Electron Devices Notes

Professor:
Monzio Compagnoni Christian

Student:
Matteo Baldo

January 16, 2017

...it's negligible...

Contents

Chapter 1

Semiconductors basics

1.1 Material distinction and energy gap

Distinction betweeen materials
In silicon-cristal electrons are subjected to periodical potential as a consequence to periodic disposition of atoms. We have bands of permitted energy level divided by gaps.
Importants bands are valence band VB ,that is the last band fully filled with electrons at 0 K, and conduction band CB ,that is the first band totally empty of electrons. Between this 2 bands there is the so called bandgap.
We can classificate all materials in 3 categories:
Metals - $E_{gap} = 0$ or "negative" and $\rho < 10^{-2}\Omega$cm .
Semiconductors - $E_{gap} = 1eV$ and $10^{-2} < \rho < 10^{5}\Omega$cm .
Isolators - $E_{gap} > 7eV$ and $\rho > 10^{5}\Omega$cm.

Energy gap function of T
In Si $E_{gap} = 1.12eV$ at room temperature (RT) but this value is a function of T; at high temperatures the silicon stretches and so does the periodical potential that influence the electrons.
The relation of E_{gap} with temperature is

$$E_{gap}(T) = E_{gap}(0) - \frac{\alpha T^2}{\beta + T} \tag{1.1}$$

where α and β change from material to material.
We can consider a sensitivity parameter of the temperature as

$$\frac{dE_{gap}(T)}{dT} = \frac{-2\alpha T(\beta + T) + \alpha T^2}{(\alpha + T)^2} \tag{1.2}$$

At RT for Si we have a change of 25meV over 100 degrees.

1.2 Silicon concentration

Conduction Band
Using the effective band approximation we want to know the densitiy of states in CB or VB and

after this the number of e or h in this bands.
Refering to the space of momentum we can say that

$$E - E_c = \frac{\hbar^2 k_x^2}{2m_x} + \frac{\hbar^2 k_y^2}{2m_y} + \frac{\hbar^2 k_z^2}{2m_z} \tag{1.3}$$

The iso-energetic surface in the space of momentum is an ellipsoid that is longer in the direction with effective mass higer. In the case 2 effective mass are equal we obtain a rotiation ellipsoide just as the case of Si in witch we have $m_x = m_z = 0.19m_0 = m_t$ and $m_y = 0.92m_0 = m_l$.
Observing that energy and the ellipsoide dimentions are directly proportional we can say that all the points inside the ellipsoide have less energy than border ones.
The volume of the ellipsoide is

$$\mathcal{V} = \frac{4}{3}\pi\sqrt{\frac{2m_x(E-E_c)}{\hbar^2}}\sqrt{\frac{2m_y(E-E_c)}{\hbar^2}}\sqrt{\frac{2m_z(E-E_c)}{\hbar^2}} = \frac{4}{3}\pi\frac{\sqrt{8m_xm_ym_z}}{\hbar^3}\sqrt{(E-E_c)^3} \tag{1.4}$$

So the density of states will be

$$N = \frac{\mathcal{V}}{(\frac{2\pi}{L})^3}\frac{1}{L^3} \tag{1.5}$$

Making thefirst derivative of the equation by dE we can obtain finally the density of states per unit energy

$$g_c(E) = \frac{4\pi}{h^3}\sqrt{2m_xm_ym_z}\sqrt{E-E_c}\cdot 2\cdot deg \tag{1.6}$$

where the last 2 terms are a corrective coefficients for the spin and the degeneration of the material we consider (for silicon 6).
We can say how many states are occupied only in a very specific case of thermodinamical equilibrium[1]. If this condition is satisfacted we can use the Fermi-Dirac statistic

$$f(E) = \frac{1}{1 + e^{(E-E_f)/kT}} \tag{1.7}$$

where E_f is the Fermi level defined as the level at witch the prbability of occupation of an energy state by an electron is 1/2. $f(E)$ makes a smooth transition from 1 to 0 as the energy increases; the width of the transition is governed by kT. That can be approximated with the Maxwell-Bolzmann statistic if $E - E_f >> kT$

$$f(E) \simeq e^{-\frac{E-E_f}{kT}} \tag{1.8}$$

This is a good approximation if we are at least well above E_f at least a few kT.
Concentretion of electrons in CB is (under the condition of thermodinamic equilibrium)

$$n = \int_{E_c}^{+\infty} g_c(E)f(E)dE \tag{1.9}$$

Making a change of variable as $x = (E - E_c)/kT$ and $\eta = (E_f - E_c)/kT$ we can write

$$n = \left[\frac{4\cdot 2\cdot deg\cdot\pi}{h^3}(kT)^{3/2}\frac{\sqrt{\pi}}{2}\right]\frac{2}{\sqrt{\pi}}\int_0^{+\infty}\frac{\sqrt{x}}{1+e^{x-\eta}}dx = N_c\cdot F_{1/2}(\eta) \tag{1.10}$$

[1] In thermodynamic equilibrium there are no net macroscopic flows of matter or of energy, either within a system or between systems. In a system in its own state of internal thermodynamic equilibrium, no macroscopic change occurs.

where N_c is the density of states $F_{1/2}(\eta)$ is the Fermi-Dirac integral of order 1/2.
Assuming $x - \eta >> 1$ (that is the M-B approximation) we arrive at

$$n = N_c \frac{2}{\sqrt{\pi}} e^{\eta} \int_0^{+\infty} \sqrt{x} e^{-x} dx = N_c e^{\eta} = N_c e^{-\frac{E_c - E_f}{kT}} \tag{1.11}$$

Valence band
Valence band of silicon has 3 sub-bands: heavy hole band and light hole that stay at VB level and the split-off band that stays 44meV under the VB. This 3 band mass are isotopic so the iso-energetic surface in the k space are spheres; bigger the sphere bigger the effective mass. We take into account for calculations only heavy hole band:

$$E_v - E = \frac{\hbar^2 k^2}{2m_{hh}} \rightarrow \hat{k}_{hh} = \frac{\sqrt{2m_{hh}(E_v - E)}}{\hbar} \tag{1.12}$$

as done before we calculate the volume of the sphere and the number of state per unit volume

$$\mathcal{V} = 4/3 \frac{\pi}{\hbar^3} \sqrt{8m_{hh}^3} \sqrt{(E_v - E)^3} \tag{1.13}$$

$$N = \frac{\mathcal{V}}{\frac{2\pi}{L}} \frac{1}{L^3} \tag{1.14}$$

we can calculate the density of states as

$$g_V(E) = \frac{4\pi}{3h^3} \sqrt{2m_{hh}^3} \sqrt{E_v - E} \cdot 2 \cdot deg \tag{1.15}$$

we proceed as for the CB obtaining the total concentration per unit volume of holes that is

$$p = p_{hh} + p_{lh} + p_{so} \tag{1.16}$$

but the concentration of holes in the split-off band is negligible due to the distance from the valence band (the transition of 1-f(E) that stays in a few kT).

$$p \simeq N_v e^{-\frac{E_f - E_v}{kT}} \tag{1.17}$$

Intrinsic and estrinsic Silicon
For intrinsic Si p=n and so we can calcuate E_f from this eq as

$$E_f = \frac{E_c + E_v}{2} - \frac{kT}{2} \ln(N_c/N_v) = E_i \tag{1.18}$$

E_i is the intrinsic level that stays in the middle of the gap (the correction is in the order of 0.3kT so negligible).
Putting 1.18 into p expression we obtain

$$p = n = \sqrt{N_c N_v} e^{-\frac{E_{gap}}{2kT}} = n_i \tag{1.19}$$

n_i is the intrinsic carrier concentration that has a strong dependance with T as it is involved in $N_c N_v$ and in E_{gap}.
From 1.11 by adding and subtracting E_i and dividing the exponential in 2 parts we can write

$$n = n_i e^{(E_f - E_i)/kT} \qquad p = n_i e^{(E_i - E_f)/kT} \tag{1.20}$$

this are 2 expressions valid in general not only for intrinsic semiconductors.
We can introduce the law of mass action as

$$pn = n_i^2 \tag{1.21}$$

For estrinsic semiconductors ,in the band diagram ,there is another band near to CB (E_d) (or near to VB (E_a) depending on the type of dope).
We want to extimate the position of E_f in n-doped Si. We can write that $n = p + N_d^+$ where N_d^+ are the ionized donor concentration so

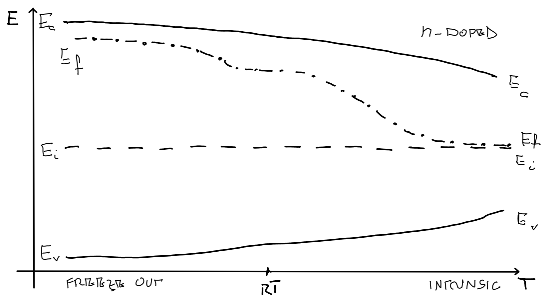
$$n = p + N_d(1 - \frac{1}{1 + \frac{1}{2}e^{(E_d - E_f)/kT}}) = p + \frac{N_d}{1 + 2e^{-(E_d - E_f)/kT}} \tag{1.22}$$

the factor 1/2 it's a spin correction coefficient. At RT $E_d - E_f >> kT$ and p concentration is negligible in comparison with N_d so $n \simeq N_d$ and from the law of mass action $p = n_i^2/N_d$. From this we can compute E_f as

$$E_c - E_f = kT \ln(N_c/N_d) \tag{1.23}$$

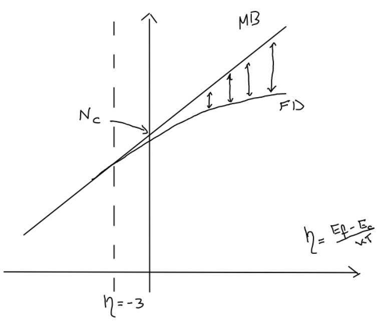
Moving E_f upwards some hypothesis falls: when E_f reachs E_d the complete ionization theory is not true anymore and from there further the M-B approximation is no longer valid.

Fermi level dependence on temperature in doped semiconductors



With high T E_g decreases and E_f will move distant to E_c following 1.23 until the hypothesis of negligible holes concentration decades and so at very high temperatures E_f will tend asimptotically to E_i.
At very low temperature E_f will increase until we reach the 0K where there is no conduction so E_f has to stay a little bit higher than E_d.

Low-temperature approssimation



In the case of a material that has its Fermi level higher than the conduction band level we cannot use the M-B approximaton. The approximation we can do is a low-temperature one that trasforms the Fermi Dirac distribution into a step so

$$n = \int_{E_c}^{E_f} \frac{48\pi}{h^3} \sqrt{2m_t^2 m_l} \sqrt{E - E_c} \tag{1.24}$$

that multiplying and dividing the risult by $(kT)^{3/2}\sqrt{\pi}/2$ we obtain

$$n = N_c 2/3 \frac{2}{\sqrt{\pi}} \left(\frac{E_f - E_c}{kT} \right)^{3/2} = N_c 2/3 \frac{2}{\sqrt{\pi}} \eta^{3/2} \tag{1.25}$$

In the graph below we can notice the difference between F-D and M-B distribution over η.

1.3 Current transport

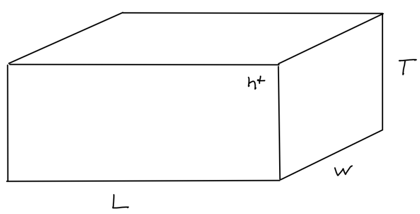
There are 2 most important mechanisms that generates current: drift and diffusion process.
Drift current is caused by the application of a electric field F; electrons are not only influenced by F but olso from scattering events so we can define a drift velocity (that is an average velocity) as

$$v_d = \mu_n F \tag{1.26}$$

So drift velocity is proportional to F with a constant μ_n called mobility that depends on doping concentration, temperature and dimensionality of the system taking in account all scattering events.It's important to note that in a bulk of doped Si if we are near the surface the mobility is much lower than the inside (dimensionality).
Mobility has a drop for Si around doping concentration of $\simeq 10^6 cm^{-3}$ beacuse at that level scattering with impurities becomes dominant in respect of thermal scattering.
Increasing too much F we arrive at the phenomena of velocity saturation ($\simeq 10^7 cm/s$) caused by scattering with high energetic optical phonons.
The density current caused by drift process is

$$J = (qn\mu_n + qp\mu_p)F = \sigma F = F/\rho \tag{1.27}$$

where σ is the conductivity of the material and ρ is the resistivity.



When we calculate the resistivity of a block of doped silicon like in figure we define $\rho_{sh} = \rho/T$ as the sheet resistivity

$$R = \rho\frac{L}{WT} = \rho_{sh}\frac{W}{L} \tag{1.28}$$

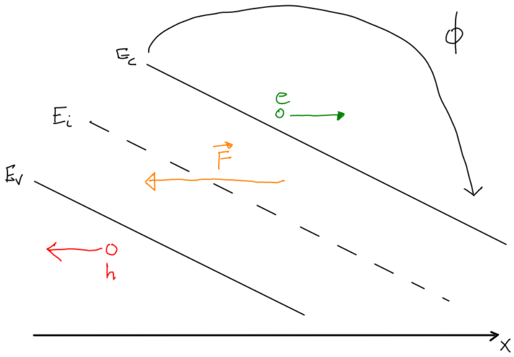
Diffusion current is driven by a gradient of concentration so the density of current is

$$J_n = qD_n\frac{dn}{dx} \quad J_p = qD_p\frac{dp}{dx} \tag{1.29}$$

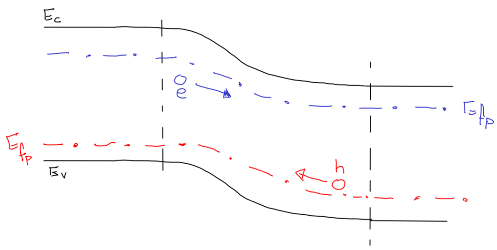
where D_n , D_p are the diffusion coefficients defined by Einstein's relations as

$$D_n = \mu_n\frac{kT}{q} \quad D_p = \mu_p\frac{kT}{q} \tag{1.30}$$

1.4 Bands "geography"



We can obtain external potential ϕ as $\phi = \frac{-E_i(x)}{q}$.
The external potential increases in the direction the bands band downwards. The electric field is in the opposite direction of the growth of potential $\vec{F} = -\frac{d\phi}{dx}$.
Electrons move by drift in the opposite direction of F and holes in the same direction of F.



Without condition of thermal equilibrium electrons move from regions of high quasi fermi level to regions of low quasi fermi level and holes vice-versa.

1.5 Poisson equation

Poisson equation is important to study the electrostatic of a system.
We can derive it from Maxwell's first law

$$F = -\vec{\nabla}\phi \rightarrow F = -\frac{d\phi}{dx} \tag{1.31}$$

and Gauss equation

$$\vec{\nabla}\vec{F} = \rho/\varepsilon_{Si} \tag{1.32}$$

from witch we obtain considering that all charges we have in a semiconductor are holes, electrons and ionized donors

$$\frac{d^2\phi}{dx^2} = -\rho/\varepsilon_{Si} = -\frac{q}{\varepsilon_{Si}}(p - n + N_d^+ - N_a^-) \tag{1.33}$$

In order to study the electrostatic of a system we have olso to know the coninutity equations of the electric field that are

$$F_{t1} = F_{t2} \Longleftrightarrow \epsilon_1 F_{n1} - \epsilon_2 F_{n2} = Q'_{int} \tag{1.34}$$

1.6 Debye length

Band banding is the phenomena caused by applying an electric field on a material at termodinamic equilibrium. This band banding causes different concentrations of e an h over the space so we can re-write the electron and holes concentrations taking in account the change of $\phi(x)$ over the space as

$$n = n_i e^{q\frac{\phi(x)-\phi_f}{kT}} \qquad p = n_i e^{q\frac{\phi_f-\phi(x)}{kT}} \tag{1.35}$$

This band banding is tipically caused by the change of doping concentration over the space.With a smooth change of N_d in a material we can assume that $n = N_d$.
Let's assume a step-like function in doping concentration in this case we cannot consider $n = N_d$ over the space beacuse this assumption will lead us to an infinite electric field.
Using Poisson equation and neglecting holes and ionized acceptors we can say that $N_d(x) = \hat{N}_d + \delta N_d(x)$ corrisponding to $\phi(x) = \hat{\phi} + \delta\phi(x)$ and assuming $\delta N_d << N_d$ we can write

$$\frac{d^2\phi}{dx^2} = -\rho/\varepsilon_{Si} = -\frac{q}{\varepsilon_{Si}}(\hat{N}_d + \delta N_d(x) - n_i e^{q\frac{\hat{\phi}+\delta\phi(x)-\phi_f}{kT}}) = -\frac{q}{\varepsilon_{Si}}(\hat{N}_d + \delta N_d(x) - \hat{N}_d e^{q\frac{\delta\phi(x)}{kT}}) \tag{1.36}$$

we can write the first order expansion of the exponential term beacuse the exponent is small ($e^{q\frac{\delta\phi(x)}{kT}} \simeq 1 + q\frac{\delta\phi(x)}{kT}$) obtaining the following differential equation

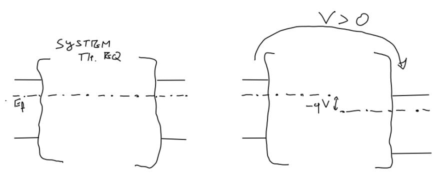
$$\frac{d^2\phi}{dx^2} = \frac{q^2\hat{N}_d}{\varepsilon_{Si}kT}\phi(x) - \frac{q}{\varepsilon_{Si}}\delta N_d(x) \tag{1.37}$$

the solution of this equation give us an exponential decrease of the potential with a length called Debye length

$$L_D = \sqrt{\frac{\varepsilon_{Si}kT}{q^2N_d}} \tag{1.38}$$

this will cause a net charge (both negative and positive near the discontinuity).
$[Look at exercitation there is the demonstration of this formula]$

1.7 Quasi-Fermi levels



For having a net current transport (we're leaving the thermodinamic equilibrium hypotesis) we have to perturbe our system by for example applying at the contact [2] a $V \neq 0$. As the figure shows we can't introduce a E_f level and so we cannot use the FD distribution. If the perturbation is weak we can recover all expressions of th.eq by introducing different Fermi levels both for e and h called quasi Fermi levels

$$n = n_i e^{\frac{E_{fn}-E_i}{kT}} \qquad p = n_i e^{\frac{E_i-E_{fp}}{kT}} \tag{1.39}$$

so we can write the law of mass action generalized as

$$pn = n_i^2 e^{\frac{E_{fn}-E_{fp}}{kT}} \tag{1.40}$$

[2]region of dispositive from witch we can externally apply a F, a good contact stays at th. eq.

Now let's try to analise the total current density of electrons. The formula has a strong dependance from the electrostatic potential (from F and n) so

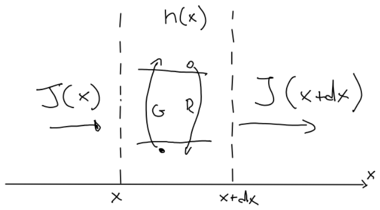
$$J_n = -qn\mu_n F + qD_n \frac{dn}{dx} = -qn\mu_n \frac{d\phi}{dx} + qD_n \frac{d}{dx}\left(n_i e^{q\frac{\phi-\phi_f}{kT}}\right) = -qn\mu_n \frac{d\phi_f}{dx} \tag{1.41}$$

But since the Fermi potential is costant the result is 0 and it's obvius beacuse we assume th.eq. with formula (1.20). If we substitute the (1.39) in n we obtain

$$J_n = -qn\mu_n \frac{d\phi_{fn}}{dx} \tag{1.42}$$

that isn't 0 beacuse ϕ_{fn} can change.
Electrons moves from the region of high quasi Fermi level to region of low quasi Fermi level, holes the opposite.

1.8 Equation of for e and h



Let's consider the figure above if we make a ballace of charges between x and x+dx we have

$$\frac{\partial n}{\partial t} dx = -J_n(x)/q + J_n(x+dx)/q + (G-R)dx \tag{1.43}$$

where G and R are generation and ricombination processes per unit time per unit volume. If we expand at the first order $J(x+dx) \simeq J_n(x) + \frac{\partial J_n}{\partial x} dx$ and simplify the expression we have

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} + G - R \qquad \frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} + G - R \tag{1.44}$$

that are the continuity equation for electrons and hole current.
In a stationary system nothing change with time so subtractring the 2 equation we have

$$\frac{1}{q}\frac{\partial (J_n + J_p)}{\partial x} = 0 \tag{1.45}$$

so the sum of the 2 contributions of current from e and h stay costant. If we neglect G and R the 2 contrbutes are costant separatly.
G and R process restabilize equilibrium of a system if we perturb the minority carrier concentration. If we disturb majority carrier the system returns in equilibrium with a very short time costant. Let's analyse the 2 case.

Majority carrier perturbation

Assume a n-doped material and let us add Δn concentration of electrons. Using poisson equation we have

$$\frac{\partial \phi^2}{\partial x^2} = -\frac{q}{\varepsilon}\left(-n - \Delta n + N_d\right) = \frac{q\Delta n}{\varepsilon} = -\frac{dF}{dx} \tag{1.46}$$

from this result we obtain a field that has a linear dependence over the space that tends to move away the excess of charge with a drift current (no diffusion the concentration is costant) so using 1.44

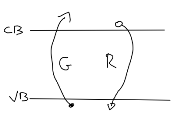
$$\frac{\partial J_n}{\partial x} = qn\mu_n \frac{\partial F}{\partial x} \tag{1.47}$$

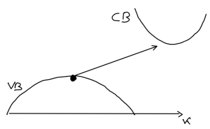
$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} = -\frac{1}{q} qn\mu_n \frac{q\Delta n}{\varepsilon} = -\frac{\Delta n}{\varepsilon\rho} \tag{1.48}$$

solving this differential equation we have a solution that has an exponential dependence with $\varepsilon\rho$ that is called the dialectric relaxation time of the material; it's the time costant that the system use to return to its equilibrium ($\simeq ps$)

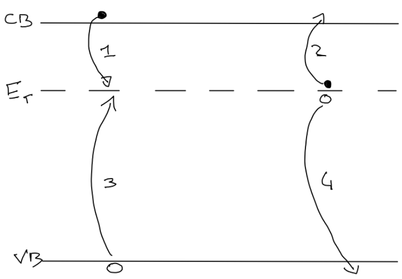
Minority carrier perturbation

The phenomena described for majority carrier concentration isn't valid for minority carrier. Considering a n-doped material like before the ρ in equation 1.48 is very high (we are considering holes now) so the real result of the field is that attracts a concentration Δn over Δp creating a quasi-neutral region where G and R slowly decrease the 2 extra concentration. In the processes of G and R in Si electrons cannot move thought the bandgap from CB to VB they need another particle to respect the conservation of momentum.





We need a difect of the cristal that creates another "band" in the bandgap this process called difect assisted and is described by Shockley-Reed-Hall theory. The difect can only be empty or filled with one electron so there can be only 4 process: 1)Difect empty captures e from CB 2) Difect filled realeses an e to CB 3) Difect filled captures hole from VB 4) Difect empty realeses an hole from VB.



For R we need 1+3 for G 2+4.
The rate of 1 can be described as

$$r_1 = N_T(1 - f(E_T))n(v_{th}\delta_n) \tag{1.49}$$

where: N_T is the total difect concentration ,f probability of the difect to be filled and the last term is a costant where δ_n is the capture crossection. The rate of process 2 can be described as

$$r_2 = N_T f(E_T) e_n \tag{1.50}$$

where e_n is a proportionality costant emission rate for e. Under th.eq. $r_1 = r_2$ and we can use the FD distribution for f. Under this condition we can calculate e_n (and using the same consideration olso for process 3 and 4 and so obtaining e_p)

$$e_n = n_i v_{th}\delta_n e^{\frac{E_T - E_i}{kT}} \qquad e_p = n_i v_{th}\delta_p e^{\frac{E_i - E_T}{kT}} \tag{1.51}$$

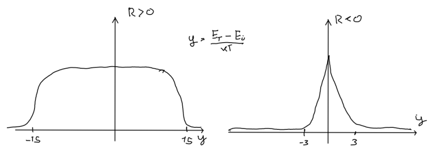
There is an exponential dependance if the difect is colser to the conduction band the process of release is higher; vice-versa the process of capture doesn't have this dependance beacuse the e has only to realese energy.

Under stationary condition $R = r_1 - r_2 = r_3 - r_4$ from this we obtain $f(E_T)$ and we get the final result

$$R = \frac{pn - n_i}{\tau_0 \left[p + n + 2n_i Cosh(\frac{E_T - E_i}{kT})\right]} \tag{1.52}$$

where $\tau_0 = 1/(v_{th} N_T \delta)$ and we've assumed $\delta_n = \delta_p = \delta$.
Re-writing the numerator using $n = n_i e^{(E_{fn} - E_i)/kT}$ $p = n_i e^{(E_{fp} - E_i)/kT}$ we can obtain that if $E_{fn} > Efp \rightarrow R > 0$ and if $E_{fn} < Efp \rightarrow R < 0$.

Looking at the graph of R in the 2 case major of less than zero we can make a few considerations



$R > 0$ relevant part in the denominator is p+n but over 15 the Cosh becomes dominant so we get the drop.

$R < 0$ Cosh is always dominant so we have from the beginning a rapid drop beacuse p+n are few and the system is generating couples.

Phisical explanation can be given from a particular case.
We want to study only recombination so $R = \nu_1 = \nu_3 = (1 - f(E_T))n = f(E_T)p$ so we can caluclate $f(E_T) = n/(n+p)$ and so $R = N_T v_{th} \delta \frac{pn}{p+n} = \frac{pn}{\tau_0(n+p)}$ that is a semplification of the general formula we have a flat recombination rate beacuse 1,3 processes have only to loose energy the drop is caused by the growing relevance of 2 or 4 moving E_T near to CB of VB.
The best position fot the difect is on the middle of the gap.
Maximum G is when p=n=0 so $G = -n_i/2\tau_0$.

1.9 Quasi-neutral condition

Under quasi-neutral condition $\Delta n + n_0 = n$ and $\Delta p + p_0 = p$ assuming low injection level ($\Delta n << p_0 + n_0$) we can neglect in the eq of R all the terms with "Δ" so we obtain

$$R = \frac{\Delta n}{\tau_n} = \frac{(p_0 + n_0)\Delta n}{\tau \left[p_0 + n_0 + 2n_i Cosh(\frac{E_T - E_i}{kT})\right]} \tag{1.53}$$

if we suppose $E_T - E_i = 0$ $\tau_n \simeq \tau_0 = 1/(N_T v_{th} \delta)$
So the continuity equation for electrons becomes

$$-\frac{d\Delta n}{dt} = R \rightarrow \Delta n \propto e^{-t/\tau_n} \tag{1.54}$$

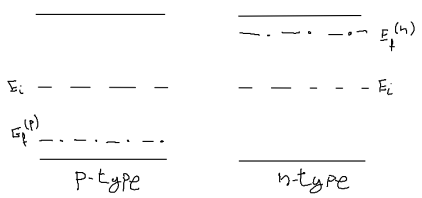
where τ_n is dipendent on the quality of the material

Chapter 2

Diodes

Diodes or pn junction are basically 2 zones of opposite doping polarity attached toghether. To study pn junctions we have to make some simplyfing assumptions : we consider 1D materials with constant doping concentration and a step-like change of doping.

2.1 Built-in potential



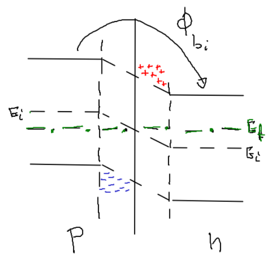
Let's consider the 2 regions isolated from each other and under th.eq. We don't have a single Fermi level so depending on the zone we have

$$E_i - E_f^{(n)} = kT\ln(\frac{N_a}{n_i}) \qquad E_f^{(n)} - E_i = kT\ln(\frac{N_d}{n_i}) \quad (2.1)$$

where $E_f^{(n/p)}$ is the Fermi level in that zone.
When we put the 2 materials toghether there will be first a diffusion current and than olso a drift current due to the band banding and the exposed charge; we will reach th.eq when the total current will be 0.
Under th.eq we will have a costant Fermi level all over the space and a band banding at the interface. Far from the intarface we will reach zones where the bands are like the isolated case. The drop of electrostatic potential in the junction under th.eq is called built in potential or ϕ_{bi} and is

$$q\phi_{bi} = (E_f^{(n)} - E_f^{(p)})|_{n.e} = kT\ln(\frac{N_a N_d}{n_i^2}) \qquad (2.2)$$

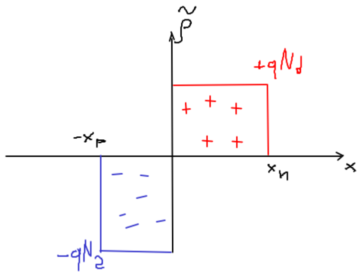
that is close to the E_g of the material.



2.2 Depletion approximation

Beacuse the distance $E_f - E_c$ increase near the junction we have a exponential deacrease of n so in the transition region we can say that $n << N_d$ and so $p << N_a$. From this we can say that the transition region is depleted of free carriers beacuse their concentration is negligible so this region is called depletion region.

2.3 Electrostatics of a pn junction

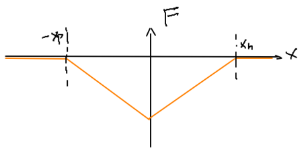


We have to solve poisson equation in the depletion region so using depletion approximation and considering complete ionization

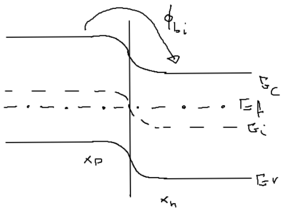
$$\frac{d\phi^2}{dx^2} = -\frac{q}{\varepsilon}(N_d - N_a) \tag{2.3}$$

and considering the concentration of fixed charges like the graph we can split the Poisson eq in 2 parts for $-x_p < x < 0$ we have $\frac{d\phi^2}{dx^2} = \frac{q}{\varepsilon}N_a$ and for $0 < x < x_n$ we have $\frac{d\phi^2}{dx^2} = -\frac{q}{\varepsilon}N_d$. If we integrate both side of this equation and remembering that $F(x_n) = 0$ we can deduce the electric field in the junction as

$$\int_x^{x_n} d\frac{d\phi}{dx} = \int_x^{x_n} -\frac{q}{\varepsilon}N_d dx \tag{2.4}$$



$$F(-x_p < x < 0) = -\frac{qN_a}{\varepsilon}(x_p + x) \qquad F(0 < x < x_n) = -\frac{qN_d}{\varepsilon}(x_n - x) \tag{2.5}$$



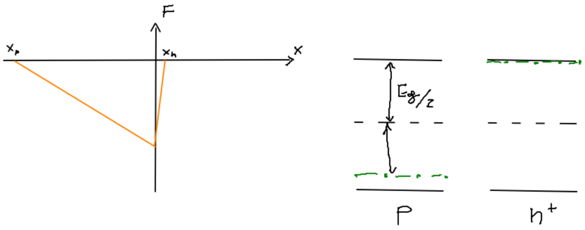
for continuity equations of the electric field the F must be continuos in 0 so from this condition we get that $N_a x_p = N_d x_n$ that means that the total charge in the depletion area is 0 (from Gauss law olso). Integrating again both parts of the equation of the field we finally obtain the potential over the space that has a parabolic dependance

$$\phi(-x_p < x < 0) = \phi(-x_p) + \frac{qN_a}{2\varepsilon}(x + x_p)^2 \qquad \phi(0 < x < x_n) = \phi(x_n) - \frac{qN_d}{2\varepsilon}(x_n - x)^2 \tag{2.6}$$

To know x_n and x_p we have to set some boundary condition: ϕ must be continuos in 0, and $N_a x_p = N_d x_n$ from this 2 condition we get

$$x_n = \sqrt{\frac{2\varepsilon}{q}\phi_{bi}(\frac{1}{N_a} + \frac{1}{N_d})} \times \frac{N_a}{N_a + N_d} \qquad x_p = \sqrt{\frac{2\varepsilon}{q}\phi_{bi}(\frac{1}{N_a} + \frac{1}{N_d})} \times \frac{N_d}{N_a + N_d} \qquad W = \sqrt{\frac{2\varepsilon}{q}\phi_{bi}(\frac{1}{N_a} + \frac{1}{N_d})} \tag{2.7}$$

2.4 Unilateral Junction

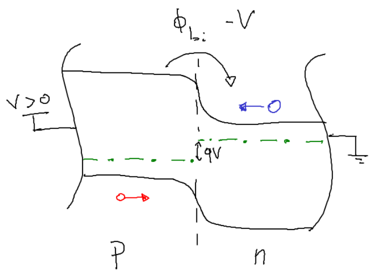


Let's consider a junction n^+p so the depletion area is almost all in the less doped material and so the electric field. We can make some approximation that the depletion area is almost in the p zone and that the Fermi level of the n^+ zone is almost at E_c so this lead us to

$$W \simeq x_p \qquad q\phi_{bi} \simeq E_g/2 + kT\ln(\frac{N_a}{n_i}) \tag{2.8}$$

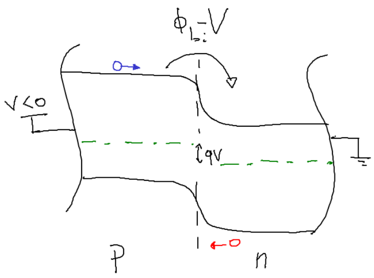
2.5 Bias

Forward Bias



Applying a external voltage like in figure we are reducing the total voltage drop over the junction. There will be a net flow of electrons from left to right that is a current by diffusion that will be large beacuse we have a lot of charges.

Invers Bias

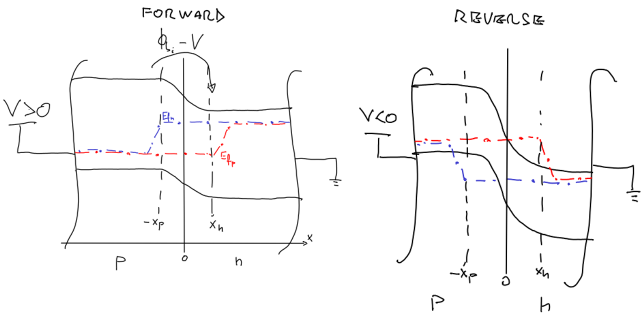


In this case $V < 0$ electrons and holes move for drift the total voltae drop is will be greater than the built in. The flow of current is given by minority carrier so small I.

2.6 The V-I characteristic of pn

First we want to investigate on foward bias electrostatic but we have to make some other simplyfing assumption: the contacts of our junction are at th.eq. and looking to the n region moving closer to the contact we will enter in a region where electrons are majority carrier. Beacuse we are injecting holes in that region we will have $p = p_{n0} + \Delta p$ so a region of quasi-equilibrium.
If we assume stationarity and that G and R process are negligible and so the $J_n = n\mu_n \frac{dE_{fn}}{dx} = cost$ so beacuse n isn't costant all over the space $\frac{dE_{fn}}{dx}$ will change over the

space. The final graph is this



we have 2 region of quasi-equilibrium and a transition region that is depleted from free charges. This is the same result of electrostatic without bias so the depletion region will be

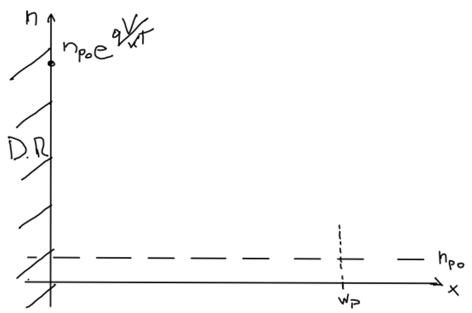
$$W = \sqrt{\frac{2\varepsilon}{q}(\phi_{bi} - V)(\frac{1}{N_a} + \frac{1}{N_d})} \tag{2.9}$$

E_{fn} remains costant until $-x_p$ but n changes beacuse E_c moves upwards.
For reverse bias we have to follow the same path of forward bias arriving at the graph above. Bands are flat near the contact beacuse we suppose (1) quasi-neutral region under (2) low level of injection with (3) costant doping and (4) negligible gradient of quasi-Fermi levels; if one of this supposition falls the bands will no more be flat.

Now we want to know the quasi-Fermi level in the minority region. We know that at $x = x_n$ $n \simeq N_d$ and for the law of mass action generalized $p = \frac{n_i^2}{n} e^{\frac{E_{fn}-E_{fp}}{kT}} \simeq p_{n0} e^{\frac{qV}{kT}}$ and at $x = -x_p$ with the same passages we obtain $n \simeq n_{p0} e^{\frac{qV}{kT}}$. Now we can solve the continuity equation only in the quasi-neutral region.
Beacuse the bands are flat we have only diffusion so $J_n = qD_n \frac{dn}{dx}$ and $G - R = \frac{-\Delta n}{\tau_n}$ from this two in the coninuity equations under stationary condition we get

$$\frac{d^2 \Delta n}{dx^2} - \frac{\Delta n}{L_n^2} = 0 \tag{2.10}$$

where $L_n = \sqrt{D_n \tau_n}$ = diffusion length of electrons in the quasi neutral p-region. Introducing the system of coordinate like in figure (where W_p is the end of the quasi-neutral p region) we can solve this differential equation.



Solving that as a polynomial we get $\lambda^2 - \frac{1}{L_n^2} = 0 \leftarrow \lambda = \pm\frac{1}{L_n}$ so we get two exponential terms that are $\Delta n(x) = Ae^{\frac{x}{L_n}} + Be^{-\frac{x}{L_n}}$ using the boundary condition $\Delta n(0) = n_{p0}(e^{\frac{qV}{kT}} - 1)$ and $\Delta n(W_p) = 0$ we arrive at

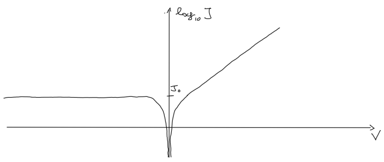
$$\Delta n(x) = n_{p0}(e^{\frac{qV}{kT}} - 1)\frac{\sinh(\frac{W_p - x}{L_n})}{\sinh(\frac{W_p}{L_n})} \tag{2.11}$$

Having only diffusion processes we can know the current but we are intrested only in $J_n(0)$ at the edge of the quasi-neutral region so

$$J_n(0) = (qD_n \frac{d\Delta n}{dx})|_{x=0} = -\frac{qD_n n_{p0}}{L_n \tanh(W_p/L_n)}(e^{\frac{qV}{kT}} - 1) \tag{2.12}$$

so in general the total current in the pn junction is

$$J = [\frac{qD_n n_{p0}}{L_n \tanh(W_p/L_n)} + \frac{qD_p p_{n0}}{L_p \tanh(W_n/L_p)}](e^{\frac{qV}{kT}} - 1) = J_0(e^{\frac{qV}{kT}} - 1) \tag{2.13}$$

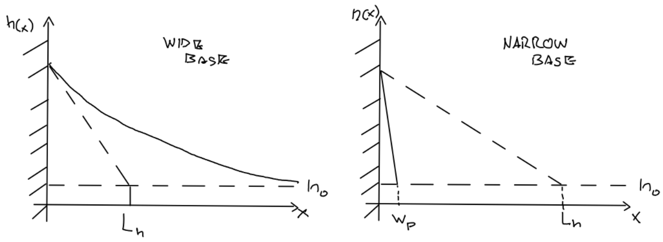


If $V >> kT/q$ we can neglect 1 and plot a log-log scale of this characteristic that is

$$\log(J) = log(J_0) + \frac{qV}{kT \ln(10)} \tag{2.14}$$

One important parameter of the characteristic is the slope in foward bias that is $kT/q \ln(10) = 60mV/dec$ @RT. The dominant contribute to the current is from the zone less doped like in the electrostatic.

2.7 Wide and narrow base diode

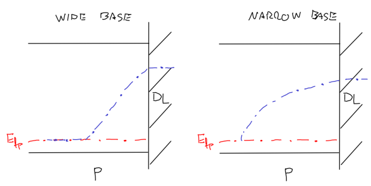


If we have a wide-base diode if $W_p >> L_n$ so we can approximate as follow the expression of charge and of current density

$$\Delta n(x) = n_{p0}(e^{\frac{qV}{kT}} - 1)e^{\frac{-x}{L_n}} \qquad J_n = \frac{qD_n n_{p0}(e^{\frac{qV}{kT}} - 1)}{L_n} \tag{2.15}$$

If we have a narrow-base diode if $W_p << L_n$ so we can approximate as follow the expression of charge and of current density

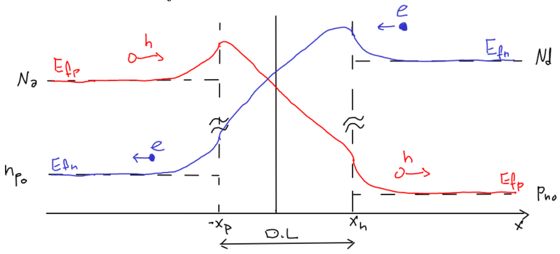
$$\Delta n(x) = n_{p0}(e^{\frac{qV}{kT}} - 1)\frac{W_p - x}{W_p} \qquad J_n = \frac{qD_n n_{p0}(e^{\frac{qV}{kT}} - 1)}{W_n} \tag{2.16}$$



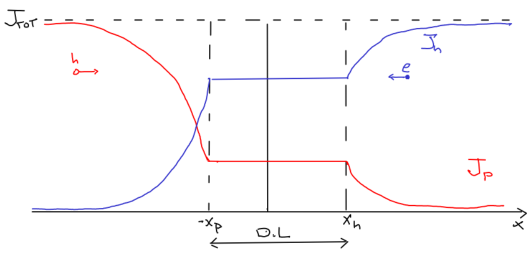
In a wide base diode n(x) decrease exponentially in the p region therefore E_{fn} will decrease linearly with a slope of kT/L_n and after the intersection with E_{fn} it will be costant as this last one. For a narrow base the decrease of E_{fn} will be logaritmic beacuse n(x) decreases linearly. In the quasi neutral p (n) region we have less electrons (holes) than in all other zone so the highest resistivity. The limit for conduction is given by the quasi-neutral region where we have less concentration of minority carriers; this is why the pn junction is called a minority carriers device , the limit for current transport is given by minority carriers.

2.7.1 Current profile

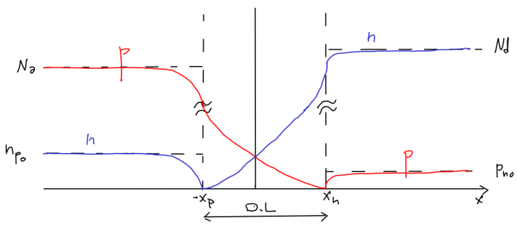
Wide base forward bias
The exponential behaviour of the minority will be followed by a consequent increase of majority charges in order to mantain the region quasi-neutral. In the depletion layer a connection between the 2 region. Minority carriers moves for diffusion but majority carrier moves for drift beacuse there is a small gradient of E_{fn} that let electrons move from right to left.



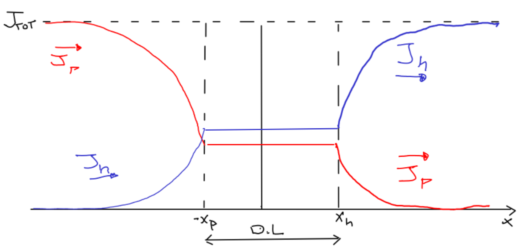
To know J_n in n region we have to remember that under stationary condition (so in quasi neutral region) $J_n + J_p = const$. Olso we remember that under stationary conditions and neglecting G-R both current contributions are constant. From this considerations we can draw the graph of J_n, J_p over the space.



Wide base reverse bias
Same path of before : minority carrier in quasi neutral regions decrease before the depletion layer so this phenomena will be copied by the majority carrier in order to mantain quasi-neutrality.



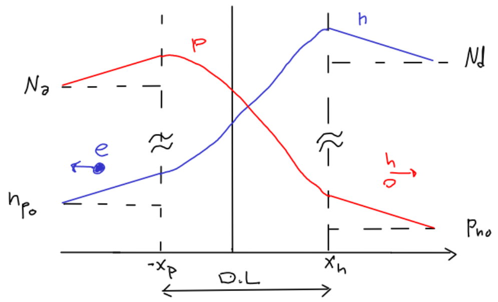
Knowing $J_n + J_p = const$ in stationary conditions and that if we neglect olso G-R process we have both currents constat we obtain the following graph.



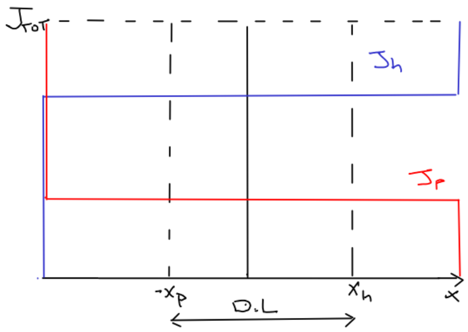
The flow of holes becomes zero at n contact but than in grows for generation processes.
We have the same behaviour of the current both in reverse and in forward bias but J_{tot} changes of orders of magnitude. As before minority carriers move by diffusion, majority by drift and in the depletion layer we have movement by drift.

Narrow base forward bias

Linear increase of minority carriers so olso majority in order to mantain the region quasi-neutral.

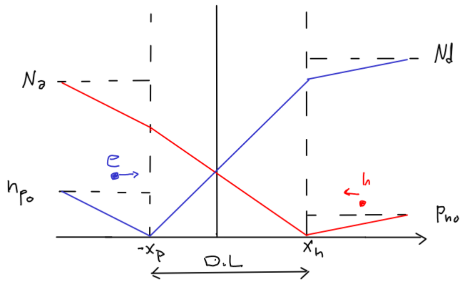


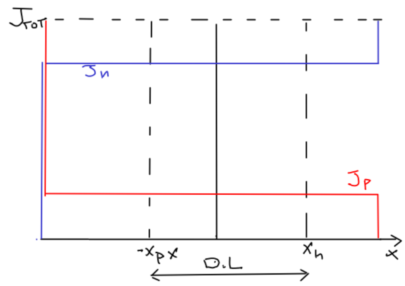
The derivate of a linear dependence is constat so J_n and J_p remain constat. There is no space to G-R process beacuse the device is much shorter than L_n so all generation and ricombination process are at the contacts.



Narrow base reverse bias

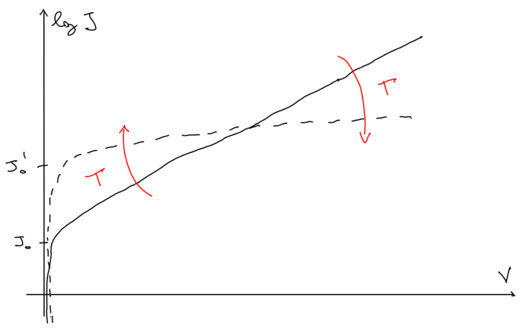
With the same path of before we arrive at





2.8 T change

Let's consider a pn junction in foward bias $J = J_0(e^{\frac{qV}{kT}} - 1) \simeq J_0 e^{\frac{qV}{kT}}$. If we increase the T the slope of the straight part will increase for the exponential term (be careful to units in the graph the line becomes flatter) but olso $J_0 = \frac{qD_n n_i^2}{N_a L_n}$ will increase for his strong dependance with temperature caused by $n_i^2 \propto T^3 e^{-\alpha/T}$.



So with the increase of T we can have both an increase or a decrease of voltage corrisponding a fixed J. We have to find the typical regime of our device so from $J = J_0 e^{\frac{qV}{kT}}$ we extract the voltage as $V = \frac{kT}{q}\ln(J/J_0)$ with J a fixed current. Now we cand derivate obtaining

$$\frac{dV}{dT} = \frac{k}{q}\ln(J/J_0) + \frac{kT}{q} J_0/J \frac{-J\frac{dJ_0}{dT}}{J_0^2} = V/T - \frac{kT}{q}\frac{dJ_0}{dT}\frac{1}{J_0} \tag{2.17}$$

for semplicity we can hilight J_0 dependences on temperature writing $J_0 = aT^{\gamma}e^{-E_g/kT}$ so we obtain that $\frac{dJ_0}{dT} = J_0\gamma/T + J_0[-\frac{dE_g}{dT}\frac{1}{kT}] + J_0\frac{E_g}{kT^2}$ so coming back at voltage we obtain

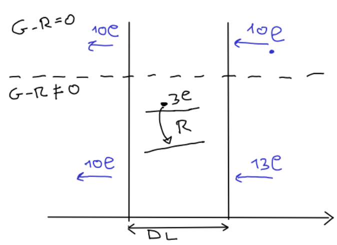
$$\frac{dV}{dT} = (V - \frac{E_g}{q})\frac{1}{T} - \frac{k\gamma}{q} + \frac{1}{q}\frac{dE_g}{dT} \tag{2.18}$$

Since our typical $V \simeq 0.6/0.7$ there are all negative terms so we are in the firs zone where we find a decrease of voltage w.r.t an increase of the temperature (at RT $dV/dT \simeq -1.9mV/K$)

2.9 Second order effects on current

There are some phenomena that we have to consider in order to make more realistic the J-V graph of a pn junction.

2.9.1 Low current regime



We have always neglect the G-R in the depletion layer but now we have to make some considerations. Under reverse bias $E_{fn} < E_{fp}$ so there will be some generation processes, knowing the quasi-Fermi level and using the law of mass action generalized we can write the SRH R coefficient as

$$R = \frac{n_i^2(e^{(qV)/(kT)} - 1)}{\tau_0(p + n + 2n_i Ch((E_t - E_i)/(kT))} \tag{2.19}$$

so the sign of the voltage applied V change the sign of R. Refering to forward bias the impact of R is that from majority region there will be more electron moving to the depletion region beacuse of minority constrain of the pn junction. So we aspect a larger current under foward bias where

$R \simeq \frac{n_i^2(e^{\frac{qV}{kT}}-1)}{\tau_0[p+n]}$ to get the worst case we take p=n so R will be maximum and we get $R = \frac{n_i(e^{\frac{qV}{2kT}})}{2\tau_0}$.
Under forward bias so we can define

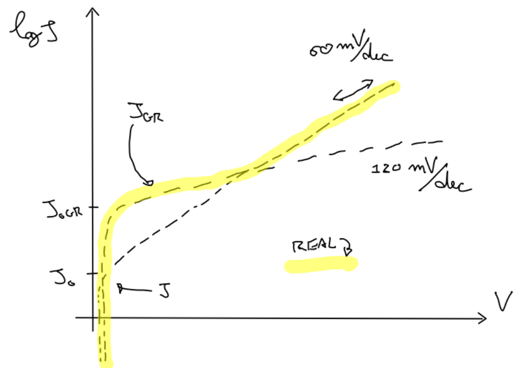
$$J_R = qR_{max}W_d = qW_d\frac{n_i(e^{\frac{qV}{2kT}})}{2\tau_0} \tag{2.20}$$

Under reverse bias $R \simeq -\frac{n_i^2}{2n_i\tau_0} = -\frac{n_i}{2\tau_0}$ and so as for forward bias we can define

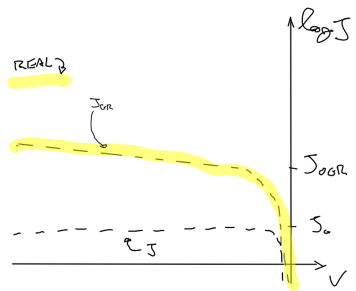
$$J_G = -q\frac{n_i}{2\tau_0}W_d \tag{2.21}$$

In general for both cases we can define

$$J_{GR} = \frac{qn_i}{2\tau_0}W_d(e^{\frac{qV}{2kT}}-1) = J_{0GR}(e^{\frac{qV}{2kT}}-1) \tag{2.22}$$



Similar to ideal but J_{0GR} is bigger and we have a factor 2 at the exp so there is a slight difference at very low current. Note that $J_{0GR} \propto 1/\tau_0$ so is very process dependent and olso that the ideal J_0 has a quadratic dependance on n_i (and this term has a linear dependance) so at high temperature we recover the ideal diode characteristic.
With reverse bias the current is dominated by J_{RG} and have a slight increase due to the dependence of W_d with V.



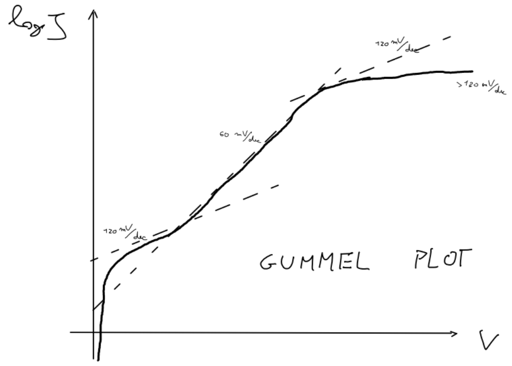
2.9.2 High current regime

At high current regime there are 2 problems that create distortion in the ideal characteristic.

High injection regime
At high current the hypotesis of low injection regime decades so taking p quasi-neutral region we have $n = n_0 + \Delta n \simeq \Delta n$ and $p = p_0 + \Delta n \simeq \Delta n$ so with the law of mass action generalized we obtain $pn = n_i^2 e^{\frac{qV}{kT}} = \Delta n^2$ and so $\Delta n = n_i e^{\frac{qV}{2kT}}$ from this through continuity equation (see 2.6) we get that

$$J \propto n_i e^{\frac{qV}{2kT}} \tag{2.23}$$

so an additional solpe of 120mv/dec.
Resistive drops
E_f gradient is no more negligible so we have parasitic resistance due to the law $J_n = n\mu_n \frac{dE_{fn}}{dx}$.
To introduce this non ideality we can say that $J = J_0 e^{\frac{qv}{mkT}}$ where m is a factor of non ideality.
What it follows is the Gummel plot of the pn junction that include all non idealities.



2.10 Small signal model

The pn junction is, of course, a non linear device if we assume that is polarized at $\overline{V} > 0$ and we change that voltage of σV we are interested in its response. We have 3 small signal parameters: conductance, depletion capacitance and diffusion capacitance.

2.10.1 Conductance

Changing the voltage of σV we have a variation of the current so a condctance that we can calculate as

$$g_m = \frac{\partial I}{\partial V} = \frac{I}{kt/q} \tag{2.24}$$

2.10.2 Deplition capacitance

Changing δV we remove a part of the depletion region as a consequence of majority injected in the junction that neutralize the fixed charges in the DL. We have a variation of the charge stored in the device as a consequence of a variation of the voltage so a capacitance.
We can introduce

$$C_{dep} = \frac{\partial Q_{DL}}{\partial V} = -\frac{\partial}{\partial V}(qN_d x_n) = \frac{\varepsilon_{si}}{W_d} \tag{2.25}$$

The minus sign is place in order to have a positive C and this formula is valid olso for non constant doping concentration.

2.10.3 Diffusion capacitance

Olso in the quasi-neutral region we have charges stored that change if σV change due to the excess of minority. This capacitance is called diffusion capacitance and is defined as

$$C_{diff} = \frac{\partial Q_{diff}}{\partial V} \tag{2.26}$$

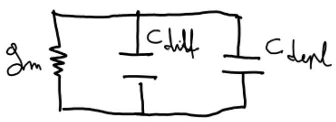
we have therefore define Q_{diff}. Assuming forward bias and so neglecting n_{p0} we can write $Q_{diff} = \int_0^{W_p} q\Delta n(x)dx$.
For a wide base diode we have

$$Q_{diff} = q\Delta n(0)\frac{L_n^2}{L_n}\frac{D_n}{D_n} = J_n\tau_n \tag{2.27}$$

For a narrow base diode

$$Q_{diff} = q\Delta n(0)\frac{W_p^2}{2W_p}\frac{D_n}{D_n} = J_n t_p \tag{2.28}$$

where t_p is defined as electron transit time through the quasi-neutral p region.



It's called like
this beacuse from the current equation we can derive that $J_n = qD_n\Delta n(0)/W_p = q\Delta n(x)v_{diff} = q\Delta n(0)\frac{W_p - x}{W_p}v_{diff}$ from this we obtain $v_{diff} = \frac{D_n}{W_p - x}$ and from this integrating from 0 to W_p in dx we obtain exactly $t_p = \frac{W_p^2}{2D_n}$.
Going back
to the diffusion capacitance we have for a wide base diode

$$C_{diff}^{wide} = g_m\tau_n \tag{2.29}$$

and for a narrow base

$$C_{diff}^{narrow} = g_m t_p \tag{2.30}$$

In the end the small signal model is showed in figure.
In reverse bias the modulation of the quasi-neutral change is negligible and so is the diffusion capacitance.

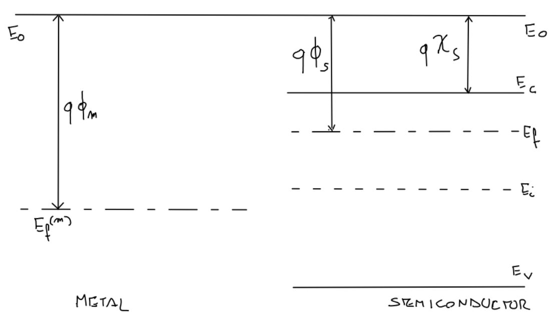
Chapter 3

Metal-semiconductor junction

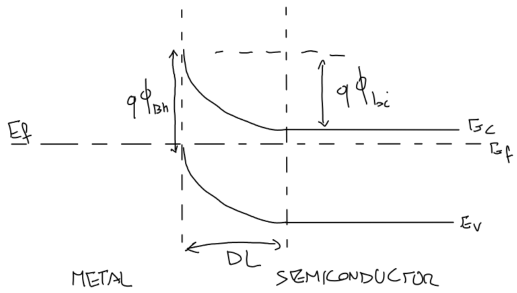
Metal-semiconductor junctions can be divided in 2 categories: ohmic contacts ,that are what we've called "a good contact", and rectifing devices ,that work like diodes. A good contact is a low parasitic resistance device with 2 pin through wich can flow a lot of J with a small V. Rectifing devices let current flow only in one direction, for this type of device is needed a low doped semiconductor and a metal with a proper work function; this type of devices are called Schottky diodes.

3.1 Schottky diode

We will study the electrostatic of this device as we've done with diodes. We initially suppose the 2 zones isolated and under thermodinamic equilibrium. Let's take the semiconductor n-doped.
For the metal we're not interested in the conduction or valence band but only in the position of the Fermi level. All type of material's bands are refered to the vacuum level. The distance between the Fermi-level of the metal and the vacuum level E_0 is calle work-function ϕ_m. For silicon we can define the electron affinity that is the distance between the conduction band and the vacuum level χ_s. In this way we can allign the two materials like in figure.

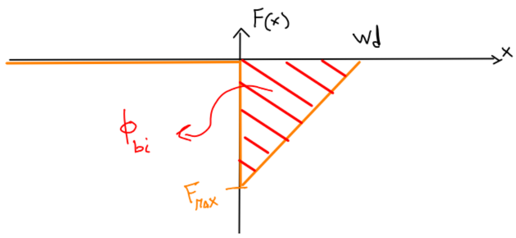


Requirement for a Schottky diode are $E_f^{(m)} < E_f^{(s)}$ and a low doped semicondctor. If
we put the 2 materials toghether the distance between the difference $q\phi_m - q\chi_s = q\phi_{bn}$ is preserved in the semiconductor the bands goes up to align E_f and we therefore have a built in potential ϕ_{bi}. There is a diffusion process from semiconductor to metal and not in the opposite direction beacuse electrons of metal don't have enough energy.



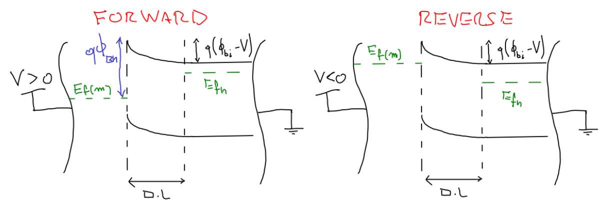
We can olso draw a graph of the electric field where we can view the deplition layer and the built in potential as area.
This device is an ideal unilateral $p^+ - n$ junction; we have the depliton layer only in the n side. We can recove the expression from the diode

$$W_d = \sqrt{\frac{2\varepsilon_{si}}{q}\frac{1}{N_d}\phi_{bi}} = x \tag{3.1}$$



3.1.1 Bias

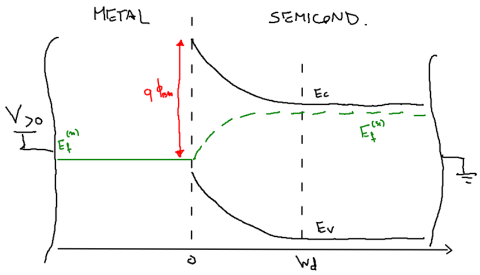
We will always consider the semiconductor part of the device n-doped and grounded; a voltage will be applied to the metal. As for the pn junction we consider foward bias if the voltage applied is grater than 0 and reverse bias if the voltage applied is less than zero.



Reverse bias
With reverse bias $E_{f(m)}$ becomes higher the depletion layer increases as the total voltage drop over the device.
Foward bias
With foward bias $E_{f(m)}$ becomes lower, far from the contact we will have in the semiconductor a quasi-neutral region ,near the contact a transition zone with a depletion region. The total voltage drop decreases as the width of the depletion region.
We have two major difference with the pn junction: the flow of electrons from the semiconductor to the metal don't have a minority constrain (the metal is full of electrons) so we will expect a high current but the transport of holes from metal to semiconductor has this

limitation (the semiconductor is n-doped). J_p will be small, negligible with respect to electrons flow.
Metal-semiconductor junction is a majority carrier device we won't deal with minority carrier in current transport phenomena. We will always neglect G-R processes beacuse are relevant only at low current regime.

3.1.2 Schottcky model

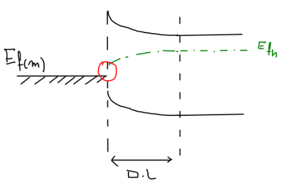


We don't know the behaviour of E_{fn} in the depletion layer but we know that in that zone we have the lower concetration of electrons so it can't be flat. Under stationary condition we can say that $J_n = qn\mu_n F + qD_n \frac{dn}{dx}$ = constant and, as boundary condition, that $n(W_d) = N_d$ $n(0) = N_c e^{\frac{q\phi_{bn}}{kT}}$. F is given the only parameter we don't know is n and $\frac{dn}{dx}$ so solving the continuity equation for electrons with this boundary conditions we can get that

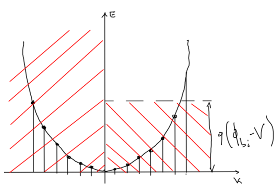
$$J = J_0(e^{\frac{qV}{kT}} - 1) \tag{3.2}$$

that is a relation identical to the pn junction but with a different J_0.
However this relation does not mach the experimental results. This is due to an implicit assumption that we've made using $n(0) = N_c e^{\frac{q\phi_{bn}}{kT}}$ as boundary condition; this happens only if the interface is always at thermodynamic equilibrium.
This model is valid for low mobility semicondutors (not for Si, Ge and GaAs).

3.1.3 Bethe's model



E_{fn} can arrive at the interfae higher than $E_{f(m)}$.
We have a very thin interface that is of the order of fractions of nm there we can't describe current transport with drift and diffusion theory beacuse they're dominated by scattering events that occurs in tens of nanometers. We have to use a thermoionc transport model.
The electrons that pass from semiconductor to metal are that with an energy greater than $q(\phi_{bi} - V)$ so an energy higher than the potential barier of the depletion layer.
From energy dispersion relation $E = E_c + \frac{\hbar^2 k^2}{2m_x} + \frac{\hbar^2 k^2}{2m_y} + \frac{\hbar^2 k^2}{2m_z}$ we define $E_x = \frac{\hbar^2 k^2}{2m_x}$ the energy related to x transport and let's take into account only one ellipsoide.



Refering to the interface between the depletion layer and the quasi neutral region, half of the E_x parabola can be neglected beacuse we want electrons that move from right to left (semiconductor to metal). We want olso to consider only electrons with energy higher than $q(\phi_{bi} - V)$ so with a $k_x > \overline{k}$ so we can neglect other states.

The equivalent current of this electrons will be

$$J_{s-m}^{(1)} = \sum_{k_x > \overline{k}} 2\frac{q}{L^3} v_x(k_x) f(k_x, k_y, k_z) \tag{3.3}$$

the 2 factor is a spin correction. Adding a corrective term we can transform the summation into an integral

$$J_{s-m}^{(1)} = \frac{1}{(\frac{2\pi}{L})^3} \int_{\overline{k_x}}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{q}{L^3} v_x(k_x) f(k_x, k_y, k_z) dk_x dk_y dk_z \tag{3.4}$$

remembering that $\hbar k_{x,y,z} = m_{x,y,z} v_{x,y,z}$ therefore $dk_{x,y,z} = \frac{m_{x,y,z}}{\hbar} dv_{x,y,z}$, we can change integration variable as

$$J_{s-m}^{(1)} = \frac{2q}{(2\pi)^3} \int_{\overline{v_x}}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v_x f(k_x, k_y, k_z) \frac{m_x m_y m_z}{\hbar^3} dv_x dv_y dv_z \tag{3.5}$$

for f we can use M-B approximation $f \simeq e^{-\frac{E-E_f}{kT}}$ using energy dispersion relation

$$E = E_c + \frac{1}{2} m_x v_x + \frac{1}{2} m_y v_y + \frac{1}{2} m_z v_z \tag{3.6}$$

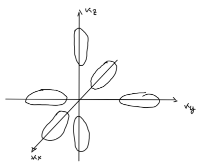
the M-B equation becomes

$$f \simeq e^{-\frac{E_c - E_{fn}}{kT}} \cdot e^{-1/2 \frac{m_x v_x^2}{kT}} \cdot e^{-1/2 \frac{m_y v_y^2}{kT}} \cdot e^{-1/2 \frac{m_z v_z^2}{kT}} \tag{3.7}$$

so solving the 3 integral (the first is a gaussian $= \frac{kT}{m_x} e^{-\frac{m_x \overline{v_x}^2}{kT}}$ second and third $= \sqrt{\frac{2kT\pi}{m_{y,z}}}$) we get

$$J_{s-m}^{(1)} = \frac{4\pi q}{h^3} \sqrt{m_x m_y} (kT)^2 e^{\frac{-\phi_{bn}}{kT}} e^{\frac{qV}{kT}} = A \cdot \sqrt{m_x m_y} T^2 e^{\frac{-\phi_{bn}}{kT}} e^{\frac{qV}{kT}} \tag{3.8}$$

where $A = \frac{4\pi q k^2}{h^3} = 120 \frac{A^2}{cm^2 K^2}$ it's called the Richardson constant.



Taking into account all the ellipsoide we get

$$J_{s-m} = A \frac{2m_t + 4\sqrt{m_t m_l}}{m_0} T^2 e^{\frac{-\phi_{bn}}{kT}} e^{\frac{qV}{kT}} = A^* T^2 e^{\frac{-\phi_{bn}}{kT}} e^{\frac{qV}{kT}} \tag{3.9}$$

with $A^* = A \cdot 2.05$.

From m-s we can say that under th.eq the 2 current must be equal and that the barrier is costant equal to $q\phi_{bn}$ so $J_{m-s} = A^* T^2 e^{\frac{-\phi_{bn}}{kT}}$.

So finally the total current throught the device will be

$$J = A^* T^2 e^{\frac{-q\phi_{bn}}{kT}} (e^{\frac{qV}{kT}} - 1) = J_{0,th} (e^{\frac{qV}{kT}} - 1) \tag{3.10}$$

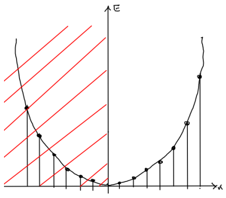
ϕ_{bn} it's a crucial parameter for current flow in metal semiconductor junction. We have a $J_{0,th} \simeq 10^5 A/cm^2$ orders of magnitude higher with respect to the pn junction and a turn on voltage of 0.3-0.4 V.

This model is valid for high mobility semiconductors.

3.1.4 Universal model

We can connect the two models in an universal one.

Starting from Schottcky model we can use his assumption of $J_n = qn\mu_n F + qD_n \frac{dn}{dx} = const$ and $n(W_d) = N_d$ but we have to change n(0). We can say something about $J_n(0)$ as the current density without scattering.



We can calculate that term using
energy dispersion relation but beacuse we are at the interface we will take all the left state of the parabola (we don't have a barrier) so we get

$$J_n(0) = A^* T^2 e^{-\frac{E_c - E_{fn}(0)}{kT}} \tag{3.11}$$

that is the correct boundary
condition. If we muliply and divide this expression by N_c we obtain

$$J_n(0) = A^* T^2 \frac{n(0)}{N_c} \tag{3.12}$$

this for the electron flow from semiconductor to the metal. For the opposite direction we use the th.eq. condition and we get

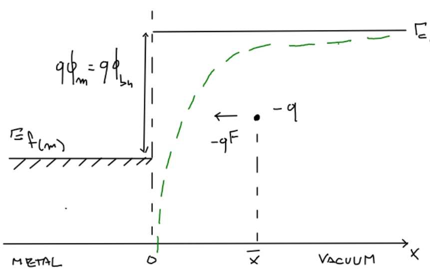
$$J_n(0) = A^* T^2 \frac{n_0}{N_c} \tag{3.13}$$

with n_0 electron concentration at the interface under th.eq.
So the final boundary condition is

$$J_n(0) = \frac{A^* T^2}{N_c} [n(0) - n_0] \tag{3.14}$$

We are using a drift-diffusion approach with thermionic boundary condition.
Solving this system we get an expression valid for all semiconductors that can be simplified in the 2 models depending on what type of semiconductor we have.
If mobility is large than it's easy to move electrons but difficult to remove them form the interface that becomes the bottom neck of the system. With low mobility it's difficult to move electrons but easy to remove them form the interfece so they don't pile up there.

3.1.5 Schottky effect



The Schottky effect is a pure electrostatic
effect that take place in the Bethe's model.
To study this
effect we start from a metal-vacuum "junction" placing in the vacuum a single electron at a certain distance $\overline{x}$.
We have an
electrostatic induction at the surface of the metal that attracts the electron in the vacuum near the surface.
In order to calculate the electric field that attracts the single electron we can use the image method placing a positive charge +q at $-\overline{x}$ and removing the metal.
From this assumption we get

$$-qF = \frac{1}{4\pi\varepsilon_0} \frac{-q^2}{(2\overline{x})^2} \tag{3.15}$$

and so the electric field

$$F = \frac{q}{16\pi\varepsilon_0 x^2} \tag{3.16}$$

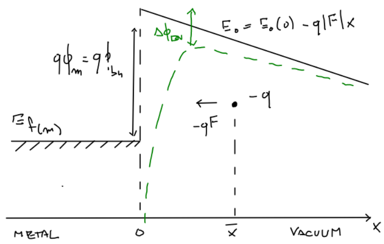
from witch we can calculate ,integrating both part of the equation from x to $+\infty$, the potential as

$$\phi(x) = \phi(+\infty) + \frac{q}{16\pi\varepsilon_0 x} \tag{3.17}$$

Multipling for -q we can obtain the energy

$$E_0(x) = E_0(+\infty) - \frac{q^2}{16\pi\varepsilon_0 x} \tag{3.18}$$

and so the profile in green in the graph.



If we consider that in the vacuum there is a costant electric field (and so $E_0(x) = E_0(0) - q|F|x$) like in figure beacuse the sistem is linear we obtain that

$$E_0(x) = E_0(0) - q|F|x - \frac{q^2}{16\pi\varepsilon_0 x} \tag{3.19}$$

and the behaviour in red in figure. We have change of the peak and so of the barrier that block our electrons. The Δ of the barrier is called Schottky barrier lowering.
We can calculate the new peak deriving by dx the $E_0(x)$ function finding the x_{max} obtaining the energy at that point

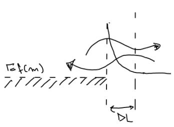
$$E_0(x_{max}) = E_0(0) - \sqrt{\frac{q^3|F|}{4\pi\varepsilon_0}} \tag{3.20}$$

So the difference in the barrier is $\Delta\phi_{bn} = \sqrt{\frac{q^3|F|}{4\pi\varepsilon_0}}$.
Now let's apply this effect on our metal semiconductor junction: when an electron travels throught the depletion layer from the semiconductor to the metal we have a barrier lowering of $\Delta\phi_{bn} = \sqrt{\frac{q^3|F_{max}|}{4\pi\varepsilon_{Si}}}$ (we place F_{max} beacuse the maximum effect we have is with the max F and the narrower distance).
In our device we have a change of the J-V characteristic

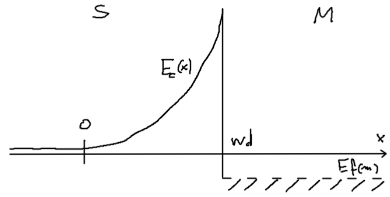
$$J = A^* T^2 e^{\frac{-q(\phi_{bn} - \Delta\phi_{bn})}{kT}} (e^{\frac{qV}{kT}} - 1) \tag{3.21}$$

that creates an increment of J_0 and 2 opposit behaviour for reverse and forward bias: with $V < 0$ the current increases increasing $|V|$, with $V > 0$ J decrease this due to the dependance of $\Delta\phi_{bn} \propto \sqrt{F_{max}} \propto V_{rv-bias}$.

3.2 Ohmic contact



In order to have an ohmic contact we need a high doped semiconductor. With N_d very high the depletion layer becomes narrower and increasing the doping concentration we arrive at a condition when there is a strong factor of J due to quantum-mechanical tunneling. That process does not depend on the type of bias so we don't have anymore a rectifing behaviour.



Taking into account the figure as reference system from the studied pn junction we know the behaviour of $E_c(x) = \frac{q^2 N_d}{2\varepsilon_{Si}} x^2$ so we can calculate T ,transparancy coefficient as

$$T = e^{-2\int Im\{E_c(x)\}dx} = e^{-2\int \sqrt{\frac{2m^*(E_c-E)}{\hbar^2}}dx} \tag{3.22}$$

we have to remember that in this case m^* is the effective mass for tunneling process. So if we consider $E = E_f$

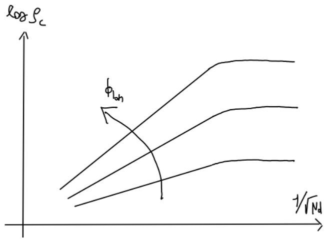
$$T = e^{-2\int \sqrt{\frac{2m^* E_c}{\hbar^2}}dx} = e^{-2\sqrt{\frac{2m^*}{\hbar^2}}\sqrt{\frac{q^2 N_d}{2\varepsilon_{Si}}}W_d^2/2} \tag{3.23}$$

That substituiting W_d with its expession we get

$$T = e^{-2\sqrt{\frac{2m^*}{\hbar^2}}\sqrt{\frac{q^2 N_d}{2\varepsilon_{Si}}\frac{2\varepsilon_{Si}}{qN_d}(\phi_{bi}-V)}} = e^{-q\frac{(\phi_{bi}-V)}{E_{00}}} \tag{3.24}$$

where $E_{00} = q\frac{\hbar\sqrt{N_d}}{2\sqrt{m^*\varepsilon_{si}}}$

check this paragraph below

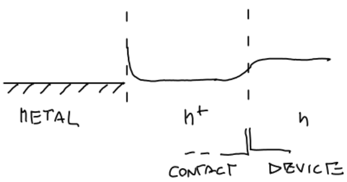


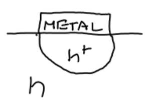
One important parameter of this device is the contact resistivity defined as

$$\rho_c = (\frac{\partial J}{\partial V})^{-1}|_{V=0} = \frac{E_{00}}{q}e^{q\phi_{bi}/E_{00}} \tag{3.25}$$

This parameter depends on the barrier hight and on $\sqrt{N_d}$. In the graph shows the $\log(\rho_c) - 1/\sqrt{N_d}$ dependance. It's a straight line until the concentration becomes too low and so tunneling effect is no more relevant. To achive large current flow we need to perturbe only al little bit the device so it stays always near thermodynamic equilibrium.

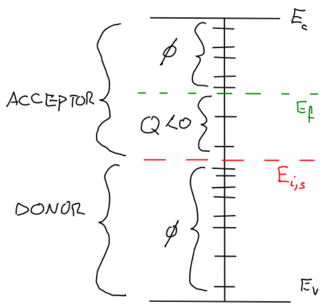
The two figures below show how a contact is done in a device.



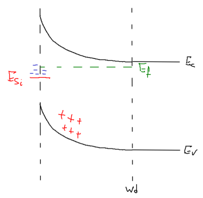


3.3 Interface states

We've always considered Si as a periodic infinite cristal but in metal-semiconductor junction we have te interface that interrupts the sequence of atoms. Some atoms of Si can't share all theyr 4 valence electrons creating some dangling bonds.



Electrons remains localize close to theyr silicon atoms this creates spurius energy level at the interface. In the bulk region electrons and holes are only in the conduction band or in the valence band but at the interface can stay in a lot of states between this two bands. The spurious states in the upper part of the bandgap have an acceptor behaviour (negativly charged when empty and neutral when filled) the others in the bottom part have a donor behaviour (neutral when filled and positivly charged when empty).
The energy that divide this two type of states is the E_{is}.
All states above E_f will be filled all the other states empty. So the previous band-diagram is incoherent with the Gauss'law since we have some exposed charge and no band banding.



We have a negative charge so $\phi < 0$ the bands band upward and the distance $E_f - E_{is}$ becomes narrower. The bands banding create a depletion region that expose a positive charge equal to the interface states' charge.
The distance between $E_c - E_f$
is not set only by doping concentration but olso from interface states.
We define interface state density $N_{is} = [cm^{-2}eV^{-1}]$
in order to find the total charge introsuced by the interface

$$|Q_{is}| = N_{is}(E_f - E_{is})q \tag{3.26}$$

The distance $E_c(0) - E_{is} \simeq E_{gap}/2$ so writing $\Delta E_{is} = E_c(0) - E_{is} \simeq E_{gap}/2$ we can say that

$$|Q_{is}| = N_{is}[\Delta E_{is} - (E_c(0) - E_f)]q \tag{3.27}$$

We know from the analysis of the pn junction that the charge in the depletion leyer is

$$Q_{dep} = qN_dW_d = \sqrt{2\varepsilon_{si}qN_d\frac{E_c(0) - E_c(W_d)}{q}} \tag{3.28}$$

where the last term is the voltage drop in the depletion layer.
For Gauss law this 2 terms have to be equal so solving the equation we get the level of the conduction band at the interface

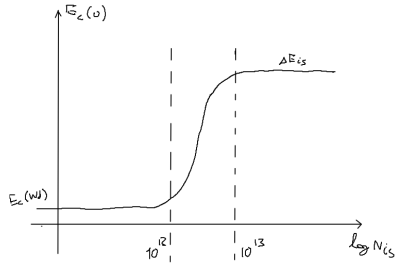
$$E_c(0) = \Delta E_{is} + \frac{\varepsilon_{si}N_d}{q^2N_{is}^2} - \sqrt{(\frac{\varepsilon_{si}N_d}{q^2N_{is}^2})^2 + \frac{2\varepsilon_{si}N_d}{q^2N_{is}^2}(\Delta E_{is} - E_c(W_d))} \tag{3.29}$$

This effect changes the barrerier hight and so theflow of current throught the device. ϕ_{bn} becomes strongly dependent on N_{is}

3.3.1 Limit cases

Starting from the equation $|Q_{is}| = Q_{dep}$ we have 2 limit cases.
$N_{is} = 0$ in this case $E_c(0) = E_c(W_d)$ so we don't have a band banding we restore the ideal case.
$N_{is} \to +\infty$ in this case we notice that the first term has to be finite and so $\Delta E_{is} - (E_c(0) - E_{is}$ has to be zero that means that $E_c(0) = E_f + \Delta E_{is}$.
This case means we are moving E_{is} up to E_f that is totally indipendent from doping concentration but depends only from the interface states. This condition is called "Fermi level pinning at the surface".

The plot of $E_c(0) - \log(N_{is})$ make a transition between $10^{12} - 10^{13}$ as order of magnitude.

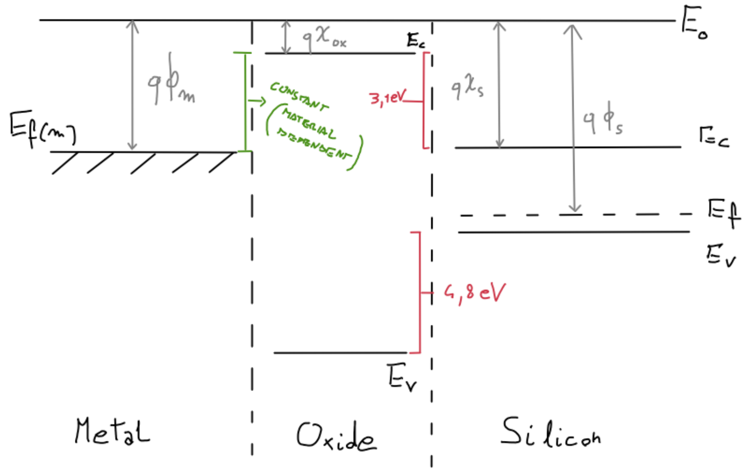


Chapter 4

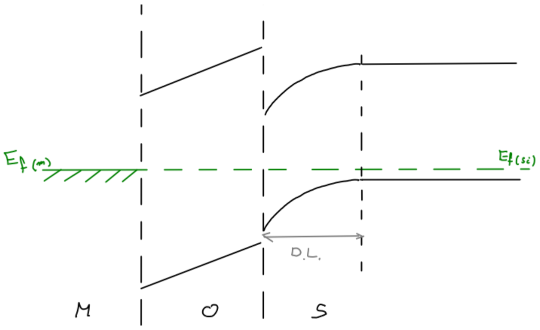
MOS capacitor

We start our analysis with a 1D device, we will consider the substrate a monocristal of constant p-doped silicon.
We will consider for oxide SiO_2 that has a high energy gap ($\simeq$8-9eV), a low concentration of free carriers, a large resistivity, a reasonably low concentration of spurious difect ($\simeq 10^{10}cm^{-3}$) but a low dialectric constant ($\simeq 3.9$).This last propriety is the cause why recently SiO_2 is being substituted with high-k materials.
In our analysis we will consider the dialectric ideal (this means we have to solve only the Poisson eq. for the electrostatic beacuse there will be no current flow) and as gate material a metal.

4.1 Working regimes



We start considering the 3 material isolated from each other under th.eq.
We will refer $E_{f(m)} \simeq E_c^{Si}$. we have olso to introduce the vacuum level and so electron affinity for oxide and silicon, metal work function, and silicon work function. From this data we can say that the metal we need is Al in order to mantain $E_{f(m)} \simeq E_c^{Si}$.
We olso underlined in red the potential barrier created by the oxide for holes and electrons.

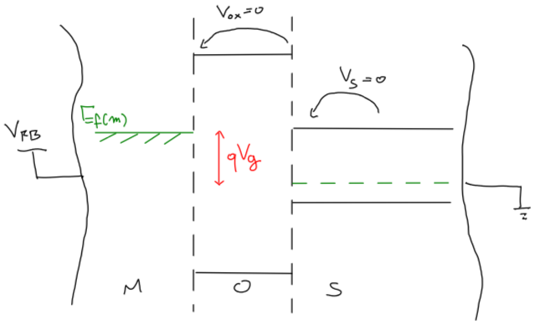


We put the 3 materials in contact after a while we will have a single Fermi-level all over the device and we can say that far from the contact p-Si will be at equilibrium. $E_{f(m)}$ will move downwards toghether with E_c^{ox} since the distance $E_{f(m)} - E_c^{ox}$ is constant (material dependent) as E_{gap}^{ox} . We have discontinuity of F, for Gauss law $\varepsilon_{ox} F_{ox} = \varepsilon_{Si} F_{Si} \rightarrow F_{ox} = 3F_{Si}(0)$. We have a linear band banding in the oxide for the hp of idealty of the dialectric and a parabolic band banding (that means a linear F) in the Si.

We have 2 voltage drops over the silicon and over the metal that define $V_s + V_{ox} = \phi_{bi}$. If we apply a $V > 0$ at the gate we push $E_{f(m)}$ downwards creating a higher band banding both in oxide and silicon, this band banding will increase both V_{ox} and V_s so assuming that the bulk keeps th.eq we can write the following equation

$$V_G + \phi_{bi} = V_s + V_{ox} \tag{4.1}$$

4.1.1 Flat band regime



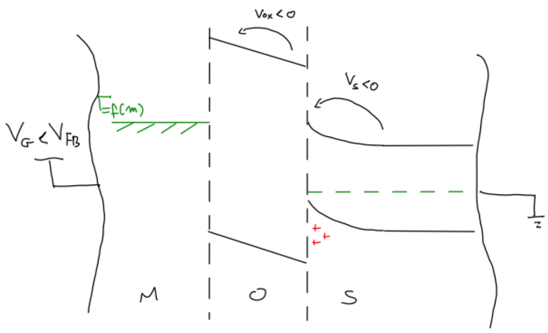
If we apply $V_G = -\phi_{bi}$ from the previous equation we have that $V_s = V_{ox} = 0$ and a separation $E_{f(m)} - E_{f(Si)} = qV_G$. This is the so called Flat-band condition of the mos capacitor

$$V_{FB} = -\phi_{bi} \tag{4.2}$$

Is defined as the voltage to apply at the gate to have all bands flat. now we can re-write the balance of voltages as

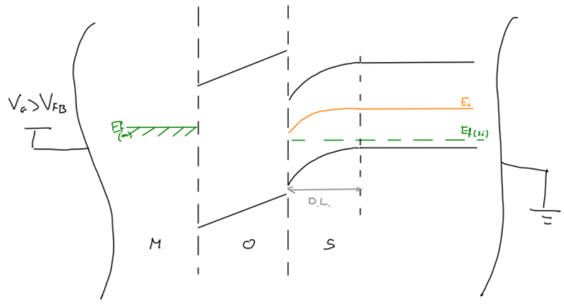
$$V_G - V_{FB} = V_s + V_{ox} \tag{4.3}$$

4.1.2 Accumulation regime



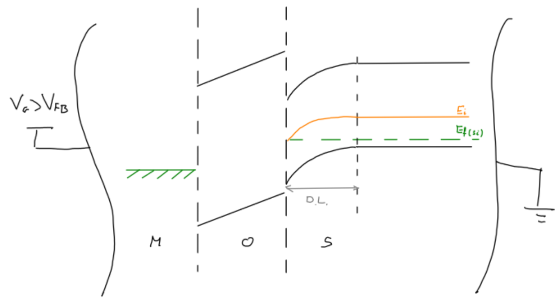
If $V_G < V_{FB}$ bands band upwards we have more holes concentration at the interface between Si-Ox than the doping concentration.
This is calle the accumulation regime.

4.1.3 Deplition regime



With $V_G > V_{FB}$ we have in silicon a depletion region so we are in the depletion regime of the mos capacitor.

4.1.4 Weak inversion condition

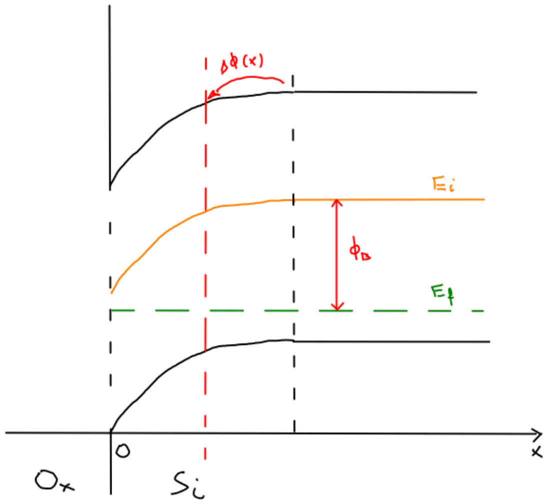


If we further increase $V_G >> V_{FB}$ we arrive at the condition that the intrinsic Fermi level at the surface reaches the Fermi level ($E_f = E_i$).
From now on if we slightly increase the voltage we will have more electrons than holes at the interface. This is the condition for weak inversion condition.

4.1.5 Strong inversion regime

If we continue to increase the voltage applied we will reach the strong inverson regime that is characterized by $(E_i - E_f)|_{interface} = -(E_i - E_f)|_{bulk}$ so the concentration of electrons at the interface is equal to the doping concentration in the bulk.

4.2 Electrostatics



We have to solve the Poisson equation $\frac{d^2\phi}{dx^2} = -\frac{q}{\varepsilon_{Si}}(p - n - N_a^-)$ only in the substrate.
We define ϕ_B the electrostatic potential of the bulk far away from the interface and we take E_f that is costant as the reference energy level 0.
We introduce olso $\Delta\phi(x) = \phi(x) - \phi_B$ that is the change of the electrostatic potential wrt the bulk;$\Delta\phi(W_d) = 0$ and $\Delta\phi(0) = V_s$.

Now we have to write $p - n - N_a^-$ as function of $\phi(x)$.
For electrons we can say that $n = n_i e^{(E_f - E_i)/kT}$ if we multiply and divide by q we get $n = n_i e^{q\phi(x)/kT}$ same for holes as $p = n_i e^{-q\phi(x)/kT}$. Electrons and holes in the bulk are $n_0 = n_i e^{q\phi_b/kT}$ and $p_0 = n_i e^{-q\phi_b/kT}$. From this 4 equation we can write that

$$n = n_0 e^{q\Delta\phi/kT} \qquad p = p_0 e^{-q\Delta\phi/kt} \tag{4.4}$$

In the bulk we can say ythat $p_0 = n_0 + N_a$ so

$$N_a = p_0 - n_0 \tag{4.5}$$

Now Poisson equation becomes

$$\frac{d^2\phi}{dx^2} = -\frac{q}{\varepsilon_{Si}}(p_0 e^{-q\Delta\phi/kt} - n_0 e^{q\Delta\phi/kT} - p_0 + \frac{n_i^2}{p_0}) \tag{4.6}$$

Until now the only approximation we have made is the M-B. Now we consider the complete ionization approximation $p_0 \simeq N_a$. And we make a change of variable

$$\frac{d^2\phi}{dx^2} = -\frac{q}{\varepsilon_{Si}}(N_a e^{-q\Delta\phi/kt} - \frac{n_i^2}{N_a} e^{q\Delta\phi/kT} - N_a + \frac{n_i^2}{N_a}) \tag{4.7}$$

$$d(\frac{d\phi}{dx}) = -\frac{q}{\varepsilon_{Si}}(N_a e^{-q\Delta\phi/kt} - \frac{n_i^2}{N_a} e^{q\Delta\phi/kT} - N_a + \frac{n_i^2}{N_a})dx \tag{4.8}$$

$$\frac{d\phi}{dx} d(\frac{d\phi}{dx}) = -\frac{q}{\varepsilon_{Si}}(N_a e^{-q\Delta\phi/kt} - \frac{n_i^2}{N_a} e^{q\Delta\phi/kT} - N_a + \frac{n_i^2}{N_a})dx\frac{d\phi}{dx} \tag{4.9}$$

Now we integrate one time this expression. The first member is like f(x)f'(x) so the result is $-1/2(\frac{d\phi}{dx})^2$. Integrating olso the second member and gathering the common terms the equation we get

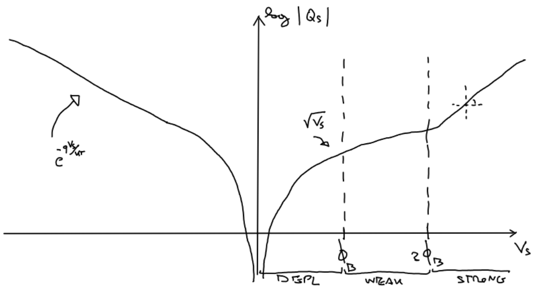
$$(\frac{d\Delta\phi}{dx})^2 = \frac{2kTN_a}{\varepsilon_{si}}\left(e^{-q\Delta\phi/kT} + \frac{q\Delta\phi}{kT} - 1 + \frac{n_i^2}{N_a^2}(e^{q\Delta\phi/kT} - \frac{q\Delta\phi}{kT} - 1)\right) = F^2(x) \qquad (4.10)$$

For Gauss law $F_s = F(0) = -\frac{Q_s}{\varepsilon_{si}}$ (and olso that $\Delta\phi(0) = V_s$)from this relation and the previous equation we get

$$Q_s = \pm\sqrt{2kTN_a\varepsilon_{si}}\left(e^{-qV_s/kT} + \frac{qV_s}{kT} - 1 + \frac{n_i^2}{N_a^2}(e^{qV_s/kT} - \frac{qV_s}{kT} - 1)\right)^{1/2} \qquad (4.11)$$

We already know the total charge in the substrate for all working regimes.

4.2.1 Dominant terms in the solution of Poisson equation



To better undestand this formula let's analyse all working regime separately and draw a $\log|Q_s| - V_s$ graph.

In accumulation regime when $V_s < 0$ we have that $Q_s \simeq \sqrt{2kTN_a\varepsilon_{si}}e^{-\frac{qV_s}{2kT}}$. So in the graph we see a straight line dependence. We can say that the term

$$e^{\frac{-qV_s}{kT}} \rightarrow h_{sub} \qquad (4.12)$$

in the solution of the Poisson equation is the term that takes into account for hole concentration in the substrate.

In the depletion regime $0 < V_s < 2\phi_b$ we have to consider two dominant term: $\frac{qV_s}{kT}$ and $\frac{n_i^2}{N_a^2}e^{\frac{qV_s}{kT}}$. If we evaluate the second term at $2\phi_b$ we get 1 so it's negligible and therefore the first term is predominant. In this regime

$$Q_s \simeq -qN_a\sqrt{\frac{2\varepsilon_{Si}}{q}\frac{1}{N_a}V_s} \qquad (4.13)$$

that is the charge of the depletion layer.

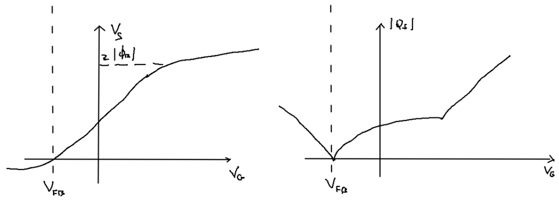
In strong inversion $Q_s \simeq \sqrt{2kTN_a\varepsilon_{si}}\frac{n_i^2}{N_a^2}e^{\frac{qV_s}{2kT}}$ so we can say that the term

$$\frac{n_i^2}{N_a^2}e^{\frac{qV_s}{kT}} \rightarrow e_{sub} \qquad (4.14)$$

in the solution of the Poisson equation is the term that takes into account electrons in the substrate.

4.2.2 Total charge in function of gate bias

We want $Q(V_g)$; to have that expression we need to remember that $V_g - V_{fb} = V_s + V_{ox}$ but we can write $V_{ox} = F_{ox}t_{ox} = \frac{\varepsilon_{Si}F_{Si}}{\varepsilon_{ox}}t_{ox} = -Q(V_s)/C_{ox}$ so $V_g - V_{fb} = V_s - \frac{Q_s}{C_{ox}}$ using this expression and $Q_s = Q_s(V_s)$ we can get what we want.
We will plot olso the 2 graphs $V_s - V_g$ and $|Q_s| - V_g$.



We will take as reference point the flat band condition where $Q_s = 0$ and $V_s = 0$.

For $V_s < 0$ we said that $Q_s \simeq \sqrt{2kTN_a\varepsilon_{si}}e^{-\frac{qV_s}{2kT}}$ so we get that

$$V_g - V_{fb} = V_s - \frac{\sqrt{2kTN_a\varepsilon_{si}}}{C_{ox}}e^{-\frac{qV_s}{2kT}} \tag{4.15}$$

we can neglect V_s beacuse it's negligible wrt the exponential term so we get

$$V_s \simeq -\frac{2kT}{q}\ln\left(\frac{C_{ox}(V_{fb} - V_g)}{\sqrt{2kTN_a\varepsilon_{si}}}\right) \qquad Q_s \simeq C_{ox}(V_{fb} - V_g) \tag{4.16}$$

a small variation of V_g let V_s move not so much due to the exponential increase of the charge in the bulk and a consequent big field in the oxide; we are not far from a metal plate capacitor.

For $0 < V_s < \phi_b$ we get that $Q_s \simeq \sqrt{2qN_a\varepsilon_{Si}V_s}$ so we get the following equation (that we don't solve)

$$V_g - V_{fb} = V_s + \frac{\sqrt{2qN_a\varepsilon_{Si}V_s}}{C_{ox}} \tag{4.17}$$

For weak inversion $\phi_b < V_s < 2\phi_b$ we can consider both dominants terms getting the following equation

$$V_g - V_{fb} = V_s + \frac{\sqrt{2\varepsilon_{si}kTN_a}}{C_{ox}}\left(\frac{qV_s}{kT} + \frac{n_i^2}{N_a^2}e^{qV_s/kT}\right)^{1/2} \simeq V_s + \frac{\sqrt{2\varepsilon_{si}kTN_a}}{C_{ox}}\sqrt{\frac{qV_s}{kT}}\left(1 + \frac{n_i^2}{2N_a^2}e^{qV_s/kT}\frac{kT}{qV_s}\right) \tag{4.18}$$

using Taylor expansion we get an exponential dependance on V_g of the e charge (that is the second term). Form this dependace we will get the sub-th current of the mos capacitor.

For strong inversion we get

$$V_g - V_{fb} = V_s + \frac{\sqrt{2\varepsilon_{si}kTN_a}}{C_{ox}}\frac{n_i}{N_a}e^{qV_s/2kT} \tag{4.19}$$

remembering that $\frac{n_i^2}{N_a^2}e^{q2|phi_b|/kT} = 1$ we can say that $\frac{n_i}{N_a} = e^{-q2|\phi_b|/2kT}$ so we can write that

$$V_g - V_{fb} = V_s + \frac{\sqrt{2\varepsilon_{si}kTN_a}}{C_{ox}}e^{\frac{q(V_s - 2|\phi_b|)}{2kT}} \tag{4.20}$$

So we get V_s and making a rought approx of $V_s \simeq 2|\phi_b|$ olso Q_s

$$V_s = 2|\phi_b| + \frac{2kT}{q} \ln\left(\frac{C_{ox}(V_g - V_{fb} - V_s)}{\sqrt{2\varepsilon_{si}kTN_a}}\right) \qquad Q_s = -C_{ox}(V_g - V_{fb} - 2|\phi_b|) \tag{4.21}$$

The approximation means that entering in storng inversion we get a maximum depletion layer and a maximum depletion charge

$$W_d^{max} = \sqrt{\frac{2\varepsilon_{si}}{qN_a}2|\phi_b|} \qquad Q_{dep}^{max} = -\sqrt{2\varepsilon_{si}qN_a2|\phi_b|} \tag{4.22}$$

we increase the charge of electrons not that of the depletion layer.
The V_g at witch we enter in the strong inversion is $V_g - V_{fb} = 2|\phi_b| + \frac{|Q_{dep}^{max}|}{C_{ox}}$ that is called threshold voltage of the mos capacitor that is

$$V_t = V_{fb} + 2|\phi_b| + \frac{|Q_{dep}^{max}|}{C_{ox}} \quad \rightarrow Q_{inv} = Q_s - Q_{dep}^{max} = -C_{ox}(V_g - V_t) \tag{4.23}$$

4.3 Small signal capacitance

We want to study the small signal capacitance of the MOS capacitor that is a non linear device so we want $C_g = -\frac{dQ_s}{dV_g}$ the change of charge over the change of voltage.
We can guess the behaviour of C_g using the expression $V_g - V_{fb} = V_s - Q_s/C_ox$ and calculating the derivative wrt $d(-Q_s)$ we get

$$\frac{dV_g}{d(-Q_s)} = \frac{dV_s}{d(-Q_s)} + \frac{1}{C_ox} \tag{4.24}$$

that is

$$\frac{1}{C_g} = \frac{1}{C_s} + \frac{1}{C_{ox}} \tag{4.25}$$

The total gate capacitance can be expressed by the series of 2 capacitive terms one fixed (C_{ox}) one variable (that is the substrate capacitance C_s). From this we know that C_g has an upper limit of C_{ox}.
Let's analise the small signal capacitance in the various regime.

4.3.1 in accumulation regime

We derivate the charge of the accumulation region obtaining

$$C_s = \sqrt{2kTN_a\varepsilon_{si}}e^{-\frac{qV_s}{2kT}}\frac{q}{2kT} = Q_s/\frac{2kT}{q} \tag{4.26}$$

Knowing $Q_s(V_g)$ we get

$$C_s = \frac{C_{ox}(V_{fb} - V_g)}{2kT/q} \tag{4.27}$$

This is $>> C_{ox}$ if $V_{fb} - V_g >> 2kT/q$. So from flat band to minus infinite the capacitance increases and becomes similar to C_{ox}. We are close to a metal plate capacitor.

4.3.2 in flat band condition

We have to include in the Q expression both electrons and holes so
$Q_s = \sqrt{2kTN_a\varepsilon_{si}}(e^{\frac{-qV_s}{kT}} + \frac{qV_s}{kT} - 1)^{1/2}$.
If we make the Taylor expansion of the exponential term around 0 we get $Q_s = -V_s\sqrt{\frac{\varepsilon_{si}q^2N_a}{kT}}$ so

$$C_s = \frac{d(-Q_s)}{dV_s} = \frac{\varepsilon_{si}}{\sqrt{\frac{\varepsilon_{si}kT}{q^2N_a}}} = \frac{\varepsilon_{si}}{L_D} \tag{4.28}$$

we can recognise the Dybae lenght in the expresison we can interpretate this result as this ; the holes (majority) screen a perturbation of the electric field form the gate like a step increase of doping concentration.
This capacitance and the oxide capacitance are comparable.

4.3.3 in depletion regime

In depletion regime we have a charge $Q_s = -\sqrt{2\varepsilon_{Si}qN_aV_s}$ that is the usual expression for the charge in a depletion layer. From the well known realtion $C = \varepsilon_{si}/t_{ox}$ we get

$$C_g = C_{ox}/\sqrt{1 + 2C_{ox}^2\frac{V_g - V_{fb}}{\varepsilon_{si}qN_a}} \tag{4.29}$$

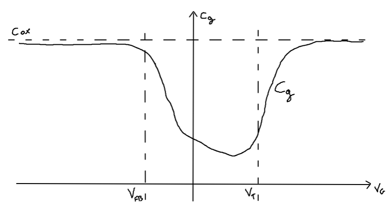
This capacitance decrease quickly near the flat band condition beacuse the contribution related to holes disappears over a few kT/q.

4.3.4 in strong inversion

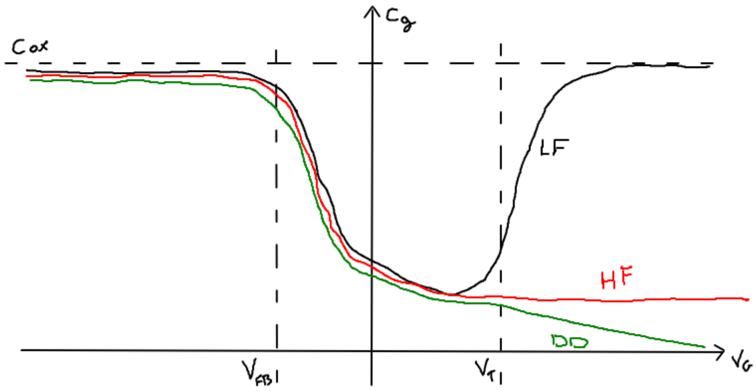
We have the charge $Q_s = \sqrt{2\varepsilon_{Si}kTN_a}\frac{n_i}{N_a}e^{\frac{qV_s}{2kT}}$ so making the derivate we reach

$$C_s = \frac{|Q_s|}{2kT/q} = C_{ox}\frac{(V_g - V_{fb} - 2|\phi_b|)}{2kT/q} \tag{4.30}$$

Increasing V_g we get a larger capacitance we reach a condition similar to a metal plate capacitance.



4.4 Frequency regimes



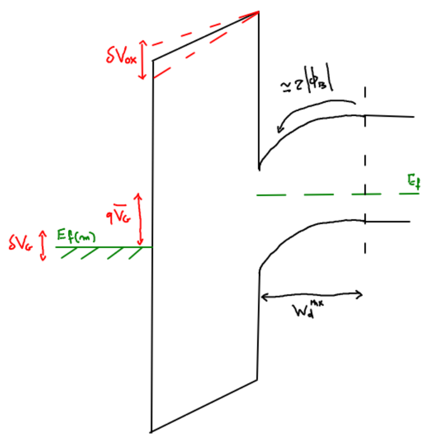
Let's look more colsely to the substrate during the different regimes when we apply a small signal to the gate.

In accumulation regime
$\overline{V_g} < V_{fb}$ if we move the gate of $\delta V_g > 0$ we are removing h from the silicon surface to the substrate. Vice-versa with $\delta V_g < 0$ we move h from the substrate to the silicon surface.
This is a modulation of majority carriers concentration so it happens in a very short time in the order of the dialectric relaxation time.

In weak inversion
$V_{fb} < \overline{V} < V_t$ we have a depletion region with a certain voltage drop V_s and an exposed negative charge at the surface.
With a $\delta V_g > 0$ we slightly increase the width of the depletion layer so we are removing holes from the edge of the depletion layer to the substrate. With $\delta V_g < 0$ we reduce the depletion layer so we add h from the substrate to the depletion layer to neutralize the exposed charge.
Olso here we modulate the majority carrier so this is a process that have as time constant the dialectric relaxation time.

Strong inversion
Depletion layer at the maximum expansion at the silicon surface we have a lot of electrons. Appling δV_g we modulate e concetration this electrons cannot came from the contact that exchange only majority they have to be created by generation processes.

4.4.1 Low frequency regime

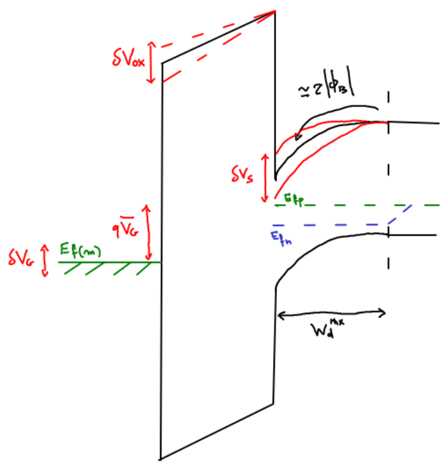


The time constant
for the generation process in strong inversion can be extracted from the generation current $J_g = q\frac{n_i}{2\tau_0}W_d^{max}$ that is the total current that can be created per unit time per unit area. The charge of the depletion layer is $Q_{dep} = qN_aW_d^{max}$ the time for generate this charge is

$$\tau_g = \frac{Q_{dep}}{J_g} = 2\tau_0 N_a/n_i \tag{4.31}$$

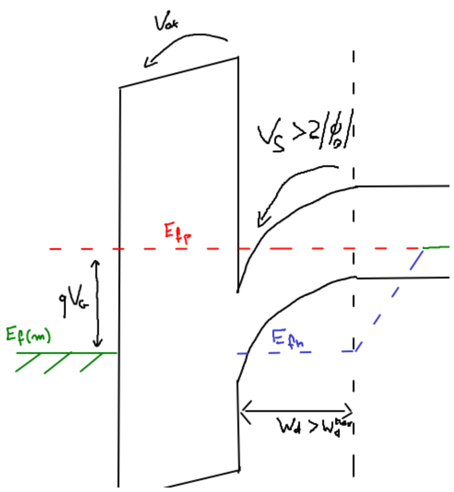
τ_g is technology dependent and very big in the order of seconds. Only if the signal has a frequence much lower than τ_g ($f < 1Hz$) we have the ideal charachteristic seen before.
Looking at the band diagram a movement of δV_g moves $E_{f(m)}$ upwards or downwards. With a $\delta V_g > 0$ we have to increase the total voltage drop but the dependance with V_{sub} is very weak so $\delta V_s \simeq 0$ and we modulate only the voltage drop over the oxide (under charge sheet approximation) $\delta V_g \simeq V_{ox}$

4.4.2 High frequency regime



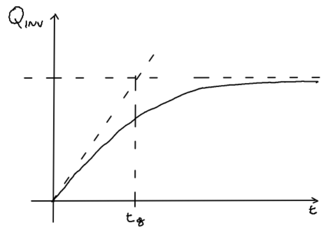
For signals with high frequency ($f > 1Hz$) we cannot change electron concentration so we can only add or remove holes at the substrate making the depletion layer bigger or smaller. The graph of C_g remains constant form the threshold voltage on. The total capacitnace depends on the frequency of the small signal we are appling. Here the modulation of δV_g can't modulate the electron concentration at the surface of silicon. There is a modulation of the depletion layer so δV_s but for the continuity of the electric field olso δV_{ox} have to change in a comparable manner. With the downwards band banding we have to keep constant the distance with E_{fn} with the conduction band so our substrate is no longer at th.eq. This triggers GR processes that tends to move the system back in the th.eq of low freqeuncy. E_{fn} is constant over the depletion layer and merges with E_{fp} in the quasi neutral region of the device. We have a flow of electrons from the substrate to the surface similar to the pn-j reverse current.

4.4.3 Deep deplition



If we consider a step increase of the gate voltage form 0 to $\overline{V}_g > V_t$ we don't give the time to the system to create electrons so we get a super wide depletion layer $W_d >> W_d^{max}$. This is called deep depletion condition of the MOS capacitor and it's a large signal regime. If we apply on this step increase a δV_g this will modulate the width of the depletion layer (in low and high freqence too) beacuse we don't have electron at the interface. The capacitance is $C_s = C_{dep} = \varepsilon_{si}/W_d$ so the total capacitance will decrease over V.
W_d is related to V_s so $W_d > W_d^{max}$ means $V_s > 2|\phi_b|$ the conduction band can bend even below the fermi level but we have to keep the condition of no electrons at the interface so we have the split in the two Fermi

level for electrons and holes. Small voltage drop ove the oxide due to big voltage drop over the substrate. GR processes will pull the system to the LF regime in order to restore the equilibrimu so with a Q_{inv} of electrons at the surface.



Let's study the behaviour in time of this charge under the simplifing assumption that Q_s remains constant: so by S-R-H statisic

$$\frac{dQ_{inv}}{dt} = -qG[W_d(t) - W_d^{max}] = -q\frac{n_i}{2\tau_0}[W_d(t) - W_d^{max}] \quad (4.32)$$

but $Q_s = Q_{inv} + Q_{dep} = Q_{inv} - qN_aW_d(t)$
from here we get that $W_d(t) = (Q_{inv} - Q_s)/(qN_a)$ where Q_s is a constant so we get the following differential equation

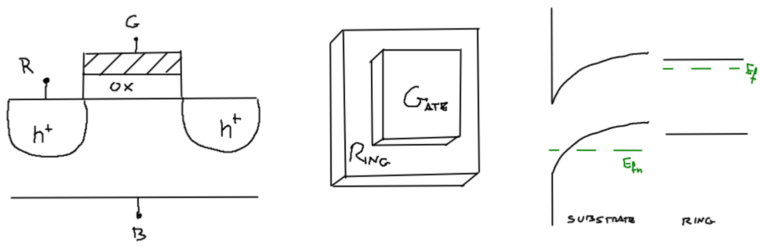
$$\frac{dQ_{inv}}{dt} = \frac{n_i}{2\tau_0 N_a}Q_{inv} + \frac{qn_i}{2\tau_0}[\frac{Q_s}{qN_a} + W_d^{max}] \quad (4.33)$$

that is a simple first order differential equation which solution is

$$Q_{inv}(t) = [Q_s + qN_aW_d^{max}][1 - e^{t/t_g}] \qquad t_g = \frac{2\tau_0 N_a}{n_i} \quad (4.34)$$

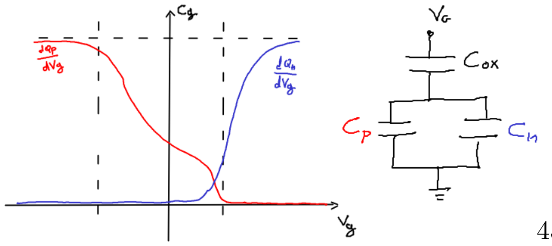
so an exponential increase of the charge with a time constant t_g

4.5 MOS capacitor with ring



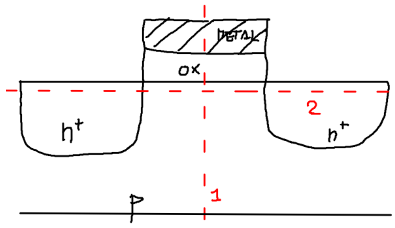
We want to modify the structure adding a contact that can provide a lot of electrons. This contact is calle ring it's n^+ doped. With $V_g > V_t$ E_{fn} goes down but the Fermi level in the ring is higher than it so there is a flow of electrons from the ring to the substrate with a time constant similar to the transient time from the 2 zones of the electrons.

4.5.1 Split C-V curve



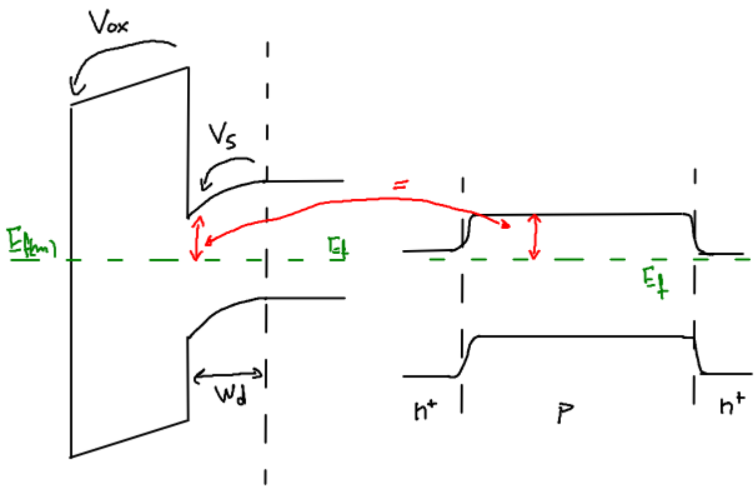
With this structure it's possible to measure the current of e and h indipendently and split the 2 contributions. Let's assume that in accumulation regime and weak inversion we modulate only dQ_p no dQ_n flow so the capacitance is only $\frac{dQ_p}{dV_g} = C_g$ beacuse $\frac{dQ_n}{dV_g} = 0$. At stong inversion we modulate only electrons beacuse we reach the maximum width of the depletion layer so $\frac{dQ_p}{dV_g} = 0$ and $\frac{dQ_n}{dV_g} = C_g$. In this way we are splitting the total gate capacitance in 2 contributions due to holes and electrons.

4.5.2 Electrostatic



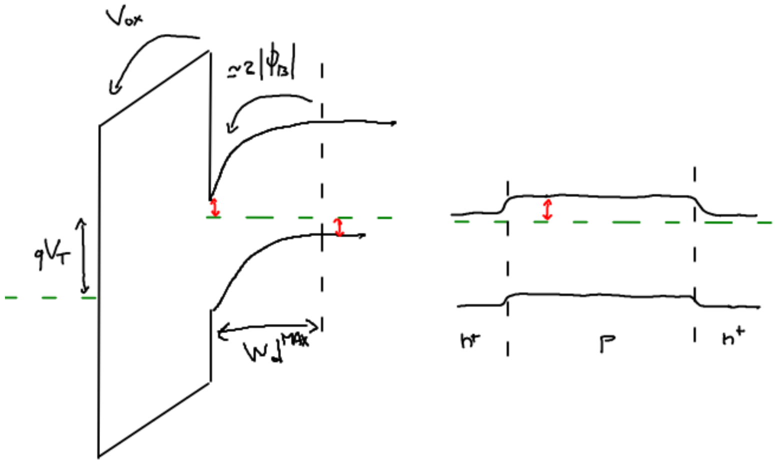
Beacuse the MOS capacitor is an intrinsic 2D device we have to study the electrostatics in the 2 directions of the graph.
We are not intrested in polrization with $V_r < 0$ beause this will make the pn junction in forward bias.

$\rightarrow V_r = V_g = 0$



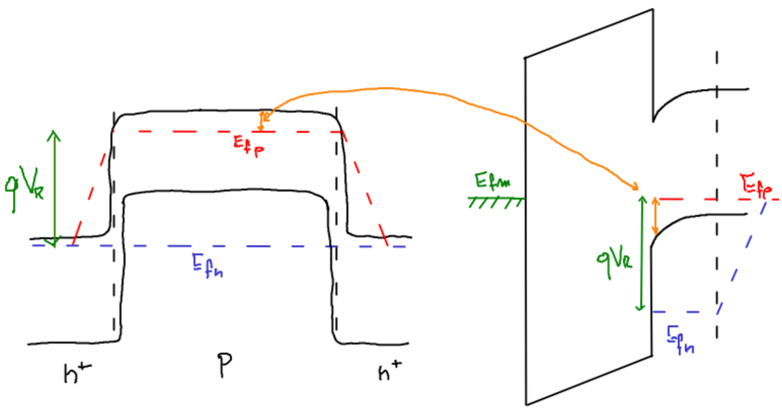
We are under th.eq we are not perturbing the device. In crossection 1 we have weak inversion. In crossection 2 we have 1 single Fermi level for th.eq from crossection 1 we know the distance between the conduction band and E_f in the substrate. The region under the gate in graph 2 is depleted

$\rightarrow V_r = 0 \quad V_g = V_t$



First graph is the one we have already seen. In crossection 2 we make the same consideration as before.

$\rightarrow V_r > 0 \quad V_g = 0$

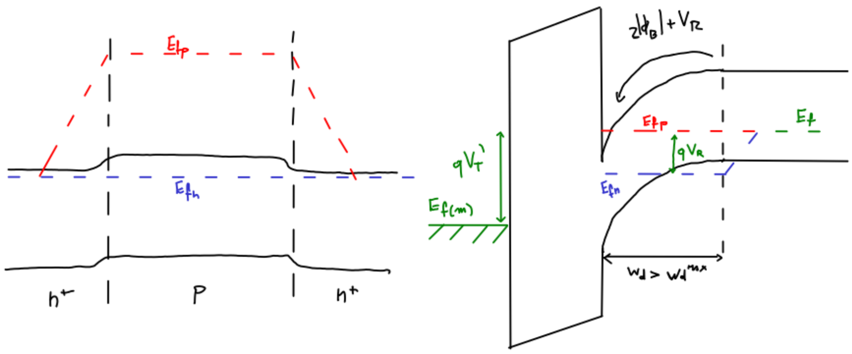


We have no change in crossection 1 wrt the first case. For second crossection we have to set the distance $E_c - E_f$ under the gate in accord to the first graph. In the n^+ region we are applying a potential that bands the band downwards by an amount qV_r giving a large band banding and so a separation of the Fermi levels.
The quasi Fermi level for electrons has to stay constant for simmetry E_{fp} changes.
Form this we have to modify graph 1 where E_{fn} is moved downwards by an amount qV_r (doing this we do not change the electrostatic beacuse the electron concetration was already negligible so no other modifications in the band banding are needed).

$\rightarrow V_r > 0 \quad V_g = V_t$

Crossection 1 same as the case of $V_r = 0 V_g = V_t$ with the same step of before we get the separation of the quasi Fermi levels in the first graph and the same graph for crossection 2 of the previous case.

In this case we are no more in strong inversion due to the shift downwards of E_{fn} of qV_r. To gain the strong inversion regime we have to apply a stornger gate bias $V_g = V_t'$ to have $V_s = 2|\phi_b| + V_r$ and of course a wider depletion layer.

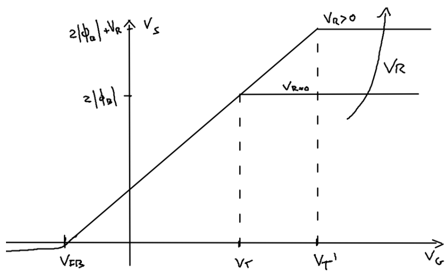


We prolong the depletion regime applying $V_r > 0$ changing the on-set of the strong inversion regime to

$$V_t' = V_{fb} + 2|\phi_b| + V_r + \frac{\sqrt{2\varepsilon_{si} q N_a (2|\phi_b| + V_r)}}{C_{ox}} \tag{4.35}$$

adding and removing $\frac{\sqrt{2\varepsilon_{si} q N_a 2|\phi_b|}}{C_{ox}}$ we obtain

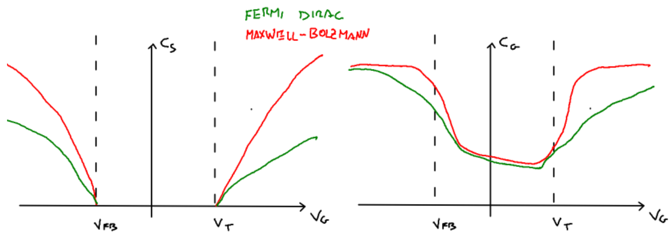
$$V_t' = V_t(V_r = 0) + V_r + \frac{\sqrt{2\varepsilon_{si} q N_a}}{C_{ox}}[\sqrt{2|\phi_b| + V_r} - \sqrt{2|\phi_b|}] \tag{4.36}$$



The charge in the substrate will be from the relation $V_g - V_{fb} = V_s - Q_s/C_{ox}$

$$Q_{inv} = Q_s - Q_{dep}^{max} = -C_{ox}(V_g - V_{fb} - 2|\phi_b| - V_r - |Q_{dep}^{max}|/C_{ox}) = -C_{ox}(V_g - V_t') \tag{4.37}$$

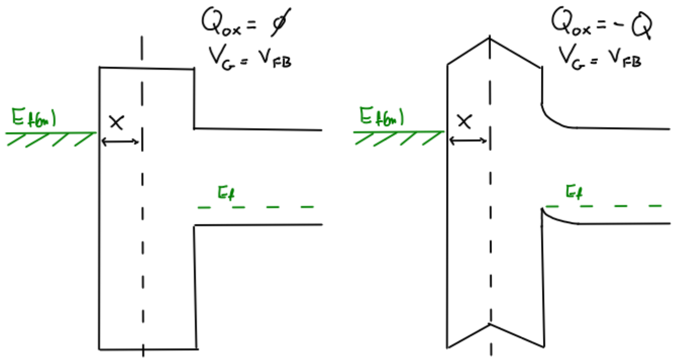
In accumulation and strong inversion the MB approximation isn't valid and gives us a big error wrt the FD. Below there are the graph change for the 2 distribution for $C_s - V_g$ and $C_s - V_g$.



4.6 Effects of non idealities

4.6.1 Charge in the oxide

Charge in the oxide is caused due to spurious states inside the dialectric or difects that trap charges furthermore ageing increase the density of this imperfections.
Let's assume a MOS capacitor at the flat band condition with a negative charge sheet at distance x from the gate.



The negative charge decreases the potential so bands upwards the bands. We have a linear banding beacuse the charge is concentrated in one point.
We have a layer of hole at the the substrate and olso a positive charge on the gate (electrostatic induction tends to nihil the charge). The important fact is that $V_s \neq 0$ with the flat-band voltage.
To recover the flat band state we have to decrease $E_{f(m)}$ so increase the gate voltage; the new flat band voltage will be $V'_{fb} = V_{fb} + \Delta V_g$. With Gauss law we get

$$\Delta V_g = -\frac{Q_{ox}}{\varepsilon_{ox}} x \tag{4.38}$$

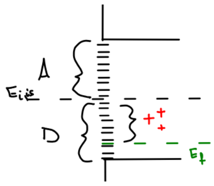
This expression can be re-writed introducing the parameter $C_{xg} = \varepsilon_{ox}/x$ the capacitance between x and the gate as

$$\Delta V_g = -\frac{Q_{ox}}{C_{xg}} \tag{4.39}$$

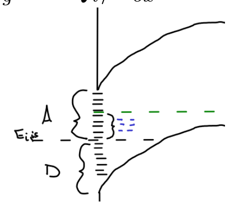
The maximum value for this voltage variation is when $x = t_{ox} \rightarrow C_{xg} = C_{ox}$. This is reasonable result lower the distance higher the electrostatic induction.
The above results are general and valid for all working regimes of the MOS capacitor.
The effect of this phenomena results in the C-V curve as a shift rightwards if the charge is negative, leftwards if the charge is positive, of ΔV_g.
If we get a distribution of charge per unti volume of ρ we get

$$\Delta V_g = \int_0^{t_{ox}} \frac{x\rho}{\varepsilon_{ox}} dx \tag{4.40}$$

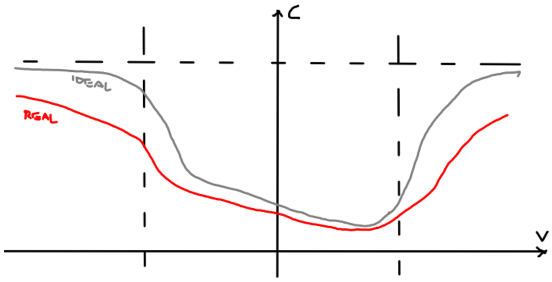
4.6.2 Interface states



First we will consider the flat-band condition.
Acceptor states are all empty so no charge as the donor states below E_f that are all filled. There are some donor empty that gives us a positive charge (that make the potential higher and so decrease the bands ??). Following the results of the charge sheet in the oxide we can restore the ideal case by appling a $\Delta V_g = -Q_i/C_{ox}$.

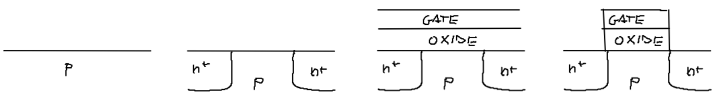


In strong inversion we have a negative charge exposed that can be compensated with an extra voltage as before but with opposite sign.
In accumulation and flat-band regime we have a negative increase so a shift leftwards of the C-V curve for the stron inversion vice versa so we have a strech of the total C-V characteristic.

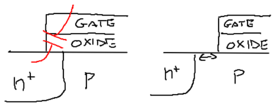


4.7 Impact of polysilicon gate

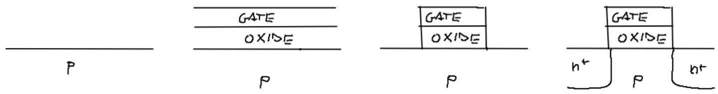
The benefit of usign the polisilicon gate comes from a tecnological issues.
First way to do ring-MOS was by create a p-substrate, create the n^+ regions, let the oxide deposit all over, deposit the gate and then remove by etching the excess of gate and oxide. This type of process is a non self allign process.



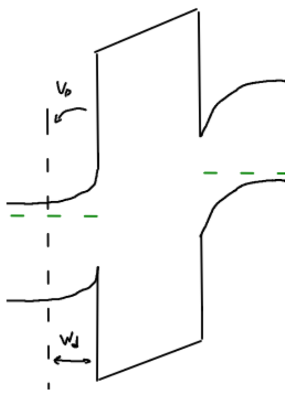
Processes tollerances define where the gate is placed. If the gate is over the n^+ zone there will be a big parassite capacitance but on the other hand with the gate far from that zone there will be no control of electrons by the gate.



A better process is : p-substrate, deposit the oxide, deposit the gate, etching excess and than n-dope. This is a self alligned process.



To dope we have to do a heat threatment at 800C so this process does not works with metals (melting point issues) but works with polysilicon.
Olso another improvmen comes from the threshold voltage $V_t = V_{fb} + 2|\phi_b| + |Q_{dep}^{max}|/C_{ox}$ that with a poly-p-doped gate is $\simeq 0$ but with a n-doped-poly is reduced to $V_t = |Q_{dep}^{max}|/C_{ox}$ **check this last**.
The problem with this gate material is that we have to consider the charge as for unit volume not as untit area (we don't have that much electrons as in a metal) so doing we get a depletion layer olso in the gate of $W_d = \sqrt{\frac{2\varepsilon_{si}V_p}{qN_d^g}}$ and so an additional voltage drop V_p and an additional capacitance.



The voltage ballance is now $V_g - V_{fb} = V_s + V_{ox} + V_p$ deriving wrt $-Q_s$ we get

$$\frac{1}{C_{ox}} = \frac{1}{C_s} + \frac{1}{C_{ox}} + \frac{1}{C_g} \tag{4.41}$$

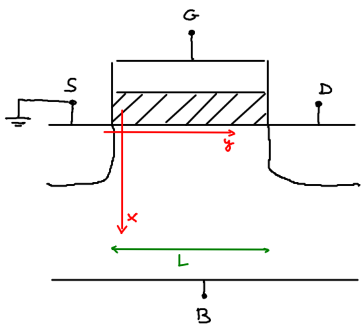
The C-V curve doesn't change until strong inversion when this phenomena becomes dominant. We can write that

$$\frac{1}{C_g} = \frac{2kT/q}{Q_p} + \frac{1}{C_{ox}} + \frac{Q_p}{qN_d^g\varepsilon_{si}} \tag{4.42}$$

check expression above the last term becomes dominant in strong inversion.
This is why we now use again metal for gate (with particular mix with high melting point).

Chapter 5

MOS transistor



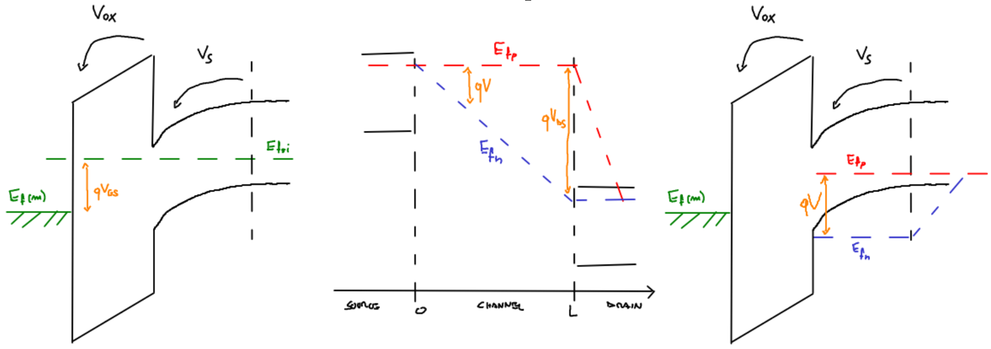
We start our analysis of the MOS-transistor with the hypotesis of an ideal dialectric (SiO_2) and constant doping concentrations.
This is an intrinsically bi-dimensional device we will study it in the crossection showed in figure and assuming that all is omogenuos in the ortogonal direction.
The most important region of the device is the "channel region" of lenght L that is the one under the gate between the 2 n-doped regions. We will consider in our analysis the 2 x and y axes placed as in figure.
We want to study the electrostatic of this device changing the bias of the four pin; withouth lack of generality since we are intrested in the voltage difference we ground the source for all our analysis.

Working principles

If V_{gs} is low (until depletion regime) there is no interation between source and drain but if we move to higher V_{gs} (inversion region) we get a lot of electrons from source to drain for the inversion layer so a flow of current it's possible under the gate. The behaviour of the device is like a resistor.
Holes have a minor role in the behaviour of the device this is why the MOS transistor, also called field effect unipolar transistor, is an unipolar device.

5.1 Gradual channel approximation

To study the electrostatic of this device we have to ideally solve the bi-dimentional Poisson equation but we can simplify our analysis studing only long channel MOS: we define a channel "long" if we can use the gradual channel approximation.
The gradual channel approximation says that the electrostatic in channel region is a quasi 1-D electrostatic and so we can approximate every vertical crossection as a MOS capacitor. If L is large the electric field in the y direction is very small and the dominant electric fiel is in the vertical direction.
We can use the resuslts we have from the MOS capacitor.



Assuming $V_{gs} > 0$ and $V_{ds} > 0$ along the first crossection we get what we've sudied with the MOS capacitor but in the second crossection we have a split of the 2 quasi-Fermi levels due to the shift downwards of the drain bands of an amount of qV_{ds} so the substrate is no longer at th. eq. and so also the frist graph have to be changed. E_{fp} remains constant over the channel but E_{fn} changes.
Placing the 0 level to E_{fp} that doesn't change we introduce the parameter

$$V = -\frac{E_{fn}(y)}{q} = V(y) \tag{5.1}$$

V goes from 0 at the source to V_{ds} at the drain.

5.2 Poisson equation

Looking at the last band diagram we plotted the poisson equation it's very similar to the one of an MOS capacitor the only difference is in the n term that now is

$$n = \frac{n_i^2}{N_a} e^{q\Delta\phi/kT} e^{-qV/kT} \tag{5.2}$$

The last exponential takes into account the decrease of the quasi-Fermi level for electrons that is far from the conduction band wrt the MOS capacitor case.
If $n(0) = \frac{n_i^2}{N_a} e^{qV_s/kT} e^{-qV/kT}$ let's consider $n(0) = N_a$ we get $e^{-2|\phi_b|/kT} e^{-qV/kT} e^{qV_s/kT} = 1$ and so

$$-2|\phi_b| - V_s + V = 0 \qquad V_s = 2|\phi_b| + V \tag{5.3}$$

check all up passage.

Repeating all calcualtion done for the expression we get is similar to the one with the mos capacitor

$$\left(\frac{d\Delta\phi}{dx}\right) = \frac{2kTN_a}{\varepsilon_{si}}[e^{-q\Delta\phi/kT} + \frac{q\Delta\phi}{kT} - 1 + \frac{n_i^2}{N_a^2}\left(e^{-qV/kT}(e^{q\Delta\phi/kT} - 1) - \frac{q\Delta\phi}{kT}\right)] \tag{5.4}$$

we change the term that takes into account the electron concentration.
With this result we know $F_x^2 = \left(\frac{d\Delta\phi}{dx}\right)^2$ and that $F_x(x,y) = F_x(\Delta\phi, V)$.

5.3 Charge in the substrate

At the silicon surface the field is $F_s = F(0,y) = F(V_s, V)$ from this with the Gauss law we can say that the charge in the substrate is

$$Q_s \simeq -\sqrt{2\varepsilon_{si}kTN_a}\left(\frac{qV_s}{kT} + \frac{n_i^2}{N_a^2}e^{q(V_s-V)/kT}\right)^{1/2} \tag{5.5}$$

Now to divide the contribution coming from the inversion charge and from the depletion layer we have to refer to the different regimes.

5.3.1 Charge in weak inversion

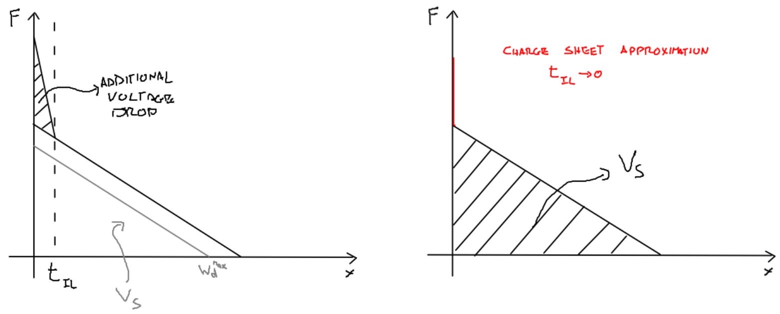
In weak inversion we have that

$$Q_{inv} = -\sqrt{\frac{\varepsilon_{si}qN_a}{2V_s}}\frac{kT}{q}\frac{n_i^2}{N_a^2}e^{q(V_s-V)/kT} \qquad Q_{dep} = -\sqrt{2\varepsilon_{si}qN_aV_s} \tag{5.6}$$

In the y axes E_c, E_v are flat and $F_y^{weak.i} \simeq 0$ so current transport is a diffusion process and the flow of holes is negligible.

5.3.2 Charge in strong inversion

The expression seen before for the depletion charge is no longer valid in strong inversion becuse the concentration of electrons isn't negligible and the inversion layer charge chanege the F profile



To recover that expression we can make a charge sheet approximation that is considering the thikness of the inversion layer so small to be negligible to have a graph like in figure above.
Making this approximation we conclude that

$$Q_{dep} - \sqrt{2\varepsilon_{si}qN_aV_s} \qquad Q_{inv} = Q_s - \sqrt{2\varepsilon_{si}qN_aV_s} \tag{5.7}$$

In stong inversion the electrons concentration is no more negligible and at the drain $E_{fn} < E_{fp}$ and bands move downwards form source to drain as the quasi-Fermi level for electrons so there will be a drift flow of electrons and a relevan electric fiels in the y direction.

5.4 Current behaviour

In order to arrive to $V_s(V_{gd}, V_{ds})$ we have to use not only $V_{gs} + V_{fb} = V_s + V_{ox} = V_s - Q_s/C_{ox}$ but olso the continuity equation for electrons beacuse the substrate isn't at th.eq.
We assume the device stationary and the GR processes negligible so J_n is constant in the channel region. From the interface with the oxide to the substrate the current will decrease due to the decrease of electron concentration but form S-D J will remains constant beacuse the direction is parallel to the current flow so we can say that

$$I_{ds} = -W \int_0^{t_{inv}} J_n(x)dx \tag{5.8}$$

where t_{inv} is the thikness of the inversion layer and we simply multiply by W beacuse we supposed that everything is constant in that direction. The minus sign is becuse we want positive the current that enter in the drain.
So taking into account drift and diffusion process we have

$$I_{ds} = -W \int_0^{t_{inv}} n\mu_n \frac{dE_{fn}}{dx} dx \tag{5.9}$$

we know that $\frac{dE_{fn}}{dx} = -q\frac{dV}{dy}$ so solving the integral we get

$$I_{ds} = -\mu_n W \frac{dV}{dy} Q_{inv} \tag{5.10}$$

This is the differential form of the drain current in MOS transistor and making another step of integration we come to

$$I_{ds} = -\mu_n \frac{W}{L} \int_0^{V_{ds}} Q_{inv} dV \tag{5.11}$$

the integral form for drain current. This expression are general.
Let's now foucs on storng inversion regime.
In this regime we've seen that using the charge sheet approximation $Q_{inv} = Q_s - \sqrt{2\varepsilon_{si} q N_a V_s}$ if we combine this equation with $Q_{dep} - \sqrt{2\varepsilon_{si} q N_a V_s}$ and with $V_{gs} + V_{fb} = V_s - Q_s/C_{ox}$ we get that $V_s \simeq 2|\phi_b| + V$ and $Q_s = -C_{ox}(V_{gs} - V_{fb} - V_s)$ if we use this two equation with the first one (inversion charge) we get that

$$Q_{inv} \simeq -C_{ox}(V_{gs} - V_{fb} - 2|\phi_b| - V) + \sqrt{2\varepsilon_{si} q N_a (2|\phi_b| + V)} \tag{5.12}$$

we will not solve this equation but we will consider some different regimes

5.4.1 Ohmic regime $V_{ds} << 2|\phi_b|$

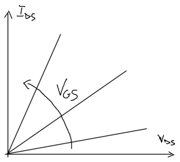
If V_{ds} is small olso V will be small so we can neglect that contribution

$$Q_{inv} \simeq -C_{ox}(V_{gs} - V_{fb} - 2|\phi_b| - \frac{Q_{dep}^{max}}{C_{ox}}) \simeq -C_{ox}(V_{gs} - V_t) \tag{5.13}$$

so putting this result in the integral form for the current and solving the integral we get

$$I_{ds} = \mu_n C_{ox} \frac{W}{L} (V_{gs} - V_t) V_{ds} \tag{5.14}$$

This is the expression for ohmic regime the current is proportional to V_{ds}.



The inversion charge does not change from source to drain and beacuse V_{ds} is very small the electrostatic is almost uniform along the channel.
It's like having a resistor in the channel
that can be modulate by V_{gs}. We can define a resistivity for the channel as

$$R_{ch} = (\frac{\partial I_{ds}}{\partial V_{ds}})^{-1} = \frac{1}{\mu_n C_{ox} (V_{gs} - V_t)} \frac{L}{W} = \rho_{sh}^{ch} \frac{L}{W} \tag{5.15}$$

We espect the field constatn and so equal to V_{ds}/L.

5.4.2 Parabolic regime $V_{ds} < 2|\phi_b|$

Now V_{ds} is no more negliglibe but enought small to make a Taylor expansion of the second terms as

$$Q_{inv} \simeq -C_{ox}(V_{gs} - V_{fb} - 2|\phi_b| - V) + \sqrt{2\varepsilon_{si} q N_a 2|\phi_b|}(1 + \frac{1}{2}\frac{V}{2|\phi_b|}) = -[...] + \sqrt{\frac{\varepsilon_{si} q N_a}{4|\phi_b|}} V = -[...] + |Q_{dep}^{max}| + C_{dep} V \tag{5.16}$$

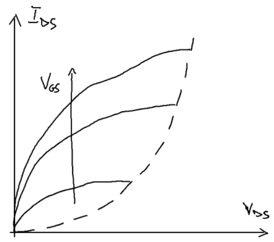
and so re-writing the equation

$$Q_{inv} \simeq -C_{ox}(V_{gs} - V_{fb} - 2|\phi_b| - \frac{Q_{dep}^{max}}{C_{ox}} - V(1 + \frac{C_{dep}}{C_{ox}})) = -C_{ox}(V_{gs} - V_t - mV) \tag{5.17}$$

The term -mV is an increase of the threshold voltage from source to drain.
The inversion charge is decreasing from source to drain; putting this equation in the integral we get

$$I_{ds} = \mu_n C_{ox} \frac{W}{L} [(V_{gs} - V_t) V_{ds} - \frac{m V_{ds}^2}{2}] \tag{5.18}$$

we have a parabolic behaviour of the current. If V_{ds} is small we can neglect the quadratic term and so recovering the previous condition.



5.4.3 Saturation point

We can call the current and voltage of the vertex of the parabola saturation's one

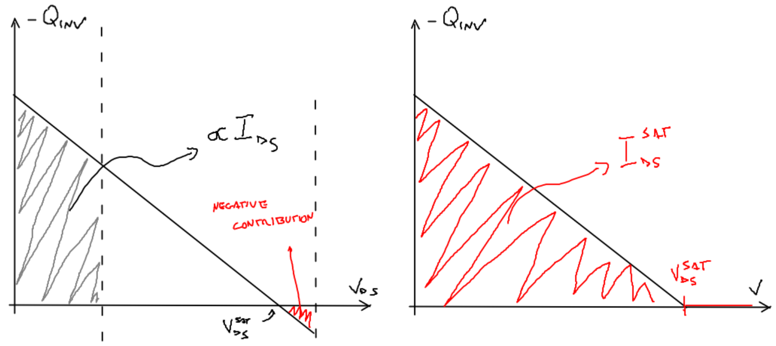
$$V_{ds}^{sat} = \frac{V_{gs} - v_t}{m} \qquad I_{ds}^{sat} = \mu_n C_{ox} \frac{W}{L} \frac{(V_{gs} - V_t)^2}{m} = \mu_n C_{ox} \frac{W}{L} \frac{(V_{ds}^{sat})^2}{m} \tag{5.19}$$

Increasing V_{ds} we increase the electric field and decrease the charge (so we get a more resistive channel). At V_{ds}^{sat} the charge at the drain is

$$Q(y = L) = -C_{ox}(V_{gs} - V_t - m\frac{V_{gs} - V_t}{m}) = 0 \tag{5.20}$$

this is called the pinch off conditon of the transistor. If from the saturation point on we keep using the previous equation we get a reduction of the current that has no sense.

Mathematical explanation of dicendent parabola



Let's consider $-Q_{inv}(V_{ds})$ we calculated the area of this expression to find the current so the area under the rect in the graph is proportional to the current. When $V_{ds} = V_{ds}^{sat}$ we are at the vertex of the parabola and so the current is proportional to the area of the triangle.
From there on we add negative area contribution that has no sense it's like accumuling holes in the inversion layer. At the drain we were at the limit of strong inversion regime if we increase V_{ds} over the saturation point we enter at the drain at weak inversion regime where electrons concentration can be apporoximated as negligible Q_{inv}. The real graph of the inversion charge becomes zero over the saturation voltage so from the vertex on the current remains constant.

Physical explanation of dicendent parabola

pag 75 tutta riguardare non ho capito
The parameter $m = 1 + \frac{C_{dep}}{C_{ox}}$ involves doping concentration and oxide thikness and express how much of V_g drops over the substrate; for a good mos m=1,1-1,4.

5.5 Over the saturation point

We still don't know E_{fn} along the channel or the voltage drop in function of y.
From the differential from of the drain current we have $I_{ds} = -\mu_n W \frac{dV}{dy} Q_{inv}$ we can integrate this expression up to a potential V in a point y

$$\int_0^y I_{ds} dy = \int_0^V -\mu_n W Q_{inv} dV \tag{5.21}$$

$$I_{ds} y = \int_0^V \mu_n W C_{ox}(V_{gs} - V_t - mV) dV = \mu_n W C_{ox}[(V_{gs} - V_t)V - \frac{mV^2}{2}] \tag{5.22}$$

but we know from the parabolic regime another expression for
$I_{ds} = \mu_n \frac{W}{L} C_{ox}[(V_{gs} - V_t)V_{ds} - \frac{mV_{ds}^2}{2}]$ so we can get V(y) as

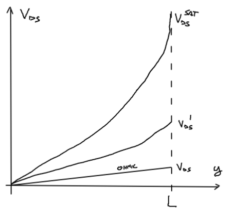
$$V(y) = \frac{V_{gs} - V_t}{m} - \sqrt{\left(\frac{V_{gs} - V_t}{m}\right)^2 - \frac{2y}{L}\left(\frac{V_{gs} - V_t}{m}\right)V_{ds} + \frac{y}{L}V_{ds}^2} \tag{5.23}$$

from that we know E_{fn} beacuse $V = -E_{fn}/q$.
This result is general but we will study it in the various regime; in ohmic region V_{ds} is small so we can neglect the quadratic term and extract the $\frac{V_{gs}-V_t}{m}$ term from the sqare root, in this way making a Taylor expansion we get

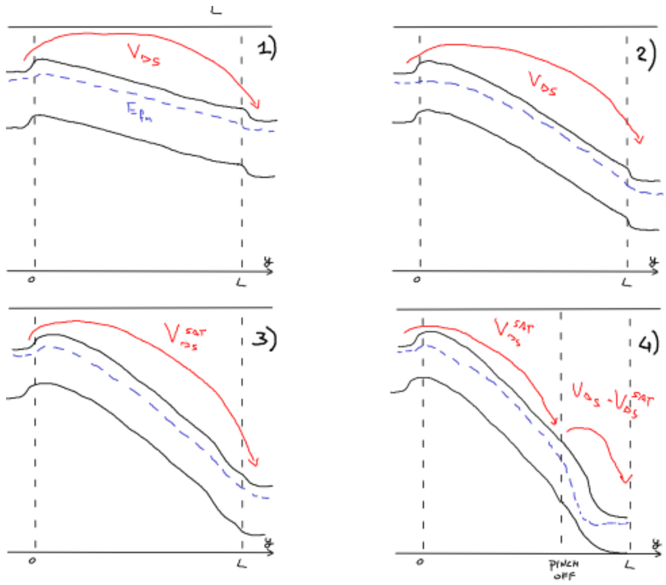
$$V(y) \simeq \frac{V_{gs} - V_t}{m} - \frac{V_{gs} - V_t}{m}\sqrt{1 - \frac{2y}{L}V_{ds}/\left(\frac{V_{gs} - V_t}{m}\right)} \simeq \frac{V_{gs} - V_t}{m} - \frac{V_{gs} - V_t}{m}\left(1 - \frac{y}{L}V_{ds}/\left(\frac{V_{gs} - V_t}{m}\right)\right) \tag{5.24}$$

$$V \simeq \frac{V_{ds}}{L}y \tag{5.25}$$

resonable result when the channel is uniform and we said that $\frac{dV}{dy}$ has to be constan so V linear function of y. We get a constant drift contribution over the channel **Drift?**.

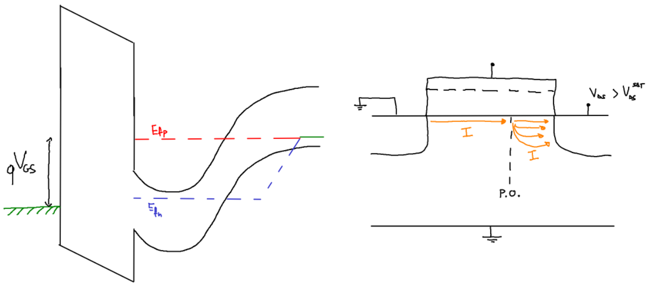


Increasing the drain voltage we enter in the parabolic regime and the charge decrease beacuse the differential form of the current must be always valid $\frac{dV}{dy}$ must increase we need higher electric field. We increase the slope of the rect form source to drain.
At pinch off point we have no charge in L so at V_{ds}^{sat} we reach L with vertical tangent. This condition means infinite electric field in y direction so F_y is no more negligible and the gradual channel approximation is no longer valid. We have to study the 2D Poisson equation. We will not study it mathematically but let's see the results.

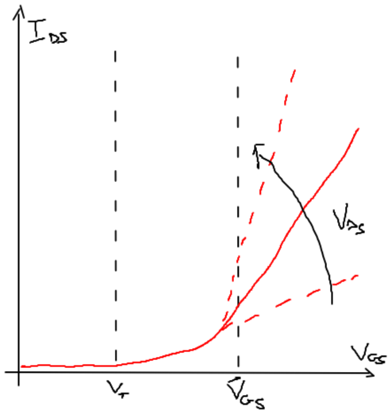


In ohmic regime V_{ds} is small so we get the 1) graph.
Increasing V_{ds} we increase the slope of E_{fn} that change like figure 2).

We reach the pinch off point large F very large tangent at L but not infinite like in figure 3). Using gradual channel approximation we sad that the drain over the saturation voltage loses control of the band banding this is unreasonable; in reality the pinch off point moves leftwards for voltages higher than saturation at draid leaving a potential V_{ds}^{sat} from source to pinch-off and form pinch-off to drain $V_{ds} - V_{ds}^{sat}$. like in figure 4). This effect acts like a reduction of the L of the channel.
In the x axis the band diagram is purely 2D band-diagram; the maximum electron concentration is no more at the interface but it's spread in the vertical direction.
Beyond the pinch-off we have a depleted region.



5.6 Transcaracteristic



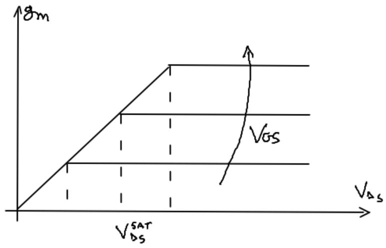
We assume V_{ds} set and we change V_{gs} and we want to know the current profile.
First region we get $V_{gs} < V_t$ so we are in accumulation or weak inversion region so the current is negligible.
When we reach $V_{gs} = V_t$ we enter in the on-state regime we are at saturation beacuse we have V_t at the source but going to the drain the thrashold voltage increase so we don't have threshold there. $I_{ds}^{sat} = \mu_n C_{ox} \frac{W}{L} \frac{(V_{gs}-V_t)^2}{2m}$ so we have a parabolic dependance.
We increase V_{gs} until we reach the threshold condition also at the drain at the voltage $V_t' = V_t + mV_{ds} = \overline{V_{gs}}$ in this case the current is $I_{ds} = \mu_n C_{ox} \frac{W}{L} [(V_{gs} - V_t)V_{ds} - \frac{mV_{ds}^2}{2}]$ so a linear dependance.

5.7 Small signal parameters

5.7.1 Transonductance g_m

If we slightly move V_{gs} we change the current flow so we can define a small signal transconductance as $g_m = \left(\frac{\partial I_{ds}}{\partial V_{gs}}\right)_{V_{ds}}$ that in ohmic-parabolic and saturation regime is

$$g_m^{ohmic} = \mu_n C_{ox} \frac{W}{L} V_{ds} \qquad g_m^{sat} = \frac{1}{2}\mu_n C_{ox} \frac{W}{L} \frac{V_{gs} - V_t}{m} \tag{5.26}$$

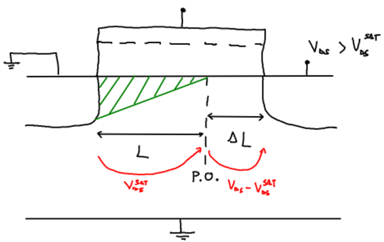


5.7.2 Transconductrance g_d

We can define also $g_d = \left(\frac{\partial I_{ds}}{\partial V_{ds}}\right)_{V_{gs}}$ that for ohmic and parabolic regime is

$$g_d^{ohmic} = \mu_n C_{ox} \frac{W}{L}[(V_{gs} - V_t) - mV_{ds}] \qquad g_d^{sat} = 0 \tag{5.27}$$

This last result is obviously an approximation becuse we implicitedly used the gradual channel approximation that is not valid in saturation regime.
Using 2D Poisson equation we get that the pinch-off point moves from drain to source if $V_{ds} > V_{ds}^{sat}$; this phenomena is equivalent to a reduction of the L of the transistor that means an increase of current.



For the sake of semplicity we suppose $\Delta L \propto V_{ds}$ with direct proportionality so

$$\Delta L = \frac{V_{ds} - V_{ds}^{sat}}{F_p} \tag{5.28}$$

where F_p is the average field in the ΔL region. Moving to larger V_{ds} than V_{ds}^{sat} we can write

$$I_{ds} = I_{ds}^{sat} \frac{L}{L - \Delta L} = I_{ds}^{sat} \frac{1}{1 - \frac{\Delta L}{L}} \tag{5.29}$$

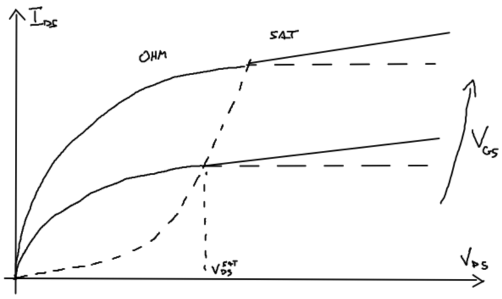
refering to a long channel device ΔL is very small compared to L so we can make a Taylor expansion obtaining

$$I_{ds} = I_{ds}^{sat}(1 + \frac{\Delta L}{L}) = I_{ds}^{sat}(1 + \frac{V_{ds} - V_{ds}^{sat}}{LF_p} \tag{5.30}$$

we get a linear increase of the current increasing V_{ds} and olso a larger slope ot the rect with higher currents.
So the real output resistance is

$$g_d = \left(\frac{\partial I_{ds}}{\partial V_{ds}}\right)^{-1}_{V_{gs}} = \frac{LF_p}{I_{ds}^{sat}} \tag{5.31}$$

that is a strong function of L.



5.7.3 Capacitive terms

In this section we will introduce only intrinsic capacitive terms.
We define Q_c total charge in the inversion layer as

$$Q_c = W \int_0^L Q_{inv} dy = [C] \tag{5.32}$$

if we change the integration variable by multipling and dividing by dV we get

$$Q_c = W \int_0^{V_{ds}} Q_{inv} dV \frac{dy}{dV} \tag{5.33}$$

from the differential form of the drain current we know that $I_{ds} = -\mu_n W \frac{dV}{dy} Q_{inv}$ so using this expression to derive $\frac{dV}{dy}$ we get

$$Q_c = -\mu_n \frac{W^2}{I_{ds}} \int_0^{V_{ds}} C_{ox}^2 (V_{gs} - V_t - mV)^2 dV \tag{5.34}$$

solving the integral and unsing the expression fro the current we get the final result

$$Q_c = -WLC'_{ox} \frac{2}{3} \frac{m^2 V_{ds}^2 + 3(V_{gs} - V_t)^2 - 3mV_{ds}(V_{gs} - V_t)}{2(V_{gs} - V_t) - mV_{ds}} \tag{5.35}$$

This formula is general and valid for all on-state regimes but we can apporximate it for the ohmic regime when $V_{ds} \simeq 0$ and for the saturation regime where $V_{ds}^{sat} = (V_{gs} - V_t)/m$ obtaining

$$Q_c = -WLC_{ox}(V_{gs} - V_t) \qquad [ohmic] \tag{5.36}$$

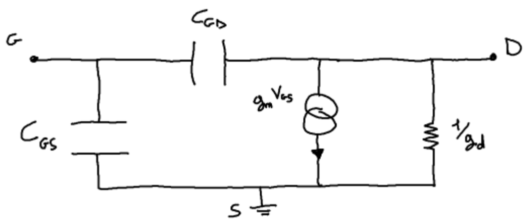
$$Q_c = \frac{2}{3} Q_c^{omhic} \qquad [saturation] \tag{5.37}$$

With this two expressions we can derive the gate and the drain capacitance as

$$C_g = -\frac{\partial Q_c}{\partial V_{gs}}|_{V_{ds}} = WLC_{ox}[1 - \frac{m^2 V_{ds}^2}{3(2(V_{gs}-V_t) - mV_{ds})^2}] \tag{5.38}$$

$$C_d = \frac{\partial Q_c}{\partial V_{ds}}|_{V_{gs}} = 2/3W/LC_{ox}[1 - \frac{)V_{gs}-V_t)^2}{[2(V_{gs}-V_t) - mV_{ds}]^2}] \tag{5.39}$$

With this parameter we can complete the small signal model of the MOS transistor



We have only to link the values C_g, C_d to the capacitance C_{gd}, C_{gs} we can do this considering a modulation of the drain/source with the other terminals grounded and probe what impedance we see; doing this procedure we get

$$C_g = C_{gs} + C_{gd} \qquad C_d = C_{gd} \tag{5.40}$$

In ohmic regime we get that $C_{gs} = C_{gd} = 1/2WLC_{ox}$ while in saturation regime $C_d = C_{gd} \simeq 0$,$C_{gs} = 2/3WLC_{ox}$; the phisical reason of this is that the drain loses control of the inversion charge only source provides electrons at the channel.

5.7.4 Frequency responce

The frequqncy responce is related to the transit time of the electrons in the channel so eh can define it as

$$\tau_t = \int_0^L \frac{dy}{v_d} \tag{5.41}$$

recovering the definition used for the narrow base diod we get $\tau_t = Q_c/I_{ds}$;beacuse Q_c isn't a linear function of I_{ds} we can only use a small signal model so

$$\tau_t = -\frac{\partial Q_c}{\partial I_{ds}} = -\frac{\partial Q_c}{\partial V_{gs}} \frac{\partial V_{gs}}{\partial I_{ds}} = C_g \frac{1}{g_m} \tag{5.42}$$

So in ohmic and saturation regimes we get

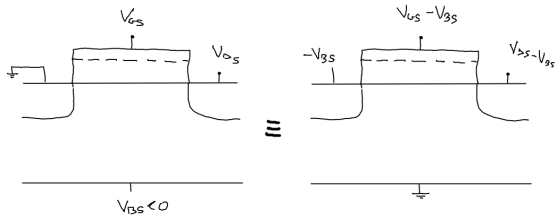
$$\tau_{ohm} = \frac{L^2}{\mu_n V_{ds}} \qquad \tau_{sat} = \frac{2}{3} \frac{L^2}{\mu_n V_{ds}} \tag{5.43}$$

Smaller L we get smaller distance and grater electric field so grater velocity (over a smaller space).
Using the cutting frequency of the transistor we find that

$$f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})} = \frac{g_m}{2\pi C_g} = \frac{1}{2\pi \tau_t} \tag{5.44}$$

5.8 Body effect

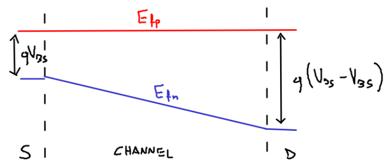
Until now we've always considered the bulk grounded now we consider the situation if we apply a $V_{bs} < 0$ (with a positive voltage we will have the 2 junction forward biased). Total voltage under the channel is now $V_{gs} - V_{bs}$. We can restore the initial condition of substrate grounded if we add at all the other pins a voltage of $-V_{bs}$ like in figure



In this case we get that $V_{gs} - V_{bs} - V_{fb} = V_s - Q_s/C_{ox}$ from this we get that che inversion charge is (under charge sheet approximation)

$$Q_{inv} = Q_s - Q_{dep} = -C_{ox}(V_{gs} - V_{bs} - V_{fb} - V_s) + \sqrt{2\varepsilon_{si} q N_a V_s} \tag{5.45}$$

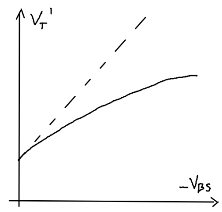
In ohmic regime $V_s \simeq 2|\phi_b| - V_{bs}$ so we get a change of the band diagram where olso at the source E_{fn} is under E_{fp}



making a little bit of calculation we get that in ohmic regime

$$Q_{inv} - C_{ox}(V_{gs} - V_{fb} - 2|\phi_b| - \frac{\sqrt{2\varepsilon_{si} q N_a (2|\phi_b| - V_{bs})}}{C_{ox}}) = -C_{ox}(V_{gs} - V_t') \tag{5.46}$$

V_t' is the new threshold voltage of the system and we get $I_{ds} = \mu_n C_{ox} \frac{W}{L}(V_{gs} - V - t')V_{ds}$ this expression is equal to the one with the bulk grounded but with another V_t. For the other regimes we get the same resuslts of bulk grounded with V_t' and for the parabolic regime m is modified.

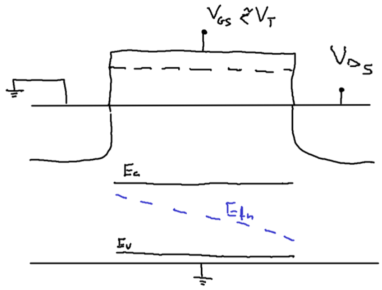


For better undestand the behaviour of $V_t'(-V_{bs})$ we can study

$$\frac{dV_t}{dV_{bs}}|_{V_{bs}=0} = [\frac{\sqrt{2\varepsilon_{si} q N_a}}{C_{ox}} \frac{1}{2} \frac{1}{\sqrt{2|\phi_b| - V_{bs}}}]|_{V_{bs}=0} = \frac{C_{dep}}{C_{ox}} = m - 1 \tag{5.47}$$

5.9 Subtheshold regime

5.9.1 Current expression



In this regime $V_{gs} \lesssim V_t$ so we don't have strong inversion anywhere in the channel, the electron concentration is negligible for electrostatics that is dominated by the depletion charge. Bands are almost flat in the channel (also if electorn concentation change beacuse it's neglibile).
We get a diffusion current of electrons the drift flow is negligible. Solving the poisson equation we got that $Q_{inv} = \sqrt{\frac{\varepsilon_{si} q N_a}{2V_s}} \frac{kT}{q} \frac{n_i^2}{N_a^2} e^{q(V_s - V)/kT}$ we can use this expression to derive the current in this regime so

$$I_{ds} = -\mu_n \frac{W}{L} \int_0^{V_{ds}} -\sqrt{\frac{\varepsilon_{si} q N_a}{2V_s}} \frac{kT}{q} \frac{n_i^2}{N_a^2} e^{q(V_s - V)/kT} dV \tag{5.48}$$

We can extract form the integral the square root and V_s beacuse they are not function of y of V so

$$I_{ds} = \mu_n \frac{W}{L} \sqrt{\frac{\varepsilon_{si} q N_a}{2V_s}} \frac{kT}{q} \frac{n_i^2}{N_a^2} e^{qV_s/kT} \int_0^{V_{ds}} e^{-qV/kT} dV \tag{5.49}$$

Solving the integral we get the following expression

$$I_{ds} = \mu_n \frac{W}{L} \sqrt{\frac{\varepsilon_{si} q N_a}{2V_s}} \frac{kT}{q} \frac{n_i^2}{N_a^2} e^{qV_s/kT} \left(1 - e^{-qV_{ds}/kT}\right) \tag{5.50}$$

we want a dependance on V_{gs} not on V_s.
As done for the MOS capacitor we use $V_{gs} - V_{fb} = V_s + \sqrt{2\varepsilon_{si} q N_a V_s}/C_{ox}$ we can assume that $V_s \lesssim 2|\phi_b|$ so we can make a Taylor expansion of $\sqrt{V_s}$ over $2|\phi_b|$ getting

$$\sqrt{V_s} \simeq \sqrt{2|\phi_b|} + \frac{1}{2} \frac{V_s - 2|\phi_b|}{\sqrt{2|\phi_b|}} \tag{5.51}$$

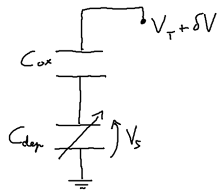
(remember f(a)+f'(a)(x-a)). Replacing this expression in the previous formula we get

$$V_{gs} = V_{fb} + V_s + \sqrt{2\varepsilon_{si} q N_a 2|\phi_b|}/C_{ox} + \frac{1}{2} \frac{V_s - 2|\phi_b|}{\sqrt{2|\phi_b|}} \sqrt{2\varepsilon_{si} q N_a}/C_{ox} \tag{5.52}$$

adding and removing $2|\phi_b|$ we can write that expression as

$$V_{gs} = V_t + (V_s - 2|\phi_b|)m \qquad \rightarrow \qquad V_s = 2|\phi_b 2|\phi_b| + \frac{V_{gs} - V_t}{m} \tag{5.53}$$

The physical meaning of this can be explained considering a system like in fiure



where we can say that $V_s = 2|\phi_b| + \delta\frac{C_{ox}}{C_{ox}+C_{dep}} = 2|\phi_b| + \frac{\delta}{m}$.
Turning back the the expression for the current we can write it in this way now

$$I_{ds} = \mu_n C_{ox}(m-1)\frac{W}{L}(\frac{kT}{q})^2 e^{q(V_{gs}-V_t)/kT}\left(1-e^{-qV_{ds}/kT}\right) \tag{5.54}$$

That is the expression for the current in the subtheshold regime. If V_{ds} is more than a few kT the last exponential can be neglected.
The current density J_n is a pure diffusion process so we can say that $J_n = qD_n\frac{dn}{dy}$ and so

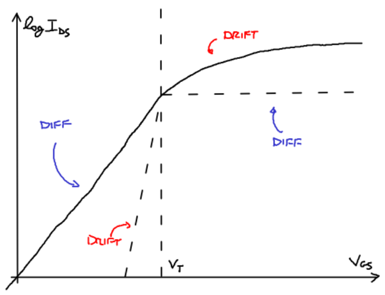
$$I_{ds} = -w\int_0^{t_{inv}} qD_n\frac{dn}{dy}dx = WD_n\frac{d}{dy}\left(\int_0^{t_{inv}} -qndx\right) = WD_n\frac{d}{dy}Q_{inv} \tag{5.55}$$

This means that Q_{inv} has to change linearly over y so that

$$I_{ds} = -WD_n\frac{Q_{inv}(0)-Q_{inv}(L)}{L} \tag{5.56}$$

Since the change of tha charge is linear E_{fn} must drop logaritmically.

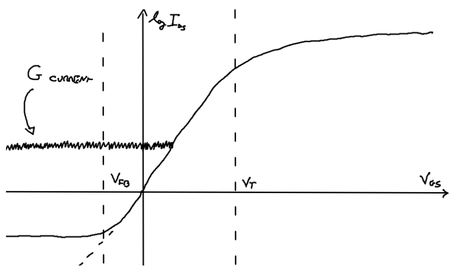
5.9.2 Transcaracteristics



Let's plot now $I_{ds}(V_{gs})$ in a logaritmic plot. We obtain for the subtheshold regime a rect with a slope defined as sub-theshold slope or STS of

$$STS = \left(\frac{\partial\log(I_{ds})}{\partial V_{gs}}\right)^{-1} = \frac{kT}{q}\ln(10)m \tag{5.57}$$

At room temperature STS=(60 mV/dec)m.
Form V_t on the diffusion contribution to the current saturates and the drift becomes dominant.
If we move to low current we break the assumption of flat band condition V_s becomes indipendent on V_{gs} so our expression is no more valid and the current will saturate to a lower value.
At low current we have to take also in account the GR processes in the channel. We have a net generation process beacuse $E_{fn} < E_{fp}$ and olso form the reversed bias junctions form drain-bulk. Generated electrons will assist the current flow. Tipically before breaking the assumption of flat-band the current saturate due to the G process.



Improvment for MOS

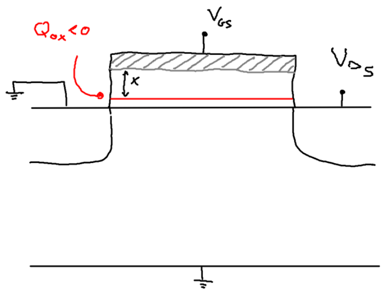
If we reduce t_{ox} the gate will have better control of the substrate and better $STS = \frac{kT}{q}\ln(10)(1 + (\varepsilon_{si}t_{ox})/(\varepsilon_{ox}W_d^{max}))$ and olso lower $V_t \propto |Q_{dep}^{max}|/(\varepsilon_{si}/t_{ox})$.
It's olso good to decrease N_a to better deplete the substrate getting better STS and lower V_t

5.10 Temperature dependance

[da fare che non ho sbatti ora diocane]

5.11 Effects of non idealities

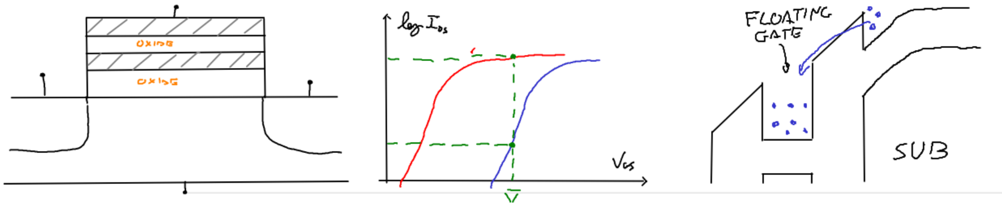
5.11.1 Charge in the oxide



We will consider a sheet of negative charge in the oxide at distance x from the gate. In MOS-capacitor we recover the ideal behaviour with the application of an extra voltage on the gate of

$$\Delta V_g = -\frac{Q_{ox}}{C_{xg}} \tag{5.58}$$

this consideration is also good for MOS-transistor. We have only a rigid shift of the transcaracteristic by an amount $\Delta V_g s$ that will be rightward if the charge is negative or lefward if positive.
This phenomena is used adding a strate of poly-silicon in the oxide called floating gate like in figure.
Depending at the charge given at a certain $\overline{V}$ we get different values of current. The charge in the floating gate is changed by quantum mechanical tunneling or other processes.

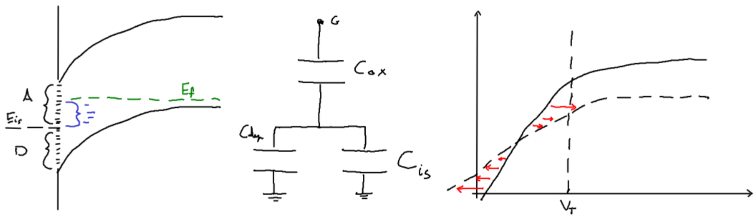


5.11.2 Interface states

We get the same consideration of the MOS-capacitor.
Now the change of charge at the surface is modulating the slope of the sub-threshold slope. It's like having another capacitance in parallel with C_dep so

$$STS = \frac{kT}{q}\ln(10)\left(1 + \frac{C_{dep} + C_{i,s}}{C_{ox}}\right) \tag{5.59}$$

We don't have a strech of the characteristic when $E_f = E_{i,s}$



5.11.3 Parassitic resistance of n^+ regions

5.11.4 Parassitic capacitances

5.12 Short channel devices

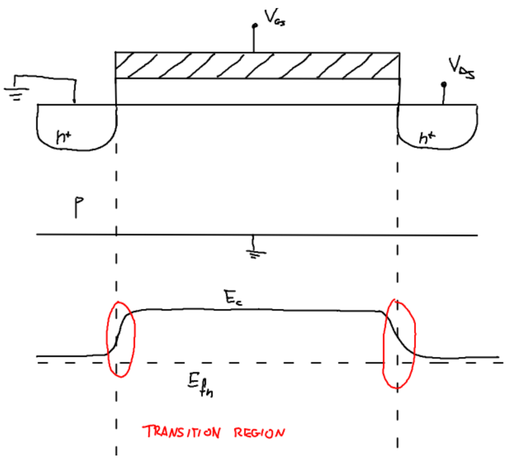
Reducing L we can increase the number of components per unit area and so decrease the costs per component and logic function. Furthermore

$$I_{ds} \propto \frac{W}{L} \qquad C \propto WL \qquad \rightarrow \qquad \frac{I_{ds}}{C} = \frac{dV}{dt} \propto \frac{1}{L^2} \tag{5.60}$$

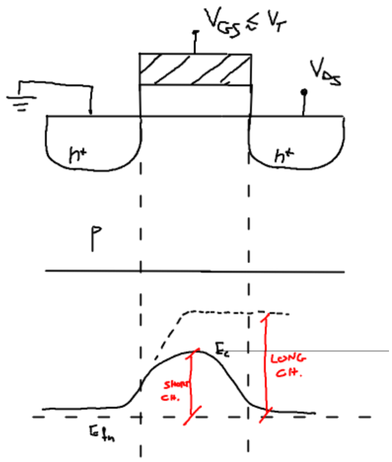
If we start reducing L leaving all the other parameters unchanged (like the oxide thikness or the doping concentrations) at the beginning the performance will increase but after a while they will become worst; this beacuse some secondary effects of the long channel MOS becomes dominants degrading the good performance of the transistor.

5.12.1 2D electrostatic in channel coming from n^+ regions

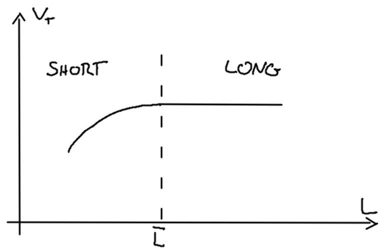
In a long channel device with drain bias equal to 0 and $V_{gs} \lesssim V_t$ bands are flat in the channel and the electron concentration is negligible.



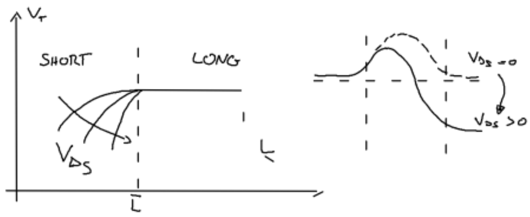
In our analysis we've always neglected the presence of the source and drain regions (that untill now only setted the value of E_{fn} at the boundary of the channel). We have 2 transition regions that we have neglected beacuse L is long and the thikness of this 2 region is negligible wrt L so they had a negligible impact.
If we reduce only L the electrostatics of the channel changes ; E_c rises but not as much as before beacuse there is the influence of the drain transition region. We have everywhere an impact of the electrostatic of the source and drain regions we don't have a zone that can be controlled only by the gate.



There are relevant electric field in y direction (gradient of E_{fn}) so an intrinsecally 2D electrostatic.
The distance $E_{fn} - E_c$ is narrower than before so we have more electrons that means we are reducing the V_t; this effect is called the short channel effect. Transistor with different L have different V_t. Variability on L will reflect on variability in V_t.

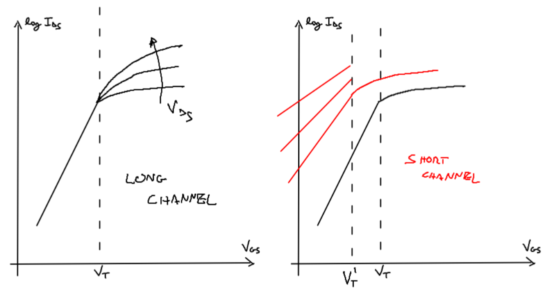


If now we increase V_{ds} in short channel we change the electrostatic and the band banding moving E_c^{max} to a lower value and near to the source as in figure.



We are further reducing the V_t of the transistor; this phenomena is called drain induced barrier lowering or DIBL. The effect is a decrease of the output resistance in the saturation regime.

In the transcharacteristic changing V_{ds} we change V_t increasing the on-state current exponentially and we have olso a worstening of the STS that has a flatter dependance beacuse increasing V_{ds} we decrease the barrier and F_x becomes negligible wrt F_y.



This goes on until we have no barrier at all the gate is no more able to turn off the device; this condition is called punch-throught regime.
From the 2D Poisson equation we get that the difference of V_t wrt a long channel device is

$$\Delta V_t = \frac{24t_{ox}}{W_d^{max}}[\sqrt{\phi_b(\phi_b - V_{ds})} - 0.4 \cdot 2|\phi_b|]e^{-\pi L/2(W_d^{max}+3t_{ox})} \tag{5.61}$$

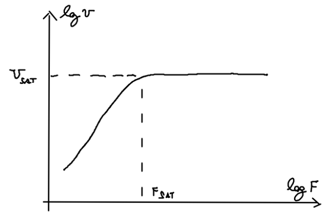
the importat term is the exponential.
To have a $\Delta V_t < 100mV$ we need to have

$$L > 2(W_d^{max} + 3t_{ox}) \tag{5.62}$$

if this condition is respected we have a long channel device if not a short channel. So we don't have an absolute value but a dependance con t_{ox}, N_a to separate long and short channel mosfets.
Reducing t_{ox} we get a better control of the gate on the substate. With higher doping concentration the transitions regions are smaller but we get also a worst STS.

5.12.2 Velocity saturation

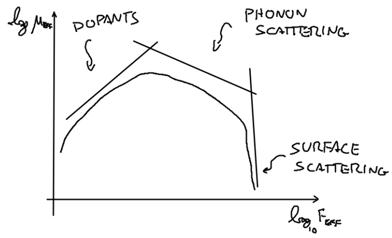
Now let's forget about previous effects seen for short channel devices (we're making a sort of sovrapposition of effects).
Reducing L we are increasing F but at very high F we the relation between the electric field and the drift velocity isn't linear ; at fields of $10^4 V/cm$ the velocity of electrons saturates at a value of $10^7 cm/s$.



To express this dependence we can use the following formula

$$v_d = \mu_{eff} \frac{F}{[1 + \left(\frac{F}{F_{sat}}\right)^n]^{1/n}} \tag{5.63}$$

where n is a coefficient equal to 1 for holes in a p-mos, 2 for electrons in n-mos.
The term μ_{eff} it's an effective mobility and depends on F_x; since the transport of electrons is very close to the surface there is a strong interaction with the interface Si-Oxide that causes surface-scattering events. The mobility in that region is much lower than that in the bulk and depends a lot on the vetical fields. We introduce olso an effective vertical electric field as $F_{eff} = \frac{Q_{dep}+0.5Q_{inv}}{\varepsilon_{si}}$ that is related to the effective mobility as in the graph.



Now for the sake of semplicity we will consider n=1 in n-mos. We want to introduce velocity saturation expression in the current $I_{ds} = -\mu_n W Q_{inv} dV/dy$. Taking into account only the on-state regime we know that $V_s = 2|\phi_b| + V$ and so that $dV_s/dy = dV/dy = -F_y$ in strong inversion regime so we arrive at

$$\mu_n \frac{d_V}{dy} = -\mu_n F_y = -v_d \qquad \rightarrow \qquad I_{ds} = W Q_{inv} v_d \tag{5.64}$$

So now we can introduce the velocity saturation equation in the expression (attention at the denominator we want a positive number so we take —F—)

$$I_{ds} = -W Q_{inv} \frac{\mu_{eff} \frac{dV}{dy}}{1 + \frac{1}{F_{sat}} \frac{dV}{dy}} \tag{5.65}$$

making some simple calculations we can write this equation like

$$I_{ds} = -[\mu_{eff} W Q_{inv} + \frac{I_{ds}}{F_{sat}}] \frac{dV}{dy} \tag{5.66}$$

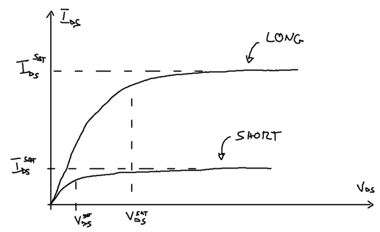
We can take dy to the first member and integrate from 0-L and form 0-V_{ds}; by assuming the quasi-1D model the only term that we really have to integrate is Q_{inv} that we can write like $Q_{inv} = -C_{ox}(V_{gs} - V_t - mV)$ **check this formula**.
Solving the integral we get

$$I_{ds} = \frac{\mu_n C_{ox} \frac{W}{L}[(V_{gs} - V_t)V_{ds} - \frac{mV_{ds}^2}{2}]}{1 + \frac{1}{F_{sat}} \frac{V_{ds}}{L}} \tag{5.67}$$

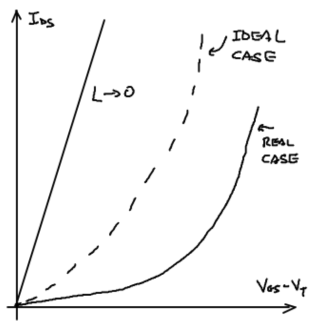
We get the same expression of the ideal case but with the ddenominator that scales the current. With L long the denominator $\simeq 1$.
Decreasing L we get a decrease of the current wrt the long channel case.
This relation has got a maximum at

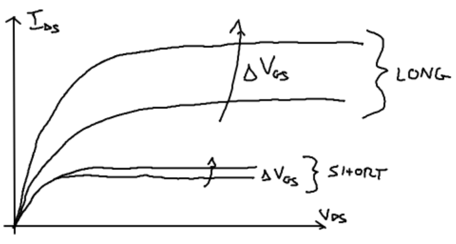
$$V_{ds}^{sat} = \frac{2(V_{gs} - V_t)/m}{1 + \sqrt{1 + \frac{2(V_{gs} - V_t)}{F_{sat} mL}}} \tag{5.68}$$

so we are reaching the saturation regime before than a long channel device and we get a higer slope of the rect after that point (this is olso caused by the DIBLE). With a short L we get lower saturation voltage and saturation current.



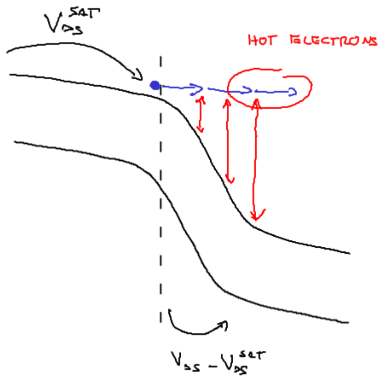
If we consider the inversion charge at the drain interface at V_{ds}^{sat} in a short channel device isn't zero we have still strong inversion. Before reaching pinch-off point at the drain we reach the saturation velocity; this 2 processes can lead our transistor in saturation region.
Ideally with shorter L we expect a steeper parabola in the transcaracteristic but we get considering the velocity saturation we get a high error. In an extreme case with $L \to 0$ $I_{ds}^{sat} \simeq WC_{ox}(V_{gs} - V_t)v_d^{sat}$ we get a linear dependance with V_{gs} and we get a lower sensitivity of the current to gate changes.





Hot electrons

The saturation velocity v_{sat} is the maximum average velocity in a drift-diffusion framework and so with scattering events. We can avoid scattering events by using steeper band banding.



Electrons that before scattering reaches a lot of kinetics energy (2-3eV) are called hot electrons. This electrons are so full of energy that can create impact ionization with atoms and so GR processes.
They can olso inteact with silicon atoms at the surface breaking bonds and so creating some interface states. They are able to overcome the barrier of the oxide and reach the gate or remain trapped in the oxide (they are used for floating gates charge processes).

5.13 Scaling rules

There are some sets of scaling rules that we can take as guideline to scale our devices.

5.13.1 Constant field (Dennard)

Main idea is to avoid the 2D electrostatic regime keeping the ratio between the vertical and the horizontal fields constant $F_x/F_y = const.$
In order to do this we need to scale all vertical and horizontal dimensions of the same factor as the voltages.
For the doping concentrations we can say form the 2D Poisson equation in sub-threshold regime that

$$\frac{\partial^2\phi}{\partial x} + \frac{\partial^2\phi}{\partial y} = \frac{qN_a}{\varepsilon_{si}} = -\frac{\partial F_x}{\partial x} - \frac{\partial F_y}{\partial y} \tag{5.69}$$

and so scaling all dimension by a factor k we get that to have the same electrostatic conditions

$$\frac{qN_a'}{\varepsilon_{si}} = -\frac{\partial F_x/k}{\partial x} - \frac{\partial F_y}{\partial y/k} = -k\frac{\partial F_x}{\partial x} - k\frac{\partial F_y}{\partial y} \tag{5.70}$$

from wch we get that $N_a' = kN_a$ so the doping concentrations must increase by a factor k.
So the rules for constant fields scaling are

- Dimensions (L,W,t_{ox},X_j) $\rightarrow \cdot 1/k$
- Voltages (V_{gs}, V_{ds}, V_{bs}) $\rightarrow \cdot 1/k$
- Doping concentrations (N_a, N_d) $\rightarrow \cdot k$

Device parameters changes

Electric field remains constant as the carriers velocity.
Capacitances scale as 1/k.

Taking into account the drain current in parabolic regime $I_{ds} = \mu_n C_{ox} \frac{W}{L}[(V_{gs} - V_t)V_{ds} - \frac{mV_{ds}^2}{2}]$ the current is scaled as 1/k but we have to consider the V_t dependance. The threshold voltage has to be reduced by a factor k in accord with the second rule but we know that

$$V_t = V_{fb} + 2|\phi_b| + \frac{\sqrt{2\varepsilon_{si} q N_a 2|\phi_b|}}{C_{ox}} \qquad C_{ox} \to \cdot k \qquad N_a \to \cdot k \tag{5.71}$$

but $2|\phi_b|$ does not scale by a factor k we have a logaritmic dependance on the doping concentration ($2|\phi_b| = kT/q \ln(N_a/n_i)$). We are not free to modify as we want V_t that is related with the energy gap and on $2|\phi_b|$. To overcome this we have to use non-costant doping concentrations.
The width of the depletion layer $W_d = \sqrt{\frac{2\varepsilon_{si}}{q} \frac{1}{N_a} 2|\phi_b|}$ do not scale with k but less for it's dependance with $2|\phi_b|$. So scaling the device with this set of rules we are slowly moving to the short channel regime according to the formula

$$L > 2(W_d^{max} + 3t_{ox}) \tag{5.72}$$

beacuse L scales as 1/k ,t_{ox} scales as 1/k but due to the dependance of W_d^{max} on $2|\phi_b|$ it scales less than 1/k.
Scaling ,in theory , also V_t of k since the scaling process does not affect the STS we are exponentially increasing the off-power dissipation.
The transit time $\tau = L^2/(\mu_{eff} V_{ds})$ is scaled of 1/k so we are getting a better frequency responce.

Circuit parameters changes

Delay of logic gates:
CV/I_{ds} capacitances, current and voltage scale as 1/k so we get that the delay of logic gates decreases by a factor 1/k.
Power dissipation:
VI_{ds} both scale as 1/k so the power dissipation scales as $1/k^2$ we have less dissipated power per logic function.
Density of power dissipation:
$\frac{VI_{ds}}{WL}$ the power dissipated per unit area remains constant (but we have more logic function in that area)
Integration density:
1/WL number of device per unit area increase by a factor k

The negative aspects of this set of scaling rules are that we are slowly moving to short channel regime, we have higher off-power dissipation and we lose compatibility to one thecnology to the other scaling the voltages.

5.13.2 Generalized scaling rules (Baccarjni)

We accept an increase of F in order to move away from the 2D electrostatic regime (moving near to the velocity saturation regime).

With the same path of before we define the 3 rules of this set

- Dimensions (L,W,t_{ox},X_j) $\rightarrow \cdot 1/k$
- Voltages (V_{gs}, V_{ds}, V_{bs}) $\rightarrow \cdot \alpha/k$
- Doping concentrations (N_a, N_d) $\rightarrow \cdot \alpha k$

with $1 < \alpha < k$

Parameters change

Carrier velocity:
In the case of linear or low field regime $\cdot\alpha$, in velocity saturation $\cdot 1$.
Current:
In low field $\cdot\alpha^2/k$, in velocity saturation $\cdot\alpha/k$.
Capacitances:
The electric field is not involved so we get 1/k for every regime.
Delay of logic gates:
In low field $\cdot 1/k$, in saturation $\cdot 1/k$.
Power dissipation:
In low field $\cdot\alpha^3/k$, in saturation $\cdot\alpha^2/k$.
Densisty of power dissipation:
In low field $\cdot\alpha^3/k^2$, in saturation $\cdot\alpha^2/k^2$ (we get problems form heat dissipation).

Increasing the doping concentration of more than k we get that W_d^{max} decreases more so we don't reach short channel regime.
With $\alpha = k$ we keep the same voltages and so compatibility with the older processes.

Increasing F we are increasing hot electrons that can tunnel throught the gate that is becoming thinner so we get an exponential increase of the gate current with this process. In order to avoid this we use high-k material instead of $Si0_2$ increasing the thikness of the oxide but preserving the same capacitance so

$$t_{ox}^{hk} = \frac{\varepsilon_{hk}}{\varepsilon_{ox}} t_{ox} \tag{5.73}$$

However high-k materials are not so good in terms of charge inside the oxide, spurios states and interface states wrt SiO_2.