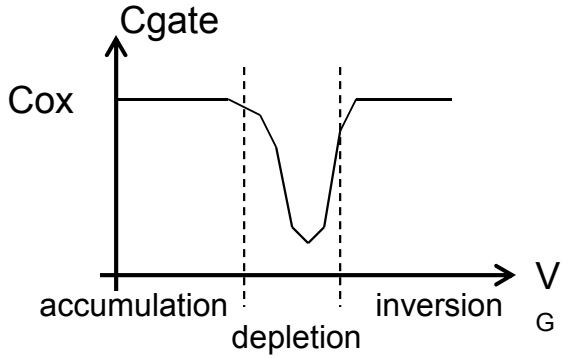
- In Accumulation
  - Gate capacitance = Oxide capacitance
  - $C_{ox} = \varepsilon_{ox}/t_{ox}$ [F/cm^2]



- In Depletion
  - Gate capacitance has 2 components
  - 1) oxide capacitance
  - 2) depletion capacitance of the substrate depletion region
    - Cdep = ε_{si}/x_d, x_d = depth of depletion region into substrate
  - Cgate = Cox || Cdep = Cox Cdep / (Cox+Cdep) < Cox

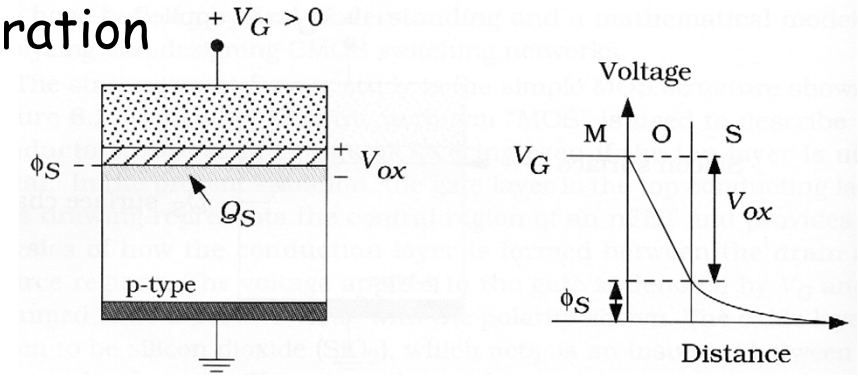


- In Inversion
  - free carries at the surface
  - Cgate = Cox



Inversion Operation

- MOSFET "off" unless in inversion
  - look more deeply at inversion operation
- Define some stuff
  - Qs = total charge in substrate
  - V_G = applied gate voltage
  - Vox = voltage drop across oxide
  - ϕ_s = potential at silicon/oxide interface (relative to substrate-ground)
  - Qs = - Cox V_G
  - V_G = Vox + ϕ_s
- During Inversion (for nMOS)
  - V_G > 0 applied to gate
  - Vox drops across oxide (assume linear)
  - ϕ_s drops across the silicon substrate, most near the surface

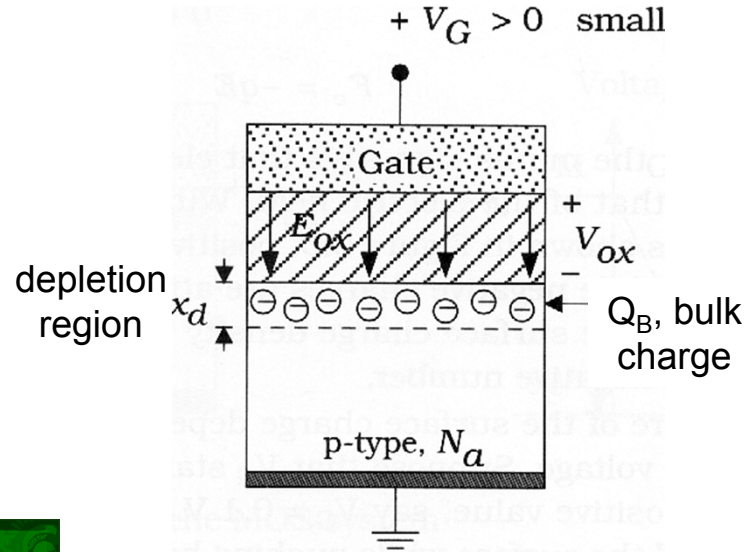


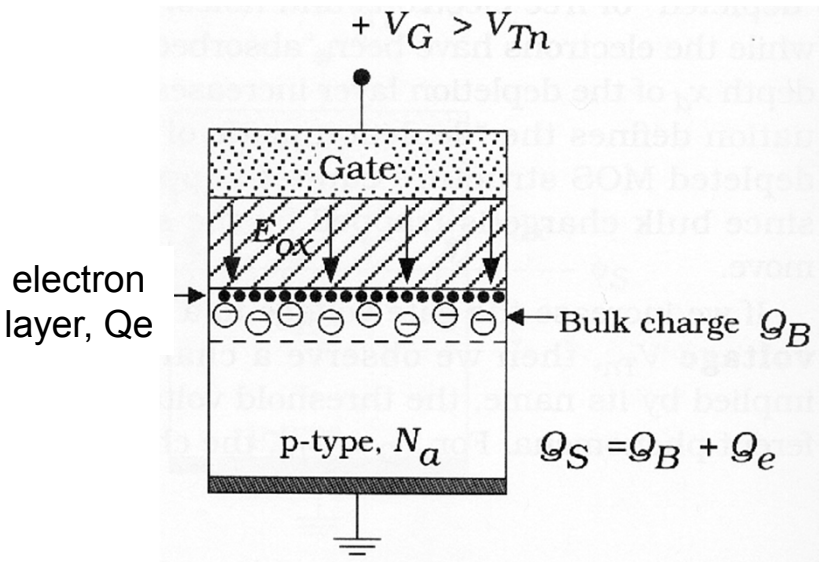
Surface Charge

- Q_B = bulk charge, ion charge in depletion region under the gate
  - $Q_B = -q N_A x_d$, x_d = depletion depth
  - $Q_B = -(2q \varepsilon_{Si} N_A \phi_s)^{1/2} = f(V_G)$
  - charge per unit area
- Qe = charge due to free electrons at substrate surface
- $Qs = Q_B + Qe < 0$ (negative charge for nMOS)

recall from pn junction, when $N_D >> N_A$

$$W \cong x_n = \left[\frac{2\varepsilon(\Psi_0 + V_R)}{qN_A}\right]^{1/2}$$





Surface Charge vs. Gate Voltage

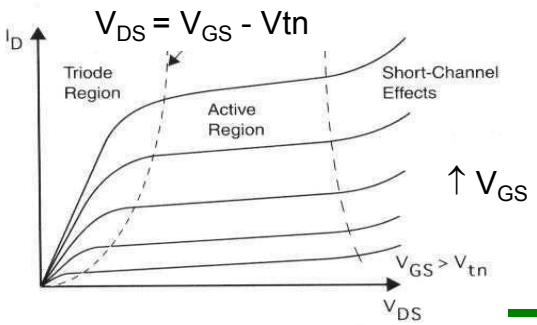
- Surface Charge vs. Gate Voltage
  - V_G < Vtn, substrate charge is all bulk charge, Qs = Q_B
  - V_G = Vtn, depletion region stops growing
    - x_d at max., further increase of V_G will NOT increase x_d
    - Q_B at max.
  - V_G > Vtn, substrate charge has both components, Qs = Q_B + Qe
    - since Q_B is maxed, further increases in V_G must increase Qe
    - increasing Qe give more free carriers thus less resistance

- Threshold Voltage
  - **Vtn defined as gate voltage where Qe starts to form**
  - Qe = -Cox (V_G-Vtn)
  - Vtn is gate voltage required to
    - overcome material difference between silicon and oxide
    - establish depletion region in channel to max value/size

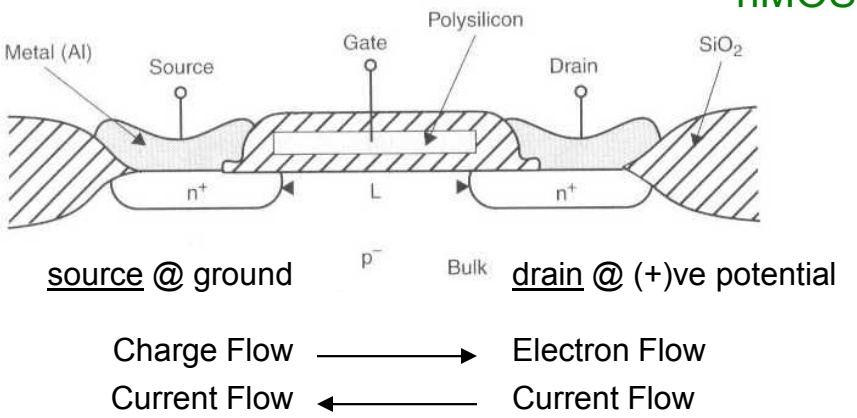


Overview of MOSFET Current

- Gate current
  - gate is essentially a capacitor ⇒ **no current through gate**
  - gate is a control node
    - V_G < Vtn, device is off
    - V_G > Vtn, device is on and performance is a function of V_{GS} and V_{DS}
- Drain Current (current from drain to source), I_D
  - Source = source/supply of electrons (nMOS) or holes (pMOS)
  - Drain = drain/sink of electrons (nMOS) or holes (pMOS)
  - V_{DS} establishes an E-field across the channel (horizontally)
    - free charge in an E-field will create a drain-source current
    - is I_D drift or diffusion current?
- MOSFET I-V Characteristics



nMOS



nMOS Current vs. Voltage

- **Cutoff Region**
  - $V_{GS} < Vtn$

  $\Rightarrow I_D = 0$ (Not quiet--there is leakage "subthreshold current")

- **Linear Region**
  - $V_{GS} > Vth$, $V_{DS} > 0$ but very small
    - $Qe = -Cox\,(V_{GS}-Vtn)$
    - $I_D = \mu_n Qe\,(W/L)\,V_{DS}$

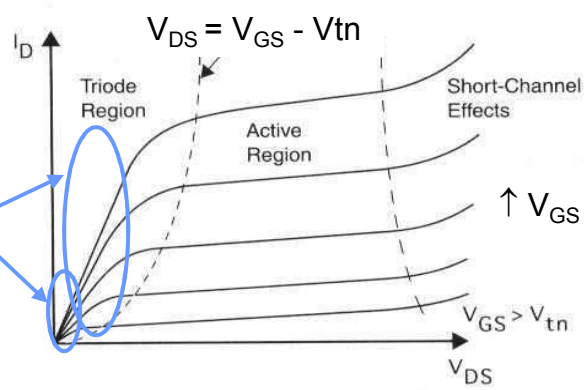
  $\Rightarrow I_D = \mu_n Cox\,(W/L)\,(V_{GS}-Vtn)\,V_{DS}$

- **Triode Region**
  - $V_{GS} > Vth$, $0 < V_{DS} < V_{GS}-Vth$
    - surface potential, ϕ_s , at drain now $f(V_{GS}-V_{DS}=V_{GD}) \Rightarrow$ less charge near drain
    - assume channel charge varies linearly from drain to source
      - at source: $Qe = -Cox\,(V_{GS}-Vtn)$, at drain: $Qe = 0$

  $\Rightarrow$ $$I_D = \frac{\mu_n C_{OX}}{2}\frac{W}{L}\left[2(V_{GS}-V_t)V_{DS}-V_{DS}^2\right]$$

General Integral for expressing ID
- channel charge = f(y)
- channel voltage = f(y)
- y is direction from drain to source

$$I_D = \alpha \int_0^{V_D} Q_I(y)\delta V(y)$$



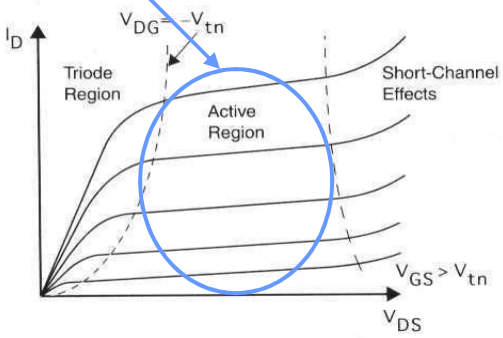
nMOS Current vs. Voltage

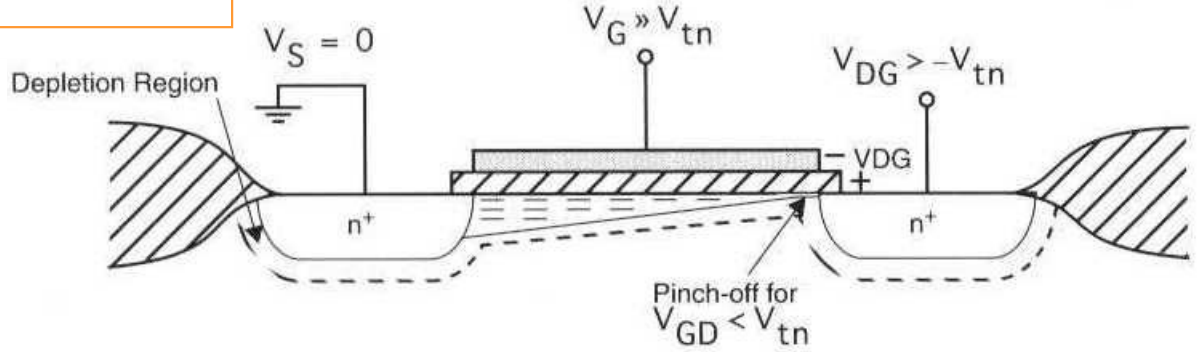
- Saturation Region (Active Region)
  - V_{GS} > Vtn, V_{DS} > V_{GS}-Vtn
    - surface potential at drain, ϕ_{sd} = V_{GS}-Vtn-V_{DS}
    - when V_{DS} = V_{GS}-Vtn, ϕ_{sd} = 0 $\Rightarrow$ channel not inverted at the drain
      - channel is said to be pinched off
    - during pinch off, further increase in V_{DS} will not increase I_D
      - define saturation voltage, Vsat, when V_{DS} = V_{GS}-Vtn
    - current is saturated, no longer increases
    - substitute Vsat=V_{GS}-Vtn for V_{DS} into triode equation

$\Rightarrow$ square law equation

$$I_D = \frac{\mu_n C_{OX}}{2}\frac{W}{L}(V_{GS} - V_t)^2$$

$$I_D = \frac{\mu_n C_{OX}}{2}\frac{W}{L}\left[2(V_{GS} - V_t)V_{DS} - V_{DS}^2\right]$$





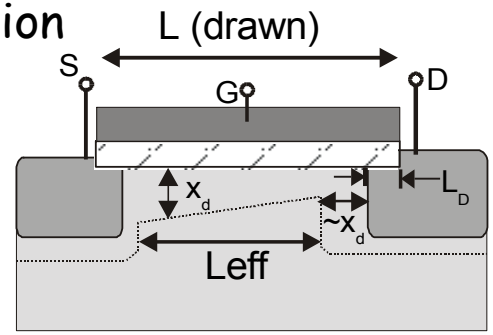
gure 3.19 When V_{DS} is increased such that $V_{GD} < V_{tn}$, then the channel becomes pinched off ... the drain end.

Other Stuff

- Transconductance
  - process transconductance, k' = $\mu_n C_{ox}$
    - constant for a given fabrication process
  - device transconductance, β_n= k' W/L
- Surface Mobility
  - mobility at the surface is lower than mobility deep inside silicon
  - for current, I_D, calculation, typical μ_n = 500-580 $cm^2/V\text{-}sec$
- Effective Channel Length
  - effective channel length reduced by
    - lateral diffusion under the gate
    - depletion spreading from drain-substrate junction

$$Leff = L(drawn) - 2L_D - X_d$$

$$X_d = \sqrt{\left(\frac{2\varepsilon_s\left(V_D - \left(V_G - V_t\right)\right)}{qN_A}\right)}$$



Second Order Effects

- Channel Length Modulation
  - Square Law Equation predicts I_D is constant with V_{DS}
  - However, I_D actually increases slightly with V_{DS}
    - due to effective channel getting shorter as V_{DS} increases
    - effect called channel length modulation
  - Channel Length Modulation factor, λ
    - models change in channel length with V_{DS}
  - Corrected I_D equation

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

    - Veff = V_{GS} - Vtn
- Body Effect
  - so far we have assumed that substrate and source are grounded
  - if source not at ground, source-to-bulk voltage exists, $V_{SB} > 0$
  - $V_{SB} > 0$ will increase the threshold voltage, Vtn = f(V_{SB})
  - called Body Effect, or Body-Bias Effect



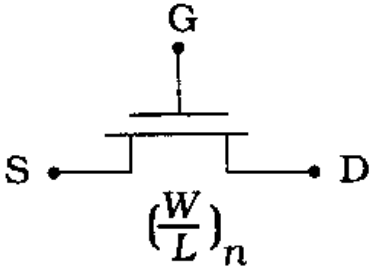
pMOS Equations

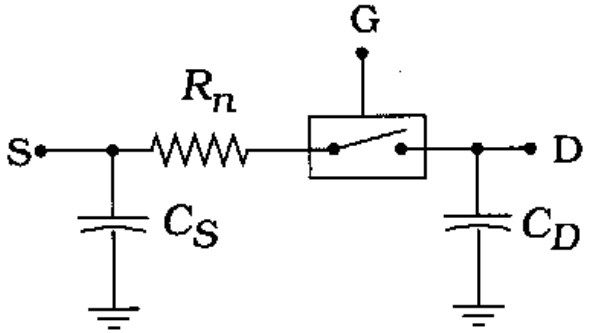
- analysis of nMOS applies to pMOS with following modifications
  - physical
    - change all n-type regions to p-type
    - change all p-type regions to n-type
      - substrate is n-type (pWell)
    - channel charge is positive (holes) and (+)ve charged ions
  - equations
    - change V_{GS} to V_{SG} (V_{SG} typically = VDD - V_G)
    - change V_{DS} to V_{SD} (V_{SD} typically = VDD - V_D)
    - change Vtn to |Vtp|
      - pMOS threshold is negative, nearly same magnitude as nMOS
  - other factors
    - lower surface mobility, typical value, μ_p = 220 cm^2/V-sec
    - body effect, change V_{SB} to V_{BS}



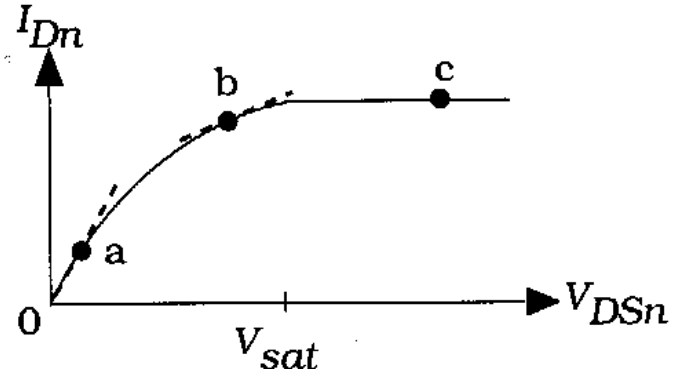
MOSFET RC Model

- Modeling MOSFET resistance and capacitance is very important for transient characteristics of the device
- RC Model





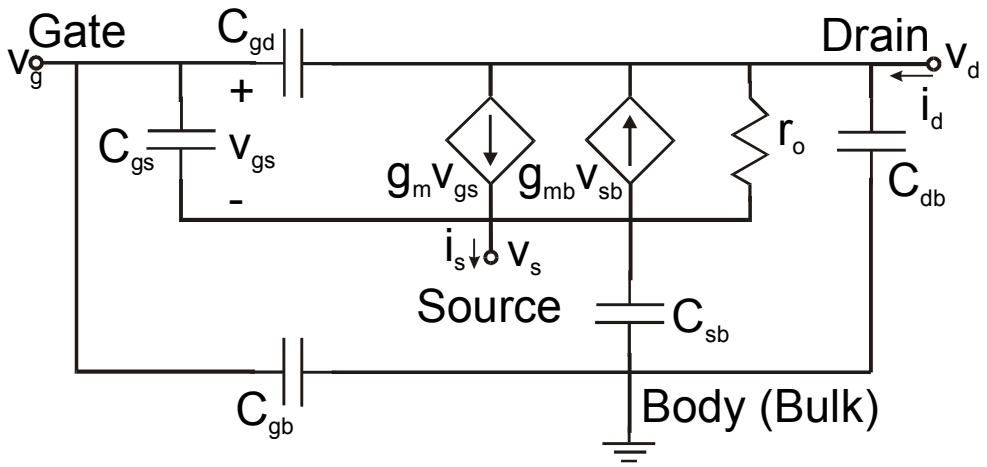
- Drain-Source (channel) Resistance, Rn
  - $Rn = V_{DS} / I_D$
    - function of bias voltages
  - point (a), linear region
    - $Rn = 1/[\beta_n(V_{GS}-Vtn)]$
  - point (b), triode region
    - $Rn = 2/\{\beta_n[2(V_{GS}-Vtn)-V_{DS}]\}$
  - point (c), saturation region
    - $Rn = 2V_{DS} / [\beta_n (V_{GS}-Vtn)^2]$

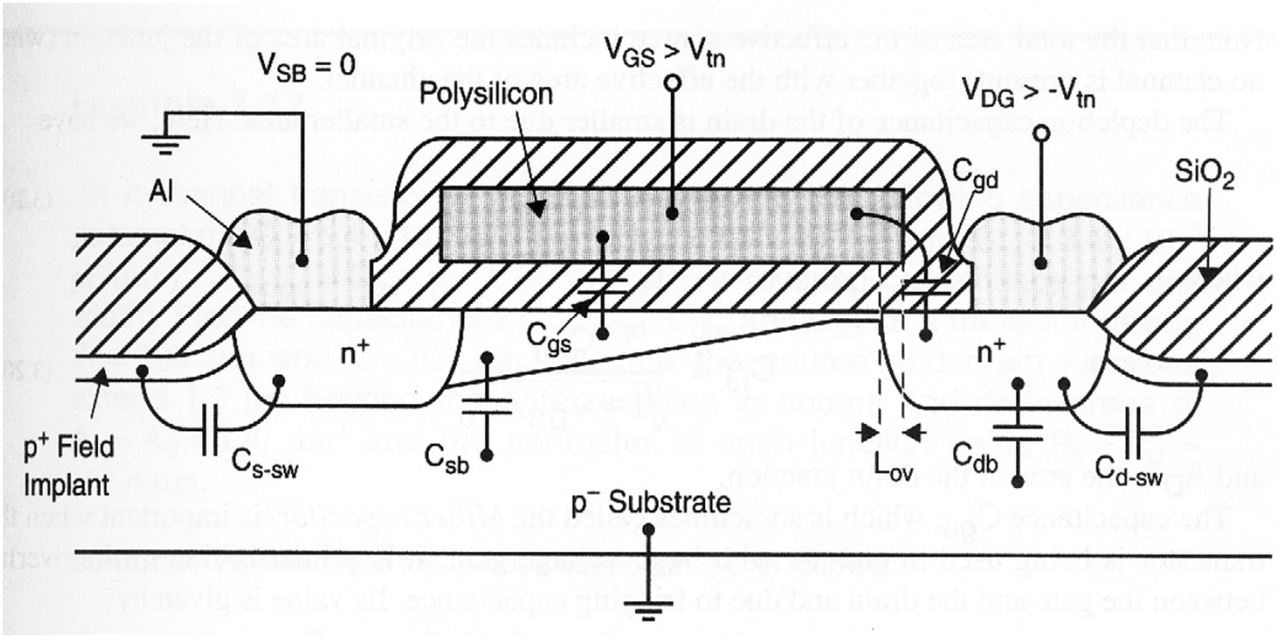


- general model equation
  - $Rn = 1/[\beta_n(V_{DD}-Vtn)]$

MOSFET Capacitances -Preview

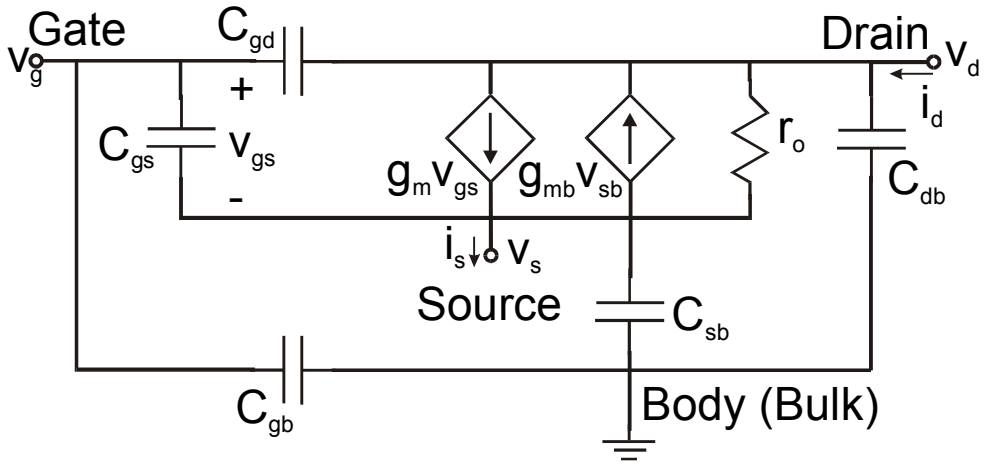
- Need to find C_S and C_D
- MOSFET Small Signal model
  - Model Capacitances
    - **Cgs**
    - **Cgd**
    - Cgb
    - **Cdb**
    - **Csb**
    - no Csd!
- MOSFET Physical Capacitances
  - layer overlap
  - pn junction





MOSFET conductances –Preview#

- How do you determine g_m and r_o ?



- This is a "small-signal" MOSFETmodel or "AC" model. I.e., it is a linearized version of the Ids equations about an operating "DC' point." One has

In the triode region: $I_D = \frac{\mu_n C_{OX}}{2}\frac{W}{L}\left[2(V_{GS}-V_t)V_{DS}-V_{DS}^2\right]$, or

In the statuation region (assumed): $I_D = \frac{\mu_n C_{OX}}{2}\frac{W}{L}(V_{GS}-V_t)^2$

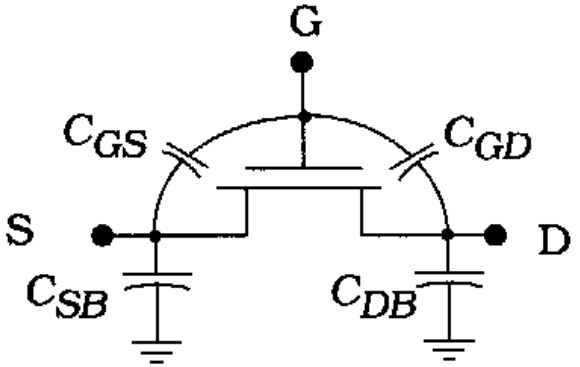
Then (where * means the DC operating point)

$$g_m := \frac{\partial I_{DS}}{\partial Vgs}\Big|_*$$

$$r_o = 1/g_o = 1/\frac{\partial I_{DS}}{\partial Vds}\Big|_*$$

RC Model Capacitances

- Why do we care?
  - capacitances determine switching speed
- Important Notes
  - models developed for saturation (active) region
  - models presented are simplified (not detailed)
- RC Model Capacitances
  - Source Capacitance
    - models capacitance at the Source node
    - $C_S = C_{GS} + C_{SB}$
  - Drain Capacitance
    - models capacitance at the Drain node
    - $C_D = C_{GD} + C_{DB}$



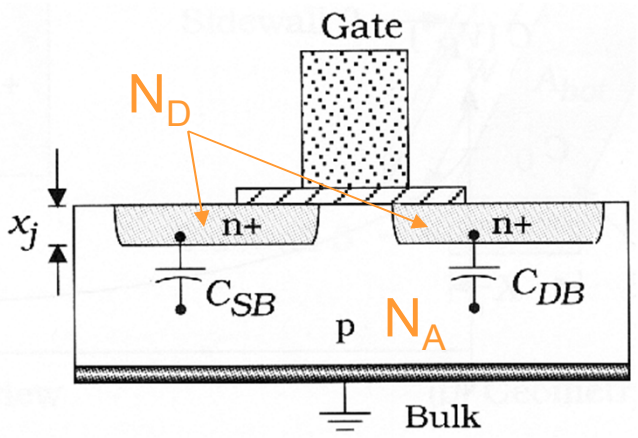
(a) nFET

What are C_{GS}, C_{GD}, C_{SB}, and C_{DB}?

MOSFET Parasitic Capacitances

- Gate Capacitance
  - models capacitance due to overlap of Gate and Channel
    - C_G = Cox W L
  - estimate that C_G is split 50/50 between Source and Drain
    - $C_{GS} = \frac{1}{2} C_G$
    - $C_{GD} = \frac{1}{2} C_G$
  - assume Gate-Bulk capacitance is negligible
    - models overlap of gate with substrate outside the active tx area
    - $C_{GB} = 0$
- Bulk Capacitance
  - C_{SB} (Source-Bulk) and C_{DB} (Drain-Bulk)
    - pn junction capacitances

$$C_j = \left(C_{jo} \Big/ \sqrt{1 + \frac{V_R}{\Psi_0}} \right) \qquad C_{jo} = A \left[\frac{q \varepsilon N_A N_D}{2\Psi_0 (N_A + N_D)} \right]^{1/2}$$



What are V_R, Ψ_0, N_A, and N_D?

MOSFET Junction Capacitances

- Capacitance/area for pn Junction

$$C_j = C_{jo} \Big/ \left(1 + \frac{V_R}{\Psi_0}\right)^{m_j} \qquad C_{jo} = \left[\frac{q\varepsilon N_A}{2\Psi_0}\right]^{1/2} \qquad \Psi_0 = V_T \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

m_j = grading coefficient (typically 1/3) assuming N_D (n+ S/D) >> N_A (p subst.)

- S/D Junction Capacitance
  - zero-bias capacitance
    - highest value when V_R = 0, assume this for worst-case estimate
    - $C_j = C_{jo}$
  - $C_{S/Dj} = C_{jo} A_{S/D}$, $A_{S/D}$ = area of Source/Drain
    - what is $A_{S/D}$?
    - complex 3-dimensional geometry
      - bottom region and sidewall regions
  - $C_{S/Dj}$ = Cbot + Csw
    - bottom and side wall capacitances

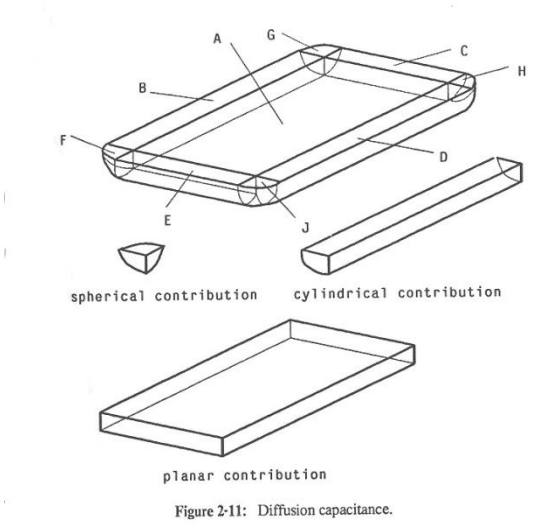
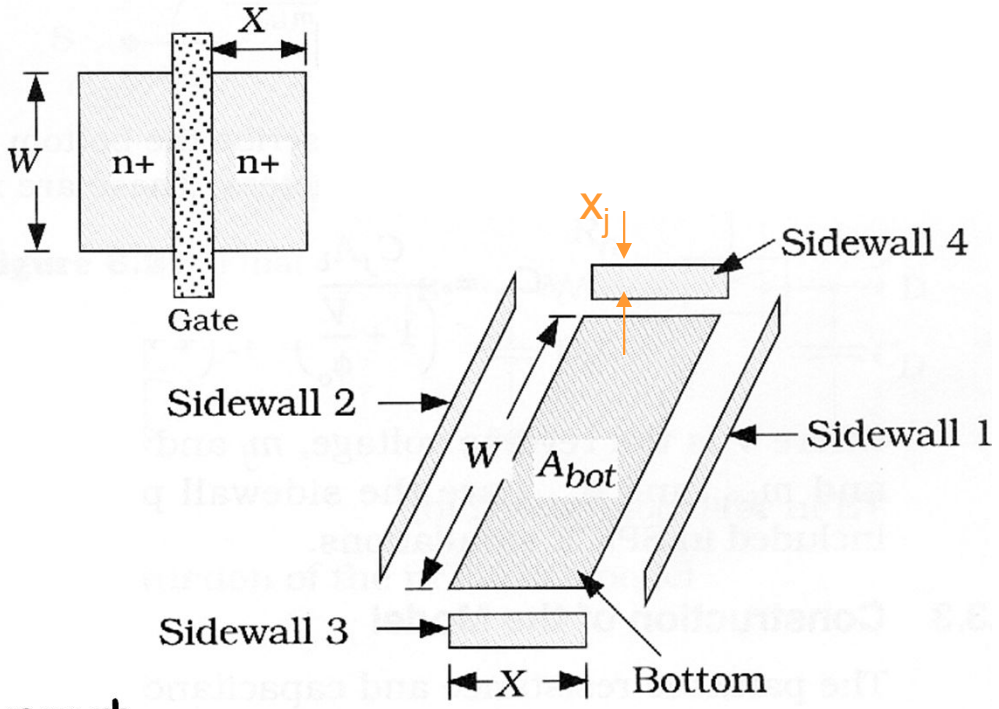


Figure 2-11: Diffusion capacitance.

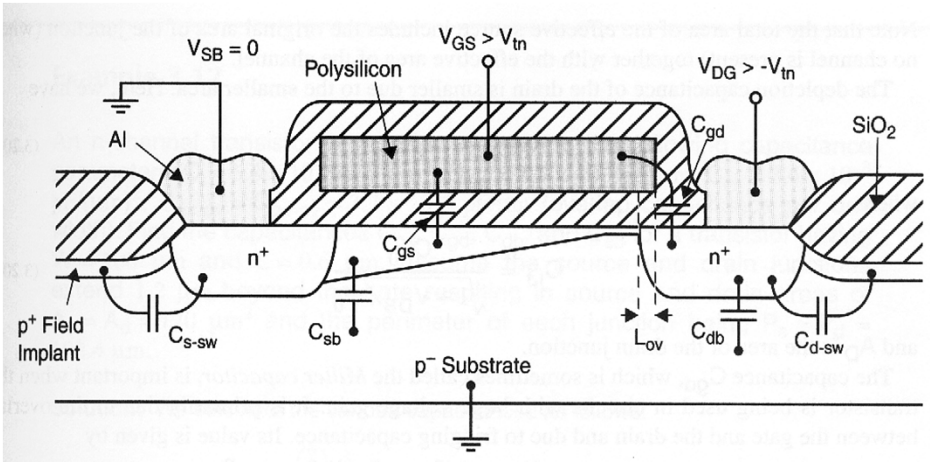
Junction Capacitance

- Bottom Capacitance
  - $C_{bot} = C_j A_{bot}$
    - $A_{bot} = X W$
- Sidewall Capacitance
  - $C_{sw} = C_{jsw} P_{sw}$
    - $C_{jsw} = C_j x_j$ [F/cm]
      - x_j = junction depth
    - P_{sw} = sidewall perimeter
      - $P_{sw} = 2 (W + X)$
- Accounting for Gate Undercut
  - junction actually under gate also due to lateral diffusion
  - $X \Rightarrow X + L_D$ (replace X with $X + L_D$)
- Total Junction Cap
  - $C_{S/Dj} = C_{bot} + C_{sw} = C_j A_{bot} + C_{jsw} P_{sw} = C_{S/Dj}$

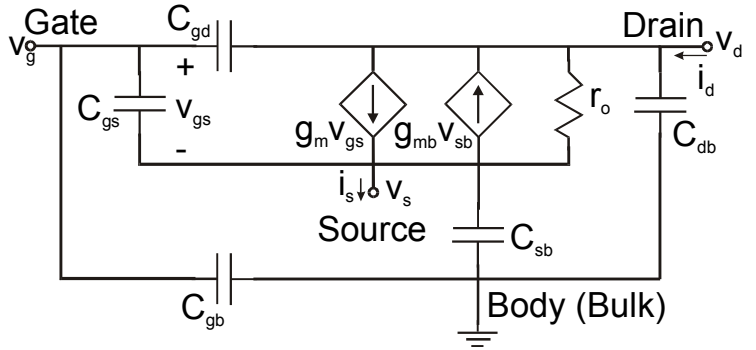


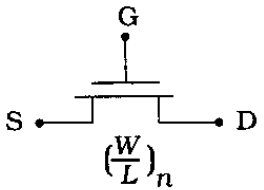
MOSFET Bulk Capacitances

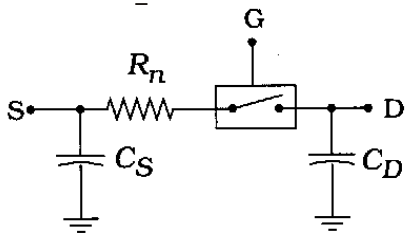
- General Junction Capacitance
  - $C_{S/Dj}$ = Cbot + Csw
- C_{SB} (Source-Bulk)
  - $C_{SB} = C_j A_{Sbot} + C_{jsw} P_{Ssw}$
- C_{DB} (Drain-Bulk)
  - $C_{DB} = C_j A_{Dbot} + C_{jsw} P_{Dsw}$



- RC Model Capacitances
  - Source Capacitance
    - $C_S = C_{GS} + C_{SB}$
  - Drain Capacitance
    - $C_D = C_{GD} + C_{DB}$







Junction Areas:examples

- Note: calculations assume following design rules
  - poly size, L = 2λ
  - poly space to contact, 2λ
  - contact size, 2λ
  - active overlap of contact, 1λ

$\Rightarrow$ W = 4λ
X1 = 5λ, X2= 2λ, X3 = 6λ

- Non-shared Junction with Contact
  - Area: X1 W = (5)(4) = $20\lambda^2$
  - Perimeter: 2(X1 + W) = 18λ
- Shared Junction without Contact
  - Area: X2 W = (2)(4)λ^2 = $8\lambda^2$
  - Perimeter: 2(X2 + W) = 12λ
    - much smaller!
- Shared Junction with Contact
  - Area: X3 W = (6)(4)λ^2 = $24\lambda^2$
  - Perimeter: 2(X3 + W) = 20λ
    - largest area!

