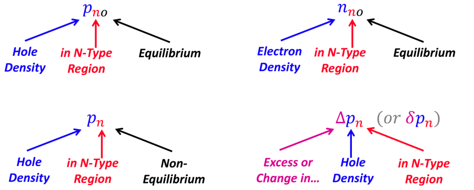
SEMICONDUCTORS & DEVICES

Jean Mégret (based on the summary of Simon Erne)
Professor C.R. Bolognesi – FS 2020

BASICS

Carrier density naming convention:



CONSTANTS & MATERIAL PARAMETERS

Planck's const.	h	$=$	$6.625 \cdot 10^{-34}$	Js
		$=$	$4.136 \cdot 10^{-15}$	eVs
Boltzmann const.	k	$=$	$1.38 \cdot 10^{-23}$	J/K
		$=$	$8.67 \cdot 10^{-5}$	eV/K
Thermal Voltage V_{Th} @300K	$\frac{kT}{q}$	$=$	$26mV$	
	kT	$=$	$25.9meV$	
	$\frac{q}{kT}$	$=$	$38.61\frac{1}{V}$	
elementary charge	q	$=$	$1.602 \cdot 10^{-19}$	C
free electron mass	m_0	$=$	$9.1 \cdot 10^{-31}$	kg
vacuum permeability	ε_0	$=$	$8.854 \cdot 10^{-14}\frac{F}{cm}$	
		$=$	$8.854 \cdot 10^{-12}\frac{F}{m}$	

$1eV = 1.602 \cdot 10^{-19}J$

Silicon @300K

intrinsic carrier conc.	n_i	$=$	$9.65 \cdot 10^9$	cm^{-3}
	n_i^2	$=$	$9.3 \cdot 10^{19}$	cm^{-6}
eff. DOS in conduct. band	N_C	$=$	$2.86 \cdot 10^{19}$	cm^{-3}
eff. DOS in valence band	N_V	$=$	$2.66 \cdot 10^{19}$	cm^{-3}
band gap	E_G	$=$	$1.12\ eV = 43.3\ kT$	
rel. permeability Si	ε_s	$=$	$11.9 \cdot \varepsilon_0$	
rel. permeability $Si0_2$	ε_{ox}	$=$	$3.9 \cdot \varepsilon_0$	
Thermal velocity	v_{th}	$\approx$	$10^7\frac{cm}{s}$	
eff. hole mass	m_p^*	$=$	$1.040 \cdot m_0$	

DOS ≔ Density of states

Galiumarsenid (GaAs) @300K

eff. DOS in valence band	N_V	$=$	$7.0 \cdot 10^{18}$	cm^{-3}
band gap	E_G	$=$	$1.42\ eV$	
eff. hole mass	m_p^*	$=$	$0.427 \cdot m_0$	

electrostatics

- Force is the negative gradient of potential energy:
  electrons: $F_e = -\frac{dE_{pot,e}}{dx} = -\frac{dE_C}{dx} = -q \cdot \vec{E}$
  holes: $F_h = -\frac{dE_{pot,h}}{dx} = -\frac{d(-E_V)}{dx} = q \cdot \vec{E}$
- conductivity: $\sigma = \frac{j_{dr}}{E} = q(n\mu_n + p\mu_p)$ $[(\Omega\ cm)^{-1}]$
- resistivity: $\rho = 1/\sigma$ $[\Omega\ cm]$
- resistance: $R = \rho\frac{Length}{Area}$ $[\Omega]$
- Electrical current: $\mathrm{I} = \rho\frac{dQ}{dt}$ $[A = \frac{C}{s}]$

Electric Field: $\vec{E} = -\frac{dV}{dx}, (E = \frac{F}{Q} = \frac{kq}{r^2})$

Potential: $V = -\int_0^l E\ dx = -\frac{1}{q}(E_V - E_{ref})$

Potential Energy: $E_C - E_{ref} = -q \cdot V$

in the band diagram an electric field is represented by band banding

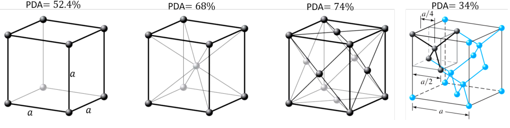
CRYSTALS

Coordination Number:

The number of nearest neighbour atoms (if bonded or not)

- simple cubic:	6	- body centred cubic:	8
- diamond:	4	- face centred cubic:	12

simple cubic	body centered cubic	face centered cubic	diamond
1 atoms/cell	2 atoms/cell	4 atoms/cell	8 atoms/cell
$r = a/2$	$r = \sqrt{3}a/4$	$r = \sqrt{2}a/4$	$r = \sqrt{3}a/8$
PDA= 52.4%	PDA= 68%	PDA= 74%	PDA= 34%



DENSITIES

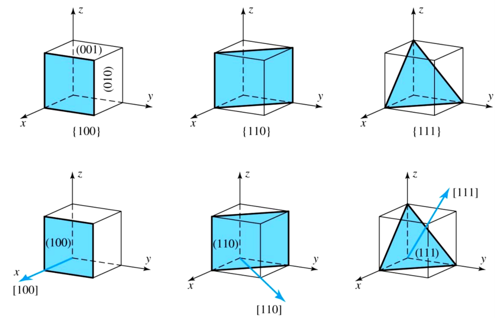
- Volume Packing Density

$$\mathrm{VPD} = \frac{V_{atoms}}{V_{cube}} = \frac{\left(\#_{corners}\cdot\frac{1}{8}+\#_{faces}\cdot\frac{1}{2}+\#_{center}\right)\cdot\frac{4\pi}{3}r^3}{a^3} = VD \cdot \frac{\frac{4\pi}{3}r^3}{a^3}$$

a ≔ unit cell length, $2r$ ≔ distance between nearest neighbours

- Volume Density
  $VD = \frac{\#_{atoms}}{V_{cube}}$
- Surface Density
  $SD = \frac{\#_{atoms}}{S_{cube}}$
- Surface Packing Density
  $SPD = \frac{S_{atoms}}{S_{square}}$

Miller Indices



Crystals have different properties in different directions.

$V_{cone} = \frac{h\pi r^2}{3}$, $A_{hexagon} = \frac{3\sqrt{3}s^2}{2}$, $V_{sphere} = \frac{4}{3}\pi r^3$, $A_{circle} = \pi r^2$

SEMICONDUCTORS BASICS

At 0K, the valence band is full, and the conduction band is empty. There is zero net current and the material is therefore isolating. Each atom is surrounded by a complete shell of 8 e-. For temperatures $> 0K$ the thermal energy **excites electrons** into an available ($4N$) empty state of the conduction band and leaves an empty state (**hole**) in the valence band behind. The atom which lost the electron is now **electronegative**.

steady state/equilibrium:

zero current flow ⇒ fermi level must be flat, intrinsic fermi level might bend due to nonuniform doping

free electron density: n_0
free hole density: p_0

intrinsic carrier concentration: $n_i = n_0 = p_0$

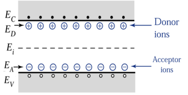
The intrinsic carrier concentration depends exponentially on temperature and the Activation Energy (for bond breaking, it is somewhat related to the bandgap of the material).

$n_i = Ke^{-\frac{E_a}{kT}}$

Doping ≙ defects

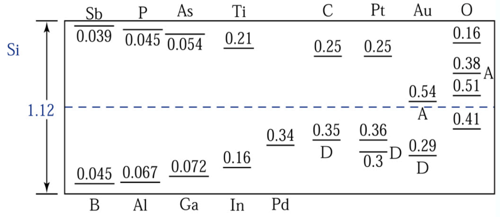
N-type: N_D Donors (majority: electrons)
Donors introducing an energy state E_D

P-type: N_A Acceptors (majority: holes)
Acceptors introducing an energy state E_A



If a material is doped with $N_A = N_A - N_D$ if $N_A > N_D$ else $N_D = N_D - N_A$

Overall the solid is still neutral when doped.
Shallower dopants (i.e. E_D (E_A) is close to the conduction (valence) band) show higher ionization (fully ionized at smaller temperatures).



Assumption: complete ionization | S3A3a

fully ionized donors: $N_D \approx n_0$ — N_D ≔ density of Donors
fully ionized acceptors: $N_A \approx p_0$ — N_A ≔ density of Acceptors

To check reasonability of assumption (analogous for N_A):
We want to compare the number of ionized donors with the total number of donors.
Let N_D^+ be the number of ionized states, n_d be the number of non-ionized atoms and E_D be the donor energy level.
$n_0 = N_D^+$ (we neglect therm. generation since $N_D^+ \gg n_i$)
$N_D = n_d + N_D^+$
On one hand, we know that:
$n_0 = N_C e^{-(E_C - E_F)/kT}$
On the other hand, we can equate the probability of finding an electron in the donor state times the total number of donors to n_d:

$$n_d = N_D * \mathcal{P}(e^- in\ E_D) = \frac{N_D}{1 + e^{(E_F - E_D)/kT}} \approx N_D e^{-(E_F - E_D)/kT}$$

Where we used the Maxwell Boltzmann approximation. So:

$$\frac{N_D^+}{N_D} = \frac{n_0}{N_D^+ + n_d} = \frac{1}{1 + \frac{n_d}{n_0}} = \frac{1}{1 + \frac{N_D}{N_C}e^{(E_D - E_F)/kT}}$$

If this ratio is larger than 0.95 = 95%, we can assume complete ionization.

At equilibrium we have:

Electroneutrality equation: $\boldsymbol{n_0 + N_A = p_0 + N_D}$
Mass action law: $\boldsymbol{n_0 \cdot p_0 = n_i^2}$

So at very low temperatures, we have little ionization.
With rising temperature, the number of ionized donors rises up (as shown above) to $n_0 = N_D$. When it reaches this number, n_0 doesn't grow any more with the temperature for a while.
At very high temperatures, the intrinsic carrier concentration n_i dominates again over the different dopings ($n_i > N_D^+$). This means the material becomes intrinsic again. And $n_0 = n_i = p_0$.

If assumption of complete ionization isn't fulfilled:
⇒ $n_0^2 - n_0(N_D - N_A) - n_i^2 = 0$

$$\rightarrow n_0 = \frac{N_D - N_A}{2} + \sqrt{\left(\frac{N_D - N_A}{2}\right)^2 + n_i^2}$$

$$\rightarrow p_0 = \frac{N_A - N_D}{2} + \sqrt{\left(\frac{N_A - N_D}{2}\right)^2 + n_i^2}$$

direct band gap:
if the minimum of the conduction band is at the same place as the maximum of the valence band.

indirect band gap:
if it is not a direct semiconductor. This means a phonon is also needed actually get an electron from valence to conduction band. Problematic in optical systems.

POPULATION OF ELECTRON STATES

Fermi Dirac Statistics

$F(E)$ = probability of finding an electron with energy E.

$$F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}$$

- $F(E_F) = 0.5$
- probability of finding a hole $= 1 - F(E)$
- in a range of $\pm 2kT$ around E_F the most action takes place

Maxwell Boltzmann Approximation

The Maxwell Boltzmann Approximation holds, if the Fermi Level is far away from the band edges ⇔ low doping

$F(E) \simeq e^{-\frac{(E - E_F)}{kT}}$, $E \gg E_F$

DOS at the conduction/valence band edge

$\boldsymbol{N_C} = \frac{4\sqrt{2}(\pi \cdot m_n^* \cdot kT)^{3/2}}{h^3}$ $\boldsymbol{N_V} = \frac{4\sqrt{2}(\pi \cdot m_p^* \cdot kT)^{3/2}}{h^3}$

m^* ≡ effective masses of carriers in the CB and VB

effective mass

near band extremum: $\frac{\partial E}{\partial k} = 0$, $\frac{\partial^2 E}{\partial k^2} \approx \frac{1}{m^*}$

- the greater the curvature of bands vs. momentum k, the lower the effective mass
- the smaller the energy gap, the lighter the electron effective mass

concentration of electrons/holes

$n_0 = \int_{E_C}^{\infty} f(E) \cdot D_C(E)dE$
$p_0 = \int_{-\infty}^{E_V}(1 - f(E)) \cdot D_V(E)dE$

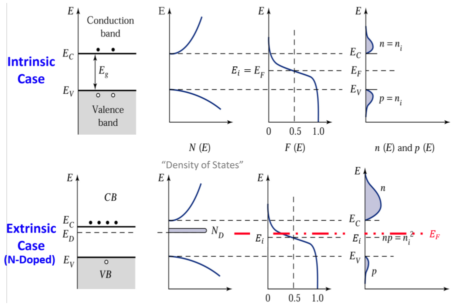
Density of States:
$D_C = \frac{8\pi\sqrt{2}}{h^3}(m_n^*)^{3/2}\sqrt{E - E_C}$
$D_V = \frac{8\pi\sqrt{2}}{h^3}(m_p^*)^{3/2}\sqrt{E_V - E}$

with Maxwell Boltzmann Approximation *(3kT away from band edge)*

$n_0 = N_C e^{-(E_C - E_F)/kT}$ $E_C - E_F = -kT \cdot \ln(n_0/N_C)$
$p_0 = N_V e^{-(E_F - E_V)/kT}$ $E_V - E_F = kT \cdot \ln(p_0/N_V)$

in equilibrium:

$n_i^2 = n_0 \cdot p_0 = N_C N_V \cdot e^{-(E_C - E_V)/kT} = N_C N_V \cdot e^{-(E_G)/kT}$
$n_i = N_C \cdot e^{-(E_C - E_i)/kT} = N_V \cdot e^{(E_V - E_i)/kT}$
N-Type: $n_0 = n_i e^{(E_F - E_i)/kT}$ $\rightarrow p_0 = n_i^2/n_0$
***P*-Type:** $p_0 = n_i e^{(E_i - E_F)/kT}$ $\rightarrow n_0 = n_i^2/p_0$



FERMI LEVEL

$$E_i - E_F = \begin{cases} 0 & \text{intrinsic} \\ -kT\cdot\ln(N_D/n_i) & \text{N − type doping} \\ +kT\cdot\ln(N_A/n_i) & \text{P − type doping} \end{cases}$$

intrinsic fermi level: $E_i = \frac{1}{2}(E_C + E_V) + \frac{1}{2}kT\cdot\ln\left(\frac{N_V}{N_C}\right)$
$\approx \frac{1}{2}(E_C + E_V)$

extrinsic fermi level:
$E_F - E_C = \quad kT\cdot\ln(n_0/N_C)$ N-type doping
$E_F - E_V = -kT\cdot\ln(p_0/N_V)$ P-type doping
$E_F - E_i = -kT\cdot\ln(p_0/n_i) = kT\cdot\ln(n_0/n_i)$
E_F is always noted depending on another Energy level

If E_F is above E_C the material is said to be degenerate.

EXTRINSIC SEMICONDUCTORS: APPROXIMATIONS

n-type $(N_D \gg N_A)$

- $N_D \gg n_i \quad\Rightarrow\quad n_0 + N_D \approx N_D, \quad p_0 \approx \frac{n_i^2}{n_0} = \frac{n_i^2}{N_D}$
- majority: $e^- \Rightarrow E_F$ closer to E_C

p-type $(N_D \ll N_A)$

- $N_A \gg n_i \quad\Rightarrow\quad p_0 + N_A \approx N_A, \quad n_0 \approx \frac{n_i^2}{p_0} = \frac{n_i^2}{N_A}$
- majority: *holes* $\Rightarrow E_F$ closer to E_V

Fermi level must be flat at equilibrium. Otherwise, there would be transport mechanisms (current) which violates the concept of equilibrium.

GENERATION / RECOMBINATION

Generation and Recombination work to restore equilibrium conditions:

Excess of Minority Carriers → **Recombination**
Depletion of Minority Carriers → **Generation**

steady state:

In steady state, electrons are continually generated due to **thermal Energy.** In average we get:

G = generation rate = recombination rate = R
$G = \beta(n\cdot p) = \beta\cdot n_i^2 = R$

for equilibrium: $n = n_0, \quad p = p_0$
for non-equilibrium: $n = n_0 + \Delta n, \quad p = p_o + \Delta p$
generation & recombination in pairs → $\Delta n = \Delta p = G_L\tau_{n,p}$

In steady state, the change in the Semiconductor conductivity is: $\Delta\sigma = q(\mu_n + \mu_p)G\tau_p$

Note: In steady state, non-equilibrium the carrier concentrations are constant.

LOW / HIGH – LEVEL INJECTION

low-level injection: $\Delta n \ll n_{n0}$

Number of carriers generated are small compared to the background doping density of the material. In other words: injected minority carriers concentration at the depletion region edge is less than the majority carrier concentration.

→ P-type: $n_p(x) \ll p_{p0}$
→ N-type: $p_n(x) \ll n_{n0}$
⇒ $n_n \approx N_D, p_p \approx N_A$

→ minority carrier recombination rates are linear

high-level injection: $\Delta n \gg n_{n0}$

number of carriers generated carriers is large compared to the background doping density of the material. In other words: injected minority carrier concentration exceeds the majority carrier concentration:

→ P-type: $n_p(x) > N_A$
→ N-type: $p_n(x) > N_D$

→ minority carrier recombination rates are proportional to the number of carriers squared

DIRECT RECOMBINATION

Direct recombination across the bandgap results in the **emission** of a **photon** of energy: $\boldsymbol{E_G = h\cdot\nu}$

Thermal Generation Rate: $G_{th} = R_{th} = \beta(n_0\cdot p_0) \quad [cm^{-3}s^{-1}]$
Recombination Rate: $R \;= \beta(n\cdot p) \quad [cm^{-3}s^{-1}]$

External Generation Rate: G_L
Total Generation: $G = G_L + G_{th}$

Net Generation Rate: N-Type $\frac{dp_n}{dt} = G - R = G_L + G_{th} - R$

in steady state, non-equilibrium $\left(\frac{dp_n}{dt} = 0\right)$ we find:

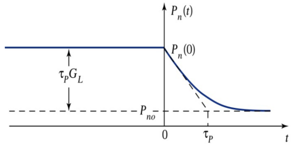
Net Recombination Rate: $U \equiv G_L = R - G_{th} \cong \frac{\Delta p}{1/(\beta n_{no})} = \frac{\Delta p}{\tau_p}$

→ N-Type $R_p = \frac{\Delta p_n}{\tau_p}$
→ P-Type $R_n = \frac{\Delta n_p}{\tau_n}$

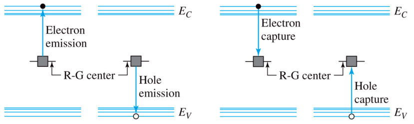
Minority Carrier Lifetime:

→ N-Type holes $\tau_p = \frac{1}{\beta n_{no}} \quad [s]$
→ P-Type electrons $\tau_n = \frac{1}{\beta p_{po}} \quad [s]$

The minority carrier lifetime describes how fast the excess carrier concentration decays back toward equilibrium, when excitation ends. Note that it is determined by **majority carrier concentration.**



INDIRECT RECOMBINATION



Recombination through a ‘G-R center’ aka ‘Trap’.

G-R Centers are most effective when their energy level $\boldsymbol{E_t}$ is **near** $\boldsymbol{E_i}$ of the bandgap. The capturing rate $\boldsymbol{U}$ is:

$$U \approx v_{th}\sigma_0 N_t \cdot \frac{\Delta p}{1 + \left(\frac{2n_i}{n_{n0}}\right)\underbrace{\cosh\left(\frac{E_t - E_i}{kT}\right)}_{=1 \text{ for } E_t = E_i}} = \frac{\Delta p}{\tau_p}$$

Minority carrier lifetime

$$\tau_p = \frac{1 + \left(\frac{2n_i}{n_{n0}}\right)\cosh\left(\frac{E_t - E_i}{kT}\right)}{v_{th}\sigma_0 N_t} \quad [s] \quad \xrightarrow{E_t \approx E_i} \approx \frac{1}{v_{th}\sigma_0 N_t} = \tau_n$$

Density of Recombination centers: N_t
Recombination center cross-section: σ

steady state equilibrium

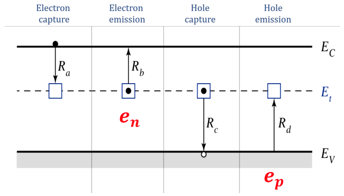


Abb: steady state equilibrium

- $\frac{1}{2}m_n v_{th}^2 = \frac{3}{2}kT$
- $f(E_t) = f = \frac{1}{1 + e^{(E_t - E_F)/kT}}$
- Emission Rate

electron: $e_n = \frac{v_{th}\sigma_n n(1-f)}{f} = v_{th}\sigma_n n_i e^{(E_t - E_i)/kT}$
hole: $e_p = v_{th}\sigma_p n_i e^{(E_i - E_t)/kT}$

Electron Capture Rate: $R_a = n\cdot N_t(1-f)\cdot v_{th}\cdot\sigma_n$
Electron Emission Rate: $R_b = e_n\cdot N_t\cdot f$

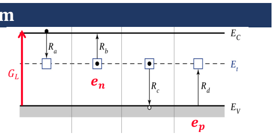
Hole Capture Rate: $R_c = p\cdot v_{th}\cdot\sigma_n\cdot N_t\cdot f$
Hole Emission Rate: $R_d = e_p\cdot N_t(1-f)$

in **steady state equilibrium:**
$R_a = R_b \quad \& \quad R_c = R_d$

steady state non-equilibrium

$$\frac{dn_n}{dt} = G_L - (R_a - R_b) = 0$$

$$\frac{dp_n}{dt} = G_L - (R_c - R_d) = 0$$



$$\Rightarrow G_L = R_a - R_b = R_c - R_d \equiv U$$

Direct vs. Indirect Recombination:

Direct and indirect recombination occur in **parallel**, as competitive mechanisms. Very often, one mechanism is faster and is characterized by **shorter recombination lifetime ⇒ mechanism is dominant.**

CARRIER TRANSPORT

Total carrier transport:

$$\begin{array}{lll} & \text{drift} & \text{diffusion} \\ \text{electrons:} & J_n = nq\mu_n\vec{E} & + \; qD_n\frac{dn(x)}{dx} \\ \text{holes:} & J_p = pq\mu_p\vec{E} & - \; qD_p\frac{dp(x)}{dx} \end{array} \quad \left[\frac{A}{cm^2}\right]$$

total current = drift current + diffusion current = electron +hole current

equilibrium: *(no net current)*
$J_n = 0, \quad J_p = 0$

DIFFUSION

A **concentration gradient in particle** and a random **thermal motion** (i.e. equal probability to move in any direction) leads to a diffusion of the particles.

thermal equilibrium: zero net current
average thermal velocity: $\frac{1}{2}mv_{th}^2 = \frac{3}{2}kT$

Fick’s First Law of Diffusion *(3D)*

$$J_{diff} = -D\cdot\nabla N = -D\left(\frac{\partial N}{\partial x}\vec{x}_u + \frac{\partial N}{\partial y}\vec{y}_u + \frac{\partial N}{\partial z}\vec{z}_u\right)$$

Diffusivity: D *(diffusion constant)*

Einstein relations:

electron: $D_n = \frac{kT}{q}\mu_n = V_t\,\mu_n \quad \left[\frac{cm^2}{s}\right]$
holes: $D_p = \frac{kT}{q}\mu_p = V_t\,\mu_p$

cannot be used for heavily doped semiconductors (Maxwell Boltzmann doesn’t hold)

Net Flux: $F = F_{Right} - F_{Left}$

$$\boxed{J_{diff\ n,p} = \begin{cases} -qF = \;\; qD_n\frac{dn(x)}{dx} & \text{for electron} \\ -qF = -qD_p\frac{dp(x)}{dx} & \text{for holes} \end{cases} \quad \left[\frac{A}{m^2}\right]}$$

Diffusion length:

for electrons: $L_n = \sqrt{D_n\tau_n}$
for holes $L_p = \sqrt{D_p\tau_p}$

DRIFT

Electrons move in the **opposite direction** of the E-Field.

$$\boxed{J_{drift\ n,p} = \begin{cases} -q\,n\,v_{dr,n} = q\,n\,\mu_n\,E & \text{for electrons} \\ \;\; q\,p\,v_{dr,p} = q\,p\,\mu_p\,E & \text{for holes} \end{cases}}$$

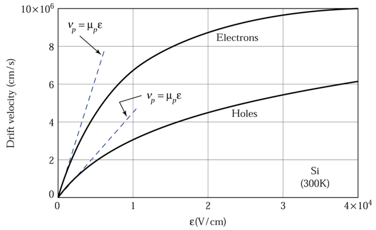
μ_n / μ_p ≔ electron / hole mobility $\left[\frac{cm^2}{V\cdot s}\right]$

drift velocity:

electrons: $v_{drift,n} = -\mu_n\,E \quad \left[\frac{cm}{s}\right]$
holes: $v_{drift,p} = \;\;\mu_p\,E$

Holes usually move slower than electrons $(\mu_n > \mu_p)$

total drift current: $j_{dr,tot} = j_{dr,n} + j_{dr,p} = \sigma E$
mobility/conductivity: $\sigma = n\,q\,\mu_n + p\,q\,\mu_p$



e^- velocity cannot increase indefinitely as in vacuum (saturation of velocity).

SIMPLIFICATIONS

Equilibrium

At equilibrium there is no net current!

$$\begin{cases} J_n = J_{dr,n} + J_{diff,n} = 0 \\ J_p = J_{dr,p} + J_{diff,p} = 0 \end{cases} \Leftrightarrow \begin{array}{l} E = \frac{kT}{q}\frac{1}{n(x)}\frac{dn(x)}{dx} \\ E = \frac{kT}{q}\frac{1}{p(x)}\frac{dp(x)}{dx} \end{array}$$

steady state

steady state → $\frac{dn}{dt} = \frac{dp}{dt} = 0$, no electrical field $(E = 0)$

for electrons *(as minority carriers)*: $D_n\frac{d^2n}{dx^2} + G_n - \frac{\Delta n}{\tau_n} = 0$

for holes *(as minority carriers)*: $D_p\frac{d^2p}{dx^2} + G_p - \frac{\Delta p}{\tau_p} = 0$

CONTINUITY EQUATION

The conservation of carriers results in the continuity equations, where G is the generation and R the recombination rate.

For electrons:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{dx} + (G_n - R_n) \qquad J_n = nq\mu_n\vec{E} + qD_n\frac{dn}{dx}$$

$$\frac{\partial n_p}{\partial t} = n_p\mu_n\frac{\partial \vec{E}}{dx} + \mu_n\vec{E}\frac{\partial n_p}{\partial x} + D_n\frac{\partial^2 n_p}{\partial x^2} + \left(G_n - \frac{n_p - n_{p0}}{\tau_n}\right)$$

For holes:

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{dx} + (G_p - R_p) \qquad J_p = pq\mu_p\vec{E} - qD_p\frac{dp}{dx}$$

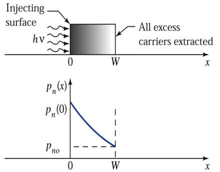
$$\frac{\partial p_n}{\partial t} = -p_n\mu_p\frac{\partial \vec{E}}{dx} - \mu_p\vec{E}\frac{\partial p_n}{\partial x} + D_p\frac{\partial^2 p_n}{\partial x^2} + \left(G_p - \frac{p_n - p_{n0}}{\tau_p}\right)$$

steady state

Boundary Conditions:
$p_n(0) = const, \quad p_n(W) = p_{n0}$

General solution:

$$p(x) = p_{n0} + (p_n(0) - p_{n0})\left[\frac{\sinh\left(\frac{W-x}{L_p}\right)}{\sinh\left(\frac{W}{L_p}\right)}\right]$$



For $W \to \infty, L_p \ll W$

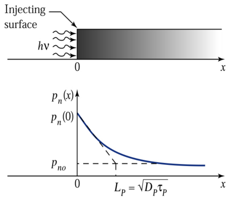
all quantities are time independent

$$\frac{\partial p_n}{\partial t} = 0 = D_p\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p}$$

Boundary Conditions:
$p_n(0) = const, \quad p_n(x \to \infty) = p_{n0}$

minority carrier diffusion length:

$$\boxed{L_p = \sqrt{D_p\tau_p}}$$

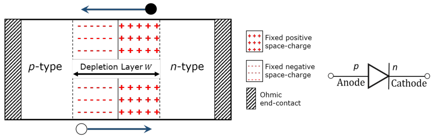


$$p_n(x) = p_{n0} + (p_n(0) - p_{n0})\exp\left(-\frac{x}{L_p}\right)$$

For W short, $L_p \gg W$, linearize → no recombination

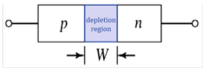
$$p_n(x) = p_{n0} + (p_n(0) - p_{n0})\left(1 - \frac{x}{W}\right)$$

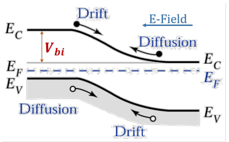
PN JUNCTION



equilibrium

The **fermi level** through a PN-Junction remains **constant (flat) at equilibrium.** The drift current will exactly oppose the diffusion current **(zero net current)** and therefore a Voltage (built in Voltage V_{bi}) is applied over the junction.





Minority carriers:

$$n_{p0} = \frac{n_i^2}{N_A} = N_D \cdot e^{-\frac{qV_{bi}}{kT}} \qquad p_{n0} = \frac{n_i^2}{N_D} = N_A \cdot e^{-\frac{qV_{bi}}{kT}}$$

$$n_{n0} = n_{p0}e^{\frac{qV_{bi}}{kT}} \qquad p_{p0} = p_{n0}e^{\frac{qV_{bi}}{kT}}$$

The electric filed points from the n-side to the p-side.

Built in Voltage V_{bi}

In general we know:

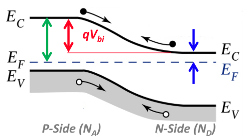
- $n_0 = N_C e^{-(E_C - E_F)/kT}$

From the graph follows:

- $(E_C - E_F)_P = (E_C - E_F)_N + qV_{bi}$
- $\Leftrightarrow e^{-\frac{(E_C-E_F)_P}{kT}} = e^{-\frac{(E_C-E_F)_N}{kT}} e^{-\frac{qV_{bi}}{kT}}$

Simply multiplying both sides by N_C gives us:

- $n_{p0} = n_{n0} \cdot e^{-\frac{qV_{bi}}{kT}}$
- $\Leftrightarrow \frac{n_i^2}{N_A} = \mathrm{N_D} \cdot e^{-\frac{qV_{bi}}{kT}}$



Reordering the terms we get:

$$\boxed{V_{bi} = \frac{kT}{q}\ln\left(\frac{N_A N_D}{n_i^2}\right) = \frac{kT}{q}\ln\left(\frac{p_{p0}}{p_{n0}}\right) = \frac{kT}{q}\ln\left(\frac{n_{n0}}{n_{p0}}\right)}$$

Remarks:
- The built in Voltage V_{bi} only depends only on the doping level at the depletion region edge (magnitude of $N_A \& N_D$)
- Voltmeter cannot measure the built in voltage because in order to measure it, it needs to take some current from the circuit (measures small current over a high series impedance). But at equilibrium, there is no current, so no measurement possible.

QUASI FERMI LEVEL

Under bias (e.g. illumination), the equilibrium fermi level splits into **2 distinct „Quasi Fermi Levels"** in each region of the diode and the np-product is in-/decreased. This is caused by a slow recombination rate.

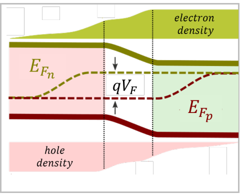
$$\boldsymbol{n(x) = N_C e^{-(E_C(x) - E_{F_n})/kT}}$$

$$\boldsymbol{p(x) = N_V e^{-(E_{F_p} - E_V(x))/kT}}$$

$$\boldsymbol{n(x)p(x) = n_i^2 e^{(E_{F_n} - E_{F_p})/kT}}$$

$$\boldsymbol{= n_i^2 e^{qV_F/kT}}$$

$$\boldsymbol{= N_C N_V e^{(E_{F_n} - E_{F_p})/kT}}$$



SHOCKLEY BOUNDARY CONDITIONS

Minority Carrier Concentration:

$$n_p(-x_p) = N_D \cdot e^{-\frac{q(V_{bi}-V_F)}{kT}} = n_{p0} \cdot e^{\frac{qV_F}{kT}} = n_{p0} + \Delta n_p$$

$$p_n(x_n) = N_A \cdot e^{-\frac{q(V_{bi}-V_F)}{kT}} = p_{n0} \cdot e^{\frac{qV_F}{kT}} = p_{n0} + \Delta p_n$$

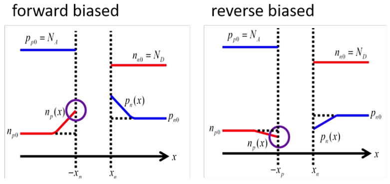
Remember: n_{p0}, p_{n0} are the minority carrier concentrations $\Rightarrow n_{p0} = \frac{n_i^2}{N_A}, p_{n0} = \frac{n_i^2}{N_D}$

$$\frac{dn}{dx} = \frac{n_{p0}(e^{(qV_F)/kT} - 1)}{L_n} \qquad \frac{dp}{dx} = \frac{p_{n0}(e^{(qV_F)/kT} - 1)}{L_p}$$

applied voltage:

$$V_F = -V_R = \frac{kT}{q}\ln\left(\frac{n_p(-x_p)}{n_{p0}}\right) = \frac{kT}{q}\ln\left(\frac{p_n(x_n)}{p_{n0}}\right)$$

Plotted on an x-axis log scale:



Forward bias → minority carrier injection
Reverse bias →minority carrier extraction

ELECTROSTATICS – PN JUNCTION

1-D Poisson-equation: $\frac{dE}{dx} = \frac{\rho}{\varepsilon_r\varepsilon_0} = \frac{\rho}{\varepsilon_s}$

Electric field: $E(x) = \int_x^{x_n}\frac{\rho(y)}{\varepsilon_s}dy \quad$ for $\quad 0 < x < x_n$

charge density: $\rho = \begin{cases} -q \cdot N_A & \text{P'region} \\ q \cdot N_D & \text{N'region} \end{cases}$

***p*-region** $-x_p \le x \le 0$	***n*-region** $0 \le x \le +x_n$
$\frac{dE}{dx} = \frac{\rho}{\varepsilon_s} = -\frac{qN_A}{\varepsilon_s}$	$\frac{dE}{dx} = \frac{\rho}{\varepsilon_s} = \frac{qN_D}{\varepsilon_s}$
$E(x) = \int_{-x_p}^{x}\frac{\rho(y)}{\varepsilon_s}dy$ $= -\frac{qN_A(x + x_p)}{\varepsilon_s}$	$E(x) = \int_x^{x_n}\frac{\rho(y)}{\varepsilon_s}dy$ $= \frac{qN_D(x - x_n)}{\varepsilon_s}$

$$|E_{max}| = |E(x = 0)| = \frac{qN_A(x_p)}{\varepsilon_s} = \frac{qN_D(x_n)}{\varepsilon_s}$$

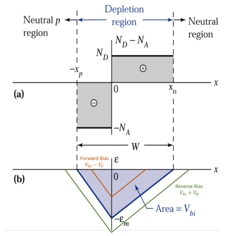
S5a4: If we are at equilibrium then: $nq\mu_n\vec{E} = -qD_n\frac{dn(x)}{dx}$

Or even $\vec{E} = -\frac{dV}{dx} = -\frac{1}{q}\frac{d(E_V - E_F)}{dt}$

Charge Neutrality:

$$\int_{-x_p}^{0} N_A(x)dx = \int_0^{x_n} N_D(x)dx$$

$N_A x_p = N_D x_n$ (const. doping)



Note: If you increase the doping level then:
→ W decreases
→ E_{max}increases

$$V_{bi} = \int_{-x_p}^{x_n} E(x)\,dx$$

$$= (\Phi_{xn} - \Phi_{xp})$$

Note that :

$$V_{bi} = \frac{qN_A(x_p)^2}{2\varepsilon_s} + \frac{qN_D(x_n)^2}{2\varepsilon_s} = \frac{1}{2}E_{max}W \Rightarrow \boxed{E_{max} = \frac{2V_{bi}}{W}}$$

Depletion Width:

$$\boxed{\boldsymbol{W = x_p + x_n = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)V_{bi}}}}$$

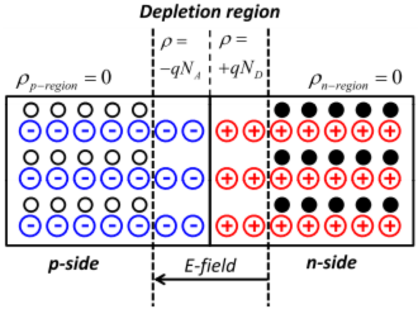
for reverse bias: $V_{bi} \mapsto V_{bi} + V_R$, for forward bias: $V_{bi} \mapsto V_{bi} - V_F$

$$x_n = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{N_A}{N_D(N_A + N_D)}\right)V_{bi}} \qquad x_p = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{N_D}{N_A(N_A + N_D)}\right)V_{bi}}$$

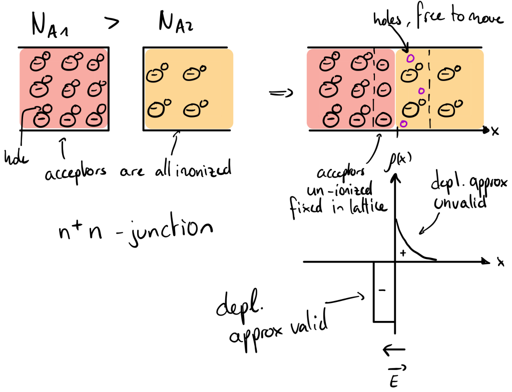
Due to the reverse bias the area is now increased to $V_{bi} + V_R$, accordingly the depletion length increases.
For high doping levels, W is very narrow.

DEPLETION APPROXIMATION

The depletion approximation is the fact that we can approximate the charge densities as being "box-like". This approximation is valid in the depletion regions where the acceptors/donors are uncovered (un-ionized).
The approximation is usually valid if both sides of the junction are of different types



Example of isotype junction:



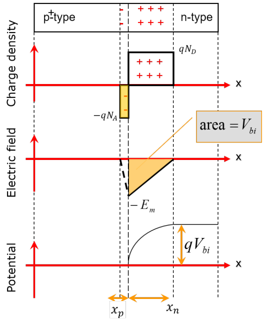
ONE SIDED JUNCTION

Remark: **high-level injection**
→ n^+p: $n_p(x) > N_A$
→ p^+n: $p_n(x) > N_D$

For a one sided junction the **lightly doped side** determines the depletion length W.

***p^+n* junction:**

If $N_A \gg N_D$ then we call the junction p^+n.
→ $x_p \ll x_n \approx W$
→ $W \approx \sqrt{\frac{2\varepsilon_s}{qN_D}(V_{bi} + V_R)}$
→ $E_{max} = \frac{qN_D}{\varepsilon_s}W$



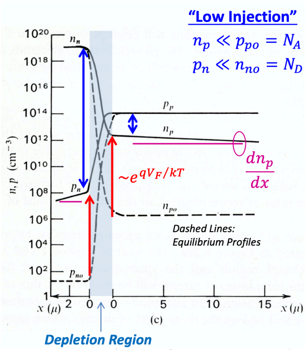
***n^+p* junction:**

$N_A \ll N_D$
→ $W \approx \sqrt{\frac{2\varepsilon_s}{qN_A}(V_{bi} + V_R)}$

CURRENT IN THE PN-JUNCTION (LONG DIODE)

We will now try to understand how the current is generated inside a PN-Junction, and derive it's IV-characteristics, which are the ones of a diode.

We assume a long diode (i.e. undepleted regions are much larger than L$_{p,}$: $L \ll W$). There is zero field in the undepleted regions so only drift current in those regions. Due to minority carrier injection, there is a minority carrier gradient, so a diffusion current. In the following, we look at the n-side.



The change in minority carrier is

$$p_n(x > x_n) = p_{n0} + \Delta p_n e^{-\frac{x-x_n}{Lp}}$$
$$= p_{n0} + p_{n0}\left(e^{\frac{qV_F}{kT}} - 1\right)e^{-\frac{x-x_n}{Lp}}$$

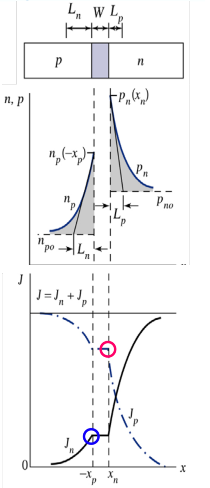
Where we inserted the Shockley boundary condition for $\Delta p_n = p_n(x_n) - p_{n0}$. Using the formula for diffusion current in x_n, we get:

○ $J_p(x_n) \quad = -qD_p\left[\frac{dp_n}{dx}\right]_{x_n}$

$$= -qD_p\left[p_{n0}\left(e^{\frac{qV_F}{kT}} - 1\right)e^{-\frac{x-x_n}{Lp}} \cdot \left(-\frac{1}{L_p}\right)\right]_{x_n} = \frac{qD_p p_{n0}}{L_p}\left(e^{\frac{qV_F}{kT}} - 1\right)$$

We can add up both contributions to get the total current.

$\Rightarrow J = J_n + J_p$

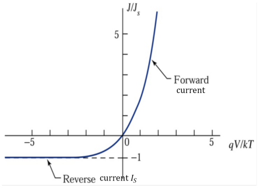


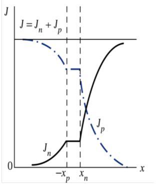
$$\boxed{\mathbf{J} = qD_n\frac{dn}{dx} - qD_p\frac{dp}{dx} = \underbrace{\left[\frac{qD_n n_{p0}}{L_n} + \frac{qD_p p_{n0}}{L_p}\right]}_{J_S = Saturation\ Current} \cdot \left(e^{\left(\frac{qV_F}{kT}\right)} - 1\right)}$$

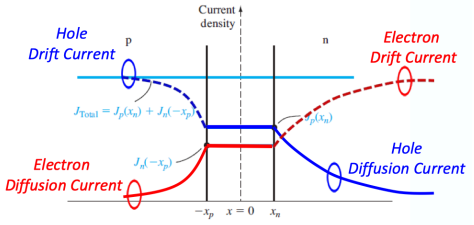
$$\boxed{\Rightarrow I = I_S\left(e^{\left(\frac{qV_F}{kT}\right)} - 1\right) \approx I_S e^{\left(\frac{qV_F}{kT}\right)}}$$

If we use the approximation we neglect the tiny reverse saturation current.

The current $J = J_n + J_p$ has to be constant throughout the whole depletion region. In the depletion region we have to consider the diffusion and the drift current whereas we only have to consider the diffusion current outside the depletion region.







SHORT DIODE (FORWARD BIAS)

The Diode is shorter than the diffusions length ($L \gg W$), and since the boundary condition must be fulfilled, it forces the charge density to equilibrium at the end of the Diode ($= W$) we get $p_n(W) = p_{n0}$.

This means we have a linear decay in minority carriers.

$$p_n(x) = p_{n0} + (p_n(x_n) - p_{n0})\left(1 - \frac{x - x_n}{W}\right)\ for\ x > x_n$$

This means that the respective contributions in diffusion current of the minority and majority carriers stay **constant throughout the non-depleted region**! In opposition to the long diode case, where the minority carrier diffusion current would decay exponentially, reciprocally to the majority carrier diffusion current.

With help of the continuity equation in steady state we find:

$$J_{diff,p} \quad = -qD_p\left[\frac{dp_n}{dx}\right]_{x=W_n} = q[p_n(0) - p_{n0}]\frac{D_p}{L_p}\frac{\cosh\left(\frac{W_n - x}{Lp}\right)}{\sinh\left(\frac{W_n}{Lp}\right)}$$

Weak Recombination Limit:

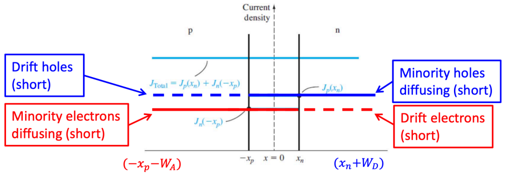
$$W_n/L_p \ll\ 1 \Rightarrow \sinh\left(\frac{W_n}{L_p}\right) \approx W_n/L_p$$
$$\Rightarrow J_{diff,p}\big|_{x=W_n} \ = \ q \cdot p_{n0}\frac{D_p}{W_n}\left(e^{\frac{qV_F}{kT}} - 1\right)$$

analogously for $J_{diff,n}$

we get in total:

$$\boxed{\Rightarrow J = \left(\frac{qD_p p_{n0}}{W_n} + \frac{qD_n n_{p0}}{W_p}\right)\left(e^{\frac{qV_F}{kT}} - 1\right) = J_s(e^{\frac{qV_F}{kT}} - 1)}$$

J_s is increased compared to the long diode (W << L).



forward bias $V_F = \mathrm{E_{Fn}} - \mathrm{E_{Fp}} > 0$

A forward bias corresponds to connecting the positive terminal to the anode (*p*-type region) and negative terminal to the cathode (*n*-type region)

Applying a forward bias V_F **reduces** band bending and **reduces** the recombination length.

→ Bias diminishes the electric field.

→ "-" terminal pushes the electrons from the p-side to the n-side. "+" terminal attracts those electrons to the p-side.

→ injection of minority carriers ($n_p > n_{p0}, p_n > p_{n0}$)

→ excess in minority carrier concentrations at the depletion region edge

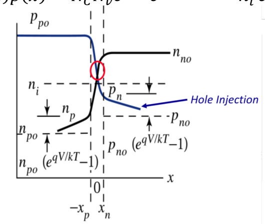
The voltage across the diode is reduced:

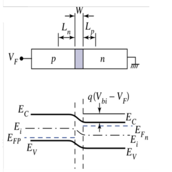
replace $V_{bi} \mapsto V_{bi} - V_F$

Shockley bound. cond.: $n_p = n_{n0}e^{-q(V_{bi} - V_F)/kT} = n_{p0}e^{\frac{qV_F}{kT}}$

$$p_n = p_{n0}e^{\frac{qV_F}{kT}}$$

$$n(x)p(x) = N_C N_v e^{-\frac{E_g}{kT}} e^{\frac{E_{Fn} - E_{Fp}}{kT}} = n_i^2 e^{\frac{E_{Fn} - E_{Fp}}{kT}} = n_i^2 e^{q\frac{V_F}{kT}} \geq n_i^2$$





Recombination in depletion Region

$$U \quad = \frac{v_{th}\sigma_0 N_t n_i^2(e^{qV_F/kT} - 1)}{p_n + n_n + 2n_i\cosh\left(\frac{E_t - E_i}{kT}\right)} \qquad \text{Note: } \cosh(\sim 0) \approx 1$$

$$p_n n_n = n_i^2 e^{qV_F/kT}$$

$$U_{max} \quad = \frac{v_{th}\sigma_0 N_t n_i^2(e^{qV_F/kT} - 1)}{2n_i(e^{qV_F/2kT} + 1)} = \frac{1}{2}v_{th}\sigma_0 N_t n_i e^{qV_F/2kT}$$

U_{max} for $p_n = n_n = n_i e^{\frac{qV_F}{2kT}}, V_F > 3\ kT/q$

$$J_{rec} = \int_0^W qUdx \cong \frac{qW}{2}v_{th}\sigma_0 N_t n_i e^{\frac{qV_F}{kT}} = \frac{qWn_i}{2\tau_r}e^{qV_F/2kT}$$

Total forward current:

$$\boxed{\begin{aligned} J_{FT} &= J_S\left(e^{\frac{qV_F}{kT}} - 1\right) + J_{rec} \\ &= \left[\frac{qD_n}{N_A L_n} + \frac{qD_p}{N_D L_p}\right] n_i^2(e^{\frac{qV_F}{kT}} - 1) + \frac{qWn_i}{2\tau_r}e^{\frac{qV_F}{2kT}} \\ &\quad \text{ideal} \qquad\qquad\qquad\qquad \text{recombination forward current}\end{aligned}}$$

Ideal current increases more rapidly than the recombination current and eventually dominates.

reverse bias $V_R = \mathrm{E_{Fp}} - \mathrm{E_{Fn}} > 0$

A reverse bias corresponds to connecting the positive terminal to the cathode (*n*-type region) and the negative terminal to the anode(*p*-type region).

Applying a reverse bias V_R **increases** band banding and **increases** the recombination length

→ Bias increases the electric filed.

→ "-" terminal repels the electrons to the n-side.

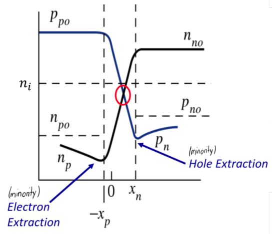
→ deficit in minority carrier concentrations at the depletion region edge ***(carrier extraction)*** ($n_p < n_{p0}, p_n < p_{n0}$)

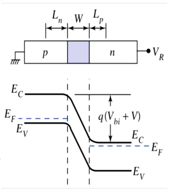
The voltage across the diode is increased:

replace $V_{bi} \mapsto V_{bi} + V_R$

E.g.: $n_p = n_{n0}e^{-q(V_{bi} + V_R)/kT} = n_{p0}e^{-\frac{qV_R}{kT}}$

$$n(x)p(x) = n_i^2 e^{-q\frac{V_R}{kT}} \leq n_i^2$$





Carrier concentrations in the depletion region are lowered with respect to equilibrium.

Generation in depletion Region

Under reverse bias:$n(x)p(x) < n_i^2$. Since the semiconductor, will always try to restore equilibrium, for a carrier deficit, generation takes place. So electron holes pairs are generated int the depletion region, which gives rise to a "generation current", that adds to J$_S$.

To approximate the total generation, we integrate the maximum generation over W.

Generation Rate:

$$G = -U \ = \frac{v_{th}\sigma_n\sigma_p N_t n_i^2}{\sigma_p\left[n_i e^{(E_i - E_t)/kT}\right] + \sigma_n\left[n_i e^{(E_t - E_i)/kT}\right]}$$
$$= \frac{v_{th}\sigma_0 N_t n_i}{\left[e^{(E_i - E_t)/kT}\right] + \left[e^{(E_t - E_i)/kT}\right]}$$
$$= \frac{v_{th}\sigma_0 N_t n_i}{2\cosh\left(\frac{E_t - E_i}{kT}\right)} = \frac{n_i}{\tau_g} \qquad \text{Note: } \cosh(\sim 0) \approx 1$$

$$J_{gen} = \int_0^W qGdx \cong \frac{qn_i}{\tau_g}W$$

Total reverse current:

$$\boxed{\begin{aligned} J_{RT} = J_S + J_{gen} &= \left[\frac{qD_n}{N_A L_n} + \frac{qD_p}{N_D L_p}\right] n_i^2 + \frac{qWn_i}{\tau_g} \\ &\quad \text{ideal} \qquad\qquad \text{recombination reverse current}\end{aligned}}$$

Remarks:

W increases with the square root of $V_R + V_{bi}$

Narrower bandgap materials have high n_i and $\boldsymbol{J_S}$ will dominate.

Wider bandgap materials have small n_i so $\boldsymbol{J_{gen}}$ might dominate.

This is counterintuitive to the fact that smaller bandgap means easier generation, but we look at n$_i$ not E$_G$.

DIODE NON-IDEALITIES

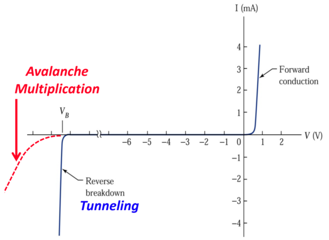
We have seen forward bias recombination, reverse bias generation. We now see 2 breakdown mechanisms in reverse bias.

Band-to-Band Tunneling (Zener)

The high reverse bias increases the electric field such that the electrons **tunnel** (quantum mechanically) **across the bandgap** and thus **increase** the **current exponentially.** The current at which it happens is the breakdown voltage V_B. B2B tunnelling dominates (occurs for smaller V$_B$) if **both sides are heavily doped** and when the **bandgap is relatively small. => depletion width very thin**

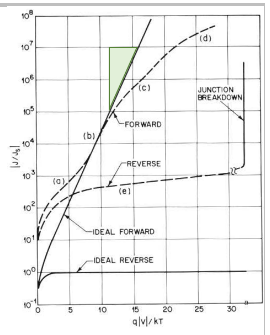
Avalanche Multiplications/Impact ionization

If the electric field (reverse bias V_R) becomes high enough that carriers acquire enough **kinetic energy** to break covalent bonds in the depletion layer, they **generate new electron-hole pairs** via collisions, thus increasing the current rapidly. Since large W and small E_G implies a high probability of collision and generation of e/h pairs, for high impact ionization, we want a **thick depletion region W**, this happens for **lower doping levels** and a **narrow bandgap.**



We can **summarize** all behaviours:

a) Recombination in depletion region
b) ideal injection ($\eta = 1;\ 60\frac{mV}{dec}\ slope$)
c) high-level injection, (minority carriers approach majority concentration, $\eta = 2$)
d) series resistance effects ($\Delta V_F = r_s \cdot IF$) where ΔV_F is the deviation from the ideal characteristic
e) generation in depletion region
f) Junction breakdown mechanisms



$$J \approx J_s e^{\frac{qV_F}{kT}} \qquad for\ V_F > 3kT$$
$$J \approx -J_s \qquad for\ V_F < -3kT$$

An **ideal diode characteristic** has a **slope of 60 mV/dec**

For non idealities we consider the ideality factor η

$$\Rightarrow J = J_s\left(e^{\left(\frac{qV_F}{\eta kT}\right)} - 1\right)$$

To calculate η: $\eta = \frac{g}{60}$ where g is the gradient in $\frac{mV}{dec}$

Or, (see s8a3a) take 2 points $(J_1; V_{F1}), (J_2; V_{F2})$

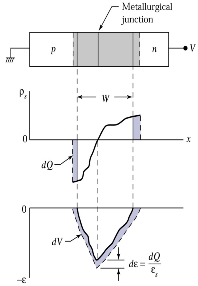
$$\frac{J_1}{J_2} = e^{\left(\frac{q(V_{F1}-V_{F2})}{\eta kT}\right)} \Leftrightarrow \eta = \frac{V_{F1} - V_{F2}}{\frac{kT}{q}\ln\left(\frac{I_{F1}}{I_{F2}}\right)}$$

DEPLETION/JUNCTION CAPACITANCE

$$\boldsymbol{C_j} \equiv \frac{\mathbf{dQ}}{\mathbf{dV}} = \frac{\mathrm{dQ}}{\mathrm{W}\frac{\mathrm{dQ}}{\epsilon_s}} = \frac{\epsilon_s}{W}$$

$$\frac{1}{C_j^2} = \frac{1}{\epsilon_s^2}\frac{2\epsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(V_{bi} - V_F)$$

Assuming the doping levels are constant this capacitance can be used to determine the built in voltage. By simply measuring, with a capacitance meter, the voltage at which $\frac{1}{c_j^2} = 0$.



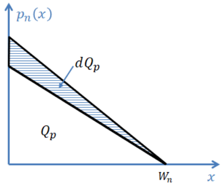
DIFFUSION CAPACITANCE

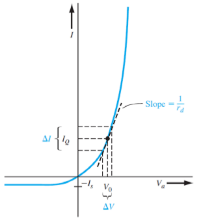
Given a 1-sided short diode. For a small decrease in V_F we have an excess charge dQ_p.

$$C_d = \frac{dQ_p}{dV_F} = \frac{dQ_p}{dJ}\frac{dJ}{dV_F} =$$
$$\tau \cdot \frac{1}{r_d} = \frac{W_n^2}{2D_p}\frac{J}{\frac{kT}{q}} \quad \left[\frac{F}{cm^2}\right]$$

Where we used:

$$J = \left(\frac{qD_p p_{n0}}{W_n}\right)\left(e^{\frac{qV_F}{kT}} - 1\right)$$
$$Q_p = \frac{W_n}{2}\cdot\frac{JW_n}{D_p}$$
$$\tau := t\ to\ drain\ Q_p\ with\ J$$
$$\frac{1}{r_d} := ss\ conductance$$

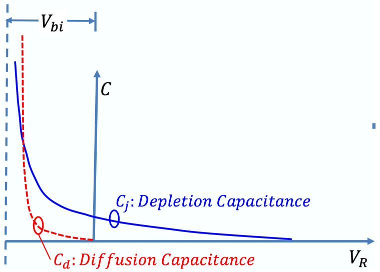




JUNCTION VS DIFFUSION CAPACITANCE

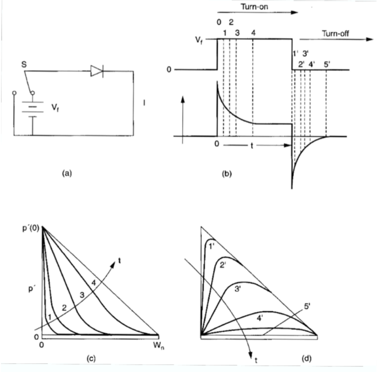
-The junction (depletion) capacitance C_j dominates in reverse bias. It would become infinite for a forward bias of V_{bi}, but the depletion approximation model fails for strong forward bias.

-In forward bias, the diffusion capacitance C_d due to minority carrier charge storage eventually becomes dominant: it is proportional to current and grows exponentially (faster than the power law of C_j).



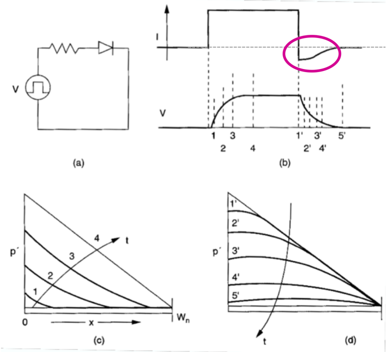
DIODE CHARGE STORAGE (SWITCHING)

Zero resistance: Shockley boundary conditions appear directly at the edges of the depletion region.(c) and d)). Remember: I (see b)) is proportional to the derivative of the carrier densities.



Turn-on   Turn-off

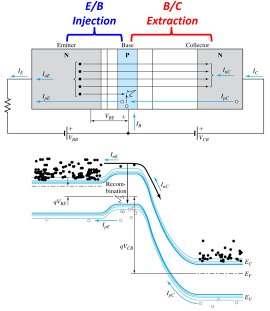
Non-zero resistance: Now we have an RC pair so instantaneous change in voltage across the diode is not possible. So Shockley boundary conditions aren't instantaneous(c) and d)).



Turn-on   Turn-off

BJT (BIPOLAR JUNCTION TRANSISTOR)

Operating Principle for normal active mode



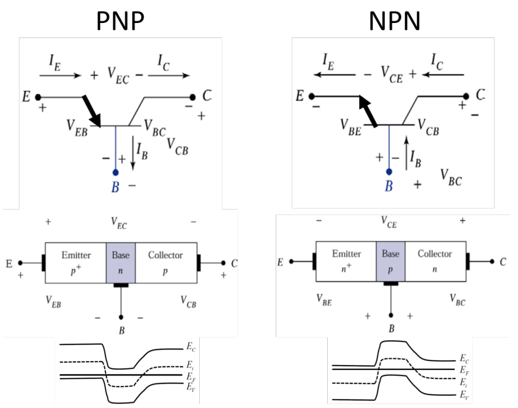
Emitter/Base forward bias injects minority carriers in the base.

In the base recombination may occur. The consumed e/h pairs are replaced by the base contact.

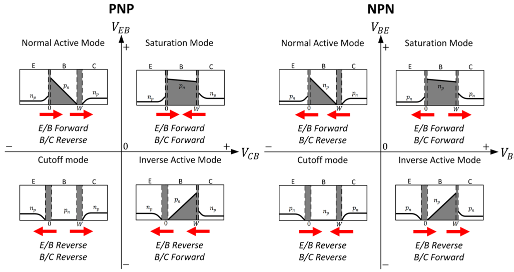
The electrons which did not recombine are extracted by the reverse-biased Base/ Collector junction. Ideally this current is independent of V_{CB}

Note: The BJT is a minority carrier device

Q: Why doesn't the current flow out in the base contact?
A: The base layer is thin so carriers will easily pass thought it. Minority carriers are extracted by the B/C junction, because they are driven by their gradient caused by the reverse bias. But if the collector was open circuit, all the current would flow through the base contact.
Q: Why do we need 2 types of BJTs?
A: Combined together, they provide circuit design flexibility. NPN are faster (larger bandwidth) because they rely on electron transport (higher mobility and drift velocity than h).



Operation Modes



When the minority carrier density is higher (bending up) in the Base @ the E/B or B/C then it is forward biased, if it is smaller (bending down), then it is reversed.

NPN Modes:

Normal Active Mode:

Current flows from collector to emitter. The transistor acts as a voltage controlled current source $I_C(V_{BE})$. The collector current I_C is **prop. to** the base current $I_B = \frac{I_C}{\beta}$.

Emitter injects e^- into the base which sucked off by the collector. The V_{BE} controls the number of injection e^-.

Inverse Active Mode:

Like Normal Active Mode, but current flows from emitter to collector. The gains (α & β) are much smaller.

Emitter and Collector change roles (e^- are injected via the collector)

Saturation Mode:

The transistor acts like a **short circuit** ⇔ **On Mode**. Current flows almost freely from Collector to Emitter

Base is flooded with e^- from both sides and the current cannot be controlled by V_{EB} any longer. The current is the max current of the normal active mode.

Cutoff Mode:

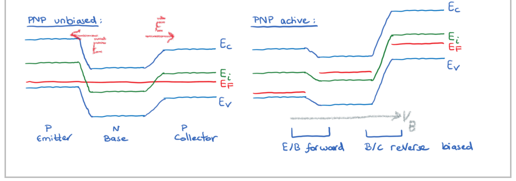
The transistor acts like an **open circuit** ⇔ **Off Mode**. No current flows from Collector to Emitter.

Because of the reverse bias over Base/Emitter junction no e^- will be injected into the base. No current can flow.

For PNP BJT the Emitter and Collector change roles respectively the Emitter injects holes instead of e^- → current direction changes.

drawing band diagrams in modes
1. draw the equilibrium (unbiased) band diagram

2. If the biased voltage is in the same direction as the electric field of the unbiased BJT, then the potential difference grows, if the biased voltages is in the opposite direction than the unbiased BJT, then the potential difference is reduced.



Shockley Boundary Condition – Carrier Concentration

At the end of the depletion region it holds that:

$n_p(-x_p) = n_{p0}\exp\left(\frac{qV_F}{kT}\right)$ & $p_n(x_n) = p_{n0}\exp\left(\frac{qV_F}{kT}\right)$

The normed carrier concentration thus has to be equal on both sides of the depletion region (for E/B & B/C)

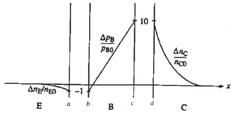
$$\frac{n_p(-x_p)}{n_{p0}} = \frac{p_n(x_n)}{p_{n0}} = \exp\left(\frac{qV_F}{kT}\right)$$

We can rewrite it in form of a change:

$n_p(-x_p) = n_{p0} + \Delta n \Rightarrow \Delta n = n_{p0}\left(\exp\left(\frac{qV_F}{kT}\right) - 1\right)$

$\Rightarrow \frac{\Delta n}{n_{p0}} = \left(\exp\left(\frac{qV_F}{kT}\right) - 1\right)$ & analog : $\frac{\Delta p}{p_{n0}} = \left(\exp\left(\frac{qV_F}{kT}\right) - 1\right)$

NPN – Inverse Active Mode

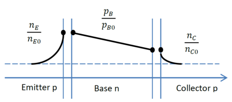


normed concentration change:

$\frac{\Delta n_E}{n_{E0}} = \frac{n_E(a) - n_{E0}}{n_{E0}} = e^{\left(\frac{qV_{EB}}{kT}\right)} - 1$

$\frac{\Delta n_C}{n_{C0}} = e^{\left(\frac{qV_{CB}}{kT}\right)} - 1$

PNP – Saturation Mode

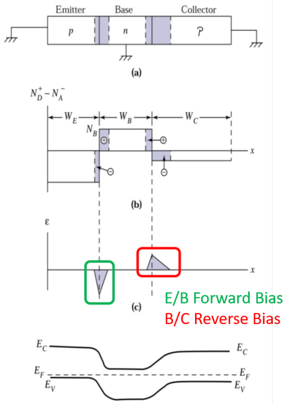


IDEAL BJT

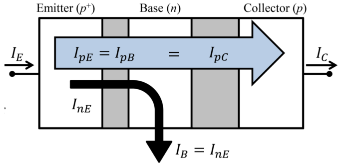
Assumptions:

- No Generation/Recombination in the Base Layer
- no B/C junction reverse leakage
- Shockley Boundary conditions (injection from Emitter to Base and from Base into Emitter)
- I_c doesn't depend on V_{BC}

ideal PNP BJT | Equilibrium:



PNP (no recombination in Base)



There are two currents present: a hole current and an electron current.

$I_E = I_{pE} + I_{nE} = I_C + I_B = \frac{1}{1-\alpha} I_B$ $\quad I_C = I_{pC} + I_{nC} = \alpha I_E = \beta I_B$

Currents ($V_{BC} = 0$)

$p_{n0} = n_i^2/N_{D,B}$ & $n_{E0} = n_i^2/N_{A,E}$

$$J_{pB} = \frac{qD_{pB}}{W}\frac{dp_{nB}}{dx} = \frac{qD_{pB}}{W} p_{n0}\left(e^{qV_{EB}/kT} - 1\right) = J_{pC} = J_{pE}$$

$$J_{nE} = \frac{qD_{nE}}{L_{nE}}\frac{dn_E}{dx} = \frac{qDn_E}{L_{nE}} n_{E0}\left(e^{qV_{EB}/kT} - 1\right)$$

$$I_C = A_E \cdot J_{pC} = I_S\left(e^{qV_{EB}/kT} - 1\right)$$

$$I_B = A_E \cdot J_{nE} = \frac{I_S}{\beta}\left(e^{\frac{qV_{EB}}{kT}} - 1\right)$$

Common Emitter Current Gain (forward)

$\beta_{fwd} = \frac{I_C}{I_B} = \frac{I_{pC}}{I_{nE}} = \frac{I_{pE}}{I_{nE}}$

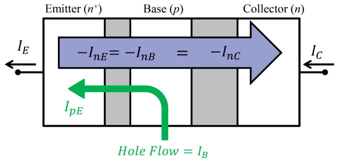
we assume that all junctions have equal areas and can therefore write:
with: $p_{n0} = n_i^2/N_{D,B}$ & $n_{E0} = n_i^2/N_{A,E}$

$\beta_{fwd} = \frac{J_{pE}}{J_{nE}} = \frac{D_{pB}}{D_{nE}}\frac{L_{nE}}{W_B}\frac{N_{A,E}}{N_{D,B}}$

For a useful BJT we want a high β and therefore $N_{A,E} \gg N_{D,B}$

Please note, that this only holds for forward active mode. In the same manner we can derive the common emitter current gain for the inverse active mode $\beta_{rev} = \frac{I_E}{I_B}$ (S7.4)

NPN (no recombination in Base)



$I_E = I_{pE} + I_{nE} = I_C + I_B$

Currents ($V_{BC} = 0$)

$$J_B = J_{pE} = -qD_{pE}\frac{dp_E}{dx} = -\frac{qD_{pE}}{L_{pE}} p_{E0}\left(e^{qV_{BE}/kT} - 1\right)$$

$$J_{nC} = J_{nB} = qD_{nB}\frac{dn_B}{dx} = -\frac{qD_{nB}}{W_B} n_{B0}\left(e^{qV_{BE}/kT} - 1\right)$$

PROPERTIES OF BJT

NPN	PNP

Emitter Efficiency

$\gamma := \frac{I_{E_n}}{I_{E_p}+I_{E_n}} = \frac{I_{E_n}}{I_E}$	$\gamma := \frac{I_{E_p}}{I_{E_p}+I_{E_n}} = \frac{I_{E_p}}{I_E}$

Common Emitter Current Gain

$$\beta := \frac{I_C}{I_B} = \frac{I_C}{I_E - I_C} = \frac{\alpha I_E}{I_E - \alpha I_E} = \frac{\alpha}{1-\alpha}$$

ideal:	$\beta_{fwd} = \frac{I_{nC}}{I_{pE}} = \frac{I_{nE}}{I_{pE}}$ $= \frac{D_{nB}}{D_{pE}}\frac{L_{pE}}{W_B}\frac{N_{DE}}{N_{AB}}$	$\beta_{fwd} = \frac{I_{pC}}{I_{nE}} = \frac{I_{pE}}{I_{nE}}$ $= \frac{D_{pB}}{D_{nE}}\frac{L_{nE}}{W_B}\frac{N_{A,E}}{N_{D,B}}$

$\beta_{rev} = \frac{D_{nB}}{D_{pC}}\frac{L_{pC}}{W_B}\frac{N_{DC}}{N_{AB}}$	$\beta_{rev} = \frac{D_{pB}}{D_{nE}}\frac{L_{nE}}{W_B}\frac{N_{AC}}{N_{DB}}$

Common Base Current Gain

$$\alpha := \frac{I_C}{I_E} = \frac{\beta}{1+\beta}$$

$\alpha \approx \frac{I_{nC}}{I_E} = \gamma \cdot \alpha_T$	$\alpha \approx \frac{I_{Cp}}{I_E} = \gamma \cdot \alpha_T$

Base Transport Factor

Fraction of carriers that succeed in crossing the base.
If the base thickness is much smaller than the base recombination length, then it holds that: $\alpha_T = 1$ because there is no recombination in the base ($I_{BB} = 0$).

$\alpha_T = \frac{I_{nC}}{I_{nE}}$	$\alpha_T = \frac{I_{pC}}{I_{pE}}$

Transconductance

For a voltage driven current source, the gain is defined as a transconductance

$$g_M = \frac{dI_C}{dV_{EB}} = I_S e^{qV_{EB}/kT} \cdot \frac{q}{kT} = \frac{I_C}{kT/q}$$

neutral (undepleted) base width $W_{neutral}$

The neutral base width is the difference between total base width and the depletion region in base resulting from both junctions.

$$W = x_p + x_n = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(V_{bi} + V_R)}, \qquad N_A x_p = N_D x_n$$

For **NPN** it is the p-side extend:

$$W = x_p + x_n = x_p\left(1 + \frac{N_A}{N_D}\right) \Leftrightarrow x_p = \frac{N_D}{N_A + N_D} W$$

we calculate x_p for both junctions: $x_{p_{EB}}, x_{p_{BC}}$

$\Rightarrow W_{neutral} - x_{p_{EB}} - x_{p_{BC}}$

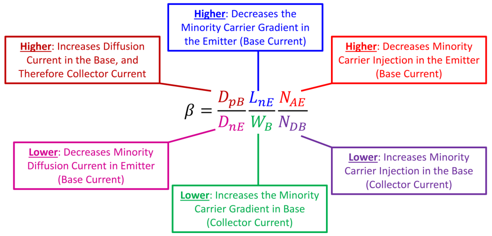
For **PNP** it is the n-side extend:

$$W = x_p + x_n = x_n\left(1 + \frac{N_D}{N_A}\right) \Leftrightarrow x_n = \frac{N_A}{N_A + N_D} W$$

we calculate x_n for both junctions: $x_{n_{EB}}, x_{n_{BC}}$

$\Rightarrow W_{neutral} - x_{n_{EB}} - x_{n_{BC}}$

Gain mechanism: (B|7.15)



$$\beta = \frac{D_{pB}}{D_{nE}}\frac{L_{nE}}{W_B}\frac{N_{AE}}{N_{DB}}$$

Higher: Decreases the Minority Carrier Gradient in the Emitter (Base Current)

Higher: Increases Diffusion Current in the Base, and Therefore Collector Current

Higher: Decreases Minority Carrier Injection in the Emitter (Base Current)

Lower: Decreases Minority Diffusion Current in Emitter (Base Current)

Lower: Increases Minority Carrier Injection in the Base (Collector Current)

Lower: Increases the Minority Carrier Gradient in Base (Collector Current)

If recombination does not play a role in the emitter (i.e. $L_{nE} \gg W_E$), use emitter thickness W_E instead of L_{nE}

NON-IDEAL BJT

BASE RECOMBINATION

Some of the injected electrons I_E recombine with holes in the base. Note that most electrons reach the collector since $L_{nB} \gg W$. The recombined holes are re-supplied by base current and therefore I_{BB} will rise and β will be reduced.

$$I_B = I_E - I_C = I_{En} + \underbrace{\left(I_{Ep} - I_{Cp}\right)}_{\text{Base Recombination}} - I_{Cn}$$

B/C REVERSE LEAKAGE

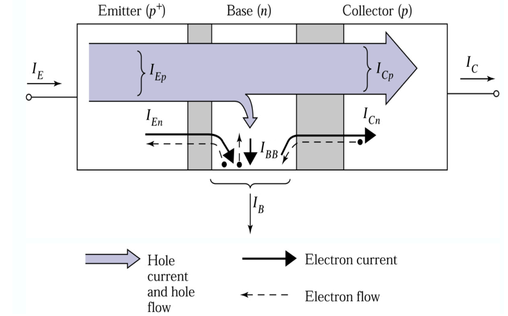
The B/C Junction is reverse biased but nonetheless a small minority hole diffusion current from collector to base exists.

$$I_{pC} = \frac{qD_{pC}p_{C0}}{L_{pC}}$$

BAND TO BAND TUNNELLING

At **high doping level** quantum mechanical tunnelling occurs.
For a npn BJT we have (equivalent for pnp) electrons from the first n-p junction will tunnel and recombine with a hole of the p-type base. Therefore: $I_{Et} = I_{Bt}$
Normal BJT operation: $I_{En} = I_{Bn} + I_{Cn}$
With tunnelling: $I_{E,tot} = I_{En} + I_{Et} = I_{Bn} + I_{Bt} + I_{Cn}$
If tunnelling becomes dominant, then I_{Cn} goes to 0. ($\beta \to 0$)

PNP



$I_B = I_E - I_C = I_{nE} + I_{BB} - I_{nC} = I_{nE} + \left(I_{pE} - I_{pC}\right) - I_{nC}$

For $W_B \gg L_B W$ there is no recombination in the base Region $\Rightarrow \alpha_T = 0, J_{BB} = 0, J_{pC} = J_{pE} = J_{pB}$

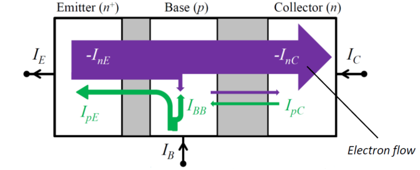
For Silicon it holds: (total reverse current)

$$J_{nC} = J_S + \underbrace{J_{gen}}_{\approx 0} = \left[\frac{qD_n n_{p0}}{L_n}\right]$$

$n_{p0} = \frac{n_i^2}{N_A}$minority carrier concentration in C

⇨ Hole current & hole flow → Electron current ← - - - Electron flow

NPN

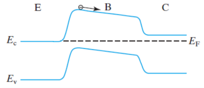


$I_B = I_{pE} + I_{BB} - I_{pC} = I_{pE} + \left(I_{nE} - I_{nC}\right) - I_{pC}$

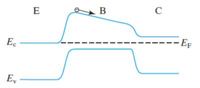
DRIFT AIDED TRANSISTOR

Carrier transport can be aided by introducing an E-Field in the base layer by grading the base doping.

The first solution can be achieved by having a different doping profile through the base.



The second solution is achieved by reducing the bandgap across the base, by incorporating e.g. some germanium atoms (smaller bandgap) in silicon



Drift/diffusion current density:

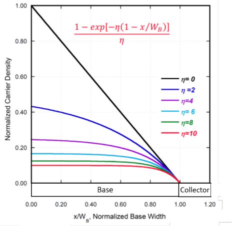
$$J_n = qD_n\frac{dn}{dx} + qn\mu_n E = qD_n\left[\frac{dn}{dx} + \frac{nE}{kT/q}\right]$$

Electron density with an E-field (NPN)

$$n(x) = -\frac{J_n W_B}{qD_n}\frac{1-\exp\left(-\eta(1-x/W_B)\right)}{\eta}$$

where: **accelerating field factor**

$$\eta = \frac{W_B \cdot E}{kT/q}$$



The electric field helps to reduce the electron density near the emitter. This reduces the stored charge $\boldsymbol{Q_B}$ and therefore the base transit time.

- → near the Emitter current is carried by drift
- → near the Collector (all) the current is carried by diffusion

$$Q_B = -q\int_0^{W_B} n(x)dx = \frac{J_n W_B^2}{D_n \eta^2}(\eta - 1 + e^{-\eta})$$

Base Transit Time (reduced)

$$\tau_B = \frac{Q_B}{J_n} = \frac{W_B^2}{D_n}\left(\frac{\eta - 1 + e^{-\eta}}{\eta^2}\right) \overset{\eta>3}{=} \frac{W_B^2}{D_n}\left(\frac{\eta-1}{\eta^2}\right)$$

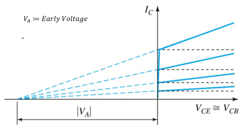
without E-Field ($\eta = 0$): $\tau_B = \frac{1}{2}\frac{W_B^2}{D_n}$

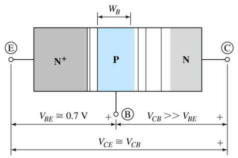
derived with l'Hôpital rule for $\lim\limits_{\eta\to 0}\tau_B$

Inverted E Field:

Consider a doping grading with η. If the doping grading is inverted, we observe $\eta' = -\eta$.

EARLY EFFECT (BASE WIDTH MODULATION)





The Collector current depends on V_{BC}. Increasing the Collector/Base reverse bias **widens** the **depletion region at the C/B junction**. The widening of the depletion region leads to a **smaller base width** $\boldsymbol{W}$ and therefore the **minority carrier gradient** in the **Base** is **enhanced** which lead to an **increased** collector current $\boldsymbol{I_C}$. To avoid this effect, the Base doping must be higher than the collector doping (i.e. (npn) $N_{AB} \gg N_{DC}$)

$$\frac{dI_C}{dV_{EC}} = \frac{I_C}{V_A + V_{EC}} =: \frac{1}{R_{out}} \qquad \rightarrow \text{high } V_A \text{ are desireable}$$

To determine V_A, determine 2 points of the IV curve then:

$$V_A = J_{C1}\cdot\frac{V_{CE,2} - V_{CE,1}}{J_{C2} - J_{C1}} - V_{CE,1}$$

Gummel number: $G_B = N_{DB}\cdot W_B$

(PNP) $I_C = qA_E n_i^2 D_{pB}\cdot\frac{1}{G_B}\left(e^{qV_{EB}/kT} - 1\right)$

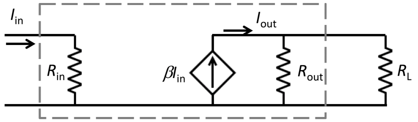
$$\frac{I_C}{dI_C/dV_{EC}} = \mathrm{V_A} + \mathrm{V_{EC}} - \frac{qG_B}{q\,dG_B/dV_{BC}} = \frac{qG_B}{C_{BC}} = \frac{qN_{DB}W_B}{C_{BC}} \quad , qG_B = Q_B$$

Note: High early voltage V_A requires a high base Gummel number

SSE AND POWER GAIN B|8.18

First determine the operating point (I_C) with large signal circuit.

Collector current: $I_C = I_S\exp\left(\frac{qV_{BE}}{kT} - 1\right)$



Power Gain $\boldsymbol{G}$

$$G = \frac{P_{out}}{P_{in}} = \frac{I_{out}^2 R_L}{I_{in}^2 R_{in}} \xrightarrow{R_{out}\gg R_L} \beta^2\frac{R_L}{R_{in}} = \beta^2\frac{R_L q I_B}{kT}$$

Transconductance

$$g_M = \frac{\partial I_C}{\partial V_{EB}} = \frac{I_C}{kT/q}$$

$$\frac{1}{R_{out}} = \frac{dI_C}{dV_{EC}}$$

$$R_{in} \approx \frac{I_C}{\beta(kT/q)}$$

It is desirable that the Output Resistance R_{out} is as large as possible, such that the Device act like an Ideal Current Source, i.e. be able to feed a constant current to the load regardless of the load resistance. For $R_{out}\to\infty$ the Early Voltage acts as $V_A\to\infty$ and therefore the Early Effect is negligible.

If we cannot neglect the Early Effect, or R_{out} is finite, then:

$$G_A = \frac{P_{out}}{P_{in}} = \frac{I_{out}^2 R_L}{I_{in}^2 R_{in}} = \frac{(\beta I_{in})^2}{I_{in}^2}\left(\frac{R_{out}}{R_{out}+R_L}\right)^2\frac{R_L}{R_{in}}$$

$$= \beta^2\cdot\frac{R_L}{R_{in}}\left(\frac{\frac{R_{out}}{RL}}{\frac{R_{out}}{R_L}+1}\right) = \beta^2\cdot\frac{R_L}{R_{in}}\left(\frac{[V_A+V_{CE}]/I_C R_L}{[V_A+V_{CE}]/I_C R_L + 1}\right)^2$$

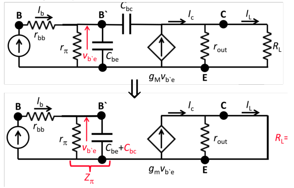
Here, we face a trade-off: high Power Gain requires high Early voltage, high Early voltage requires high Gummel Number. But a high Gummel number reduces the Current Gain.

Waisted power: $P_D = (V_{CE}\cdot I_C) - P_{out}$

Intrinsic voltage gain: $A_V = g_m\cdot R_{out} = \frac{V_A}{kT/q}$

BJT BANDWIDTH

Small Signal Equivalent Circuit



$$\boldsymbol{I_C = g_m v_{be}}$$

$$\boldsymbol{I_b = \frac{v_{be}}{Z_\pi}}$$

Input Impedance:

$$\boldsymbol{Z_\pi = r_\pi || \underbrace{(C_{be} + C_{bc})}_{:=C_\pi}}$$

The Common Emitter current gain cut-off frequency f_{T0} represents the frequency at which the **current gain= 1** with a short-circuit load ($R_L = 0$).

$$\beta(\omega) = \frac{I_C}{I_B} = \frac{g_m r_\pi}{1 + j\omega r_\pi C_\pi} = \frac{\beta_0}{1 + j\omega r_\pi C_\pi} = \frac{\beta_0}{1 + j(f/f_\beta)}$$

This is a lowpass RC filter

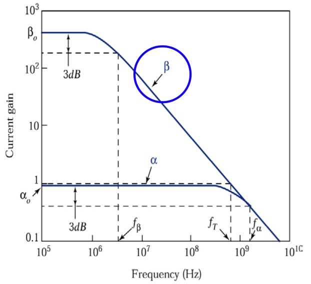
with: $\beta(0) = \beta_0 = g_m r_\pi \qquad f_\beta = \frac{1}{2\pi\, C_\pi r_\pi}$

We determine $|\boldsymbol{\beta(\omega)}| \overset{!}{=} \mathbf{1} \quad\Rightarrow\quad \boxed{f_{T0} = \beta_0 f_\beta = \frac{g_m}{2\pi C_\pi}}$

Note: $\alpha(\omega) = \frac{I_C}{I_E} = \frac{\beta(\omega)}{1+\beta(\omega)} = \frac{\alpha_0}{1+j(f/f_\alpha)}$

with: $\alpha_0 = \beta_0/(\beta_0+1) \qquad f_\alpha = (\beta_0+1)f_\beta$

$f_\beta = (1-\alpha_0)f_\alpha = f_T/\beta_0$



additional delay terms

Previously we assumed that the collector current is an instantaneous function of V_{BE} . But in fact, the Minority Carriers must **diffuse** across the base. This causes a Time Delay called the **Base Transit Time** $\boldsymbol{\tau_B}$. Additionally they must **travers the depletion region**, which adds a **Collector Signal Delay** $\boldsymbol{\tau_C}$. The time delays are incorporated through the exp function since in Laplace domain, time delay T is e^{-sT}.

$$\alpha(\omega) = \frac{\alpha_0}{1+j(f/f_\alpha)}\cdot\underbrace{e^{-j\omega\tau}}_{Delay\ Term} \qquad \text{Note: } e^{j\omega\tau}\approx\frac{1}{1+j\omega\tau} \quad , if\ \omega \ll 1/\tau$$

$$\Rightarrow \alpha(\omega) = \frac{\alpha_0}{1 + jf\left(\frac{1}{f_\alpha}+\frac{1}{f_\tau}\right)} = \frac{\alpha_0}{1 + jf\left(\frac{1}{f_{\alpha\tau}}\right)}$$

the new alpha Cut-off Frequency is:

$$\frac{1}{f_{\alpha\tau}} = \frac{1}{f_\alpha} + \frac{1}{f_\tau} = \frac{1}{f_\alpha} + \frac{1}{2\pi\tau} \qquad , f_{\alpha_{new}} = f_{\alpha\tau}$$

Cut-off Frequency

The cut-off frequency f_T ($|\beta(f_T| \equiv 1$) is given as:

$$\boxed{f_T = \sqrt{\beta_0^2 - 1}\cdot f_{\beta\tau} \cong \alpha_0 f_{\alpha\tau} = \frac{1}{2\pi\tau_T}}$$

where we used: $f_{\alpha\tau} = (\beta_0+1)f_{\beta\tau} = f_T/\alpha_0$ and $\beta_0^2 \gg 1$

$f_{\beta\tau} = (1-\alpha_0)f_{\alpha\tau} = f_T/\beta_0$

Total transit time $\tau_T = \frac{C_\pi}{g_m} + \tau_B + \tau_C + \cdots$

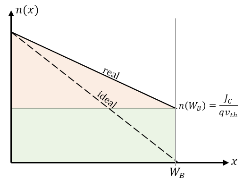
Delay Times

Fundamental Transistor Delay	C_π/g_m
Base Transit Time (ideal)	$\tau_B = Q_B/J_C = W_B^2/2D_n$
Collector Signal Delay	$\tau_C = W_C/2v_{Sat}$
Emitter Charging Time	$C_{BE}(R_E + R_C + r_\pi)$
Collector Charging Time	$C_{BC}(R_C)$

REAL BASE TRANSIT TIME

In reality, the velocity at which electrons can leave the base and enter the collector is limited by the thermal velocity v_{th}. The collector current density at the B/C boundary is given as:

$$J_C = q\cdot n(W_B)\cdot v_{th}$$



First we determine another expression for J_C:

$$J_C = qD_n\frac{dn}{dx} = qD_n\frac{[n(0) - n(W_B)]}{W_B}$$

The new base transit time τ_B' is defined as the total minority charge Q_B divided by J_C

total minority charge: $Q_B = \left(n(W_B)\cdot W_B + \frac{W_B}{2}[n(0) - n(W_B)]\right)q$

$$\tau_B' = \frac{Q_B}{J_C} = \frac{q\cdot n(W_B)\cdot W_B}{J_C} + \frac{1}{2}\frac{q\cdot W_B[n(0) - n(W_B)]}{J_C}$$

we use both definitions of J_C and we get:

$$\tau_B' = \frac{q\cdot n(W_B)\cdot W_B}{q\cdot n(W_B)\cdot v_{th}} + \frac{1}{2}\frac{q\cdot W_B[n(0)-n(W_B)]}{q\cdot D_n\frac{[n(0)-n(W_B)]}{W_B}}$$

$$\boxed{\tau_B' = \frac{W_B}{v_{th}} + \frac{1}{2}\frac{W_B^2}{D_n} = \frac{W_B}{v_{th}} + \tau_{B,ideal}}$$

Therefore the real transit time $\tau_B' > \tau_{B,ideal}$ because more carriers can be stored and the slope isn't as steep as before. The same principle can be applied when, for example, the diffusivity isn't constant throughout the base. Then

Power Gain Cut-off Frequency f_{max}

Power Gain: $G_p = \frac{1}{f^2}\cdot\frac{f_T}{8\pi R_B C_{BC}}$

The power cut-off frequency f_{max} is defined where $G_p = 1$

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_B C_{BC}}}$$

Conclusion:

Fast means **high frequencies**, therefore we want to increase f_T which corresponds to decreasing the delay terms and therefore we need **high collector current** levels to be fast! But we know that high collector currents mean high current gain and this leads to a **high power dissipation**. So high-speed bipolar integrated circuits have high power dissipation.

HBT | HETEROJUNCTION BIPOLAR TRANSISTOR

Different materials are used in the Base and Emitter, therefore different intrinsic carrier concentrations.

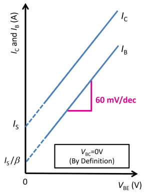
$$\beta_{HBT} = \frac{D_{pB}}{D_{nE}}\frac{L_{nE}}{W}\frac{n_{iB}^2/N_{DB}}{n_{iE}^2/N_{AE}} = \beta_{BJT}\cdot\frac{n_{iB}^2}{n_{iE}^2} \qquad \text{with: } \frac{p_{nB0}}{n_{pE0}}$$

$$= \beta_{BJT}\cdot\underbrace{\frac{(N_{CB}N_{VB})}{(N_{CE}N_{Ve})}}_{negligible}\frac{e^{-E_{GB}/kT}}{e^{-E_{GE}/kT}} = \beta_{BJT}\cdot e^{(E_{GE}-E_{GB})/kT}$$

⇒ Gain through different band gaps

To achieve a high β we want to have $n_{iB} > n_{iE}$ what corresponds to an higher bandgap in the emitter region.

GUMMEL CHARACTERISTICS

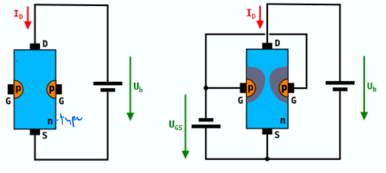


The Gummel plot reflects the quality of the emitter-base junction, while $\boldsymbol{V_{BC}}$ is kept constant ($\boldsymbol{V_{BE} = 0}$). We can read off the plot the common-emitter current gain $\boldsymbol{\beta}$, the common-base current $\boldsymbol{\alpha}$.

FET

Field effect transistors (**FET**) are a type of transistors where the conductivity of a majority carrier channel between two contacts (source and drain) is modulated by a gate electrode.

JFET



The depletion of reverse-biased PN junctions narrows the channel (pinches the channel) and modulates current flowing between the source and drain. Low input gate current. Normally-ON devices.

MOSFET

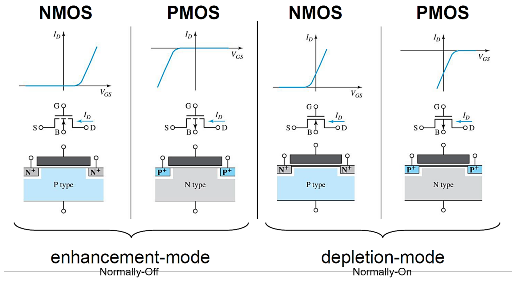
MOSFET=Metal Oxide Semiconductor Field Effect Transistor

MOSFET's are **majority carrier devices!** Therefore electrical current in an **N-Channel** transistor is carried by **electrons**, whereas in an **P-Channel** transistor the current is carried by **holes**.

NMOS & PMOS have different Gate Lengths due to different mobility of electrons & holes.
The NMOS/PMOS pair is designed so that their speed match each other.
MOSFETS require less space than BJTs.

Two varieties of MOSFET's:

1. a channel is present at equilibrium
   → **Normally-On ⇔ Depletion-Mode**
2. no channel is present at equilibrium
   → **Normally-Off ⇔ Enhancement-Mode**



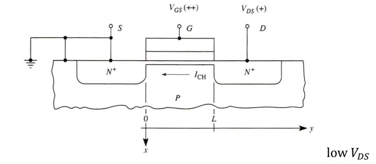
Sheet Resistance

Consider a uniform quadratic layer with a resistivity ρ, a thickness X and width/length both $= L$.
The Sheet Resistance is independent of L and defined as:

$$R_S[\Omega/square] = \frac{\rho L}{A} = \frac{\rho L}{XL} = \frac{\rho[\Omega m]}{X[m]}$$

Operating Principle (E-Mode, N-Channel)

The vertical field V_{GS} applied through an oxide insulator modulates the carrier density in the channel and thus its conductivity. First let V_{DS} be quite small.



Channel Charge Density:

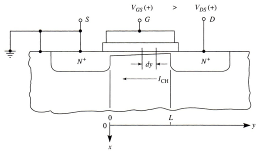
$$Q_n = -\text{qnX} = -\frac{qn\rho}{R_S} = -\frac{qn}{qn\mu_n R_S} = -\frac{1}{\mu_n R_S}$$

The gate and oxide work as a simple capacitor:

$$Q_n = -C_{OX}(V_{GS} - V_T)$$

Sheet Resistance:

$$R_S = \frac{1}{\mu_n C_{OX}(V_{GS} - V_T)}$$



V_{DS} increases
→ channel Voltage $V(y)$ vary from 0 @ Source to V_{DS} @ Drain $(y = L)$
→ Sheet Resistance will vary across the channel

Resistance of channel element (Length dy, Width Z at Position y):

$$\Rightarrow dR = \frac{dy}{Z} R_S(y) = \frac{\text{dy}}{Z\mu_n C_{OX}\big(V_{GS} - V_T - V(y)\big)}$$

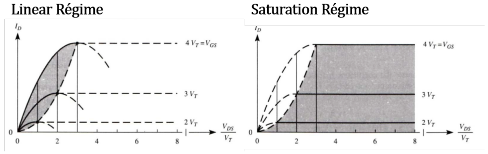
The current is therefore:

$$I_{CH} = \frac{dV}{dR}$$

$$I_{CH} = \int_0^L I_{CH} dy = Z\mu_n C_{OX} \int_0^{V_{DS}} V_{GS} - V_T - V(y) dV$$

$$\boxed{I_{CH} = I_D = \frac{\mu_n C_{OX}}{2}\frac{Z}{L}[2(V_{GS} - V_T)V_{DS} - V_{DS}^2]}$$

This equation defines inverted parabolas:



Linear Régime

Saturation Régime

Linear region

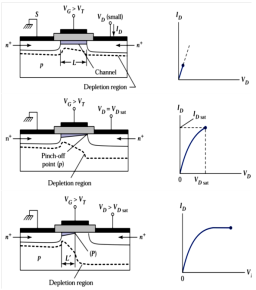
$0 \le V_{DS} \le V_{Dsat}$ $\quad ,V_{Dsat} = V_{GS} - V_T$
$V_{GS} > V_T$

$$I_D = \frac{\mu_n C_{OX}}{2}\frac{Z}{L}[2(V_{GS} - V_T)V_{DS} - V_{DS}^2]$$

Saturation region

$V_{DS} > V_{Dsat}$ $\quad ,V_{GS} > V_T$

$$\boxed{I_{Dsat} = \frac{\mu_n C_{OX}}{2}\frac{Z}{L}[(V_{GS} - V_T)^2]}$$



Note that the Depletion region widens at the Drain side.
Channel length modulation will make I_D slightly increase in the saturation region (instead of being constant).

In the saturation regime, we define:

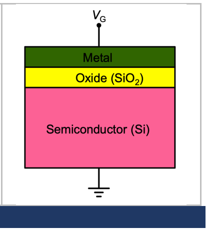
Transconductance:	**Material Constant:**
$g_m = \frac{dI_{Dsat}}{dV_{GS}} = 2K(V_{GS} - V_T)$	$K = \frac{\mu_n C_{OX}}{2}\frac{Z}{L}$

Note:
- g_m varies linearly with V_{GS} whereas g_m depends exponentially on V_{BE} in a BJT
- NMOS devices show higher g_m since they rely on electron mobility rather than PMOS, which rely on hole mobility $\mu_n > \mu_p$

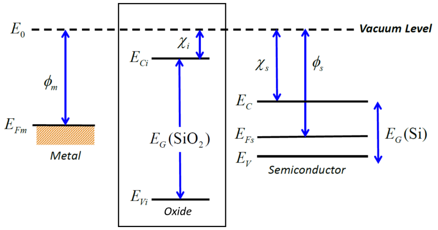
Q: Why does current still flow, though the channel completely disappear in the saturation regime?
A: If we argue per contradiction: having no current means constant carrier density across the channel, but this would mean constant channel width. Contradiction with original assumption. Physically, the pinched off region has a longitudinal electric field that goes to infinity, this supports a drift current even though the carrier (e$^-$) density is vanishing.

MOSCAP

To analyse the MOSFET, we first have a look at the MOS-Capacitor, which illustrates the operation principle between the gate and the channel.
MOSCAP is a MOS structure consisting of an oxide between metal and semiconductor.



BAND DIAGRAM



Vacuum Level: reference energy level $E_0\ [eV]$
Work function ϕ: energy difference from Fermi-level to $E_0[V]$
Metal: $q\phi_m = E_0 - E_{Fm}[\text{eV}]$
SC: $q\phi_S = E_0 - E_{FS}[\text{eV}]$
Electron affinity χ: $q\chi = E_0 - E_C\ [eV]$
bulk potential ψ_B: energy difference between Fermi-level and intrinsic Fermi-level $[V]$ i.e.
$\psi_b = (E_i - E_F)/q$

$$\boxed{\psi_B = \begin{cases} \frac{-kT}{q}\ln\left(\frac{N_D}{n_i}\right) < 0 & \text{NType} \\ \frac{+kT}{q}\ln\left(\frac{N_A}{n_i}\right) > 0 & \text{PType} \end{cases}}$$

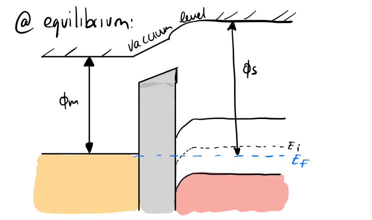
surface potential ψ_S: energy difference between bulk potential and intrinsic Fermi-level @ the oxide interface $(\psi(0) = \psi_S)\ [V]$

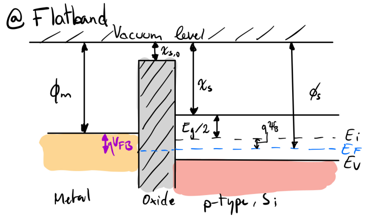
Note: Conversion energy difference ↔ voltage

$$E_a - E_b[eV] \ \Rightarrow\ \frac{E_a - E_b}{q}\ [V] \quad | \quad q\phi[eV] \ \Rightarrow\ \phi[V]$$

FLATBAND VOLTAGE

Under equilibrium the Fermi-Level must again be constant (flat) through the whole structure (zero current flow). At equilibrium, E_C, E_V will usually be bent (Shockley boundary condition).





We define the **Flatband Voltage V_{FB}** as the Gate Voltage V_G that makes the bands flat. If there is no charge at the oxide-semiconductor interface, this is equivalent to the difference of the workfunctions.

$$\phi_s = \chi_s + E_g/2q + \psi_B = \chi_s + (E_C - E_F)/q$$
$$= \chi_s - kT/q \cdot \ln{(n_0/N_c)}$$

$$\phi_{ms} = (\phi_m - \phi_s)$$

$$\boxed{\boldsymbol{V_{FB} = \phi_{ms}} - \frac{1}{\varepsilon_s}\int_0^d x\rho(x)dx \boldsymbol{= \phi_{ms} - \frac{Q_f}{C_{ox}}} \qquad [V]}$$

Last term usually irrelevant
$\rho \coloneqq$ Charge Density
The second term is if there are only fixed Charge $Q_f[C/cm^2]$(12.25)
Remember: $E_i > E_F$ → P-Type $\qquad E_i < E_F$ →N-Type

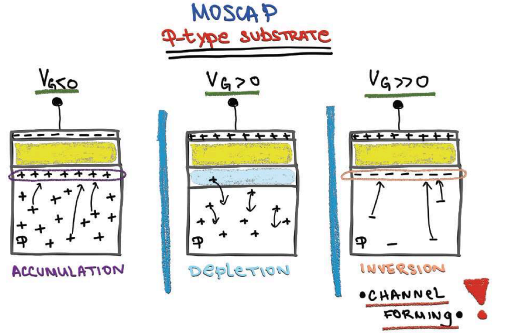
Selected Gate Electrode Materials

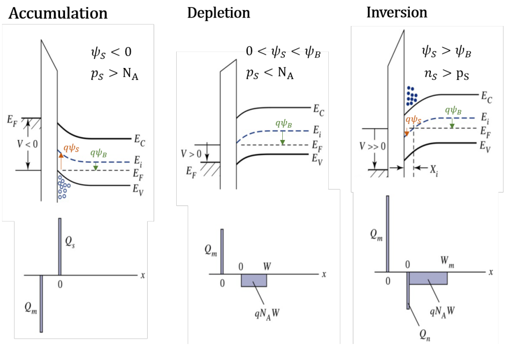
Gate Material	Work Function (eV)
n+ Polysilicon	4.0
Al	4.25
W	4.6
MoSi$_2$	4.5
PtSi	5.4
Pd$_2$Si	5.1

Selected Gate Insulators $\Delta E_C = \chi_s - \chi_i$

Insulator	ε_r	Gap (eV)	ΔE_C to Si
SiO$_2$	3.9	8-9	3.2
Si$_3$N$_4$	7.2-7.6	5.1	2.0
Al$_2$O$_3$	9.0	8.7	2.1
Ta$_2$O$_5$	26	4.5	0.5
ZrO$_2$	25	5.8	1.2
HfO$_2$	25	5.7	1.5
TiO$_2$	80	3.5	1.2

Channel Modulation by V_G





p_s ≔ hole concentration @ surface n_s ≔ electron concentration @ surface

Condition of Interest:

$$\begin{cases} \psi_s = 0 & \text{Flatband Condition} \\ \psi_s = \psi_B & \text{Midgap}, p_s = n_s = n_i \Rightarrow \text{intrinsic MOS Capacitor} \\ \psi_s \geq 2\psi_B & \text{Strong inversion} \end{cases}$$

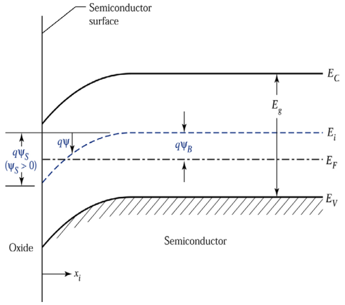
Because we need a standard non-ambiguous criterion for inversion we define inversion as:

$$\boldsymbol{\psi_s(inv) = 2 \cdot \psi_B}$$

The Fermi-level E_F **stays flat** perpendicular to the surface because there is no current flow through the oxide.

P-type substrate bends down at inversion (more electrons than holes at the oxide interface). So $V_G = V_T > 0$ because it needs to attract electrons and repel holes at the oxide interface.
N-type substrate bends up at inversion (more holes than electrons at the oxide interface). This means $V_T < 0$ because it needs to attract electrons and repel holes.

Surface Potential



We define the **Electrostatic Potential $\boldsymbol{\psi}$** such that it is zero in the bulk.

Carrier Densities:

P-Type:

$$p_p = n_i e^{(E_i - E_F)/kT} = n_i e^{(\psi_B - \psi)/kT}$$
$$n_p = n_i e^{(E_F - E_i)/kT} = n_i e^{(\psi - \psi_B)/kT}$$
$$p_p \cdot n_p = n_i^2$$

N-Type:

$$p_n = n_i e^{(E_F - E_i)/kT} = n_i e^{(\psi - \psi_B)/kT}$$
$$n_n = n_i e^{(E_i - E_F)/kT} = n_i e^{(\psi_B - \psi)/kT}$$
$$p_n \cdot n_n = n_i^2$$

ψ dependant on $x \rightarrow \psi(x)$

Electrostatic Potential:

$$\psi(x) = \psi_s \left(1 - \frac{x}{W}\right)^2 = \frac{qN_A}{2\varepsilon_S}(W - x)^2 \quad 0 \leq x \leq W$$

where: $\psi_S = \psi(0) = \frac{qN_A W^2}{2\varepsilon_S}$ $[V]$

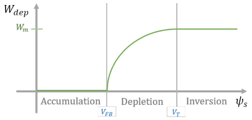
Inversion:

The surface region is inverted once we have more electrons than holes. We define this electron concentration as:

$$n_S = N_A = n_i e^{q\psi_B/kT} \Rightarrow \psi_S(inv) = 2\psi_B = \frac{2kT}{q} \ln\left(\frac{N_A}{ni}\right)$$

Depletion region:

$$\boxed{W = \sqrt{\frac{2\varepsilon_S \psi_S}{qN_A}} = 2\sqrt{\frac{\epsilon_S kT ln(N_A/n_i)}{q^2 N_A}}}$$



The surface depletion stops expanding when inversion is reached and the maximum depletion region is computed as:

$$W_m = W_{max} = 2\sqrt{\frac{\varepsilon_S \psi_B}{qN_A}} \quad \text{with } \psi_s = 2\psi_B$$

Threshold Voltage

We define the threshold voltage **for ideal MOS** as the voltage where inversion starts.

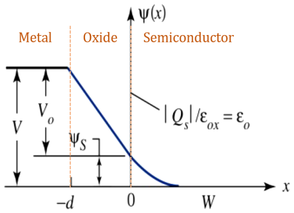
$$\boxed{V_{T,ideal} = \underbrace{\frac{qN_A W_m}{C_{ox}}}_{Voltage\ accros\ Oxide} + \psi(inv) = \frac{\sqrt{2q\varepsilon_S N_A(2\psi_B)}}{C_{ox}} + 2\psi_B}$$

For **non-ideal MOS** the threshold voltage is modiefied by the workfunction difference ϕ_{ms} and oxide charges (S12a2)

$$V_T = \underbrace{\frac{\sqrt{2q\varepsilon_S N_A(2\psi_B)}}{C_{ox}} + 2\psi_B}_{V_{T,ideal}} + \underbrace{\phi_{ms} - \frac{Q_f}{C_{ox}}}_{V_{FB}}$$

where: $C_{ox} = \varepsilon_{ox}/d$ & Q_f ≔ fixed Charge

Gate Voltage



Potential Drop across Oxide Layer:

From the continuity @ the interface it must hold $\boxed{\varepsilon_{ox}E_{ox} = \varepsilon_s E_s}$

$$V_{ox} = d \cdot E_{ox} = d \cdot \frac{\varepsilon_S}{\varepsilon_{ox}} E_s = \frac{\varepsilon_S}{C_{ox}} E_s = \frac{\sqrt{2q\varepsilon_S N_A \psi_S}}{C_{ox}} \quad [V]$$

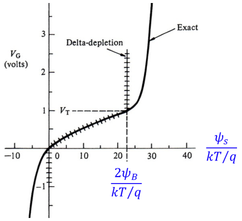
$$C_{ox} = \frac{\varepsilon_{ox}}{d} = \frac{\varepsilon_0 \varepsilon_{(SiO_2)}}{d} \quad \left[\frac{F}{cm^2}\right]$$

Surface Electric Field in SC:

$$E_s = E(0) = \frac{\varepsilon_{ox}}{\varepsilon_S} \cdot E_{ox}$$

$$E_s = \sqrt{\frac{2qN_A}{\varepsilon_S}\psi_S} = \frac{qN_A}{\varepsilon_S} W \left[\frac{V}{m}\right]$$

If there is depletion inside the gate, then all of the above can be replicated.



Gate Voltage:

$$\underbrace{Gate\ Voltage}_{V_G} = Potential\ drop\ (\underbrace{over\ SC}_{\psi_S} + \underbrace{accros\ Oxide}_{V_{ox}})$$

$$\boxed{\boldsymbol{V_{G,ideal} = \psi_s + V_{ox}}}$$

For non ideal MOS: $V_G = V_{G,ideal} + V_{FB}$

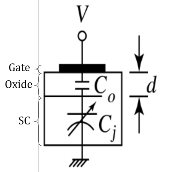
Capacitance vs. Frequency

Capacitance $C = Q/V$

Capacitance in
- Accumulation $C_0 = C_{ox} = \frac{\varepsilon_{ox}}{d}\left[\frac{F}{cm^2}\right]$
- Depletion $C_j = \frac{\varepsilon_s}{W_{max}}$

Total Capacitance:

C=$C_0 \parallel C_j = \frac{C_0 C_j}{C_0 + C_j} = \frac{\varepsilon_{ox}}{d + (\varepsilon_{ox}/\varepsilon_s)W}$

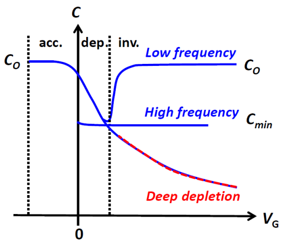


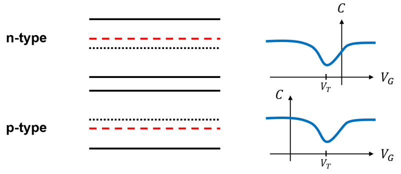
Various Oxide Charges→ they shift the threshold

Accumulation: $V_G < V_{FB}$	Majority carriers respond to AC signal at both HF & LF, $\boldsymbol{C = C_{ox} = \varepsilon_{ox}/t_{ox}}$
Depletion: $V_{FB} < V_G < V_T$	Depletion region and oxide capacitance in series, C decreases with V_G due to widening of depletion region. $C = \frac{\varepsilon_{ox}}{\left(\frac{\varepsilon_{ox}}{\varepsilon_s}\right) x_d + t_{ox}}$
Inversion: $V_T < V_G$	At LF, minority carrier generation/recomb occurs in response to AC signal ($C = C_{ox}$). At HF, the minority carriers do not respond to the AC signal. C is constant due to constant depletion region width ($W = W_m$)
Deep depletion:	DC bias is swept so rapidly that minority carriers cannot respond and therfore no inversion layer is formed. The charge on the gate is balanced by depletion of substrate.

Measurements in comparision to generation lifetime τ_g

For a p-type substrate we do a V_G sweep (from low to high) at different speeds (frequencies): (For n-type substrate we go from high V_G to low, see **s12a1**)





Deep depletion is just a very fast sweep, what happens is that the depletion width continues to grow with higher V_G so $C_d = \frac{\varepsilon_s}{W_s}$ will become smaller and smaller.

C_d is a series resistance created by the absence of charges (depletion) near the oxide interface.

SUBTHRESHOLD RÉGIME

We consider a MOSFET operating with a **very weakly inverted surface** i.e. not completely ON ($V_G < V_T$). This is called the **subthreshold regime.**
From Source to the Drain, the *npn* / *pnp* regions acts as a BJT. The current will be dominated by **diffusion.**

Carrier densities at Source/Drain side:

$n(0) = n_i e^{q(\psi_s - \psi_B)/kT}$ $\qquad n(L) = n_i e^{q(\psi_s - \psi_B - V_D)/kT}$

With $\psi_s \approx (V_G - V_T)$ we find:

$$I_D = -qA \cdot D_n \frac{dn}{dy} \cong qAD_n \frac{n(0) - n(L)}{L}$$
$$\approx \underbrace{\frac{qAD_n n_i e^{-q\psi_B/kT}}{L}}_{\kappa} e^{q\psi_S/kT} = \kappa \cdot e^{q(V_G - V_T)/kT}$$

Subthreshold Swing

The Subthreshold Swing S measures how efficiently the device can be turned on and off. S is typically about $70 - 110\, mV/Decade$

$$S = \frac{1}{\frac{\partial(\log_{10}(I_D))}{\partial V_G}} = \frac{\Delta V_G}{\log_{10} I_D|_{V_G=V_T} - \log_{10} I_D|_{V_G=0}}$$

Subthreshold Leakage Current

The subthreshold leakage current $I_D|_{VG=0}$ can be derived from the subthreshold Swing with $\Delta V_g = V_T$

$$\log I_D|_{V_G=0} = \log I_D|_{V_G=V_T} - \frac{V_T}{S}$$

CURRENT SATURATION

Channel Pinch-Off

The saturation current is given as:

$$\boxed{I_{Dsat} = \frac{\mu_n C_{OX}}{2} \frac{Z}{L} [(V_{GS} - V_T)^2]}$$

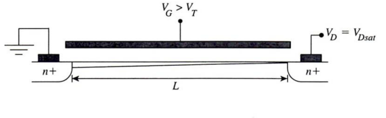
Increasing the drain voltage beyond V_{Dsat} causes the channel **pinch-off point to move towards the source**. Therefore the **effective channel length is reduced** to $(L - \Delta L)$, thus the current increases:
as seen from the FET square law

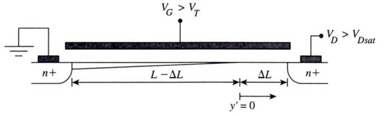
$$I_D = \frac{\mu_n C_{OX}}{2} \frac{Z}{L - \Delta L} [(V_{GS} - V_T)^2]$$

with $L - \Delta L = L(1 - \lambda V_D)$ & $K = \frac{\mu_n C_{OX}}{2} \frac{Z}{L}$

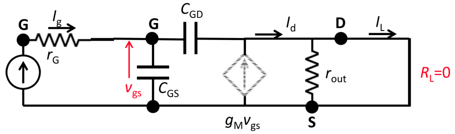
The I-V Characteristic becomes Drain Biased:

$$I_D = \frac{K}{2}(V_G - V_T)^2(1 + \lambda V_D) \qquad \lambda = \frac{1}{E_0\, L}$$





CUT-OFF FREQUENCY



$$I_d = g_m \cdot v_{gs} \qquad I_g = \frac{v_{gs}}{1/j\omega(C_{gs}+C_{gd})} \qquad A(\omega) = \frac{g_m}{j\omega(C_{gs}+C_{gd})}$$

For low Gate Voltages V_G, the capacitances from Gate to Drain and Gate to Source are almost the same ($C_{gs} = C_{gd}$). When $\boldsymbol{V_D}$ **increases**, the channel pinches-off near the Drain ant therefore $\boldsymbol{C_{gd}}$ **drops.**

Cut-Off Frequency: $(\boldsymbol{A(\omega) \equiv 1})$

$$f_T = \frac{g_m}{2\pi(C_{gs}+C_{gd})} = \frac{3}{4}\frac{\mu_n}{\pi}\frac{(V_{GS}-V_T)}{L^2}$$

where we used:

$$g_m = 2K(V_{GS}-V_T) \quad \& \quad K = \frac{\mu_n C_{ox}}{2}\frac{Z}{L}$$

SHORT CHANNEL EFFECTS

Threshold Voltage Shift

Reducing the channel length increases the transconductance g_m, the speed and device density. This downscaling leads to so called short channel effects.

Charge sharing:
A part of the region below the gate is depleted by the Source and Drain pn-junction depletion regions. The Gate voltage $\boldsymbol{V_G}$ **needed for inversion (threshold voltage** $\boldsymbol{V_T}$) thus decreases since the Gate must deplete less material to achieve inversion.
For short channel length, the subthreshold swing degrades.

$$\Delta V_T = -\frac{qN_A W_m r_j}{C_0 L}\left(\sqrt{1+\frac{2W_m}{r_j}}-1\right)$$

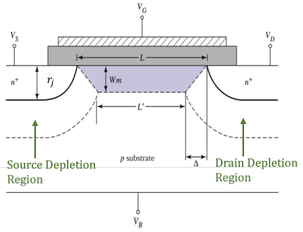
r_j ≔ Junction Depth $[\mu m]$ d ≔ Oxide thickness [Å] W_m ≔max depletion width

$$V_G = \underbrace{\frac{\sqrt{2q\varepsilon_s N_A(2\psi_B)}}{C_0} + 2\psi_B}_{V_{T,ideal}} + \underbrace{\phi_{ms} - \frac{qQ_f}{C_{ox}}}_{V_{FB}}$$

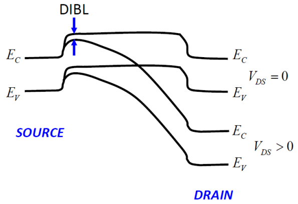
We define de minimal channel length for long channel behavior as:

$$L_{min} \geq 0.4 \cdot \left[r_j d(W_S+W_D)^2\right]^{1/3}$$

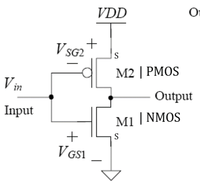
$W_{S,D}$ ≔ S/D Depletion Depths $[\mu m]$
Note: Thin Oxide d reduces the shift, whereas short Gate lengths and Deep Junctions increses the V_T shift.

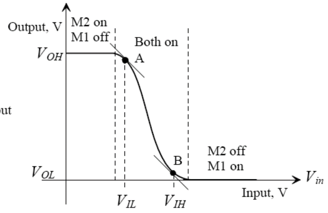


DIBL (Drain-induced barrier lowering):
For short gate length V_T decreases with increasing V_{DS} due to a reduction of the potential barrier below the Gate.



CMOS INVERTER

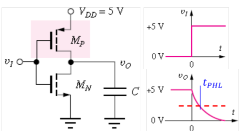




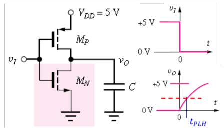
The standby power dissipation is ideally assumed to be Zero, because no DC current flows through M1 & M2.

Digital Switching Performance

Pull Down FET



Pull Up FET



As v_I goes High the PMOS turns off whereas the NMOS switches ON to discharge the load C down to logic level GND

As v_I goes LOW the NMOS turns off whereas the PMOS switches ON to charge the load C up to logic level V_{DD}

Propagation Delay

We define the Propagation Delay as the time t_{PHL} that's needed reach $V_{DD}/2$ from High.

$$t_{PHL} = \frac{8}{7}\frac{C_L}{K_n V_{DD}}$$

where we used the
- Proportionality Constant $K_n = \frac{C_i \mu_n Z_n}{L_n}$
- Lowes possible load Capacitance $C_L = C_i(L_n Z_n + L_p Z_p)$

for $L_n = L_p \& Z_n = Z_p$:

$$t_{PHL} \approx 2\frac{L_n^2}{V_{DD}\mu_n} \propto \frac{1}{f_T V_{DD}}$$

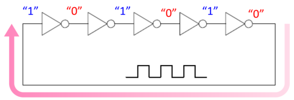
⇒ **High Cutoff Frequency = Fast Digital Switching.**
⇒ shorter Gates = higher performance

The Propagation delay from Low to $V_{DD}/2$ is defined as the time t_{PLH} and derived almost the same way but with the PMOS and will thus depend on μ_p.

This delay values are optimistic and represent the lowes values reachable because we neglectet several capacitances.

Improving Digital Switching Speed:
- Reduce C_L
- increase the W/L ration of transistors
- increase V_{DD}

Ring Oscillator



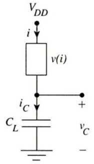
5 Stage Ring Oscillator

With an odd number of stages the circuit is unstable

$$f_{ROSC} = \frac{1}{N(t_{PHL}+t_{PLH})} \qquad , N \coloneqq \text{Stages}$$

Energy & Power Dissipation per Switching Gate

Energy Dissipation



$$E_{PMOS(LH)} = C_L \int_0^{V_{DD}} v\,dv = \frac{C_L V_{DD}^2}{2}$$

$$E_{NMOS(HL)} = \frac{C_L V_{DD}^2}{2}$$

$$E = E_{NMOS} + E_{PMOS} = C_L V_{DD}^2$$

Power Dissipation

$$P = \alpha_{0\to1} E f_{clock} + V_{DD} I_{Leakage} = \alpha_{0\to1} C_L V_{DD}^2 f_{clock} + V_{DD} I_{Leakage}$$

$\alpha_{0\to1}$ ≔ probability thate Gate switches in a given clock period
$I_{Leakage}$ ≔ Leakage current from V_{DD} to GND when Gate is not switching

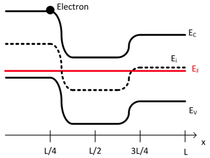
APPENDIX

Grösseneinheiten

G	Giga	10^9	**k**	Kilo	10^3	**μ**	Mikro	10^{-6}	**p**	Piko	10^{-12}
M	Mega	10^6	**m**	Milli	10^{-3}	**n**	Nano	10^{-9}	**f**	Femto	10^{-15}

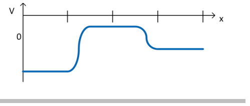
DRAWING GRAPHS

equilibrium
→ fermi level flat
$p(x) = n_i e^{(E_i(x)-E_F)/kT}$



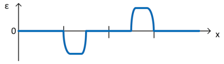
Electrostatic Potential

$V = -\frac{1}{q}\left(E_C - E_{ref}\right)$
(electrons)
$V = -\frac{1}{q}\left(E_{ref} - E_V\right)$ *(holes)*



Electric Field ε

$\vec{E} = -\frac{dV}{dx} = \frac{1}{q}\frac{dE_i}{dx} = \frac{1}{q}\frac{dE_V}{dx} = \frac{1}{q}\frac{dE_C}{dx}$

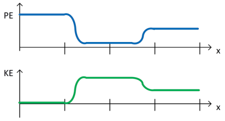


direction derived from the electrostatic potential:
$\varepsilon - Field$ same direction as holes

Energy

$E_{kin} + E_{pot} = const$

potential energy:
(electrons) $E_{pot} = E_C - E_{ref}$
(holes) $E_{pot} = E_{ref} - E_V$

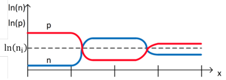


carrier concentration

$p = n_i e^{\left(\frac{E_i - E_{ref}}{kT}\right)}$

$n = n_i e^{\left(\frac{E_{ref} - E_i}{kT}\right)}$

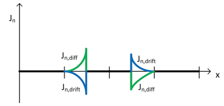
use a log-log scale



Currents

$J_{n,diff} = qD_n\frac{dn(x)}{dx}$

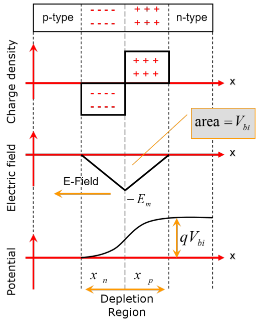
$J_{n,drift} = q\,n\,\mu_n\,\vec{E}$



more examples → Exercise Set 4.1

ELECTROSTATICS

PN Junction



NP Junction

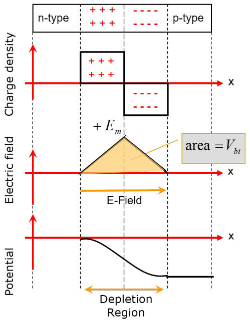


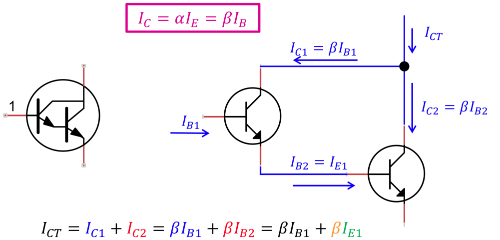
TABLE OF CONTENTS

EXERCISES INDEX

DARLINGTON PAIR

B|18

$\boxed{I_C = \alpha I_E = \beta I_B}$

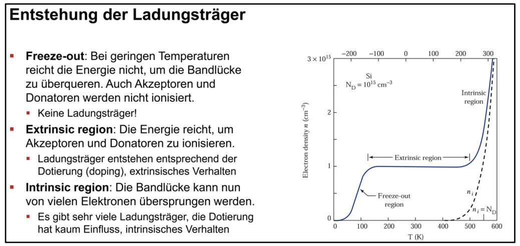


$$I_{CT} = I_{C1} + I_{C2} = \beta I_{B1} + \beta I_{B2} = \beta I_{B1} + \beta I_{E1}$$

$$I_{CT} = \beta I_{B1} + \beta\,\beta I_{B1}/\alpha = \beta I_{B1} + \beta(\beta+1)I_{B1} = \beta(\beta+2)I_{B1}$$

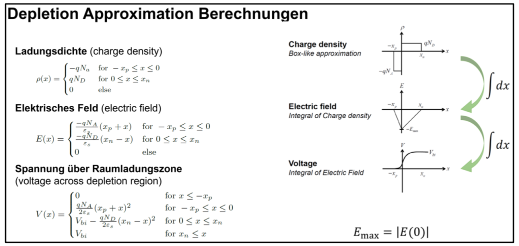
$$\boxed{\beta_{Darlington} \approx \beta^2}$$

OTHER



Entstehung der Ladungsträger

- **Freeze-out**: Bei geringen Temperaturen reicht die Energie nicht, um die Bandlücke zu überqueren. Auch Akzeptoren und Donatoren werden nicht ionisiert.
  - Keine Ladungsträger!
- **Extrinsic region**: Die Energie reicht, um Akzeptoren und Donatoren zu ionisieren.
  - Ladungsträger entstehen entsprechend der Dotierung (doping), extrinsisches Verhalten
- **Intrinsic region**: Die Bandlücke kann nun von vielen Elektronen übersprungen werden.
  - Es gibt sehr viele Ladungsträger, die Dotierung hat kaum Einfluss, intrinsisches Verhalten



Depletion Approximation Berechnungen

Ladungsdichte (charge density)

$$\rho(x) = \begin{cases} -qN_a & \text{for } -x_p \le x \le 0 \\ qN_D & \text{for } 0 \le x \le x_n \\ 0 & \text{else} \end{cases}$$

Elektrisches Feld (electric field)

$$E(x) = \begin{cases} \frac{-qN_A}{\varepsilon_s}(x_p + x) & \text{for } -x_p \le x \le 0 \\ \frac{-qN_D}{\varepsilon_s}(x_n - x) & \text{for } 0 \le x \le x_n \\ 0 & \text{else} \end{cases}$$

Spannung über Raumladungszone
(voltage across depletion region)

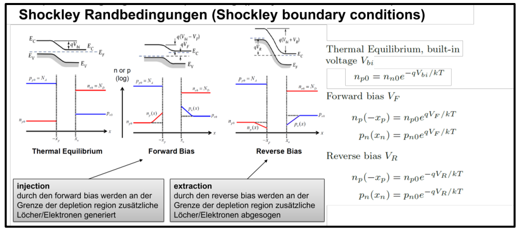
$$V(x) = \begin{cases} 0 & \text{for } x \le -x_p \\ \frac{qN_A}{2\varepsilon_s}(x_p + x)^2 & \text{for } -x_p \le x \le 0 \\ V_{bi} - \frac{qN_D}{2\varepsilon_s}(x_n - x)^2 & \text{for } 0 \le x \le x_n \\ V_{bi} & \text{for } x_n \le x \end{cases}$$

Charge density
Box-like approximation

Electric field
Integral of Charge density

Voltage
Integral of Electric Field

$E_{\max} = |E(0)|$



Shockley Randbedingungen (Shockley boundary conditions)

Thermal Equilibrium **Forward Bias** **Reverse Bias**

injection
durch den forward bias werden an der Grenze der depletion region zusätzliche Löcher/Elektronen generiert

extraction
durch den reverse bias werden an der Grenze der depletion region zusätzliche Löcher/Elektronen abgesogen

Thermal Equilibrium, built-in voltage V_{bi}

$$n_{p0} = n_{n0}e^{-qV_{bi}/kT}$$

Forward bias V_F

$$n_p(-x_p) = n_{p0}e^{qV_F/kT}$$
$$p_n(x_n) = p_{n0}e^{qV_F/kT}$$

Reverse bias V_R

$$n_p(-x_p) = n_{p0}e^{-qV_R/kT}$$
$$p_n(x_n) = p_{n0}e^{-qV_R/kT}$$

JUNCTION RESISTANCE OF A FORWARD BIASED IDEAL DIODE

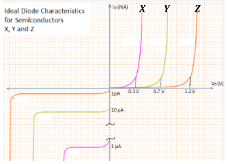
The junction resistance is defined as: $r = \frac{\partial V}{\partial I}$

→ ideal diode current with forward bias:

$$I = I_s\left(e^{\frac{qV_F}{kT}} - 1\right) \approx I_s\left(e^{\frac{qV_F}{kT}}\right) \Rightarrow V_F = \frac{kT}{q}\ln\left(\frac{I}{I_s}\right)$$

$$\Rightarrow r = \frac{\partial V}{\partial I} = \frac{kT}{q}\frac{1}{I}$$

DIODE CURRENT CHRARACTERISTICS



Bandgap:

$$V_{bi} = \frac{kT}{q}\ln\left(\frac{N_D N_A}{n_i^2}\right), \quad n_i = \sqrt{N_V N_C} e^{-\frac{E_g}{kT}}$$

→ the larger the bandgap, the higher V_{bi} to overcome to turn on the diode

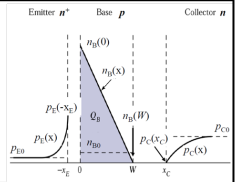
$$\rightarrow E_{gX} < E_{gY} < E_{gZ}$$

Reverse Current:

$$I = I_S\left(e^{\left(\frac{wV_D}{kT}\right)} - 1\right) \Rightarrow I_S = \left(\frac{qD_n n_{p0}}{L_n} + \frac{qD_p p_{n0}}{L_p}\right) = \left(\frac{qD_n n_i^2}{L_n N_A} + \frac{qD_p n_i^2}{L_p N_D}\right)$$

Therefore: the larger the bandgap ⇒ lower n_i^2 ⇒ lower I_S

Strombeiträge im NPN BJT



- Annahme: **Keine Rekombination** in den depletion regions sowie in der Base
- Die Ströme berechnen sich wie bei einer Diode (→ Diffusion):

$$J_B = J_{pE} = -qD_{pE}\frac{\partial}{\partial x}p_E(x)\bigg|_{x=-x_E}$$

$$J_C = J_{nB} = qD_{nB}\frac{\partial}{\partial x}n_B(x)\bigg|_{x=W}$$

$$J_E = J_B + J_C$$

- **Annahme**: Emitter ist viel länger als die minority carrier Diffusionslänge, denn dann

$$p_E(-x_E) = p_{E0}\exp\left(\frac{qV_{BE}}{kT}\right)$$

- **Annahme**: Base ist viel kürzer als die minority carrier Diffusionslänge, denn dann

$$n_B(0) = n_{B0}\exp\left(\frac{qV_{BE}}{kT}\right)$$

$$n_B(W) = n_{B0}\exp\left(\frac{qV_{BC}}{kT}\right)$$

- Aus Kapitel 1 kennen wir die Diffusionsglg.

$$p_E(x) = p_{E0} + (p_E(-x_E) - p_{E0})\exp\left(-\frac{-x-x_E}{L_p}\right)$$

$$n_B(x) = n_B(W) + (n_B(0) - n_B(W))\left(1 - \frac{x}{W}\right)$$

- Eingesetzt in die grauen Gleichungen:

$$J_B = -\frac{qD_{pE}}{L_{pE}}p_{E0}\left(\exp\left(\frac{qV_{BE}}{kT}\right) - 1\right)$$

$$J_C = -\frac{qD_{nB}}{W}n_{B0}\left(\exp\left(\frac{qV_{BE}}{kT}\right) - \exp\left(\frac{qV_{BC}}{kT}\right)\right)$$

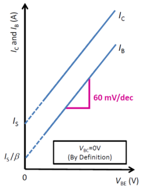
- Grundsätzlich also die selben Terme wie beim einfachen Diodenstrom, daher auch nicht auf meiner Zusammenfassung

3.2 Ideale Ströme
Gummel-Charakteristik

Umwandlung Stromdichte ↔ Strom
$I = J \cdot A$

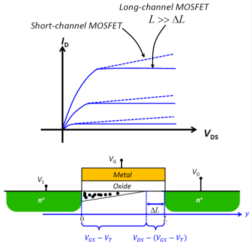
$$I_C = A_E\frac{qD_{pB}}{W_B}\frac{n_i{}^2}{N_{DB}}\left(e^{qV_{EB}/kT} - 1\right) = I_S\left(e^{qV_{EB}/kT} - 1\right)$$

$$I_B = A_E\frac{qD_{nE}}{L_{nE}}\frac{n_i{}^2}{N_{AE}}\left(e^{\underbrace{qV_{EB}/kT}} - 1\right) = \frac{I_S}{\beta}\left(e^{qV_{EB}/kT} - 1\right)$$



- Im logarithmischen Plot rechts haben die Ströme im idealen Fall eine Steigung von 60 mV/Dekade (ergibt sich aus ●)
- Eine Nichtidealität führt zu einer weniger steilen Gerade (Faktor η)
- Ein anderes β verschiebt den Anfangspunkt

Channel Length Modulation (CLM)



- In der Herleitung des Saturation Currents sind wir davon ausgegangen, dass der pinch-off-Bereich ΔL sehr klein ist verglichen mit der gesamten Channel-Länge
- Bei kurzem Channel stimmt dies nicht mehr, und da $I_D \propto \frac{Z}{L}$ steigt der Strom, wenn der **Channel kürzer wird**, also

$$I_D = \mu_n C_{ox}\frac{Z}{L - \Delta L}\frac{(V_{GS} - V_T)^2}{2}$$

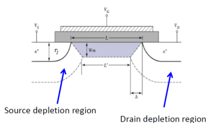
- Dies ist das Analogon zum Early-Effekt in einem BJT und führt zu einem Ausgangswiderstand

Charge Sharing:

- Zwischen dem Bulk und den Kontakten für Source und Drain bilden sich jeweils pn-junctions aus
- Wenn der Channel sehr kurz ist, «fressen» diese **pn-junction depletion regions** einen signifikanten Teil des Channels weg, wodurch er noch kürzer wird
- Dadurch sinkt die benötigte Gate-Spannung um Inversion zu erreichen, also die Threshold-Spannung V_T

$$\Delta V_T = -\frac{qN_A W_m r_j}{C_0 L}\left(\sqrt{1 + \frac{2W_m}{r_j}} - 1\right)$$

Ein **kürzerer Channel** hat viele Vorteile: Man erhöht die Transconductance g_m sowie die Geschwindigkeit und reduziert die Baugrösse. Es treten allerdings mehr Nichtidealitäten, so genannte Short Channel Effects, auf.



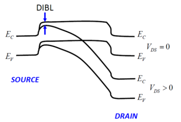
Source depletion region

Drain depletion region

Drain-induced barrier lowering (DIBL):

- Bei grosser Drain/Source-Spannung dehnt sich die Drain-Bulk pn-junction depletion region sehr weit aus und beginnt, mit derjenigen der Source zu interargieren
- Das heisst, der Channel wird praktisch ganz von den depletion regions um die Source und die Drain aufgefressen
- Die Energiehürde für ein Elektron, um über den Channel zu springen, sinkt plötzlich
- Daher sinkt die Threshold-Spannung ab.

Ein **kürzerer Channel** hat viele Vorteile: Man erhöht die Transconductance g_m sowie die Geschwindigkeit und reduziert die Baugrösse. Es treten allerdings mehr Nichtidealitäten, so genannte Short Channel Effects, auf.



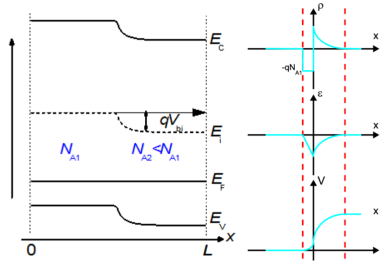
Gate Oxide Scaling

- Wir wollen den MOSFET möglichst mit der Gate-Spannung kontrollieren
- Reduzieren wir die **Dicke des Oxids** (SiO_2), so erhöht sich die Oxid-Kapazität C_{ox}. Dies verbessert die Steuerbarkeit des MOSFET.
- Sehr dünne Oxidschichten können allerdings von Elektronen durchtunnelt werden
- Um **Tunneling** zu verhindern, muss dann ein Oxid mit höherem κ (Leitfähigkeit) verwendet werden
- Die **equivalent oxide thickness** wird verwendet, um verschiedene Oxide zu vergleichen

$$EOT = d \cdot \frac{\varepsilon_{SiO_2}}{\varepsilon_{\mathrm{high}\kappa}}$$

EXERCISES

ISOTYPE JUNCTION



We consider a junction where only the doping level but not the doping type changes.

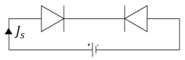
For example an acceptor doping level $N_{A1} > N_{A2}$ as shown in the figure:

- The concentration gradient leads to **diffusion** of holes **into the lower doped side.**
- Negatively charged acceptors stay on the **higher doped, left side.**
- The charge carrier density on the lower doped side is no longer set up by immobile dopants, but by holes which **diffuse.**
- The charge carrier density drops **exponentially.**

CLICKER QUESTIONS

BACK TO BACK DIODE

If there are two diodes back to back, then one of them is always reverse biased and the current flowing through the circuit is the reverse leakage current J_S



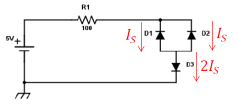
example:

Diodes $\boldsymbol{D1}$ & $\boldsymbol{D2}$ are reverse biased so they can only pass current $\boldsymbol{I_s}$

$\boldsymbol{D3}$ is forward biased:

$$\boldsymbol{I_{D3} = I_s\left(e^{\left(\frac{qV_{D3}}{kT}\right)} - 1\right) = 2I_s}$$

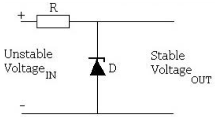
$$\boldsymbol{V_{D3} = \frac{kT}{q}\ln(3) = 28.56mV}$$



ZENER DIODE C|6.3

If we want to increase the output voltage we have to lower the doping level.

Because:
lower doping →lower electrical field in depletion region → higher reverse bias $V_R$is needed to achieve a cerain electrical field



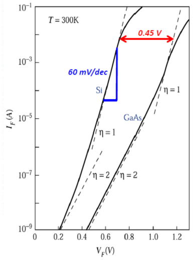
IDEALITY FACTOR C|6.4-6.6

lower = better
$\boldsymbol{\eta = 1 = \mathrm{ideal}}$

Plot: lin(V) vs log(I)

ex:
if I increase an ideal BJT ($\eta = 1$) by $60mV \rightarrow I$ increases 10x

ex:
ideality factor $\eta = 2$ corresponds to $120\frac{mV}{dec}$



MATHEMATICS

Dot product:

$$\vec{a}\cdot\vec{b} = |\vec{a}|\cdot|\vec{b}|\cdot\cos(\measuredangle[\vec{a},\vec{b}])$$

$$\measuredangle[\vec{a},\vec{b}] = \arccos\left(\frac{\vec{a}\cdot\vec{b}}{|\vec{a}|\cdot|\vec{b}|}\right)$$

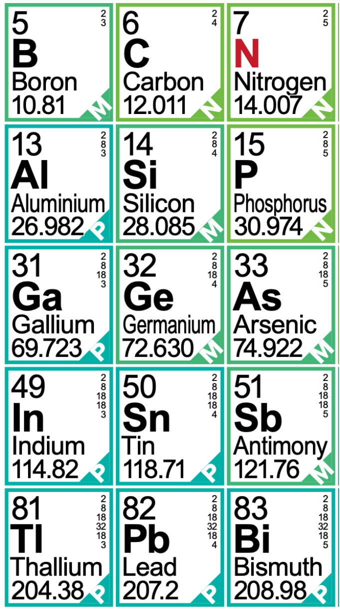
logarithm laws:

- $\log_b(P \cdot Q) = \log_b(P) + \log_b(Q)$
- $\log_b(P/Q) = \log_b(P) - \log_b(Q)$
- $\log_b(P^n) = n \cdot \log_b(P)$
- $\log_b(\sqrt[n]{P}) = \log_b(P)/n$
- $-\log_b(P) = \log_b(1/P)$
- $\log_a(P) = \frac{\log_b(P)}{\log_b(a)}$

Units

Frequency ν, f	Hertz	Hz	$\frac{1}{s}$
Pressure	Pascal	Pa	$\frac{N}{m^2} = \frac{kg}{m \cdot s^2}$
Power P	Watt	W	$\frac{J}{s} = \frac{m^2 \cdot kg}{s^3}$
Force F	Newton	N	$\frac{m \cdot kg}{s^2}$
Energy E	Joule	J	$N \cdot m = \frac{m^2 \cdot kg}{s^2}$
Drehmoment	Newton Meter $N \cdot m$		$\frac{m^2 \cdot kg}{s^2}$
El Current I	Ampere	A	$1A = 1C/s$
El. Resistance	Ohm	Ω	$1\Omega = 1V/A$
El Charge q, e	Coulomb	C	$1C = 1A \cdot s$
El Current density j			$1A/m^2$
El Charge density ρ			$1C/m^3$
El Voltage	Volt	V	$V = W/A$
El Field E			$1V/m$
Length	Angstrom Å		$10^{-8}cm$
Mass m	Kilogram	kg	$\frac{Ns^2}{m} = \frac{Js^2}{m^2}$
Temperature T	Kelvin	K	$1K = 1°C + 273$
Capacitance	Farad	F	$\frac{As}{V} = \frac{A^2 s^4}{kg \cdot m^2}$

PERIODIC TABLE, ELEMENTS OF INTEREST

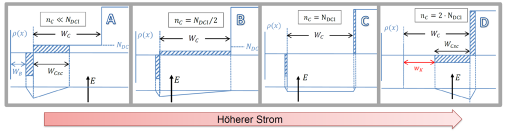


5 B Boron 10.81 2 3 M	6 C Carbon 12.011 2 4 N	7 N Nitrogen 14.007 2 5 N
13 Al Aluminium 26.982 2 8 3 P	14 Si Silicon 28.085 2 8 4 M	15 P Phosphorus 30.974 2 8 5 N
31 Ga Gallium 69.723 2 8 18 3 P	32 Ge Germanium 72.630 2 8 18 4 M	33 As Arsenic 74.922 2 8 18 5 M
49 In Indium 114.82 2 8 18 18 3 P	50 Sn Tin 118.71 2 8 18 18 4 P	51 Sb Antimony 121.76 2 8 18 18 5 M
81 Tl Thallium 204.38 2 8 18 32 18 3 P	82 Pb Lead 207.2 2 8 18 32 18 4 P	83 Bi Bismuth 208.98 2 8 18 32 18 5 P

KIRK EFFECT B|9.2

For high current levels, the charge of carriers travelling through the B/C depletion region modifies the electric field profile in the B/C depletion region.

→ at the B/C junction the E-field drops and eventually becomes 0
→ **base widens** ⇒ τ_B increases ⇒ β decreases
⇒ lowers f_T & f_{max} ⇒ lowers V_A
⇒ therefore the Kirk Effect is also referred as
"Base Spreading" or **"Base Pushout"**



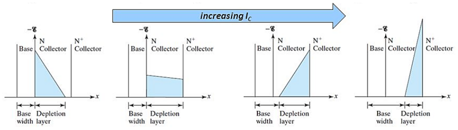
Wir beginnen mit einer one-sided junction zwischen Base und Collector, wie gehabt.

Signifikante Elekronendichte durch grossen Strom, kompensiert die Hälfte der Donator-Ladungen. Durch reduzierte effektive Dotierung verbreitert sich die depletion region

Noch höherer Strom kompensiert gesamte Donator-Ladungen. Depletion region verschwindet vollends

Es gibt einen «Elektronen-Stau» am Ende des Collectors, wodurch sich eine «umgekehrte» depletion region aufbaut.

PVK Halbleiterbauelemente
AMIV an der ETH Zürich



For high current levels the electron density n_C becomes therefore comparable to the donor density (npn BJT)

→ electron density cannot be neglected in calculations of the E_field.

(Poisson eqution:) $E(x) = \frac{q}{\varepsilon_s}(N_{DC} - n_C)x + E(0)$, $n_C = \frac{J_{nC}}{qv_{sat}}$

Kirk-Effect threshold current :
When the current gets higher than J_K, then the Kirk effect takes place and it result in a field inversion.

$$J_K = qv_{sat}\left(N_{DC} + \frac{2\varepsilon_s V_{BC}}{qW_C^2}\right)$$

The Kirk Effect can be reduced by making the collector doping N_{DC} higher or the collector width W_C smaller.
If we optimize for large J_K by increasing the collector doping, the Early Voltage decreases (=worsen) although!