

3.1 MOSFET intro

- Basic semiconductor physics
- Basic MOSFET working principle
- MOSFET types
- MOSFET applications
- Basic MOSFET fabrication process

3.2 MOS physics

- Work function
- Contact phenomena
 - Metal - metal
 - Metal - semiconductor
 - Metal - insulator - semiconductor (MIS)
- Working regimes of the MOS structure

3.3 Ideal long-channel MOSFET

- Charge models (Q-constant, Gradual-Channel)
- MOSFET working regimes

3.4 Realistic behavior

- Basics of charge carrier transport
- Modification in charge carrier transport
- Break-through phenomena
- Parasitics

In **mechanics** the kinetic energy W_{kin} of a particle with mass m and velocity v is defined:

$$W_{kin} = \frac{1}{2}mv^2$$

In **electrodynamics** for a charge q in an electrical potential V the potential energy can be defined:

$$W_{pot} = qV$$

If the charge is a single electron with the elementary charge e and the potential is 1 V than the energy equals:

$$W [J] = e \cdot V = 1.6 \cdot 10^{-19} \text{ A} \cdot \text{sec} \cdot 1 \text{ V} = 1.6 \cdot 10^{-19} \text{ J}$$

$$e = 1.60 \cdot 10^{-19} \text{ C}$$

To get easier numbers in electronic calculations the potential energy is commonly expressed in units of e, creating a new unit called "**electronvolt eV**" :

$$W [\text{eV}] = e \cdot V$$

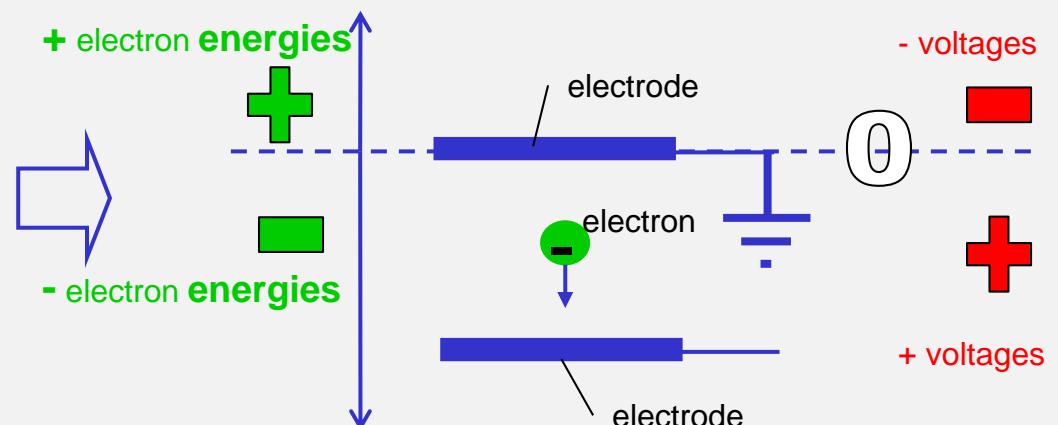
$$1 \text{ eV} = 1.6 \cdot 10^{-19} \text{ J}$$

Without scattering (ballistics) the kinetic energy of an electron, accelerated by a potential, equals:

$$W_{kin} = \frac{1}{2}mv^2 = e \cdot V$$

Due to the negative sign of electron charge:
electrons in positive potentials are bonded
and their energy has a negative sign

$$W_{pot} = q \cdot V = -|e| \cdot V$$



In **thermodynamics** a thermal energy can be defined by a temperature T [K] and equals:
(temperature is transferred from hot walls by swinging wall atoms to the particle in collisions)
 k_B is called the Boltzmann-constant and equals:

$$k_B = 1.38 \cdot 10^{-23} \text{ J/K}$$

$$W_{th} = k_B T$$

In many problems including strong mechanical scattering (as an equalization process) the kinetic energy is related to the thermal energy by:

$$W_{kin} = \frac{1}{2} mv^2 = \frac{3}{2} k_B T$$

Converting J in eV results in:

$$k_B = 1.38 \cdot 10^{-23} \text{ J/K} = 8.617 \cdot 10^{-5} \text{ eV/K}$$

So a temperature T corresponds to an energy W in eV by the transformation:

$$W[\text{eV}] = k_B \cdot T[\text{K}] = 8.617 \cdot 10^{-5} \text{ eV/K} \cdot T[\text{K}] = \frac{T[\text{K}]}{11600}$$



the thermal energy of an electron at room temperature (~300K) is about:

$$W_{th}[300\text{ K}] = \frac{300}{11600} = 0.026 \text{ eV}$$

this corresponds to a thermal velocity:

$$v_{th} = \sqrt{\frac{3k_B T}{m}} = \sqrt{\frac{3 \cdot 1.38 \cdot 10^{23} \text{ J/K} \cdot T[\text{K}]}{9.11 \cdot 10^{-31} \text{ kg}}} \approx 6741 [\text{m/sec}] \cdot \sqrt{T[\text{K}]}$$

$$v_{th} = 6741 [\text{m/sec}] \cdot \sqrt{300 [\text{K}]} = 1.17 \cdot 10^5 \text{ m/sec} \approx 10^7 \text{ cm/sec}$$



At room temperature RT a free electron reaches a thermal velocity of about ~10⁷ cm/sec by thermal energy transfer

1887 Hertz/Hallwachs: Photo effect

1892 Wien: e/m

1895 Thomson: Identification of the electron

Remember Chap-1

First Quantum Effects

1900 Planck: "quantum constant"

1905 Einstein: explanation of photo effect



Quantum Physics

1925 Born, Schrödinger, Heisenberg: Quantum theory

1926 Fermi, Dirac: Fermi-statistics

1927 Pauli: Electron theory of metals

1928 Sommerfeld: Quantum theory of solid-states

1929 Bloch: solution of SG-equation in periodical potential



1925 Presser: Selen rectifier

1926 Grondahl: Cu₂O rectifier

1928 Lilienfeld: Patent Line-IGFET

1929 Schottky: "Barrier layer" at metal/semicond boundary
Assumption of **electrons and "holes"**1931 Wilson: **band model** -> explanation of metals
semiconductors, isolators

1933 Tamm: electron states at surfaces

1934 Heil: proposal of surface-IGFET

1938 Mott: depletion theory of semiconductors

1938 Schottky: "surface layer theory", diffusion current

1938 R.Pohl: „solid-state amplifier" in ion-crystal

1942 Bethe: thermo-emission theory over barrier

1940 Ohl: p- and n-Doping, diodes

1945 Shockley: proposal for FET and failure for realization

1946 Bardeen: explanation of malfunction due to surface states

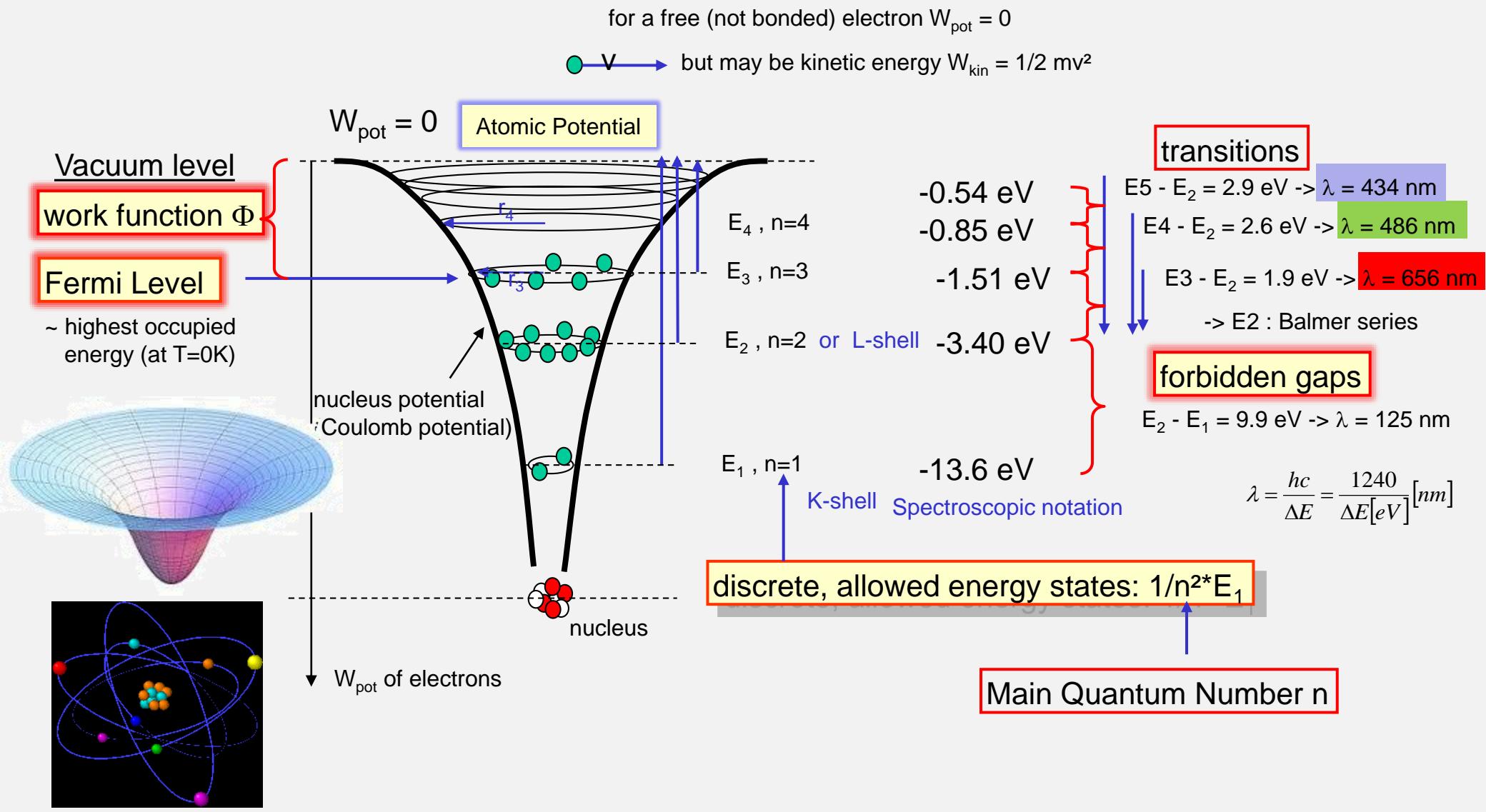
1947 Brattain, Bardeen, Shockley: discovery of Bipolartransistor (Ge-Point contact transistor)

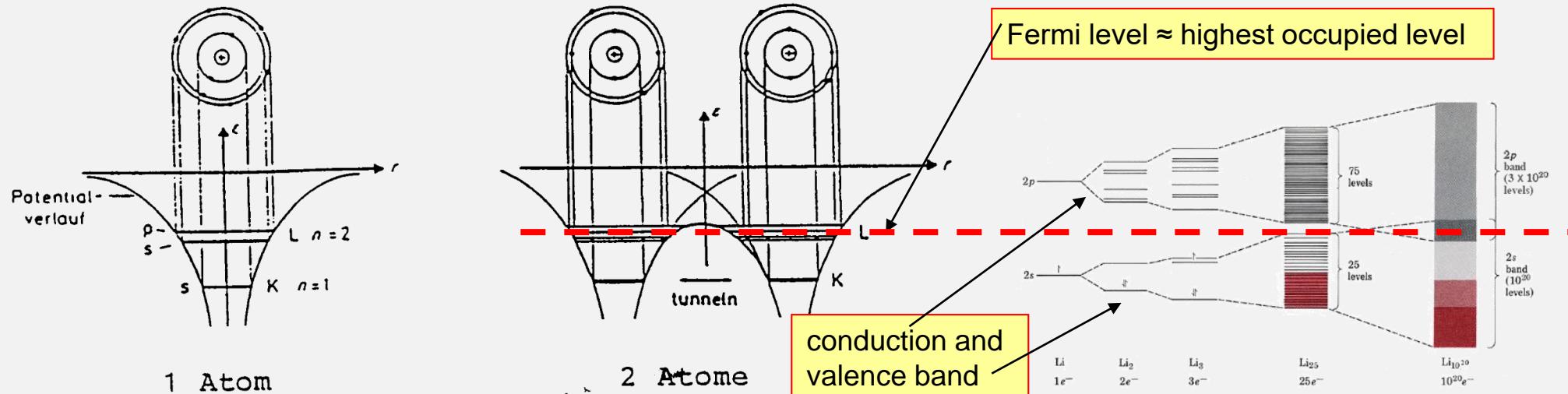
1948: first public presentation of the transistor

1949 Shockley: Theory of pn-junctions

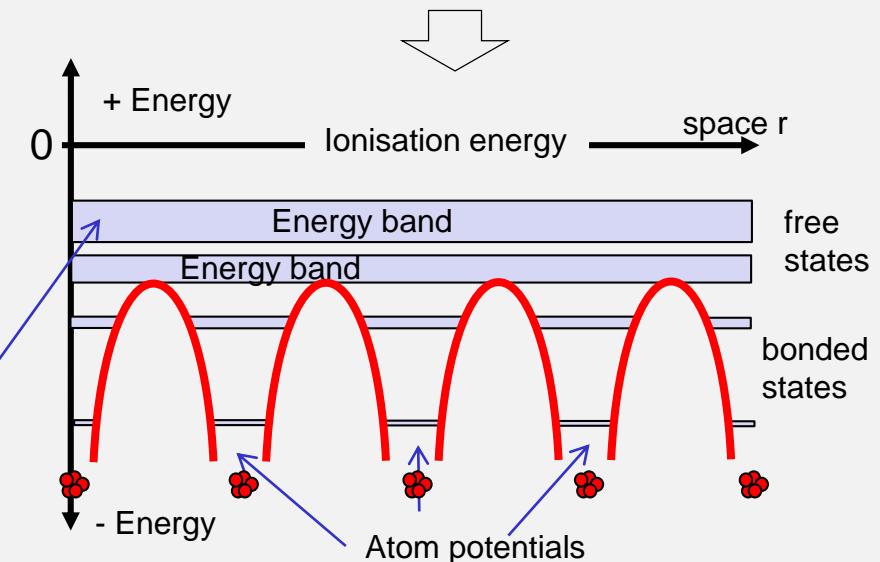


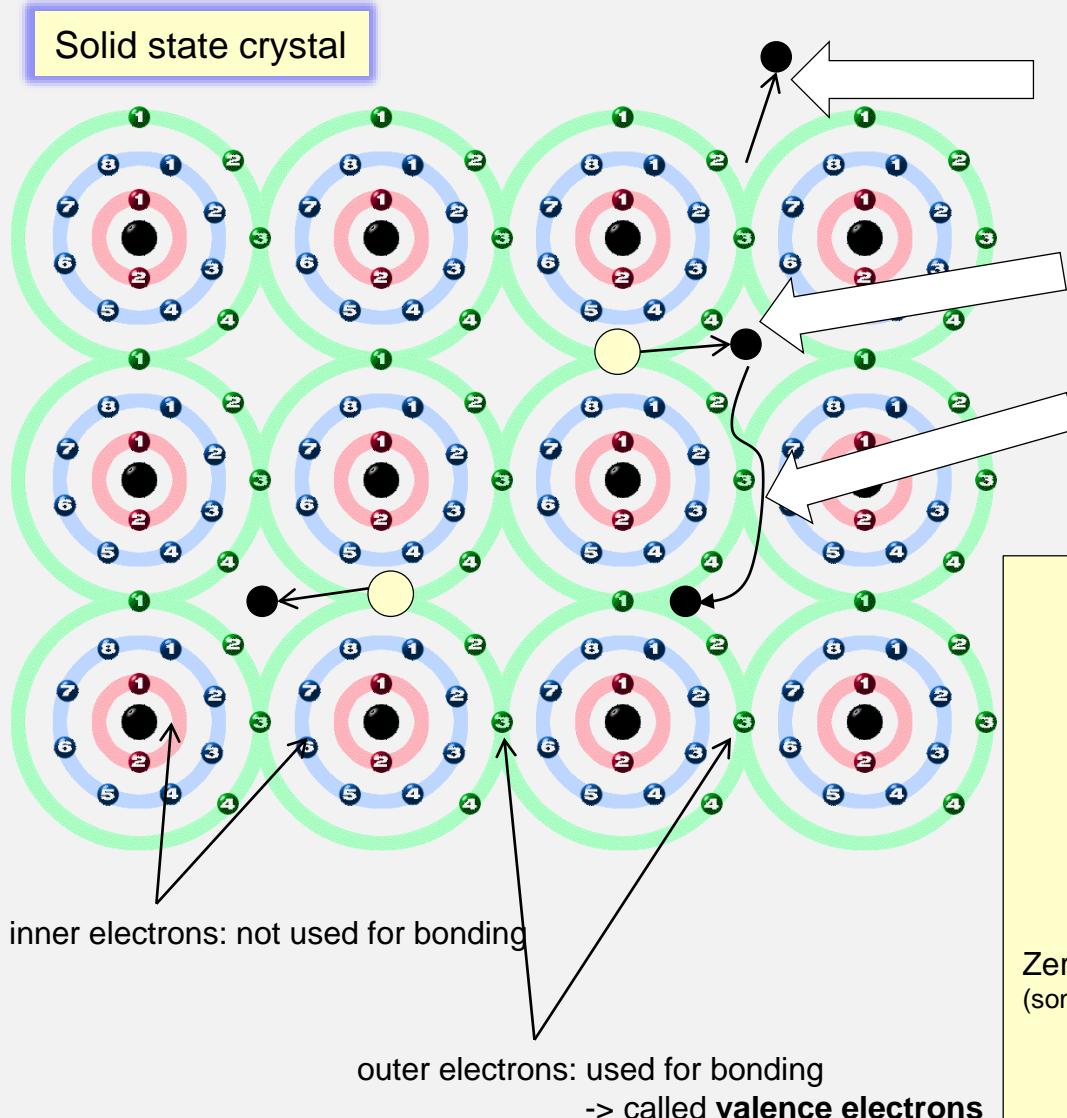
Quantum theory was well developed and used for semiconductor physics





- ▶ In overlapping atoms the energy levels must be different due to the Pauli principle.
- ▶ The separation is dependent on the number of interacting atoms.
- ▶ In a solid with 10^{22} atoms/cm³ the 10^{22} energy levels create continuous bands.



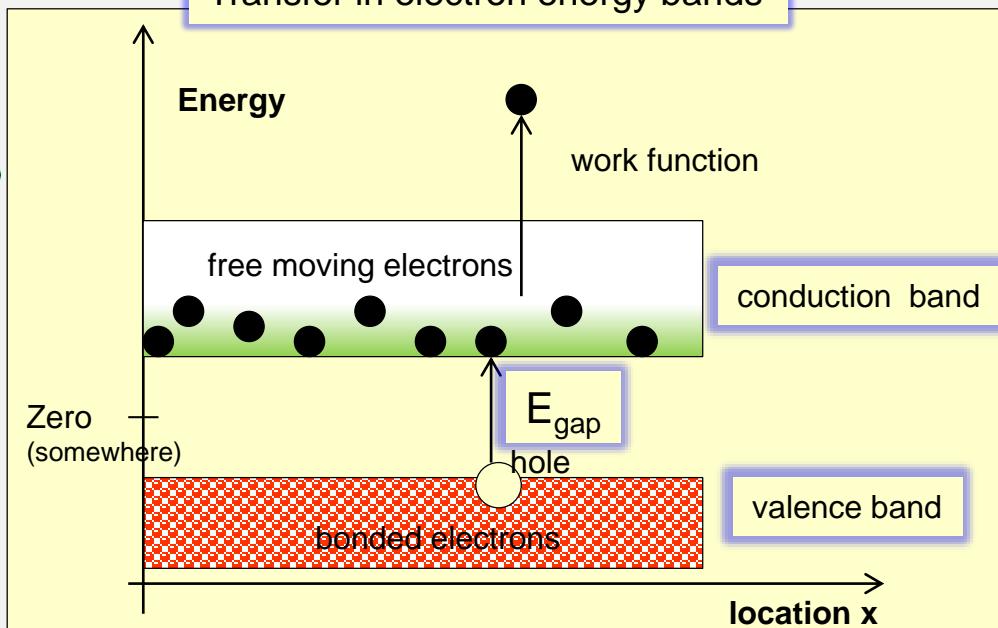


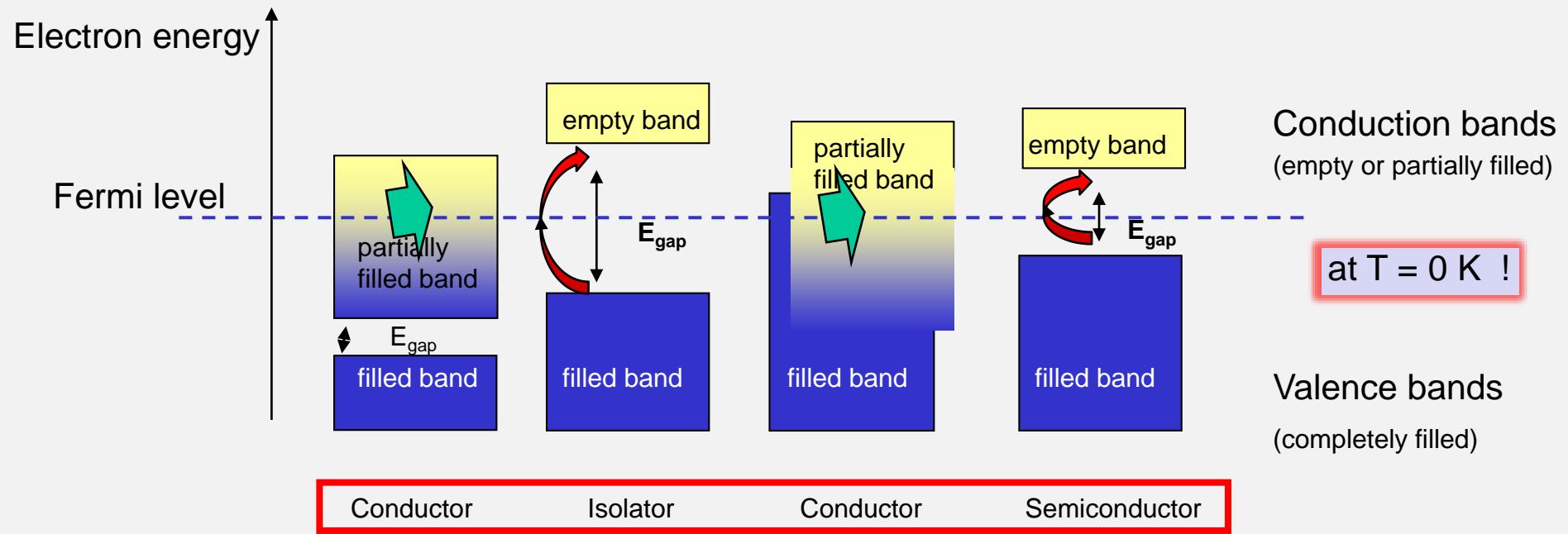
Extract (free) one electron out of the solid needs energy
-> this energy is called "**work function**"

Kick out one electron out of the bond (valence) needs energy
-> this energy is called " E_{gap} "

Non-bonded electrons move "freely" through the crystal
-> they are guided by electric fields -> "**Conduction electrons**"

Transfer in electron energy bands



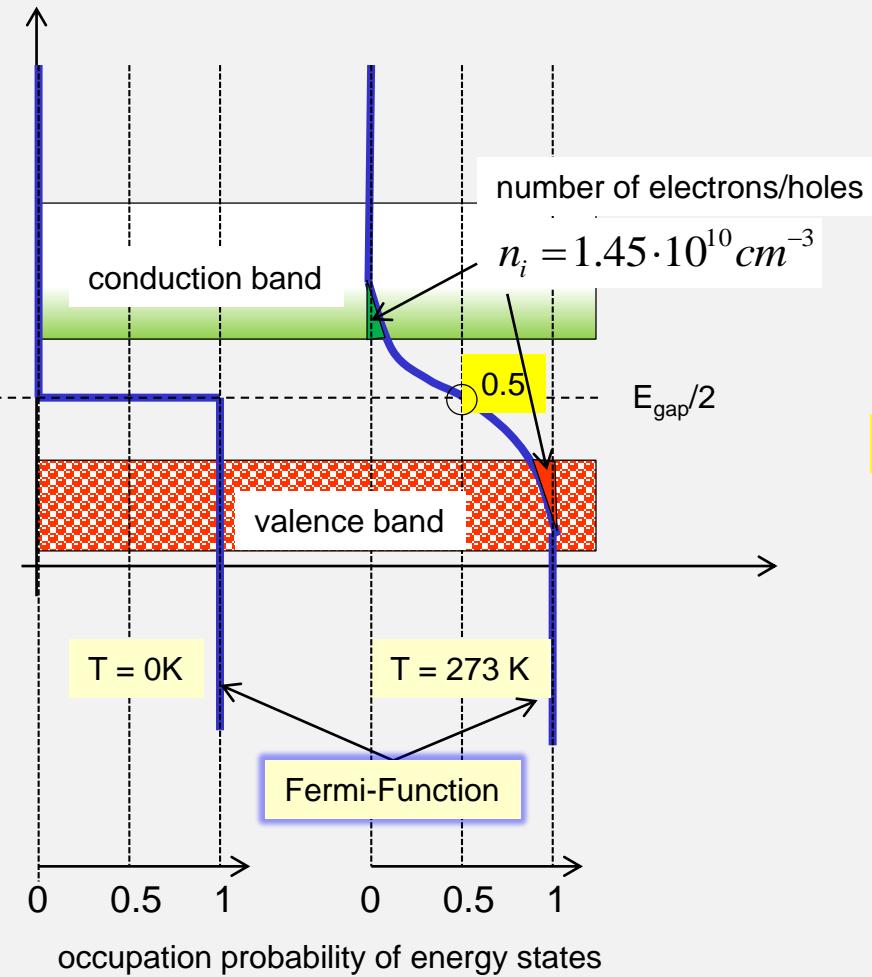


At temperatures of $T \sim 0\text{K}$ only metals or isolators exist

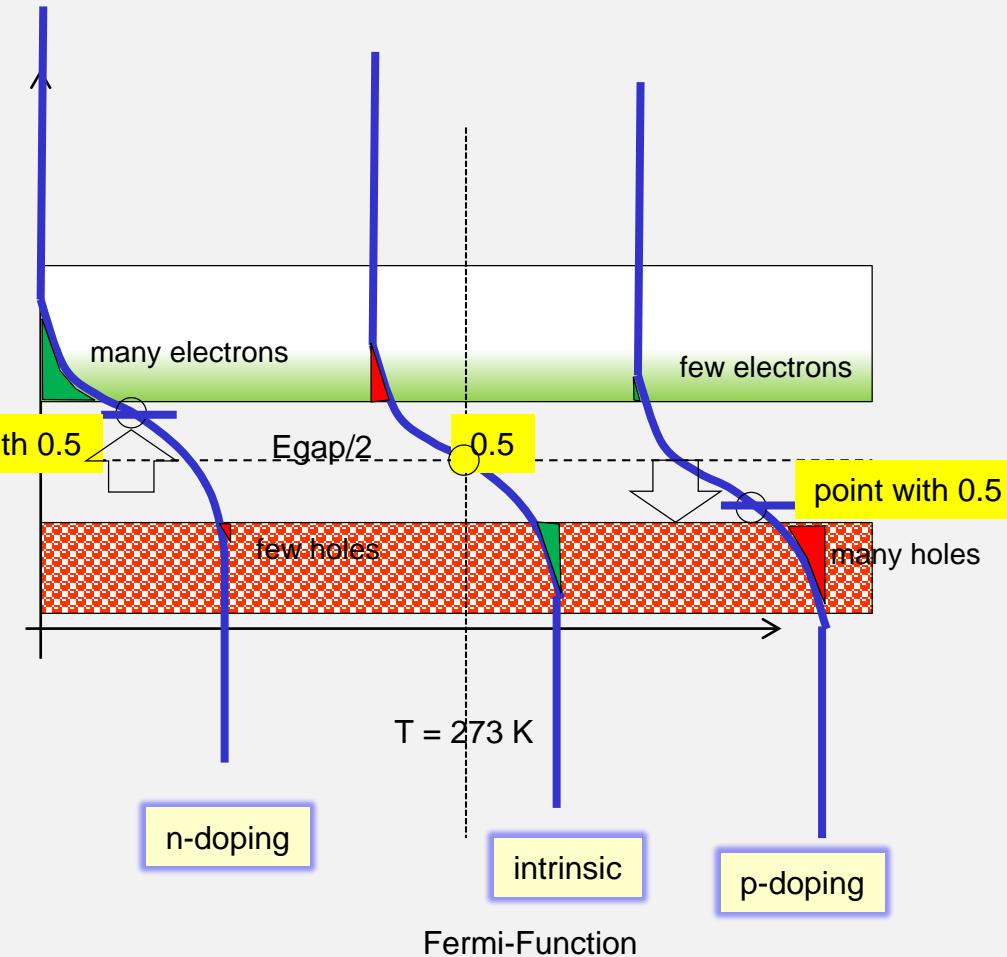


The size of the band gap determines at $T > 0\text{K}$ the number of free electrons (thermal excitement)
In the conduction band of isolators and therefore the conductivity

Better definition of Fermi-Energy



Shift of Fermi-Function with doping



Fermi energy is the level, where the electron occupation probability is 0.5

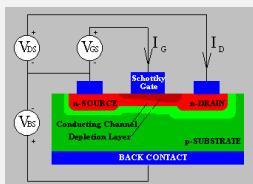
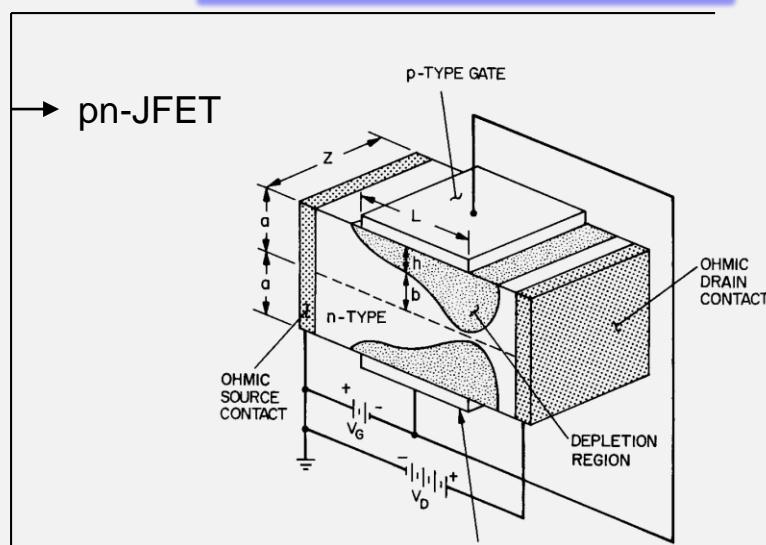
Formation of barriers
-> see chap.1

Field-effect devices

=

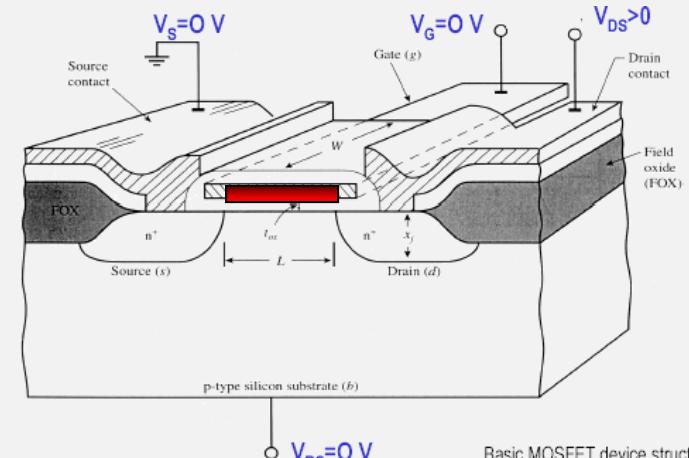
control of current by external fields

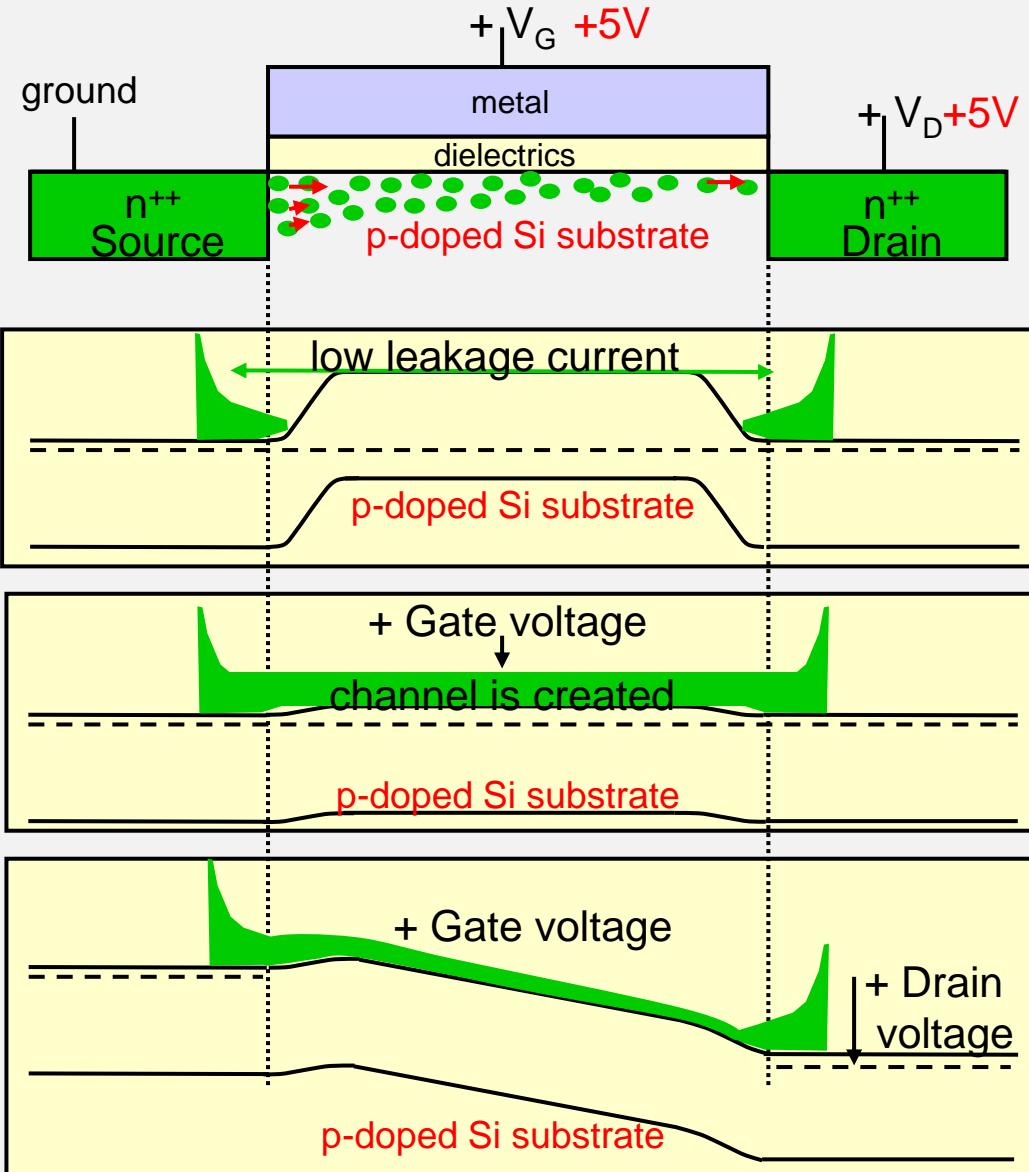
Channel cross section control



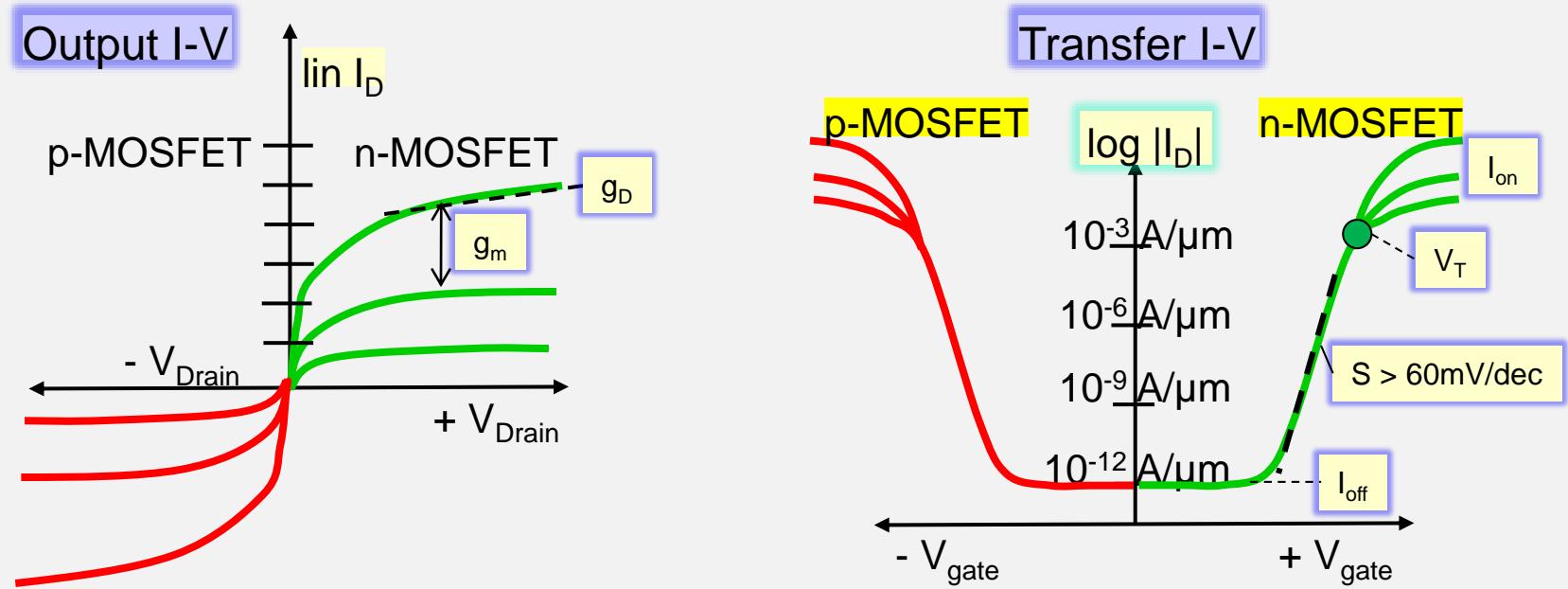
Channel conductance control

IGFET → MISFET → MOSFET
(Isolated Gate)
(Metal-Oxide-Semiconductor)
(Metal-Insulator-Semiconductor)





1. Between two contacts a barrier is realized
n+ Source, n+ Drain, p-channel
2. A capacitor on top of the channel can change the barrier height by inducing an electric field
3. If the barrier is high enough, there is no connecting path between the contacts
4. If the barrier is lowered by the Gate electrode, for the charge carriers a conducting path (channel) between Source and Drain is created.
The barrier is lowered, because a positive gate voltage on the gate capacitor forces electrons on the other plate (the semiconductor). This makes the semiconductor look like n-doped, so the p-barrier is gone.
5. If a voltage difference is applied between Source and Drain a current flow is the result



Basic I-V characteristics:

$$I_D = \mu C_{ox} \cdot \frac{w}{L} \cdot \left[(V_{GS} - V_T) \cdot V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Threshold voltage:

$$V_T = \left(\Phi_{MS} \pm \frac{Q''_{ox}}{C''_{ox}} \pm \frac{Q''_{impl}}{C''_{ox}} \right) \pm \frac{Q''_{dep\max}}{C''_{ox}} \pm |2\Psi_{bulk}| = V_{FB} \pm \frac{\sqrt{2\varepsilon_0\varepsilon_{Si} \cdot qN_{dop} \cdot (2|\Psi_{bulk}| \pm V_{BS})}}{C''_{ox}} \pm |2\Psi_{bulk}|$$

Important performance parameters:

Leakage current I_{off}

Subthreshold slope S

Threshold voltage V_T

Drive current I_{on}

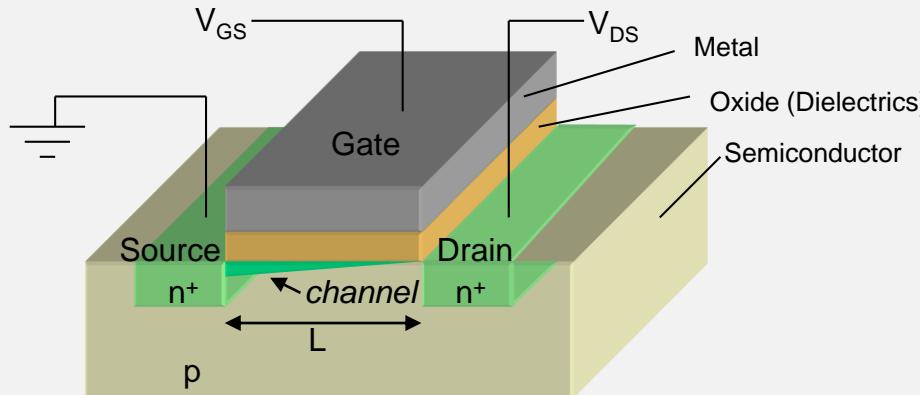
4 different types of MOSFETs:

→ **npn**

an electron channel
is created

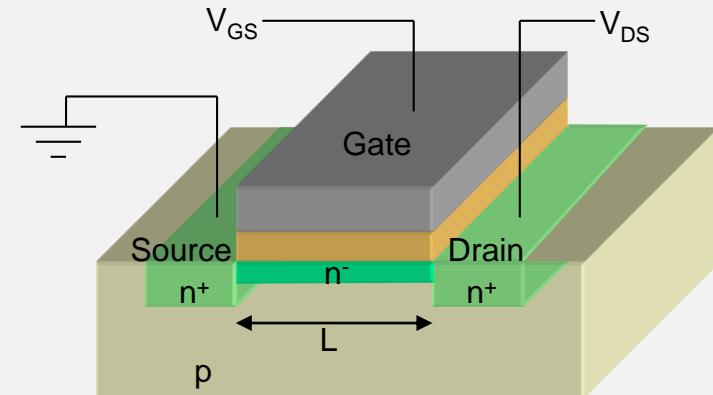
normally-off (enhancement to switch on):

barrier so high, that without external fields
the device is off and must switched on by an external field



normally-on (depletion to switch off):

barrier so low, that without external fields the device is on
and must switched off by an external field



Realization: In p-silicon two regions of n⁺ doping are realized

Realization: In p-silicon two regions of n⁺ doping are realized
and additional channel n-doping

→ **pnp**

a hole channel
is created

normally-off (enhancement):

barrier so high, that without external fields
the device is off and must switched on by an external field

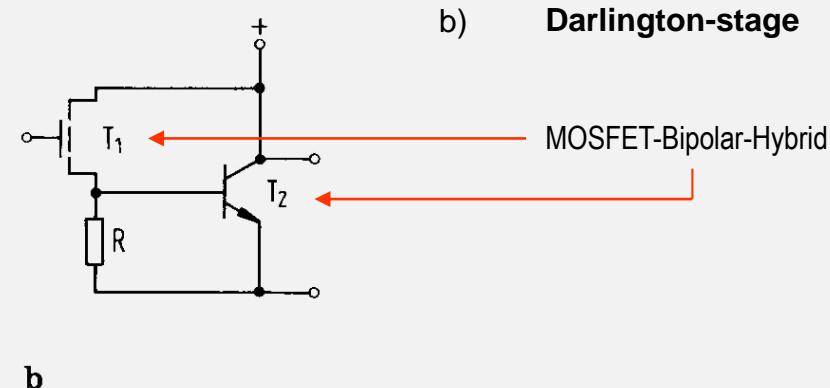
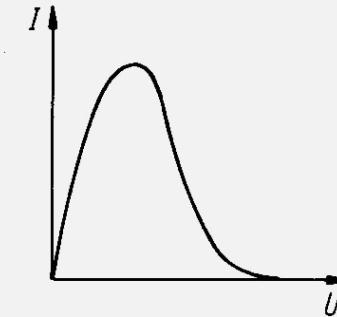
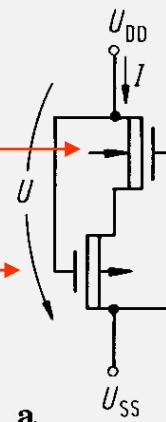
normally-on (depletion):

barrier so low, that without external fields the device is on
and must switched off by an external field

a) **λ -circuit**

Realization of special IV-characteristics

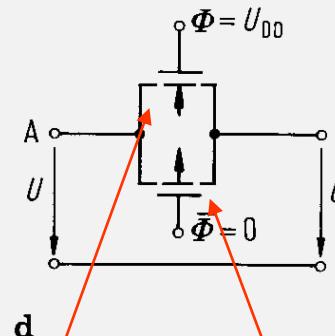
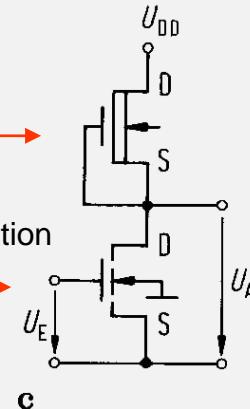
n-depl-T and p-depl-T



c) **ED-Inverter**

n-depl-T for load

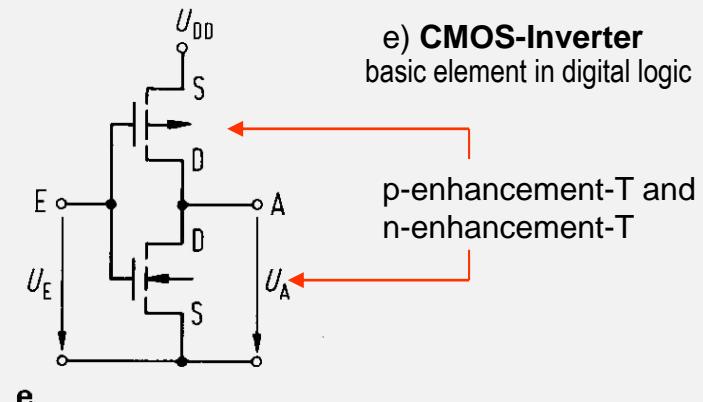
n-enhancement-T for amplification



d) **Transmission gates**
switching in analog applications

n-enhance-T and p-enhance-T parallel

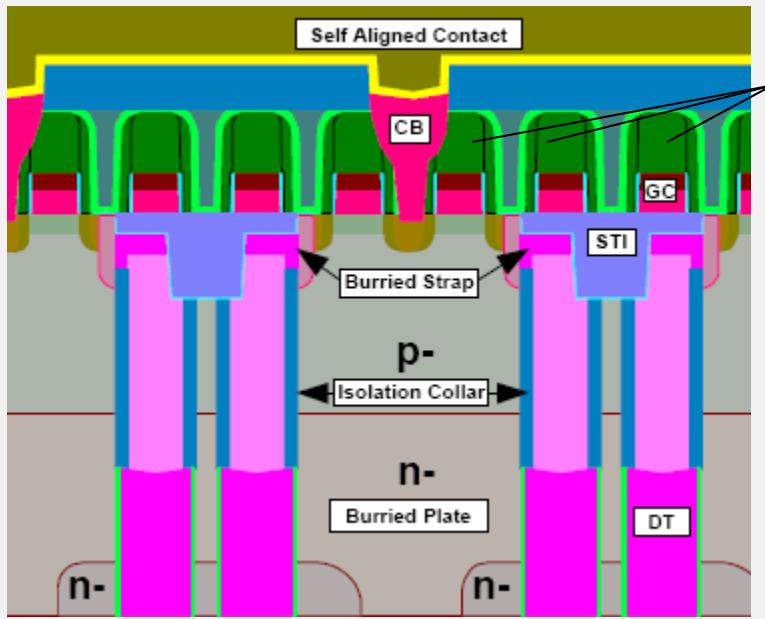
e) **CMOS-Inverter**
basic element in digital logic



Various combinations of p/n- and enhancement/depletion MOSFETs are used in logic circuits

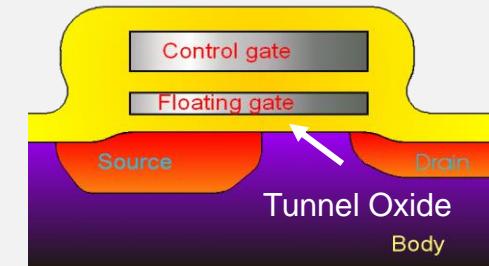
MOSFETs in actual circuits:

Dynamic Random Access Memory (DRAM)

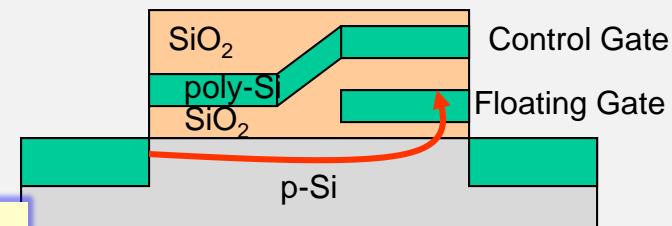


Courtesy of Qimonda, 170nm DRAM (~2000)

Electrically Erasable Programmable Read Only Memory (EEPROM)



Flash Electrically Erasable Read Only Memory (Flash-EEROM) with split-gate



* 2 MOSFETs (access + memory)
in one cell

Because of low power and small area consumption MOSFETs are the only used transistor type in High-Integration



Properties

ICs with MOSFETs

supply voltage V_{DS} : 0.8 to 20 V
Current in devices I_{DS} : 1 μ A to 10mA
Frequency: up to 10GHz

Discrete MOSFETs

V_{DS} : 5 to 100 V
 I_{DS} : up to ~2 A
Frequency: up to MHz

Power MOSFETs

V_{DS} : 5 to 1000V
 I_{DS} : up to ~100 A
Frequency: up to MHz

Applications

Server/PCs/Laptop
Consumer Electronics
Cell Phones
Communication

Amplifier
Switches

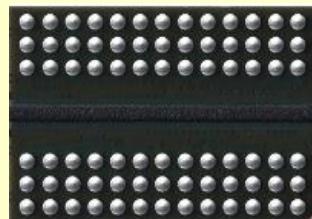
Power supplies
Switches
Motor control

Year 2019/2020

Memory (DRAM)

- * **Leading-Edge Product**
- * 20-10 nm Technology class
- * DRAM: 2,4,8,16 Gb -> 10^9 MOSFETs/chip
- * NAND: 512 Gb, 1Tb

MOSFET characteristics determine leakage current ->
clock frequency (access time)
power consumption
chip size



Courtesy of SAMSUNG
8Gb, 10nm class, year 2018



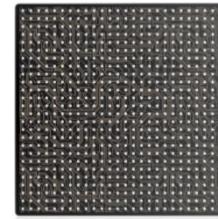
Microprocessor

- * **Leading-Edge Product**
- * 7 nm Technology
- * FINFET
- * ~ 9.90 Bill. MOSFETs/chip

MOSFETs characteristics determine leakage current ->
clock rate
power consumption
chip size

ASICs

- * **Follower product**
- * usually 2 generations behind
- * **SoC Leading-Edge**
- * 7 - 110 nm technology
- * ~ 8.5 Bill MOSFETs/chip



Analog

- * **Special design**
(Long channel)
- * ~ 100 MOSFETs
(A/D converter)
- * 0.5 μ m – 130 nm
- * ~ 30 GHz



Power

- * **Special design**
(Long channel)
- * ~ 1 MOSFET
(IGBT)



p-Si

Starting with p-doped Silicon wafer

oxide

Oxidation for device isolation

Poly-Si

Gate oxide and poly-Si deposition

gate

Poly-Si etch for gate

nitride

Oxide etch back + isolation deposition

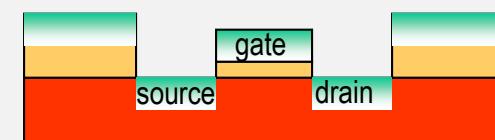
nitride

Opening of contact windows

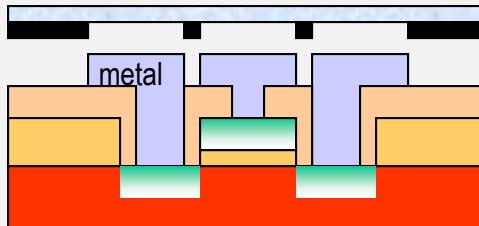


Isolation opening for device area

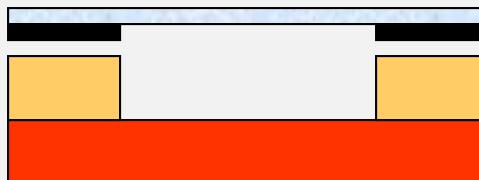
Mask 2



Overall n-doping -> self-adjusting S/D



Overall deposition of metal + patterning

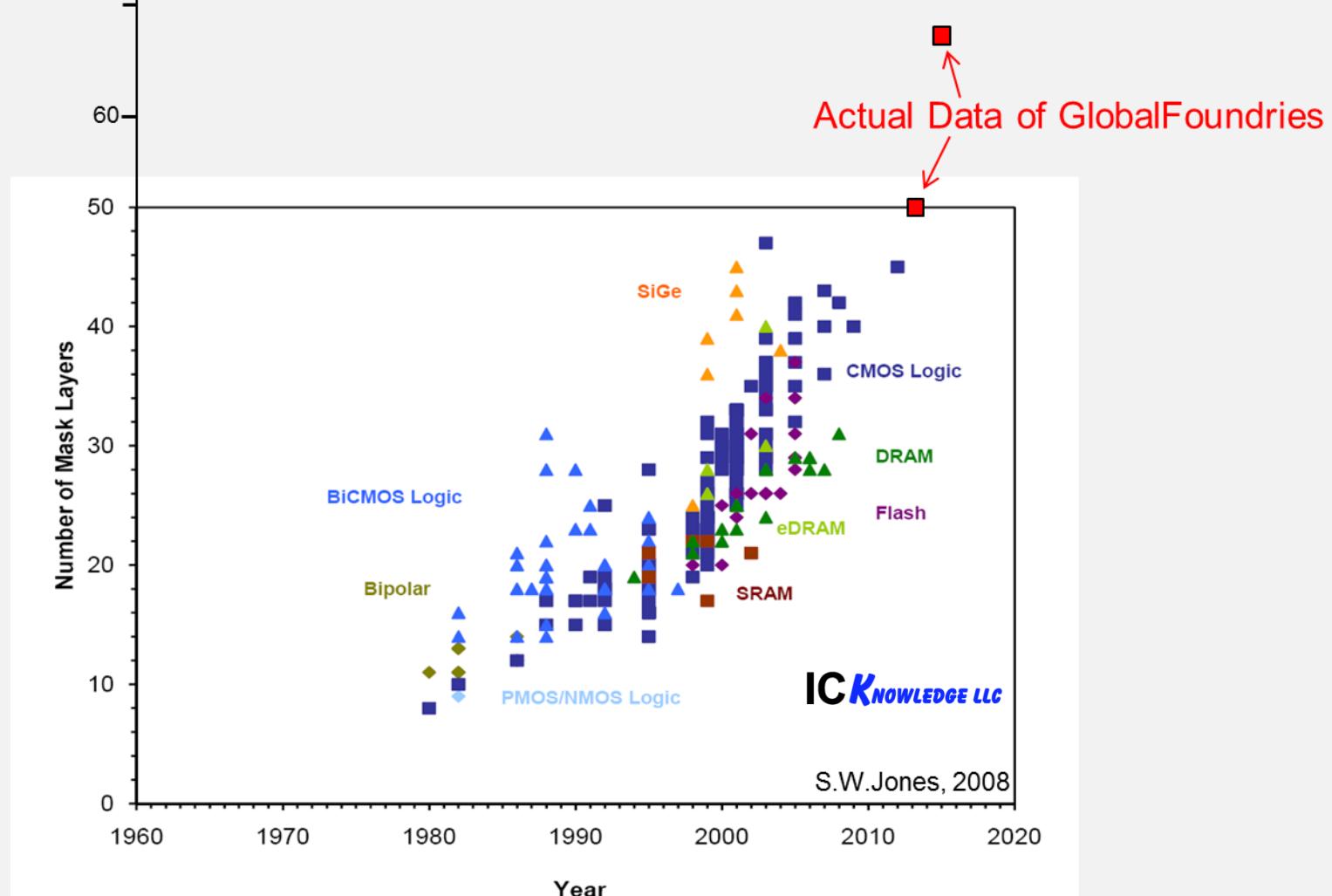


 Basic MOSFET fabrication requires only 4 masks, but reality today (CMOS 2015) is ~ 30-50 masks

Number of mask steps for fabricating a state-of art chip

but with each mask step several process steps are involved:

- * cleaning
- * resist spin-on
- * bake-out
- * exposure
- * post-bake
- * development
- * metrology
- * patterning process
- * resist strip
- * cleaning
- * metrology



3.1 MOSFET intro

- Basic semiconductor physics
- Basic MOSFET working principle
- MOSFET types
- MOSFET applications
- Basic MOSFET fabrication process

3.2 MOS physics

- Work function
- Contact phenomena
 - Metal - metal
 - Metal - semiconductor
 - Metal - insulator - semiconductor (MIS)
- Working regimes of the MOS structure

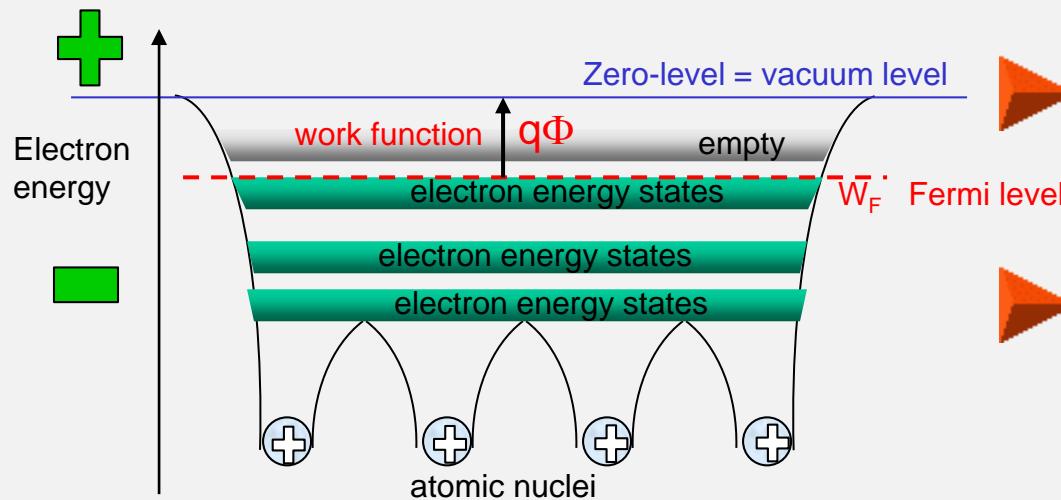
3.3 Ideal long-channel MOSFET

- Charge models (Q-constant, Gradual-Channel)
- MOSFET working regimes

3.4 Realistic behavior

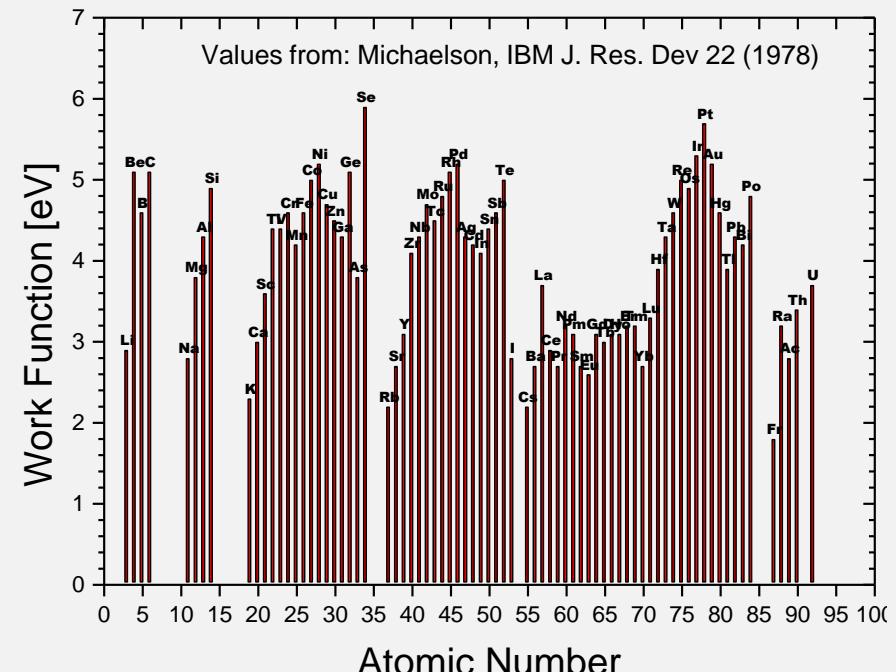
- Basics of charge carrier transport
- Modification in charge carrier transport
- Break-through phenomena
- Parasitics

Solids exhibit an equal number of positive charges in the nuclei and negatively charged electrons in the shells.



The electrons occupy allowed energy states within the solid. The highest energy level dividing occupied states by electrons from unoccupied states is called **Fermi level W_F**

The **workfunction $q\Phi$** is the energy required to extract an electron from the Fermi energy to the **vacuum level** (= far away from the solid).



The work function $q\Phi$ is between 3-5 electronvolts for most elements

For all our local calculations (=within the semiconductor) we define the midgap (intrinsic) energy W_i as reference level with zero energy

1

In Si (and other semiconductors) the level of Fermi energy ($W_F = q\Psi_{bulk}$) can be adjusted by doping

Fermi energy :
$$W_F = q\Psi_{bulk} = \pm k_B T \cdot \ln\left(\frac{N_{doping}}{n_i}\right)$$

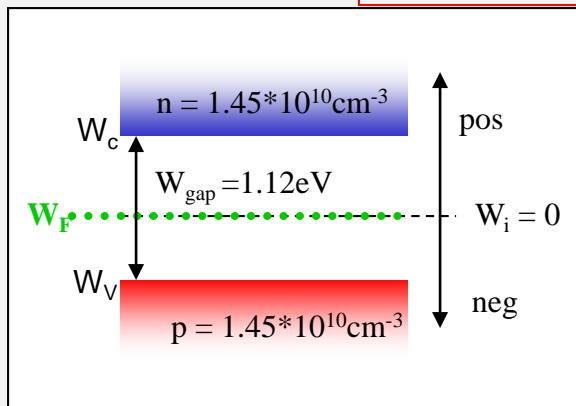
- + : n-type doping, $N_{doping} > n_i$
- : p-type doping, $N_{doping} < n_i$
- W_i : zero-point

$$n_i = 1.45 \cdot 10^{10} \text{ cm}^{-3}$$

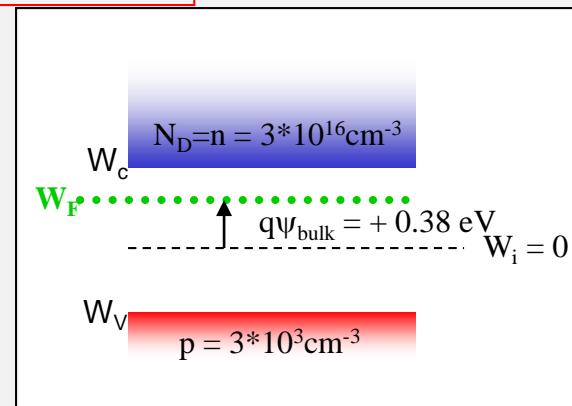
$$k_B T = 0.026 \text{ eV}$$

Notation:

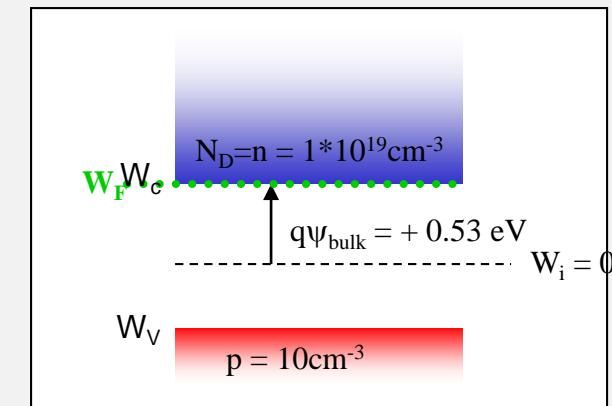
Usually greek letters
 $\Psi, \Phi, \chi, \dots [V]$ are used
for potentials



No doping $\rightarrow W_{Fermi} = W_i$



Medium n-doping $\rightarrow W_{Fermi}$ moves towards W_c



High n-doping $\rightarrow W_{Fermi}$ approaches W_c

2

In bandgap materials (semiconductor, insulator) there are no electrons at the fermi level. Because all electrons are (at the bottom) of the conduction band an additional energy **affinity** $q\chi$ is defined to characterize the extraction energy of the electrons:

$$q\chi_{Si} = 4.1 \text{ eV}$$

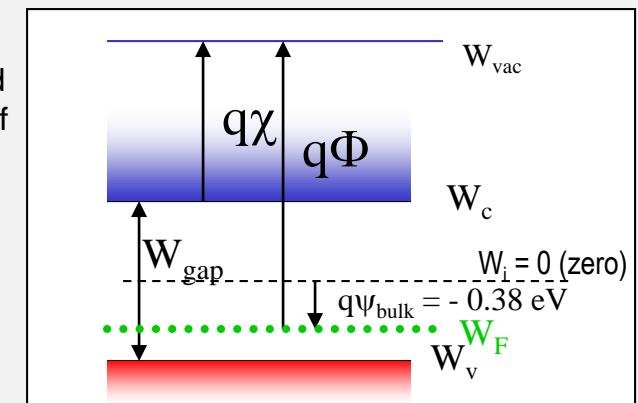
$$W_{gap} = 1.12 \text{ eV}$$

3

Calculation of the work function in silicon:

$$q\Phi_{Si} = \left(q\chi_{Si} + \frac{W_{gap,Si}}{2} \right) \pm \left| q\Psi_{bulk} \right|$$

p-dop
n-dop



In **metals** each atom (density $\sim 10^{22} \text{ cm}^{-3}$) delivers a free electron for charge transport (roughly)

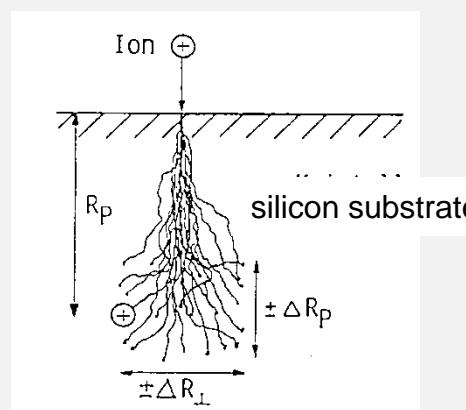
all phenomena of charge carrier transport are dominated by the huge number of free electrons -> easy to calculate

In **insulators** the number of free charge carriers is negligible ($\sim 1 \text{ cm}^{-3}$)

→ always present undesired charges (defects, interface states $< 10^{12} \text{ cm}^{-2}$) dominate real behavior -> difficult to calculate

In **semiconductors** the number of free charge carriers depends on doping ($10^{14} - 10^{20} \text{ cm}^{-3}$)

→ always present undesired charges (defects, interface states $< 10^{12} \text{ cm}^{-2}$) may have the same influence as doping
Example: doping by ion implantation -> difficult to calculate



-> 1 doping atom may generate 1000 defects and more !
-> defect density may exceed doping density



For physical calculations in charge carrier transport metals are frequently used as starting systems



so we start with metals on the next pages

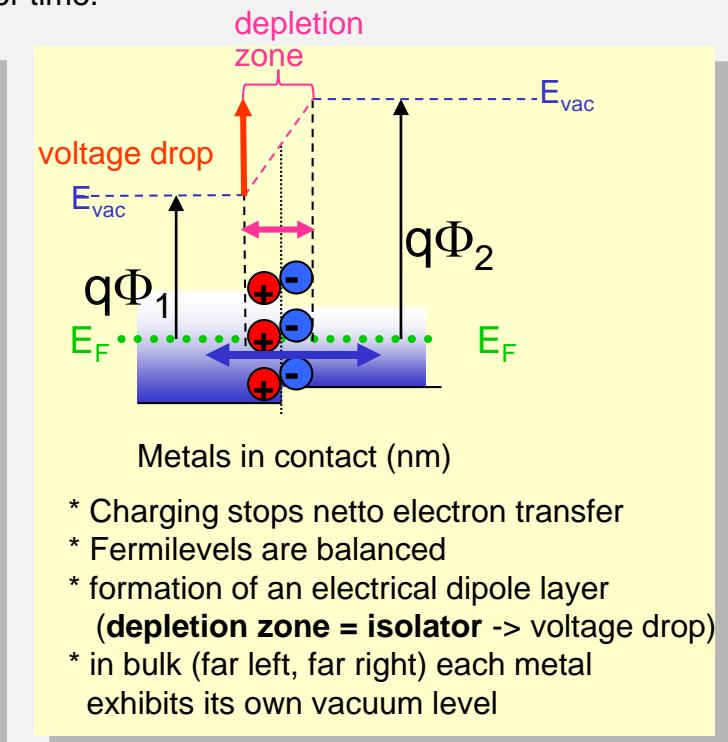
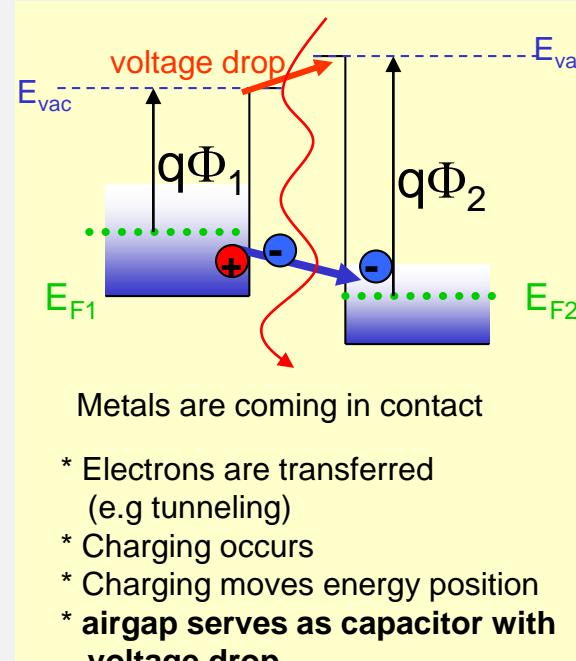
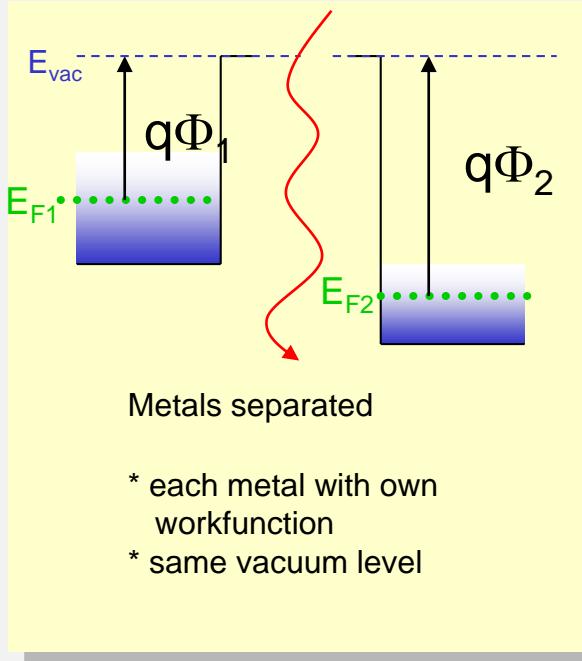
Free electrons in solids always try to occupy the lowest energy states in a system (thermodynamic law to minimize the total energy)

When the both metals come in contact, electrons will flow from the metal with higher Fermilevel (= lower workfunction) into the metal with lower Fermilevel (= higher workfunction).

This electron transfer charges the left metal positive (due to missing electrons) and the right metal negative (enrichment of electrons).

This charging shifts the energy position of both metals and creates an increasing electric field.

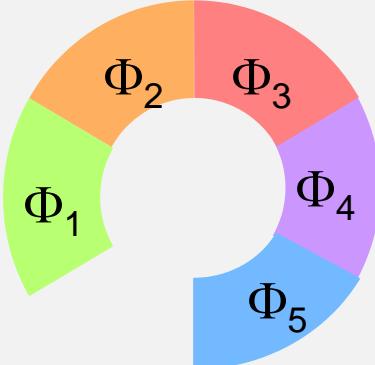
The rising electric field will finally stop netto electron transfer from left to right. In this moment the Fermilevels in both metals are balanced, thermodynamic equilibrium is reached, the same amount of electrons are exchanged per time.



The created **contact voltage** V_{con} is the difference of both work potentials:

$$V_{con} = \Phi_2 - \Phi_1$$

If a series of various metals with different work functions is in contact:



The contact voltage between the two outermost metals can be calculated

$$V_{1,5} = V_{1,2} + V_{2,3} + V_{3,4} + V_{4,5}$$

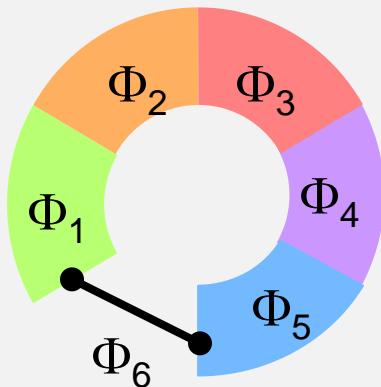
$$= (\Phi_2 - \Phi_1) + (\Phi_3 - \Phi_2) + (\Phi_4 - \Phi_3) + (\Phi_5 - \Phi_4) = (\Phi_5 - \Phi_1)$$

Ideal case:

The contact voltage between the two outermost metals is independent of the number of metals between and their work functions.

Realistic case:

Because the contact is not perfect (open bonds, atomic rearrangement, dirt) always so-called interface charges are created. These unintentional charges create in the capacitance of the contact zone an additional voltage shift, which can not be calculated and results in a dependence of inner materials.



If the two outermost metals are connected by an extra metal the sum of all contact voltages is Zero !

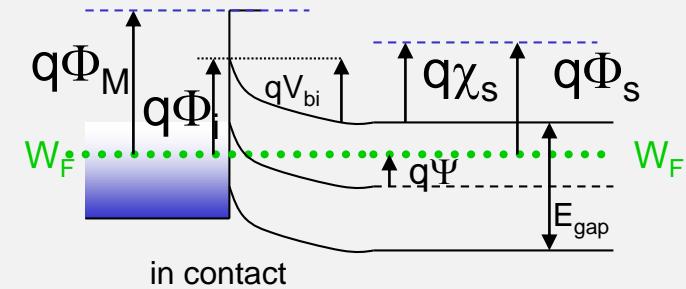
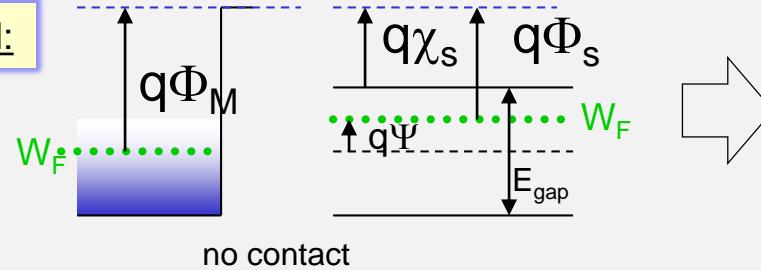
What a pity !

For current transport the barrier height is the important parameter

Two models exist to describe the barrier height:

1

Schottky Model:



$$V_{bi} = \Phi_M - \Phi_s = \Phi_M - \left(\chi_s + \frac{W_{gap}}{2q} - \Psi_b \right)$$

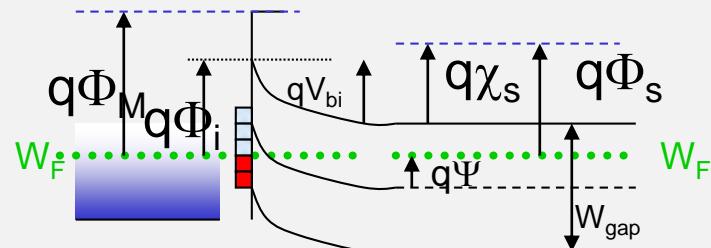
The barrier height Φ_i is determined by the difference in the work functions

This is the very ideal case !

2

Bardeen Model:

Although the workfunctions of materials in contact remain the same the barrier height between two materials is influenced by additional "interface charges" and not following the Schottky model prediction

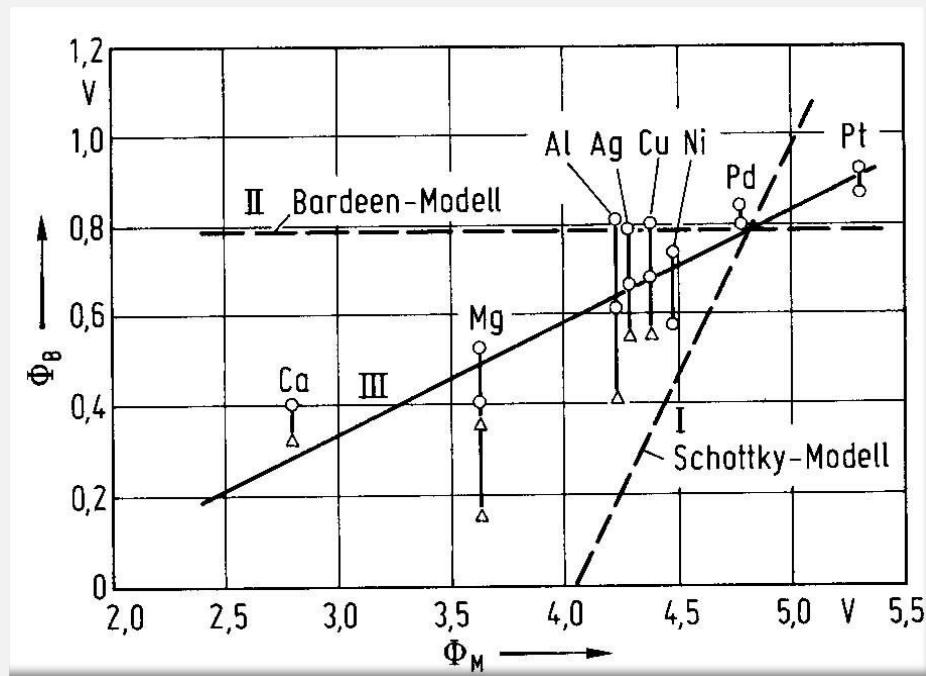


The barrier height Φ_i is fixed by the interface charges and not ideally dependent on the work functions

This is usually called Fermi level pinning !

for self-study you may have a look for example at: <http://academic.brooklyn.cuny.edu/physics/tung/Schottky/surface.htm>

Experimental values of barrier heights:



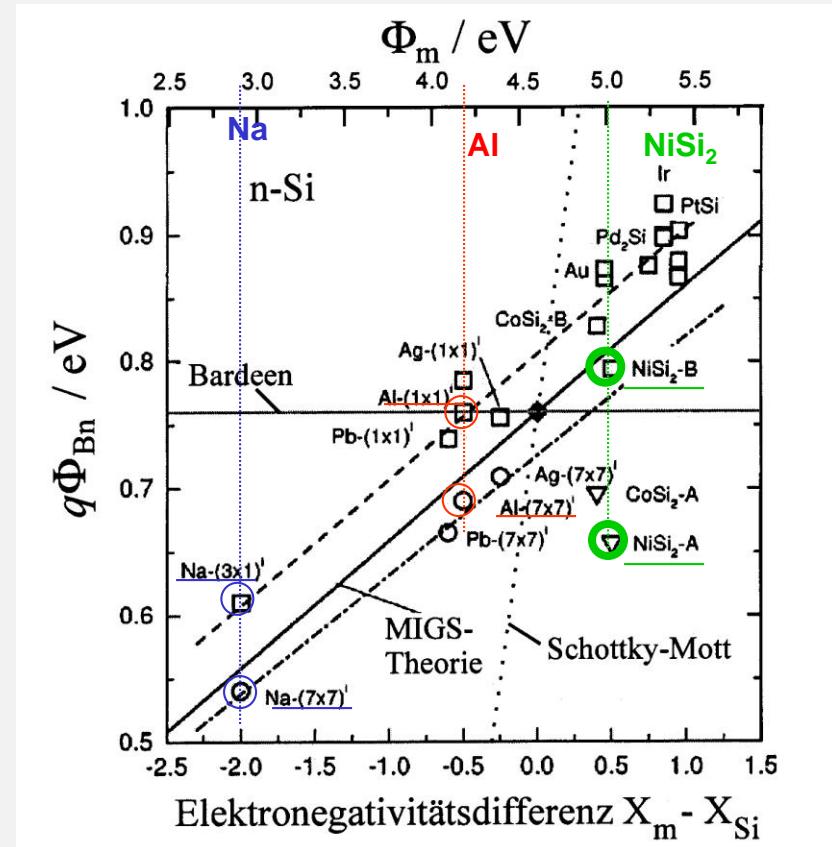
Barrier height of various metals on n-Si compared with Schottky theory and Bardeen theory (interface states)



Barrier height is a mix between both theories



The barrier height is also a material and property dependent "constant"



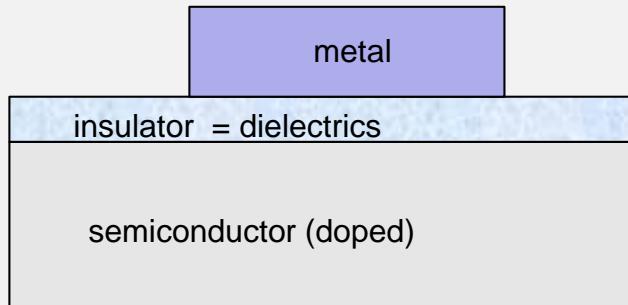
Barrier height of various metals on n-Si with different Silicon surface arrangements or different structure



Same work function, but structural dependent barrier height

Now we consider the problem of 3 stacked materials in contact:

We use the common metal - insulator - semiconductor (MIS) structure



- for many decades (from 1960 - 2005) the common metal was highly-doped poly-Si, now replaced by nearly metals like NiSi
- for many decades (from 1960 - 2005) the common insulator was SiO_2 , now replaced by high-k dielectrics like HfSiON

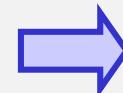


This arrangement of materials and supply voltages is called a MIS- or MOS dipole (diode, varactor)

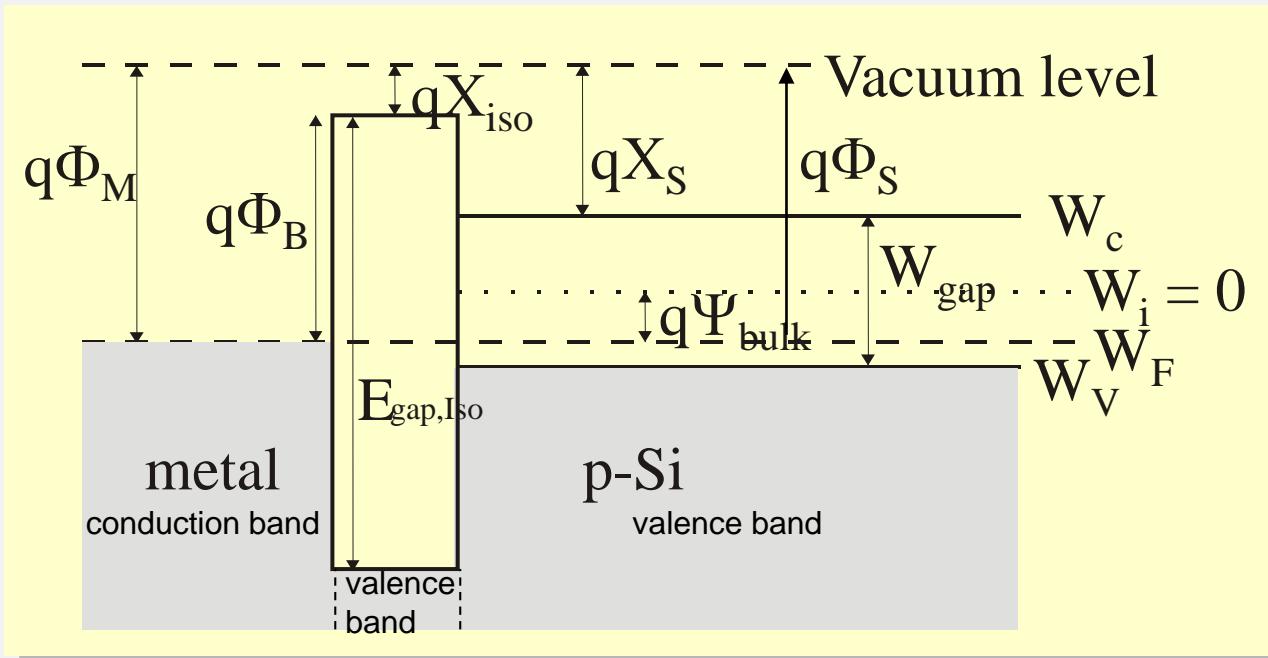
Starting point:

So-called "ideal" (= non realistic) MOS-system

- * the work function of the metal and the silicon should be the same (for example by doping)
- * no additional charges exist (interface states, oxide charges,...)



we have an initial flatband condition (no charges must be exchanged, because Fermi level (related to vacuum level) is the same in all materials)



Measurement techniques

* Photo emission -> determination of work functions

* C-V -> determination of barrier heights -> if interface states are present and (not allowed) Schottky-model applied -> changed work function

As already shown for the Schottky barrier heights the pure elemental work function cannot be used in practice to calculate the barrier heights because interface charges, interdiffusion and other effects change the real, measured barrier height.

The same effect is observed for MIS-systems, so the work function values presented in literature must be handled with care and may only give a first guess of work function calculations.

Commonly used values for affinity and work function in literature (on SiO_2/Si):

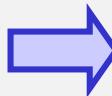
	$\chi \text{ Si}$	$\chi \text{ Ge}$	$\chi \text{ GaAs}$	$\chi \text{ GaP}$	$\chi \text{ SiO}_2$
$\chi [\text{V}]$	~ 4.05	~ 4.0	~ 4.07	~ 4.3	0.95
	n ⁺ -Si	p ⁺ -Si	Ti	Al	Pt

	n ⁺ -Si	p ⁺ -Si	Ti	Al	Pt
$\Phi [\text{V}]$	~ 4.1	~ 5.2	~ 3.9	~ 4.1	5.4

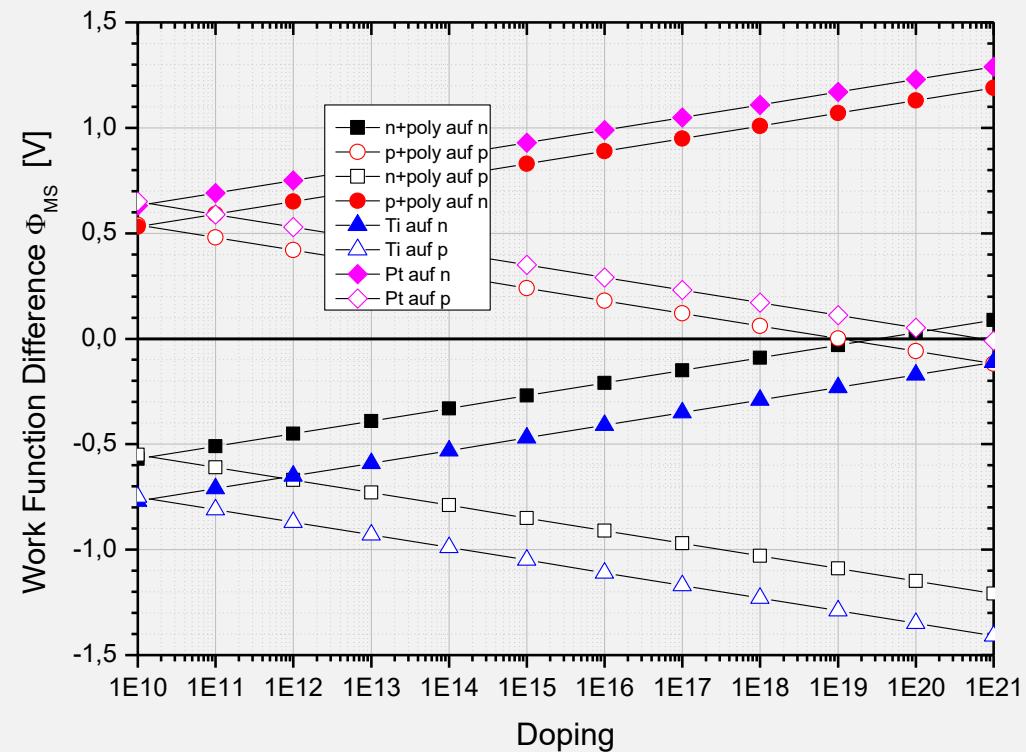
Doping dependent work function difference $\Phi_{MS} = \Phi_M - \Phi_S$

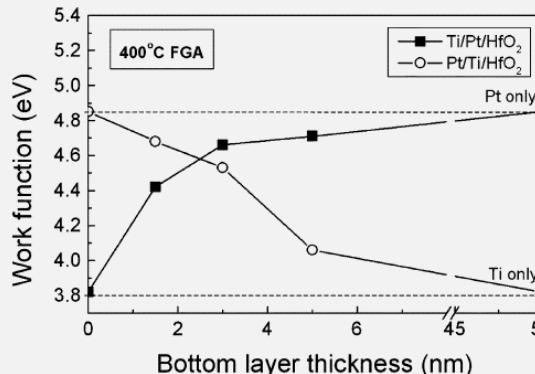
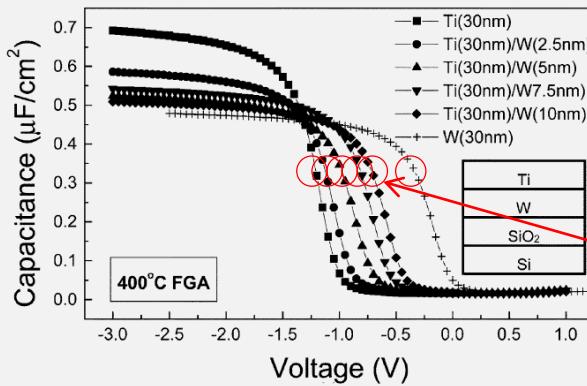
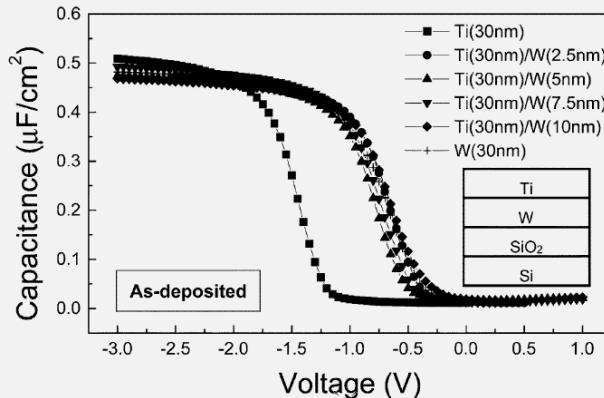
$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \left(\chi_s + \frac{W_{gap}}{2q} \pm |\Psi_b| \right)$$

↑ p-dop
↓ n-dop

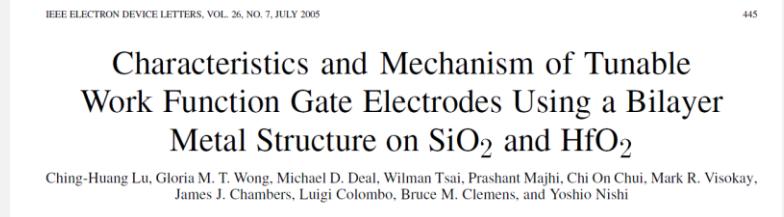


we are using now the potential/voltages instead of energies, because our way should lead us to the threshold voltage V_T of an MOSFET and the IV-characteristics





The work function of pure Ti on SiO_2/Si (~3.95 eV) is changed to the work function of pure W (~5.3 eV) by thin underlying W-layers

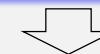


After forming gas anneal the work function is changing again, maybe due to interdiffusion of metals

Note: measured "flatband voltage" is used to extract workfunction

Also the work functions are changing on a different gate dielectrics (HfO_2 instead of SiO_2)

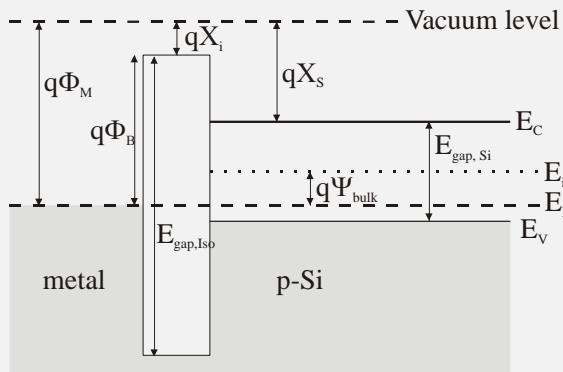
In reality the work function difference Φ_{MS} is dependent on the used materials and processes (cleaning, deposition, annealings, ...).



Calculation of work function difference Φ_{MS} (and later V_T) is only a first estimation

1) Ideal Case:

$\Phi_{\text{metal}} = \Phi_{\text{semiconductor}}$
and no technological charges



work function metal: Φ_m

$$\text{work function silicon: } \Phi_{Si} = \left(X_{Si} + \frac{E_{gap,Si}}{2 \cdot q} \oplus \Psi_{bulk} \right)$$

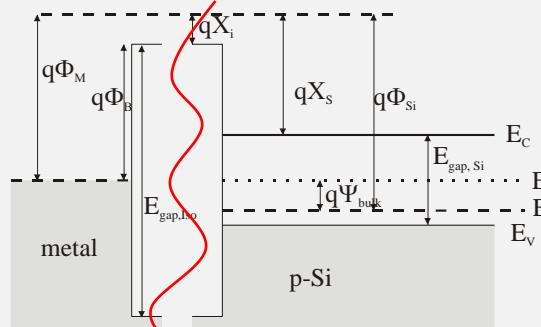
+ : p-type doping
- : n-type doping

Fermi level

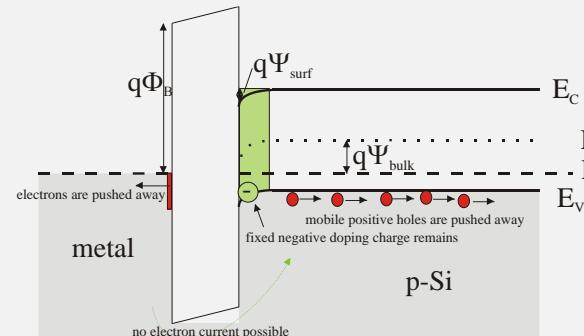
$$\Psi_{bulk} = kT \cdot \ln \left(\frac{N_{doping}}{n_i} \right)$$

2) More realistic case:

$\Phi_{\text{metal}} \neq \Phi_{\text{semiconductor}}$
and no technological charges



before contact, $\Phi_m < \Phi_{Si}$

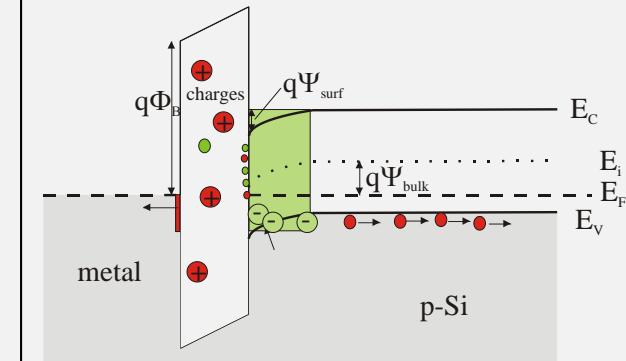


in contact

band bending in isolator and Si
-> a surface potential Ψ_{sur} appears

3) Realistic case:

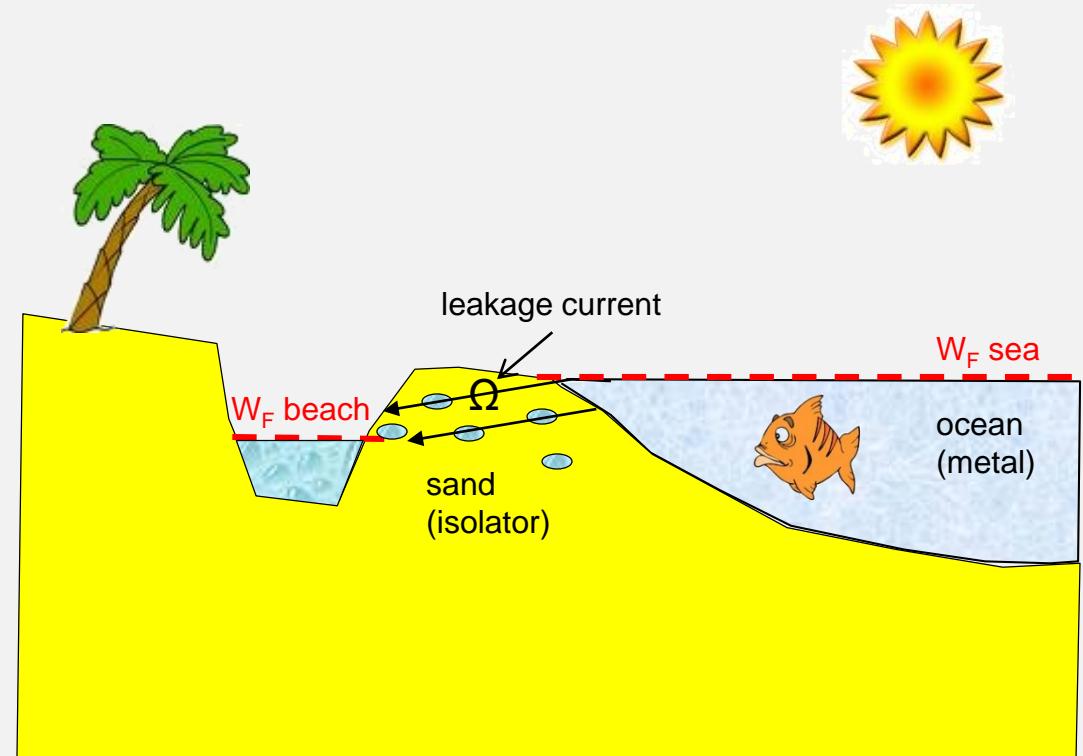
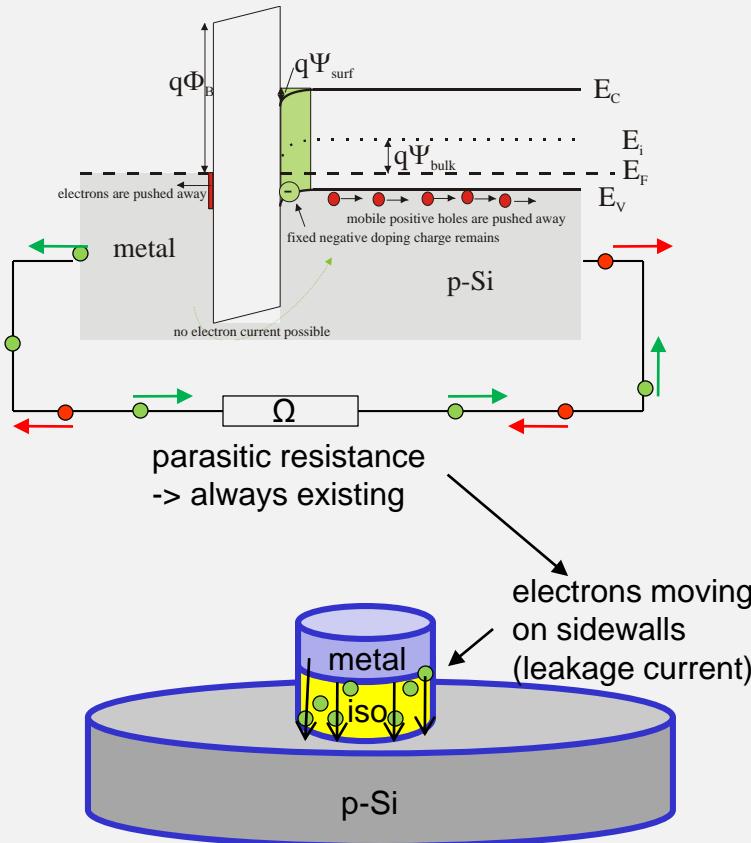
$\Phi_{\text{metal}} \neq \Phi_{\text{semiconductor}}$
and technological charges



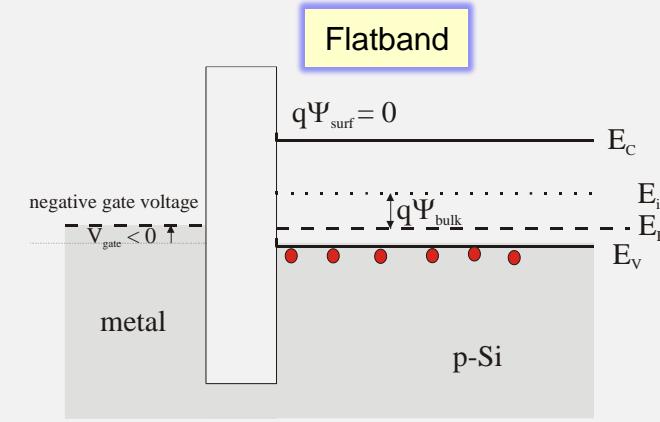
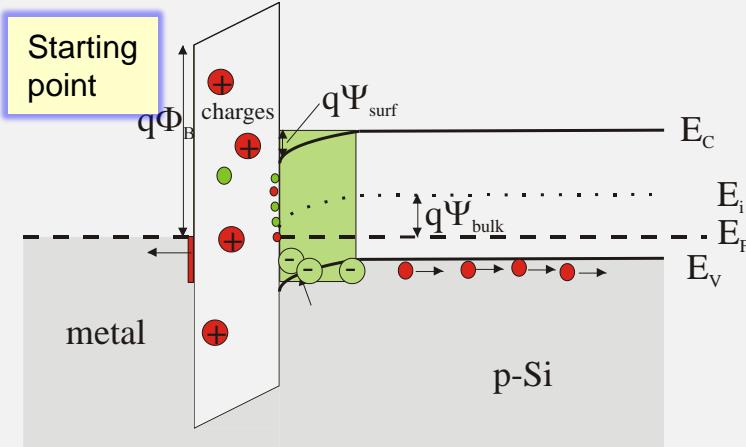
band bending in isolator and Si
-> a surface potential Ψ_{sur} appears

No one can predict the amount of band bending !

How is the electron transfer done ?

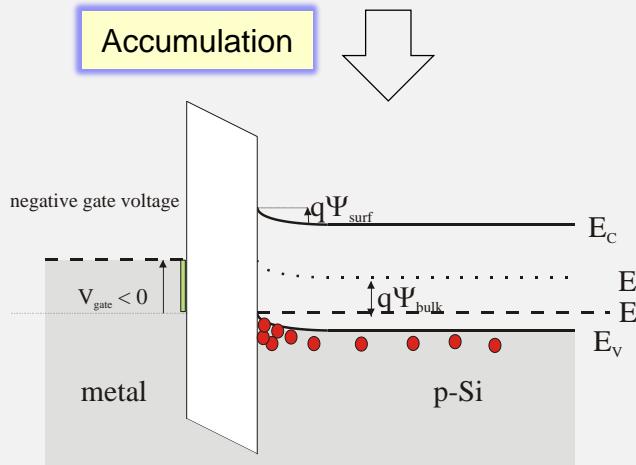


Realistic bandstructure, $V_{gate} = 0$

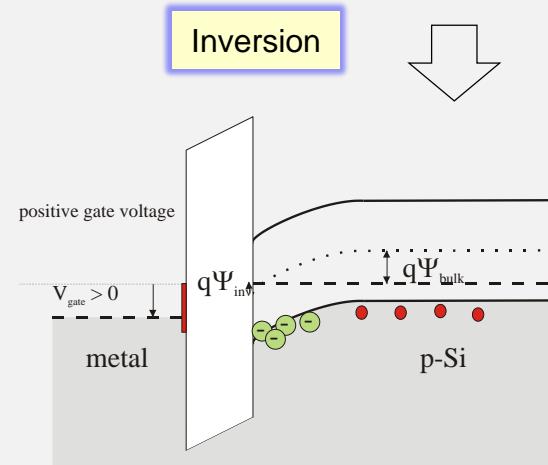


To achieve flatband condition the external voltage V_G must compensate the band bending by the work function difference Φ_{MS} and oxide charges Q_{ox}

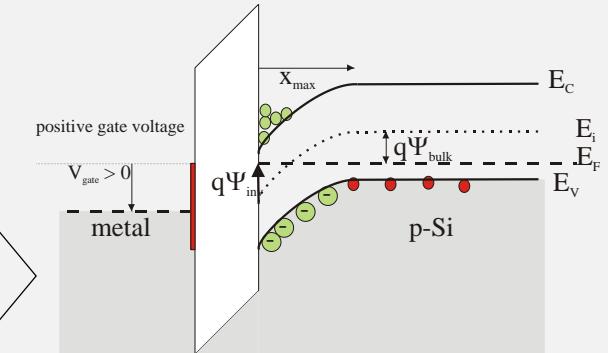
Accumulation



Inversion



Strong Inversion

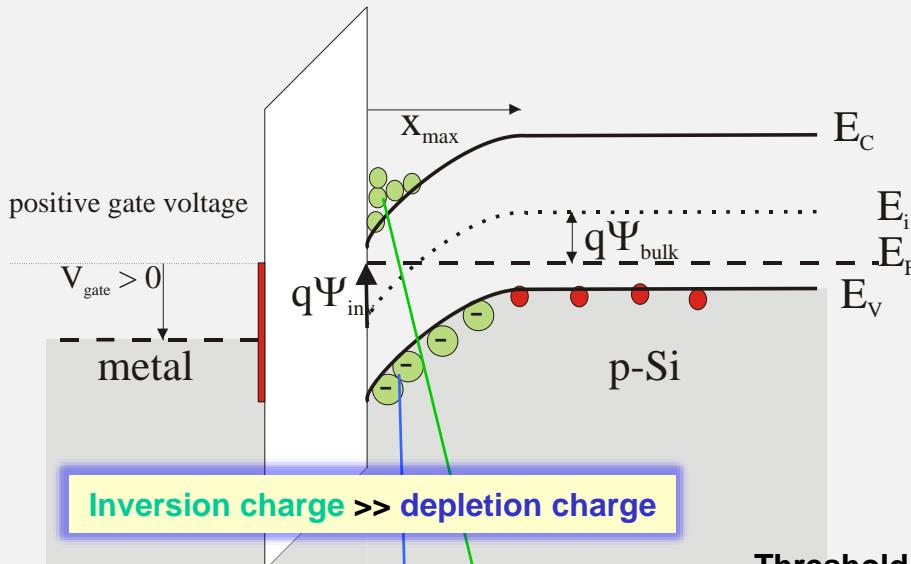


Mobile majority carriers are attracted to the SiO₂ / Si interface

Mobile majority carriers are pushed away from the SiO₂ / Si interface. The **fixed negative doping charge** remains and creates a depletion layer

Mobile minority carriers are attracted to the SiO₂ / Si interface, an inversion layer is created

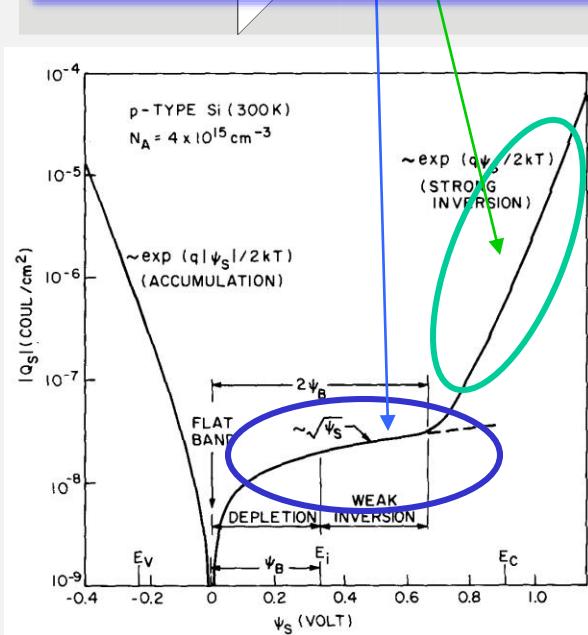
Definition of Threshold Voltage

Definition:

Strong Inversion is achieved, if the band bending is twice bulk Fermi level

In this point the potential at the SiO_2/Si interface is determined by the surface potential $2\Psi_b$ and the gate capacitor charge (depletion charge)Definition of ideal threshold voltage V_T :ideal = same work-function,
+ no oxide charges

$$V_T = \pm \frac{Q_{depl}}{C''_{ox}} \pm 2\Psi_{bulk} = \pm \frac{\sqrt{2\epsilon_0\epsilon_{Si} \cdot qN_{dop} \cdot |2\Psi_{bulk}|}}{C''_{ox}} \pm 2\Psi_{bulk}$$

+ signs: p-substrate
- signs: n-substrateThreshold voltage in reality: (with work function difference Φ_{MS} and oxide charges Q_{ox})

$$V_T = \left(\Phi_{MS} \pm \frac{Q_{ox}}{C''_{ox}} \pm \frac{Q_{impl}}{C''_{ox}} \right) \pm \frac{Q_{dep\max}}{C''_{ox}} \pm |2\Psi_{bulk}| = V_{FB} \pm \frac{\sqrt{2\epsilon_0\epsilon_{Si} \cdot qN_{dop} \cdot (|2\Psi_{bulk}| \pm V_{BS})}}{C''_{ox}} \pm |2\Psi_{bulk}|$$

Depending on the substrate doping (n- or p-type), on the nature of oxide charges (+ or -) and channel doping (p- or n-type) to adjust threshold voltage, the \pm signs have to be taken into account

Maximum width of depletion layer:



$$w_{max} = \sqrt{\frac{2\epsilon_0\epsilon_{Si} \cdot 2\Psi_{bulk}}{q \cdot N_{dop}}} \approx (50 - 500) \text{ nm}$$

Thickness of strong inversion layer:

$$d_{inv} \propto \frac{1}{V_{gate}} \sim (2 - 10) \text{ nm}$$

for a n-MOSFET with p-type substrate:

Basically V_T on the Gate electrode of the Gate capacitor must be (more) positive, because depletion and inversion charge is negative

$$V_T = \left(\left(\Phi_M - \left(\chi_{semi} + \frac{W_{gap}}{|2q|} + |\Psi_{bulk}| \right) \right) \pm \frac{Q''_{ox}}{C''_{ox}} \pm \frac{Q''_{impl}}{C''_{ox}} \right) + \frac{Q''_{dep\ max}}{C''_{ox}} + |2\Psi_{bulk}| = V_{FB} + \frac{\sqrt{2\varepsilon_0\varepsilon_{Si} \cdot qN_A \cdot (2|\Psi_{bulk}| - V_{BS})}}{C''_{ox}} + |2\Psi_{bulk}|$$

Ψ_{bulk} : + sign, because higher p-doping needs more (positive) Gate voltage to achieve 2Ψ band bending for inversion

Q_{ox} : - sign, if charges are positive $Q+$, then depletion or inversion is created and therefore less (positive) V_T needed
(usually oxide charges are positively charged, when thermal oxides at high temperatures are used)

+ sign, if charges are negative $Q-$, then band bending is towards accumulation, therefore a higher (positive) V_T is necessary for inversion
(usually, if oxides are grown or deposited at low temperatures (and plasma-assisted -> electrons create neg. charges)

Q_{impl} : + sign, if implant doping is N_A , then the p-substrate doping is increased, a higher barrier results and (positive) V_T is increased
- sign, if implant doping is N_D , then the p-substrate doping is lowered, a lower barrier results and (positive) V_T is lowered

$Q_{dep\ max}$: + sign, because higher p-substrate doping increases channel barrier and therefore (positive) threshold voltage V_T

V_{BS} : - sign. If a positive bulk voltage (Source grounded) is applied, the positive bulk bias biases the Source/bulk diode in forward direction, which reduces also the barrier of the Source-channel diode such reduces the (positive) threshold voltage. To limit the Source/bulk current V_{BS} should be lower than $2\Psi_b$.
If negative bulk voltage (Source grounded) is applied, the Source barrier is increased and in consequence (positive) threshold voltage $(-V_{BS})=+V_{BS}$.

for a p-MOSFET with n-type substrate:

Basically VT on the Gate electrode of the Gate capacitor must be (more) negative, because depletion and inversion charge is positive

$$V_T = \left(\left(\Phi_M - \left(\chi_{semi} + \frac{W_{gap}}{|2q|} - |\Psi_{bulk}| \right) \right) \pm \frac{Q''_{ox}}{C''_{ox}} \pm \frac{Q''_{impl}}{C''_{ox}} \right) - \frac{Q''_{dep\ max}}{C''_{ox}} - |2\Psi_{bulk}| = V_{FB} - \frac{\sqrt{2\varepsilon_0\varepsilon_{Si} \cdot qN_{dop} \cdot (2|\Psi_{bulk}| + V_{BS})}}{C''_{ox}} - |2\Psi_{bulk}|$$

Ψ_{bulk} : - sign, because higher n-doping needs more (negative) Gate voltage to achieve 2Ψ band bending for inversion

Q_{ox} : + sign, if charges are negative $Q-$, then depletion or inversion is created and therefore less (negative) V_T needed
- sign, if charges are positive $Q+$, then band bending is towards accumulation and higher (negative) V_T is necessary for inversion

Q_{impl} : - sign, if implant doping is N_D , then the n-substrate doping is increased, a higher barrier results and (negative) V_T is increased
+ sign, if implant doping is N_A , then the n-substrate doping is lowered, a lower barrier results and (negative) V_T is lowered

$Q_{dep\ max}$: - sign, because higher n-substrate doping increases channel barrier and therefore (negative) threshold voltage V_T

V_{BS} : - sign. If a negative bulk voltage (Source grounded) is applied, the negative bulk bias biases the Source/bulk diode in forward direction, which reduces also the barrier of the Source-channel diode such reduces the (negative) threshold voltage. To limit the Source/bulk current V_{BS} should be lower than $2\Psi_b$.
If positive bulk voltage (Source grounded) is applied, the Source barrier is increased and in consequence (more negative) threshold voltage.

$$V_T = \left(\Phi_{MS} \pm \frac{Q''_{ox}}{C''_{ox}} \pm \frac{Q''_{impl}}{C''_{ox}} \right) \pm \frac{Q''_{dep\ max}}{C''_{ox}} \pm |2\Psi_{bulk}| = V_{FB} \pm \frac{\sqrt{2\epsilon_0\epsilon_{Si} \cdot qN_{dop} \cdot (|2\Psi_{bulk}| \pm V_{SB})}}{C''_{ox}} \pm |2\Psi_{bulk}|$$

Rough estimation of contribution of the various parts in the VT-equation:

we assume:

Substrate doping: $\sim 10^{18}$



$$\Psi_{bulk} = \pm k_B T \cdot \ln \left(\frac{N_{doping}}{n_i} \right) = \pm 0.026 \text{ V} \cdot \ln \left(\frac{N_{doping}}{n_i} \right) = \pm 0.026 \text{ V} \cdot \ln \left(\frac{10^{18}}{1.45 \cdot 10^{10}} \right) \approx \pm 0.5 \text{ V}$$

$10^{17} \text{ cm}^{-3} \rightarrow \sim 0.4 \text{ V}_{bulk}$

Work-function difference Φ_{MS} : $\pm 1.0 \text{ V}$ for poly-Gate CMOS, -1 V (n-MOS), $+1 \text{ V}$ (p-MOS)

Oxide capacitance:

(we assume $t_{ox} = 1 \text{ nm}$)

$$C''_{ox} = \frac{\epsilon_0 \cdot \epsilon_{ox}}{t_{ox}} = \frac{8.85 \cdot 10^{-14} \text{ F/cm} \cdot 4}{1 \text{ nm}} = 3 \cdot 10^{-6} \text{ F/cm}^2$$

Typical oxide charges: $\sim 10^{11} \text{ q/cm}^2$



$$Q''_{ox} = 1.6 \cdot 10^{-19} \text{ C} \cdot 10^{11} / \text{cm}^2 \approx 10^{-8} \text{ A sec/cm}^2$$



$$\Delta V_{ox} = \frac{Q''_{ox}}{C''_{ox}} = \frac{10^{-8} \text{ A sec/cm}^2}{10^{-6} \text{ F/cm}^2} \approx 0.01 \text{ V}$$

Depletion charge:

$$Q''_{dep\ max} = \sqrt{2\epsilon_0\epsilon_{Si} \cdot qN_{dop} \cdot |2\Psi_{bulk}|} \approx \sqrt{2 \cdot 8.85 \cdot 10^{-14} \text{ F/cm} \cdot 12 \cdot 1.6 \cdot 10^{-19} \text{ C} \cdot 10^{18} \text{ cm}^{-3} \cdot |2 \cdot 0.5 \text{ V}|} \approx 6 \cdot 10^{-7} \frac{\text{A sec}}{\text{cm}^2}$$

Change in threshold voltage:

$$\Delta V_{dep\ max} = \frac{Q''_{dep\ max}}{C''_{ox}} = \frac{6 \cdot 10^{-7} \text{ A sec/cm}^2}{10^{-6} \text{ F/cm}^2} \approx 0.6 \text{ V}$$

all together:

$$V_T(\text{n-MOS}) = -1 \text{ V} \pm 0.01 \text{ V}_{ox} \pm 0 \text{ V}_{imp} + 0.6 \text{ V}_{dep} + 1 \text{ V}_{bulk} \sim +0.6 \text{ V}$$

$$V_T(\text{p-MOS}) = +1 \text{ V} \pm 0.01 \text{ V}_{ox} \pm 0 \text{ V}_{imp} - 0.6 \text{ V}_{dep} - 1 \text{ V}_{bulk} \sim -0.6 \text{ V}$$

ok

(symmetrical V_T)

Threshold Channel Doping

$$V_T = \left(\Phi_{MS} \pm \frac{Q''_{ox}}{C''_{ox}} \pm \frac{Q''_{impl}}{C''_{ox}} \right) + \frac{Q''_{depmax}}{C''_{ox}} \pm |2\Psi_{bulk}| = V_{FB} \pm \frac{\sqrt{2\epsilon_0\epsilon_{Si} \cdot qN_{dop} \cdot (|2\Psi_{bulk}| \pm V_{SB})}}{C''_{ox}} \pm |2\Psi_{bulk}|$$

2) Wanne- und Kanal-Implantationen

2.1 p-Wanne

- Oxidation Streuoxid (15 nm)
- DD-Messung Strukturierung LOCOS (Anlage KLA)
- FT p-Wanne (J), unkritische Ebene
- Impl. p-Wanne (B - 5,0E12 cm⁻² - 230 keV)
- Impl. n-Kanal (B - 3,6E12 cm⁻² - 25 keV)

Bem: Es erfolgt keine separate Feld-Implantation. Die Einsatzspannung des Feldoxidtransistors wird durch die Wannenimplantation eingestellt.

Using the equation of threshold voltage we can shift the threshold voltage with an implant

Example:

Let the threshold voltage be: $V_T = +0.3$ V. We want to increase threshold voltage to reduce leakage current in a DRAM access n-MOSFET up to 2.0 V. The thickness of the gate oxide is 1.6nm. Which implant is needed?

Answer:

The threshold voltage must be increased by +1.7 V, because it is an n-MOSFET we have to implant p-doping atoms, e.g. boron.

We calculate:

$$C''_{ox} = \frac{\epsilon_0 \cdot \epsilon_{ox}}{t_{ox}} = \frac{8.85 \cdot 10^{-14} F/cm \cdot 4}{1.6 nm} = 6.6 \cdot 10^{-6} F/cm^2$$

$$\Delta V_T = \frac{Q''_{imp}}{C''_{ox}} = \frac{N_A}{6.6 \cdot 10^{-6} F/cm^2} = 1.7 V$$



$$Q''_{imp} = \Delta V_T \cdot C''_{ox} \approx 1 \cdot 10^{-5} Asec/cm^2$$

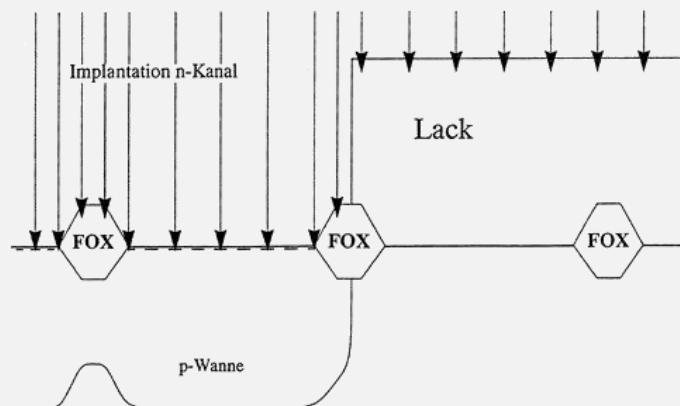


Abb. 4: Querschnitt nach Implantation n-Kanal

C6NH3-Prozeßflow

J. Kreisbeck
HL FE R PIE 1

Example of ancient CMOS-process,
SIEMENS, ~1990, 0.5 µm technology

The required implant dose is:

$$D_{N_A} = \frac{Q''_{imp}}{q} \approx \frac{10^{-5} Asec/cm^2}{1.6 \cdot 10^{-19} Asec} \approx 6 \cdot 10^{13} / cm^2$$

3.1 MOSFET intro

- Basic semiconductor physics
- Basic MOSFET working principle
- MOSFET types
- MOSFET applications
- Basic MOSFET fabrication process

3.2 MOS physics

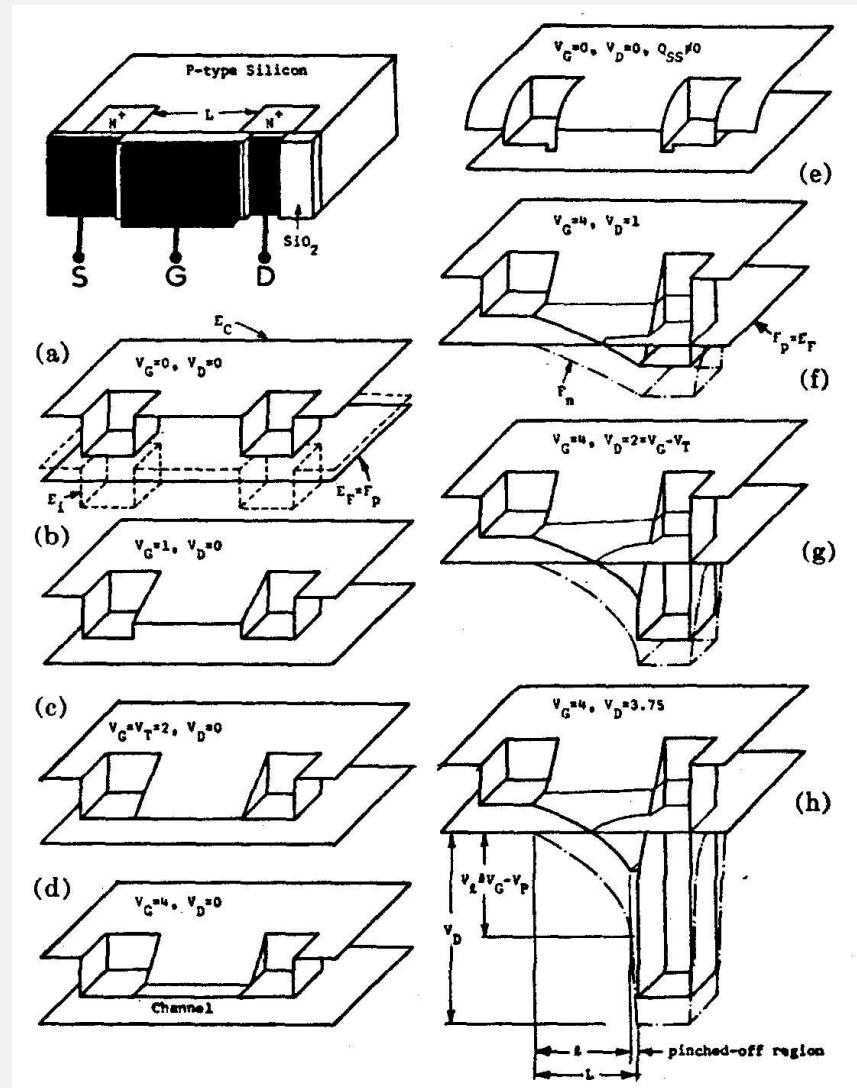
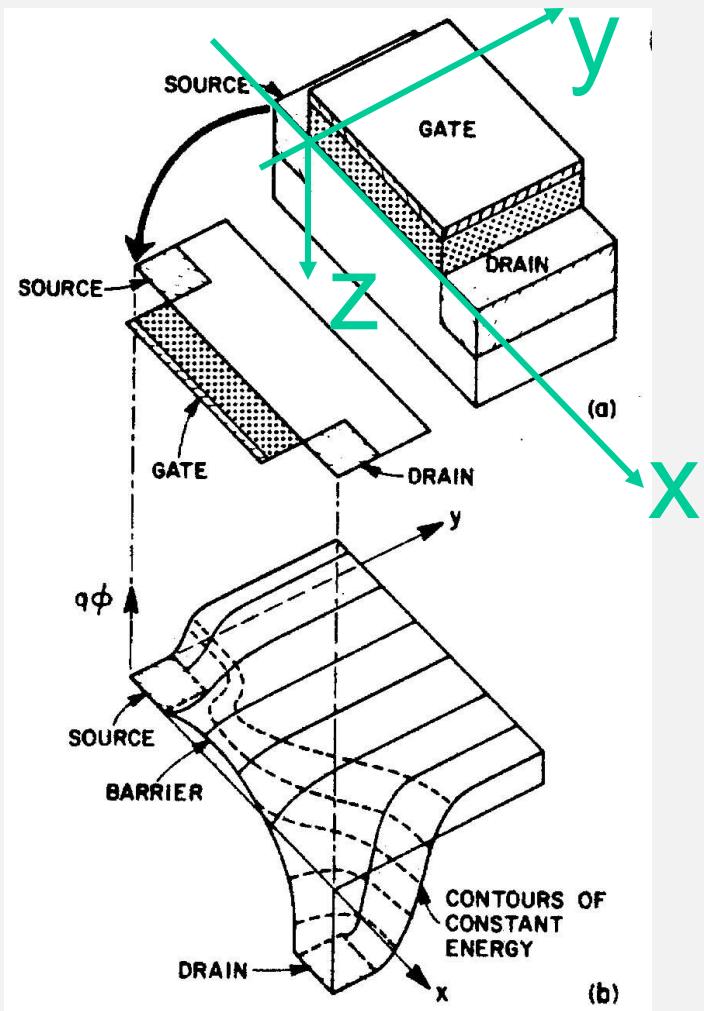
- Work function
- Contact phenomena
 - Metal - metal
 - Metal - semiconductor
 - Metal - insulator - semiconductor (MIS)
- Working regimes of the MOS structure

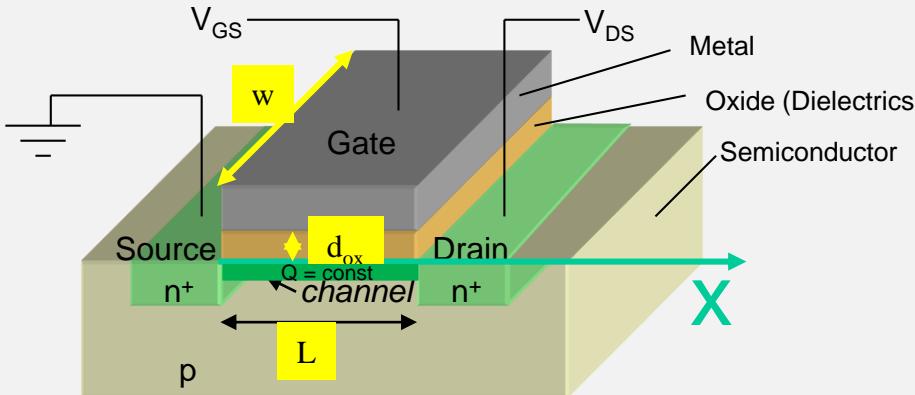
3.3 Ideal long-channel MOSFET

- Charge models (Q-constant, Gradual-Channel)
- MOSFET working regimes

3.4 Realistic behavior

- Basics of charge carrier transport
- Modification in charge carrier transport
- Break-through phenomena
- Parasitics

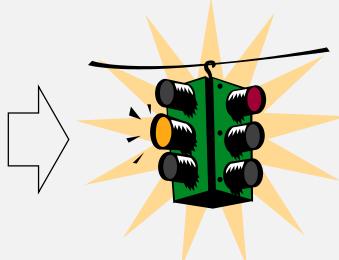




Calculating the Gate Capacitor:

$$Q = C_{gate} \cdot V_{GS} = \frac{\epsilon_0 \epsilon_r \cdot w \cdot L}{d_{ox}} \cdot V_{GS}$$

(movable) charge in the channel



$$I_{DS} = \mu \cdot C'' \cdot \frac{w}{L} \cdot V_{GS} \cdot V_{DS}$$

$= \frac{1}{R_{channel}}$

Assumption:

strong inversion (channel on):

- the inversion charge \gg depletion charge,
so the number of charges is determined by gate capacitor charge

$V_{gate} \gg V_{DS}$:

- the charge Q below the gate is not dependent on V_{DS}
or Q_{gate} is not dependent on channel position x

Calculating the Source-Drain Resistor:

Transit time:

$$t_{tr} = \frac{L}{v} = \frac{L}{\mu \cdot E_x} = \frac{L^2}{\mu \cdot V_{DS}} \quad (\text{see chap.1})$$

current:

$$I_{DS} = \frac{Q}{t_{tr}} = \frac{\epsilon_0 \epsilon_r \cdot w \cdot L}{d_{ox}} \cdot V_{GS} \cdot \frac{\mu \cdot V_{DS}}{L^2}$$

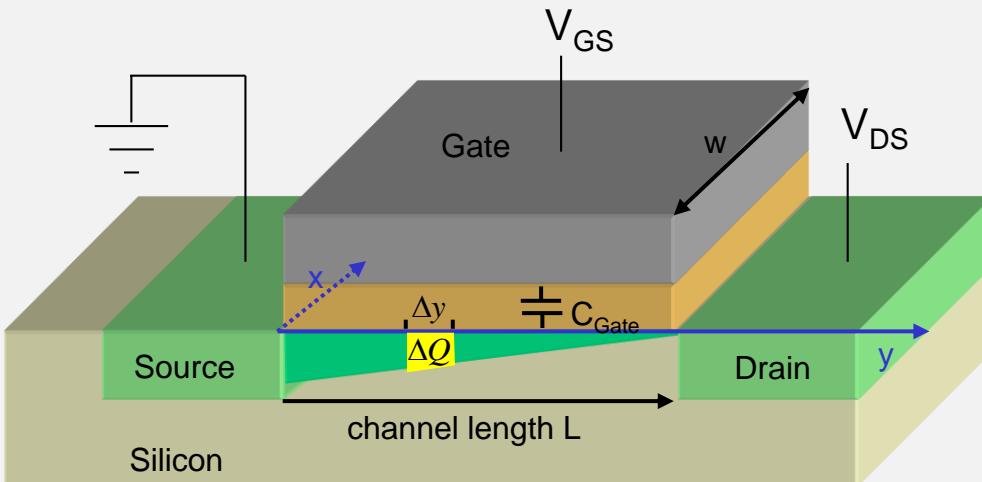
$$= \frac{\epsilon_0 \epsilon_r}{d_{ox}} \cdot \mu \frac{w}{L} \cdot V_{GS} \cdot V_{DS} = \mu \cdot C'' \cdot \frac{w}{L} \cdot V_{GS} \cdot V_{DS}$$

using the capacity/area C''

$$C''_{ox} = \frac{C_G}{w \cdot L} = \frac{\epsilon_0 \epsilon_{ox}}{d_{ox}}$$

! too simple !

(just a resistor, but this is, what we put in)



Gradual Channel Approximation

The applied Drain voltage drops along the channel from Source ($V_y=0$) to Drain ($V_y=V_D$), therefore the voltage difference $V_G - V_y$ varies along the channel and therefore the created inversion charge $Q(y)$:

$$Q(y) = C_{Gate} \cdot [(V_{GS} - V_T) - V(y)]$$

To calculate the local charge $Q(y)$ we divide the channel in small pieces Δy , containing pieces of charge ΔQ

$$\Delta Q(y) = \Delta C_G \cdot [(V_{GS} - V_T) - V(y)] \quad \text{with} \quad \Delta C_G = C_G \cdot \frac{\Delta y}{L}$$

From continuity equation:

$$j = ne \cdot v$$

we calculate the local current in the section Δy with:

$$I(y) = \rho_y \cdot v(y)$$

using the charge per length:

$$\rho_y = \frac{\Delta Q(y)}{\Delta y}$$

and the **drift velocity**:

$$v(y) = \mu \cdot E(y) = \mu \cdot \frac{dV(y)}{dy}$$



we do not use diffusion current, only drift current !



we assume mobility μ should be a constant !

All together:

$$I(y) = \rho_y \cdot v(y) = \frac{\Delta Q(y)}{\Delta y} \cdot \mu \cdot \frac{dV_y}{dy} = \frac{\Delta C_G \cdot [(V_{GS} - V_T) - V(y)]}{\Delta y} \cdot \mu \cdot \frac{dV_y}{dy} = \frac{C_G \cdot \frac{\Delta y}{L} \cdot [(V_{GS} - V_T) - V(y)]}{\Delta y} \cdot \mu \cdot \frac{dV_y}{dy}$$

Simple differential equation -> we separate length dependent parameters on one side and voltage dependent parameters on the other side.
Finally we have to integrate along the channel:

$$\int_0^L I(y) dy = \frac{C_G \cdot \mu}{L} \cdot \int_0^{V_D} [(V_{GS} - V_T) - V(y)] dV_y = \frac{C_G \cdot \mu}{L} \cdot \left[\int_0^{V_D} [(V_{GS} - V_T)] dV_y - \int_0^{V_D} V(y) dV_y \right]$$

here we split the voltage integral

resulting in:

$$I(y) \cdot y \Big|_0^L = I_D \cdot L = \frac{C_G \cdot \mu}{L} \cdot \left[(V_{GS} - V_T) \cdot V_D - \frac{V_D^2}{2} \right] \rightarrow$$

$$I_D = \frac{C_G \cdot \mu}{L^2} \cdot \left[(V_{GS} - V_T) \cdot V_D - \frac{V_D^2}{2} \right]$$

or as commonly used:

$$I_D = \mu C_{ox} \cdot \frac{w}{L} \cdot \left[(V_{GS} - V_T) \cdot V_{DS} - \frac{V_{DS}^2}{2} \right]$$

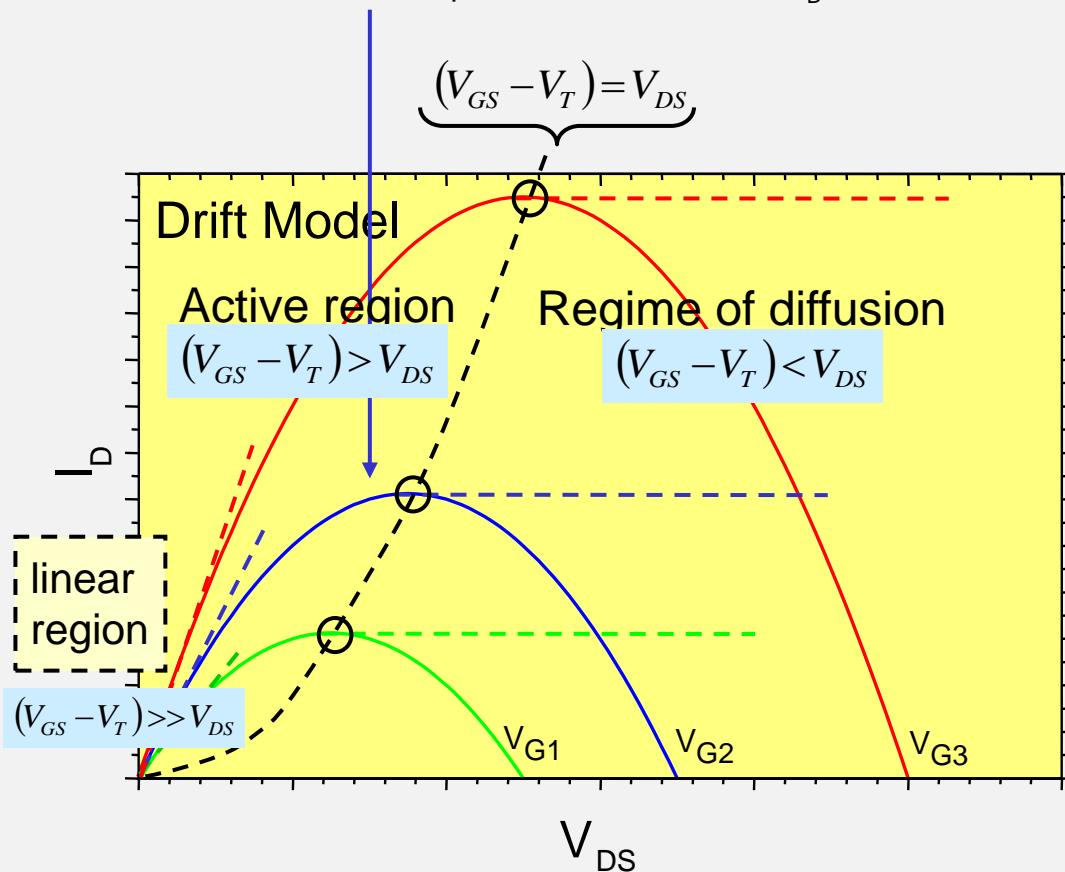
using the capacity/area C'

$$C_{ox} = \frac{C_G}{w \cdot L} = \frac{\epsilon_0 \epsilon_{ox}}{d_{ox}}$$

units: $I_D [A] = \mu \left[\frac{cm^2}{V \cdot sec} \right] \cdot C_{ox} \left[\frac{A \cdot sec}{V \cdot cm^2} \right] \cdot \frac{w}{L} [V] \cdot \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] [V^2]$

$$I_D = \mu C_{ox} \frac{w}{L} \left[(V_{GS} - V_T) \cdot V_{DS} - \frac{V_{DS}^2}{2} \right]$$

This is a parabolic behavior with V_D !



But we know, that after reaching points of maximum current I_D , the current remains constant (= saturation). This is due to **pinch-off** and resulting **diffusion**.

We can calculate the border points between drift and diffusion:

$$I_D = k \cdot \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \quad \text{using: } k = \mu C_{ox} \frac{w}{L}$$

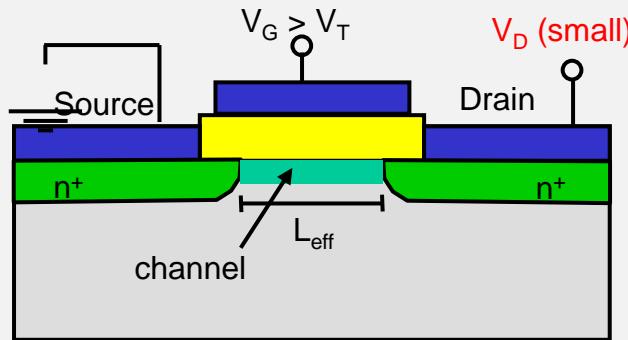
$$\rightarrow \frac{dI_D}{dV_{DS}} = k \cdot [(V_{GS} - V_T) - V_{DS}] \stackrel{!}{=} 0$$

$$\rightarrow (V_{GS} - V_T) = V_{DS}$$

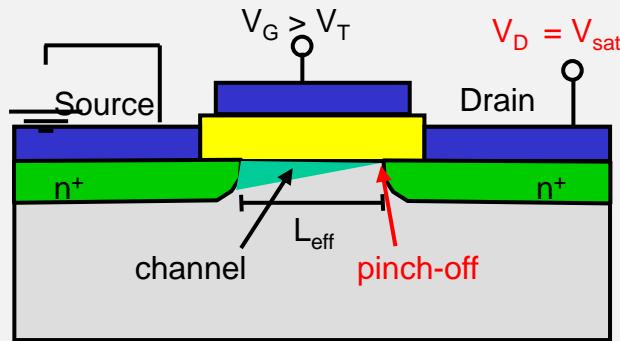
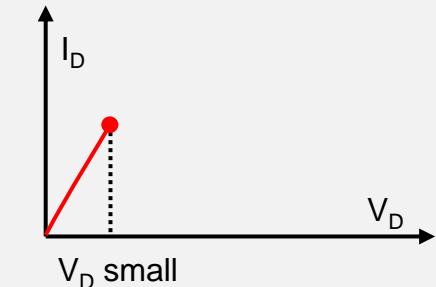
$$\rightarrow I_D = k \cdot \left[(V_{GS} - V_T) \cdot (V_{GS} - V_T) - \frac{(V_{GS} - V_T)^2}{2} \right]$$

$$I_D = \frac{k}{2} \cdot (V_{GS} - V_T)^2$$

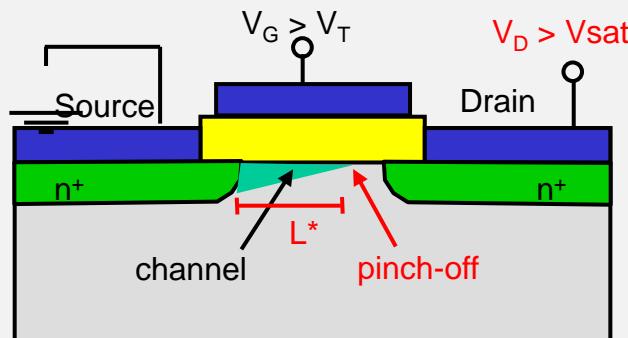
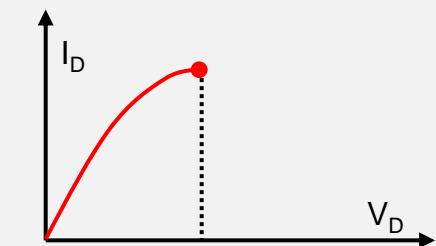
quadratic saturation characteristics of current maxima (pinch-off points)



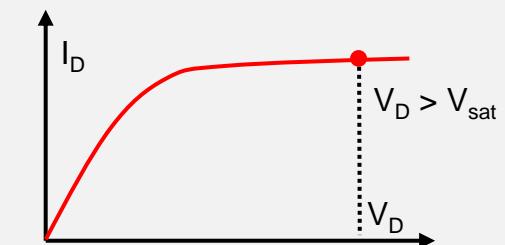
with small V_D
the channel charge
is nearly constant
from Source to Drain
-> resistor behavior

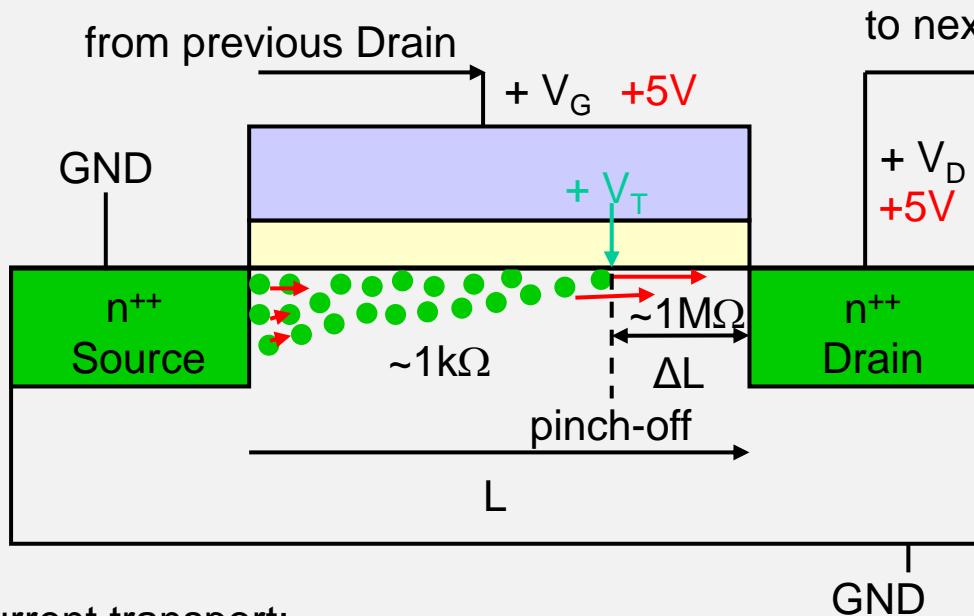


if at Drain $V_D = V_G - V_T$
no channel exists,
this is called **pinch-off**.
At this point the maximum
in current is reached:
 $V_D = V_{sat}$



if $V_D > V_{sat}$ the pinch-off
point ($V_y = V_G - V_T$) is
reached closer to Source.
The channel length L is
reduced to L^* .
-> to reach Drain the electrons
must diffuse





- ▶ Drain current of MOSFET 1 loads the Gate of following MOSFET
-> $V_{\text{Drain}} = V_{\text{Gate}}$
- ▶ - at Source the channel potential is $V_G - V_S \sim V_G$ -> a channel exists
- at Drain $V_G \sim V_D$ -> no channel exists
- ▶ no channel, but depletion/inversion
-> ΔL = isolator
-> drain-source voltage mainly drops over short ΔL -> high electric field

Current transport:

continuity equation: $\vec{j} = n \cdot e \cdot \vec{v} = \text{const}$

Many electrons start with low velocity at the Source in the channel, accelerated by drain-source voltage.
Few electrons reach the end of channel, but with higher velocity (-> $j = \text{const.}$).

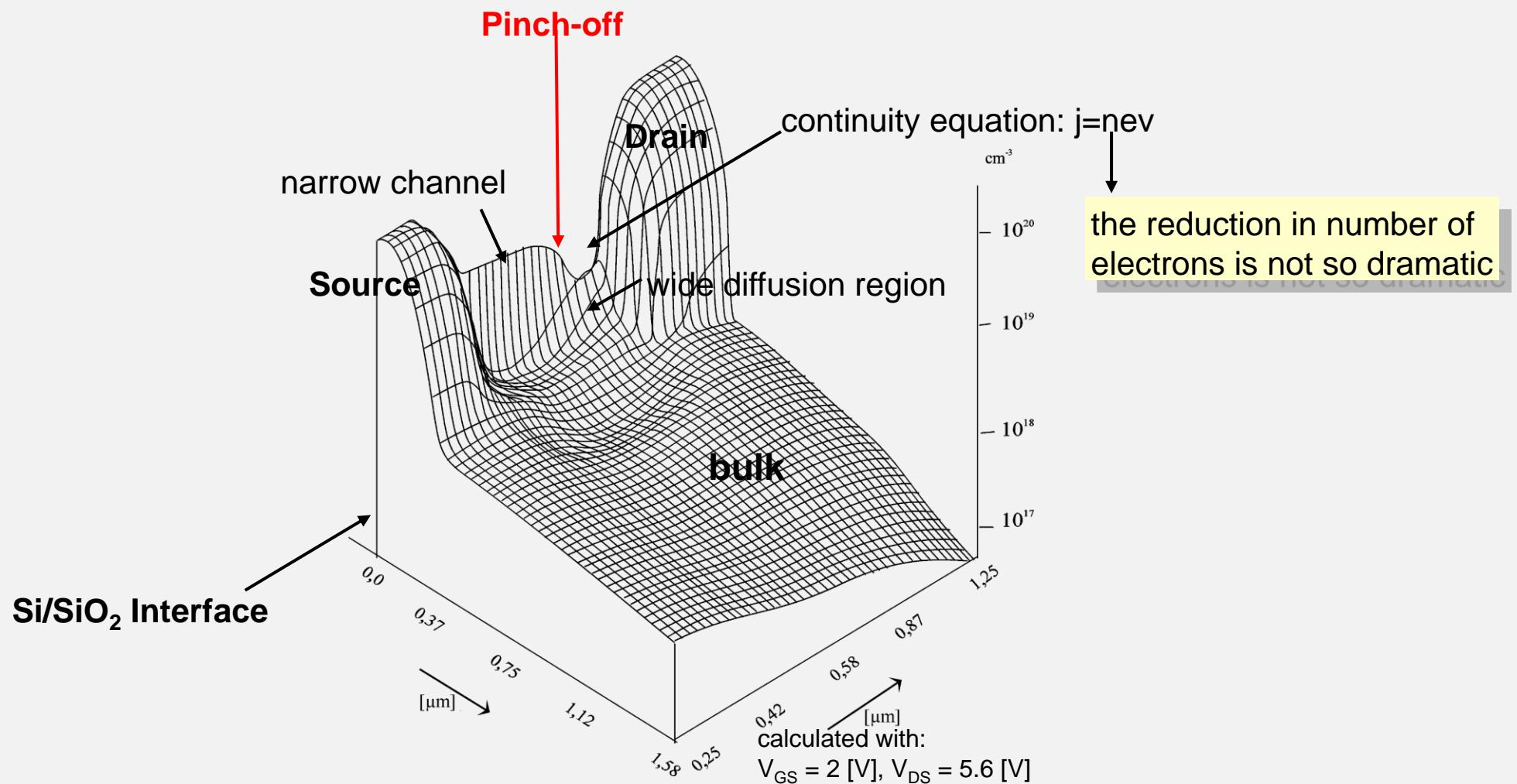
At the end of channel electrons jump in ΔL by diffusion and are highly accelerated by the high electric field.

If the electric field is too high -> **impact ionization** and avalanche break-through occurs

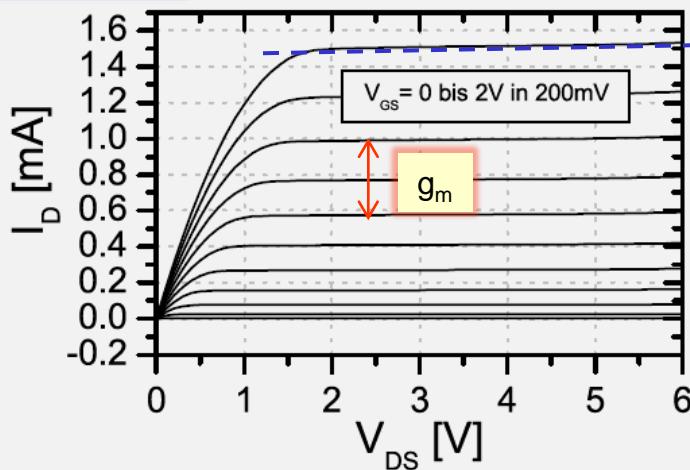
Technological solution: increase conductivity in front of Drain by an additional resistor

→ **Lightly-Doped Drain LDD**

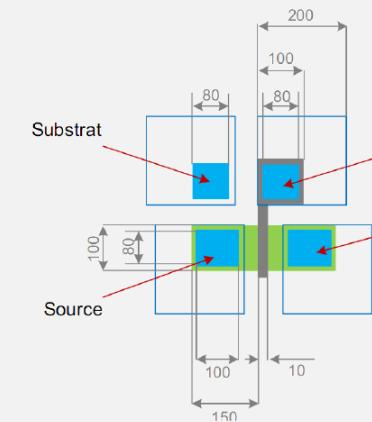
▶ today LDD structure is used for highly doped "extensions"



Output I-V

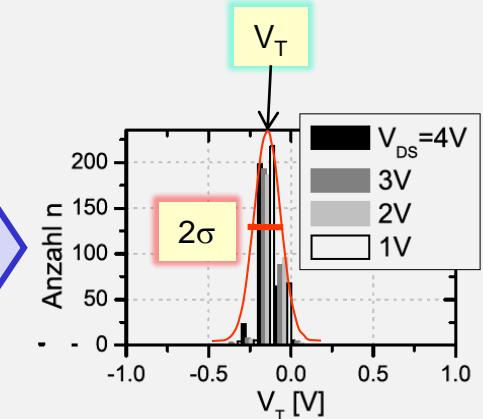
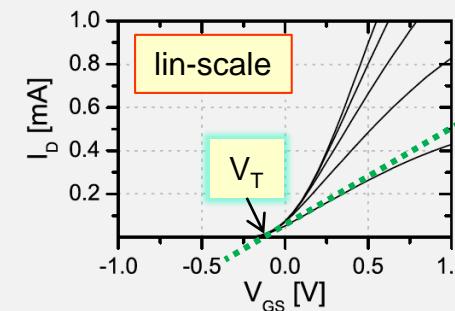
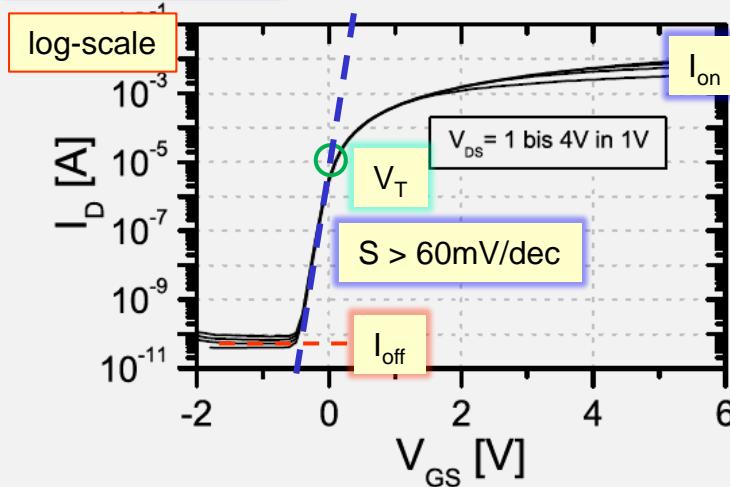


MOSFETs fabricated in our labs



Channel length: 10 μ m
w/L: 10
Tox : 2.6 nm

Transfer I-V

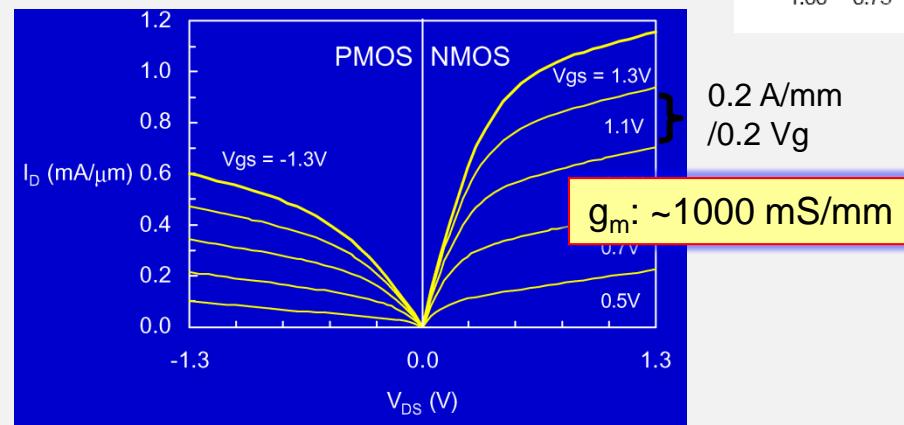
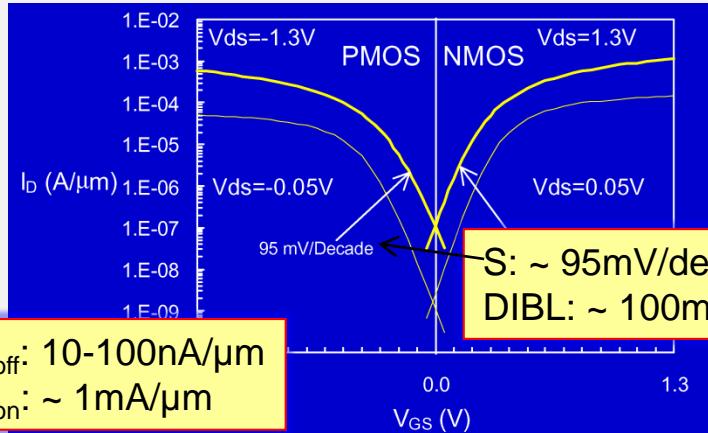


All extracted parameters exhibit statistical fluctuations due to the degree of process controllability

Typical performance values from INTEL, FUJITSU, Motorola, Hitachi:

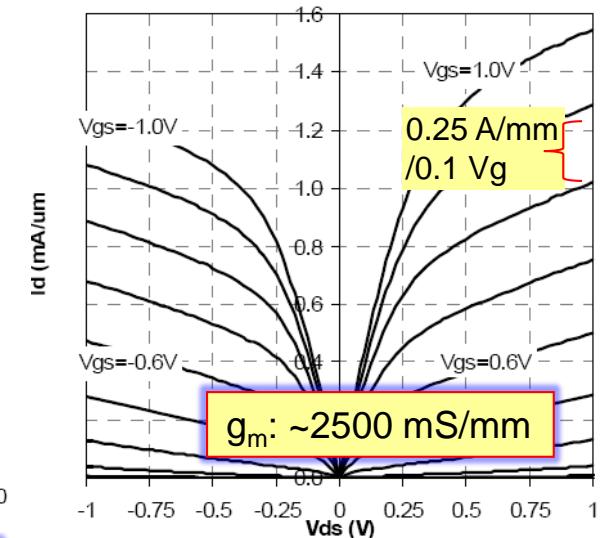
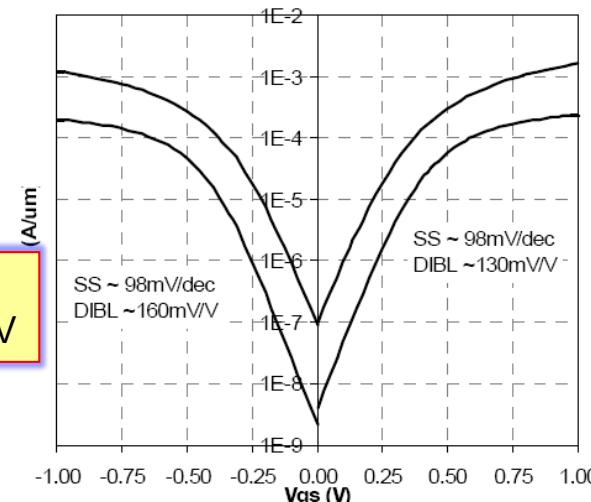
2000 Node: 130 nm

L_{gate} : ~ 70nm, t_{ox} : ~ 1.5nm
 V_{DD} : ~ 1.2 V, V_t : ~ 0.5 - 0.4 V



2009 Node: 32 nm

L : ~ 32 nm, EOT: ~ 0.9 nm high-k HfSiON
 V_{DD} : ~ 1.0 V, V_t : ~ 0.4 V



INTEL, IEDM 2008

Subthreshold

The subthreshold region is important for low-voltage and low-power applications:

- leakage current (for example in discharge in DRAM -> retention time or ULSI -> static power consumption)
- switch-on and off characteristics (e.g subthreshold slope -> switching speed)

Physics: weak inversion -> the MOSFET acts like a bipolar transistor (for example like a npn-structure)
-> current is diffusion current

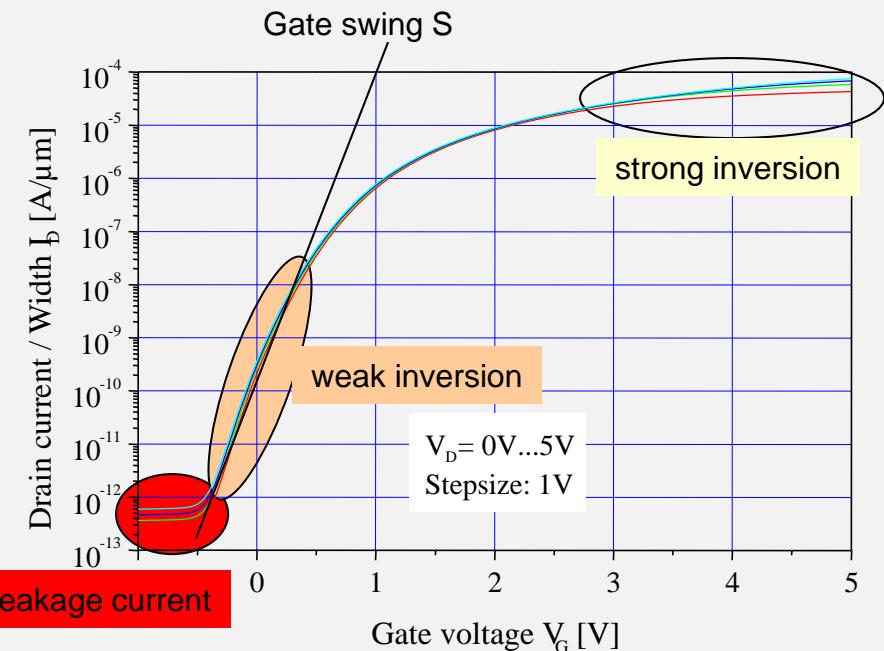
$$I_D = -qA \cdot D_n \frac{dn}{dx} = qA \cdot D_n \cdot \frac{(n_{source} - n_{drain})}{L}$$

$$I_D = I_0 \cdot (1 - \exp[-\beta V_{GS}]) \quad \beta = q/kT$$

$$\downarrow \quad I_0 = \mu \frac{w}{L} \cdot \frac{C_{ox}}{2\beta^2 \cdot \sqrt{V_{GS} - V_{FB}}} \cdot \exp[\beta \cdot (V_{GS} - V_{FB} - 2\Psi_{bulk})]$$

Gate swing S:

$$S = \frac{dV_{GS}}{d(\log I_D)} \approx \frac{\ln 10}{\beta} \cdot \left(1 + \frac{C_D}{C_{ox}} \right) \xrightarrow{\text{longchannel}} 60 \text{mV/dec}$$



Result:

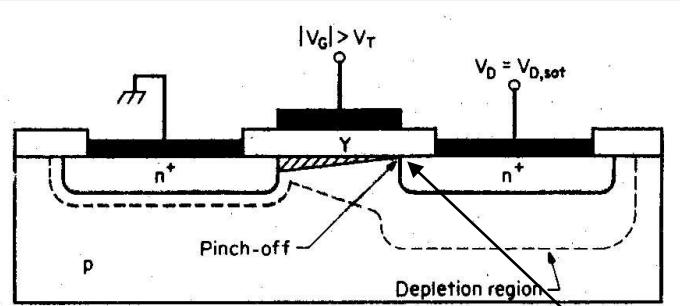
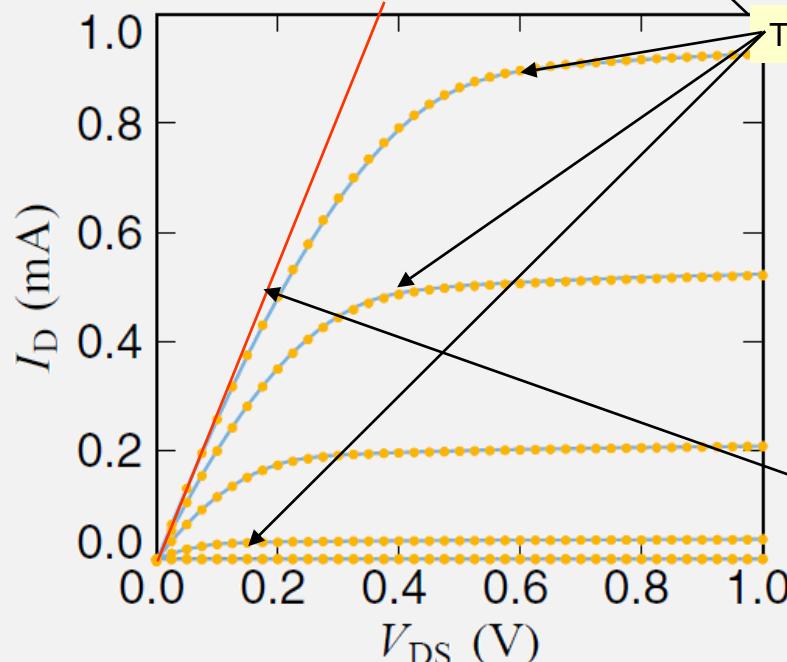
subthreshold current varies exponentially with gate voltage V_{GS}

- the best achievable value of switching (= gate voltage swing S or subthreshold slope) > 60 mV/dec
- for $V_D > 3kT/q \sim 100\text{mV}$ no dependence on V_{DS} should occur → characteristics for long channel device

Active region $V_{\text{gate}} > V_T$ and $V_{\text{DS}} > V_{\text{sat}}$

-> a channel exists from Source to Drain

-> the inversion charge density varies from Source (high density) to Drain (lower density)

► $W/L = 10/1\mu\text{m}$, $V_{GS} = 0 \dots 1\text{V}$ 

The maximum of current is reached, if the inversion charge vanishes at Drain (pinch-off)

especially, if $V_D \ll V_{\text{gate}}$ -> linear regionhigh gate voltage -> inversion channel on with high charge density
small V_D -> Channel acts as a resistor

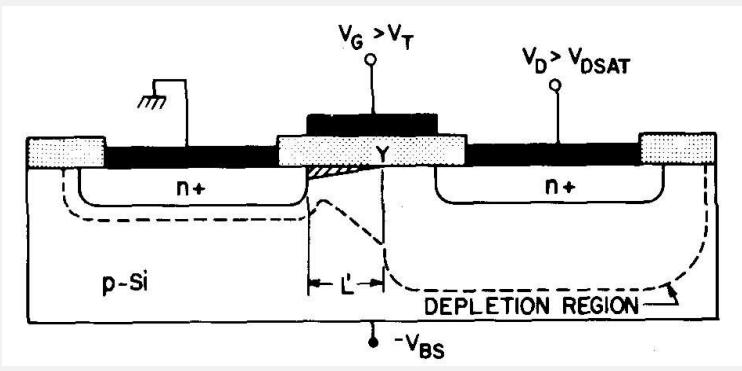
$$I_D = \mu \cdot C_{\text{ox}} \cdot \frac{W}{L} \cdot (V_{GS} - V_T) \cdot V_{DS}$$

drain conductance: $g_D = \left. \frac{\partial I_D}{\partial V_D} \right|_{V_{\text{gate}}=\text{const}}$

Typical example (see left):

$$g_D = \frac{1\text{mA}}{0.4\text{V}} = 2.5\text{mS}$$

 $R_{\text{channel}} : \sim \text{k}\Omega$

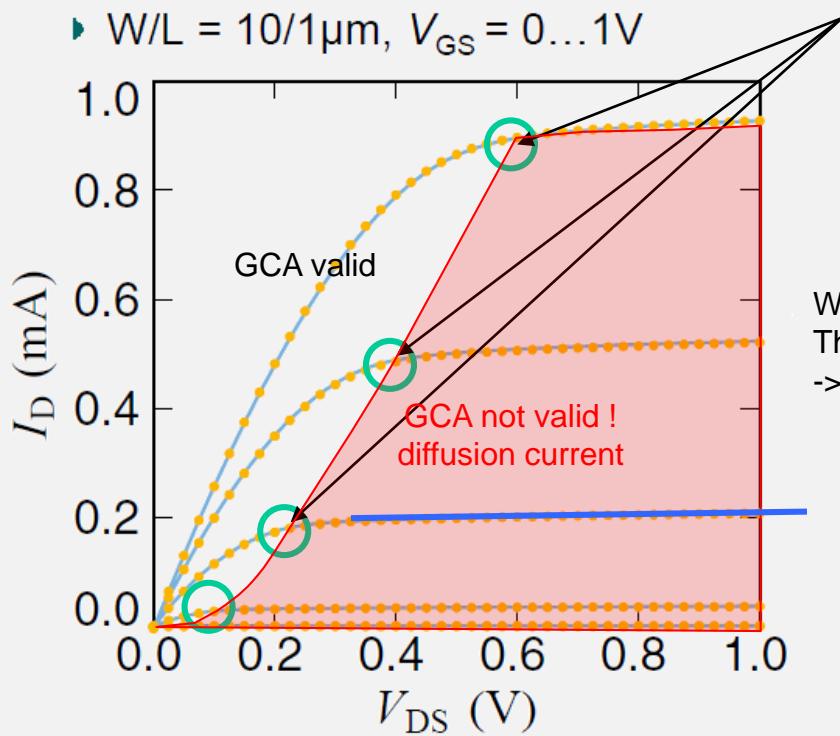


Saturation

If the Drain voltage is increased up to V_{gate} , then the voltage difference between the silicon and the gate electrode is zero and no inversion charge can be created. The channel ends (**pinch-off**).

$$I_D = \frac{1}{2} \cdot \mu \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2$$

► $W/L = 10/1\mu\text{m}$, $V_{GS} = 0 \dots 1\text{V}$



The current I_{DS} for two gate voltages is quadratic dependent on V_G

The transconductance g_m increases linear with V_G

$$g_m = \left. \frac{\partial I_D}{\partial V_{gate}} \right|_{V_{DS}=const} = \mu \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)$$

With $V_{DS} > V_{sat}$ the pinch-off point moves towards Source.
The current transport between pinch-off and Drain now is purely diffusion current
-> increasing V_D does not increase I_D , I_D remains constant.

Drain (channel) conductance: $g_D = \left. \frac{\partial I_D}{\partial V_D} \right|_{V_{gate}=const} = 0$ but not realistic !



At this point the GCA-model ends !
Modifications to the model must be done !

3.1 MOSFET intro

- Basic semiconductor physics
- Basic MOSFET working principle
- MOSFET types
- MOSFET applications
- Basic MOSFET fabrication process

3.2 MOS physics

- Work function
- Contact phenomena
 - Metal - metal
 - Metal - semiconductor
 - Metal - insulator - semiconductor (MIS)
- Working regimes of the MOS structure

3.3 Ideal long-channel MOSFET

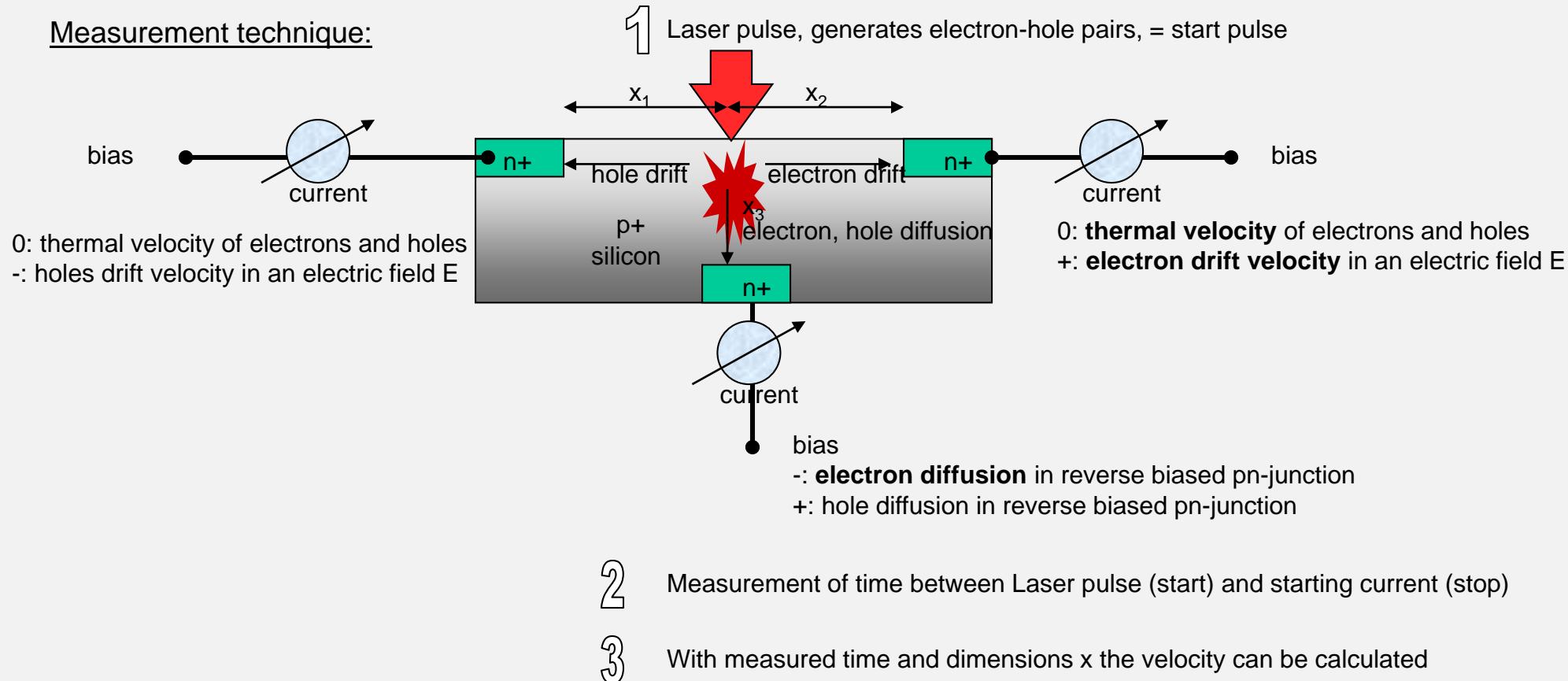
- Charge models (Q-constant, Gradual-Channel)
- MOSFET working regimes

3.4 Realistic behavior

- Basics of charge carrier transport
- Modification in charge carrier transport
- Break-through phenomena
- Parasitics

High Electric Fields induce:

- Modification in carrier transport
 - Mobility degradation due to Gate field
 - Velocity saturation due to Source-drain field
 - Channel length modulation in saturation
- Break-through effects
 - 1 Hot electrons -> Impact ionization
 - Avalanche multiplication (Drain break-through)
 - Parasitic Bipolar Transistor
 - Gate damage
 - 2 Gate break-through
 - 3 Punch-through
- Parasitics
 - Source-Drain resistors
 - Capacitors

Measurement technique:

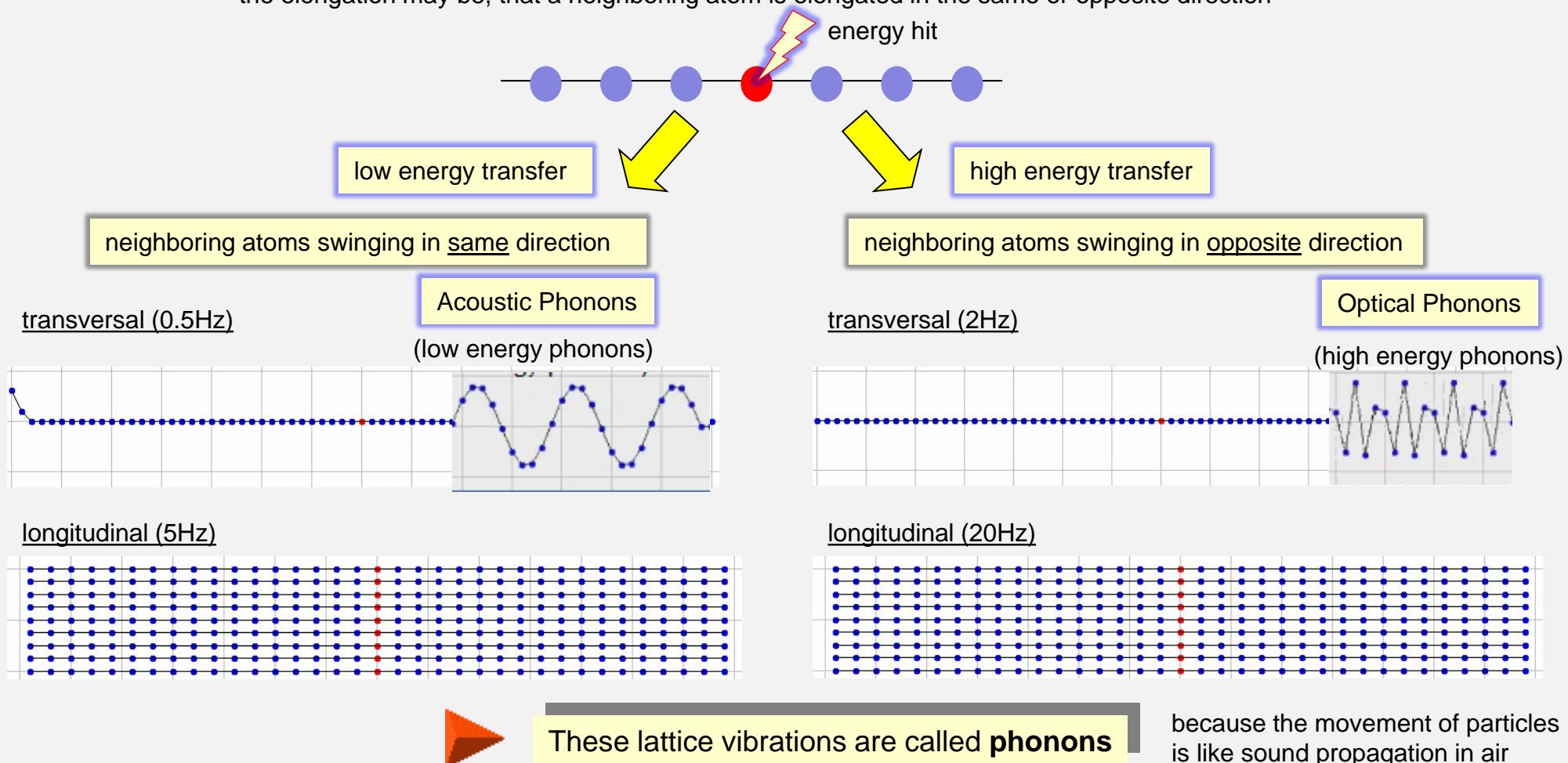
3 types of classical motion are possible:

- thermal velocity without any external fields
- drift velocity within an electric field
- diffusion velocity in a carrier density gradient

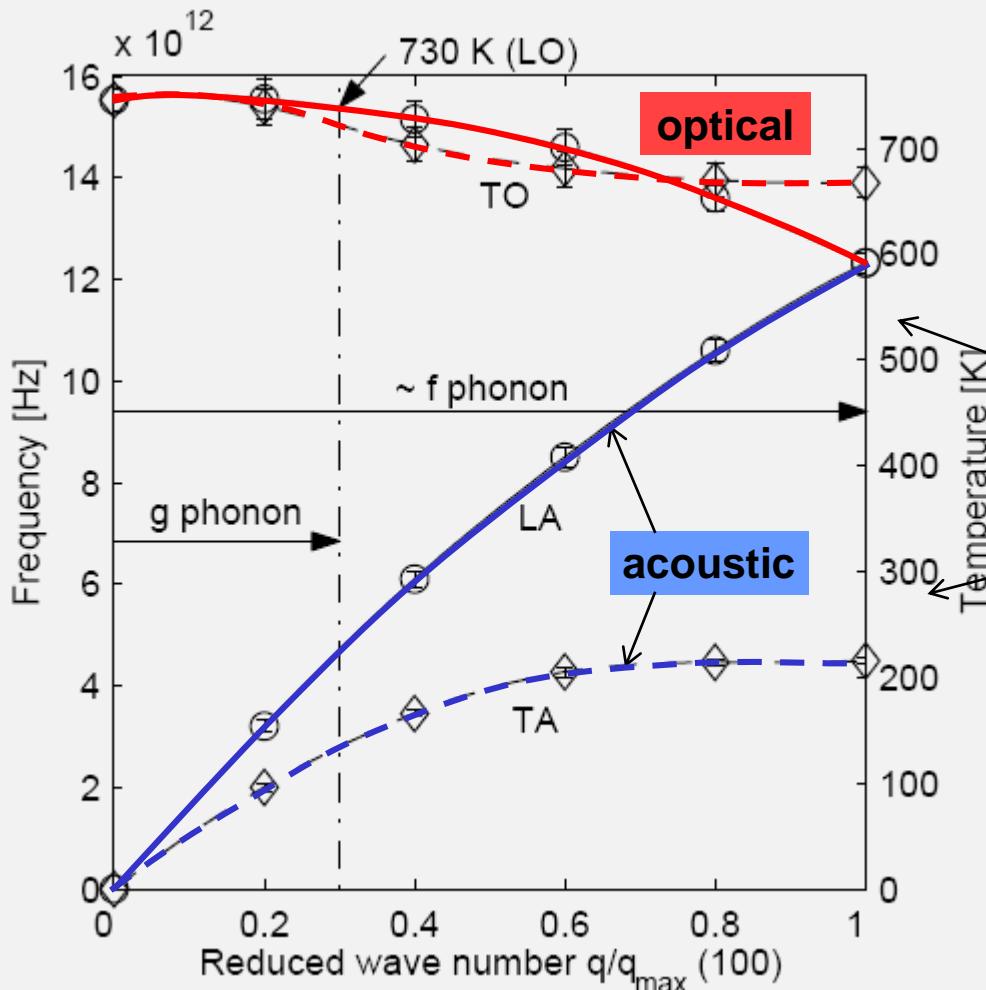
1

The environment: Phonons in the crystal

- in a perfect crystal all atoms occupy perfectly periodic lattice sites (at temperature $T = 0 \text{ K}$)
- with thermal excitement (e.g. incoming radiation) photons may kick atoms out of their position (physically speaking: photon scattering by transfer of energy and momentum)
- the elongation may be, that a neighboring atom is elongated in the same or opposite direction



Incoming radiation (or mechanical vibrations due to vibrating environment atoms (gas, solid)) transfer energy into the crystal until the energy within the crystal is the same as of the environment (= heating up until thermal equilibrium).



Lattice vibrations, where neighboring atoms are moving in the same direction are called **acoustical phonons** (due to the analogy to sound propagation).

Lattice vibrations, where neighbouring atoms are moving in the opposite direction are called **optical phonons**.

To generate optical phonons a temperature $T > 600\text{K}$ is needed
($E \sim kT > 60\text{ meV}$)

In silicon at room temperature (300K) only acoustic phonons exist
($E \sim kT \sim 26\text{ meV}$)

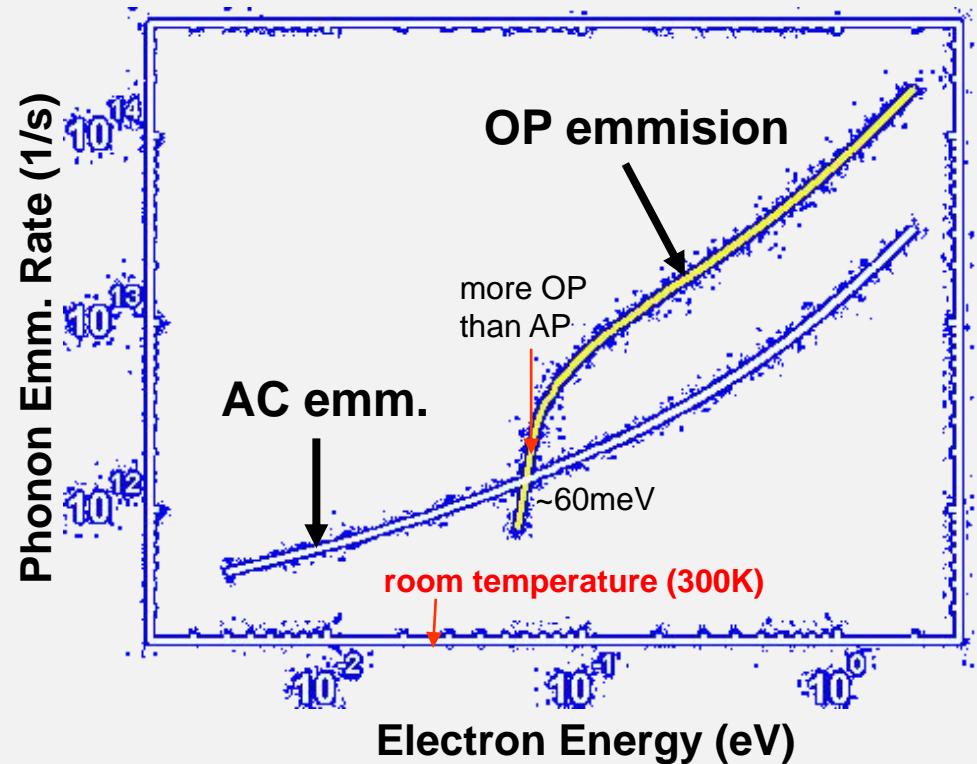
The elongations of the atoms may occure vertical to the lattice chain (**transversal**) or within the lattice chain (**longitudinal**)

2

Carrier movement in thermal equilibrium

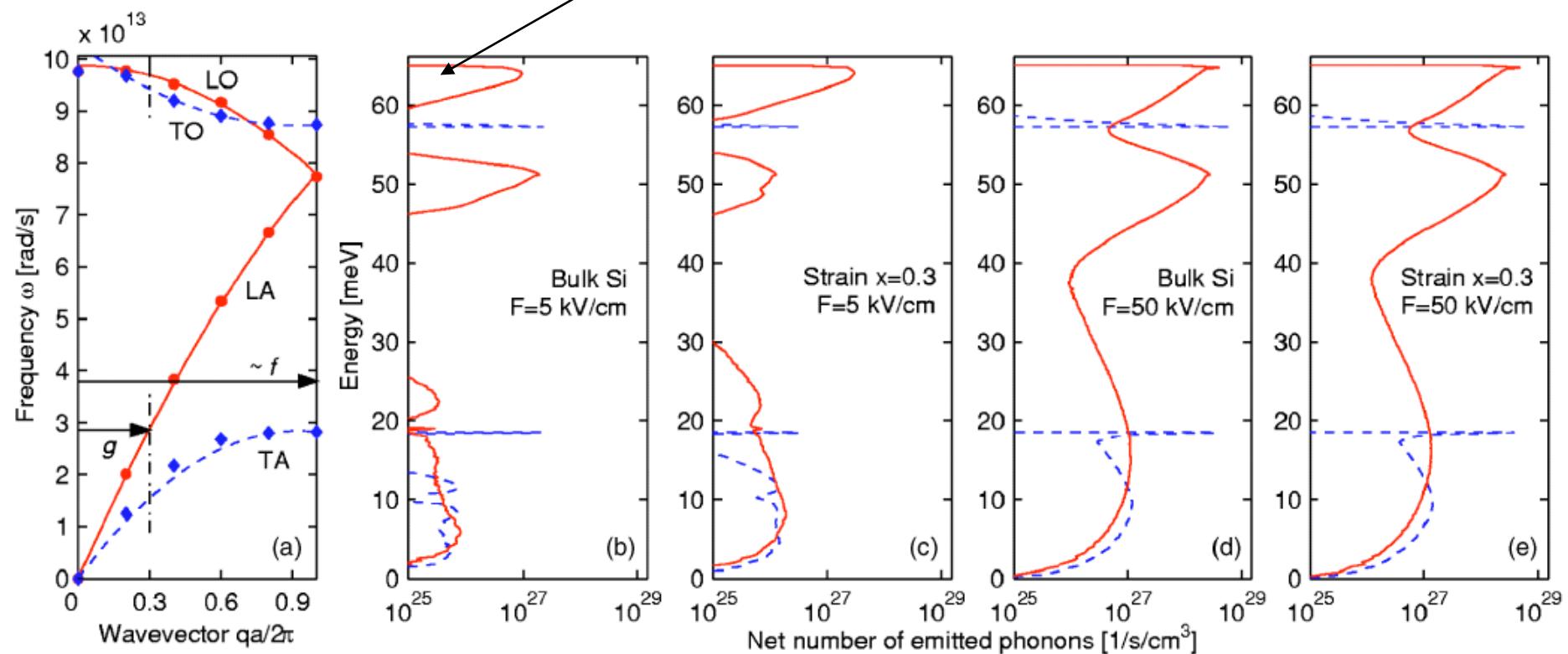
- the phonons also transfer energy and momentum to the electrons, the electrons get in motion.
On the other side the electrons transfer energy and momentum back to the phonons.
Due to this scattering events the electrons get in thermal equilibrium with the lattice energy.
- the energy spectrum of the electrons has a Maxwellian distribution.
The **main kinetic energy** $E = 1/2 mv^2$ of the electrons will be equal to the thermal energy $E = kT$.
With $E(300K) = 0.026 \text{ eV}$ the main thermal velocity of electrons can be calculated to $\sim 1 \times 10^7 \text{ cm/sec}$.

- Almost no electrons have the energy to generate optical phonons at room-temperature.
- Electrons with energies above kT (by accident) are called "hot electrons"



Result:

High energy electrons can lose energy with phonons
only in small portions of about 63 meV (highest value)



APPLIED PHYSICS LETTERS 86, 082101 (2005)

Monte Carlo simulation of Joule heating in bulk and strained silicon

Eric Pop^{a)} and Robert W. Dutton*Department of Electrical Engineering, Stanford University, Stanford, California 94305*

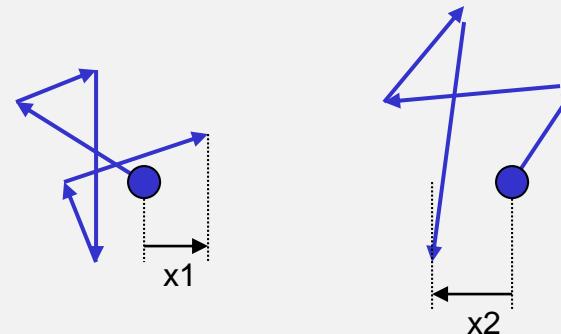
Kenneth E. Goodson

Department of Mechanical Engineering, Stanford University, Stanford, California 94305

3

Movement in an external electric field: **drift**

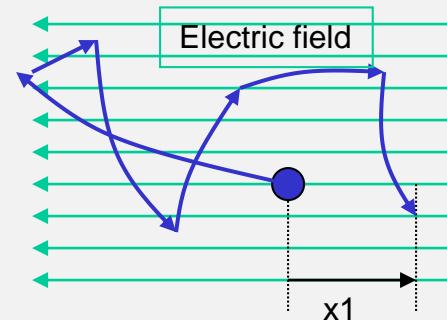
Thermal motion:
(without electric field)



In all directions:

$$\sum_i x_i = 0$$

Drift motion:
(with electric field)



In field direction:

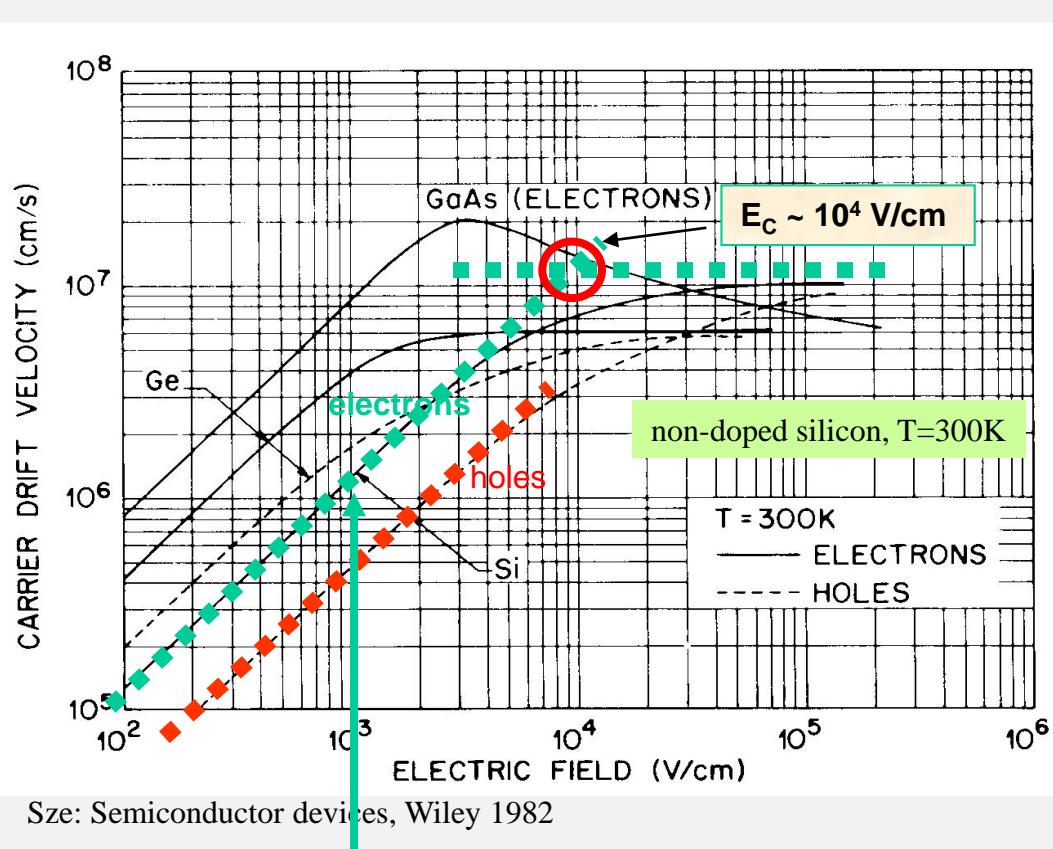
$$\sum_i x_i \neq 0$$



The velocity gain in direction of an external electric field is called drift velocity

In electric fields the charge carriers are accelerated and stopped by scattering in non-periodic events.

In macroscopic view a constant, intermediate **drift velocity** is achieved, which depends on the electric field:



Sze: Semiconductor devices, Wiley 1982

Approximation:

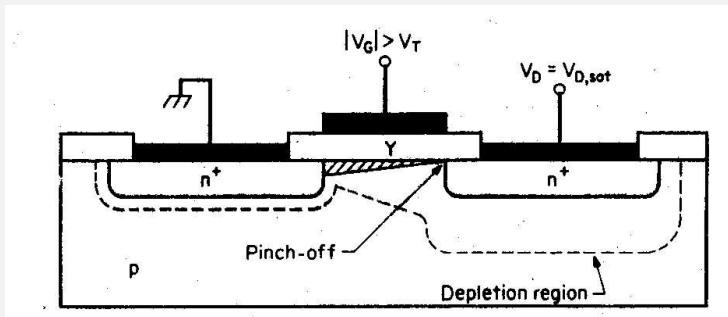
$$\mu = \frac{\Delta v}{\Delta E} \approx \frac{10^6 \text{ cm/sec}}{10^3 \text{ V/cm}} \approx 1000 \frac{\text{cm}^2}{\text{V sec}}$$

In bulk silicon the electron **low-field mobility** is $\sim 1500 \text{ cm}^2/(\text{Vsec})$

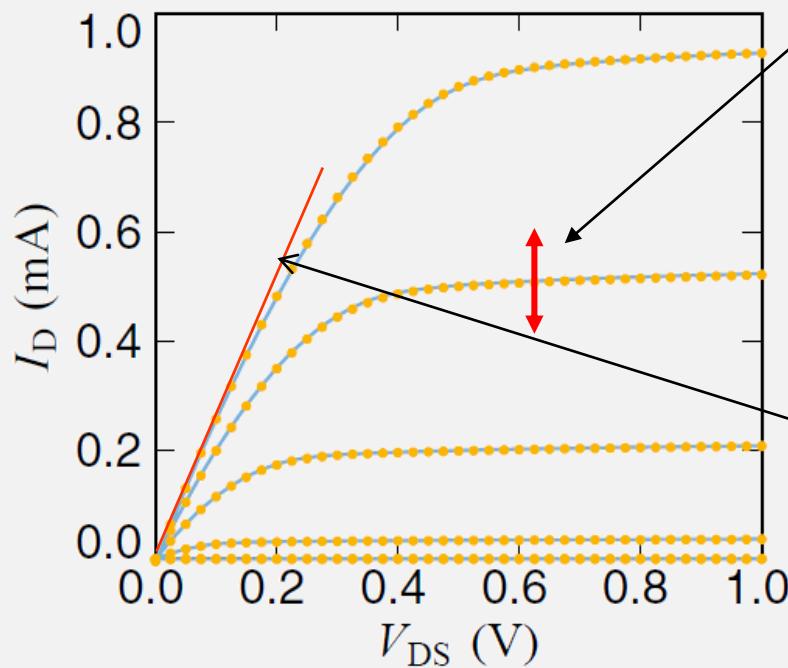
$$v = \mu * E$$

The proportionality constant is called **mobility μ**

- ▶ In semiconductor materials the mobility is constant up to $\sim 10^3 \text{ V/cm}$. For higher fields the velocity saturates and therefore mobility decreases
- ▶ The mobility for electrons is higher as for holes due to lighter mass (\sim factor 3).
- ▶ The mobility decreases ($\mu \rightarrow 0$) for high electric fields



► W/L = 10/1μm, $V_{GS} = 0 \dots 1V$



Remember:

Mobility μ makes the current gain of a MOSFET higher or lower

$$I_D = \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot \left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right]$$

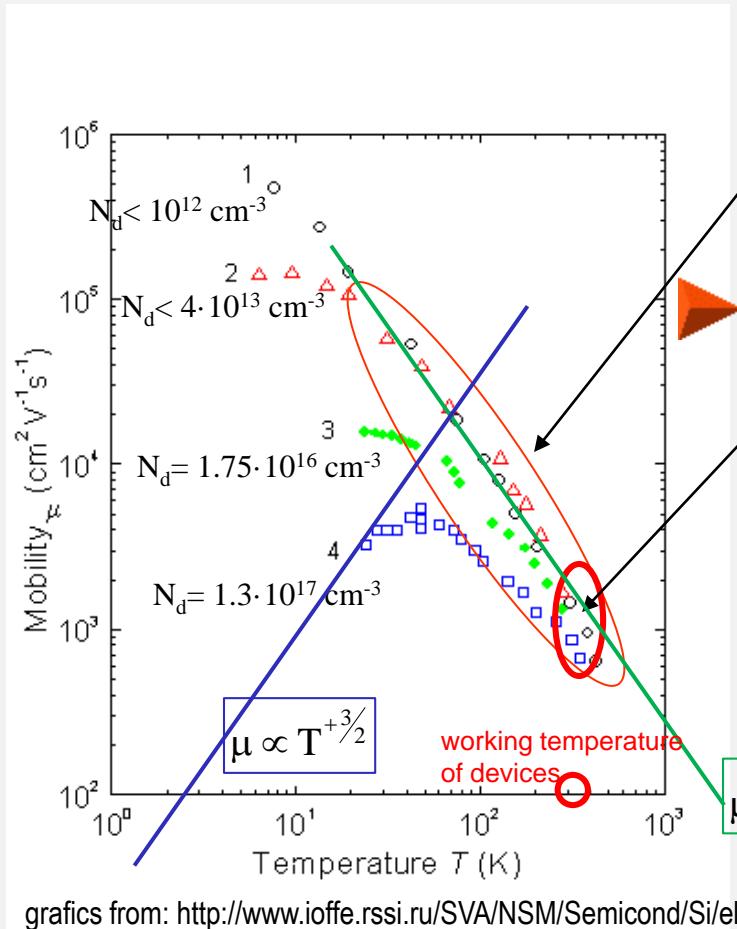
Therefore the following MOSFET can be faster charged -> speed

Mobility μ makes the conductance of a MOSFET higher or lower

$$I_D = \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot (V_{GS} - V_T) \cdot V_{DS}$$

drain conductance 1/R: $g_D = \frac{\partial I_D}{\partial V_D} \Big|_{V_{gate}=const}$

typically: $\sim \text{mS}$
or $R_{ch} \sim \text{k}\Omega$



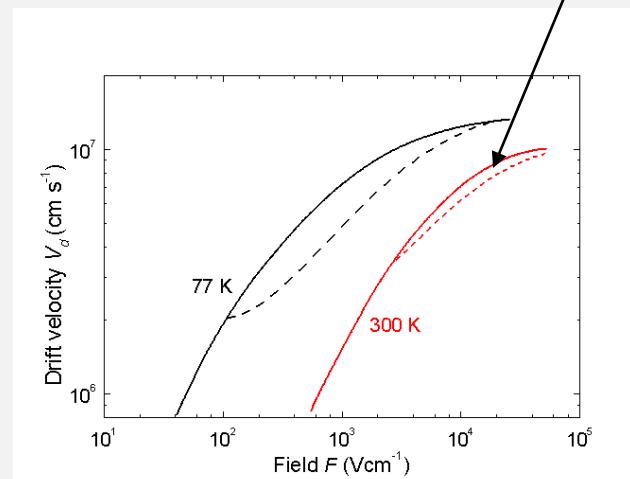
The low-field mobility in bulk silicon is strongly **dependent on temperature** (due to temperature dependent existence of phonons)

The low-field mobility in bulk silicon is **dependent on doping** (at room temperature) (due to Coulomb-scattering at lattice imperfections)

The low-field mobility in bulk silicon is little **dependent on crystal orientation**

graphics from: <http://www.ioffe.rssi.ru/SVA/NSM/Semicond/Si/electric.html>

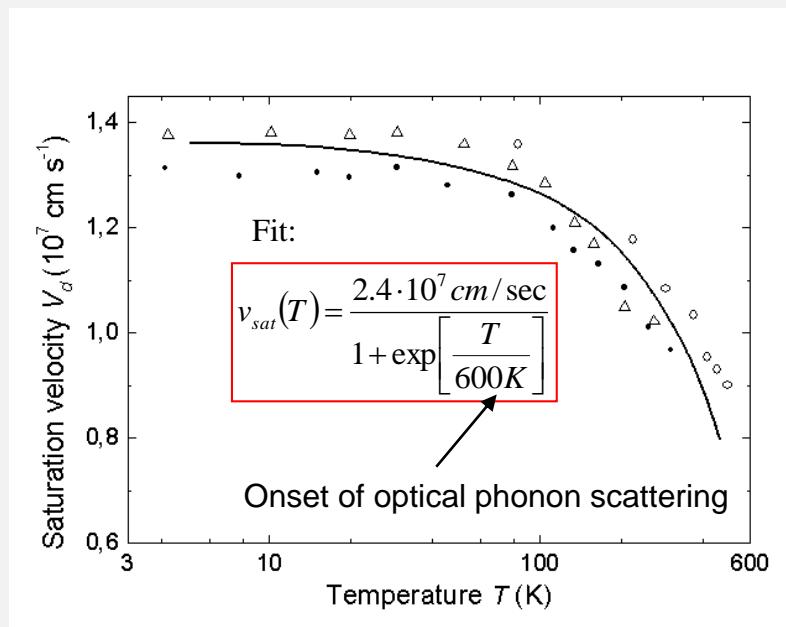
1. High purity Si ($N_d < 10^{12} \text{ cm}^{-3}$); time-of-flight technique (Canali *et al.* [1973])
2. High purity Si ($N_d < 4 \cdot 10^{13} \text{ cm}^{-3}$): photo-Hall effect (Norton *et al.* [1973])
3. $N_d = 1.75 \cdot 10^{16} \text{ cm}^{-3}$; $N_a = 1.48 \cdot 10^{15} \text{ cm}^{-3}$; Hall effect (Morin and Maita [1954]).
4. $N_d = 1.3 \cdot 10^{17} \text{ cm}^{-3}$; $N_a = 2.2 \cdot 10^{15} \text{ cm}^{-3}$; Hall effect (Morin and Maita [1954]).



Field dependences of the electron drift velocity.

Solid lines: $F \parallel (111)$, dashed lines: $F \parallel (100)$.
(Jacoboni *et al.* [1977])

Temperature dependence of the saturation electron drift velocity



graphics from: <http://www.ioffe.rssi.ru/SVA/NSM/Semicond/Si/electric.html>

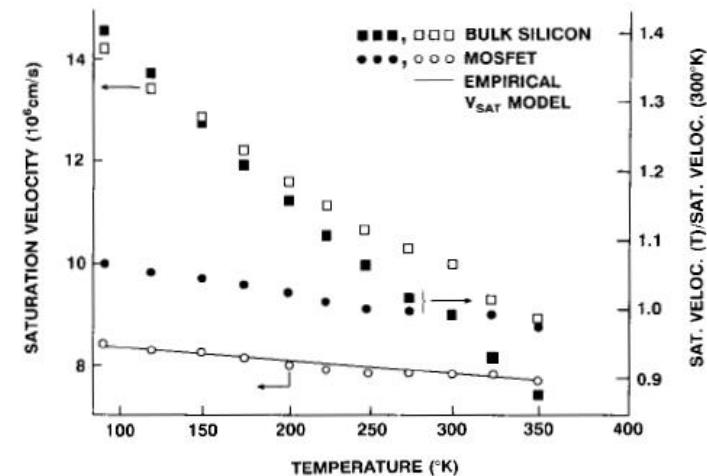
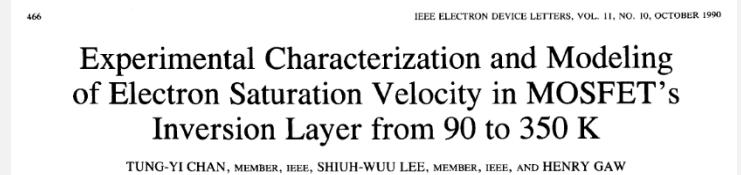
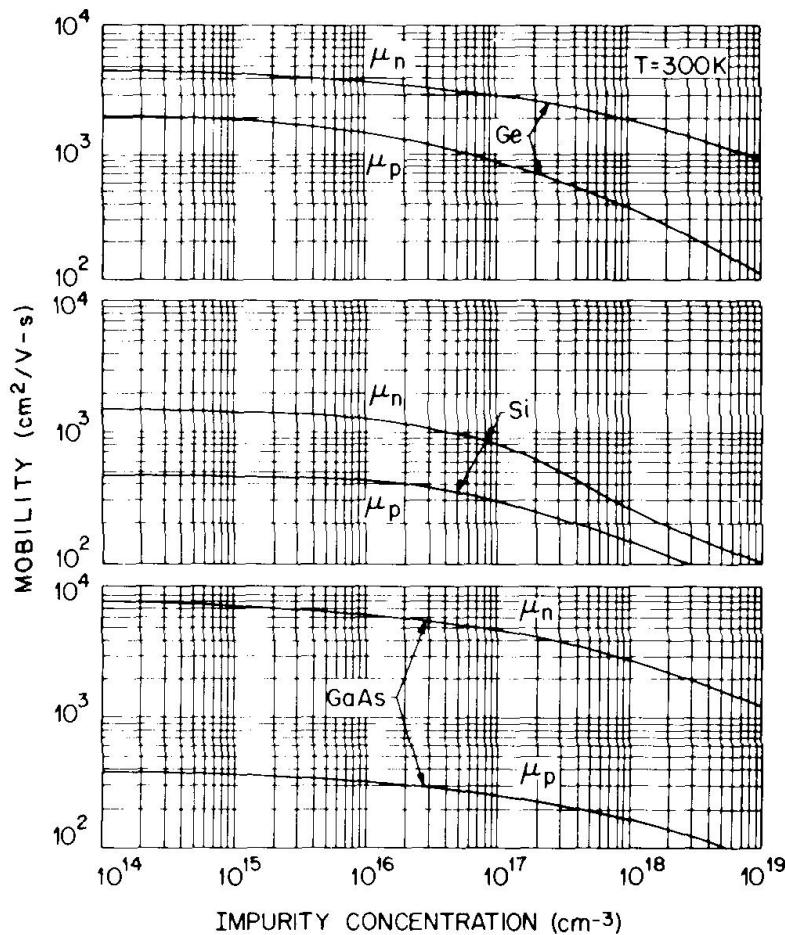


Fig. 2. Comparing electron saturation velocity and its temperature dependence in the MOSFET inversion layer and bulk silicon, and the empirical $v_{\text{sat}}(T)$ relation for CAD application.

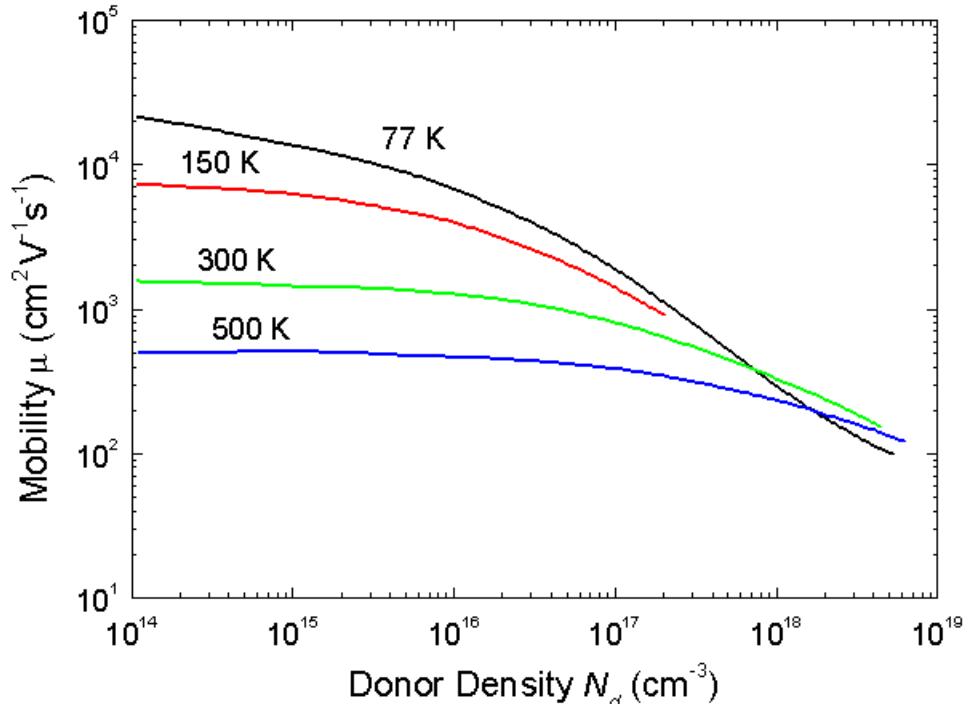
III. EXPERIMENTAL RESULTS AND DISCUSSIONS

The $g_m^{(i)}$ method has been applied to n-channel devices of a $0.25\text{-}\mu\text{m}$ CMOS technology, and the v_{sat} of electrons from 90 to 350 K is plotted in Fig. 2. The increase in v_{sat} with decreasing temperature suggests that, as in bulk silicon, carrier velocity in the inversion layer is also limited by some strongly coupled phonon scattering. For comparison, $v_{\text{sat-blk}}$ is also plotted [17]. Note that the temperature dependence of



Sze: Semiconductor devices, Wiley 1982

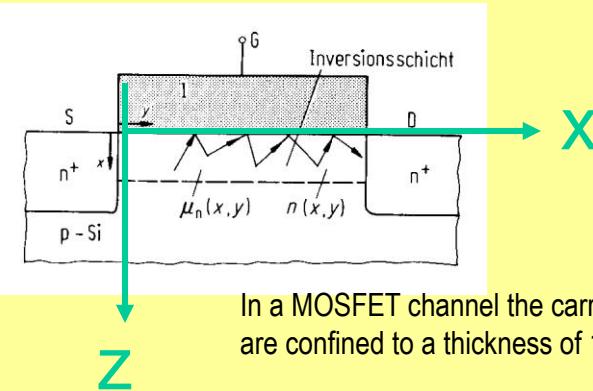
Electron drift mobility versus donor density at different temperatures
(Li and Thumber [1977]).



graphics from: <http://www.ioffe.rssi.ru/SVA/NSM/Semicond/Si/electric.html>

Mobility degradation due to the vertical gate-field

The effect:



In a MOSFET channel the carriers are confined to a thickness of 1-10 nm.

The gate field E_z (normal or transversal) makes :

- the channel narrower and
- accelerates the carriers towards the interface

charges at the interface and surface roughness increase scattering

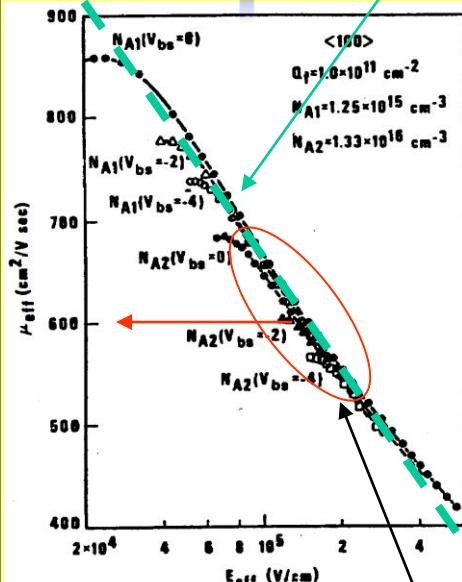
The effective normal field $E_{\perp eff}$ within the inversion layer :

In a MOSFET channel the mobility (μ_e : ~600) is reduced by a factor 2-3 compared to the bulk mobility (μ_e : ~1600)

The result:

1

Inversion layer mobility is a unique function of vertical gate field



Experimental Fit:

2

Inversion layer mobility decreases by temperature

- > safety for Power-MOSFETs
- > bad for ULSI HF-processors

$$\mu_{\perp} = \frac{\mu_{bulk}}{1 + \Theta \cdot E_{\perp eff}}$$

frequently used in CAD, although more complex formulas exist !

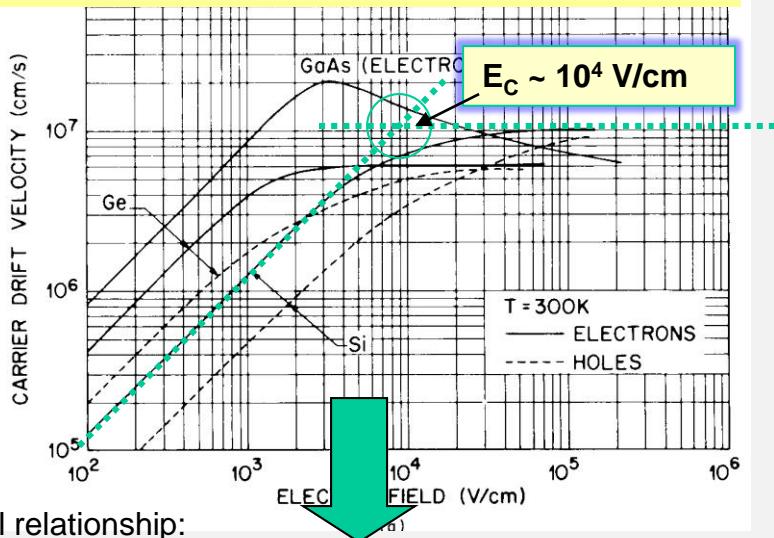
low-field channel mobility μ_0 : ~ 600 cm²/(Vsec) for electrons, ~ 250 cm²/(Vsec) for holes

mobility degradation coefficient Θ : ~ (0.03 – 0.09) cm / V

$$E_{\perp eff} \approx \frac{(Q_{bulk} + \eta Q_{chan})}{\epsilon_0 \epsilon_{Si}} \approx \frac{C_{ox}}{\epsilon_0 \epsilon_{Si}} \cdot [\eta(V_{GS} - V_T) + \gamma \cdot \sqrt{2\Psi_{bulk}}]$$

$\eta \sim 0.5$ for electrons

For a constant vertical gate field $E_{\perp eff}$ the drift velocity saturates for high lateral fields E_x and hence the mobility μ_{\perp} degrades.



Empirical relationship:

$$v_{drift} = \frac{\mu_{\perp eff} \cdot E_x}{\sqrt{1 + \left(\frac{E_x}{E_{crit}}\right)^2}}$$

not suitable for CAD

$$v_{drift} = \frac{\mu_{\perp eff}}{1 + \left(\frac{E_x}{E_{crit}}\right)} \cdot E_x$$

simplified for CAD

new mobility μ_x

$$\mu_x = \frac{\mu_{\perp eff}}{1 + \frac{V_{DS}}{L \cdot E_{crit}}}$$

The velocity saturation can easily involved in the solution for the drain current, since both, v_{drift} and I_{DS} , are calculated by integration of the position dependent channel-bulk voltage $V_{CB}(x)$ and $E(x) = d(V_{CB}/dx)$.

1

The drain current with constant μ (or $\mu_{\perp eff}$) is reduced by:

$$1 + \frac{V_{DS}}{L \cdot E_{crit}}$$

2

In addition the pinch-off voltage $V_{DS,sat}$ is reduced

$$V_{DSsat} = LE_c \cdot \left(\sqrt{1 + \frac{2(V_{GS} - V_{th})}{LE_c}} - 1 \right)$$

3

Especially in the saturation region ($V_{DS} > V_{DSsat}$)

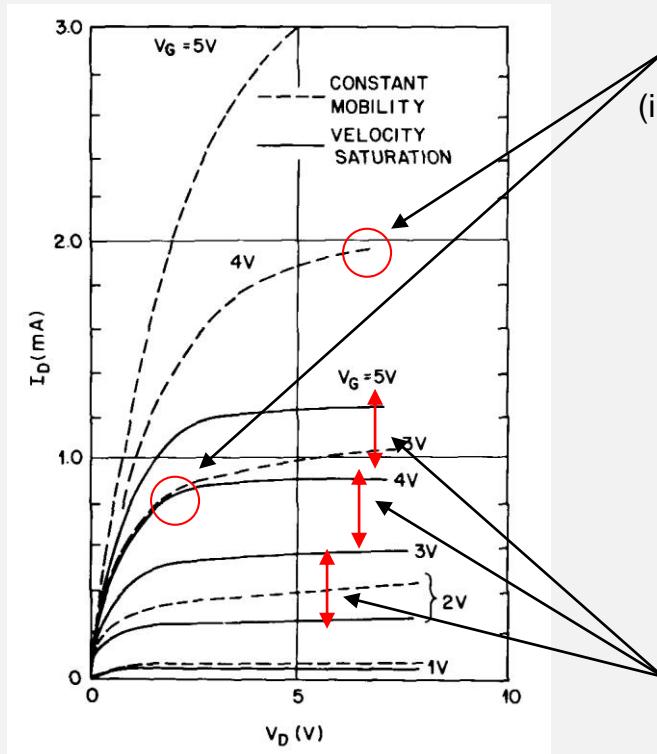
$$I_D = C_{ox} \cdot w \cdot (V_{GS} - V_{th}) \cdot v_{max}$$

the drain current is linear dependent on V_{GS} , not quadratic !

From the measured transconductance :

the value of v_{max} can be calculated.

$$g_m = \left. \frac{\partial I_{Dsat}}{\partial V_G} \right|_{V_{DS}} = w \cdot C_{ox} \cdot v_{max}$$



Reduction of the pinch-off voltage V_{DSSat} :
(indicated here for 4V ---- $\mu = \text{const.}$ and 4V $\mu \neq \text{const.}$)

$$V_{DSSat} = LE_c \cdot \left(\sqrt{1 + \frac{2(V_{GS} - V_{th})}{LE_c}} - 1 \right)$$

Especially in the saturation region ($V_{DS} > V_{DSSat}$)

$$I_D = \tilde{C}_{ox} \cdot w \cdot (V_{GS} - V_{th}) \cdot v_{\max}$$

the drain current is linear dependent on V_{GS} , not quadratic !

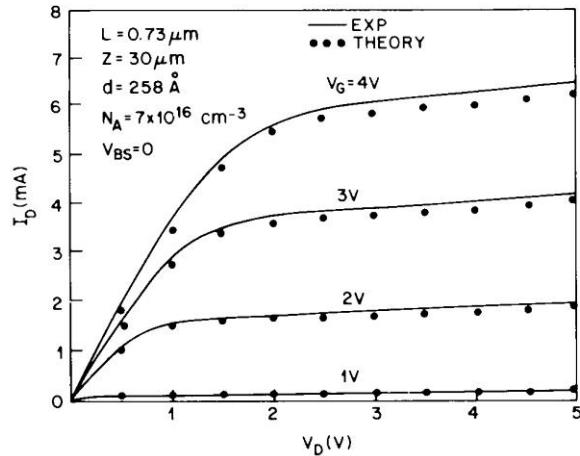
(This results in aequidistant separation of I_D (V_G)-lines)

Overall result concerning mobility μ :

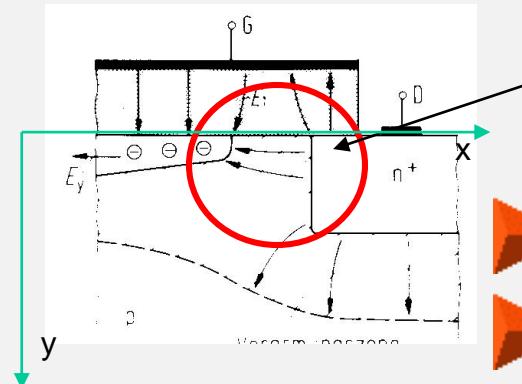
- from current equation: mobility μ determines drain current I_D (= drive current for charging next gate)
- but mobility is dependent on scattering mechanism: doping, temperature, interface "roughness"
- and decreased within external electric fields: vertical Gate field and lateral Drain field

The IV-characteristics is determined by effects on mobility μ , which must be taken into account, when simulating devices or circuits

Channel length modulation in saturation



In reality the drain current behind saturation ($V_{SD} > V_{DSSat}$) is not constant as proposed by the Gradual-Channel Approximation.



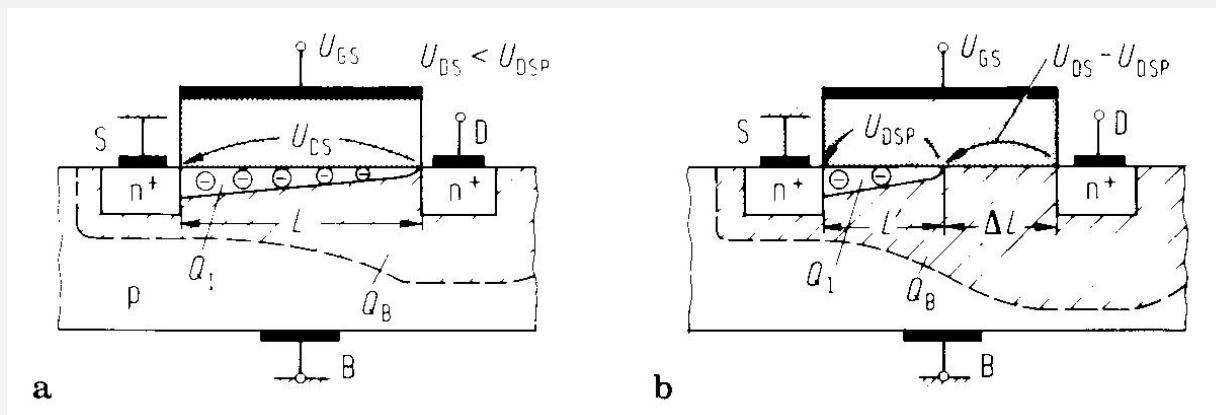
The current increases, because the GCA ($E_y \gg E_x$) is not valid.

2-dim shape of electric field

The electronic channel length < geometrical length

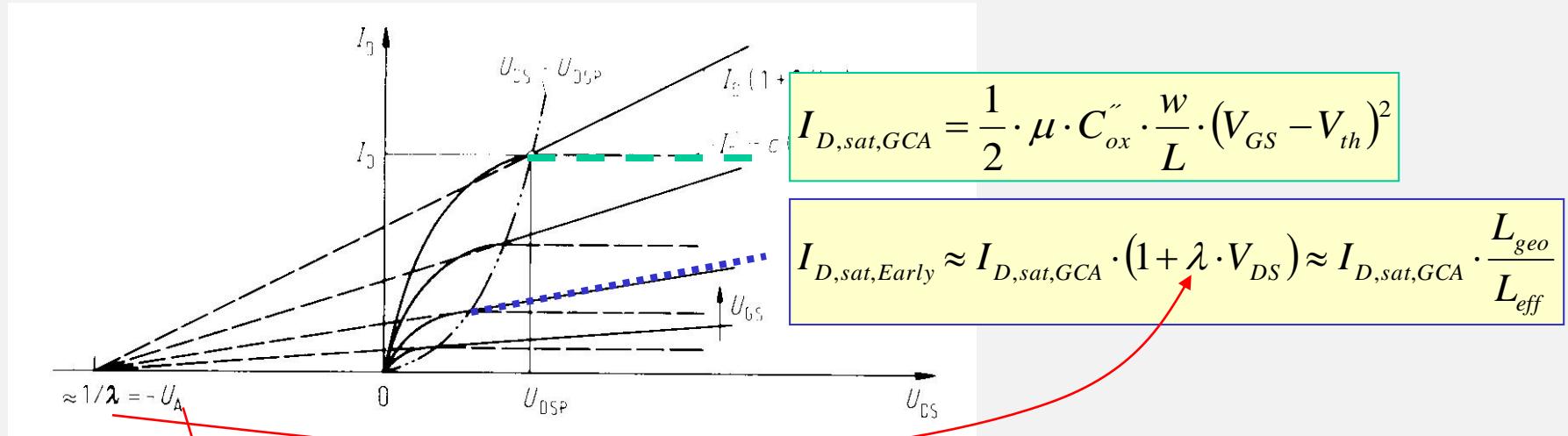
$$I_D = \mu \cdot C_{ox} \cdot \frac{W}{L} \left[(V_{GS} - V_{th})V_{DS} - \frac{V_{DS}^2}{2} \right]$$

shorter channel -> higher current



Practical Model:

In praxis it is observed that the extrapolations of all I_D lines cut the V_{DS} -axis in the same point. This point is called **Early-voltage**.



From the experimental I_D the Early-voltage V_{Early} , the Early-factor λ and the channel shortage $\Delta L = L_{geo} - L_{eff}$ can be determined.

Typical values: $V_{Early} = B \cdot L_{geo} \cdot \sqrt{N_{Dop}}$ with $B \approx 0.1 \dots 0.5 \frac{V}{\sqrt{\mu m}}$

Analytical, 1d / pseudo-2d models:

The drain Voltage V_{DS} is divided in the pinch-off voltage V_{DSsat} over the effective channel length L_{eff} and a remaining overshoot voltage $V_{over} = V_{DS} - V_{DSsat}$. From continuity conditions for potential and electric field at V_{DSsat} 2 equations with 2 parameters V_{DSsat} and L_{eff} can be deduced and solved.

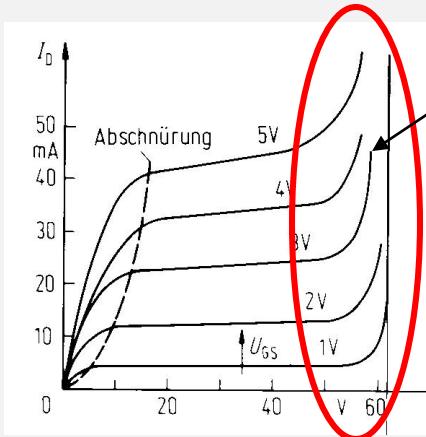
Numerical 2-d simulations:

With computers the Poisson-equation and the semiconductor equations can be solved in 2d. The result should be ok but can not be involved in circuit simulation.

Fitting approximation modells: Based on results of 2d-simulation various fitting modells are developed and fitted to reality

High Electric Fields induce:

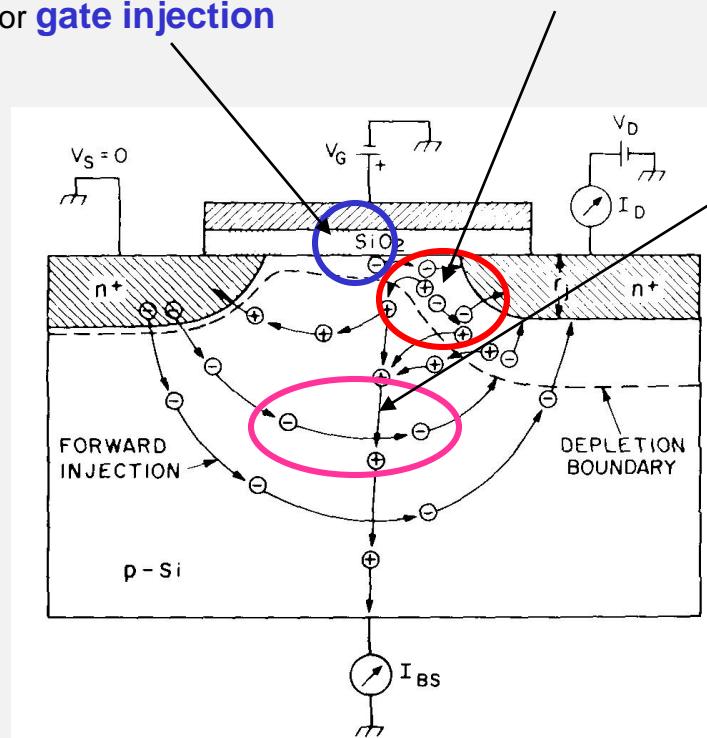
- Modification in carrier transport
 - Mobility degradation due to Gate field
 - Velocity saturation due to Source-drain field
 - Channel length modulation in saturation
- Break-through effects
 - 1 Hot electrons -> Impact ionization
 - Avalanche multiplication (Drain break-through)
 - Parasitic Bipolar Transistor
 - Gate damage
- 2 Gate break-through
- 3 Punch-through
- Parasitics
 - Source-Drain resistors
 - Capacitors



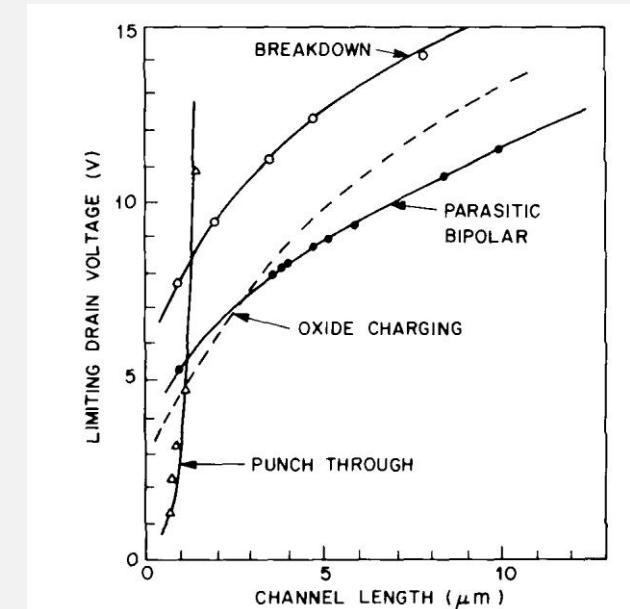
With increasing drain voltage V_{DS} high electric fields are induced, which cause several so-called **break-through effects**.

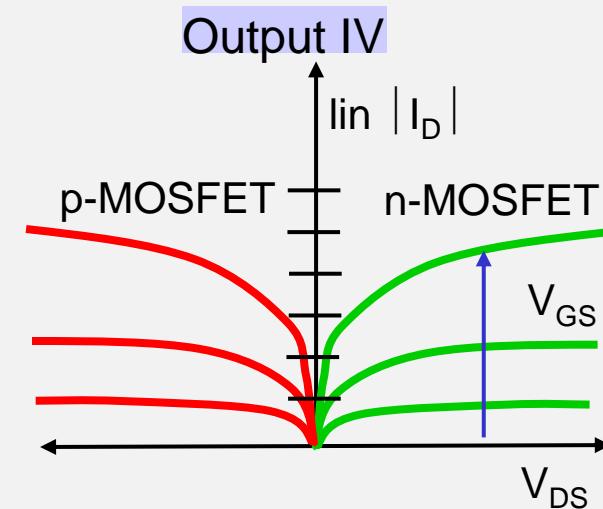
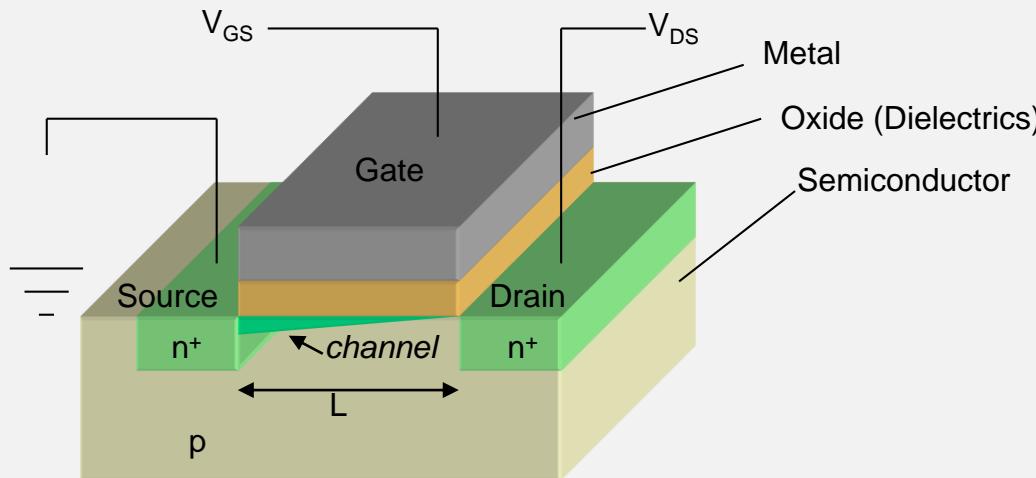
Physically these effects are caused by so-called **hot electrons**, which have much more kinetic energy as the room temperature equivalent of $kT = 0.026\text{eV}$.

These hot electrons can do **impact ionization**, trigger **parasitic bipolar action** or **gate injection**



Onset of break-through effects

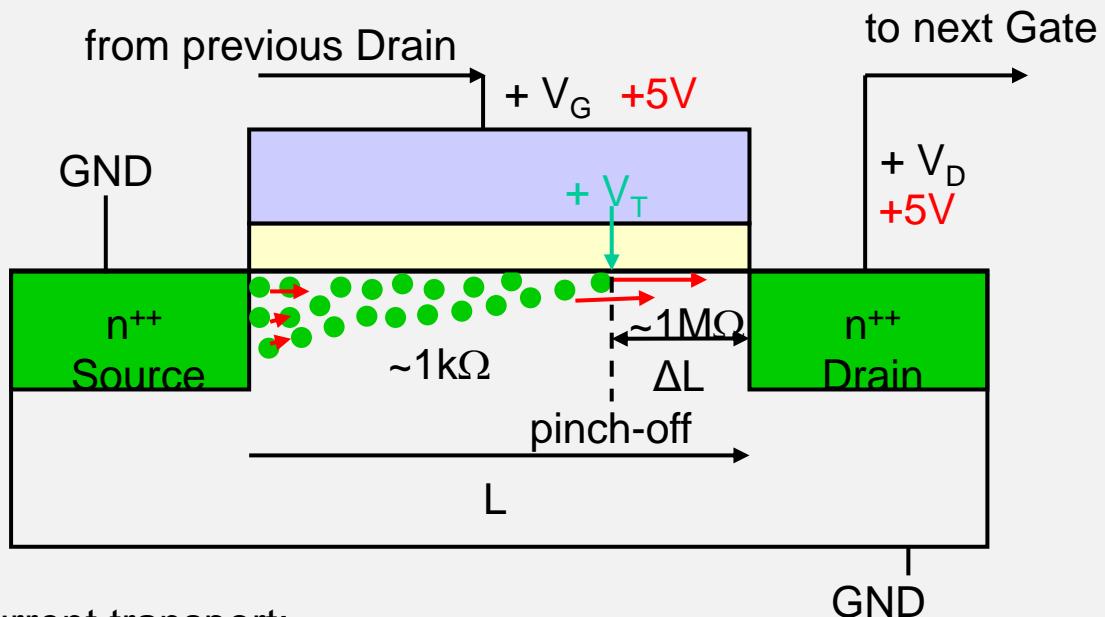




In a single MOSFET the I-V characteristics is created by any combination of Gate and Drain voltage

Analog Application: for use as amplifier (Mobil-Phones -> speech, level recognition, ...) the whole IV characteristics is used

Digital Application: for almost all applications (logic, memory) the transistor is used as a switch (on/off). Especially in logic applications the output (Drain) of the first transistor charges the Gate of the following transistor -> both voltages have the same value -> this creates special effects concerning the I-V characteristics



Current transport:

continuity equation: $\vec{j} = n \cdot e \cdot \vec{v} = \text{const}$

Many electrons start with low velocity at the Source in the channel, accelerated by drain-source voltage.
Few electrons reach the end of channel, but with higher velocity ($\rightarrow j = \text{const.}$).

At the end of channel electrons jump in ΔL by diffusion and are highly accelerated by the high electric field.

If the electric field is too high \rightarrow **impact ionization** and avalanche break-through occurs

Technological solution: increase conductivity in front of Drain by an additional resistor

\rightarrow **Lightly-Doped Drain LDD**

\rightarrow today LDD structure is used for
highly doped "extensions"

1

Impact Ionization**1 a) Bulk silicon:**

When the electric field in a semiconductor is increased above a certain value, the carriers gain enough energy so they can create electron-hole pairs by impact ionization.

The minimum energy value is the energy of the band gap (Si: 1.1 eV), but because of various scattering effects the maximum of ionization is achieved at an energy of about 3.6 eV.

The generation rate G_{imp} is proportional to the number of carriers n, p and their velocity v_n, v_p with an ionization coefficient a_n, a_p :

$$G = \alpha_n n v_n + \alpha_p p v_p$$

A physical expression for the ionization coefficient is given by:

$$\alpha(\vec{E}) = \left(\frac{q\vec{E}}{Q_{\text{ion,eff}}} \right) \cdot \exp \left[- \frac{\vec{E}_{is}}{\vec{E} \cdot \left(1 + \frac{\vec{E}}{\vec{E}_{op}} \right) + \vec{E}_{kT}} \right]$$

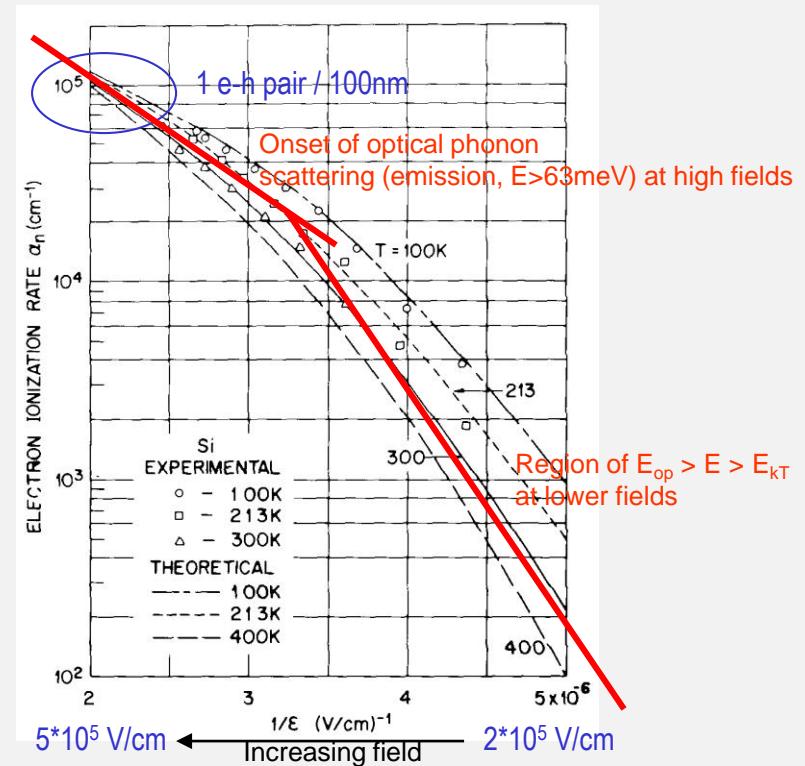
E : electrical field,

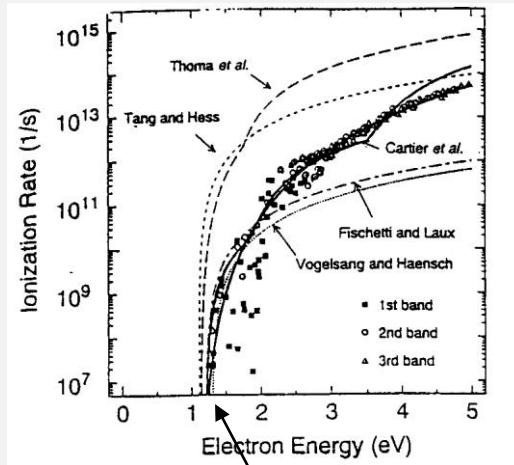
E_{is} : threshold field for ionization scattering,

E_{op} : threshold field for optical-phonon scattering,

E_{kT} : threshold field for thermal scattering,

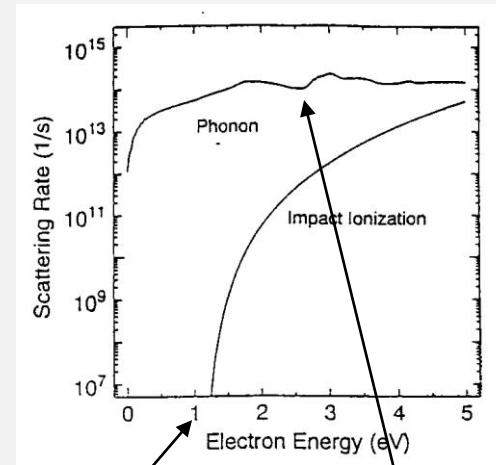
$Q_{\text{ion,eff}}$: effective ionization energy [Si: ~3.6 eV for electrons]





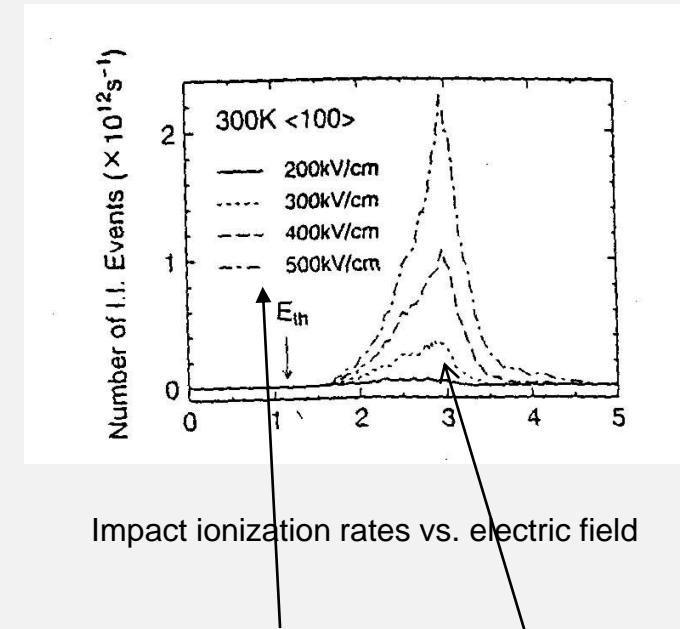
Calculated impact ionization using various models

Ionization threshold at electron energies $> E_{\text{gap}}$



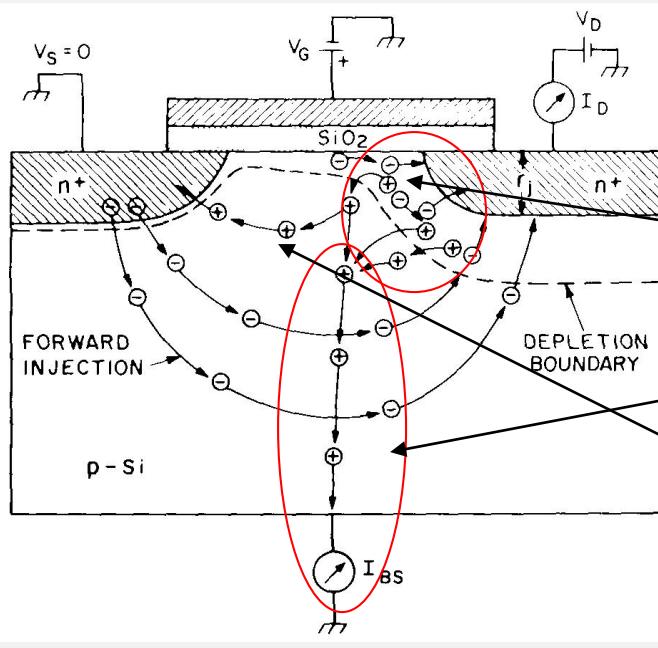
Phonon and impact scattering

Phonon scattering is dominant over ionization scattering

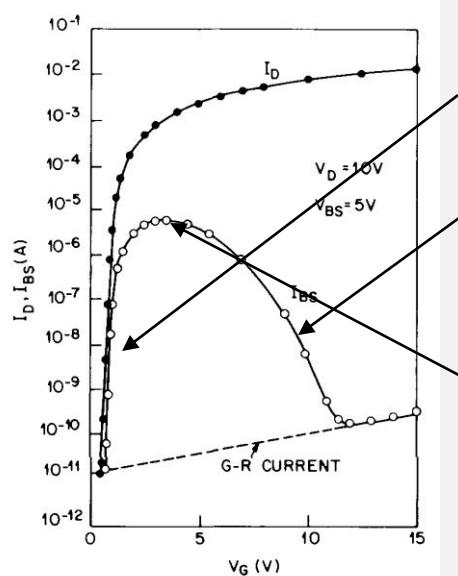
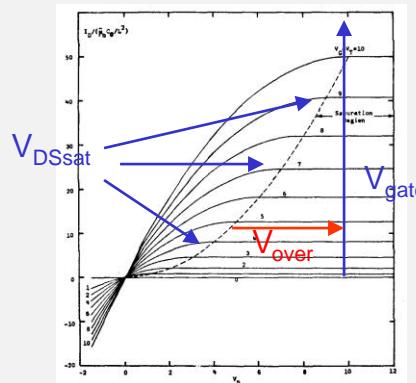


Impact ionization rates vs. electric field

Independent of electric field the optimum energy for impact ionization is around 3 eV

1b) Impact ionization within a MOSFET-channel:Physical, first-order model:

1. At sufficiently high drain voltages $V_{DS} > V_{sat}$ a pinch-off region exists
2. In the pinch-off region no inversion charge exists, the resistance of the depleted pinch-off region ΔL is higher than the channel resistance. The drain voltage V_{DS} mainly drops over the pinch-off region. The short pinch-off region is the region with the highest lateral electrical field E_x in the MOSFET.
3. Consequently, it is assumed, that the impact is homogeneously concentrated in the pinch-off region
4. With varying gate voltage the effective electric field in the pinch-off region can be varied.
5. In a long channel MOSFET the generated holes are accelerated as majority carriers towards the bulk electrode
(in a short channel MOSFET the holes may reach the Source and trigger a parasitic bipolar transistor).



For a given V_{DS} an increasing gate voltage causes an increasing channel current (see IV-characteristics) and therefore an increasing impact current (-> bulk current I_{BS}).

But with increasing gate voltage also the saturation voltage V_{DSsat} increases, resulting in a lower overshoot voltage $V_{DS} - V_{DSsat}$ and therefore the electric field in the pinch-off region decreases and therefore the impact current decreases.

Where the two factors balance a maximum in impact exists.
Without changing drain voltage V_{DS} no break-through will occur !

Quantitative Calculations

The bulk equation:

$$\alpha(\vec{E}) = \left(\frac{q\vec{E}}{Q_{imp,eff}} \right) \cdot \exp \left[- \frac{\vec{E}_{is}}{\vec{E} \cdot \left(1 + \frac{\vec{E}}{\vec{E}_{op}} \right) + \vec{E}_{kT}} \right]$$

is simplified to:
for $E_{op} > E > E_{kT}$

$$\alpha(E) = A_n \cdot \exp \left[- \frac{\vec{E}_{crit}}{\vec{E}} \right]$$

$$A_n \sim 7 * 10^5 \text{ cm}^{-1}$$

$$E_{crit} \sim 1.22 * 10^6 \text{ V/cm}$$

The generation in the pinch-off region is calculated as a substrate current I_{bulk} :

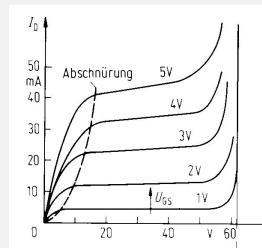
$$\frac{I_{bulk}}{I_{DS}} = \int_{L-\Delta L}^L \alpha dx = \frac{\epsilon_0 \epsilon_{Si}}{qwN_{dop}} \cdot \vec{E}_{x\max}^2 \cdot A_n \cdot \exp \left[- \frac{\vec{E}_{crit}}{\vec{E}_{x\max}} \right]$$

using:

$$\vec{E}_{x\max} = \sqrt{\frac{2qN_{dop}}{\epsilon_0 \epsilon_{Si}} \cdot (V_{DS} - V_{DSSat})}$$

the maximum field in the pinch-off region

The break-through voltage V_{DSbr} is approximately:



$$V_{DSbr} \approx \frac{V_{GS}}{1 + \frac{w_{pn}}{3d_{ox}}} + \frac{w_{pn} \cdot E_{crit}}{1 + \frac{w_{pn}}{3d_{ox}}}$$

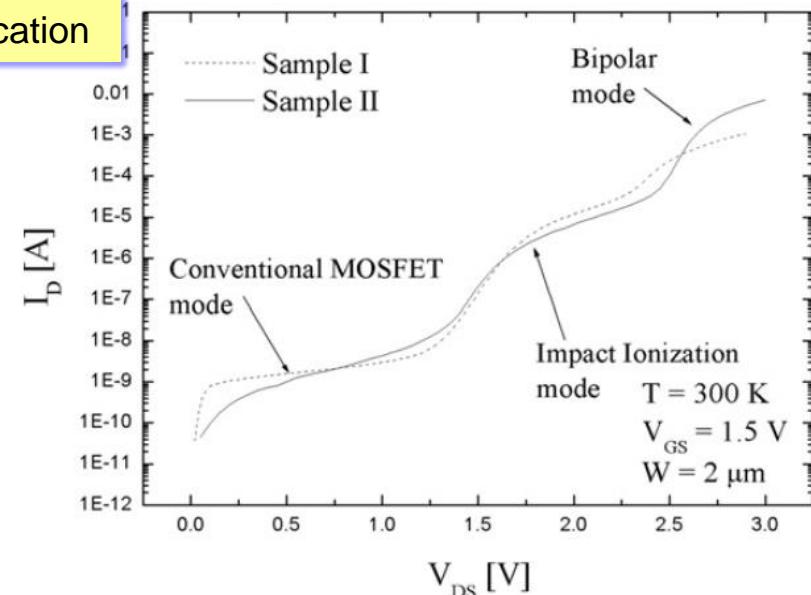
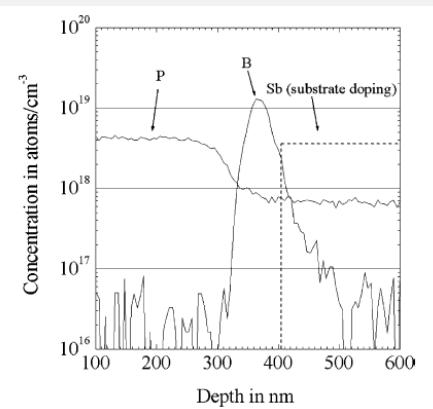
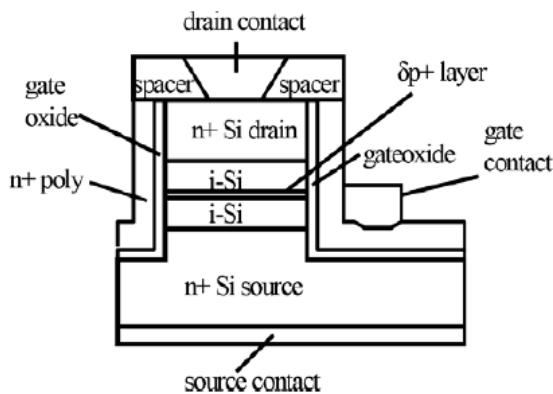
$$\begin{cases} V_{DSbr} \approx w \cdot E_{crit} \\ V_{DSbr} \approx 3d_{ox} \cdot \frac{V_{GS}}{w} \end{cases}$$

for thick gate oxides, drain-channel breakthrough

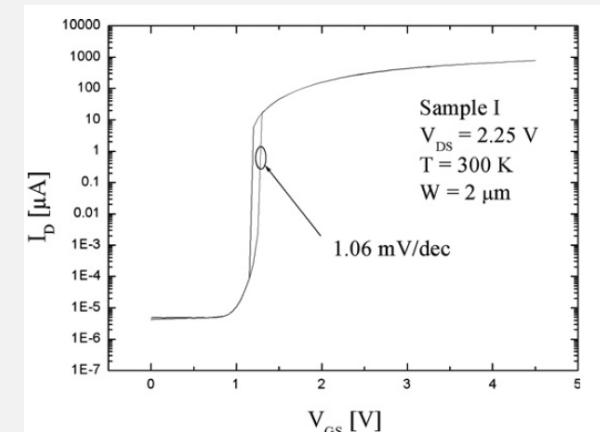
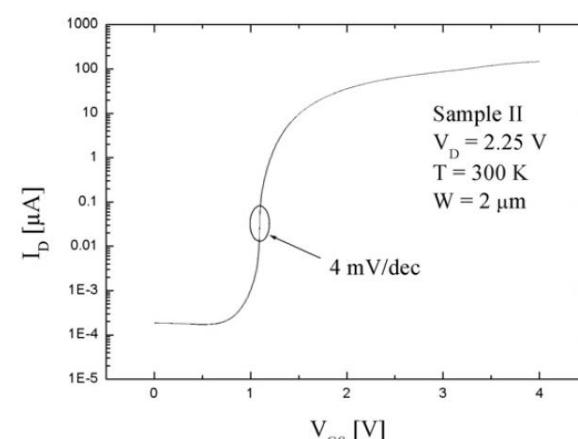
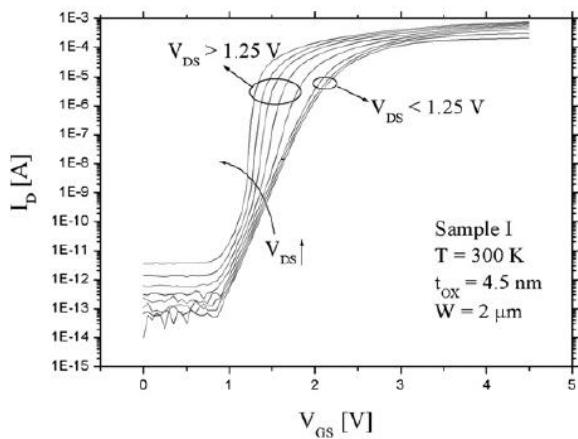
for thin gate oxides, drain-substrate breakthrough

Geometrical factors like oxide thickness or doping must be taken into account, to avoid impact, when simulating devices or circuits

Optimization of δ -doping profile enables additional bipolar amplification



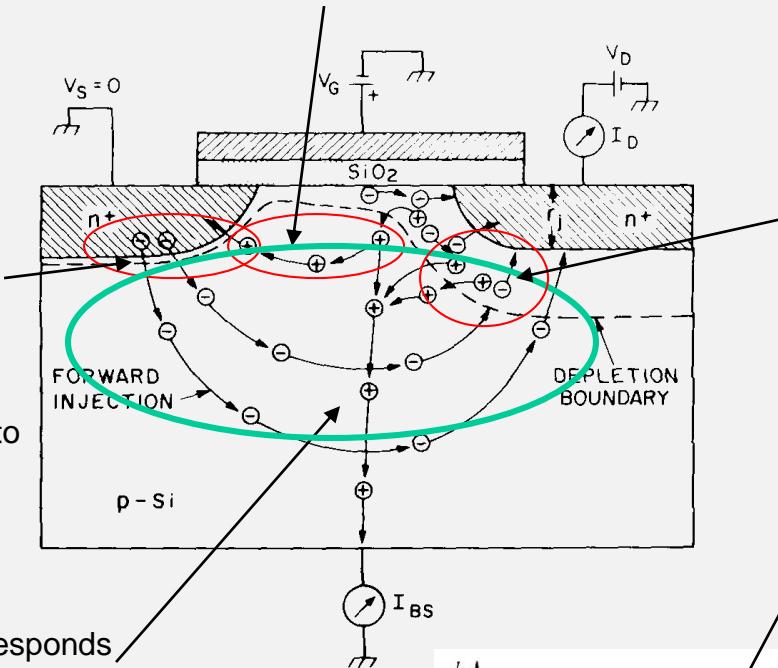
Abelein et al., Institute of Physics, UniBw Munich
Nano and Giga Challenges in Electronics and Photonics, NGC 2007



with bipolar action steep slopes and hysteresis is enabled

Parasitic Bipolar Transistor

1 Some of the generated holes may be attracted by the negatively charged Source



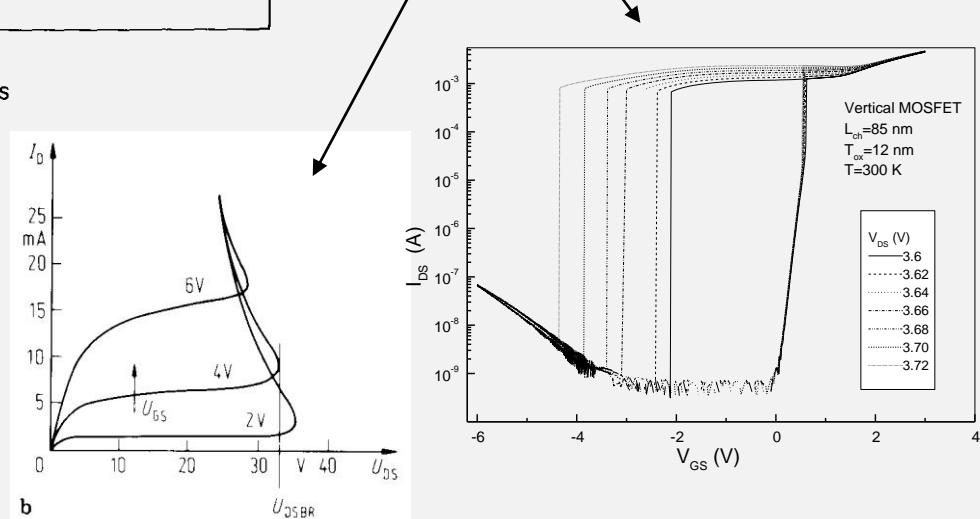
2 The existence of positively charged holes increases the substrate potential and lowers the junction barrier or even generate a forward biased junction for electrons

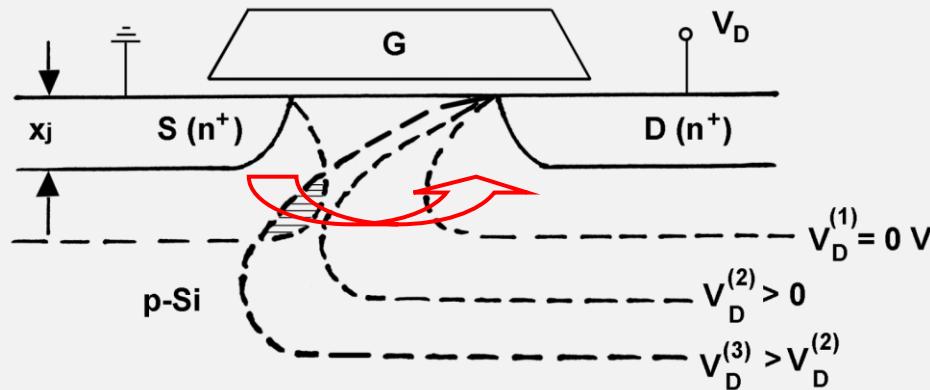
3 From the Source electrons are injected into the substrate and attracted by the positively charged Drain

4 The resulting current increase corresponds to a „hidden“, **parasitic bipolar transistor**

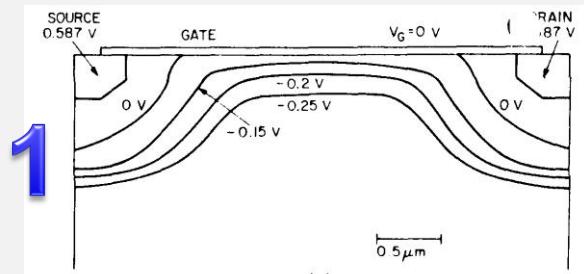
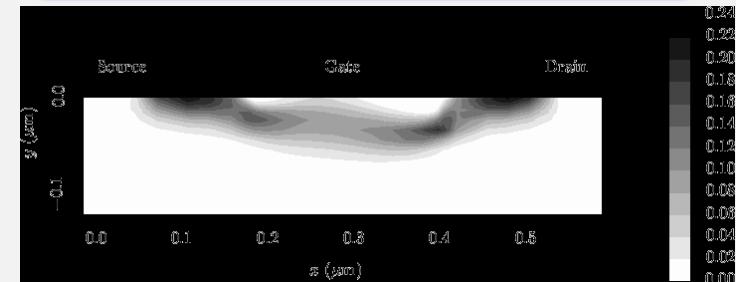
5 Some of the electrons may generate additional e-h pairs, where the new holes support the the bipolar mechanism

6 In consequence a self-controlling feed-back effect exists (**snap-back**)
The gate voltage can be switched off, but current remains !



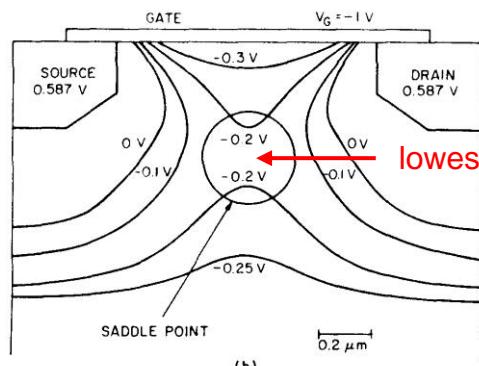
Punch-through:

Punch-through is the creation of a low barrier path through the bulk, if S/D depletion zones touch



Long-channel device:
 $L = 3 \mu\text{m}$, $N_{\text{dop}} = 1 \times 10^{15} \text{ cm}^{-3}$, $d_{\text{ox}} = 50 \text{ nm}$, $r_j = 0.33 \mu\text{m}$

$$V_s = 0.587 \text{ V}, V_{DS} = 0.587 \text{ V}, \\ V_g = 0 \text{ V}, V_{\text{bulk}} = -0.289 \text{ V}$$



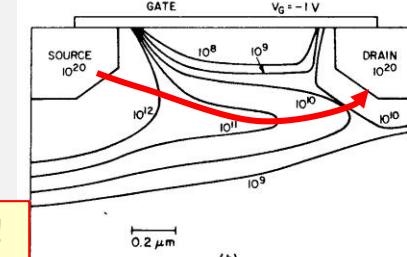
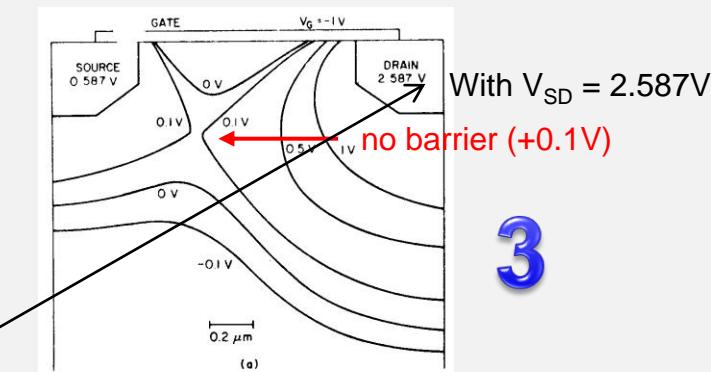
Same device, but:
 $L = 1 \mu\text{m}$, $V_g = -1 \text{ V}$, $V_{DS} = 0.587 \text{ V}$

lowest barrier (-0.2V) for electrons

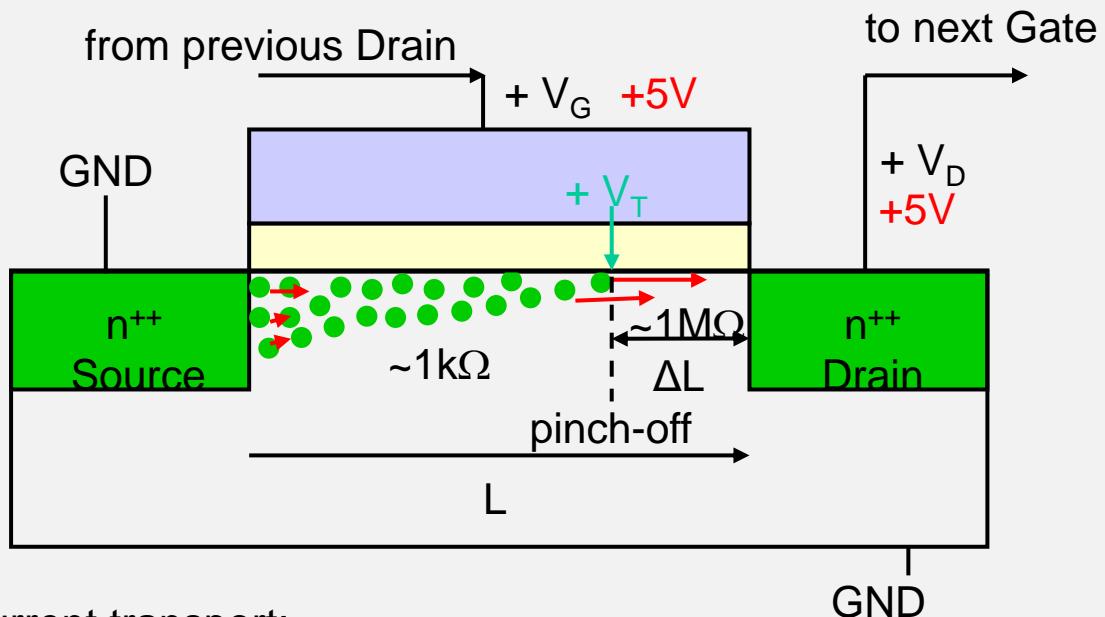
The depletion zones touch in the bulk, a saddle point of low potential is created, which can open a path for electrons

! Punch-through is not a npn-transistor effect !

(because no base current exists, no impact)



Created electron path



Current transport:

continuity equation: $\vec{j} = n \cdot e \cdot \vec{v} = \text{const}$

Many electrons start with low velocity at the Source in the channel, accelerated by drain-source voltage.
Few electrons reach the end of channel, but with higher velocity ($\rightarrow j = \text{const.}$).

At the end of channel electrons jump in ΔL by diffusion and are highly accelerated by the high electric field.

If the electric field is too high \rightarrow **impact ionization** and avalanche break-through occurs

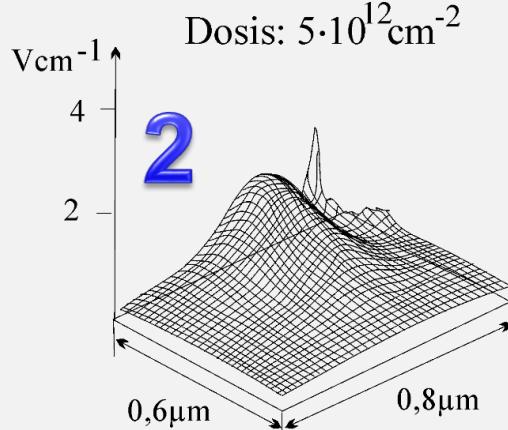
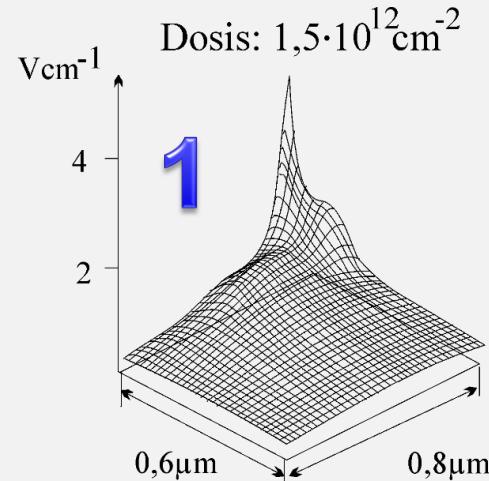
Technological solution: increase conductivity in front of Drain by an additional resistor

\rightarrow **Lightly-Doped Drain LDD**

- ▶ Drain current of MOSFET 1 loads the Gate of following MOSFET
 $\rightarrow V_{\text{Drain}} = V_{\text{Gate}}$
- ▶ - at Source the channel potential is $V_G - V_S \sim V_G \rightarrow$ a channel exists
- at Drain $V_G \sim V_D \rightarrow$ no channel exists
- ▶ no channel, but depletion/inversion
 $\rightarrow \Delta L = \text{isolator}$
 \rightarrow drain-source voltage mainly drops over short $\Delta L \rightarrow$ high electric field

\rightarrow today LDD structure is used for highly doped "extensions"

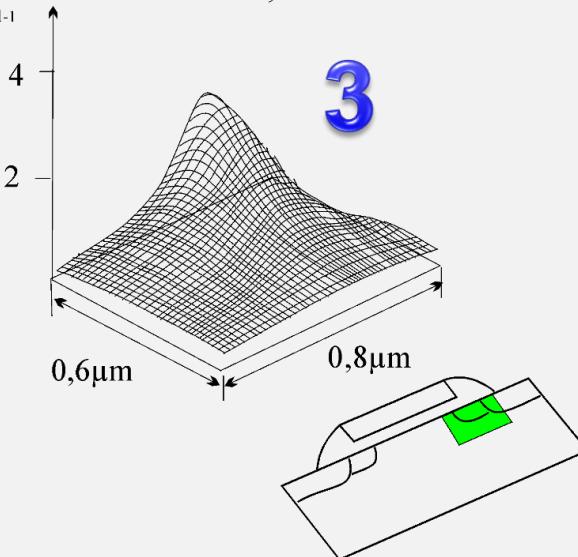
Adjustment of LDD Doping Level:



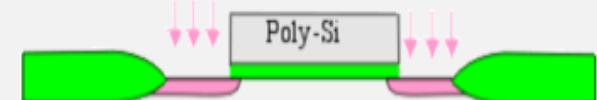
$$V_{SD} = 5 \text{ [V]}, V_G = 3 \text{ [V]}, \\ L_{eff} = 1.2 \text{ [\mu m]}, x_{LDD} = 0.2 \text{ [\mu m]}$$

Technology:

Dosis: $1,5 \cdot 10^{13} \text{ cm}^{-2}$



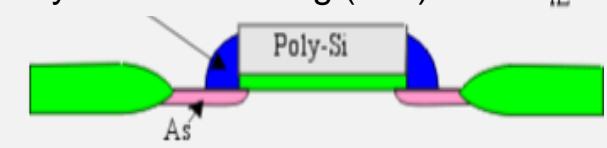
1. Self-adjusting Implantation of LDD



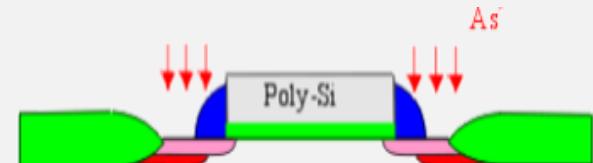
2. Maskless Deposition of Spacer Layer

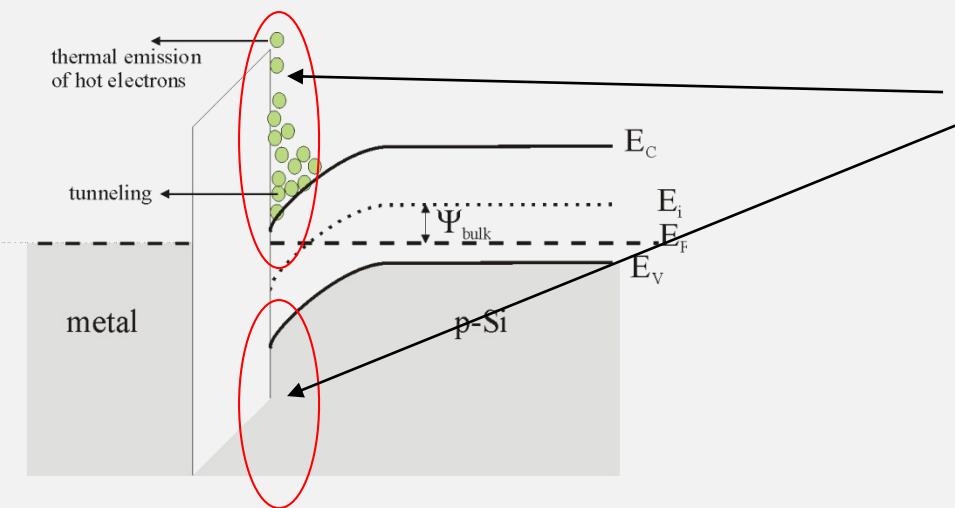


3. Maskless Creation of Spacers
by vertical Etching (RIE)



4. Self-adjusting Implantation of S/D



Gate-Injection:

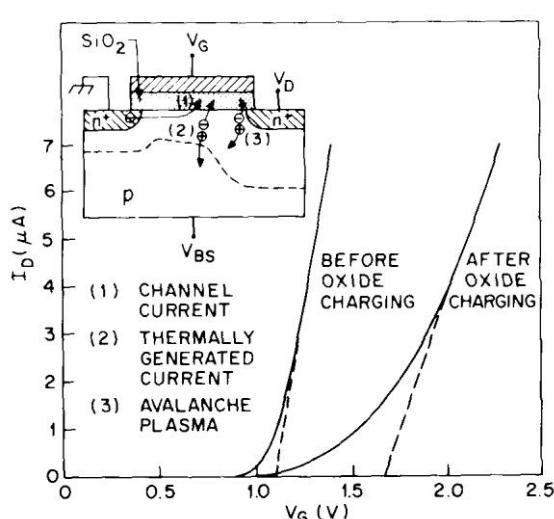
Hot carriers can overcome the barrier into the gate-oxide

3.1 eV for electrons

4.2 eV for holes

Thermal emission over the barrier is assisted by tunneling through the barrier

The resulting gate current degrades the gate oxide by creating defects and oxide charges.



Hot electrons origin from:

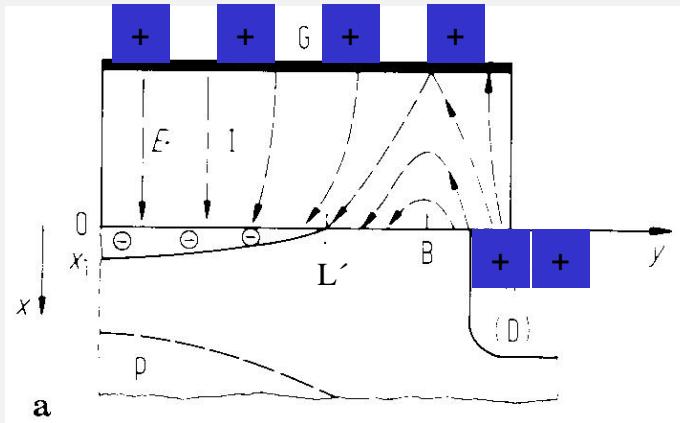
- channel hot electrons CHE (1)
- drain avalanche hot electrons DAHE (2)
- substrate generated hot electrons SGHE (3)

The result is:

- a shift of parameters,
eg. Shift of threshold voltage, transconductance -> reliability
- an early dielectric break-through

For calculating the break-through we have to calculate the Electric Field E in the gate dielectrics

Physical Models



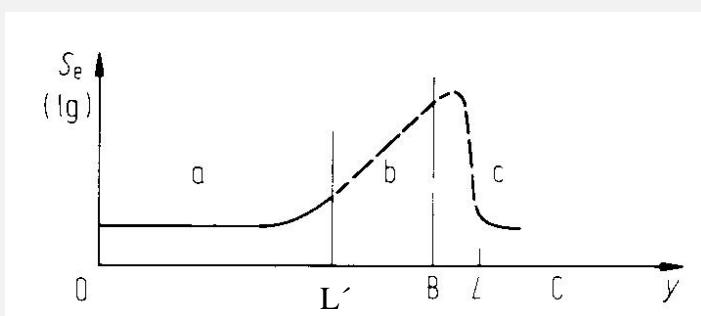
1

2-dim electric field in the pinch-off region

- highest density of electric field lines -> highest electric field
- change of field line direction at "gate field reversal point" B (near L')

2

Due to the inhomogeneous electric field in the pinch-off region
the **voltage drop across the oxide** V_{ox} must be calculated locally (y -dependent):



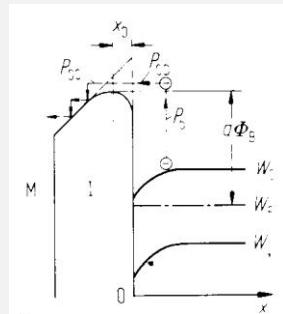
$$V_{ox}(y < L') = V_{th} + (V_{GS} - V_{th}) \cdot \sqrt{1 - \frac{y}{L'}}$$

$$V_{ox}(y > L') = V_{th} - (V_{GS} - V_{th}) \cdot \left[\frac{2}{3} \cdot \sqrt{\frac{\epsilon_{ox}}{\epsilon_{Si} d_{ox} y_{ch} L'}} \cdot (y - L')^{3/2} \right]$$

Change of field direction due to change in voltage sign

Quantitative Calculations for gate current:

Schottky-Barrier-Emission Model



1 Modified Richardson equation

$$I_G = AwT^2 \cdot \exp\left[-\frac{\Phi_B}{kT} \cdot \int_0^L \exp\left[\frac{\text{const.}}{T} \cdot \sqrt{\frac{V_{ox}(x)}{\varepsilon_0 \varepsilon_{ox} \cdot d_{ox}}}\right] dx\right]$$

2 Temperature of carrier is calculated from power:

$$kT \propto V_{DS} \cdot I_D \quad \rightarrow \text{for all electrons}$$

3 $V_{ox} = \dots$ see previous page

$$4 \quad q\Phi_B = 3.1\text{eV} - \frac{kT}{q} \cdot \sqrt{E_{ox}} - 9 \cdot E_{ox}^{2/3}$$

$$= 3.1 - 2.6 \cdot 10^{-4} \sqrt{E_{ox} [\text{V/cm}]} - 10^{-5} (E_{ox} [\text{V/cm}])^{2/3}$$

Hot Electron Gas Emission Model

-> better estimation of electron energy kT_e

The energy of the carrier distribution determines the number of injected electrons

even at $V_{SD} < E_{gap}$ some electrons gain enough energy

1a

Shifted Maxwell-distribution
(from Boltzmann transport equation)

$$T_e(x) = \frac{2q}{5k} \cdot \int_0^\infty E_x(x-u) \cdot \exp\left[-\frac{3u}{5\tau_E \cdot v_{drift}}\right] du$$

1b

Empirical:

$$\frac{T_e}{T} = 1 + \frac{\tau_E \cdot v_{drift} \cdot q}{3/2 kT} \cdot \frac{\vec{S}_n \cdot \vec{E}}{|\vec{S}_n|}$$

Energy relaxation time $\tau_E : \sim 0.5 \cdot 10^{-12} \text{ sec}$

Drift velocity $v_{drift} : \sim v_{max} \sim 10^7 \text{ cm/sec}$

Lattice temperature $T \sim \text{operating temperature}$

Channel current density $\vec{S}_n = q \cdot n(x, y) \cdot v_{drift}$

2

$$I_{gate} = \text{const} \cdot n_{ch} \cdot \sqrt{T_e} \cdot \exp\left[-\frac{\Phi_B}{kT_e}\right]$$

Lucky Electron Model

The probability determines the number of injected electrons

some lucky electrons without scattering gain enough energy to overcome the barrier

1 Probability of thermal emission

$$P_B = A^* \cdot \sqrt{kT_e} \cdot \exp\left[-\frac{\Phi_B}{kT_e/q}\right]$$

2 Probability of oxide tunneling

$$P_{tun} = \text{const} \cdot I_D \cdot \exp\left[-\frac{\Phi_B}{E_{ave} \cdot \lambda}\right]$$

3a

$$I_G = \frac{I_{DS}}{\lambda_\perp} \cdot \int_0^L P_B P_{tun} dx$$

λ_\perp Mean free path
(or probability for scattering of hot electrons with change of direction)

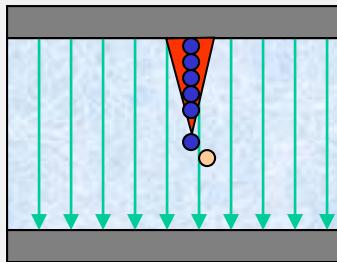
3b

$$I_G = \text{const} \cdot I_D \cdot \exp\left[-\frac{\Phi_B}{E_{ave} \cdot \lambda}\right]$$

Gate Dielectric Break-through:

1

Ideal (= no defects), material dependent break-through of bulk materials



1 Because of thermal excitation always e-h pairs are thermally generated within the semiconductor

$$n_i = \sqrt{N_C N_V} \cdot \exp\left[-\frac{E_{gap}}{2kT}\right]$$

	metal	Ge	Si	GaAs	SiO ₂
E _{gap}	-	0.67	1.12	1.42	9.0
n _i [cm ⁻³]	1e22	2.4e13	1.5e10	1.8e6	1

2 In high electric fields the carriers are accelerated and can gain enough energy for impact and avalanche multiplication

3 The avalanche track may destroy the dielectric material by generation of defects

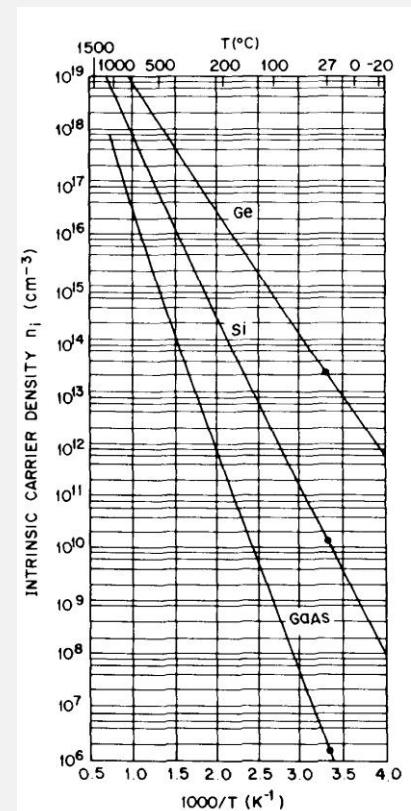
$$E_{ox,bt} \approx (5-10) \cdot 10^6 V/cm \approx (0.5-1) V/nm$$

For various insulator materials the expression holds:

$$E_{max}(ins2) = E_{max}(ins1) \cdot \frac{\epsilon_1}{\epsilon_2}$$



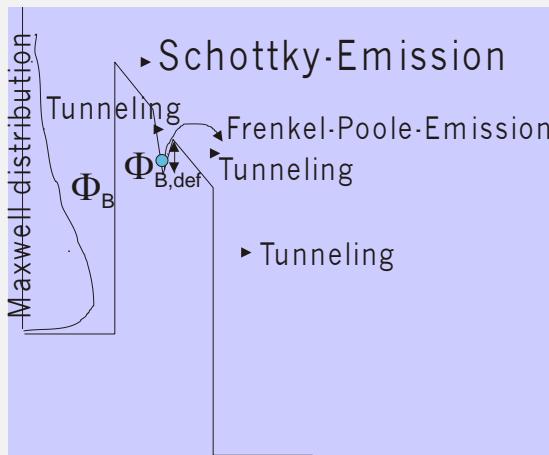
important for high-k materials -> break-through is lowered !



2

Real, material dependent break-through

In reality in every dielectric local defects exist (unsaturated bonds, ions, traps, ...), which enhance current transport



Schottky-Emission: thermal emission of electrons

$$I \propto T^2 \cdot \exp\left[\frac{a\sqrt{V}}{T} - \frac{q\Phi_B}{kT}\right]$$

Ohmic Emission: electrons in defects are hopping from defect to defect

$$I \propto V \cdot \exp\left[-\frac{c}{T}\right]$$

Frenkel-Poole-Emission: Ohmic Emission with field assistance

$$I \propto V \cdot \exp\left[\frac{2a\sqrt{V}}{T} - \frac{q\Phi_B}{kT}\right]$$

Tunneling: injected or trapped electrons

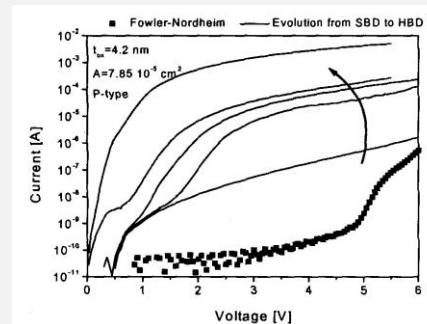
$$I \propto V^2 \cdot \exp\left[-\frac{b}{V}\right]$$

All currents charge and degrade the oxide quality and electrical parameters on a short or long term scale

In MOSFETs: To introduce safety in dependence of real defect densities an additional voltage limit is defined

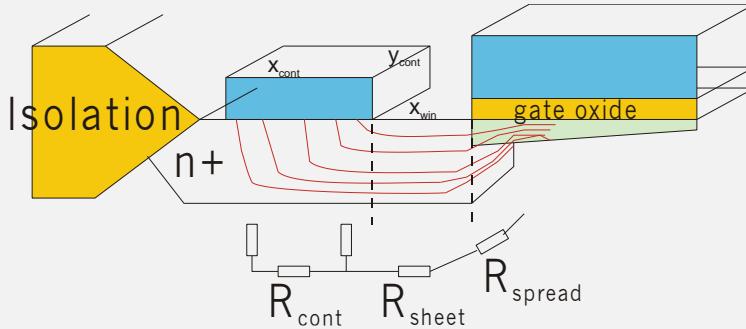
$$V_{G,\max} = d_{ox} \cdot \left(\frac{E_{bd}}{2} - \Delta E_{dec} \right)$$

With ΔE_{dec} is a change in field strength, which causes a current change of one decade



High Electric Fields induce:

- Modification in carrier transport
 - Mobility degradation due to Gate field
 - Velocity saturation due to Source-drain field
 - Channel length modulation in saturation
- Break-through effects
 - 1 Hot electrons -> Impact ionization
 - Avalanche multiplication (Drain break-through)
 - Parasitic Bipolar Transistor
 - Gate damage
- Parasitics
 - Source-Drain resistors
 - Capacitors



In a real device series resistances (Source-drain resistance) influences the device characteristics

The **contact resistance** is dependent from a material dependent specific contact resistance ρ_{cont} [Ωcm^2] and a transformation resistance

The **sheet resistance** is the area resistance of the highly doped region

The **spreading resistance** is due to the current confinement from the highly doped region to the channel region

Contact resistance (metal - highly doped region)

$$R_{cont} = \frac{\sqrt{\rho_{cont} \cdot \rho_{dop}} / r_j}{y_{cont}} \cdot \coth \left[\sqrt{\frac{\rho_{dop}}{\rho_{cont}}} \cdot \frac{x_{cont}}{r_j} \right]$$

$$\begin{aligned} & X_{cont} \gg x_{window} \rightarrow R_{cont} \approx \sqrt{\frac{\rho_{cont} \cdot \rho_{dop}}{r_j}} \\ & X_{cont} \ll x_{window} \rightarrow R_{cont} \approx \frac{\rho_{cont}}{y_{cont} \cdot r_j} \end{aligned}$$

Typically:

$$R_{cont} \approx \frac{\sqrt{10\Omega\mu\text{m}^2 \cdot 10^{-3}\Omega\text{cm}}}{100\text{nm}} \approx 1\text{k}\Omega$$

for 100nm contacts !

Sheet resistance:

$$R_{sheet} = \frac{\rho_{dop}}{r_j} \cdot \frac{x_{wind}}{y_{cont}}$$

Typically:

$$R_{sheet} = \frac{10^{-3}\Omega\text{cm}}{100\text{nm}} \cdot \frac{100\text{nm}}{100\text{nm}} \approx 100\Omega$$

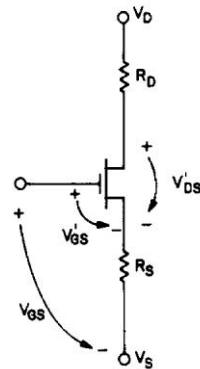
Spreading resistance:

$$R_{spread} = \frac{2\rho_{dop}}{\pi \cdot y_{cont}} \cdot \ln \left[0.75 \cdot \frac{x_j}{x_{ch}} \right]$$

Typically:

$$R_{spread} = \frac{2 \cdot 10^{-3}\Omega\text{cm}}{\pi \cdot 100\text{nm}} \cdot \ln \left[0.75 \cdot \frac{100\text{nm}}{5\text{nm}} \right] \approx 200\Omega$$

Due to the series resistance the applied voltages V are reduced by a voltage drop at source and drain and reduces the voltages V^* at the ideal device



$$V_{DS}^* = V_{DS} - I_D \cdot (R_{source} + R_{drain})$$

$$V_{GS}^* = V_{GS} - I_D \cdot R_{source}$$

In the linear region ($V_{DS} \ll V_{GS} - V_{th}$), e.g. $V_{DS} = 3kT \sim 70mV$ and $V_{GS} - V_{th} \sim 1V$

$$I_D \approx \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot (V_{GS}^* - V_{th}) \cdot V_{DS}^*$$

Using: $R_s V_{DS} \ll (R_s + R_d) \cdot (V_{GS} - V_{th})$ and: $2\mu \cdot C_{ox} \cdot \frac{w}{L} \cdot R_s \cdot V_{DS} \ll 1$

$$I_D \approx \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot \frac{(V_{GS} - V_{th})}{1 + \alpha \cdot (V_{GS} - V_{th})} V_{DS}$$

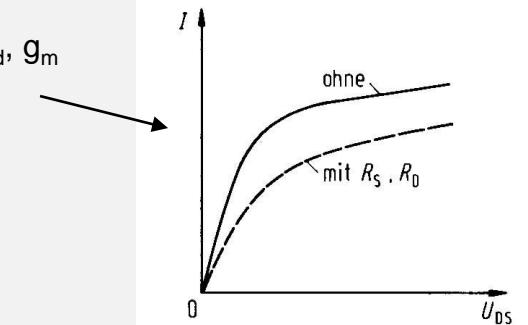
$$\alpha = \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot (R_s + R_d)$$

Experimental derivation:

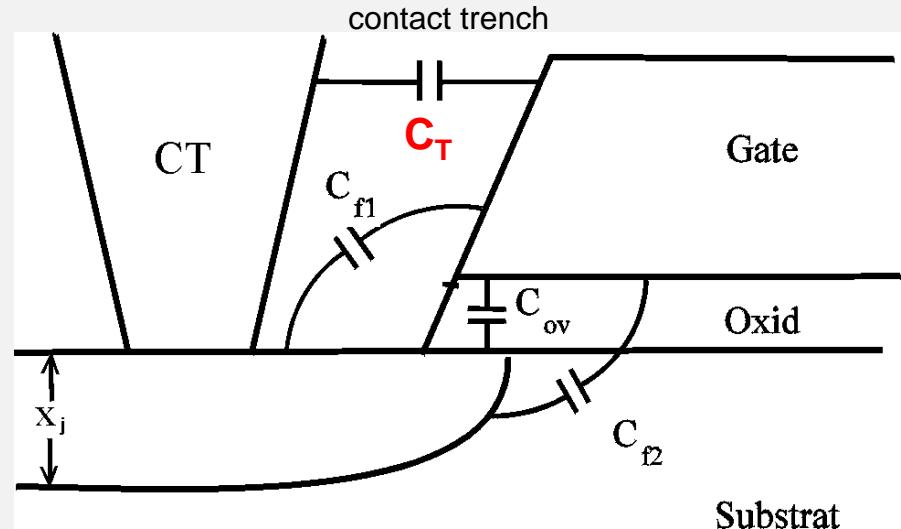
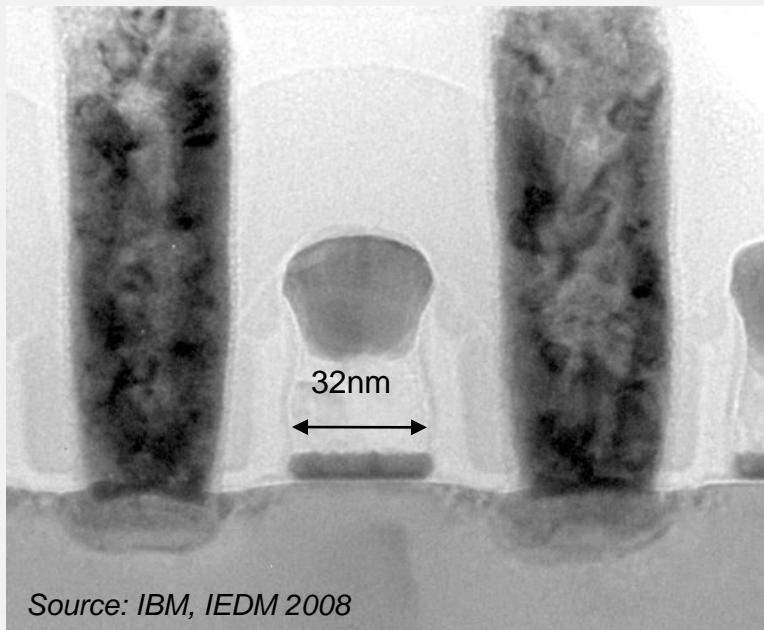
By rearranging

$$\frac{(V_{GS} - V_{th})}{I_D / V_{DS}} \approx \frac{1 + \alpha \cdot (V_{GS} - V_{th})}{\mu \cdot C_{ox} \cdot \frac{w}{L}} = X$$

degradation of I_d , g_m



Measuring V_{th} for two MOSFETs with different L
two equations for α are the result and R_s and R_d can be determined



$$C_{par} = \underbrace{\frac{\varepsilon_0 \varepsilon_{EOT} (d + \Delta)}{d_{EOT}}}_{C_{ov}} + \underbrace{\frac{\varepsilon_0 \varepsilon_{iso}}{\alpha} \ln \left[1 + \frac{x_p}{d_{ox}} \right]}_{C_{f1}} + \underbrace{\frac{2 \varepsilon_0 \varepsilon_{Si}}{\pi} \ln \left[1 + \frac{x_j}{d_{EOT}} \sin \frac{\pi \varepsilon_0 \varepsilon_{EOT}}{2 \varepsilon_{Si}} \right]}_{C_{f2}} + C_{cont}$$

High Electric Fields induce:

- Modification in carrier transport
 - Mobility degradation due to Gate field
 - Velocity saturation due to Source-drain field
- 2-dim field effects
 - Channel length modulation in saturation
- Break-through effects
 - 1 Hot electrons -> Impact ionization
 - Avalanche multiplication (Drain break-through)
 - Parasitic Bipolar Transistor
 - Gate injection
 - 2 Gate break-through
 - 3 Punch-through
- Parasitics
 - Source-Drain resistors



High electric fields modify (degrade) the MOSFET performance
High-field effects are enhanced by reducing dimensions

3.1 MOSFET intro

- Basic MOSFET working principle
- MOSFET types
- MOSFET applications
- Basic MOSFET fabrication process

3.2 MOS physics

- Work function
- Contact phenomena
 - Metal - metal
 - Metal - semiconductor
 - Metal - insulator - semiconductor (MIS)
- Working regimes of the MOS structure

3.3 Ideal long-channel MOSFET

- Charge models (Q-constant, Gradual-Channel)
- MOSFET working regimes

3.4 Realistic behavior

- Basics of charge carrier transport
- Modification in charge carrier transport
- Break-through phenomena
- Parasitics

**End
of
Chapter 3**