LS Scienza dei Materiali - a.a. 2006/06

Fisica delle Nanotecnologie – part 1

Version 5, Oct 2006

Francesco Fuso, tel 0502214305, 0502214293 - fuso@df.unipi.it
http://www.df.unipi.it/~fuso/dida

Nanotecnologie: generalità, ambito, motivazioni; stato dell'arte (elettronica)

6/10/2006 – 15-17 – room T1

Outlook

- What is *nanotechnology?*
- What are the *components* of nanotechnology?
- What are the main *driving forces* for the development of nanotechnology?
- What is the *present status* of technology?
- Survey of *conventional* electronics:
    - Electron transport (Drude model);
    - Bipolar junctions (old-style technology);
    - Planar technology and MOS-FETs
- Some *limits and problems* in miniaturization and the need for new approaches

(Nano)technology

Technology: the ability to fabricate systems **useful** for some applications
Nano: fabricated systems are "small"-sized (*hard to figure out how small…)*

i.e., the ability to manipulate matter in order to fabricate systems (or structures, or devices) with a size in the **sub-micrometer** range

Technology uses techniques, but **it is not just a technical application**: basic science is involved as well in designing new techniques and new structures with improved functionalities

(Nano)technology is strictly connected with basic science, but **it is not just investigation/interpretation** of processes in the nano-world

Nanotechnology is a "complex" (and rather vague) matter

[concepts from M.Wilson et al., Nanotechnology (Chapman&Hall, 2002)]

Components of nanotechnology

Nanotechnology shares topics with other disciplines,
but it **should not be confused** with:

-Chemistry, for the higher *control* of the involved processes;
-Materials science, for the specific interest in the small world;
-Physics, for the complexity of the systems under investigation;
-Engineering, for the specific interest in new systems;
-Biophysics, (self assembly and replication) for the *artificial* systems

Nanotechnology is an "open" and strongly interdisciplinary field

An historical example of nanotechnology

Lycurgus Cup in Roman times

Dr. Juen-Kai Wang

The glass appears green in daylight (reflected light), but red when the light is transmitted from the inside of the vessel.

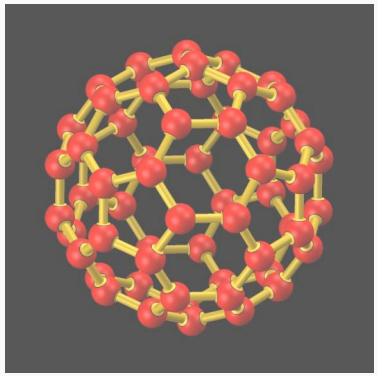
"Nanostructured" glass

The Lycurgus Cup, Roman (4th century AD), British Museum *(www.thebritishmuseum.ac.uk)*
F. E. Wagner et al., Nature ***407****, 691 (2000).*

A more recent example

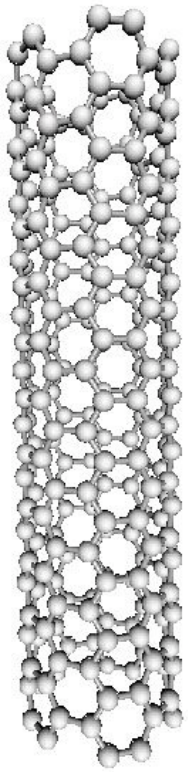
BASIC SCIENCE

NANOTECHNOLOGY



Fullerene (C_{60})

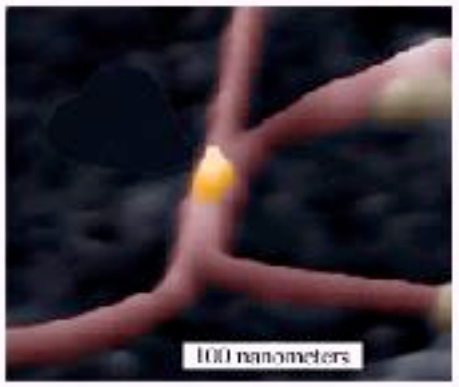
(Nobel Prize, mid 90's)



Single Wall
Carbon NanoTube
(90's)

Mesoscopic systems
(interesting for their physico-chemical properties

An artificial system made of CNT and gold nanoparticle intended to be a prototypal single-electron device
(a couple of years ago)



There is plenty of room at the bottom…I

There's Plenty of Room at the Bottom

An Invitation to Enter a New Field of Physics

by Richard P. Feynman

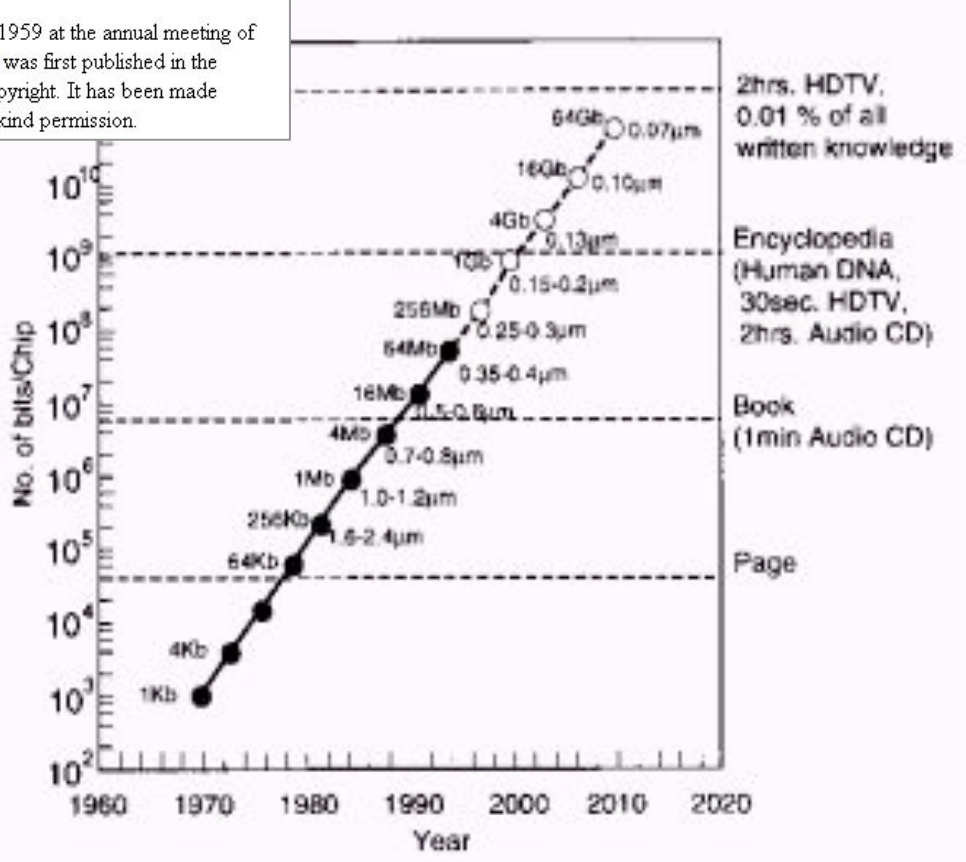
This transcript of the classic talk that Richard Feynman gave on December 29th 1959 at the annual meeting of the American Physical Society at the California Institute of Technology (Caltech) was first published in the February 1960 issue of Caltech's *Engineering and Science*, which owns the copyright. It has been made available on the web at http://www.zyvex.com/nanotech/feynman.html with their kind permission.

How do we write small?

Information on a small scale

Miniaturizing the computer

Miniaturization means increase of "power" in Information Technology

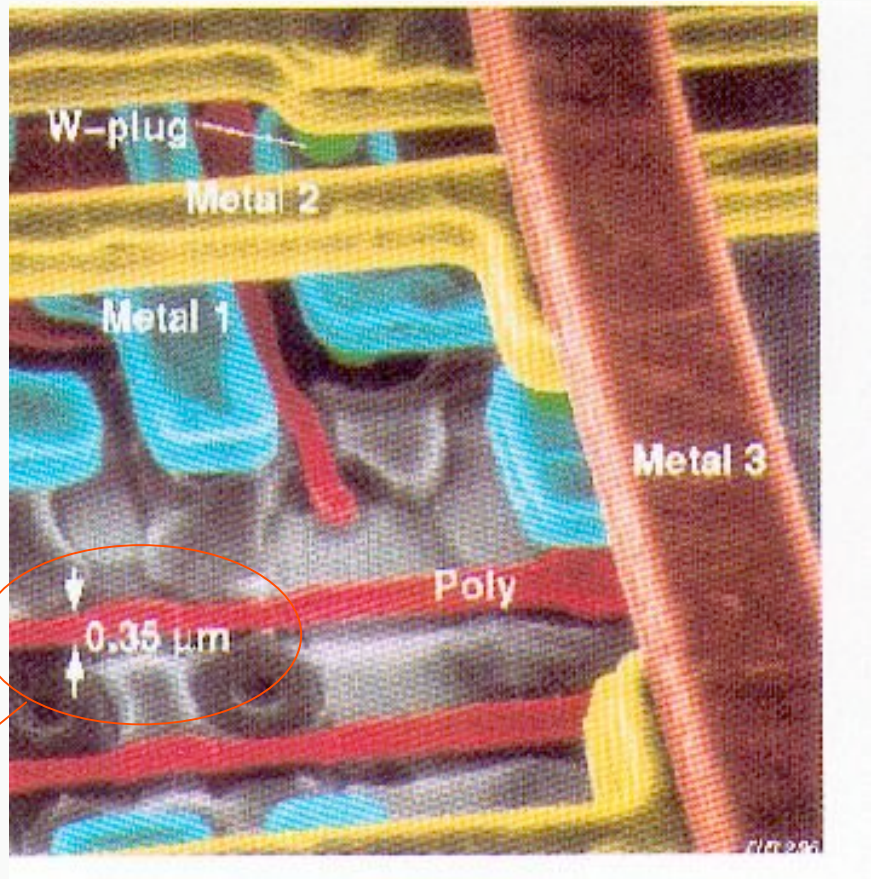


Driving forces for nanotechnology

Electronics devices:
they are typically (and *traditionally*) made of "small" structures

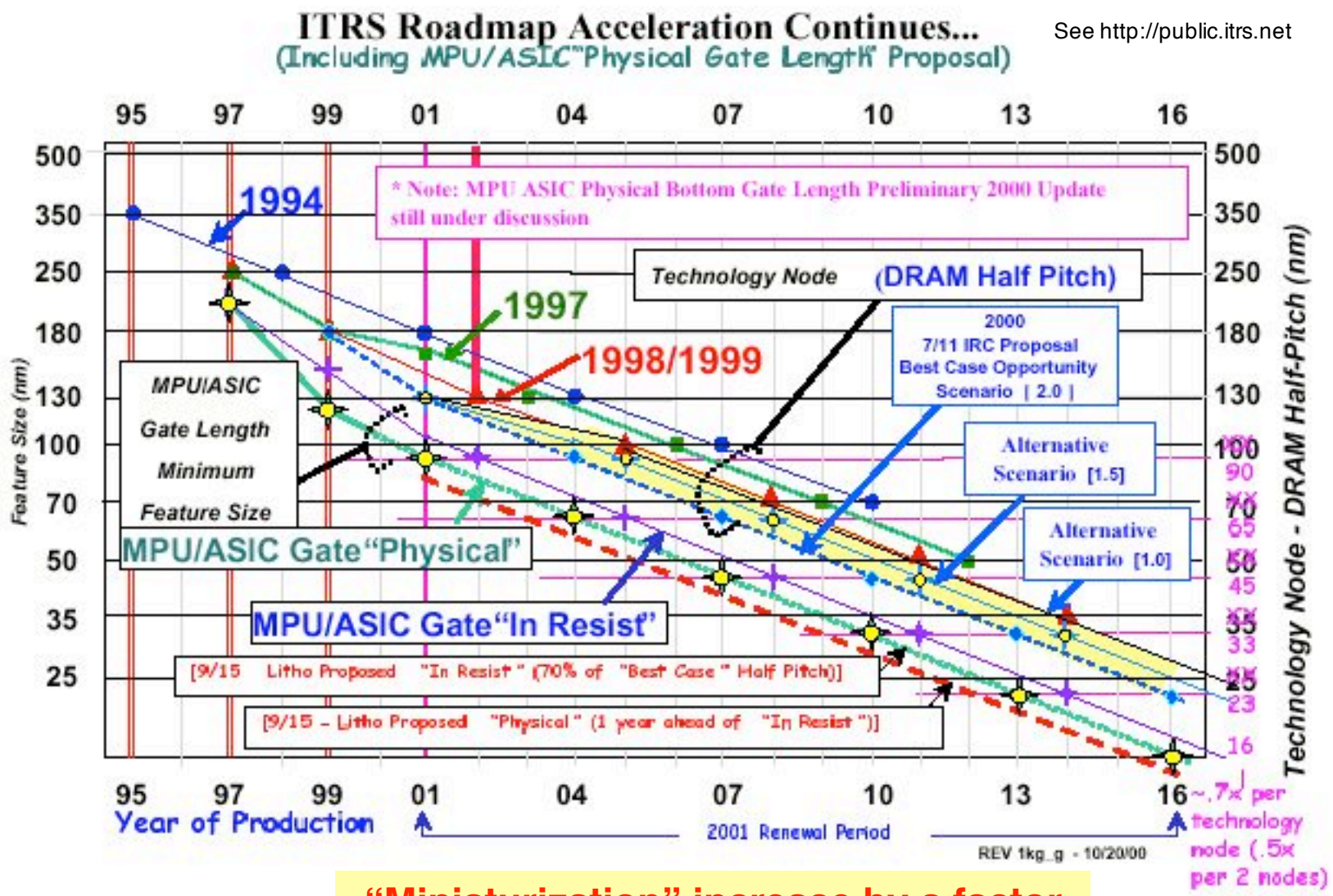
1. Thin films are deposited
2. A pattern is tranferred to the multilayered structure

Device components (resistors, capacitors, transistors, …) are so defined in an *integrated* structure



Feature size
(typ., fwhm of the smaller device features)

The “Moore’s law”



“Miniaturization” increase by a factor 3-4 in a 3-4 year period (still true?)

The present (2004) status of miniaturization (commercial)

How many transistors can dance on the head of a chip only 66 millimeters square? Over 58 million, thanks to IBM's sophisticated process technology that builds them just 90 nanometers wide. Such superior technology developments turbo-charge the G5 processor to speeds of up to 2.5GHz.

To get electronics so small requires miniaturization breakthroughs, and IBM's dedication to basic scientific research makes these advances possible. For instance, the company began researching copper as an interconnect method over 25 years ago, but the technique wasn't practical until just recently.

One in 58 Million. A transistor just 90nm wide (yellow) on substrate of SOI (blue) with copper interconnects (gray). Layers of nitride (brown) and oxide (green) insulate it from its brethren. Magnified 146,000 times.

So Small

Transistors on the PowerPC G5 hold a charge to let the system make logic decisions based on whether the transistor is on or off. Using a 90nm process for even greater performance, IBM builds these devices just .00000009 meters wide on a layer of silicon on insulator. The 58 million transistors themselves are connected by over 400 meters of copper wire that's less than 1/1000th the width of a strand of your hair. Tiny paths mean less time to complete a sequence, since the

http://www.apple.com

Feature size slightly below 100 nm (nanotechnology?)

Technical and fundamental problems

- In order to maintain the progress of Moore's Law, the 2001 ITRS envisions **more aggressive scaling than projected in prior roadmaps**. For example, dynamic random access memory chips will feature critical dimensions of 90 nanometers in 2004, which is both smaller and sooner than the 100 nanometers projected for 2005 in the roadmap published just two years ago. Similarly, microprocessor transistor gate lengths – a critical dimension that affects the processor's speed – will be just 25 nanometers in 2007, six years sooner than expected in the 1999 version of the roadmap. (Note: a nanometer is one-billionth of a meter. A human hair is 100,000 nanometers in width, and a red blood cell is 5,000 nanometers in width.)

- **We are beginning to reach the fundamental limits of the materials used in the planar CMOS process**, the process that has been the basis for the semiconductor industry for the past 30 years. Further improvements in the planar CMOS process can continue for the next five to ten years by introducing new materials into the basic CMOS structure. However as the ITRS looks forward 10-15 years, it becomes evident that even with the introduction of new materials, most of the known technological capabilities of the CMOS device structure will approach or have reached their limits. In order to continue to drive information technology advances, it becomes necessary to investigate new devices that may provide a more cost-effective alternative to planar CMOS in this timeframe.

Da www.sia-online.org

The rate of increase in miniaturization has been growing fast in information technology

Main motivations:
- Increase of "power" (computing efficiency, information storage, time response, …)
- Decrease of power consumption, usually associated with miniaturization
- Commercial reasons (a huge market!)

Technical limitations: lack of control in the manipulation, limitations inherent to the materials

Fundamental limitations: in the **techniques** (e.g., optical diffraction in lithography), in the **system operation** (e.g., *quantum* behavior)

Need for novel approaches

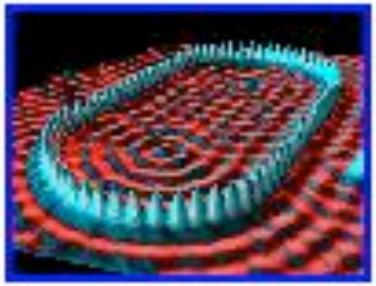
There is plenty of room at the bottom…II

Miniaturization by evaporation

Better electron microscopes

Atoms in a small world

Rearranging the atoms

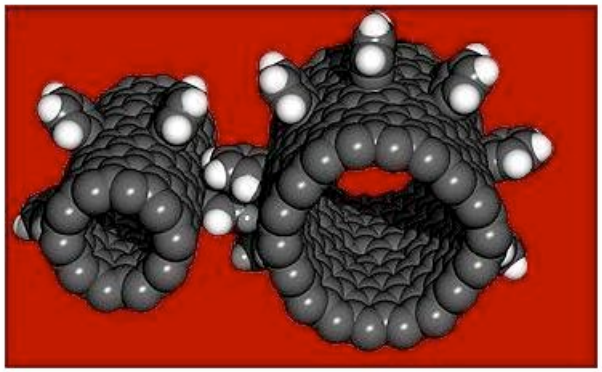


Title : *Stadium Corral*

Media : *Iron on Copper (111)*

IBM®

Miniaturization means new (*quantized*) functionalities exploitable in novel applications



A representation of nanogears made from graphitetubes billionths of a meter wide. (Picture from the NanoGallery, see references)

Nanomachines for, e.g., computation, drug dispensing, nanofluidics, …

4 • NANOTECHNOLOGY

'nanotechnology is the principle of atom manipulation atom by atom, through control of the structure of matter at the molecular level. It entails the ability to build molecular systems with atom-by-atom precision, yielding a variety of nanomachines.'

Eric Drexler (1990)

Manipulation and control of the matter at the single atom level

Our point of view I

Besides nanoelectronics, development of nanotechnologies is a crucial point in a huge variety of applicative areas, including, e.g.:
- improvement of mechanical/surface properties in coatings and structural materials (nanocomposites, …);
- realization of new and more efficient approaches for diagnostics (and even therapy) in biophysics, biomedicine (fluorescence markers, drug dispensers ,..);
- enhancement of data storage capabilities (DVD, hard-disks, …);
- enhancement of energy storage capabilities (fuel cells, …);
- realization of novel computation methods (quantum computers, …);
- design and fabrication of emitting devices (quantum-dot lasers, …);
- …

However, the “transition” from micro- to nanoelectronics appears as the most compelling and challenging area for:
- “historical” reasons;
- the inherent miniaturized nature of electronic devices;
- the ability to point out both limitations and new possibilities offered by extreme miniaturization

Our point of view II

Selected topics of interest

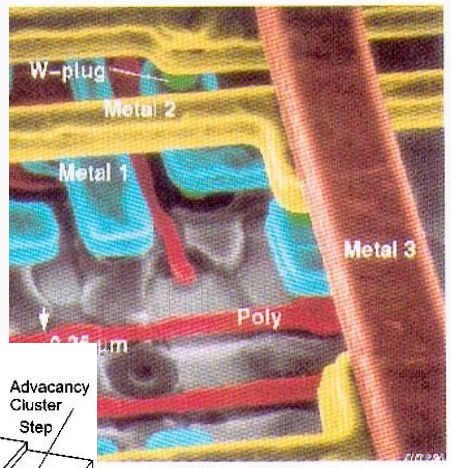
1. *A physical picture* of the **state-of-the-art** in (micro)technology for electronics

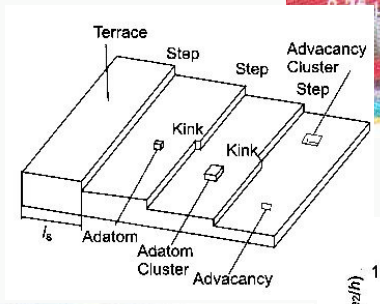
2. *Physical* methods for **fabrication** of nanostructures (i.e., evaporation, lithography, atom manipulation) and associated problems

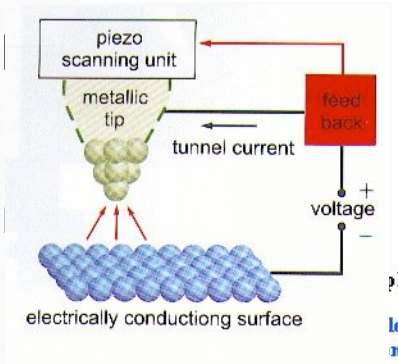
3. *Physical* tools for nanostructure **investigation** (i.e., probe microsco and new tools for fabrication

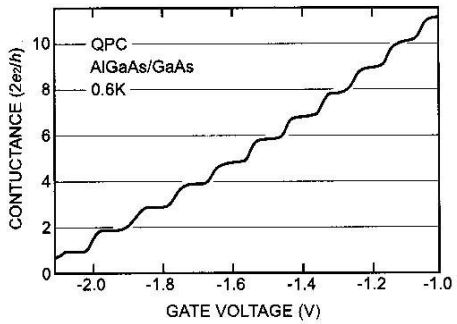
4. *Physical* **properties** of nanostructu (quantum confinement) and issues associated with miniaturization of electronic devices

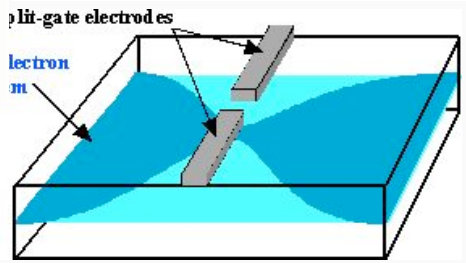
"Special" interest in *materials*



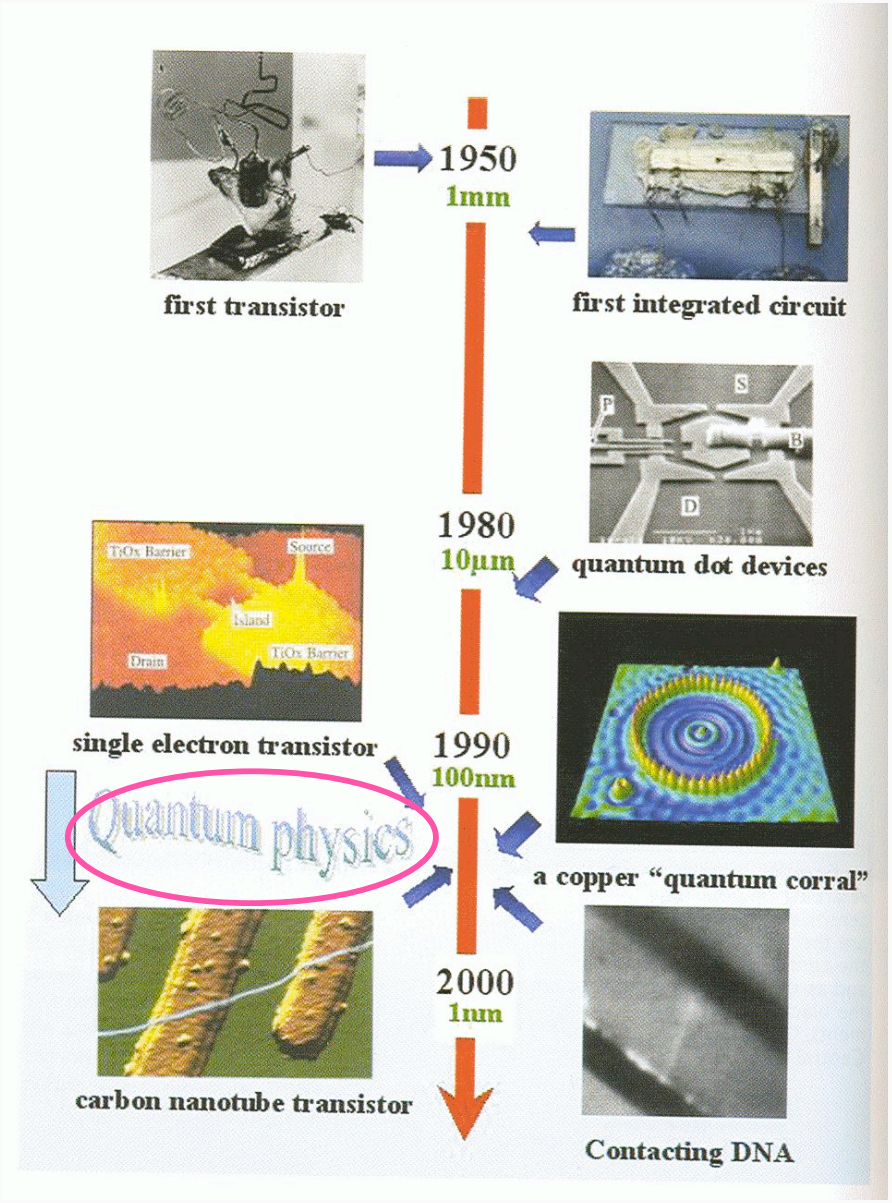








Progress in “electronics”



Bipolar transistor

Planar (thin film) technology

Optical lithography

Very Large Scale Integration

Quantum confinement

Alternative materials and technologies

???

What is electronics made of…

Electronics means the ability to control charge transport and requires:
- regions/materials where charge can move almost freely (e.g., interconnects, electrodes);
- regions/materials where charge motion is forbidden (e.g., dielectrics, capacitors);
- devices where charge motion is controlled

Conventional (micro)electronics achieves those tasks mostly by using inorganic materials (crystalline and/or amorphous conductors, semiconductors, dielectrics) and suitable architectures which, usually, can be (ideally) scaled down in size

Miniaturization may pose limitations even in conventional technologies

Basics of *conventional* electronics I

Diffusive electron transport (Drude)

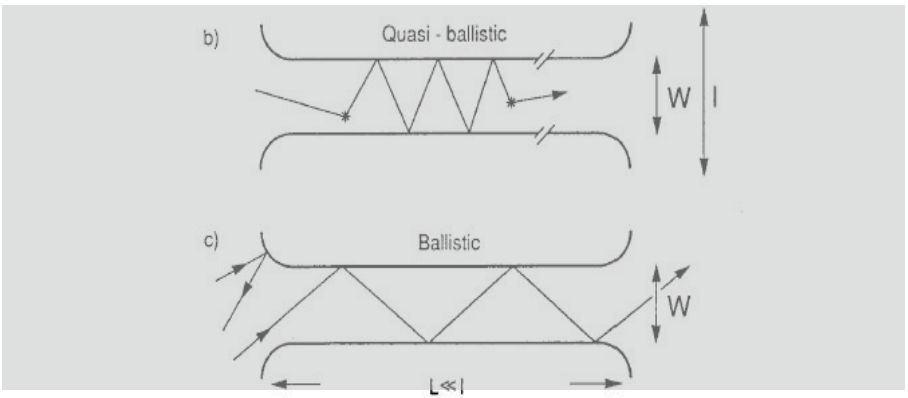


Figure 10.4: Electron trajectories characteristic of the diffusive ($\ell < W, L$), quasi-ballistic ($W < \ell < L$), and ballistic ($W, L < \ell$) transport regimes, for the case of specular boundary scattering. Boundary scattering and internal impurity scattering (asterisks) are of equal importance in the quasi-ballistic regime. A nonzero resistance in the ballistic regime results from backscattering at the connection between the narrow channel and the wide 2DEG regions. Taken from H. Van Houten et al. in "Physics and Technology of Submicron Structures" (H. Heinrich, G. Bauer and F. Kuchar, eds.) Springer, Berlin, 1988.

Classical interpretation (Drude):
Collisions betwen electrons and lattice ions lead to a friction force (and electrons do have thermal distribution of speed)
Quantum interpretation:
Collisions are replaced by loss of translational invariance in the electron wavefunctions (and the Fermi velocity must be considered)

Drift (limit) velocity:
$v_d = \tau\, eE/m^*$
but $\mathbf{J} = n\, e\, \mathbf{v_d}$
hence: $\sigma = n\, e^2\, \tau / m^*$

"Microscopic" Ohm's law:
$\mathbf{J} = \sigma\, \mathbf{E}$

τ : time interval between collisions
m^* : effective mass of the charge carriers

Diffusive means dissipative (resistance)

Transport in conductive thin films

Thin film can be considered as a *one-dimensional* example of nanotechnological system

Da M. Ohring, The materials science of thin films (Academic, 1992)

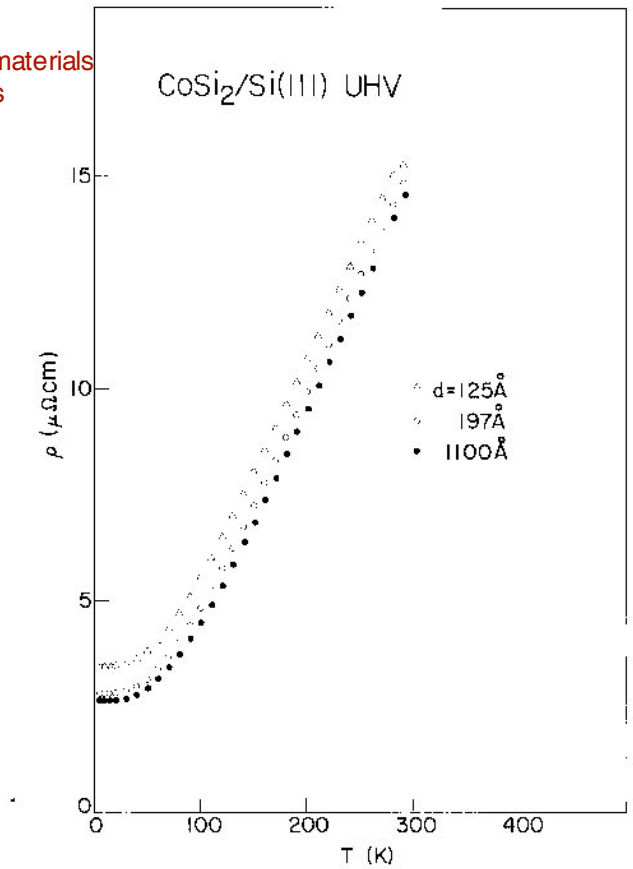


Figure 10-6. Temperature dependence of resistivity of $CoSi_2$ films. The 125-Å and 197-Å films are epitaxial. The 1100-Å film is polycrystalline. (From Ref. 10).

Neglecting (for the moment) any quantum-confinement effect

Resistivity tends to increase with decreasing film thickness due to the increased role of the electron collisions at the film interface

Interface gets importance in ruling the system behavior

Transport in dielectric thin films

Table 10-2. Conduction Mechanisms in Insulators

Mechanism	J–$\mathscr{E}$ Characteristics		Experimentally Derivable Material Constants
1. Schottky Emission	$J_S = AT^2 \exp -\frac{q\Phi_B}{kT} \exp\left[\frac{1}{kT}\left(\frac{q^3\mathscr{E}}{4\pi\varepsilon_i}\right)^{1/2}\right]$	(10-21)	Φ_B
2. Tunneling	$J_T = \frac{q^2\mathscr{E}^2}{8\pi h\Phi_B} \exp -\left[\frac{8\pi(2m)^{1/2}}{3hq\mathscr{E}}(q\Phi_B)^{3/2}\right]$	(10-22)	Φ_B
3. Space Charge Limited	$J_{SCL} = \frac{9\mu\varepsilon_i}{8}\frac{\mathscr{E}^2}{d}$	(10-23)	—
4. Ionic Conduction	$J_I = \frac{a\mathscr{E}}{kT} \exp -\frac{E_I}{kT}$	(10-24)	E_I
5. Intrinsic Conduction	$J_{In} = bT^{3/2} \exp -\frac{E_g}{2kT} \cdot \mathscr{E}$	(10-25)	E_g
6. Poole–Frenkel Emission	$J_{PF} = c\mathscr{E} \exp -\frac{E_i}{kT} \exp\left[\frac{1}{kT}\left(\frac{q^3\mathscr{E}}{\pi\varepsilon_i}\right)^{1/2}\right]$	(10-26)	E_i

a, b, c = constant.
ε_i = insulator dielectric constant.

Da M. Ohring, The materials science of thin films (Academic, 1992)

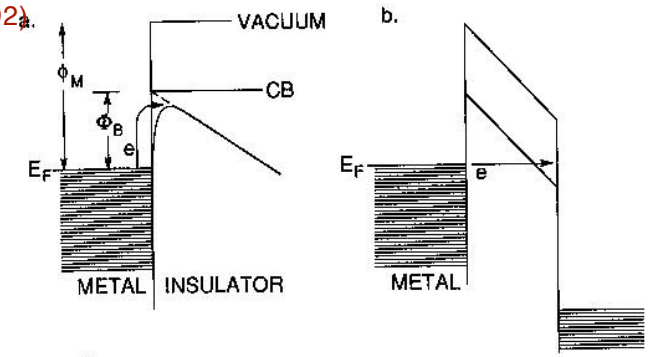


Figure 10-7. Barrier limited conduction mechanisms. (a) Schottky emission; (b) tunneling.

Barrier or thermal-activated processes tend to promote conduction in dielectric (thin) layers

Neglecting (for the moment) any quantum-confinement effect

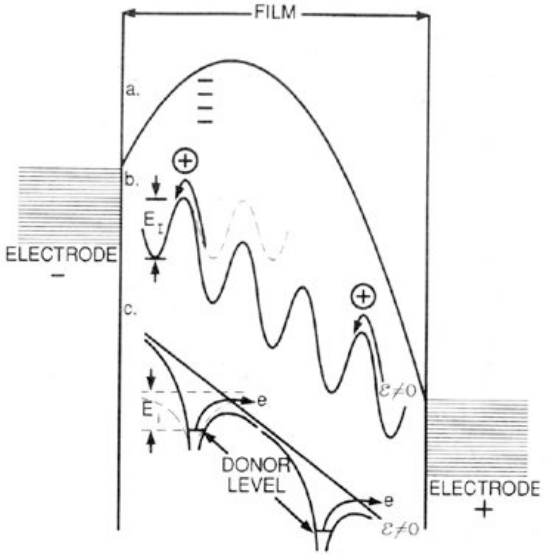


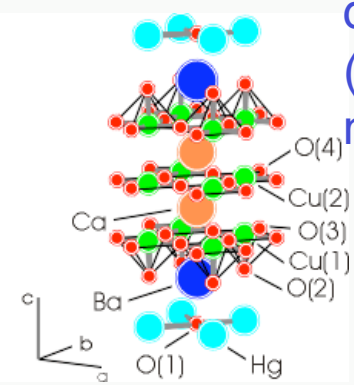
Figure 10-8. Bulk-limited conduction mechanisms. (Dotted lines refer to $\mathscr{E} = 0$) (a space-charge-limited; (b) ionic conduction of cations ⊕; (c) Poole–Frenkel.

General rule of thumbs for transport in thin films

- Processes associated with chemical/structural properties (grains, defects...)
- Processes associated with surface/volume ratio (interface, surface scattering …)

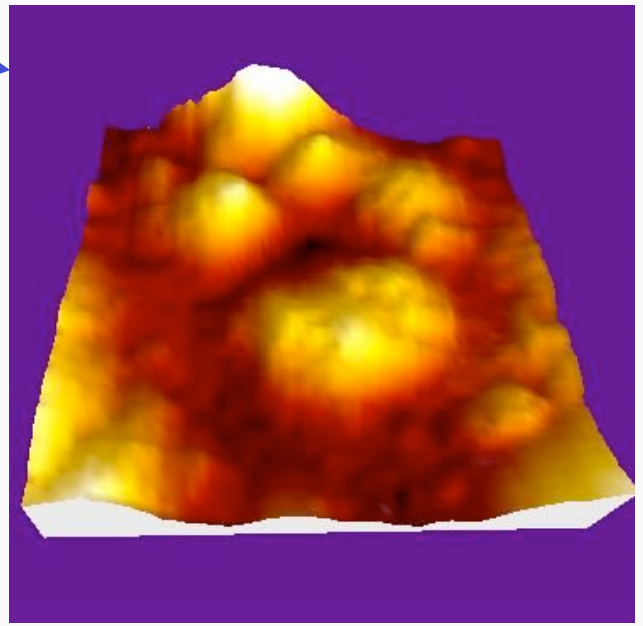
Roughly speaking:
conductors get less conductive; dielectrics get less insulating

Possible exceptions: polycrystalline materials with a complex structure
Example: high-T_C ceramic superconductors



AFM image of YBCO film deposited onto metal subs. (scan size approx 2x2 μm^2, max. height approx 180 nm)

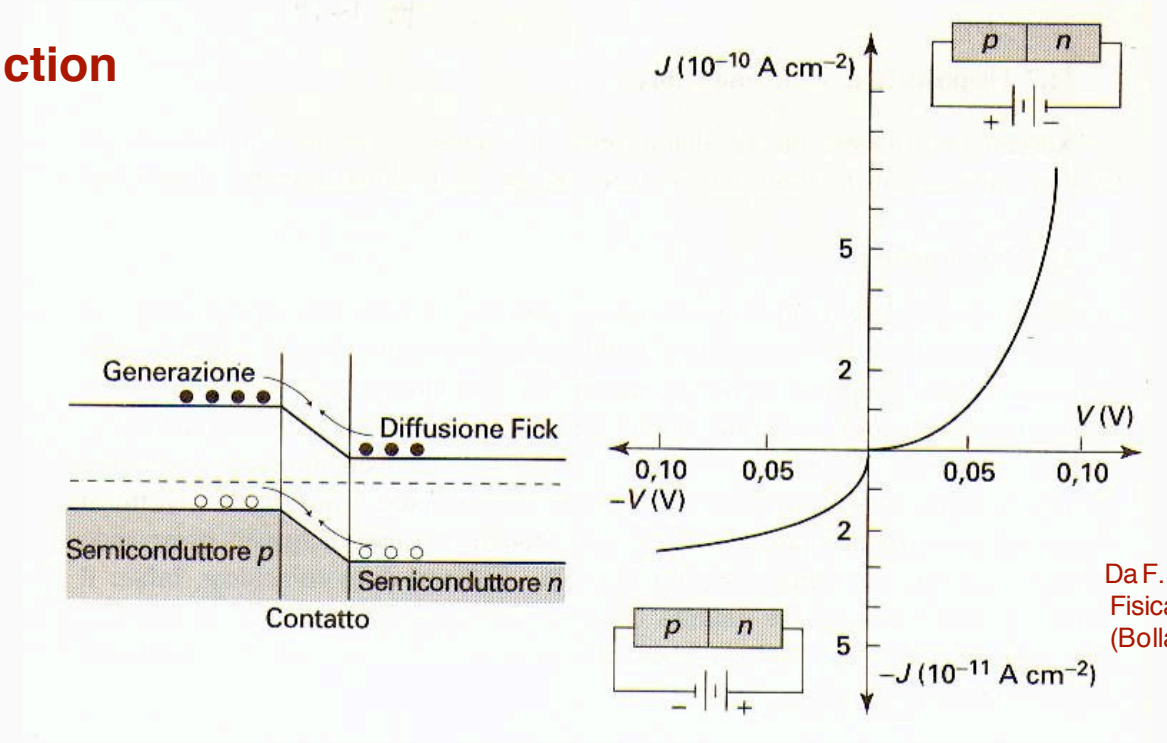
YBCO ($YBa_2Cu_3O_{7-x}$)
$T_c \sim 91$ K (at $x<0.5$)



Superconductivity requires intergrain tunneling which can be favoured in a thin film due to a larger mutual alignment of the grains

Basics of *conventional* electronics II

Bipolar junction



Da F.Bassani, U.M. Grassano,
Fisica dello Stato Solido
(Bollati Boringhieri, 2000)

In the absence of an applied field, p and n charges are redistributed so to create a junction (a charge-free region similar to a capacitor)

⇒ the junction acts as a potential barrier for charges

⇒ transport is possible only when a direct polarization is applied

⇒ a rectifying behaviour is achieved

(note: a similar behaviour occurs also in metal/semiconductor – Schottky – junctions)

An historical look at the bipolar transistor

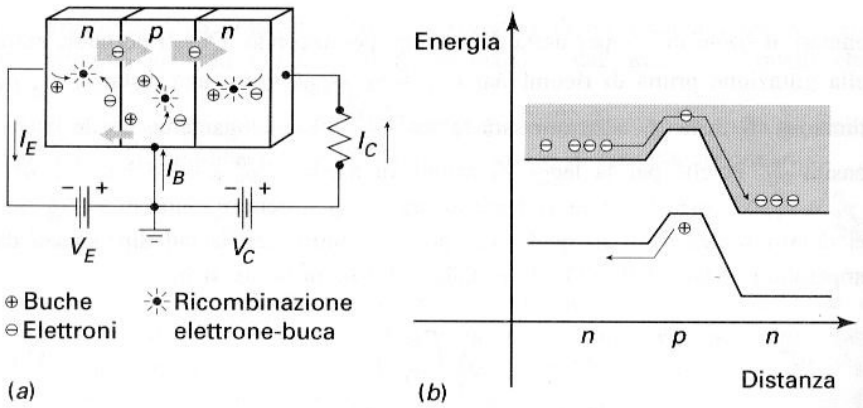


Figura 11.16
Transistor $n-p-n$, con i relativi simboli per indicare la corrente di emettitore (I_E), di collettore (I_C), e la corrente di base (I_B). (*a*) Indicazione degli stati di polarizzazione e dei flussi di corrente (buche ed elettroni). (*b*) Posizionamento delle bande in presenza di un campo (diretto per la giunzione $n-p$).

In un transistor $n-p-n$ si ha ad un estremo l'emettitore di elettroni, i quali entrano dal contatto nel semiconduttore n e all'altro estremo del secondo semiconduttore n vi è il collettore, mentre il semiconduttore p intermedio, molto più sottile degli altri, è chiamato base. All'equilibrio senza polarizzazione non si ha passaggio di corrente perché $I_g^0 = I_r^0$ a entrambe le giunzioni. Basta però applicare una differenza di potenziale tra il collettore e l'emettitore e controllare il potenziale della base per ottenere un'amplificazione di tensione. Illustriamo il funzionamento di tale transistor riferendoci alla fig. 11.16. In questo schema si hanno due circuiti. Uno è il circuito $e-b$ (emettitore-base) che è rettificante per le ragioni esposte precedentemente a proposito del diodo. L'altro è un circuito $b-c$ (base-collettore) che da solo lascerebbe passare poca corrente perché il potenziale è tale da aumentare la barriera di potenziale. In presenza del circuito precedente però molti più elettroni arrivano al semiconduttore p per l'effetto dell'abbassamento della barriera al confine $n-p$ e tali elettroni non trovano ostacoli a proseguire attraverso la zona n ed arrivare al collettore. Questo produce perciò amplificazione di potenza nel circuito $b-c$ rispetto al circuito $e-b$.

A current (e.g., BE) is used to control a current flow (e.g., CE)

Power consumption issues!

La corrente che passa per $n-p$ è I_e:

$$I_e = I_g^0(e^{\frac{eV_e}{kT}} - 1), \tag{11.64}$$

dove V_e è il potenziale dell'emettitore. La corrente che passa al collettore I_c sarà

$$I_c = I_e - I_b, \tag{11.65}$$

dove I_b, corrente di base, è piccola in ogni caso. Se la base è a terra si può ritenere $I_b \simeq 0$ e la corrente raccolta al collettore sarà data dalla (11.64). Non c'è in questo caso amplificazione di corrente tra emettitore e collettore, ma c'è grande amplificazione di tensione (o di potenza), perché la stessa corrente passa da un circuito d'ingresso a bassa impedenza (giunzione con polarizzazione diretta) ad un circuito d'uscita a grande impedenza (giunzione con polarizzazione inversa) e qui scorre attraverso una grande resistenza R_L. Dunque un transistor a base comune si comporta come un amplificatore di tensione (o di potenza).

Se il transistor è collegato con emettitore comune (a terra), si comporta come un amplificatore di corrente (vedi fig. 11.17). Come si è visto prima, quasi tutta la corrente I_e della giunzione emettitore-base (polarizzata direttamente) raggiunge il collettore, così si può scrivere

$$I_c = \alpha I_e \tag{11.66a}$$

Old-style technology

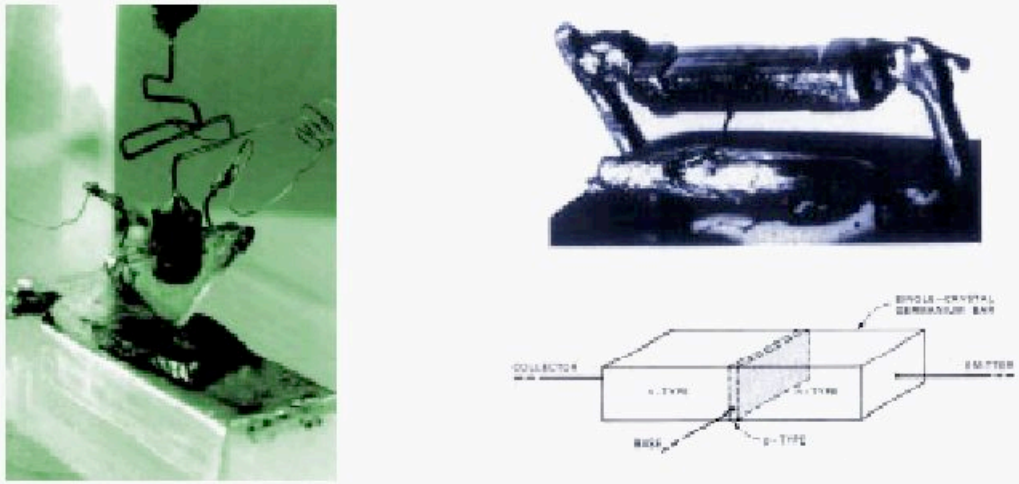


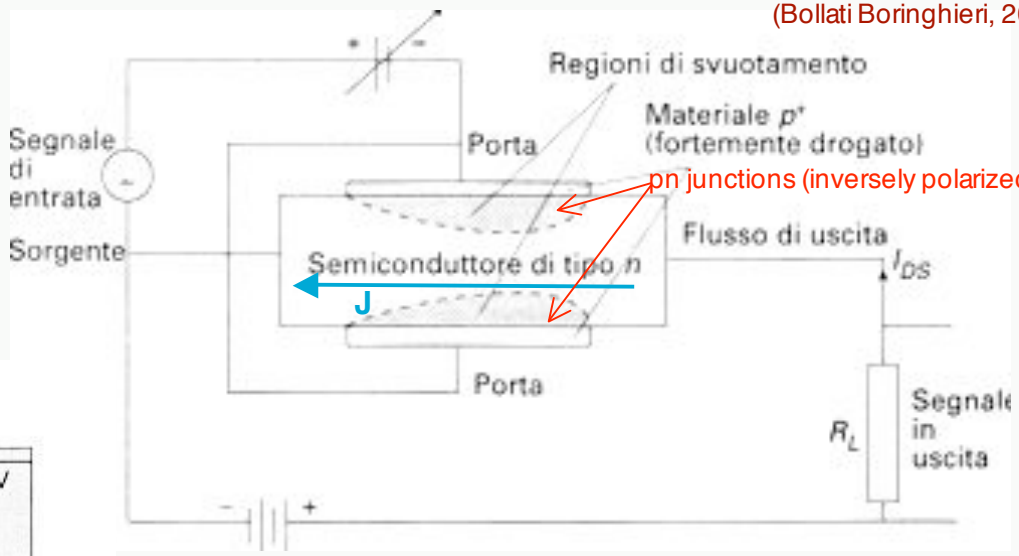
Figure 1 – The first transistors: (a) the point contact transistor of Brattain and Bardeen, 1947 (left); (b) the junction transistor of Shockley, Morgan, Sparks, and Teal, 1950 (right).

"Linear" technology did not allow for miniaturization

Basics of *conventional* electronics III

Da F.Bassani, U.M. Grassano,
Fisica dello Stato Solido
(Bollati Boringhieri, 2000)

Basic concept of a Field-Effect Transistor (FET)



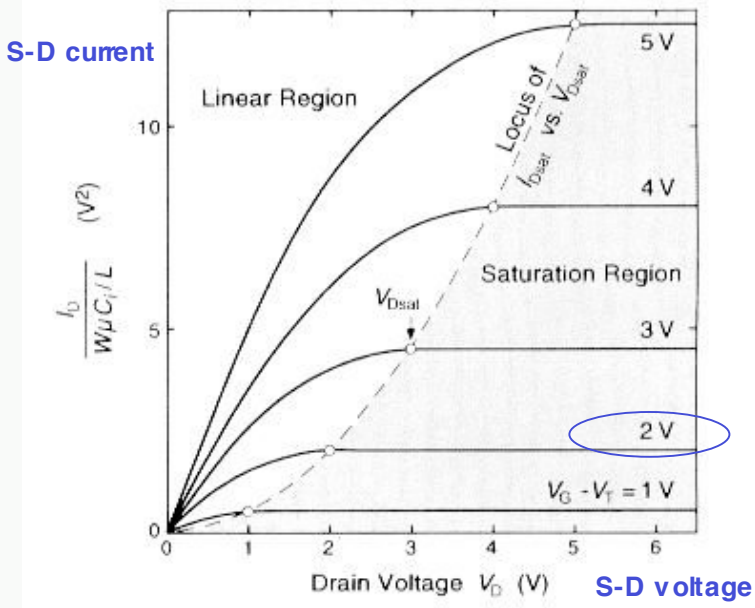


Figure 9: Idealized *I-V* curves of a MOSFET. The dashed line indicates the locus of I_{Dsat} vs. V_{dsat} [6].

The gate voltage is used to control the source-drain current (better behavior in terms of power consumption)

Towards planar technologies ('70s)

Planar technology (thin film multilayers of different materials)

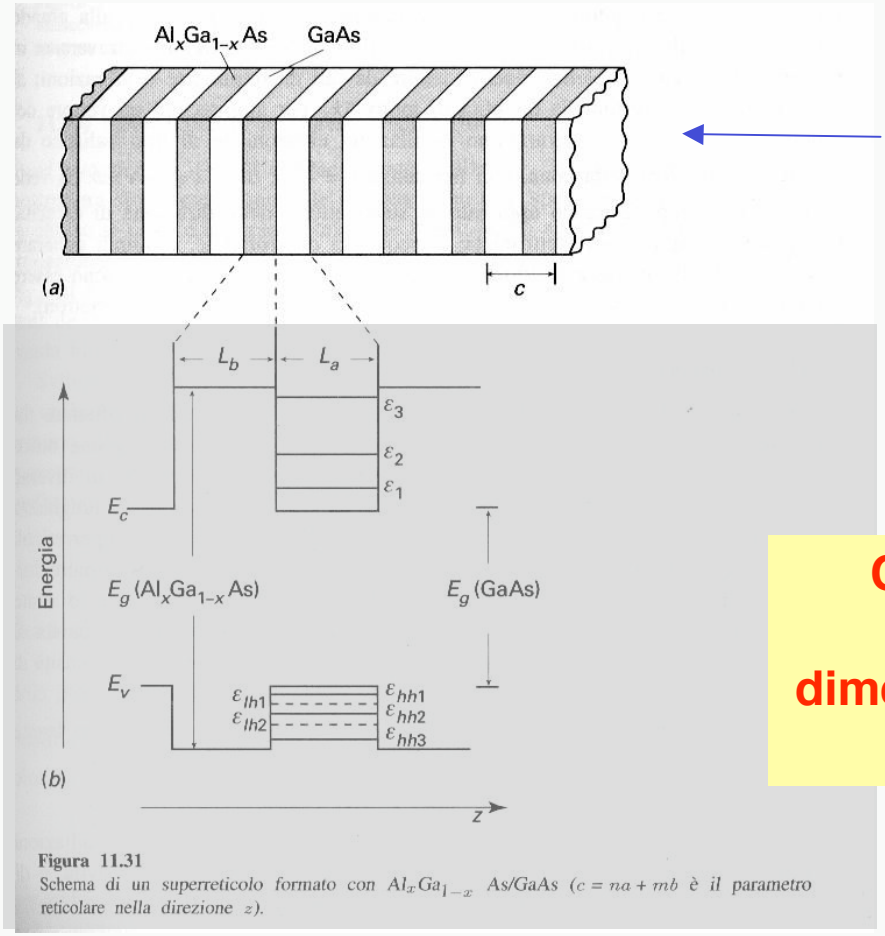


Figura 11.31
Schema di un superreticolo formato con $Al_xGa_{1-x}As/GaAs$ ($c = na + mb$ è il parametro reticolare nella direzione z).

Examples:

Semiconductor heterostructures

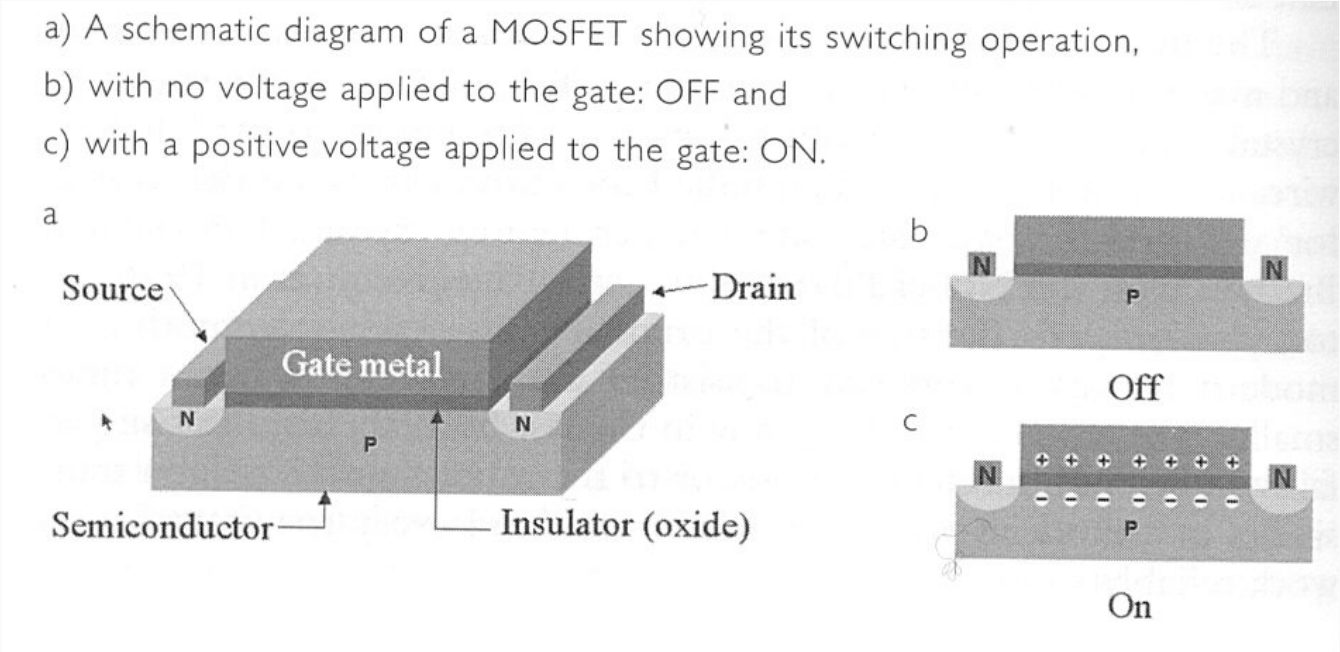
(Dielectric thin films

Thin multilayered structures)

Careful engineering of the material properties achieved by (one-dimensional) thickness control at the few (single!) atomic layer level

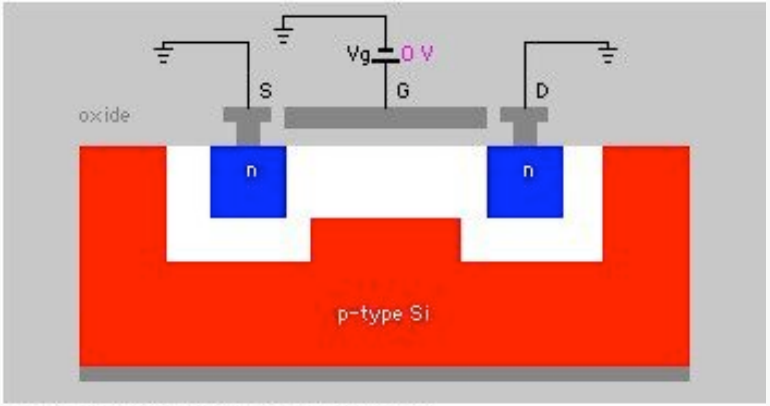
The MOS-FET I

MOS-FET architecture is compatible with planar technologies

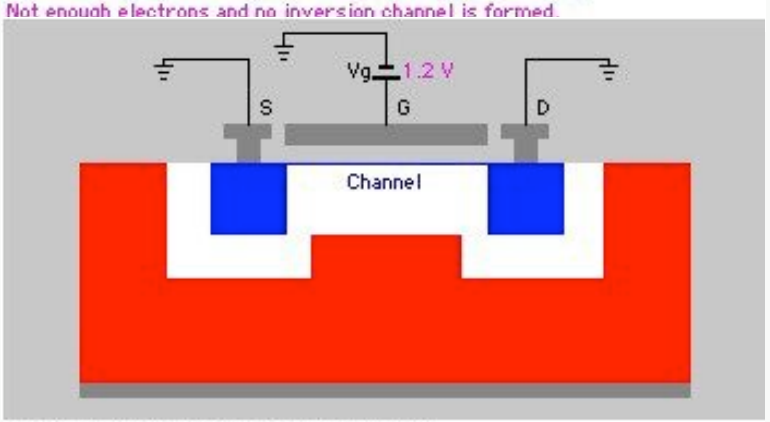


Key points:
- Metal-Oxide-Semiconductor (MOS) multilayer
- Thin dielectric (oxide) layer (E=V/d!!)
- Lateral definition of the structure
- Large scale integration possible

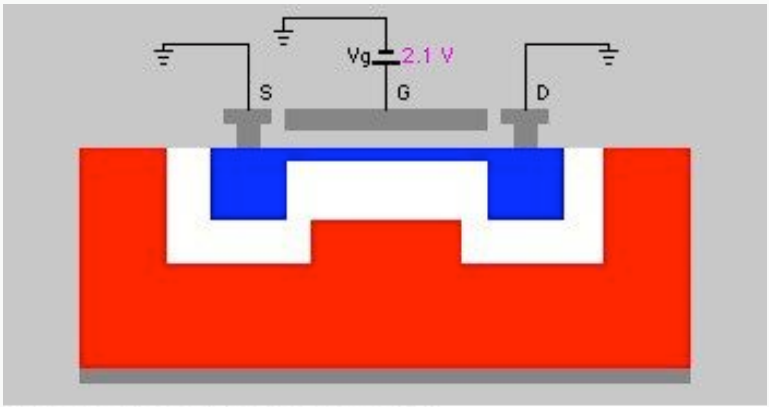
The MOS-FET II



Enhancement-mode (Normally-off) MOSFET
N-channel
Vg < Vt: gate bias is less positive than the threshold voltage.
Not enough electrons and no inversion channel is formed.



Enhancement-mode (Normally-off) MOSFET
N-channel
Vg > Vt: gate bias is more positive than the threshold voltage.
Sufficient electrons accumulate and forms the inversion channel.

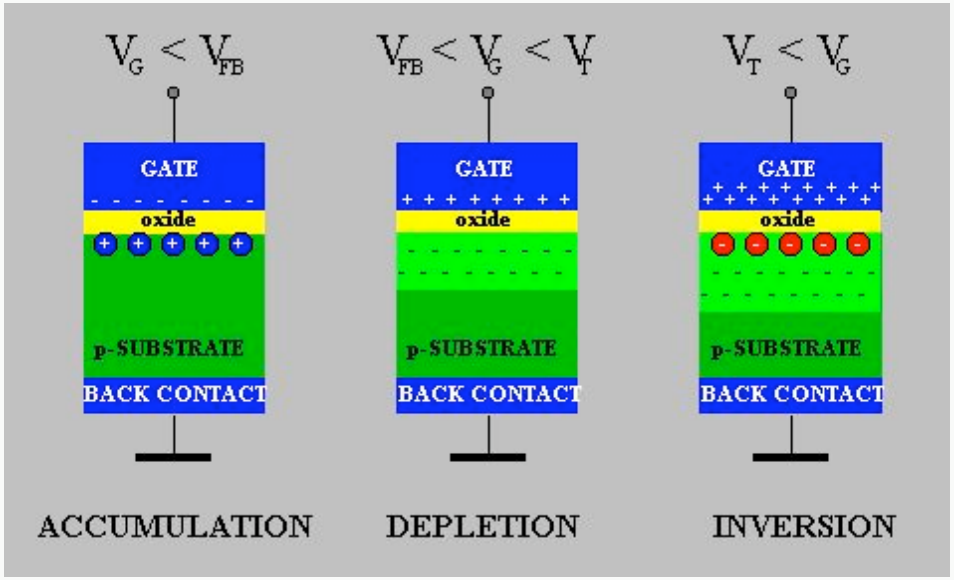


Enhancement-mode (Normally-off) MOSFET
N-channel
Vg > Vt: gate bias is more positive than the threshold voltage.
Sufficient electrons accumulate and forms the inversion channel.

In the inversion and depletion conditions (minority) charges at the interface create a channel for the transport from source to drain

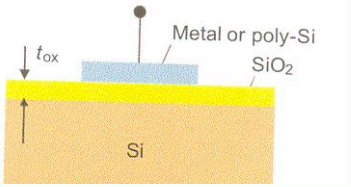
The MOS capacitor I

The electric field produced by applying a voltage to the metal (gate) rules the density of charge carrier at the semiconductor/oxide interface.
Example for p-doped semiconductor:
- accumulation: holes (positive carriers) are accumulated at the interface
- depletion: holes are depleted at the interface
- inversion: a thin layer of *almost free* electrons (negative carriers) is formed



http://people.deas.harvard.edu/~jones/es154/lectures/lecture_4/mosfet/mos_models/mos_cap/mos_cap.html

The MOS capacitor II: energy diagram



Applying an electric field:
- Fermi level is raised/depressed (depending on the sign)
- conductive/valence levels are “deformed” (band bending - depends on space)

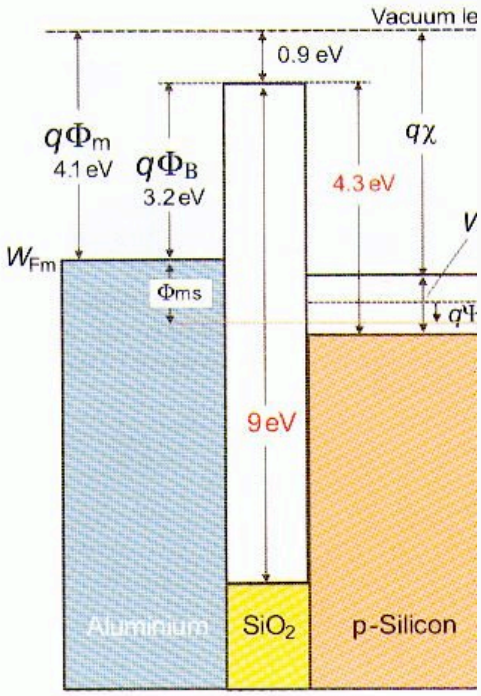
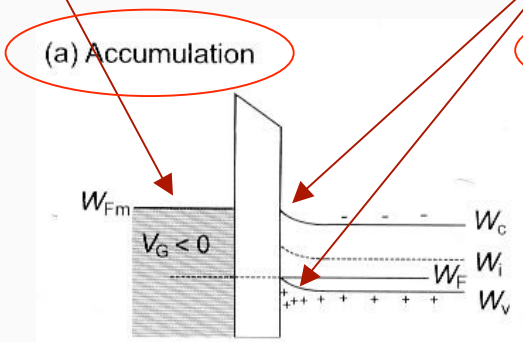
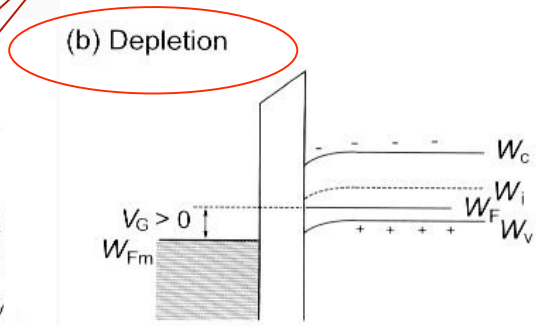
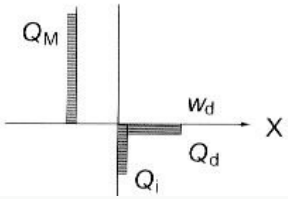
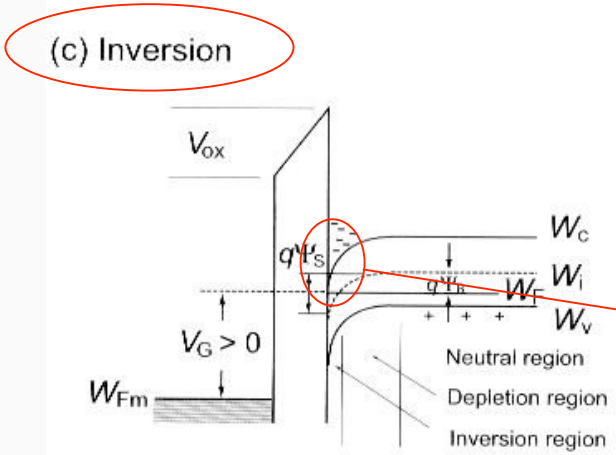


Figure 4: Energy-band diagram of the thi components of a real MOS capacitor, con ing of an Al contact, silicon dioxide and p- silicon. $q\Phi_m$ denotes the work function of metal, $q\Phi_{ms}$ the workfunction difference c versus p-Si, χ the electron affinity of the con, W_g the band energy, W_c the conducti band, W_v the valence band of silicon, $q\Psi_s$ difference between the intrinsic Fermi leve and the Fermi level W_F [5].









The cond. band crosses the “intrinsic” Fermi level ⇒ a thin, dense electron layer is formed

The MOS capacitor III

2.1 MOS Capacitor

Figure 3 shows the structure of a MOS capacitor with the three components, the metal or polysilicon contact, the silicon dioxide with a thickness t_{ox} and the silicon. The corresponding band diagram is shown in Figure 4. Due to the 9 eV bandgap of the silicon dioxide and the large band offsets relative to the silicon, the potential barrier between the conduction band of the silicon and the silicon dioxide is large (≈ 3.1 eV). This barrier crucially controls possible charge transport through the dielectric layer in the presence of an applied voltage, and thus, determines the reliability of the dielectric-semiconductor interface. Frequently poly-Si is used as a contact material instead of a metal. For p-type poly-Si the work function is $\Phi_s = \chi + W_g/2q + \Psi_B \approx 4$ eV, where χ denotes the electron affinity, W_g the band gap energy, Ψ_B the difference between the Fermi potential W_F and the intrinsic potential W_i.

The energy band diagram of an ideal MOS capacitor with a p-type semiconductor is shown in Figure 5 ($q\Phi_{ms}$ is assumed to be zero, see Figure 4). When a negative gate potential $V_G < 0$ is applied the Fermi level of the metal increases and an electric field is created in the SiO_2, indicated by the slope of the conduction band of the SiO_2, and in the silicon. Because of the low carrier concentration the Si bands bend upwards at the SiO_2 interface, leading to an **accumulation** of excess holes. In order to conserve charge, an equivalent number of electrons is accumulated at the metal side of the MOS capacitor.

When a positive potential is applied at the gate contact, its Fermi level moves down leading to band bending in the silicon in the downward direction. As a consequence, the hole concentration near the interface decreases. This status is called the **depletion condition**. Charge neutrality requires the induction of an equivalent amount of positive charge at the metal-oxide interface Q_M as negative charge in the semiconductor Q_S, explicitly,

$$Q = -Q_M \text{ with } Q_S = Q_d \qquad (1$$

where Q_d originates from the ionized donor states. A further increase of the positive gate potential, enhances band bending such that at a certain gate potential the intrinsic Fermi level crosses the Fermi level as shown in Figure 5c. Energetically, it becomes now favourable for electrons to populate the newly created surface channel. The surface behaves like an n-type semiconductor where the doping was created by inverting the original p-type silicon with an applied field. This condition is called **weak inversion** and the corresponding onset gate voltage the threshold voltage V_T. The negative charge at the semiconductor interface Q_S consists of inversion charge Q_i (electrons) and ionized acceptors Q_d (Figure 5d)

$$Q = Q_i + Q_d \qquad (2$$

As indicated in Figure 5c, three regions develop within the semiconductor: a shallow inversion region, a depletion region with a maximum depth w_d and deeper in the substrate a neutral region. A further increase of the potential yields to **strong inversion** when the concentration of the electrons exceeds the hole concentration in the substrate ($Q_i > Q_d$). Then, the gate voltage V_G can be expressed by

$$V_G = V_{ox} + \psi_S = -\frac{Q_S}{C_{ox}} + \psi_S \qquad (3$$

where C_{ox} is the oxide capacitance per unit area and ψ_S is the surface potential, reflected by the band bending in Figure 5c. The surface potential and the total induced charge at the interface can be calculated by solving Poisson's equation with appropriate boundary conditions (see e.g. [6], [7]). Under extreme accumulation and inversion conditions, when V_G and V_{ox} are significantly larger than ψ_S, then Q_S can be approximated by (a)

$$Q = -C_{ox} V_G, \text{ with } C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \qquad (4)$$

since ψ_S is always less than W_g. Eq. (4) implies that the total induced charge at the interface increases with the gate capacitance (per unit area) C_{ox}. ε_{ox} denotes the permittivity and t_{ox} the thickness of the oxide layer.

The total capacitance of the MOS-capacitor C is a series combination of the oxide capacitance C_{ox} and the semiconductor capacitance C_S. Figure 6 shows a capacitance-voltage (C-V) curve for an ideal MOS capacitor at low and high frequencies, as well as under deep depletion conditions. Whereas C_{ox} is basically independent of the gate voltage, the semiconductor capacitance changes, due to the different charge states discussed above. At zero voltage the flat band capacitance C_{FB} is given by (b)

$$C_{FB} = \frac{1}{\frac{t_{ox}}{\varepsilon_{ox}} + \frac{L_D}{\varepsilon_S}} \qquad (5)$$

where L_D is the Debye length and ε_S the silicon permittivity. (For a real capacitor a voltage must be applied to flatten the bands, because $\Phi_{ms} \neq 0$ (see Figure 4)). At negative voltages an accumulation charge builds up with a capacitance $C_S = -dQ_S/d\psi_S$ (Figure 5c). Since ψ_S is limited to 0.1 to 0.3 V in accumulation the total capacitance rapidly reaches its saturation value C_{ox}. A small positive voltage produces a depletion layer which acts as a dielectric with a width w_d in series with the oxide. Thus, the total capacitance C is given by

$$C = \frac{1}{\frac{t_{ox}}{\varepsilon_{ox}} + \frac{w_D}{\varepsilon_S}} \qquad (6)$$

(c)

decreases rapidly to a minimum C_{min}. When the gate voltage reaches the threshold voltage $V_T = 2\psi_S$ an inversion layer starts to form and C increases again. Analogous to Eq. (2), the semiconductor capacitance C_S can be broken up into a depletion charge capaci-

R.Waser (Ed.), *Nanoelectronics and Information Technology* (Wiley-VCH, 2003)

The MOS capacitor IV: capacitance

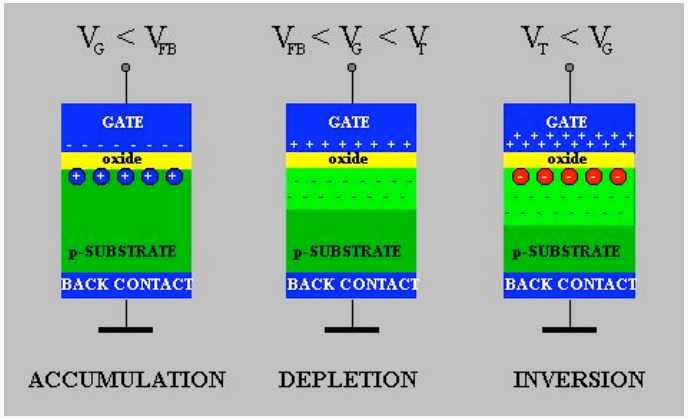
tance C_d and an inversion layer capacitance C_i. C_d and C_i are parallel capacitances in series with C_{ox}, and thus an increase of C_i increases the total capacitance as shown in Figure 6 (curve a). In contrast to the accumulation condition, under the inversion condition the surface potential ψ_S may increase to about 1.0 V. Consequently, the concomitant inversion capacitance C_i can become much larger than the depletion capacitance C_d. Under strong inversion w_d reaches its maximum when the semiconductor is effectively shielded from further penetration of the electric field by the inversion layer. C reaches its maximum value C_{ox}.

If the capacitance measurement is performed at higher frequencies (> 100 Hz), curve (b) in Figure 6 is obtained because the inversion charge arising from minority carriers cannot respond to high frequencies, unless the surface inversion channel is connected to a reservoir of minority carriers as in a MOSFET device. Thus, at high frequencies the inversion charge remains fixed at its dc value and the capacitance does not show an increase at larger V_G.

So far we have discussed only ideal MOS-structures. Real capacitors have undesirable charges within the oxide and at the dielectric/semiconductor interface. These may be mobile ionic charges, like K^+ or Na^+ ions, trapped charges in the SiO_2, fixed charges close to the interface and interface-state charges. Their densities have to be kept at a minimum. C-V measurements are sensitive to such defects, and thus are used to characterize the dielectric layers. Oxide charges will affect the threshold voltage and consequently the performance of the MOSFET. The Si/SiO_2 interface has excellent properties, making silicon the most important semiconductor material. The interface density of the state of the art thermally grown oxides is 2×10^{10} cm^{-2}/eV. However, fundamental limitations will arise when the thickness of the oxide layer becomes so thin that direct tunnelling through the ultrathin silicon oxide causes unacceptable leakage. Alternative gate dielectrics with higher permittivities solve this problem and will be discussed in this chapter.

Total capacitance is a series of the oxide and semiconductor capacitance

It depends on the gate voltage



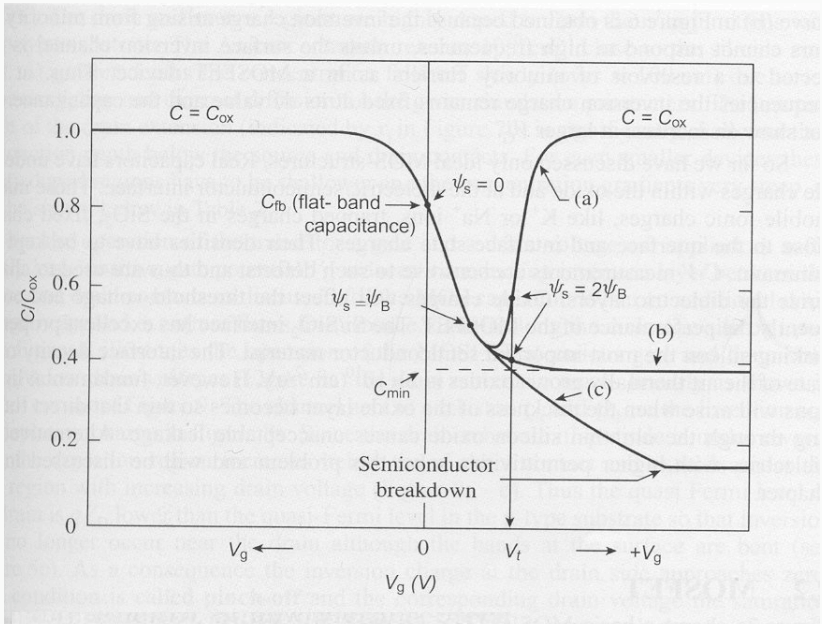


Figure 6: C-V curve of an ideal MOS capacitor under
(a) low frequency,
(b) high frequency and
(c) deep-depletion conditions [6].

The MOS-FET III

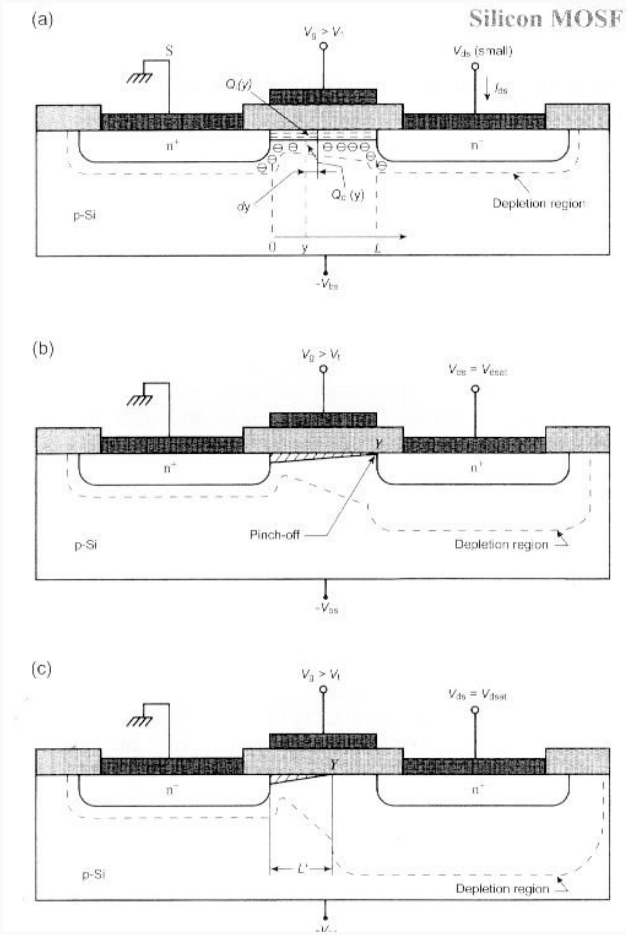
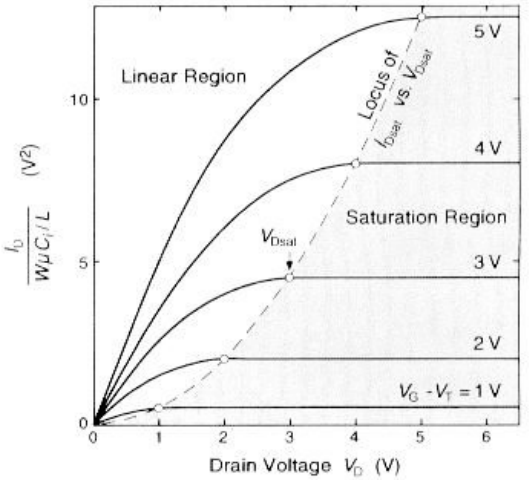


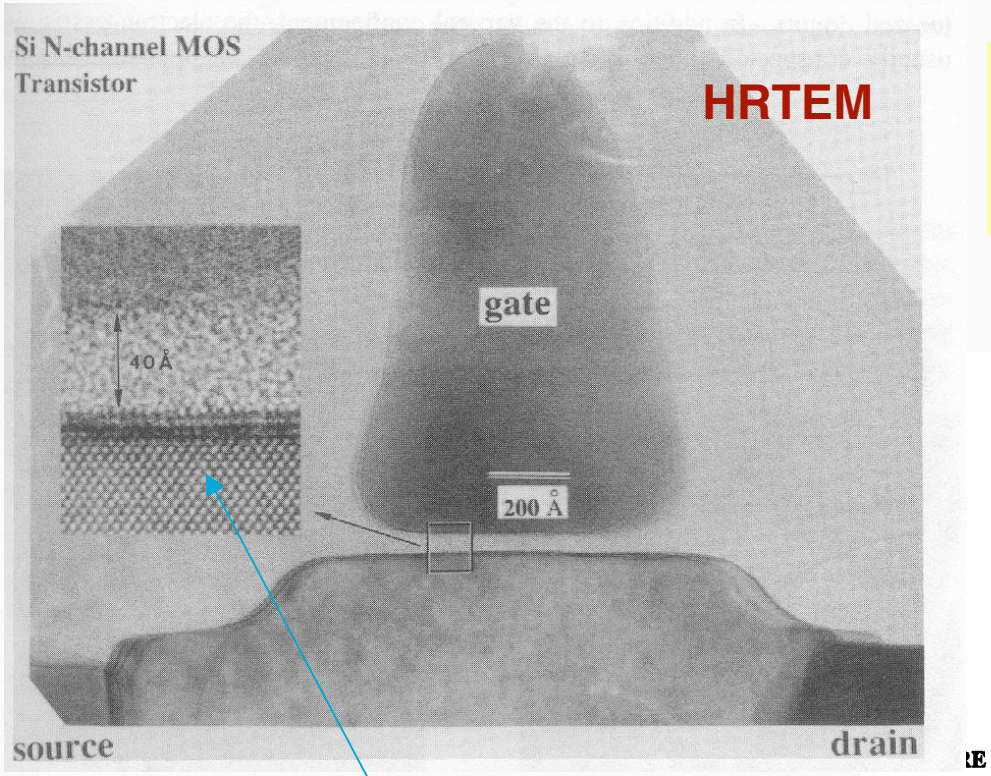
Figure 8: MOSFET operation at a gate voltage $V_g > V_t$ with increasing drain voltage V_{ds}.
(a) At low drain voltages the transistor is in the linear range. Q_i and Q_d are the inversion and the depletion charges, respectively.
(b) shows the pinch-off condition with the pinch-off point Y.
(c) saturation regime where the effective channel length is reduced to L' [6].



For the operation of the transistor a gate and a drain voltage are applied. A sufficiently large positive gate potential V_G induces a conducting inversion layer between the source and drain contacts as discussed for the MOS capacitor. When an additional drain voltage is applied, a current flows from source to drain along the dielectric/silicon interface. Figure 8 illustrates the operation of the MOSFET at various gate and drain voltages. At low drain voltages (Figure 8a) the drain current increases linearly as shown in the *I-V* curves of Figure 9. The channel acts as a resistor. Q_i and Q_d are the inversion and the depletion charge in Figure 8a. Since the drain-substrate n^+-p-diode is under reverse bias, the depletion region increases below the n^+-drain contact and extends under the gate region with increasing drain voltage (Figure 8a - c). Thus the quasi Fermi level of the drain is qV_D lower than the quasi-Fermi level in the p-type substrate so that inversion can no longer occur near the drain although the bands at the surface are bent (see Figure 5c). As a consequence the inversion charge at the drain side approaches zero. This condition is called **pinch-off** and the corresponding drain voltage the saturation drain voltage (Figure 8b). Since the channel resistance is increased, the drain current saturates (saturation region). The pinch-off point, determined by V_{Dsat}, moves towards the source contact with increasing drain voltage. The carriers now drift down the conducting channel and are injected into the surface depletion region at the pinch-off point near the drain (Figure 8c).

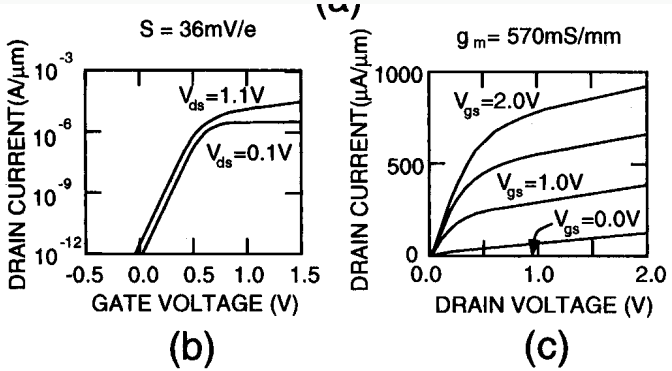
"Saturation" is achieved when no additional minority charges can be created

A “miniaturized” (sub-micron) MOS-FET



Crystalline nature of Si is detected in HRTEM

MOS-FET with nanosized features can be produced (we will see how)



E 6. (a) A high-resolution transmission electron micrograph of a cross-section of an N-MOSFET with a gate length of 0.13 μm adapted from[27] and through the courtesy of Y. Kim. The channel is less than 400 atoms long. The inset shows a lattice image of the channel region of this device. (b) and (c) represent the measured subthreshold and drain characteristics found at room temperature for an N-MOS transistor like that shown in (a). From these measurement it can be inferred that the transconductance is approximately 570 μS/μm, and the subthreshold slope, S = 36 mV per e-fold change in I_D or 84mV/decade, and the threshold voltage is V_t =0.45 V.

presently, down to 70-80 nm

G.Timp (Ed.), Nanotechnology (Springer, 1999)

Miniaturization issues

What happens if we start decreasing dimensions?

1. Reducing the gate (channel) length ⇒ **quantum conf.** (we'll mention them)
2. Reducing the involved amount of charge ⇒ **single electron** (we'll mention them)
3. Reducing thickness of oxide layer ⇒ **materials problems** (we'll mention them)
4. Reducing the overall size ⇒ **nanofabrication problems** (we'll mention them)
5. Specific issue to be mentioned now: **homogeneity and doping**

To be kept in mind:
whole set of issues must be addressed when trying to realize nanosized MOS-FETs

Field distribution and homogeneity

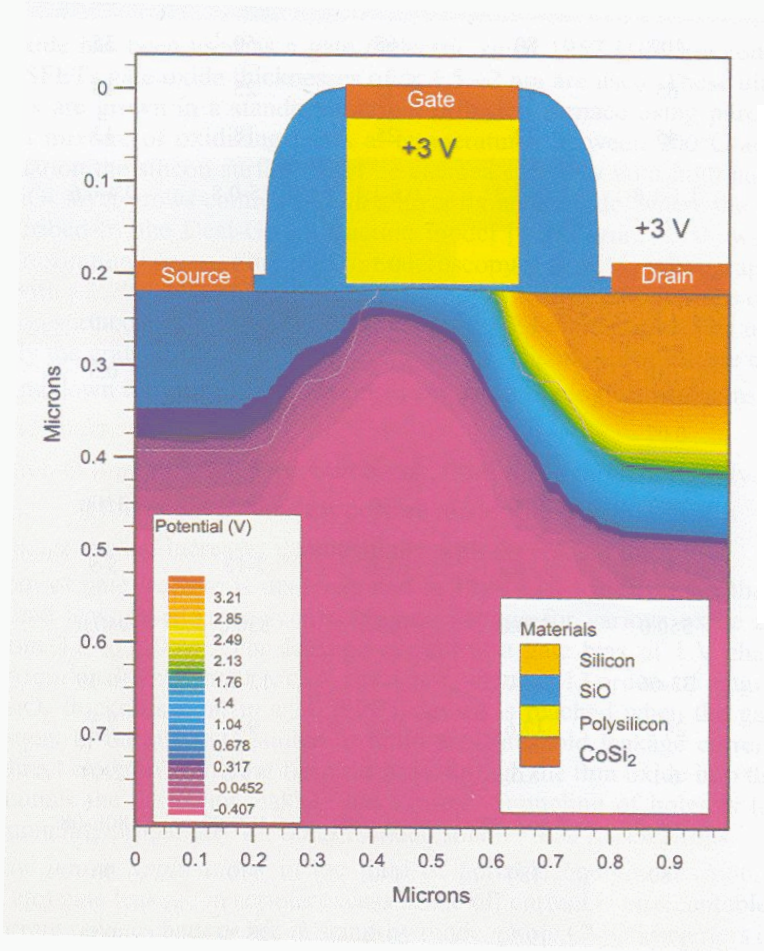
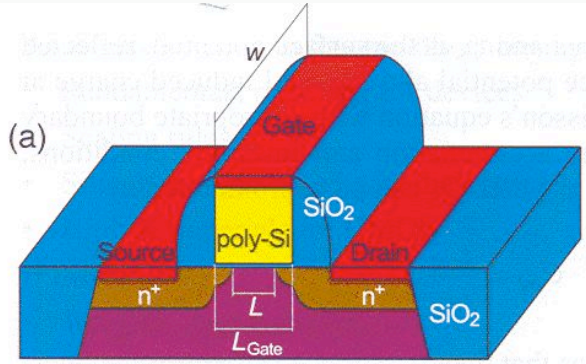


Figure 10: Simulated potential distribution of a 0.2 µm MOSFET with $V_G = 3$ V and $V_D = 3$ V. Near the drain region the potential lines are strongly affected by the drain voltage. The thin solid line indicates the n+/p-junctions.

Relatively large field gradients in small regions

Material modulation doping

Material must be precisely “manipulated” (e.g., doped, defined, …) at a very small scale



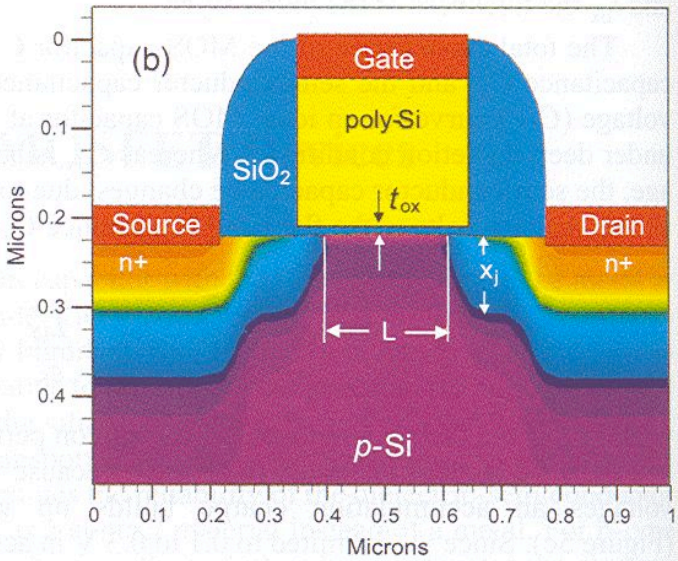


Figure 7:
(a) MOSFET device structure with the terminals, source, gate and drain. L_{Gate} denotes the (printed) gate length, L the channel length or physical gate length and w the gate width. The transistors are isolated with SiO_2 trenches on each side. The gate contact is isolated from source and drain with SiO_2 spacers on each side of the poly-silicon gate contact.
(b) Net doping profiles on a micrometer length and depth scale for a transistor with a gate length $L \cong 0.2$ μm and a gate oxide thickness t_{ox} as calculated with a device simulator (Silvaco). The colours reflect the net dopant concentrations in the Si, ranging from $\approx 10^{17}$ B cm^{-3} in the p-Si to $\approx 10^{20}$ As cm^{-3} near the source/drain silicide contacts. The depth of the n^+/p-junction at the extensions, indicated with x_j, is much shallower than the junction depth below the source and drain contacts (at the blue/dark red boundary).

Need to accurately control the material at the nanometer scale

⇩

(“Technical”) limits in miniaturization

Conclusions

✓Nanotechnology is a wide area cross-related with many scientific and technical fields

✓Electronics is an important driving force and the “transition” from micro- to nanoelectronics is a crucial topic

✓Filling the gap between micro- and nanotechnology has to face:
•Inherent limitations in scaling down the feature size;
•Material limitations due to the small size (e.g., ultra thin films);
•(Fundamental issues associated with quantum confinement)
•(Fundamental problems in the fabrication process)

✓New approaches are required for fabrication

✓New architectures are required for the device operation

✓*New functions can be achieved (as we will see!)*