

LS Scienza dei Materiali - a.a. 2007/08

## Fisica delle Nanotecnologie – part 5.1

Version 6b, Nov 2007

Francesco Fuso, tel 0502214305, 0502214293 - fuso@df.unipi.it

<http://www.df.unipi.it/~fuso/dida>

# Tecniche a scansione di sonda per nanoscopia e nanomanipolazione: STM, AFM e derivati

14/11/2007 – 12-13 – room S1

20/11/2007 – 12-13 – room S1

21/11/2007 – 12-13 – room S1

## Introduction: need for higher resolution tools I

- Due to diffraction, optical methods fail in providing the required space resolution
- Nanometer or even sub-nanometer resolution can be achieved by using electron microscopy (SEM, TEM)

... **BUT** ...

- Contrast mechanism in electron microscopy are often indirect (they imply many effects)
- Morphology can be quantitatively derived only for the in-plane features (poor info on the relative height)
- Samples must be frequently prepared (made conductive, cut in thin slices,...)
- ***Specific physical quantities (e.g., the local density of states, the magnetic or electrical polarization, the surface optical properties,...) cannot be directly measured***

**Ability to measure local (“point”) physical quantities is required to investigate nanotechnology products**

## Introduction: need for higher resolution tools II

### **ON THE OTHER HAND**

- Optical lithography is unable to provide material control at the desired level
- In a conventional context, electron beam lithography (EBL) can be used with excellent space resolution results

### **BUT**

- Both optical and electron lithographies are thought essentially for top-down approaches
- EBL involves accelerated charges, in a process inherently destructive
- Bottoms-up approaches, based, e.g., on nanoparticles, nanotubes, organics, are hardly compatible with the “aggressive” technology which has been developed for inorganics (silicon technologies)
- ***New, more flexible and more “gentle” techniques must be designed to access the full potential of nanotechnologies, at least in the laboratory environment (i.e., not necessarily suited for the industry)***

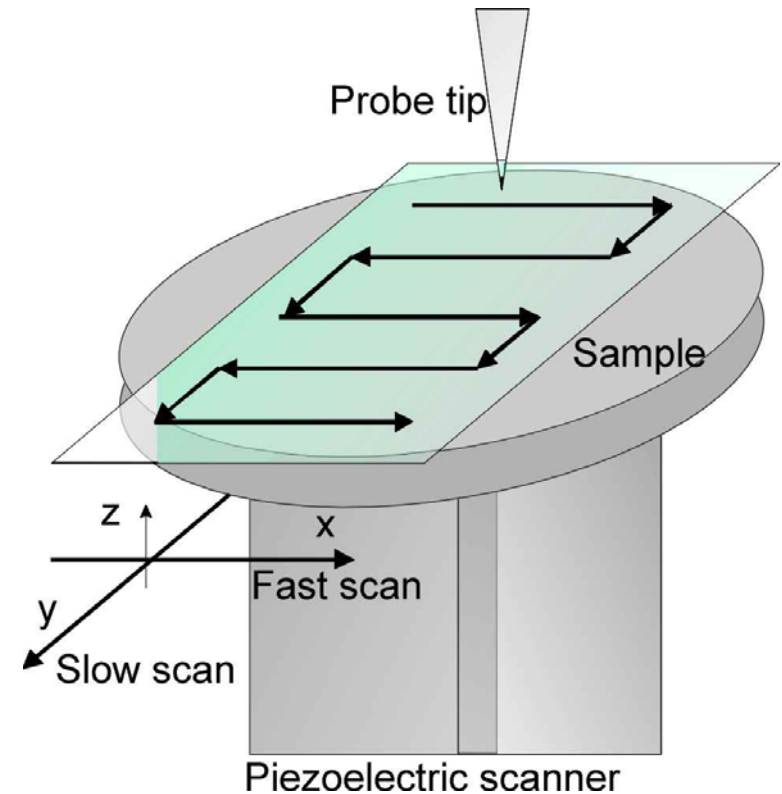
**Ability to manipulate the matter at the nanometer level is required to produce new nanotechnology**

# Basics of Scanning Probe Microscopy (SPM)

**Scanning:** piezoelectric translator  
**Probe:** tip probing local properties  
**Microscopy:** sub-micrometer resolution  
(+ system to control tip/sample distance  
+ electronics for instrument operation)

Developed starting since '80s thanks to:


- ✓ Piezo translators with sub-nm resolution;
- ✓ sub-nm probes



**Various physical quantities can be measured point-by-point during the scan and an image (i.e., a map of the quantity) can be built up**

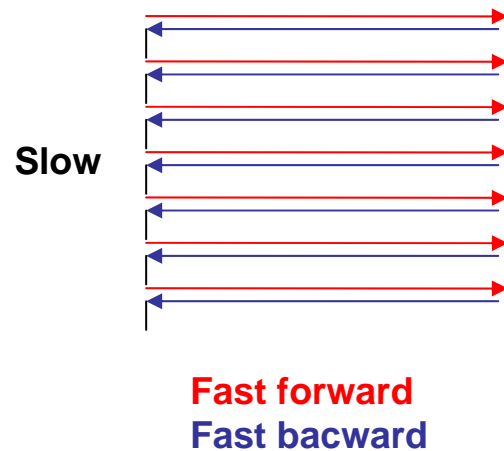
## A few examples of SPMs

Technique	Probed quantity	Resolution
SFM	STM	Electron tunneling
	AFM	Mechanical force
	SFFM	Friction force
	MFM	Local magnetization
	EFM	Local polarization
	SNOM	Optical properties
	Ap-SNOM	Optical properties
...	...	...



Depending on the probe and on its interaction with the surface, a variety of physical effects can be investigated

## A few preliminary considerations on scans



Normally, a **raster scan** is applied:

- Scan addresses an array of discrete, equispaced “pixels” (e.g., 64x64, 128x128, 256x256,...);
- Scan speed is different for one and the other directions (**fast and slow** scans, respectively);
- **Forward** (trace) and **backward** (retrace) scans along the “fast axis” are typically acquired;
- Forward/backward **comparison** is routinely used to assess the scan quality (unless it is used to derive some physical quantity, as in LFM)

The acquisition speeds depends on:

- Time response of the scanner (typically a few ms for nm-sized displacements);
- The signal-to-noise of the quantity to be probed and acquired (through  $\mu\text{s}$  to s depending on the nature of the measurement)

**Rule of thumb: fast acquisition are more suited to higher resolution  
(in order to prevent thermal and mechanical drifts)**

## A few details: piezoelectric scanner

The piezoelectric coefficient  $d$  is negative representing contraction perpendicular to the field; and positive for strain measured along the 3-direction (along which the thickness  $t$  is measured) representing expansion parallel to the electric field direction:

$$\Delta t = d_{33} V .$$

(4.4)

Although there are many ceramic compositions used today, most can be placed into two general categories: hard and soft PZT materials. Typical  $d$  coefficients for hard PZT materials are

$$d_{33} = 250 \cdot 10^{-12} \text{ m/V} , \quad d_{31} = -110 \cdot 10^{-12} \text{ m/V} ;$$

and for soft PZT materials

$$d_{33} = 600 \cdot 10^{-12} \text{ m/V} , \quad d_{31} = -270 \cdot 10^{-12} \text{ m/V} .$$

For PZT-5H

$$d_{33} = 593 \cdot 10^{-12} \text{ m/V} , \quad d_{31} = -273 \cdot 10^{-12} \text{ m/V} .$$

Displacement as small as  $\sim 0.1\text{-}0.5$  nm/V (along Z) are possible

Typical "scanner sensitivity":

$\sim 1\text{-}10$  nm/V (along Z)

$\sim 1\text{-}100$  nm/V (along XY)

Typical driving voltages up to  $\pm 250$  V

Typical min driving step size (16 bit)

$\sim 10$  mV

Typ: hollow tubes made of PZT-based ceramics with a multi-electrode configuration aimed at controlling the displacement along different directions.

Main issues:

- Linearity (possibly closed loop);
- Hysteresis;
- Distorted motion (*artifacts*).

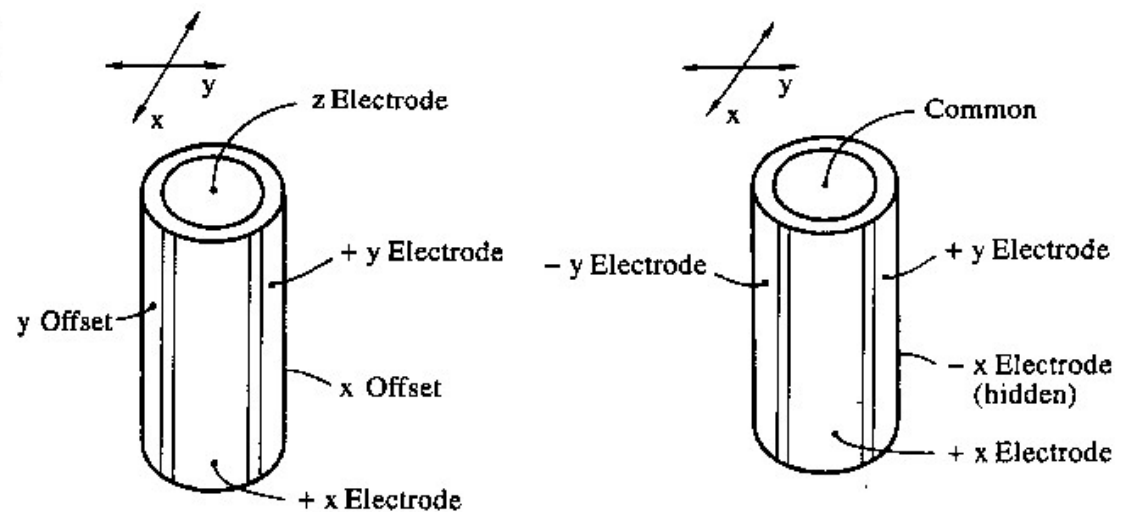


Fig. 4.5. Illustrating the voltages applied to the electrodes of the single-tube scanner

# Scanner-related artifacts I

Artifacts mean the presence of spurious information in an SPM image  
The easiest way to approach the artifact problem is in AFM (we'll see soon how it works!)

## 2.0 Scanner Artifacts

Scanners that move the probe in an atomic force microscope in the X, Y and Z directions are typically made from piezoelectric ceramics. As electromechanical transducers, piezoelectric ceramics are capable of moving a probe very small distances. However, when a linear voltage ramp is applied to piezoelectric ceramics, the ceramics move in a nonlinear motion. Further, the piezoelectric ceramics exhibit hysteresis effects caused by self-heating. Artifacts can also be introduced into images because of the geometry of the scanner. The positioning of the scanner relative to the sample can also create artifacts.

### 2.2. X-Y Calibration/Linearity

All atomic force microscopes must be calibrated in the X-Y axis so that the images presented on the computer screen are accurate. Also the motion of the scanners must be linear so that the distances measured from the images are accurate. With no correction, the features on an image will typically appear smaller on one side of the image than on the other.

## 2.6. Scanner Drift

Drift in AFM images can occur because of thermal drift in the piezoelectric scanner and because an AFM can be susceptible to external temperature changes. The most common type of drift occurs at the beginning of a scan of a zoomed-in region of an image. This artifact causes the initial part of a scan range to appear distorted. Drift artifacts are most easily observed when imaging test patterns. Drift will cause lines that should appear straight to have curvature.

Figure 16

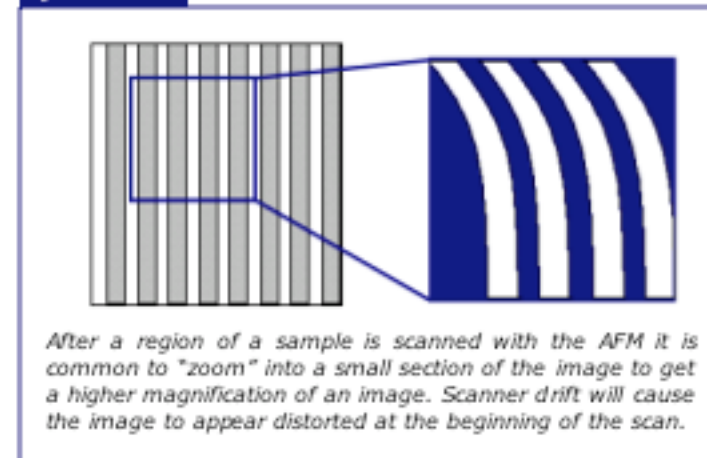
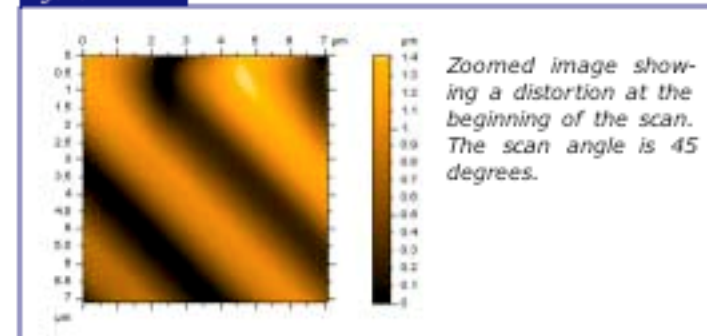


Figure 17





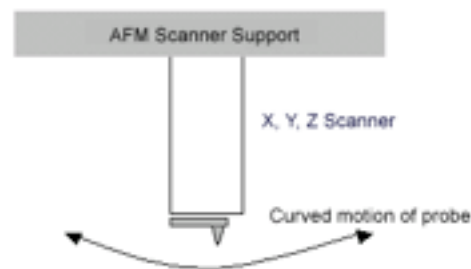
## Scanner-related artifacts II

### 2.4. Background Bow/Tilt

The piezoelectric scanners that move the probe in an atomic force microscope typically move the probe in a curved motion over the surface. The curved motion results in a "Bow" in the AFM image. Also, a large planar background or "Tilt" can be observed if the probe/sample angle is not perpendicular.

Often the images measured by the AFM include a background "Bow" and a background "Tilt" that are larger than the features of interest. In such cases the background must be subtracted from the image. This is often called "leveling" or "flattening" the image. After "leveling" the desired features are typically directly seen in the image.

Figure 12



An AFM piezoelectric scanner is often supported at the top by a mechanical assembly. Thus the motion of the probe is nonlinear in the Z axis as it is scanned across a surface. The motion can be spherical or even parabolic depending on the type of piezoelectric scanner.

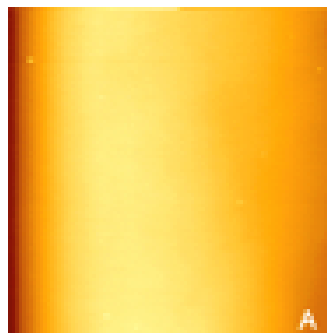


Figure 13A-B: Image (A) is an 85 X 85 micron image of a flat piece of silicon. The bow introduced into the image is seen at the edges. (B) A line profile across this image shows the magnitude of the bow.

(Image) post-processing can help  
(but can also introduce new artifacts, as well...)

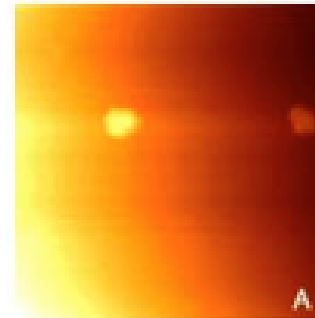
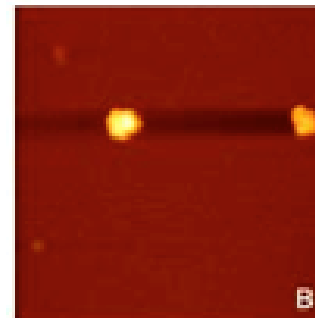
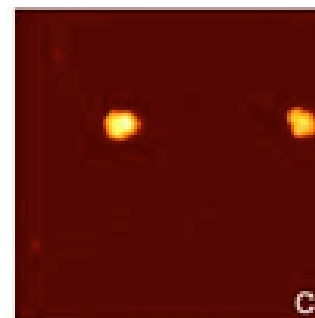


Figure 21A-C: AFM images a 1.6 X 1.6 micron image of nanospheres on a surface.

(A) The original image measured by the AFM before any image processing. Tilt is easily recognized in the image as the right side of the image appears darker than the left side of the image.



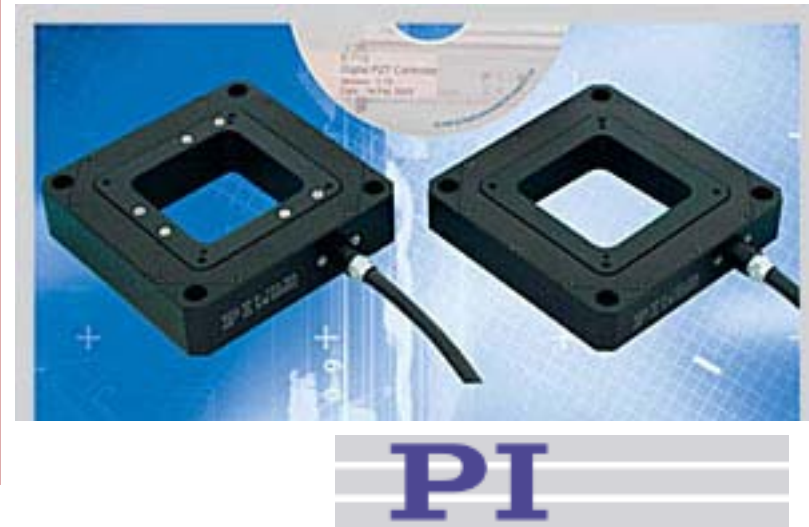
(B) The AFM image shown in "A" after a line-by-line leveling of the image with a first order background correction. The dark band in the image is caused by the image processing and is not a real structure.



(C) Particles are excluded from the background subtraction process to derive this image.

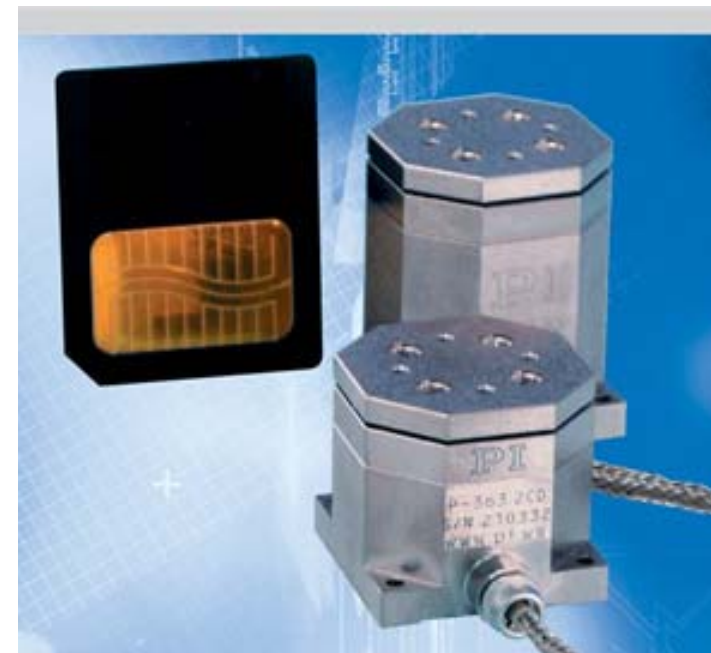
## Closed-loop scanners

- ✓ Problems associated with the scanner geometry can be solved by using different scanner configurations (e.g., non cylindrical)
- ✓ Problems associated with non-linearity, hysteresis, drifts, can be solved by using closed-loop scanners (displacement is independently measured, e.g., by capacitive or interferometric or resistive means, and a loop is applied)



### Excellent performances presently achievable

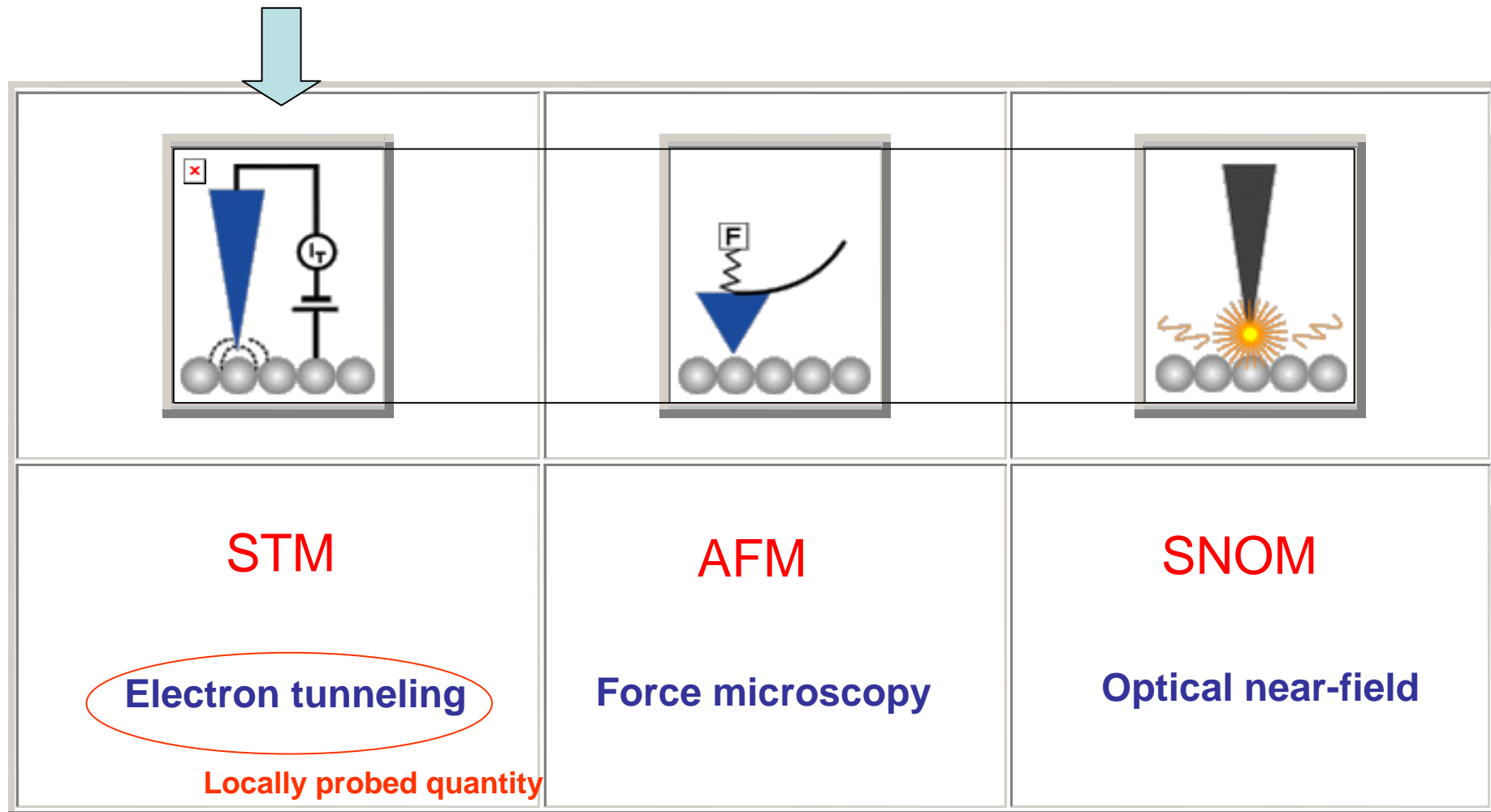
Nanometer Accuracy in 1 Millisecond with 50-Picometer Resolution  
PicoCube® systems provide resolution of 50 picometers and below. The ultra-fast XY/XYZ piezo drives offer resonant frequencies of 9.8 kHz in Z and >3 kHz in X and Y! The high resonant frequency and high-bandwidth capacitive feedback allow step and settle to 1% accuracy in as little as one millisecond.



## Outlook

1. The mother of all SPMs: Scanning Tunneling Microscopy (STM): mechanisms and instruments to investigate local electronic properties
2. SPM based on probing mechanical forces:
  - A. Atomic Force Microscopy (AFM): contact and non contact modes
  - B. Variants (lateral, electrostatic, magnetic forces,...)
3. Sub-diffraction properties of electromagnetic waves:
  - A. Scanning Near Field Optical Microscopy (SNOM);
  - B. Polarization-Modulation SNOM
4. “Lithographies” (better denoted as “nanomanipulations”) associated with:
  - A. STM;
  - B. AFM;
  - C. SNOM

# 1. Scanning Tunneling Microscopy (STM)



*Historically, STM is the first working realization of SPM, and probably the simplest*

# STM tip (probe) preparation

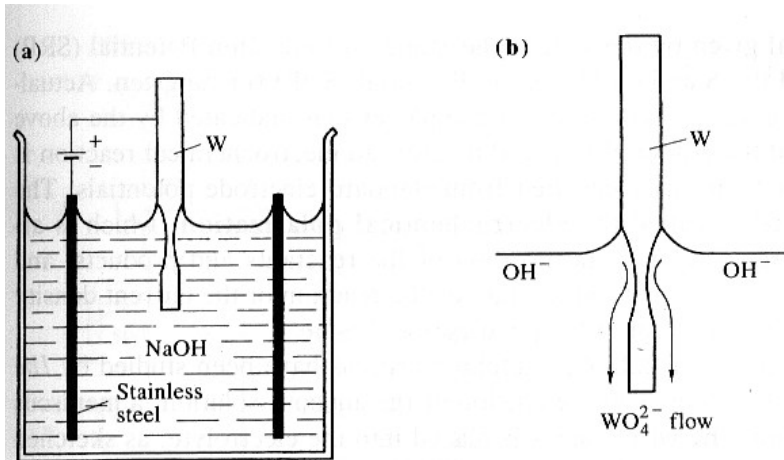
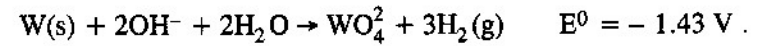
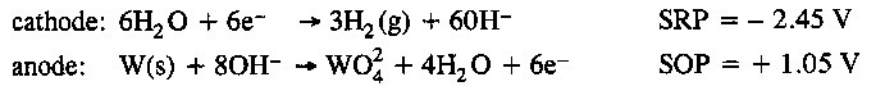
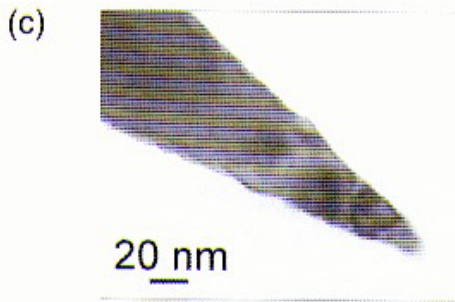
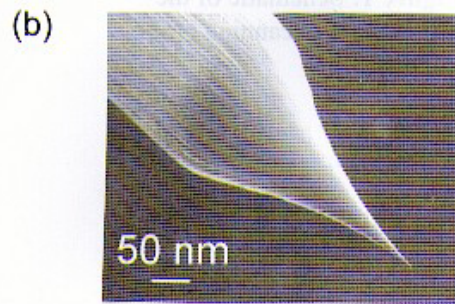


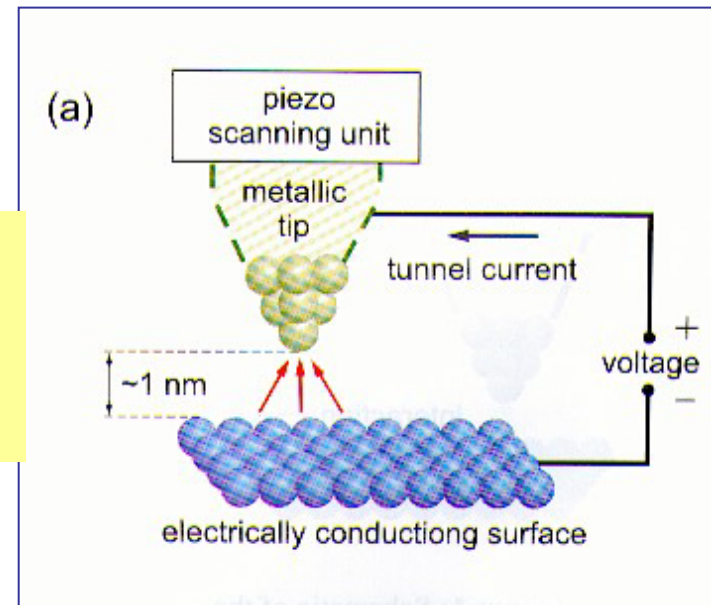
Fig. 4.12. (a) Schematic diagram of the electrochemical cell showing the tungsten wire (anode) being etched in NaOH. The cathode consists of a stainless-steel cylinder which surrounds the anode. (b) Sketch of the etching mechanism showing the "flow" of the tungstate anion down the sides of wire [4.13]



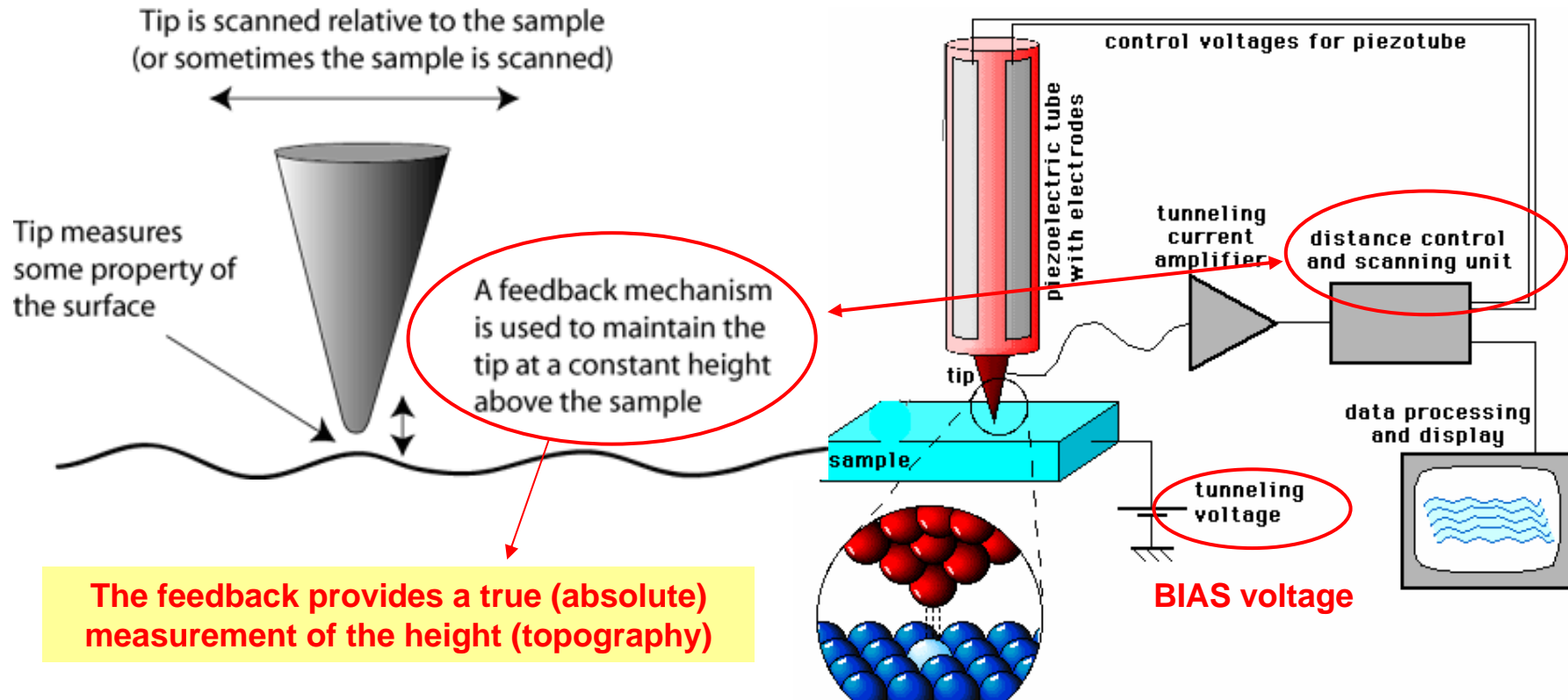
**Electrochemical etching of W or Pt/Ir typically used**



**Very sharp tips can be obtained (ideally, terminated by a "single" atom)**



# Basics of Scanning Tunneling Microscopy I



## STM:

- ✓ Probe is a conductive tip
- ✓ Sample (surface) is mostly conductive or semiconductive
- ✓ A bias voltage is applied between sample and tip
- ✓ Tip is kept at small distance from the surface (typ  $< 1$  nm)
- ✓ Tunneling current (typ in the pA range) is measured

# Basics of Scanning Tunneling Microscopy II

The invention of the scanning tunnelling microscope (STM) in 1981 was preceded by two technological achievements: the invention of the topografiner by Young *et al* in 1972, and the electron tunnelling experiments reported by Teague in his PhD dissertation in 1978 [1]. The topografiner was a device that used a piezoelectric driver to scan a surface whose probe rested approximately 100 nm above the sample surface. Using a field emission current induced by applying a positive voltage of a few kilovolts to the sample, the topografiner could produce a surface image with a lateral resolution of 400 nm [2]. Teague's experiments demonstrated vacuum tunnelling between two gold electrodes, at voltages and tunnelling gap widths used in modern day STMs. Teague's apparatus allowed fine adjustment of the gap width, but did not have a raster scan mechanism [3]. Binnig and Rohrer combined these technologies to create the STM [4]. They used the STM to determine the atomic positions in Si(111)- $7 \times 7$ , one of the most complex surface reconstructions, highlighting the STMs ability to image at the atomic scale [5].

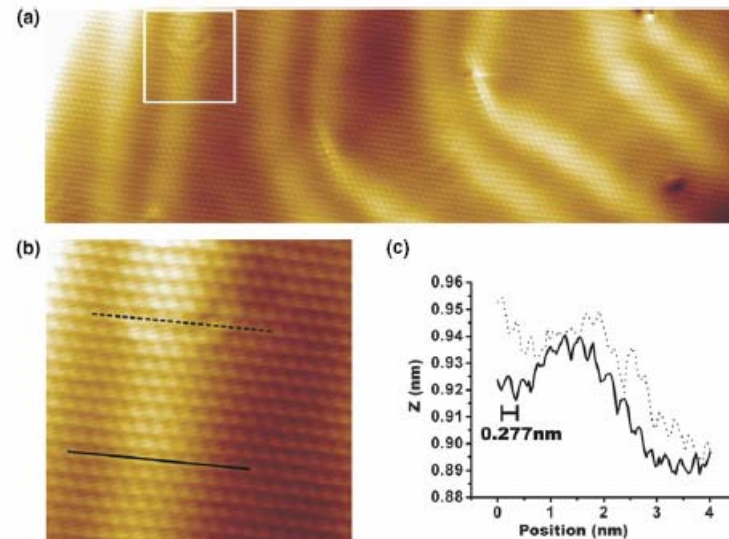


Figure 1. The high-resolution power of a LT-STM can exhibit features of a Au(111) single crystal surface on several scales: the herringbone reconstruction, surface standing waves, and gold atoms. (a) The image of Au(111) ( $9 \text{ nm} \times 30 \text{ nm}$ ) was produced at 6.5 K under tunnelling conditions of 1.1 nA and 10 mV. (b) Zooming into a section of the Au(111) surface ( $5 \text{ nm} \times 5 \text{ nm}$ ), standing waves of electronic density are observed as electrons scatter from an impurity near the top of the surface. (c) Two linescans are drawn in (b) to demonstrate the corrugation and lateral resolution of the STM. Both linescans pass over a soliton wall of the herringbone structure and show the herringbone's large amplitude. The dotted line shows an irregular pattern within the linescan, where standing waves interfere with the atomic corrugation.

**Huge potential in analyzing surfaces with extremely high space resolution**

# Reminders of tunnel effect I

Da Eisberg Resnick, Quantum Physics  
Wiley (1985)

For the barrier potential of (6-45), we know from the qualitative arguments of the last chapter that acceptable solutions to the time-independent Schrodinger equation should exist for all values of the total energy  $E \geq 0$ . We also know that the equation breaks up into three separate equations for the three regions:  $x < 0$  (left of the barrier),  $0 < x < a$  (within the barrier), and  $x > a$  (right of the barrier). In the regions to the left and to the right of the barrier the equations are those for a free particle of total energy  $E$ . Their general solutions are

$$\begin{aligned} \psi(x) &= Ae^{ik_1x} + Be^{-ik_1x} & x < 0 \\ \psi(x) &= Ce^{ik_1x} + De^{-ik_1x} & x > a \end{aligned} \quad (6-46)$$

where

$$k_1 = \frac{\sqrt{2mE}}{\hbar}$$

In the region within the barrier, the form of the equation, and of its general solution, depends on whether  $E < V_0$  or  $E > V_0$ . Both of these cases have been treated in the previous sections. In the first case,  $E < V_0$ , the general solution is

$$\psi(x) = Fe^{-k_{II}x} + Ge^{k_{II}x} \quad 0 < x < a \quad (6-47)$$

where

$$k_{II} = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \quad E < V_0$$

In the second case,  $E > V_0$ , it is

$$\psi(x) = Fe^{ik_{III}x} + Ge^{-ik_{III}x} \quad 0 < x < a \quad (6-48)$$

where

$$k_{III} = \frac{\sqrt{2m(E - V_0)}}{\hbar} \quad E > V_0$$

Note that (6-47) involves real exponentials, whereas (6-46) and (6-48) involve complex exponentials.

Since we are considering the case of a particle incident on the barrier from the left, in the region to the right of the barrier there can be only a transmitted wave as there is nothing in that region to produce a reflection. Thus we can set

$$D = 0$$

In the present situation, however, we cannot set  $G = 0$  in (6-47) since the value of  $x$  is limited in the barrier region,  $0 < x < a$ , so  $\psi(x)$  for  $E < V_0$  cannot become infinitely large even if the increasing exponential is present. Nor can we set  $G = 0$  in (6-48) since  $\psi(x)$  for  $E > V_0$  will have a reflected component in the barrier region that arises from the potential discontinuity at  $x = a$ .

We consider first the case in which the energy of the particle is less than the height of the barrier, i.e., the case:

$$E < V_0$$

In matching  $\psi(x)$  and  $d\psi(x)/dx$  at the points  $x = 0$  and  $x = a$ , four equations in the arbitrary constants  $A, B, C, F$ , and  $G$  will be obtained. These equations can be used to evaluate  $B, C, F$ , and  $G$  in terms of  $A$ . The value of  $A$  determines the amplitude of the eigenfunction, and it can be left arbitrary. The form of the probability density corresponding to the eigenfunction obtained is indicated in Figure 6-14 for a typical situation. In the region  $x > a$  the wave function is a pure traveling wave and so the probability density is constant, as for  $x > 0$  in Figure 6-10. In the region  $x < 0$  the wave function is principally a standing wave but has a small traveling wave component because the reflected traveling wave has an amplitude less than that of the

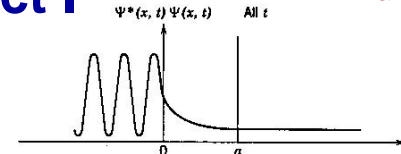


Figure 6-14 The probability density function  $\Psi^*\Psi$  for a typical barrier penetration situation.

incident wave. So the probability density in that region oscillates but has minimum values somewhat greater than zero, as for  $x < 0$  in Figure 6-10. In the region  $0 < x < a$  the wave function has components of both types, but it is principally a standing wave of exponentially decreasing amplitude, and this behavior can be seen in the behavior of the probability density in the region.

The most interesting result of the calculation is the ratio  $T$ , of the probability flux transmitted through the barrier into the region  $x > a$ , to the probability flux incident upon the barrier. This transmission coefficient is found to be

$$T = \frac{v_1 C^* C}{v_1 A^* A} = \left[ 1 + \frac{(e^{k_{II}a} - e^{-k_{II}a})^2}{16 \frac{E}{V_0} \left(1 - \frac{E}{V_0}\right)} \right]^{-1} = \left[ 1 + \frac{\sinh^2 k_{II}a}{4 \frac{E}{V_0} \left(1 - \frac{E}{V_0}\right)} \right]^{-1} \quad (6-49)$$

where

$$k_{II}a = \sqrt{\frac{2mV_0a^2}{\hbar^2} \left(1 - \frac{E}{V_0}\right)} \quad E < V_0$$

If the exponents are very large, this formula reduces to

$$T \approx 16 \frac{E}{V_0} \left(1 - \frac{E}{V_0}\right) e^{-2k_{II}a} \quad k_{II}a \gg 1 \quad (6-50)$$

as can be verified with ease. When (6-50) is a good approximation,  $T$  is extremely small.

These equations make a prediction which is, from the point of view of classical mechanics, very remarkable. They say that a particle of mass  $m$  and total energy  $E$ , incident on a potential barrier of height  $V_0 > E$  and finite thickness  $a$ , actually has a certain probability  $T$  of penetrating the barrier and appearing on the other side. This phenomenon is called barrier penetration, and the particle is said to tunnel through the barrier. Of course,  $T$  is vanishingly small in the classical limit because in that limit the quantity  $2mV_0a^2/\hbar^2$ , which is a measure of the opacity of the barrier, is extremely large.

We shall discuss barrier penetration in detail shortly, but let us first finish describing the calculations by considering the case in which the energy of the particle is greater than the height of the barrier, i.e., the case:

$$E > V_0$$

In this case the eigenfunction is oscillatory in all three regions, but of longer wavelength in the barrier region,  $0 < x < a$ . Evaluation of the constants  $B, C, F$ , and  $G$  by application of the continuity conditions at  $x = 0$  and  $x = a$ , leads to the following formula for the transmission coefficient

$$T = \frac{v_1 C^* C}{v_1 A^* A} = \left[ 1 - \frac{(e^{ik_{III}a} - e^{-ik_{III}a})^2}{16 \frac{E}{V_0} \left(\frac{E}{V_0} - 1\right)} \right]^{-1} = \left[ 1 + \frac{\sin^2 k_{III}a}{4 \frac{E}{V_0} \left(\frac{E}{V_0} - 1\right)} \right]^{-1} \quad (6-51)$$

where

$$k_{III}a = \sqrt{\frac{2mV_0a^2}{\hbar^2} \left(\frac{E}{V_0} - 1\right)} \quad E > V_0$$

Depends on  $E/V_0$  and  $a$



# Reminders of tunnel effect II

Da Eisberg Resnick, Quantum Physics Wiley (1985)

**Table 6-2.** A Summary of the Systems Studied in Chapter 6

Name of System	Physical Example	Potential and Total Energies	Probability Density	Significant Feature
Zero potential	Proton in beam from cyclotron			Results used for other systems
Step potential (energy below top)	Conduction electron near surface of metal			Penetration of excluded region
Step potential (energy above top)	Neutron trying to escape nucleus			Partial reflection at potential discontinuity
Barrier potential (energy below top)	alpha particle trying to escape Coloumb barrier			Tunneling
Barrier potential (energy above top)	Electron scattering from negatively ionized atom			No reflection at certain energies
Finite square well potential	Neutron bound in nucleus			Energy quantization
Infinite square well potential	Molecule strictly confined to box			Approximation to finite square well
Simple harmonic oscillator potential	Atom of vibrating diatomic molecule			Zero-point energy



**Single barrier tunneling**

**(Single barrier) tunneling of electrons through tip and surface (or vice versa, depending on bias polarity) is the key point of STM**

# More on tunneling current I

## Tunneling Effect in Quasiclassical Approximation

The quasiclassical qualitative condition imply that de Broglie wavelength of the particle  $\lambda$  is less than characteristic length  $L$  determining the conditions of the problem. This condition means that the particle wavelength should not change considerably within the length of the wavelength order

$$\left| \frac{d\tilde{\lambda}}{dz} \right| \ll 1 \quad (1)$$

where  $\tilde{\lambda} = \lambda / 2\pi$ ,  $\lambda(z) = 2\pi\hbar / p(z)$  – de Broglie wavelength of the particle expressed by way of the particle classical momentum  $p(z)$  [1].

Condition (1) can be expressed in another form taking into account that

$$\frac{dp}{dz} = \frac{d}{dz} \sqrt{2m(W - U)} = -\frac{m}{p} \frac{dU}{dz} = \frac{mF}{p} \quad (2)$$

where  $F = -dU/dz$  means classical force acting upon the particle in the external field. Introducing this force, we get

$$\frac{m\hbar|F|}{p^3} \ll 1 \quad (3)$$

From (3) it is clear that the quasiclassical approximation is not valid at too small momentum of the particle. In particular, it is deliberately invalid near positions in which the particle, according to classical mechanics, should stop, then start moving in the opposite direction. These points are the so called "turning points". Their coordinates  $z_1$  and  $z_2$  are determined from the condition  $E=U(z)$ .

It should be emphasized that condition (3) itself can be insufficient for the permissibility of the quasiclassical approach. One more condition should be met: the barrier height should not change much over the length  $L$ .

Let us consider the particles move in the field shown in Fig. 1 which is characterized by the presence of the potential barrier with potential energy  $U(z)$  exceeding particle total energy  $E$  and meeting all the quasiclassics conditions. In this case points  $z_1$  and  $z_2$  are the turning points.

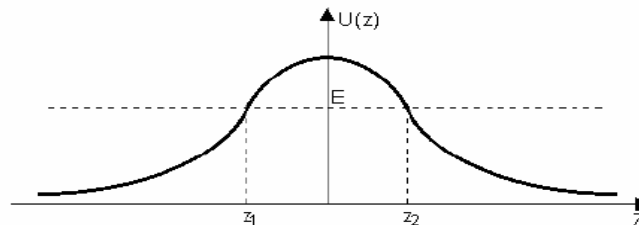


Fig. 1. Potential barrier of arbitrary shape

The approximation technique of the Schrodinger equation solution when quasiclassical conditions are met was first used by Wentzel, Kramers and Brillouin. This technique is known as WKB approximation or quasiclassical quantization method. In this textbook we do not present the Schrodinger equations solution for the given case. However, it can be found in [1], [2] and the barrier transparency in this case is

$$D(E) \propto \exp \left\{ -\frac{2}{\hbar} \int_{z_1}^{z_2} \sqrt{2m(U(z) - E)} dz \right\} \quad (4)$$

Comparing expressions (3) in chapter [Tunneling Effect](#) for transmission coefficients of rectangular barrier (precise solution of Shrodinger equation) and (4) for quasiclassical approximation, we can notice that there is no qualitative difference between them. In both cases the transparency decreases exponentially with the barrier width.

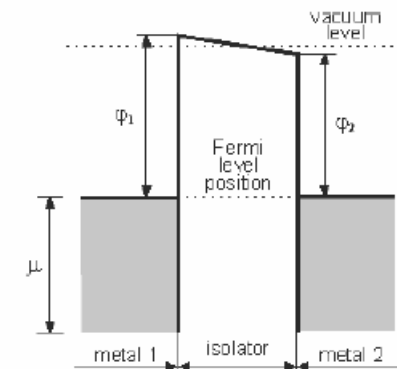


Fig. 1. Diagram of MIM system in equilibrium.  $\phi_1$  and  $\phi_2$  – work function of the left and right metals, respectively

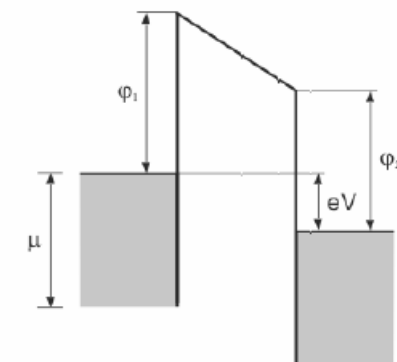


Fig. 2. Model of MIM system then positive potential is applied to the right metal

## Effect of bias voltage (asymmetric barrier)

# The MIM junction I

JOHN G. SIMMONS

**Asymmetric  
(arbitrary)  
barrier**

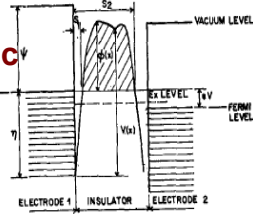


Fig. 1. General barrier in insulating film between two metal electrodes.

The purpose of this paper is to derive a single theory for the current flow through a generalized barrier. The theory is applied to the problem of the rectangular barrier, as studied by Sommerfeld and Bethe,<sup>2</sup> and by Holm.<sup>3</sup> The generalized theory permits derivation of more accurate expressions for the high and low voltages, as well as resolution of the anomalies associated with the formulae derived by Holm for intermediate voltages.

Finally, a method is described for the application of the theory to a practical barrier—that is, to a rectangular barrier with the image potential included. The hyperbolic form of the image potential is used in the generalized formula, thus eliminating the need to resort to a parabolic approximation. The result is a more accurate theoretical current-voltage relationship for a tunnel junction.

## NOTATION

- $m$  = mass of electron,
- $e$  = charge of electron,
- $h$  = Planck's constant,
- $s$  = thickness of insulating film,
- $s_1, s_2$  = limits of barrier at Fermi level,
- $\Delta s = s_2 - s_1$ ,
- $J$  = tunnel current density,
- $V$  = voltage across film,
- $V_i$  = image potential,
- $\eta$  = Fermi level,
- $f(E)$  = Fermi-Dirac function,
- $\psi$  = work function of metal electrode,
- $\phi_0$  = height of rectangular barrier,
- $\bar{\phi}$  = mean barrier height,
- $\epsilon$  = permittivity of insulating film,
- $K$  = dielectric constant, and
- $\sigma$  = tunnel resistivity ( $\Omega\text{-cm}^2$ ).

## THE TUNNEL EQUATION

When two metallic electrodes are separated by an insulating film, the equilibrium conditions require that the top of the energy gap of the insulator be positioned above the Fermi level of the electrodes. Thus, the action of the insulating film is to introduce a potential barrier between the electrodes which impedes the flow of electrons between the electrodes.

The electronic current can flow through the insulating region between the two electrodes if: (a) The electrons in the electrodes have enough thermal energy to surmount the potential barrier and flow in the conduction band. (b) The barrier is thin enough to permit its penetration by the electric tunnel effect.

Sommerfeld and Bethe, and Holm conducted analyses of these conditions for low temperatures so that thermal current could be neglected, thus restricting the electron transport between electrodes to the tunnel effect; a similar procedure is followed in this paper.

The probability  $D(E_x)$  that an electron can penetrate a potential barrier of height  $V(x)$ —the barrier is assumed to be in the  $x$  direction, as shown in Fig. 1—is given by the well-known WKB approximation<sup>4</sup>:

$$D(E_x) = \exp\left\{-\frac{4\pi}{h} \int_{s_1}^{s_2} [2m(V(x) - E_x)]^{1/2} dx\right\}, \quad (1)$$

where  $E_x = mv_x^2/2$ , and is the energy component of the incident electron in the  $x$  direction. The number  $N_1$  of electrons tunneling through the barrier from electrode 1 to electrode 2 is given by

$$N_1 = \int_0^{E_m} v_x n(v_x) D(E_x) dv_x = -\int_0^{E_m} \frac{1}{m} n(v_x) D(E_x) dE_x, \quad (2)$$

where  $E_m$  is the maximum energy of the electrons in the electrode, and  $n(v_x) dv_x$  is the number of electrons per unit volume with velocity between  $v_x$  and  $v_x + dv_x$ . For an isotropic velocity distribution, which is assumed to exist here inside the electrodes, the number of electrons per unit volume with velocity between the usual infinitesimal limits is given by

$$n(v) dv_x dv_y dv_z = (2m^4/h^3) f(E) dv_x dv_y dv_z, \quad (3)$$

where  $f(E)$  is the Fermi-Dirac distribution function. Consequently, from Eq. (3),

$$n(v_x) = \frac{2m^4}{h^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(E) dv_y dv_z = \frac{4\pi m^3}{h^3} \int_0^{E_m} f(E) dE_x. \quad (4)$$

In Eq. (4), the integrand is expressed in polar coordinates; that is,

$$v^2 = v_y^2 + v_z^2, \\ E_x = mv_x^2/2.$$

Substituting Eq. (4) in Eq. (2) yields

$$N_1 = \frac{4\pi m^3}{h^3} \int_0^{E_m} D(E_x) dE_x \int_0^{E_m} f(E) dE_x. \quad (5)$$

<sup>4</sup> D. Bohm, *Quantum Theory* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1951), p. 275.

The number  $N_2$  of electrons tunneling from electrode 2 to electrode 1 is determined in a similar manner. The tunnel probability  $D(E_x)$  is the same in either direction, and if electrode 2 is at a positive potential  $V$ , with respect to electrode 1, the Fermi-Dirac function is written as  $f(E + eV)$ ; therefore,

$$N_2 = \frac{4\pi m^3}{h^3} \int_0^{E_m} D(E_x) dE_x \int_0^{E_m} f(E + eV) dE_x. \quad (6)$$

The net flow of electrons  $N (= N_1 - N_2)$  through the barrier is

$$N = \int_0^{E_m} D(E_x) dE_x \times \left\{ \frac{4\pi m^3}{h^3} \int_0^{E_m} [f(E) - f(E + eV)] dE_x \right\}. \quad (7)$$

Writing

$$\zeta_1 = \frac{4\pi m^3 e}{h^3} \int_0^{E_m} f(E) dE_x,$$

and

$$\zeta_2 = \frac{4\pi m^3 e}{h^3} \int_0^{E_m} f(E + eV) dE_x,$$

and  $\zeta = \zeta_1 - \zeta_2$ , Eq. (7) becomes

$$J = \int_0^{E_m} D(E_x) \zeta dE_x. \quad (8)$$

## 1 → 2 electrons Net tunneling current Current-Voltage Relationship for a Generalized Barrier

Writing<sup>6</sup>  $V(x) = \eta + \varphi(x)$ , with reference to Fig. 1, Eq. (1) becomes

$$D(E_x) = \exp\left[-\frac{4\pi}{h} (2m)^{1/2} \int_{s_1}^{s_2} (\eta + \varphi(x) - E_x)^{1/2} dx\right]. \quad (9)$$

Integrating Eq. (9), using Eq. (A5) from the Appendix, yields

$$D(E_x) \approx \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}], \quad (10)$$

where  $\bar{\phi}$ , the mean barrier height above Fermi level of the negatively biased electrode, is

$$\bar{\phi} = \frac{1}{\Delta s} \int_{s_1}^{s_2} \varphi(x) dx,$$

and

$$A = (4\pi\beta\Delta s/h)(2m)^{1/2},$$

where  $\beta$  is defined in the Appendix. At 0°K,  $\zeta_1$  and  $\zeta_2$  are given by

$$\zeta_1 = (4\pi m^3 e/h^3)(\eta - E_x)$$

<sup>6</sup> By the substitution  $V(x) = \eta + \varphi(x)$ , we have inherently assumed that the width  $\Delta s$  of the barrier in the range  $E_x > V(x) > \eta$  is constant and equal to the barrier width at the Fermi level. This assumption is justified for practical barriers, because  $\Delta s$  varies slowly below the Fermi level [Fig. 4(a)], and the integral has effective values only when  $E_x \approx \eta$ .

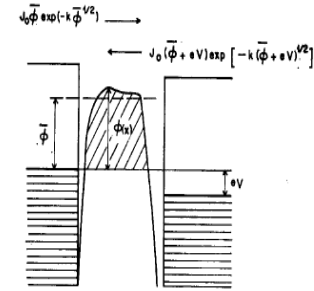


Fig. 2. Pictorial illustration of Eq. (20), showing current flow between the electrodes.

and

$$\zeta_2 = (4\pi m^3 e/h^3)(\eta - E_x - eV).$$

Hence,

$$\zeta = \begin{cases} (4\pi m^3 e/h^3)(eV) & 0 < E_x < \eta - eV \\ (4\pi m^3 e/h^3)(\eta - E_x) & \eta - eV < E_x < \eta \\ 0 & E_x > \eta \end{cases}. \quad (11)$$

**Density of states**

Substituting Eqs. (10) and (11) in Eq. (8) gives

$$J = \frac{4\pi m^3 e}{h^3} \left\{ eV \int_0^{\eta - eV} \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}] dE_x + \int_{\eta - eV}^{\eta} (\eta - E_x) \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}] dE_x \right\}. \quad (12)$$

To facilitate integration, Eq. (12) is written in the form

$$J = \frac{4\pi m^3 e}{h^3} \left\{ eV \int_0^{\eta - eV} \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}] dE_x - \bar{\phi} \int_{\eta - eV}^{\eta} \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}] dE_x + \int_{\eta - eV}^{\eta} (\eta + \bar{\phi} - E_x) \times \exp[-A(\eta + \bar{\phi} - E_x)^{3/2}] dE_x \right\}. \quad (13)$$

The first of the integrals in Eq. (13) yields

$$(8\pi m^3 e/h^3)(e/h^3) \{ [A(\bar{\phi} + eV)^{3/2} + 1] \exp[-A(\bar{\phi} + eV)^{3/2}] - [A(\bar{\phi} + \eta)^{3/2} + 1] \exp[-A(\bar{\phi} + \eta)^{3/2}] \}. \quad (14)$$

The second term in the braces is negligible compared to the first term and, usually,  $A(\bar{\phi} + eV)^{3/2} \gg 1$ ; thus Eq. (14) reduces to

$$(8\pi m^3 e/h^3 A) V(\bar{\phi} + eV)^{3/2} \exp[-A(\bar{\phi} + eV)^{3/2}]. \quad (15)$$

The second integral in Eq. (13) is of the same form as the first. Taking advantage of the approximations

# The MIM junction II

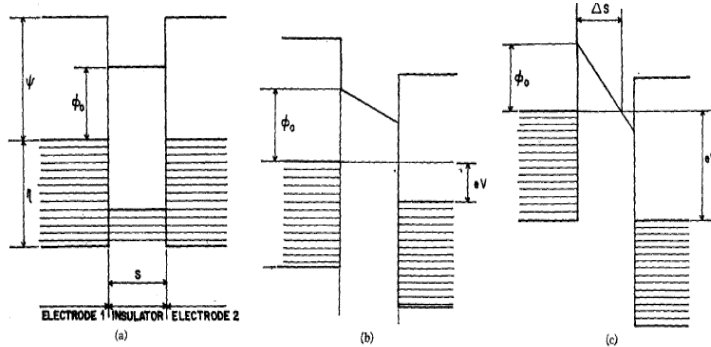


FIG. 3. Rectangular potential barrier in insulating film between metal electrodes for: (a)  $V=0$ ; (b)  $V < \varphi_0/e$ ; (c)  $V > \varphi_0/e$ .

that lead to Eq. (15), the second term integrates to

$$-(8\pi me/h^2 A^2) \bar{\varphi} \{ [A \bar{\varphi}^2 + 1] \exp(-A \bar{\varphi}) - [A(\bar{\varphi} + eV) + 1] \exp[-A(\bar{\varphi} + eV)] \}. \quad (16)$$

The third integral of Eq. (13) has the form

$$\int x^2 e^{-Ax} dx = -e^{-Ax} \left( \frac{x^2}{A} + \frac{2x}{A^2} + \frac{2}{A^3} \right), \quad (17)$$

where

$$x^2 = \eta + \bar{\varphi} - E_x.$$

The third and fourth terms in parentheses in Eq. (17) are negligible by comparison with the first two; therefore, the third integral in Eq. (13) integrates to

$$(8\pi me/h^2 A) \{ \bar{\varphi}^3 \exp(-A \bar{\varphi}) - (\bar{\varphi} + eV)^3 \exp[-A(\bar{\varphi} + eV)] \} + (8\pi me/h^2 A)(3/A) \{ \bar{\varphi} \exp(-A \bar{\varphi}) - (\bar{\varphi} + eV) \exp[-A(\bar{\varphi} + eV)] \}. \quad (18)$$

Summing Eqs. (15), (16), and (18) yields

$$J = (e/2\pi h)(\beta \Delta s)^{-2} \{ \bar{\varphi} \exp(-A \bar{\varphi}) - (\bar{\varphi} + eV) \exp[-A(\bar{\varphi} + eV)] \}. \quad (19)$$

Equation (19) can be expressed in the following form:

$$J = J_0 \{ \bar{\varphi} \exp(-A \bar{\varphi}) - (\bar{\varphi} + eV) \exp[-A(\bar{\varphi} + eV)] \}, \quad (20)$$

where

$$J_0 = e/2\pi h(\beta \Delta s)^2.$$

Equation (20) has the advantage that it can be applied to any shape of potential barrier providing the mean barrier height is known, or, alternatively, if the current-voltage characteristic of a tunnel junction is known, the mean barrier height can be determined.

Equation (20) can be interpreted as a current density  $J_0 \bar{\varphi} \exp(-A \bar{\varphi})$  flowing from electrode 1 to electrode 2 and a current density  $J_0 (\bar{\varphi} + eV) \exp[-A(\bar{\varphi} + eV)]$

flowing from electrode 2 to electrode 1, resulting in a net current density  $J$ , given by Eq. (20). (See Fig. 2.) When  $V$  is zero, a state of dynamic equilibrium can be considered to exist—that is, a current density of magnitude  $J_0 \bar{\varphi} \exp(-A \bar{\varphi})$  flowing in either direction.

### Low-Voltage Range

Although Eq. (20) can be used for very low voltages, a more convenient form can be deduced for this range. From Eq. (20),

$$J = J_0 \{ \bar{\varphi} \exp(-A \bar{\varphi}) - (\bar{\varphi} + eV) \exp[-A(\bar{\varphi} + eV)] \}. \quad (21)$$

It is observed that, since  $eV \approx 0$ ,  $\beta$  [as defined in Eq. (A6)] takes the value unity. Since  $\beta \gg eV$ , Eq. (21) can be written

$$J = J_0 \{ \bar{\varphi} - (\bar{\varphi} + eV) \exp(-A eV/2 \bar{\varphi}) \} \times \exp(-A \bar{\varphi}). \quad (22)$$

Expanding  $\exp(-A eV/2 \bar{\varphi})$ , and neglecting terms containing  $V^2$  and higher orders, Eq. (22) becomes

$$J = J_0 \{ \bar{\varphi} - (\bar{\varphi} + eV)(1 - A eV/2 \bar{\varphi}) \} \exp(-A \bar{\varphi}) = J_0 eV [A \bar{\varphi}^2/2 - 1] \exp(-A \bar{\varphi}). \quad (23)$$

Since  $A \bar{\varphi}^2/2 \gg 1$ , Eq. (23) reduces to

$$J = J_L \bar{\varphi}^2 V \exp(-A \bar{\varphi}), \quad (24)$$

where

$$J_L = [(2m)^2/\Delta s^2](e/h)^2.$$

Since  $eV$  is very small,  $\bar{\varphi}$  is considered to be the zero-voltage mean barrier height. Thus, in this case, Eq. (24) expresses  $J$  as a linear function of  $V$ ; that is, the junction is Ohmic for very low voltages.

### APPLICATION OF THE TUNNEL EQUATIONS

Consider a rectangular potential barrier [Fig. 3(a)]. This was the type of barrier studied by Sommerfeld and

Bethe, and by Holm. Sommerfeld and Bethe considered the low-voltage and high-voltage cases, and Holm the intermediate case. Equations for each are derived for these cases, using Eqs. (20) and (24).

### Low-Voltage Range: $V \approx 0$

From Fig. 3(a),

$$\Delta s = s,$$

and

$$\bar{\varphi} = \varphi_0.$$

Substituting these values in Eq. (24) gives

$$J = [3(2m\varphi)^2/2s^2](e/h)^2 V \times \exp[-(4\pi s/h)(2m\varphi_0)^2]. \quad (25)$$

This result is in agreement with the Sommerfeld-Bethe result for low voltages.

### Intermediate-Voltage Range: $V < \varphi_0/e$

From Fig. 3(b),

$$\Delta s = s,$$

and

$$\bar{\varphi} = (\varphi - eV/2).$$

Substituting these values in Eq. (20),

$$J = J_0 \{ (\varphi - eV/2) \exp[-A(\varphi - eV/2)] - (\varphi + eV/2) \exp[-A(\varphi + eV/2)] \} = \frac{e}{2\pi h(\beta s)^2} \left\{ \left( \varphi - \frac{eV}{2} \right) \exp \left[ -\frac{4\pi \beta s}{h} (2m)^2 \left( \varphi - \frac{eV}{2} \right) \right] - \left( \varphi + \frac{eV}{2} \right) \exp \left[ -\frac{4\pi \beta s}{h} (2m)^2 \left( \varphi + \frac{eV}{2} \right) \right] \right\}. \quad (26)$$

It is now necessary to discuss the error associated with Eq. (26). It can be shown, using Eq. (A6), that  $\beta$  is given by

$$\beta = 1 - (eV)^2/96(\varphi_0 + \eta - E_x - eV/2)^2.$$

If this value of  $\beta$  is substituted in Eq. (26), the maximum error in the exponents in using the approximate integral Eq. (A5) is approximately 1% and occurs when  $V = \varphi_0/e$  and  $E_x = \eta$ . For values of  $eV < \varphi_0/e$ , the error reduces rapidly. If  $\beta$  is chosen to be unity, the error in the value of the exponents is approximately 6% at  $V = \varphi_0/e$ . However, since the error reduces rapidly for values of  $V < \varphi_0/e$ , the error is only 1% at  $V = 0.75\varphi_0/e$ , and  $\beta$  can, therefore, be chosen to be unity to a reasonable approximation.

With  $\beta = 1$ , Eq. (26) becomes<sup>14</sup>

<sup>14</sup> Note added in proof. Simmons has shown [see J. G. Simmons, J. Appl. Phys. 34, 238 (1963)] that, at relatively low voltages, Eq. (27) reduces to

$$J = J_L(V + \beta V^2),$$

where

$$\beta = [(Ae)^2/96\varphi] - [A^2/32\varphi^2],$$

which is in good quantitative agreement with the experimental

$$J = \left( \frac{e}{2\pi h s^2} \right) \left\{ \left( \varphi_0 - \frac{eV}{2} \right) \exp \left[ -\frac{4\pi s}{h} (2m)^2 \left( \varphi_0 - \frac{eV}{2} \right) \right] - \left( \varphi_0 + \frac{eV}{2} \right) \exp \left[ -\frac{4\pi s}{h} (2m)^2 \left( \varphi_0 + \frac{eV}{2} \right) \right] \right\}. \quad (27)$$

Equation (27) differs from Holm's result [Eq. (16) of reference 3]. In addition to the terms of Eq. (27), Holm includes an additional term,

$$\frac{(2m)^2}{h^2} eV \left( \varphi_0 + \frac{eV}{2} \right) \exp \left[ -\frac{4\pi s}{h} (2m)^2 \left( \varphi_0 + \frac{eV}{2} \right) \right] \quad (28)$$

which is in error. Holm recognizes that there are inconsistencies in his result (Sec. VIII of reference 3), for two reasons: (1) As  $V \rightarrow 0$ , his equation does not reduce to the Sommerfeld-Bethe relationship; that is, the equation does not predict the low-voltage Ohmic characteristic [Eqs. (24) and (25)]. (2) According to his equation, the resistance of the junction initially increases with increasing voltage.

Holm suggests that these anomalies are due to the approximate nature of  $D(E_x)$ . This is not the case, however, for the anomalies are removed when the extraneous term, Eq. (28), above, in his equation [Eq. (16) of reference 3] is neglected. (See Fig. 6.)

### High-Voltage Range: $V > \varphi_0/e$

Figure 3(c) illustrates the energy diagram for this case; from this figure,

$$\Delta s = s\varphi_0/eV,$$

and

$$\bar{\varphi} = \varphi_0/2.$$

Substituting these values in Eq. (20) yields

$$J = \left( \frac{2e^2(F/\beta)^2}{8\pi h \varphi_0} \right) \left\{ \exp \left[ -\frac{4\pi \beta}{eF} m^2 \varphi_0^2 \right] - \left( 1 + \frac{2eV}{\varphi_0} \right) \times \exp \left[ -\frac{4\pi \beta}{eF} m^2 \varphi_0^2 \left( 1 + \frac{2eV}{\varphi_0} \right) \right] \right\}, \quad (29)$$

where  $F = V/s$  = the field strength in the insulator.

It is now necessary to determine the correction factor  $\beta$  appearing in Eq. (29). From Eq. (A6),

$$\beta = 1 - \left[ \frac{(eV/s)^2}{8\Delta s} \int_0^{\Delta s + s\varphi_0/eV} (\Delta s/2 - x)^2 dx / (\varphi_0/2)^2 \right] = 1 - 1/24 = 23/24.$$

Therefore, for this case,  $\beta$  is independent of  $V$ . Sub-

results of Knauss and Breslow [see H. P. Knauss and R. A. Breslow, Proc. IRE 50, 1843 (1962)].

Parabolic dep.

Linear dep.

# Measured tunneling current in STM

stituting this value in Eq. (29) gives

$$J = \frac{2.2e^3 F^2}{8\pi h \varphi_0} \left\{ \exp \left[ -\frac{8\pi}{2.96heF} (2m)^{1/2} \varphi_0^{1/2} \right] - \left( 1 + \frac{2eV}{\varphi_0} \right) \times \exp \left[ -\frac{8\pi}{2.96heF} (2m)^{1/2} \varphi_0^{1/2} \left( 1 + \frac{2eV}{\varphi_0} \right)^{1/2} \right] \right\}. \quad (30)$$

For very high voltages (that is, where  $V > (\varphi + \eta)/e$ ), the Fermi level of electrode 2 lies below the bottom of the conduction band of electrode 1. Under these conditions, electrons cannot tunnel from electrode 2 to electrode 1, since there are no empty levels available to them. The situation is reversed, however, for electrons tunneling from electrode 1 to electrode 2, since all of the available levels in electrode 2 are empty. The situation is analogous to that of field emission from a metal electrode, and, for this condition (that is, where  $V > (\varphi + \eta)/e$ ), the second term in Eq. (30) is negligible; thus,

$$J = \frac{2.2e^3 F^2}{8\pi h \varphi_0} \exp \left[ -\frac{8\pi}{2.96heF} (2m)^{1/2} \varphi_0^{1/2} \right].$$

**Field emission**

This equation is similar to the Sommerfeld-Bethe<sup>2</sup> relationship for high voltages, except for the multiplicative factor 2.2; there is also a slight difference in the numerator, where 3 is replaced by 2.96. These differences arise because of the variation of  $\Delta s$  below the Fermi level.<sup>6</sup> Because of the dominant influence of the exponential term, and since  $J$ , in Eq. (30), is a rapidly varying function of  $s$ ,  $\varphi$ , and  $V$ , this difference is considered to be insignificant.

Thus, using formulas (2)–(5), we can compute the tunnel current at given system parameters and plot current-voltage characteristics. Fig. 4 shows theoretical tunneling current-applied voltage plot in case of carbon electrode 1 ( $\phi_1 = 4,7$  eV) and platinum electrode 2 ( $\phi_2 = 5,3$  eV) at  $\delta_z = 5$  Å and contact area  $S = 10^{-17}$  m<sup>2</sup>.

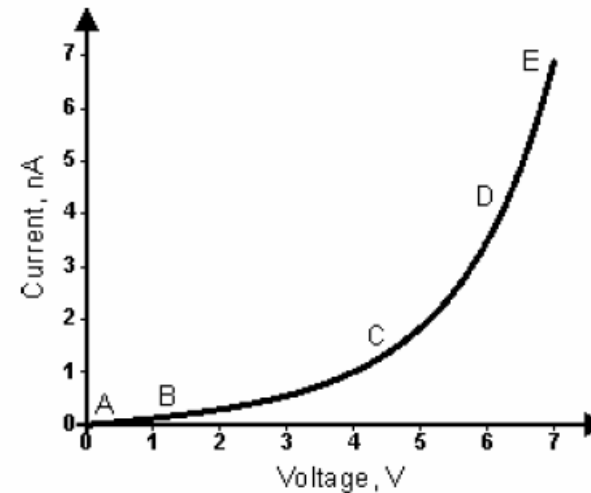


Fig. 4. Current-voltage characteristic for carbon electrode 1 and platinum electrode 2 at  $\delta_z = 5$  Å and contact area  $10^{-17}$  m<sup>2</sup>. Parts of  $J(V)$  curve correspond to the following expressions: AB – (22), BC – (23), CD – (24), DE – (25).

- ✓ J-V curve depends on materials
- ✓ J is representative of local density of states

# Quantum approach to tunneling in STM

Many papers address the problem of tunneling within a quanta-mechanical frame, as, e.g.:

INSTITUTE OF PHYSICS PUBLISHING  
Nanotechnology 17 (2006) R57–R65  
doi:10.1088/0957-4484/17/8/R01

NANOTECHNOLOGY

TUTORIAL

## Bardeen's tunnelling theory as applied to scanning tunnelling microscopy: a technical guide to the traditional interpretation

Alex D Gottlieb<sup>1</sup> and Lisa Wesoloski<sup>2</sup>

[10] Briggs G A D and Fisher A J 1999 STM experiment and atomistic modelling hand in hand: individual molecules on semiconductor surfaces *Surf. Sci. Rep.* **75** 1–81

[11] Hofer W A, Foster A S and Shluger A L 2003 Theories of scanning probe microscopes at the atomic scale *Rev. Mod. Phys.* **75** 1287–331

[12] Drakova D 2001 Theoretical modeling of scanning tunnelling microscopy, scanning tunnelling spectroscopy and atomic force microscopy *Rep. Prog. Phys.* **64** 205–90

PHYSICAL REVIEW B  
VOLUME 31, NUMBER 2

Theory of the scanning tunneling microscope

VOLUME 6, NUMBER 2      PHYSICAL REVIEW LETTERS  
TUNNELLING FROM A MANY-PARTICLE POINT OF VIEW  
J. Bardeen

For low bias across the tunnelling junction, we obtain the following Tersoff–Hamann formula<sup>11</sup> for the current from (22):

$$\begin{aligned}
 I &= \pm \frac{2\pi e}{\hbar} \rho_{\text{tip}}(\mu) \left( \frac{\hbar^2}{m} 2\pi \right)^2 A^2 \sum_{\psi_n; \mu_a < \epsilon_n < \mu_b} |\psi_n(\mathbf{0})|^2 \\
 &= (\mu_t - \mu_s) \frac{eh^3}{m^2} A^2 \rho_{\text{tip}}(\mu) \rho_s(\mathbf{0}, \mu), \quad (32)
 \end{aligned}$$

where

$$\rho_s(\mathbf{r}, \mu) = \frac{1}{|\mu_s - \mu_t|} \sum_{\psi_n; \mu_a < \epsilon_n < \mu_b} |\psi_n(\mathbf{r})|^2$$

is the 'local density of sample states' per unit volume near the point  $\mathbf{r}$  and per unit energy near the quasi-Fermi energy  $\mu$ . Thus, the Tersoff–Hamann formula predicts that the conductance of an STM at low temperature and low bias is proportional to the local density of sample states at the centre of the tip and the quasi-Fermi energy.

Quantum approaches underline the dependence on matrix elements and wavefunctions (better, density of states)

# Bardeen (quantum) approach to tunneling

## 2.2. Bardeen Approach

Another way of describing electron tunneling comes from Bardeen's approach which makes use of the time dependent perturbation theory. The probability of an electron in the state  $\Psi$  at  $E_\Psi$  to tunnel

into a state  $\chi$  at  $E_\chi$  is given by Fermi's Golden Rule

$$w = \frac{2\pi}{\hbar} |M|^2 \delta(E_\Psi - E_\chi)$$

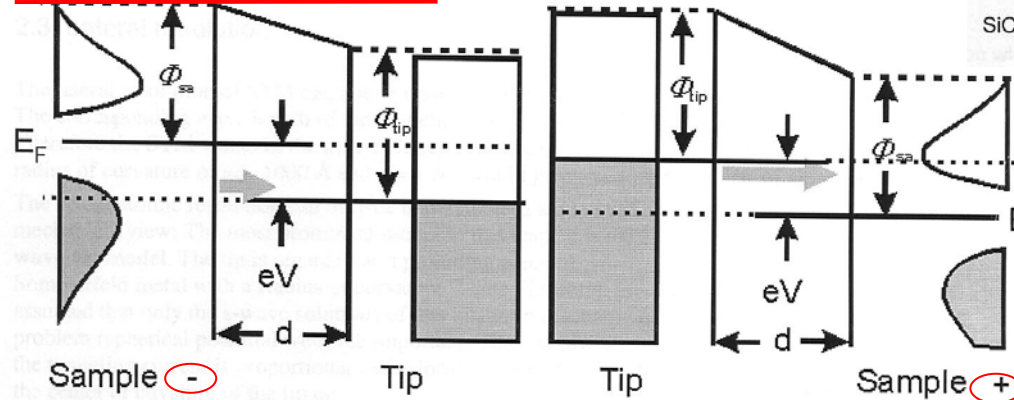
The tunneling matrix element is given by an integral over a surface in the barrier region lying between the tip and the sample:

$$M = \frac{\hbar}{2m} \int_{z=z_0} \left( \chi^* \frac{\partial \Psi}{\partial z} - \Psi \frac{\partial \chi^*}{\partial z} \right) dS$$

Applying a bias voltage  $V$  and approximating the Fermi distribution as a step function ( $kT \ll \Delta E_{\text{resolution}}$ ), the current is

$$I = \frac{4\pi e^{eV}}{\hbar} \int_0^{eV} \rho_{\text{sa}}(E_F - eV + \varepsilon) \rho_{\text{tip}}(E_F + \varepsilon) |M|^2 d\varepsilon \quad (**)$$

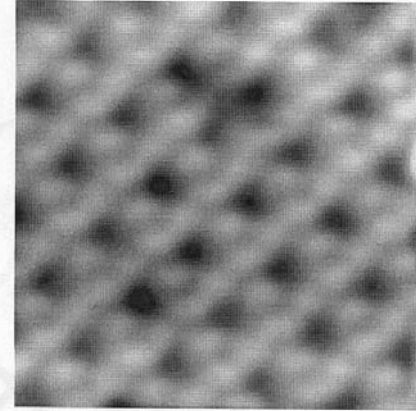
Hence the current is given by a combination of the local densities of states of the sample and the tip, weighted by the tunneling matrix element  $M$ .



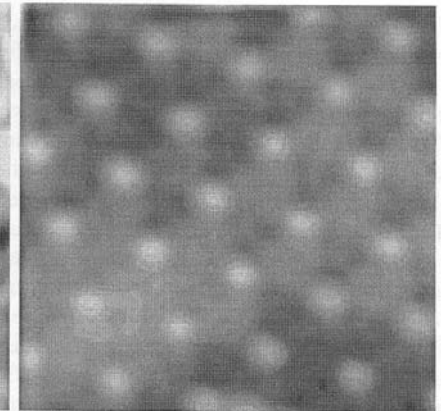
Schematic of electron tunneling with respect to the density of states of the sample.

$\delta(E_\Psi - E_\chi)$  means that an electron can only tunnel if there is an unoccupied state with the same energy in the other electrode (thus inelastic tunneling is not treated). In case of a negative potential on the sample the occupied states generate the current, whereas in case of a positive bias the unoccupied states of the sample are of importance. Therefore, as shown below, by altering the voltage, a complete different image can be detected as other states contribute to the tunneling current. This is used in tunneling spectroscopy. It should finally be mentioned that the probability of tunneling (expressed by  $M^2$ ) is larger for electrons which are close to the fermi edge due to the lower barrier.

**Quantum mech.  
Treatment**



Imaging the occupied states of SiC(0001)3x3



Imaging the unoccupied states of SiC(0001)3x3

**Tunneling gets a local dependence (in the in-plane directions) because of Local Density Of States (LDOS)**

# Local Density Of States I

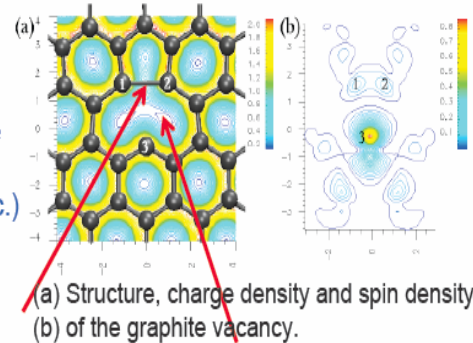
- We can even visualize each contribution in space by using the explicit shape of the basis function (times the corresponding Mulliken charge)



- We can visualize the spatial distribution of the total electron density

## Plane waves:

- Fourier transform – electron density in the real space
- Similar to real space calculations, we can project electron density to particular atom-type state (s, p, d, etc.)
- Visualization of the total electron density is not very informative
- Spin density (for magnetic systems) is more useful



(a) Structure, charge density and spin density of the graphite vacancy.  
 Electron density is larger between atoms 1-2 than between atoms 2-3 or 1-3. The trigonal symmetry is broken due to Jahn-Teller distortion

- Density of states: the definition:  $g(E) = \sum_n \delta(E - E_n)$

- DOS is just the number of energy levels between  $E$  and  $E + dE$  divided by the infinitesimal energy interval  $dE$

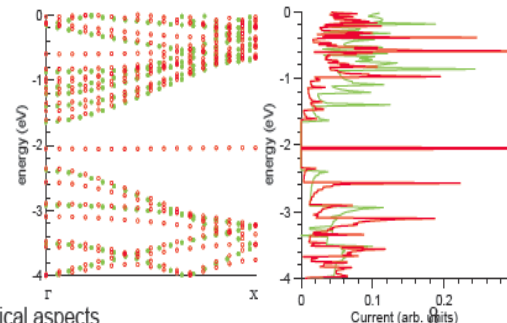
$$g(E) = \frac{1}{dE} \sum_E 1$$

- Can be calculated easily if we know the spectrum
- General concept: applicable to both finite and infinite systems
- Discrete function in finite and continuous function of  $E$  in infinite systems

- In periodic (infinite systems) when the concept of  $k$ -space works DOS can be defined as:

$$g(E) = \sum_k \delta(E - E(k))$$

- The sum runs over all possible values of  $k$ , but in actual calculations we restrict  $k$  to the first Brillouin zone.)



November 28, 2006

Practical aspects

- Local density of states: the definition:  $g(\vec{r}, E) = \sum_n |\psi_n(\vec{r})|^2 \delta(E - E_n)$

- LDOS is the charge density at point  $r$  resulting exclusively from states in the energy interval from  $E$  to  $E + dE$ .
- If we sum LDOS over all atom in the system, we will get the total DOS because of the normalization condition on the wave function
- In periodic systems without defects DOS and LDOS are equivalent

- In a real space approach with orthogonal basis functions LDOS on atom  $j$  is:

$$g(j, E) = \sum_n |c_j^n|^2 \delta(E - E_n)$$

- For a non-orthogonal basis set:

$$g(j, E) = \sum_n \sum_\alpha c_{j,\alpha}^n \sum_{i,\beta} c_{i,\beta}^n S_{\alpha,\beta}^{i,j} \delta(E - E_n)$$

- Partial (s, p, etc.) LDOS projected to the particular state atom-type state:

$$g(j, \alpha, E) = \sum_n c_{j,\alpha}^n \sum_{i,\beta} c_{i,\beta}^n S_{\alpha,\beta}^{i,j} \delta(E - E_n)$$

- We can also calculate partial LDOS in plane-wave calculations by projecting the electron density to the corresponding atom-type state

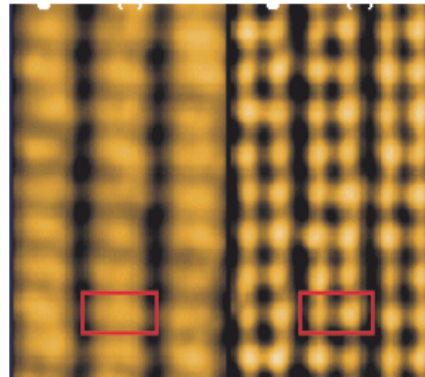
Arkady Krasheninnikov,  
 University of Helsinki and  
 Helsinki University of Technology



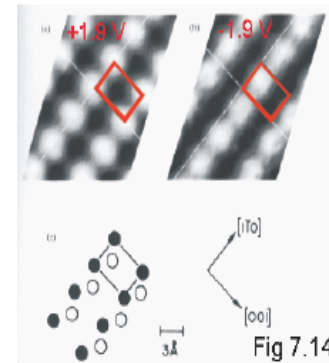
# Local Density Of States II

- STM images of semiconductors:

Si(100)(2x1)



GaAs(110)(1x1)



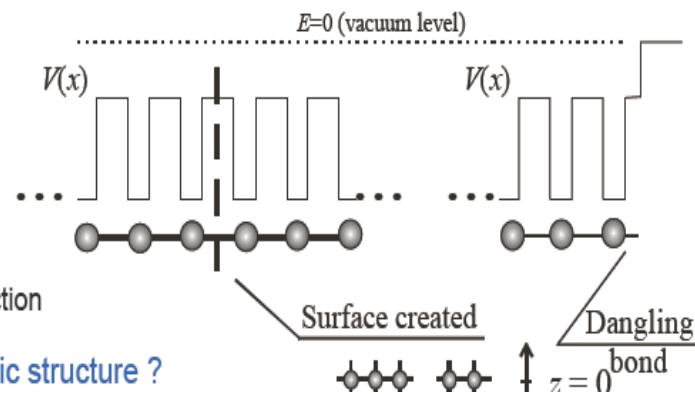
- The images are different at positive and negative voltages!
- What we measure is not simply the positions of atoms, but something else!

- The electronic structure of solids is different near the surface from that in the bulk

- New states due to

- Dangling bonds
- Loss of periodicity in z direction

- How to calculate the electronic structure ?



# Lateral resolution in STM

## 2.3. Lateral resolution

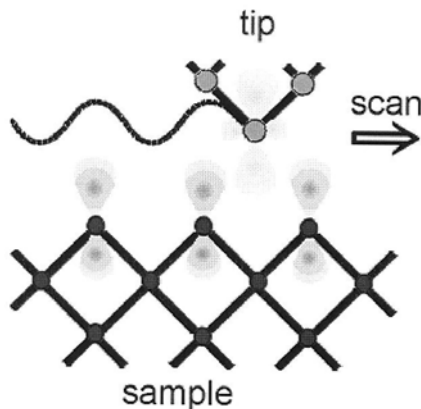
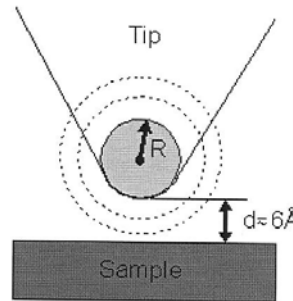
The lateral resolution of STM can not be understood in terms of a Fraunhofer diffraction resolution. The corresponding wave length of the tunneling electron would be  $\lambda > 10 \text{ \AA}$ .

Therefore the STM works in the near-field regime. The overall geometric curvature of the tip with a radius of curvature of e.g.  $1000 \text{ \AA}$  and  $\kappa = 1 \text{ \AA}^{-1}$  would give rise to a resolution of about  $50 \text{ \AA}$ .

The actual atomic resolution can only be understood in a quantum mechanical view: The most prominent model in this respect is the **s-wave-tip model**. The tip is regarded as a protruding piece of Sommerfeld metal with a Radius of curvature  $R$  (see Figure). It is assumed that only the s-wave solutions of this quantum mechanical problem (spherical potential well) are important. Thus, at low bias the tunneling current is proportional to the local density of states at the center of curvature of the tip  $r_0$ :

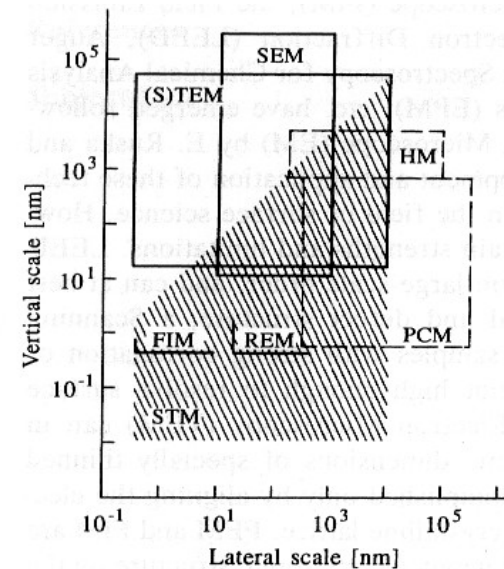
$$I \propto \sum_{E_k = E_f - eV}^{E_f} |\Psi_{\mu}(r_0)|^2 = eV \rho_{sa}(r_0, E_f)$$

In this model only the properties of the sample contribute to the STM image which is quite easy to handle. But it cannot explain the atomic resolution.



Interaction which causes a high corrugated tunneling distribution

Calculations and experiments showed that there is often a  $d_z^2$  like state near the fermi edge present at the apex atom which also predominantly contributes to the tunneling current. It is understood that this state (and also the  $p_z$  like state) is advantageous for a „sharp, tip. Since the tunneling current is a convolution of the tip state and the sample state, there is a symmetry between both: By interchanging the electronic state of the tip and the sample state, the image should be the same (reciprocity principle). This can also explain the fact that the corrugation amplitude of an STM image is often larger than that of the LDOS of the sample (measured by helium scattering). In this case the tip traces a fictitious surface with a  $d_z^2$  like state. The state of the tip atom is dependent on the material and the orientation. As the tip is quite difficult to handle, it is one of the most difficult problems in a STM experiment.



**Fig. 1.1. Comparison of the resolution range of STM with that of other microscopes [1. 1]. [HM: High-resolution optical Microscope. PCM: Phase Contrast Microscope. (S)TEM: (Scanning) Transmission Electron Microscope. FIM: Field Ion Microscope. REM: Reflection Electron Microscope]**

**“Atomic” resolution achievable in STM**

# Tunneling current and STM

## 2.1 Theoretical Fundamentals of the Scanning Tunneling Microscope

How does a measuring instrument function that allows us to see single atoms? In the case of a scanning tunnelling microscope a fine metallic tip is used as the probe (called tunnelling tip) (see Figure 3). This tip is approached toward the surface until a current flows when a voltage is applied between the tip and the sample surface. This happens at distances in the order of 1 nm. The current is called tunnel current since it is based on the quantum-mechanical tunnel effect. After a tunnelling contact is established, the tip is moved over the surface by a piezoelectric scanning unit, whose mechanical extension can be controlled by applying appropriate voltages. The scanning unit is typically capable of scanning an area of a few nm up to several  $\mu\text{m}$ . This allows us to obtain a microscopic image of the spatial variation of the tunnel current. Hence the name scanning tunnelling microscope.

- A metallic tip is moved as probe towards a conducting surface up to a distance of about 1 nm
- With an applied voltage a current flows due to the tunnel effect (tunnel current)
- The spatial variation of the tunnel current is measured by scanning over the sample surface
- A microscopic image of the surface is produced

At this stage we have to ask what kind of atomic-scale structures can be made visible by the scanning tunnelling microscope utilising the tunnel effect? These structures must by nature correspond to electrical states from or into which the electrons can tunnel. In the tunnelling process, the electrons must tunnel through the vacuum barrier between tunnelling tip and sample, which represents a potential barrier. The tunnel effect allows a particle (here an electron) to tunnel through this potential barrier even though the electron's energy is lower than the barrier height. The probability of such a process decreases exponentially with the geometrical distance between the tip and the sample and with increasing barrier height. An experimental apparatus making use of the tunnel effect must therefore minimise the potential barrier to be tunnelled through. This is realised in the scanning tunnelling microscope configuration by moving the tip very close (about 1 nm) to the surface. The electrons can then pass between the surface and the tip. The direction of the tunnel current is fixed by applying a voltage between sample and tip.

In order to explain and interpret the images of the surface states obtained in this way, efforts to develop a theory were made very soon after the invention of the scanning tunnelling microscope. One of the possible theoretical approaches is based on Bardeen's idea of applying a transfer Hamiltonian operator to the tunnelling process [2]. This had the advantage of adequately describing the many-particle nature of the tunnel junction. In the model, a weak overlap of the wave functions of the surface states of the two electrodes (tunnelling tip and sample surface) is assumed to allow a perturbation calculation. On this basis, Tersoff and Hamann developed a simple theory of scanning tunnelling microscopy [3], [4] Hence follows the tunnel current:

$$I \sim V \cdot \rho_{\text{tip}}(W_F) \cdot \rho_{\text{sample}}(r_0, W_F) \quad (1)$$

The tunnelling tip is assumed to be a metallic s-orbital as shown schematically in Figure 4. In addition, it is assumed that low voltages  $V$  (i.e., much smaller than the work function) are applied.  $\rho_{\text{tip}}(W_F)$  is the density of states of the tip and  $\rho_{\text{sample}}(r_0, W_F)$  is that of the sample surface at the centre  $r_0$  of the tip orbital and at the Fermi energy  $W_F$ . Eq. (1) shows that at low voltage the scanning tunnelling microscope thus images the electronic density of states at the sample surface near the Fermi energy. However, this result also means that the scanning tunnelling microscope images do not directly show the atoms, but rather the electronic states bound to the atoms. As can be seen in Eq. (1), the tips density of states enters in the measurement in the same way as the density of states of the sample. It is therefore desirable to know the exact electronic state of the tip, but unfortunately, in practice, every tip is different and the details remain unknown.

- Weak overlap of the wave functions of the surface states of the two electrodes (tunnelling tip and sample surface)
- Tunnelling tip approximated as an s-orbital
- Low voltages ( $V \ll$  work function)

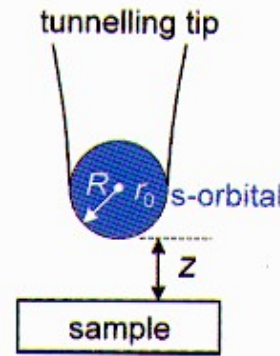


Figure 4: Schematic representation of the tunnelling geometry in the Tersoff-Hamann model.

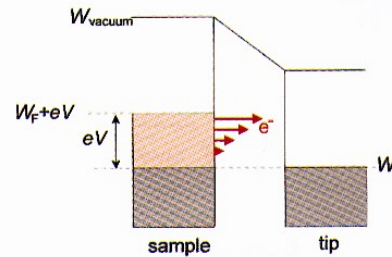


Figure 5: At high voltages not only the states near the Fermi energy  $W_F$  contribute to the current but all states whose energy ranges between  $W_F$  and  $W_F + eV$ .

Extension to "high" bias

- The tunnel current is proportional to the local density of states of the sample
- The scanning tunnelling microscope images the electronic local density of states of the sample near the Fermi energy.

In a first approximation the density of surface states decreases exponentially into the vacuum with the effective inverse decay length  $k_{\text{eff}}$

$$k_{\text{eff}} = \sqrt{\frac{2m_e B}{\hbar^2} + |k_{\parallel}|^2} \quad (2)$$

$m_e$  is the electron mass and  $k_{\parallel}$  is the parallel wave vector of the tunnelling electrons.  $B$  is the barrier height, which is approximately a function of the applied voltage  $V$  and the work functions  $\Phi_{\text{sample}}$  and  $\Phi_{\text{tip}}$  of the sample and tip [5], respectively:

$$B = \frac{\Phi_{\text{tip}} + \Phi_{\text{sample}}}{2} - \frac{|eV|}{2} \quad (3)$$

The tunnel current thus decreases exponentially with the tip-sample distance  $z$ :

$$I \sim \exp[-2k_{\text{eff}} z] \quad (4)$$

The exponential current-voltage dependence is quite essential for the high measurement accuracy of a scanning tunnelling microscope, since even small changes in distance may cause a large change in the tunnel current. Thus the tip just needs one microtip, which is only about 0.1 nm closer to the surface than the next one, and still all current flows over only the closest microtip. Thus even apparently wide tips can yield atomic resolution via one microtip.

The description of the tunnel current by Eq. (1) however, has an important restriction: it strictly speaking only applies to low voltages  $V$ . In particular for the investigation of semiconductor surfaces voltages of the order of 2 to 3 V are required due to the existence of a band gap. Thus the theory must be extended. The simplest extension yields:

$$I \sim \int_{W_{F,\text{tip}}}^{W_{F,\text{tip}} + eV} \rho_{\text{tip}}(W) \rho_{\text{sample}}(W + eV) T(W, V) dW \quad (5)$$

$T(W, V)$  is a transmission coefficient which depends on the energy of the electrons and the applied voltage. The tunnel current is composed of the product of the density of states of the tip and sample at all the different electron energies that are allowed to participate in the tunnelling process (Figure 5). For example, an image measured at  $-2$  V applied to the sample, consequently shows all occupied sample states with an energy between the Fermi energy and  $2$  eV below the Fermi energy. Tunnelling at a positive voltages analogously provides a measurement of the empty surface states in an energy interval determined again by the voltage.

In order to illustrate this effect more clearly, in the following the InP(110) surface will be presented. On InP(110) surfaces two electrical states exist near the surface: an occupied state below the valence band edge and an empty state above the conduction band edge (Figure 6). All the other states are located geometrically deeper in the crystal or energetically deeper in the bands. They thus contribute little to the tunnel current.

Da R. Waser Ed., Nanoelectronics and information technology (Wiley-VCH, 2003)

# Modes of operation in STM

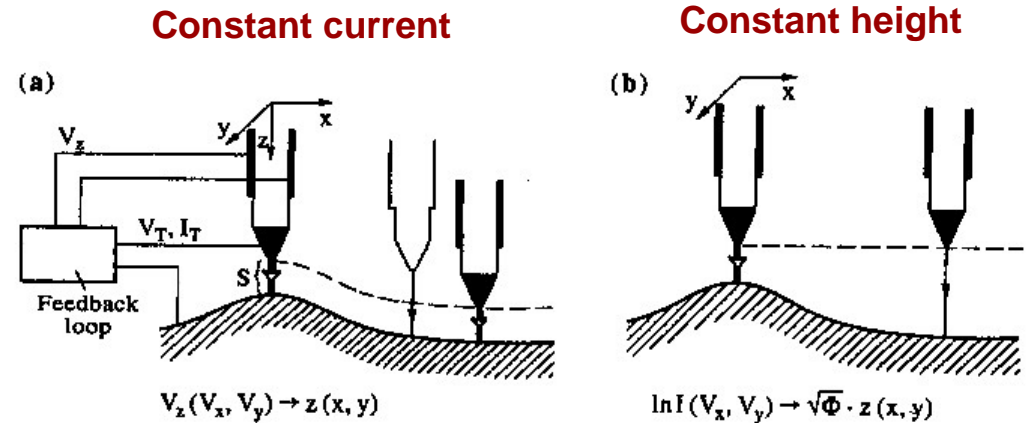
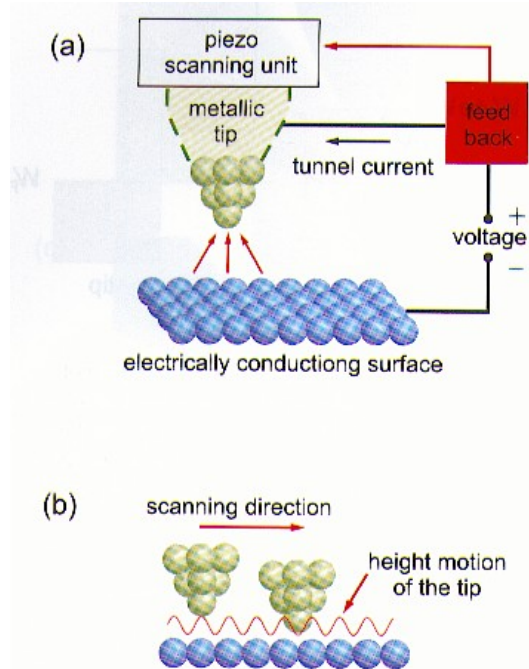


Fig. 1.2a, b. Schematic view of two modes of operation in STM [1. 10].  $S$  is the gap between the tip and the sample,  $I$  and  $V_T$  are the tunneling current and bias voltage, respectively, and  $V_z$  is the feedback voltage controlling the tip height along the  $z$  direction. (a) constant-current mode and (b) constant-height mode

## 2.2 Operating Modes of the Scanning Tunneling Microscope

Up to now, the theoretical background of a scanning tunnelling microscope has been presented, but nothing has been said about the experimental operation of a scanning tunnelling microscope. The simplest way to obtain a scanning tunnelling microscope image is to directly measure the variation of the tunnel current as a function of the scanning position while keeping the distance between tip and sample surface constant. A so-called current image is then obtained. Instead of directly recording the atomic variation of the current, however, the usual procedure is to keep the tunnel current constant while scanning over the surface. This is done by changing the distance between tip and surface using a feedback loop (Figure 8). In order to get an image, the voltage required at the piezoelectric crystal to adjust the distance is recorded. One obtains a so-called constant-current STM image.

**Feedback loop used to keep constant the tunneling current → “absolute” topography map**

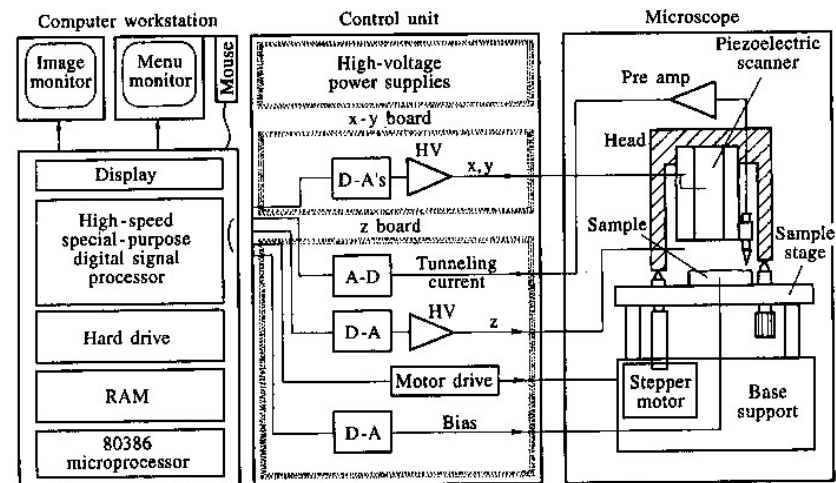


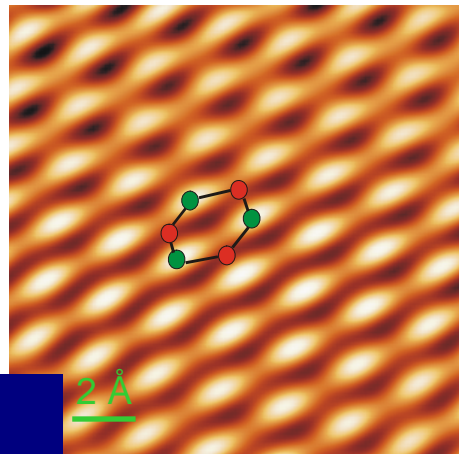
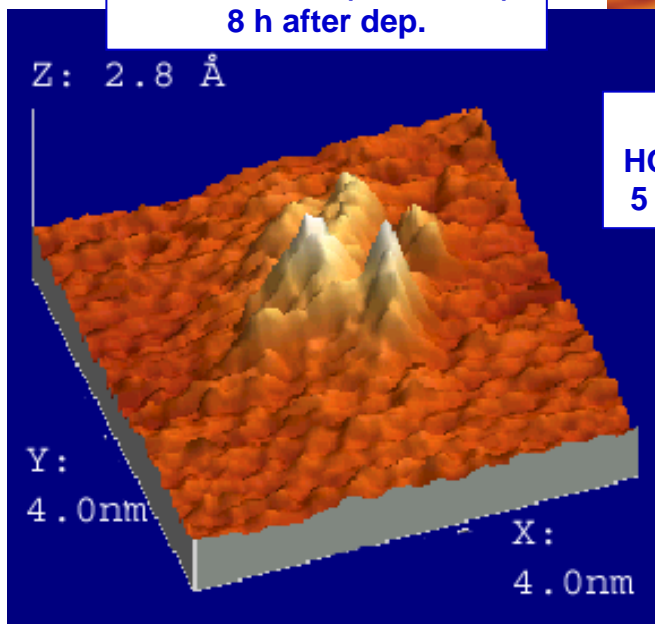
Fig. 4.27. Block diagram of the STM control, data acquisition and display system indicating all of the equipment connections

# Atomic resolution in STM I

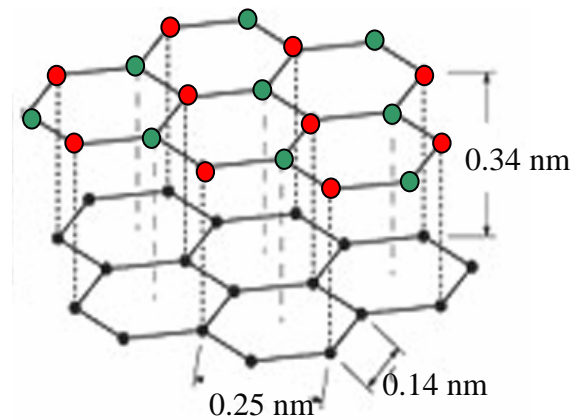
Highly Oriented Pyrolytic Graphite (HOPG) substrates well suited as test samples

A few examples

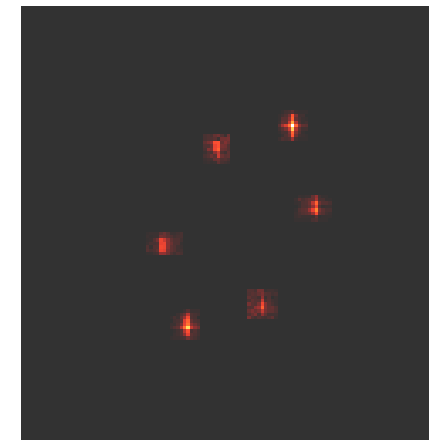
STM image - 3D view  
Cs on HOPG (bias 0.5 V)  
8 h after dep.



STM image – plan view  
HOPG substrate,  $p < 10^{-8}$  mbar  
5 nm x 5 nm scan (bias 0.5 V)



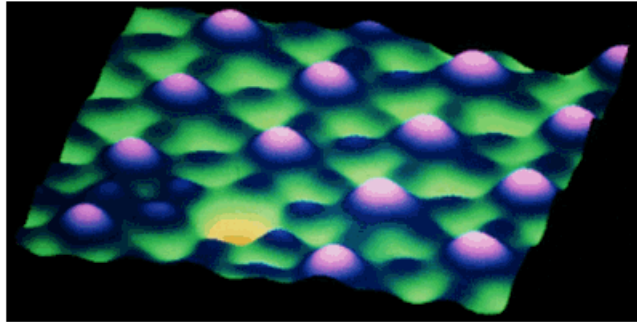
Surface electron-related features can be easily observed



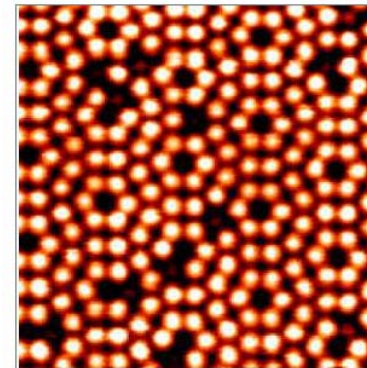
2D-FFT

# Atomic resolution in STM II

Defect I Lattice

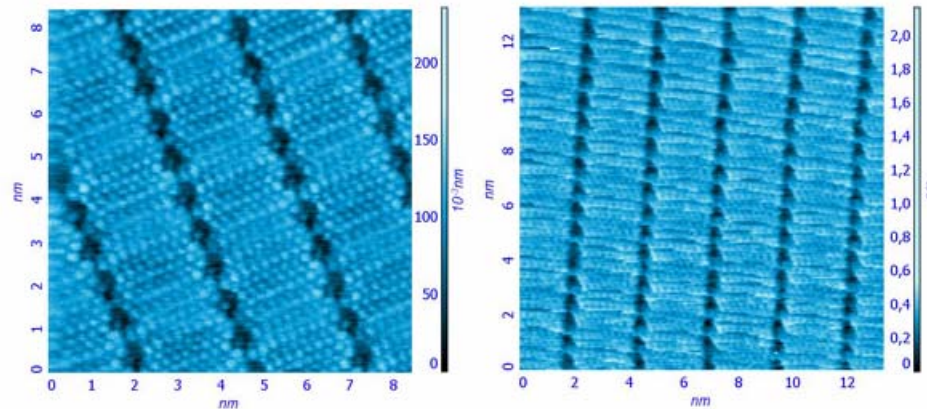


STM image showing single-atom defect in iodine adsorbate lattice on platinum. 2.5nm scan courtesy of Purdue University.



STM

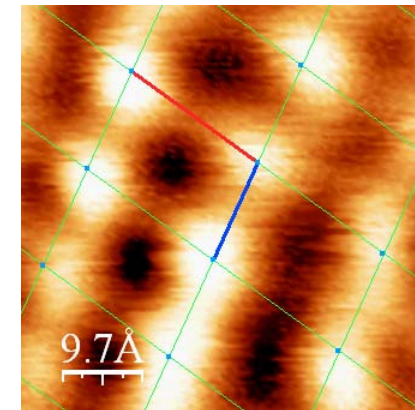
High performance STM image showing atomic resolution on Si(111)7x7



Stearic acid self-assembled molecules.

Mode: [Constant Current mode](#)  
 SPM Model: [Solver P47-PRO](#)  
 Scan size: 8.3x8.3 nm  
 Source MDT-file: [download \(259.70 Kb\)](#)

Stearic acid self-assembled molecules on the HOPG surface. Measurements were carried out in liquid-solid interface. The left image was obtained with Solver P47 Pro SPM, the right image - using the STM scanning measuring head (SF8012).



Nonanethiol molecules on Au/Cr/Si

**Atomic resolution  
 achievable on different  
 materials (even non  
 conductive)**

## Atomic resolution in STM III

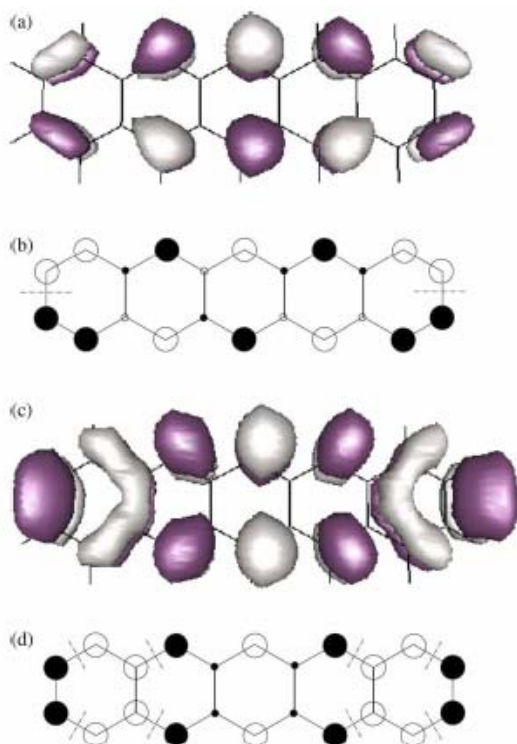
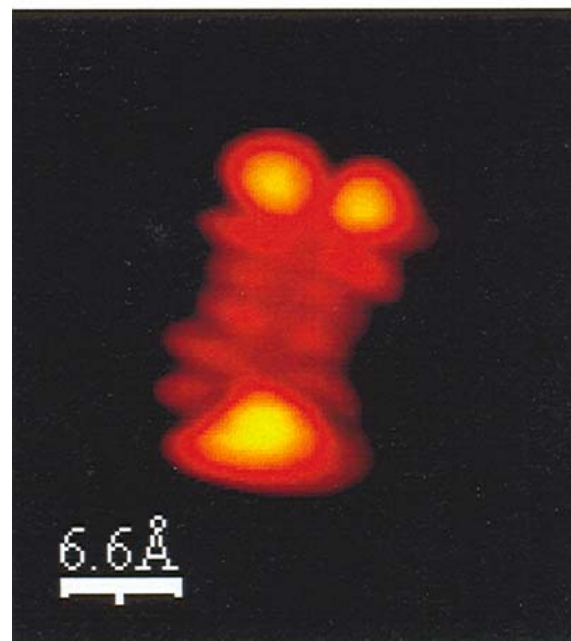
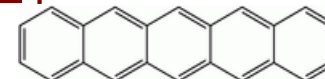


Fig. 2. Single particle wavefunctions of isolated pentacene molecule. The dark color represents a positive sign, the light color a negative sign. (a) and (b) show the HOMO, (c) and (d) the LUMO. (a) isosurface at  $\pm 0.05/\text{\AA}^3/2$ ; (b) LCAO coefficients: small circles 0.01–0.09, large circles 0.12–0.3, the short lines show regions of high gradients; (c) isosurface at  $\pm 0.03/\text{\AA}^3/2$ , (d) similar to (b). Wavefunctions plotted with the program gOpenMol [28].



A single pentacene molecule



**Imaging the local density of states enables imaging “molecular orbitals”**

ELSEVIER

Computational Materials Science 29 (2004) 362–370

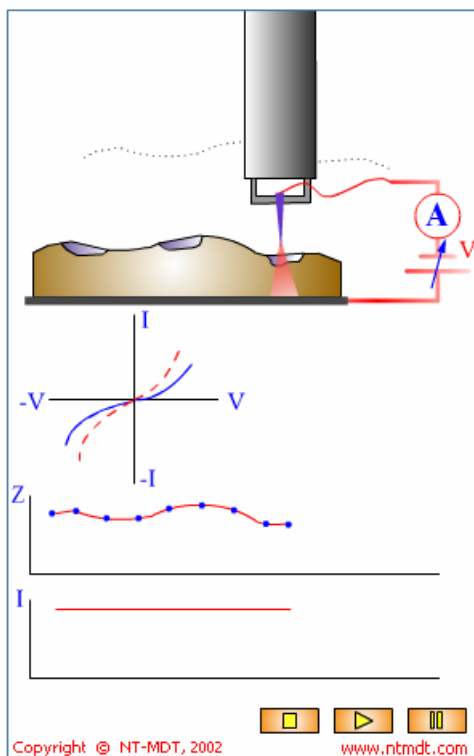
www.elsevier.com/locate/commatsci

Structural and electronic properties of pentacene molecule and molecular pentacene solid

R.G. Endres <sup>a,\*</sup>, C.Y. Fong <sup>b</sup>, L.H. Yang <sup>c</sup>, G. Witte <sup>d</sup>, Ch. Wöll <sup>d</sup>

# STM spectroscopy I

## I(V) Spectroscopy.



In I(V) Spectroscopy (or Current Imaging Tunneling Spectroscopy, CITS) a normal topographic image is acquired at fixed  $I_0$  and  $V_0$ . At each point in the image feedback loop is interrupted and the bias voltage is set to a series of voltages  $V_i$  and the tunneling current  $I_i$  is recorded. The voltage is then returned to  $V_0$  and the feedback loop is turned back on. Each I-V spectra can be acquired in a few milliseconds so there is no appreciable drift in the tip position. This procedure generates a complete current image  $I_i(x,y)$  at each voltage  $V_i$  in addition to the topographic image  $z(x,y)|_{V_0}$ .

CITS data can be used to calculate a current difference image

$\Delta I_{V_i, V_j}(x,y)$  where  $V_i$  and  $V_j$  bracket a particular surface state, producing an atomic resolved, real space image of a surface state. This technique, for example can be used in UHV to image filled ad-atom states or the dangling bond states for silicon reconstructions.

### References

1. G. Binnig and H. Rohrer: Surf. Sci. 126 (1983) 236. Rep. Prog. Phys. 55, 1165-1240 (1992).

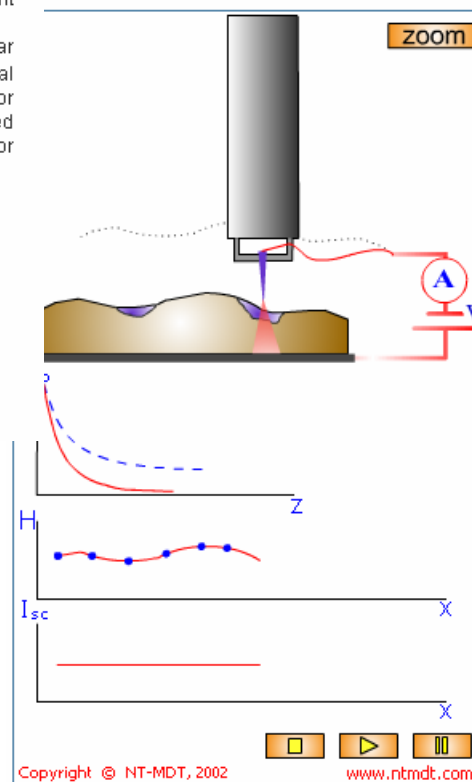
Download [Flash model](#)

Copyright © NT-MDT, 2002

www.ntmdt.com

At one point over the surface,  $I$  vs  $V$  and  $I$  vs  $Z$  curves can be acquired

## I(z) Spectroscopy.



The I(z) Spectroscopy is related to LBH spectroscopy and can be used for providing an information about the z-dependence of the microscopic work function of the surface. Next important use of the I(z) Spectroscopy is concerned with for testing of the STM tip quality.

The tunneling current  $I_T$  in STM exponentially decays with the tip-sample separation  $z$  as

$$I_T \sim \exp(-2kz),$$

where the decay constant is given by

$$2k = 2(2mU/h^2)^{1/2}$$

$U$  is the average work function  $U_{av} = (U_s + U_t)/2$ , where  $U_t$  and  $U_s$  are the tip and sample work functions, respectively.

In the I(z) Spectroscopy, we measure the tunnel current versus tip-sample separation at each pixel of an STM image. For  $U_{av} = 1$  eV  $2k = 1.025 \text{ \AA}^{-1} \text{ eV}^{-1}$ .

Sharp I(z) dependence helps in determining of tip quality. As is empirically established if tunnel current  $I_T$  drop to one-half with  $Z < 3 \text{ \AA}$  the tip is considered to be very good, if with  $Z < 10 \text{ \AA}$ , then using this tip it is possible to have an atomic resolution on HOPG. If this takes place with  $Z > 20 \text{ \AA}$  this tip should not be used and must be replaced.

### References

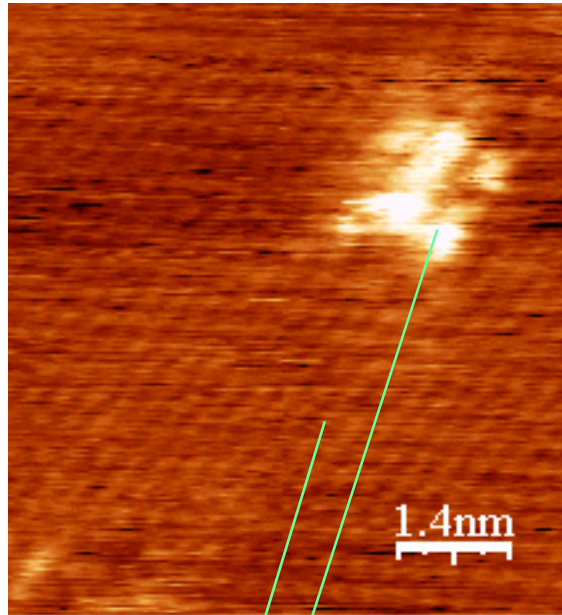
Copyright © NT-MDT, 2002

www.ntmdt.com

STM offers analytical tool with unique space resolution capabilities



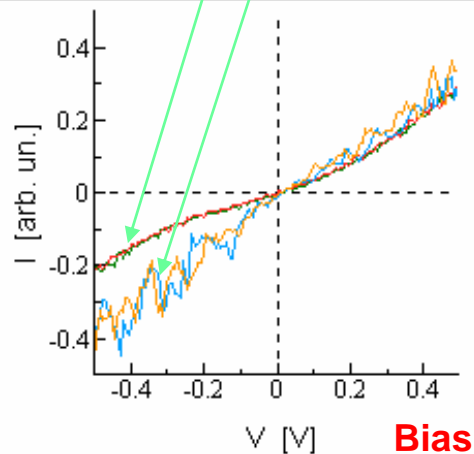
## STM spectroscopy II: an example



I-V curves can be acquired at different positions

STM image -  
plan view  
Cs on HOPG  
(bias 0.5 V)  
8 h after dep.

Tunneling  
current  
(conversion  
factor  $10^8$ )



Typical STM I-V  
curves of different  
regions  
(covered/uncovered)

Possibility to discriminate  
“conductivity” of small-sized  
samples with an excellent  
space resolution

## Observed quantities in STM

Thus, using formulas (2)–(5), we can compute the tunnel current at given system parameters and plot current-voltage characteristics. Fig. 4 shows theoretical tunneling current-applied voltage plot in case of carbon electrode 1 ( $\phi_1 = 4,7 \text{ eV}$ ) and platinum electrode 2 ( $\phi_2 = 5,3 \text{ eV}$ ) at  $\delta_z = 5 \text{ \AA}$  and contact area  $S = 10^{-17} \text{ m}^2$ .

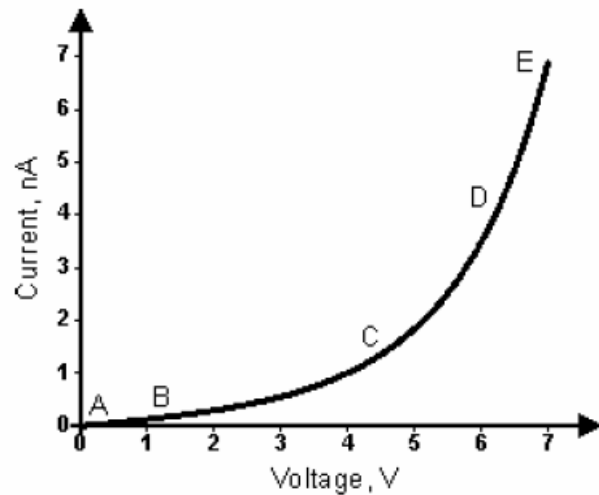


Fig. 4. Current-voltage characteristic for carbon electrode 1 and platinum electrode 2 at  $\delta_z = 5 \text{ \AA}$  and contact area  $10^{-17} \text{ m}^2$ .

Parts of  $J(V)$  curve correspond to the following expressions:  
 AB – (22), BC – (23), CD – (24), DE – (25).

✓ J-V curve depends on materials

✓ J is representative of local density of states

### Current-Voltage Characteristic

Measurement of relation between tunneling current and probe-sample voltage is carried out in  $J(V)$  spectroscopy mode. The  $J(V)$  spectroscopy is based on the dependence of tunneling current on number of electron states  $N$ , forming a tunneling contact of conductors, in the energy range from the Fermi level  $\mu$  to  $\mu - eV$  (Fig. 1), which at  $T = 0$  gives (see (7) in [chapter "John G. Simmons Formula"](#))

$$J \propto N = \int_{\mu - eV}^{\mu} \xi(E_z) dE_z \quad (1)$$

Thus the tunneling current dependence  $J(V)$  at constant tip-sample separation  $\delta_z$  represents an allocation of torn bonds as well as other electron states corresponding different energies, i.e. energy band structure of either tip or surface. Function  $\xi(E_z)$ , which was introduced in (6) of [chapter "John G. Simmons Formula"](#), depends on electron state density of phase space plane which is normal to tunneling direction at given  $E_z$ .

Using expression (1) and  $J(V)$  curve at constant tip-sample separation  $\delta_z$ , it is possible to compute the density of electronic states:

$$\frac{dJ}{d(eV)} \propto \xi(\mu - eV) \quad (2)$$

Thus, inspection of  $J(V)$  and its derivative  $dJ/d(eV)$  curves allows to investigate energy levels distribution with atomic resolution. It is possible to determine a conductivity type, in particular for semiconductors – to detect the valence band, conductivity band and impurity band. [1]–[3].

# Tunneling current spectroscopy

According to (2) and (3) from [chapter "John G. Simmons Formula in a Case of Small, Intermediate and High Voltage \(Field Emission Mode\)"](#) tunneling conductivity  $G = dJ/dV$  does not depend on applied voltage  $V$  in case  $eV \ll \bar{\phi}$ .

$$G = \frac{\gamma\sqrt{\bar{\phi}}}{\delta_z} \exp(-A\delta_z\sqrt{\bar{\phi}}) \quad (3)$$

at  $eV < \phi_2$  relation between  $G$  and  $V$  is parabolic

$$G \approx \gamma\sqrt{\bar{\phi}} \exp(-A\sqrt{\bar{\phi}})(1 + 3\sigma V^2) \quad (4)$$

On Fig. 2, 3 experimental dependences  $J(V)$ ,  $G(V)$ , which were measured for Pt and HOPG samples and Pt-Ro probe using STM Solver P47, is shown. Experimental data are in good agreement with the theoretical predictions (1)–(4).

**Local density of states can be readily probed**

**STM useful to reconstruct surface properties at the atomic level**

$$J \propto N = \int_{\mu-eV}^{\mu} \xi(E_z) dE_z$$

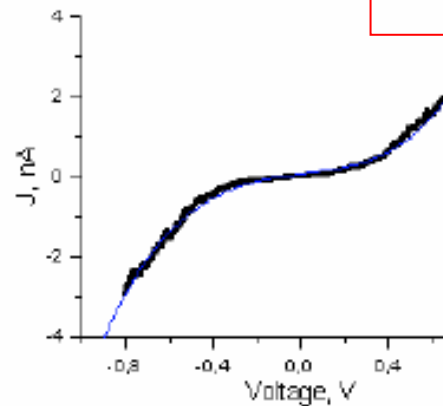


Fig. 1a. Experimental (points) and theoretical dependences  $J(V)$  for Pt

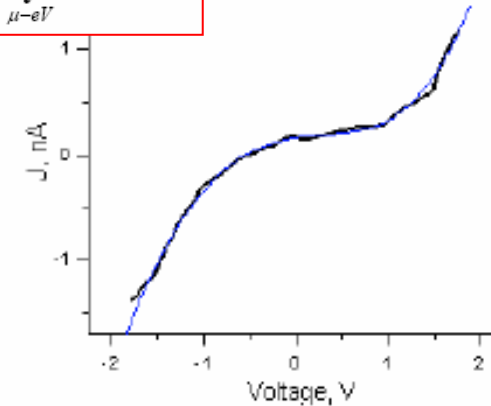


Fig. 2a. Experimental (points) and theoretical (solid line) dependences  $J(V)$  for HOPG

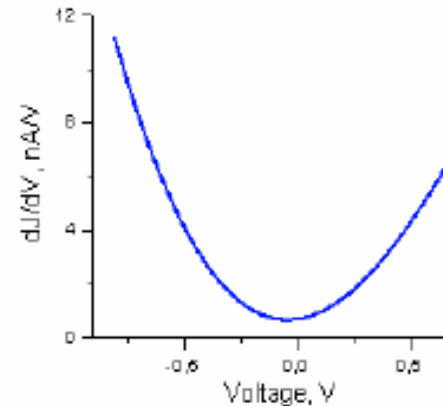


Fig. 1b. Experimental dependence  $G(V)$  for Pt

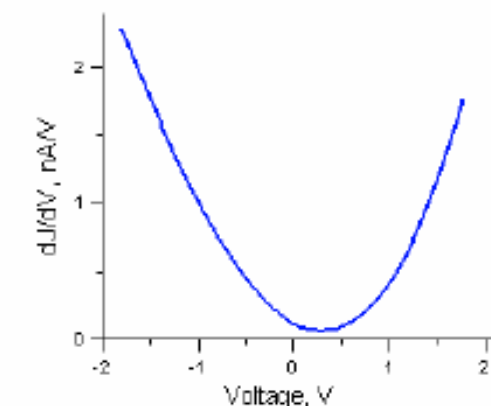


Fig. 2b. Experimental dependence  $G(V)$  for HOPG

$$\frac{dJ}{d(eV)} \propto \xi(\mu - eV)$$

# Current distance spectroscopy

## Current-Distance Characteristic

Measurement of relation between tunneling current and tip-sample distance is carried out in  $I(\delta_z)$  spectroscopy mode. According to (11) from [chapter "John G. Simmons Formula"](#), in absence of a condensate, a typical current-height relation is an exponential current decay (Fig. 1) with a characteristic length of a few angstrom [1], [2].

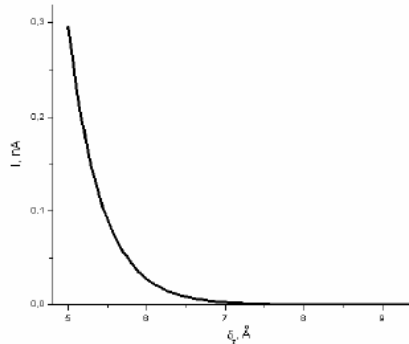


Fig. 1. Theoretical  $I(\delta_z)$  curve for Pt sample and Pt-Ro probe

Inspecting an experimental  $I(\delta_z)$  curve, it is possible to estimate the potential barrier height  $\bar{\phi}$ . If tip-sample bias  $V$  is small enough, then

according to [chapter "John G. Simmons Formula in a Case of Small, Intermediate and High Voltage \(Field Emission Mode\)"](#), tunneling current can be written as

$$I = \frac{\gamma S \sqrt{\bar{\phi}} V}{\delta_z} \exp(-A \delta_z \sqrt{\bar{\phi}}) \quad (1)$$

where  $\gamma = \frac{e\sqrt{2m}}{4\beta\pi^2\hbar^2}$ ,  $A = 2\beta\sqrt{\frac{2m}{\hbar^2}}$ ,  $S$  - contact area,  $m$  - free electron mass,  $e$  - elementary charge,  $\hbar$  - Planck's constant.

From (1)  $\bar{\phi}$  can be expressed through some analytical function of  $d \ln(I)/d\delta_z$ . Finding the natural logarithm of (1) and differentiate the result by  $\delta_z$  one can obtain

$$\frac{d \ln(I)}{d\delta_z} = -\frac{1}{\delta_z} - A\sqrt{\bar{\phi}} \quad (2)$$

If  $\delta_z \gg 1/A\sqrt{\bar{\phi}}$ , then  $\bar{\phi}$  can be expressed from (2) by following way

$$\bar{\phi} = \frac{1}{A^2} \left( \frac{d \ln(I)}{d\delta_z} \right)^2 \approx 10^{-20} \left( \frac{d \ln(I)}{d\delta_z} \right)^2 eV \quad (3)$$

where  $d \ln(I)/d\delta_z$  is expressed in  $m^{-1}$ .

Emphasize, that in most cases the condition  $\delta_z \gg 1/A\sqrt{\bar{\phi}}$  realizes practically always. For instance, if  $\bar{\phi} = 4eV$ , then expression (3) will be correct at  $\delta_z \gg 0,5 \text{ \AA}$ .

Experimental  $I(\delta_z)$  curve for Pt-flim which was measured using Pt-Ro probe in STM (Solver P47) on air, is shown on Fig. 2.

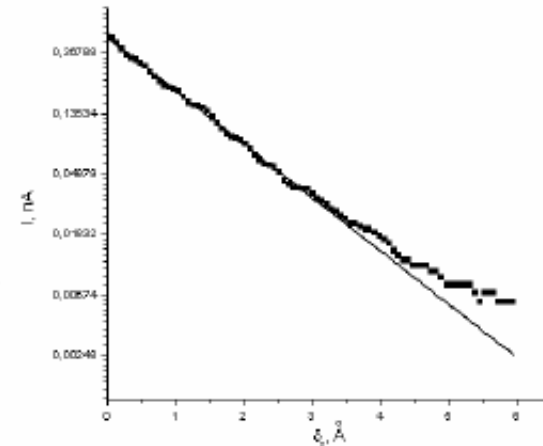


Fig. 2. Experimental  $I(\delta_z)$  curve (semilogarithmic scale). Solid line - approximation  $I = 0,47 \exp(-10\delta_z/11)$ .

**Exponential decrease  
useful to point out surface  
electronic properties**

# Density of states and workfunction imaging

## Measurements of the Electronic States Density

In chapter [Current-Voltage Characteristic](#) we considered the spectroscopy of electronic states. Meanwhile, at given  $eV$  it is possible to measure the electronic states distribution across the sample surface.

The electronic states density distribution measurement is performed in parallel with surface topography imaging in the  $J = const$  mode. Instead of constant tip bias  $V_0$ , the alternating voltage  $V = V_0 + b \sin \omega t$  is applied between sample and tip, where  $b \sin \omega t$  – the alternating signal having amplitude  $b \ll V_0$  (Fig. 1). Then, the net tunneling current is proportional to the following

$$J \propto N_0 + N_a \quad (1)$$

Thus, the total current flowing through the tunneling gap is equal to  $J = J_0 + J_a$ , where  $J_a$  – alternating component. Because  $J_0$  is held constant during the scan and  $eb = const$ , the alternating tunneling current amplitude is proportional to the electronic states density  $\xi(\mu - eV_0)$ . Hence, measuring the alternating current amplitude during scan allows for the mapping of the electronic states density in standard units. Since  $b \ll V_0$ , then  $J_a/b$  is actually  $dJ/dV$ .

The frequency  $\omega$ , as mentioned above, should be much more than the reciprocal feedback integrator time constant and be limited by maximum permissible scan frequency.

## Work-Function Distribution Study

In chapter [Current-Distance Characteristic](#) we considered that it is possible to estimate average electron work-function of electrodes using current-distance experimental curves. However, these measurements give work-function values only in small region where one electrode is located over another. In STM there is another method which is

The work function distribution measurement is performed in parallel with surface topography imaging in the  $J = const$  mode. In this case, the Z-axis piezo tube motion is determined not only by the feedback signal but also by application of an alternating signal producing motion law  $\Delta Z = a \cos(\omega t)$ . Accordingly, the tip-sample separation is  $\delta_z = \delta_{z0} + a \cos(\omega t)$ , where parameter being  $a \ll \delta_{z0}$ ,  $\delta_{z0}$  – tip-sample separation held. If voltage applied between tip and sample is small  $V \approx 0$ , then according to designations introduced, expression (2) from [chapter 1.2.2](#) can be transformed to the following

$$J = \frac{\gamma \sqrt{\phi} V}{\delta_{z0} + a \cos(\omega t)} \exp\left(-A \sqrt{\phi} (\delta_{z0} + a \cos(\omega t))\right) \approx J_0 \left[1 - A \sqrt{\phi} a \cos(\omega t)\right] \quad (1)$$

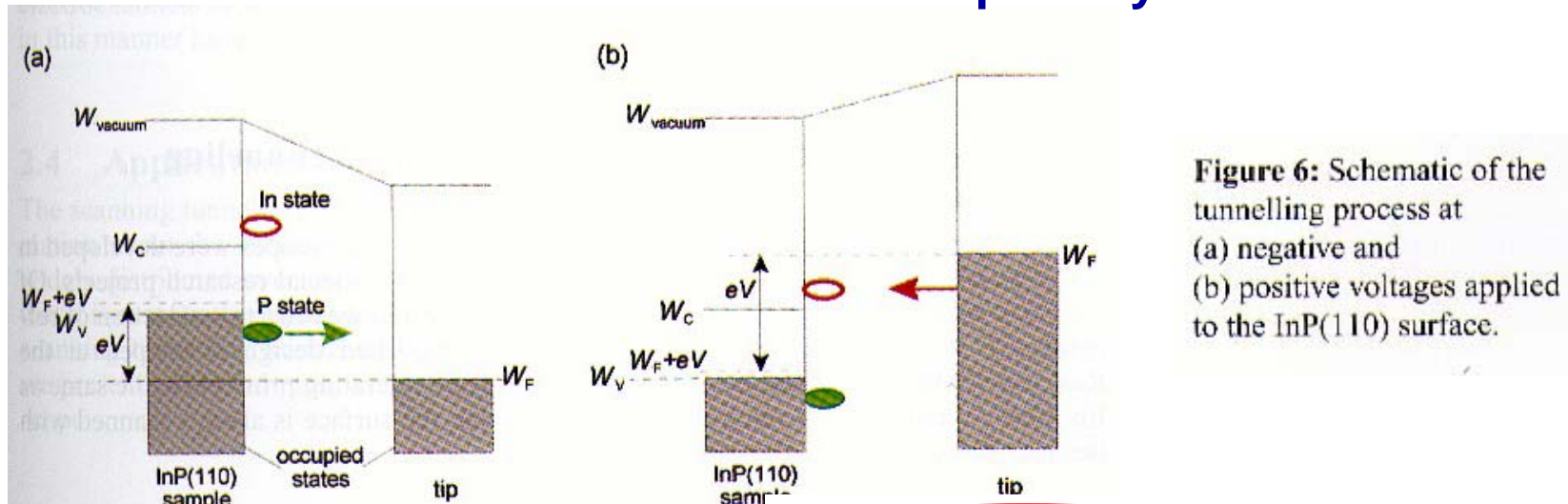
where  $J_0 = J(\delta_{z0})$ .

Thus, total current flowing through the tunneling gap in this case is equal to  $J = J_0 + J_a$ , where  $J_a$  – alternating tunneling current. Because  $J_0$  is held constant during the scan, the alternating tunneling current amplitude is proportional to the square root of the tip and the sample work function half-sum. Assuming that tip work function is constant during

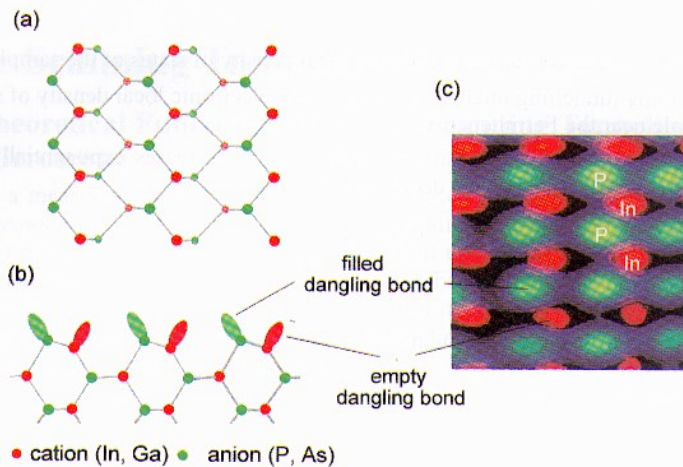
The frequency  $\omega$ , as mentioned above, should be much more than the reciprocal feedback integrator time constant and be limited by maximum permissible scan frequency.

**Modulation techniques allow for imaging spectroscopic properties**

# STM of semiconductors: bias polarity effects



**Figure 6:** Schematic of the tunnelling process at (a) negative and (b) positive voltages applied to the InP(110) surface.



**Figure 7:**  
 (a) Schematic top view and (b) side view of the (110) surfaces of III-V compound semiconductors.  
 (c) Superposition of two scanning tunnelling microscope images measured at positive (red) and negative (green) voltage. The density of state maxima correspond to the surface states at the In and P atoms, respectively.

In the special case of the InP(110) surface, the occupied surface state is spatially located above the P atoms, whereas the empty state is bound to the In atoms (Figure 7a,b). The P and In atoms are alternately arranged in zigzag rows. At negative sample voltages, the scanning tunnelling microscope probes the occupied states located at the P sublattice, whose electrons tunnel into the empty states of the tunnelling tip (Figure 6a). Conversely, only the empty surface states at the In sublattice are probed at positive voltages applied to the sample (Figure 6b) [6] – [8]. If the voltage polarity is changed every scan line, i.e. the occupied and the empty states are probed each alternating scan line, the two resulting images can be superimposed and the zigzag rows of alternating "In" and "P" atoms become visible (Figure 7c).

Apart from the spatial distribution of the density of states, its energy dependence is also of interest, and it should be possible to determine this dependence from current-voltage characteristics using Eq. (5). In order to do so, however, information is required about the transmission coefficient, which turns out to be a great obstacle even if approximations [9] are used. Therefore, in most cases, an experimentally viable approach is used, in which the density of states is approximated as follows [10], [11]:

$$\rho_{\text{sample}}(eV) \approx (dI/dV)/(IV) \quad (6)$$

It is thus possible to experimentally measure the density of states as a function of the energy relative to the Fermi level.

**In semiconductors:  
 contrast mechanisms related to bias polarity**

# STM of semiconductors: resonant levels

In semiconductors:  
(local) aps and energy levels detectable

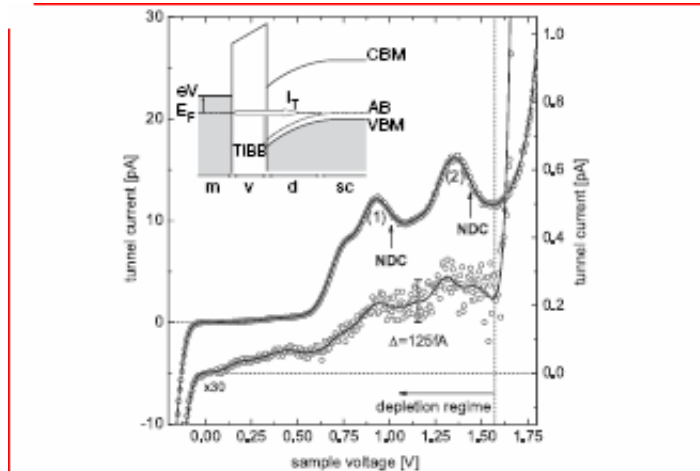


FIG. 2.  $I(V)$  characteristics of a buried carbon acceptor (upper curve) and the undisturbed surface (lower curve) obtained by spatially resolved  $I(V)$  spectroscopy (STS). The raw data are plotted as points, averaged data as lines. The  $I(V)$  curve taken on the undisturbed surface is scaled by a factor of 30 with respect to the one taken on the C acceptor. Note that the spectroscopic features indexed with (1) and (2) are followed by negative differential conductivity (NDC). Noise level for both curves is smaller than 125 fA. The inset sketches the energetic conditions of the depletion layer regime in the voltage interval from 0 V to 1.57 V in a rigid-band model: (m) metallic tip, (v) vacuum barrier, (d) depletion layer, (sc) bulk semiconductor, (TIBB) depletion barrier height, (VBM) valence band maximum, (CBM) conduction band minimum, (AB) acceptor band, and ( $I_T$ ) tunneling channel.

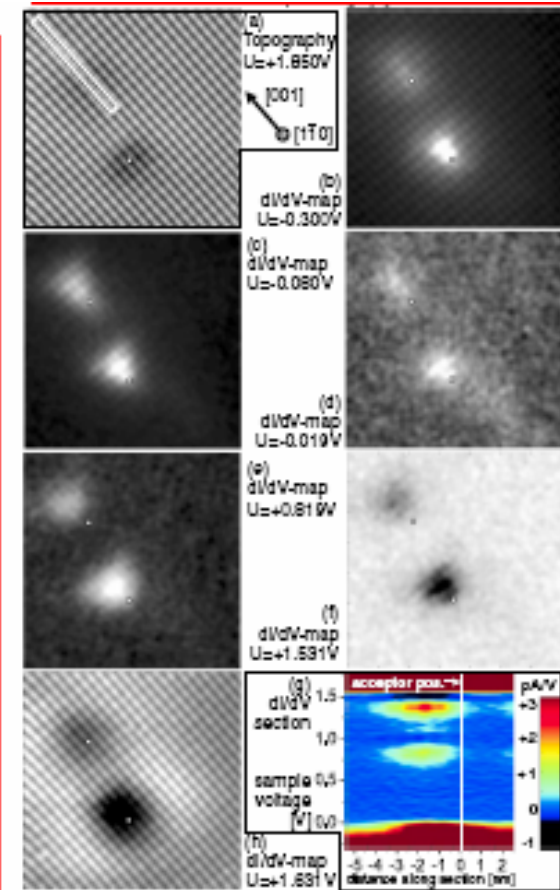
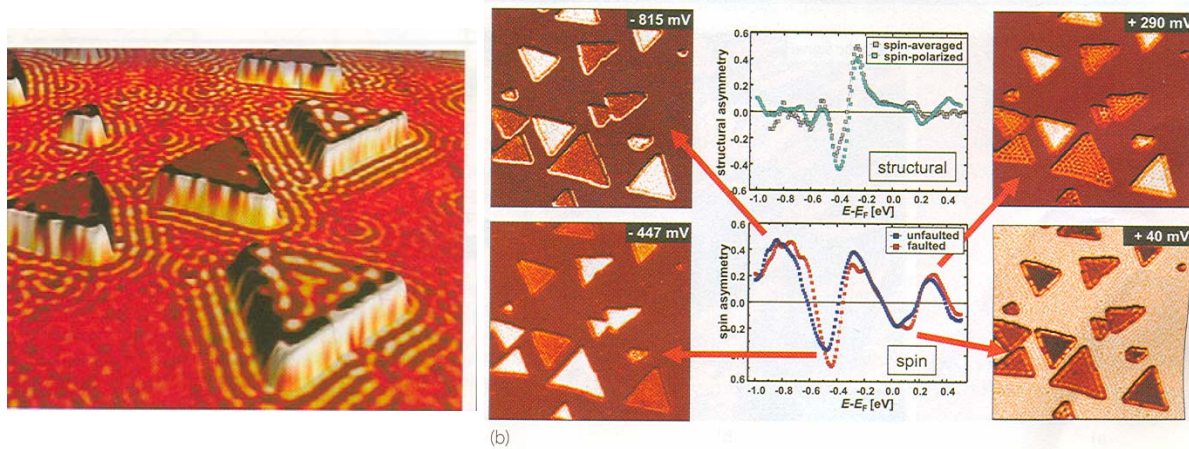


FIG. 3 (color online). Spatially resolved  $I(V)$  spectroscopy of two buried zinc acceptors near the  $(1\bar{1}0)$  surface: (a) simultaneously recorded topographic image at set point voltage +1.85 V and current 1 nA. (above) Differential conductivity spectra and excitation spectra  $d^2I/dV^2$  are acquired above one of the triangular contrasts (solid line) and above the undisturbed surface (dashed line) in the upper right corner of the imaged area.  $dI/dV$  maps: (b)–(e) Regions of enhanced conductivity shown in white, (f) NDC shown in black, (h) charge density oscillations around the acceptor atoms.  $dI/dV$  section (g) showing energy dependent change in contrast extension. The cross section is indicated by a white rectangle in (a). Projected dopant atom positions are marked by white circles.

# An STM variant: sensitivity to spin

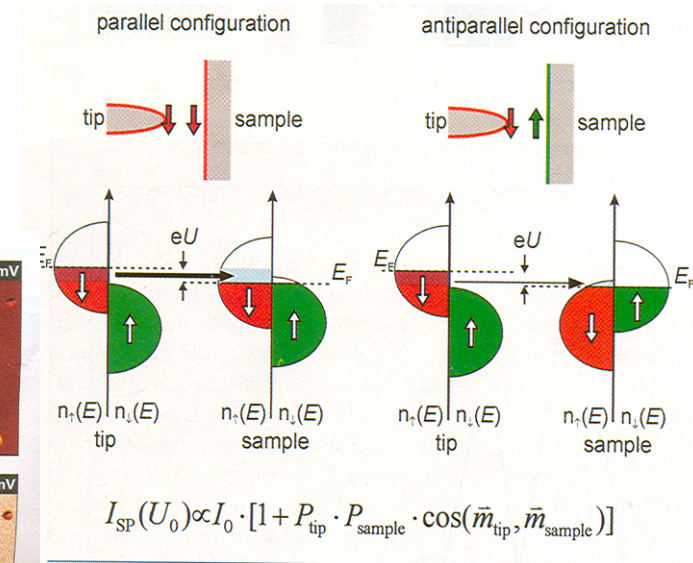
## Mapping spin structures on the atomic scale [DOI: 10.1051/EPN:2007004]

Roland Wiesendanger,  
 Institute of Applied Physics, University of Hamburg, D-20355 Hamburg • Germany  
 Email: wiesendanger@physnet.uni-hamburg.de • Web: www.nanoscience.de



(b)

europysicsnews



▲ Fig. 1: Principle of Spin-Polarized Scanning Tunneling Microscopy (SP-STM): the spin-polarized tunnelling current flowing between a magnetic tip and a magnetic sample depends on the relative alignment of the local magnetization of tip and sample as well as on the spin polarization of electronic states of tip and sample contributing to the tunnelling current.

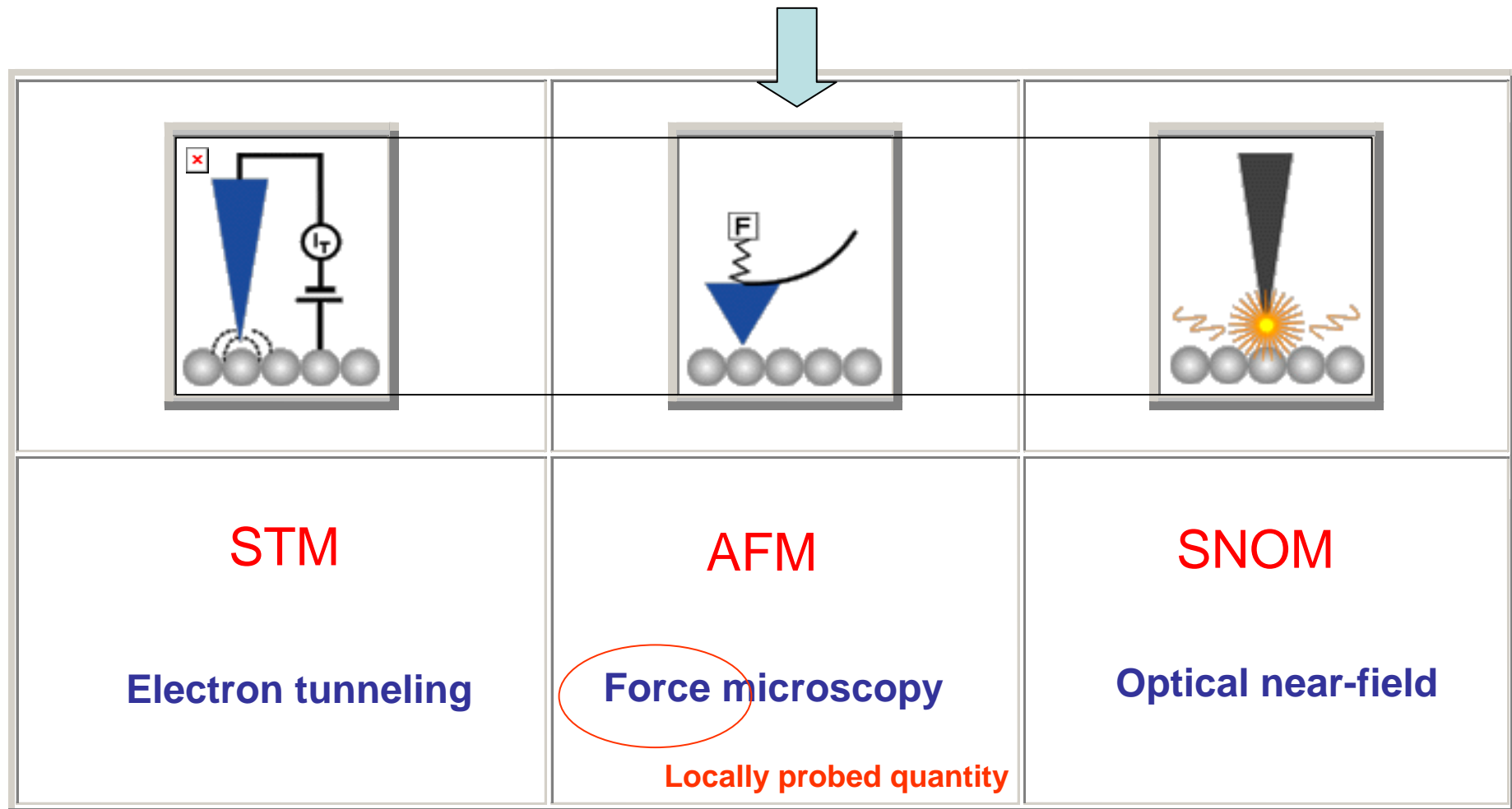
16 • volume 38 • number 2

▼ Fig. 3: Nanoscale Co islands of two atomic layers height grown on a Cu(111) substrate: (a) Perspective-view STS image exhibiting simultaneously the sample topography and the electronic standing wave patterns on the Cu (111) surface as well as the quantum mechanical electronic confinement states within the Co islands. (b) Top-view SP-STs images of the Co-islands on Cu(111) obtained at different samples biases corresponding to energies below or above the onset of the Cu(111) surface state. The sample bias voltages have been chosen in such a way that the energy-dependent spin asymmetry of the tunnel junction and therefore the magnetic contrast in the SP-STs images becomes large. Notice that the spin asymmetry not only varies in magnitude but even changes sign as a function of energy. At zero-crossings of the spin asymmetry curve no magnetic contrast is obtained [27,28].

**Magnetic-induced variations of electronic properties investigated by using a “magnetic” STM probe**

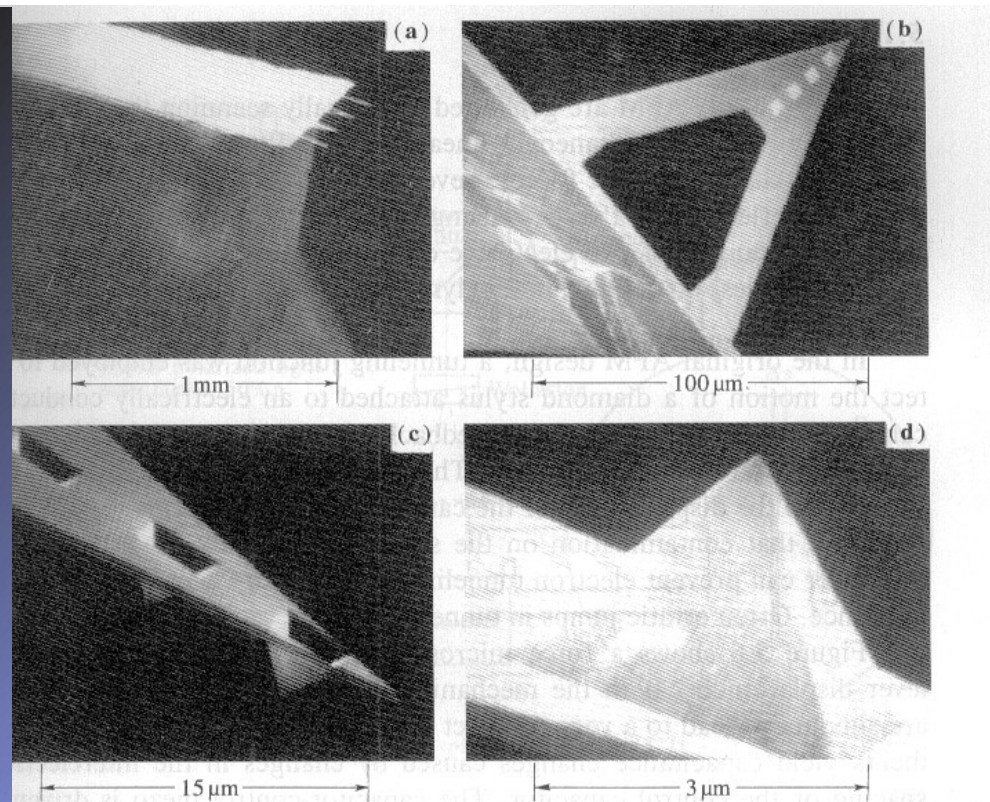
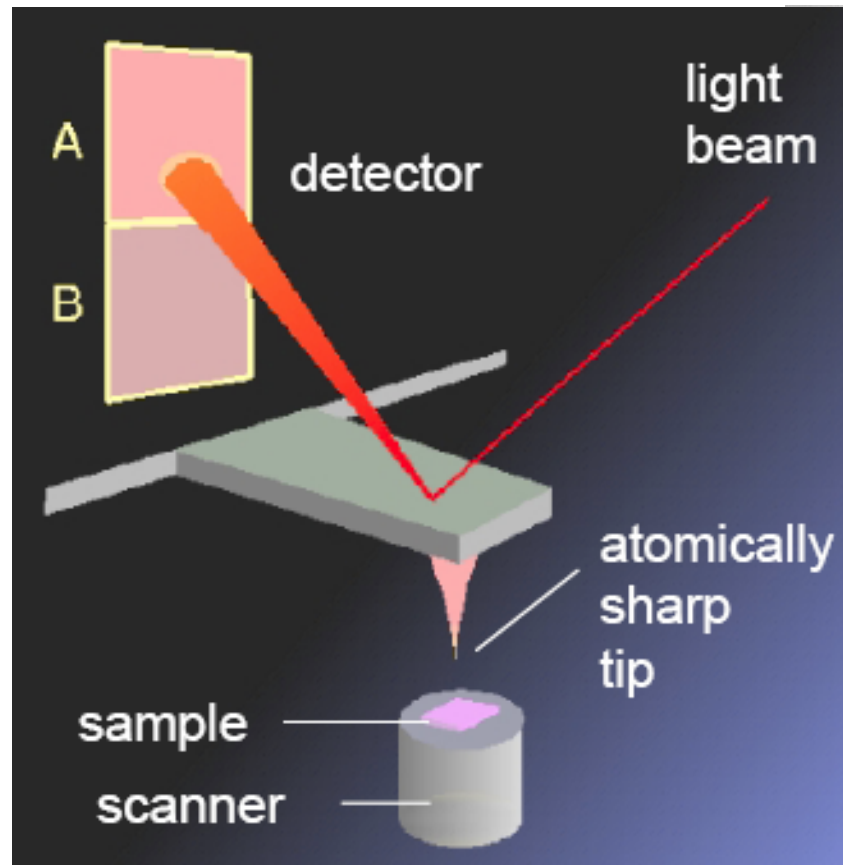


## 2. Scanning force microscopy (AFM and relatives)



*AFM is probably the most straightforward (and easy to understand/interpret) probe microscopy*

## AFM probes

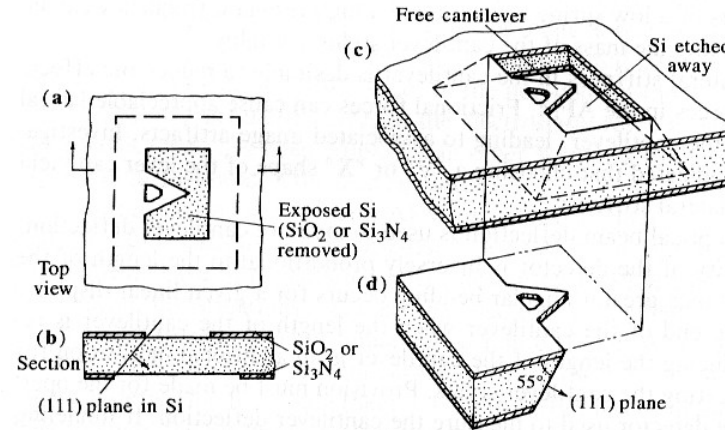
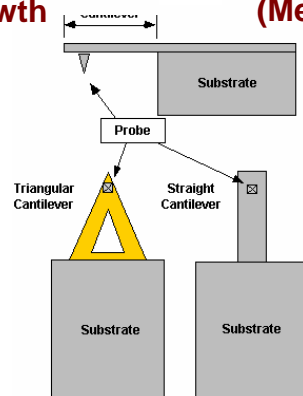
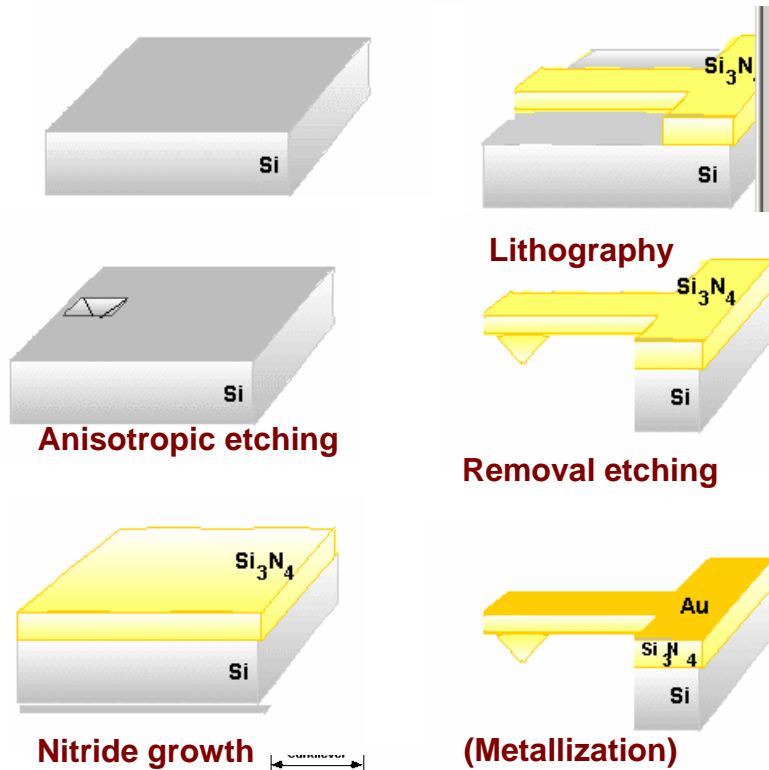


5.5a-d. SEM micrographs of  $\text{Si}_3\text{N}_4$  cantilevers with integrated pyramidal tips. (a) The  $\text{Si}_3\text{N}_4$  film is attached to the surface of a glass block with dimensions of  $2 \times 3 \times 0.7 \text{ mm}^3$ . Four cantilevers protrude from the edge of the block. (b) Four pyramidal tips can be seen at the end of this V-shaped cantilever. (c) The pyramidal tips are hollow when viewed from the back side. (d) Each tip has very smooth sidewalls, and the tip appears to terminate virtually at a point, with less than 30 nm radius [5.4]

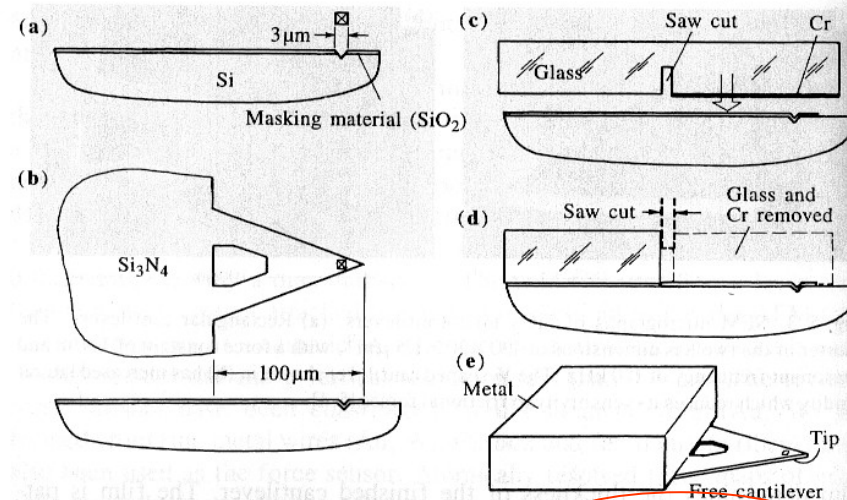
**The local character of AFM relies on the availability of suitable probes**

# Cantilever/tip fabrication: examples

The first step in the fabrication of an AFM tip is the etching of a single-crystal silicon wafer with specific crystalline orientation. This results in the forming of square pyramidal tips with characteristic angles.



**Fig. 5.2a-d.** Fabrication of thin-film microcantilevers. (a) A thin film of SiO<sub>2</sub> or Si<sub>3</sub>N<sub>4</sub> is formed on the surface of a (100) Si wafer and patterned to define the shape of the cantilever and to create openings on the top and bottom of the wafer. (b) The windows are aligned along (111) planes. (c) Anisotropic etching of the exposed Si with KOH undercuts the cantilever and self-terminates at the (111) planes as shown. (d) A small Si chip is cut from the wafer to serve as a pedestal for mounting the cantilever in the AFM [5.4]

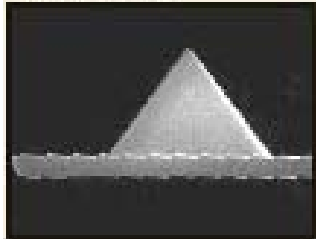


**Fig. 5.4a-e.** Fabrication of Si<sub>3</sub>N<sub>4</sub> microcantilevers with integrated pyramidal tips. (a) to (e) illustrate the steps in the fabrication process, see text [5.4]

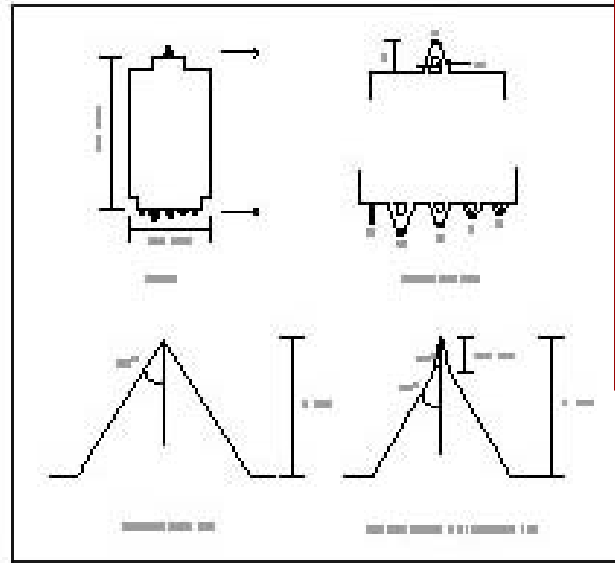
# Examples of commercial cantilevers

## FEATURES:

- Compatible with all major AFM brands.
- Typical radius of curvature: sharpened tips: < 20 nm, unsharpened tips: < 50 nm.
- Available with gold coating for high reflectivity.
- Recessed corners for easy sample approach.
- The widest range of spring constants commercially available on a single chip.



*Thermomicroscopes Microlevers are ideal for all contact imaging modes, force modulation microscopy, and liquid operation. The range in force constants enable users to image soft samples in contact as well as high load force vs. distance spectroscopy.*



Many different cantilevers are commercially available

They are different for:

- Dimensions and shape, typ 0.1-0.5 mm;
- Elastic constant (materials and design, typ 0.05-50 N/m;
- Tip coating (conductive, super-hard, etc.)

Typical Mechanical Characteristics

Cantilever type	A - triangular	B - rectangular	C - triangular	D - triangular	E - triangular	F - triangular
Standard mode of operation	Contact					
Cantilever length	180 $\mu\text{m}$	200 $\mu\text{m}$	320 $\mu\text{m}$	220 $\mu\text{m}$	140 $\mu\text{m}$	85 $\mu\text{m}$
Cantilever width	18 $\mu\text{m}$	20 $\mu\text{m}$	22 $\mu\text{m}$	22 $\mu\text{m}$	18 $\mu\text{m}$	18 $\mu\text{m}$
Cantilever thickness	0.6 $\mu\text{m}$	0.6 $\mu\text{m}$	0.6 $\mu\text{m}$	0.6 $\mu\text{m}$	0.6 $\mu\text{m}$	0.6 $\mu\text{m}$
Force Constant	0.05 N/m	0.02 N/m	0.01 N/m	0.03 N/m	0.10 N/m	0.50 N/m
Resonant Frequency	22 kHz	15 kHz	7 kHz	15 kHz	38 kHz	120 kHz

Ordering Information

Microlevers	Sharpened		Unsharpened	
	Gold coated*	Uncoated	Gold coated*	Uncoated
Half wafer - (250 chips)	MSCT-AUHW	MSCT-NGHW	MLCT-AUHW	MLCT-NGHW
Unmounted - (25 chips)	MSCT-AUNM	MSCT-NGNM	MLCT-AUNM	MLCT-NGNM
Mounted - (25 chips)	MSCT-AUMT-A	MSCT-NGMT-A	MLCT-AUMT-A	MLCT-NGMT-A
Mounted - (25 chips)	MSCT-AUMT-BF	MSCT-		F

\* Not for use with AutoProbe ME systems

Microlevers™

GENERAL PURPOSE CANTILEVERS

Cantilever choice depends for instance on:

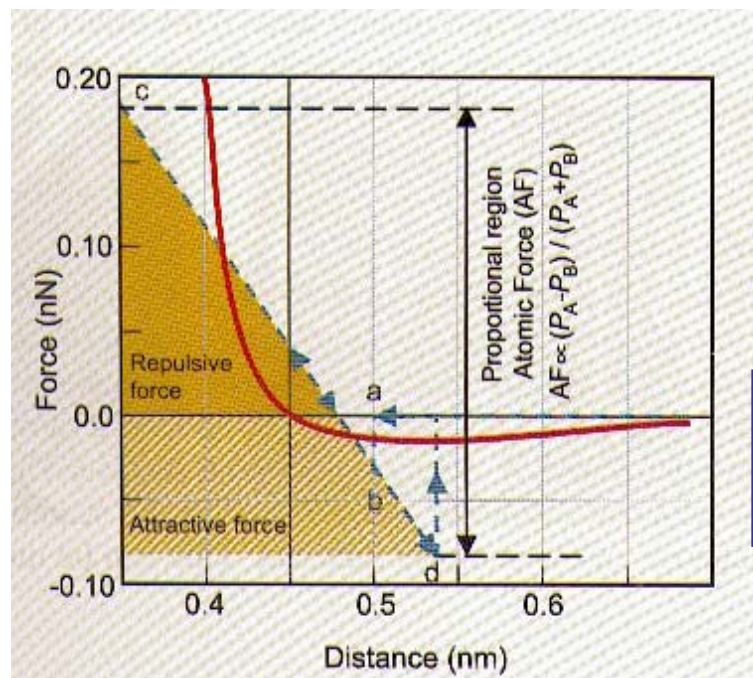
- Operation mode (contact/non contact);
- Quantities to be probed (e.g., if an electric field is needed, a conductive tip has to be used);
- Possible material manipulation (e.g., nanoindentation requires super-hard tips)

## Basics of tip/sample interaction

When the tip is approached to the sample (at sub-nm distance!), forces depend roughly on van der Waals interaction between the apical tip atoms and the surface

At “large” distance forces are weakly attractive, at “short” distance they are repulsive

Surface topography (height variations) can be sensed by monitoring the force, i.e., the cantilever deflection



When tip/sample distance is kept in the repulsive region, **contact operating mode** is achieved

When tip/sample distance is kept (*mostly*) in the attractive region, **non-contact operating mode** is achieved

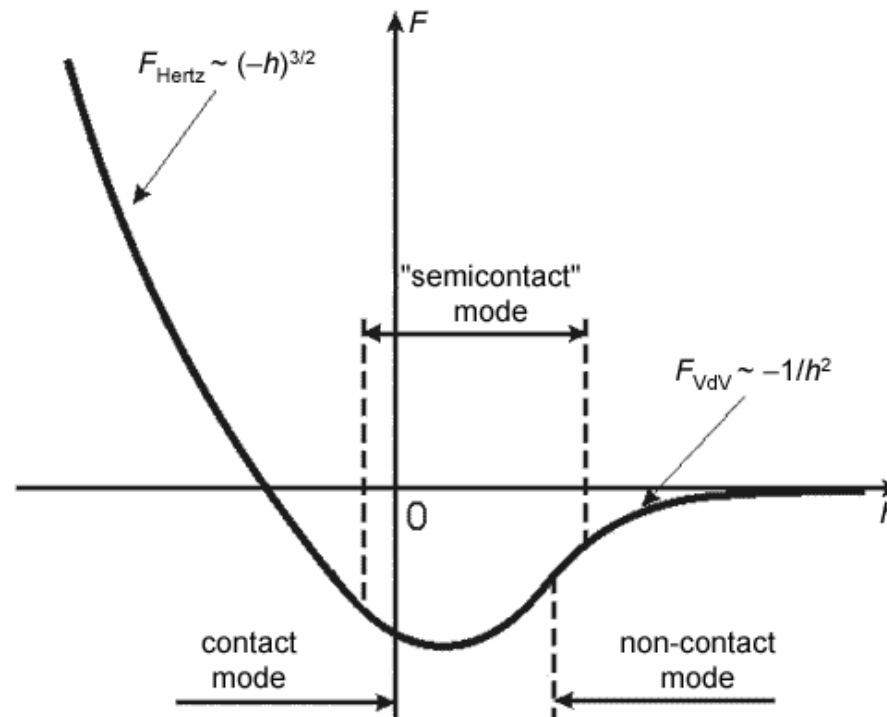
# Tip/sample forces

Sample investigation is available thanks to the forces acting between a cantilever and a surface. They are quite different. One or another force dominate at different tip-sample separations.

- During contact and the surface deformation by the cantilever, the elastic repulsion force dominates; this approximation is called the **Hertz model** and is considered in the chapter "[Elastic interactions. The Hertz problem](#)".
- At tip-sample separations of the order of several tens of angstrom the major interaction is the intermolecular interaction called the **Van der Waals force** (see chapter "[The Van der Waals force](#)").
- At the same distance between the tip and the sample and in the presence of liquid films, the interaction is influenced much by **capillary and adhesion forces**. The range of capillary forces considered in the chapter "[Capillary forces](#)" is determined by the liquid film thickness.
- At larger separations the **electrostatic interaction** starts to dominate. It is described in chapter "[Electrostatic force microscopy](#)".
- At separations of the order of a thousand of angstroms **magnetic forces** considered in the chapter "[Magnetic force microscopy](#)" prevail.



Various kind of forces do cooperate in providing the typical force vs distance behavior experienced in AFM



# A few words on elastic forces

When the cantilever and the sample are in contact, elastic forces start to act giving rise to both the sample and tip deformations which can affect the acquired image. To properly interpret the results and choose the measuring mode one should have a clear idea of elastic interactions in contact and "semicontact" modes.

Such consideration is necessary in order to:

- avoid tip or sample damage during scanning - even at low loading force the pressure in a contact zone can exceed the strength limit because contact area is very small.
- reconstruct properly the sample surface topography basing on the acquired image profile in case when surface features are of the same size as the tip curvature radius.
- analyze forces in the "semicontact" mode at a moment of the tip contact with the surface which directly affect the cantilever oscillation and are one of the damping reasons.

*Elastic deformations in the contact zone (the Hertz problem).*

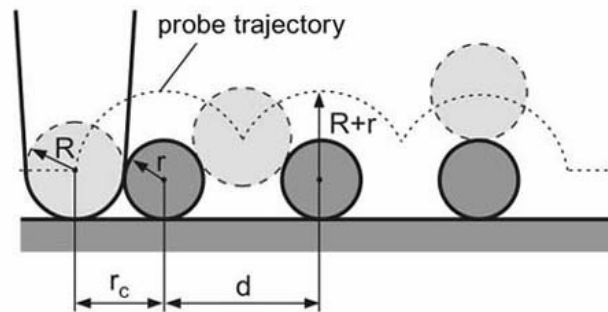
Let us consider first only the elastic force. **The Hertz problem** is deformations determination at local contact of bodies under load  $F$  action.

We have to adopt some simplifying assumptions [1].

1. Suppose that both the cantilever and sample materials are isotropic, i.e. their elastic properties are described only by two pairs of parameters – Young's moduli  $E, E'$  and Poisson ratios  $\mu, \mu'$ . (For the anisotropic materials the number of such independent elastic characteristics can reach 21).
2. Assume that in the vicinity of the contact point the undeformed parts of bodies surface in perpendicular planes orthogonal to the plane in the given point (**Fig. 1**) are described by two curvature radii  $r_1, r_2$  (for the tip) and  $r_1', r_2'$  (for the studied sample area).
3. Deformations are small compared to surfaces curvature radii.

The change in the probe vertical position during scanning in the contact mode produces profile which can differ much from the real surface topography. One of the reasons for that is the elastic deformation of the tip and the sample. For example, the decrease in the organic molecules vertical dimensions was experimentally established. Because these materials are very soft, the probe "indents" protrusions on their surfaces (See [Appendix 3](#) and [chapter 2.5.1](#)).

The second reason for the difference between scan profile and real surface geometry is the tip-sample convolution. Its consideration is important when studying small (of the order of the tip curvature radius) surface features. A finite tip dimension results in the lack of the ability to probe narrow cavities on the sample surface thus decreasing their depth and width. Similarly, convex features image appears wider. The convolution phenomenon is best understood from **Fig. 1**.



**Fig. 1.** Tip convolution during scanning. The scan profile can differ much from real surface geometry.

The Hertz problem solution relates the loading force  $F$  and the penetration depth  $h$ :

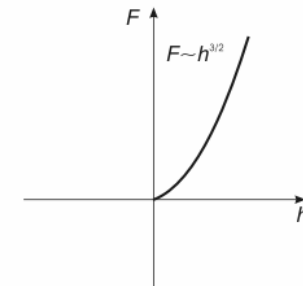
$$F = \frac{Ka^3}{R} = Kh^{\frac{3}{2}} \frac{1}{R^{\frac{1}{2}}} \quad (3)$$

Accordingly, the pressure is the following function of the force:

$$P = \frac{F}{\pi a^2} = \frac{1}{\pi} \sqrt[3]{\frac{FK^2}{R^2}} \quad (4)$$

The given solution for the case of two spherical bodies contact includes one important special case of the flat sample contact with the tip having curvature radius  $R$  ( $r = R, r' = \infty$ ).

Let us depict the Hertz problem solution, i.e. the dependence of the penetration depth (horizontal axis) upon the loading force (vertical axis) for positive  $h$ . In **Fig. 3**, the rising branch corresponds to the Hertz problem solution.

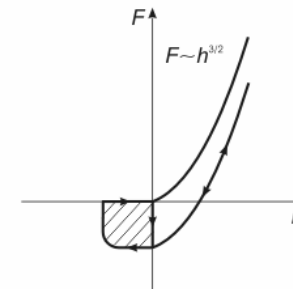


**Fig. 3.** Force  $F$  depending on penetration depth  $h$  (graph of the Hertz problem solution).

If one assumes the simultaneous effect of convolution and deformation, it becomes clear how much the image profile can differ from the real topography. In [Appendix 4](#) it is demonstrated that the acquired image needs to be analyzed and even computer processed in order to obtain the sample real topography.

*Non-elastic conservative contact forces.*

As tip makes contact with the sample, some other forces arise besides the elastic one. For example, the [Van der Waals interaction](#) (revealed not only when two bodies touch but within some distance between them) leads to the contact pressure decrease because Van der Waals forces in contrast to elastic ones are attractive but not repulsive.



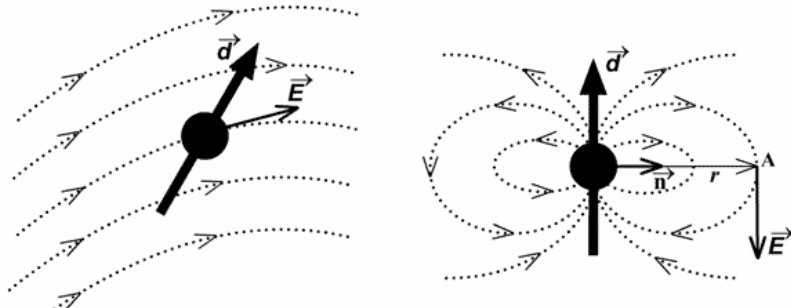
**Fig. 2.** Plots of force  $F$  vs. penetration depth  $h$ . Shown are the Hertz problem solution as well as solution with a hysteresis loop accounting for the nonconservative forces.

# A few words on van der Waals I

The Van der Waals force or the intermolecular attractive force has three components of slightly different physical nature but having the same potential dependence on the intermolecular distance  $\sim 1/r^6$ . This lucky circumstance allows to compare directly constants of interaction that correspond to three Van der Waals force components because proportions between them will be held constant at different  $r$  magnitudes. Constants at  $1/r^6$  multiplier will differ for various materials.

$$W = W_{\text{orient}} + W_{\text{ind}} + W_{\text{disp}} \sim 1/r^6 \quad (1)$$

All three Van der Waals force components are based on dipoles interaction, therefore we should remember two basic formulas:



the energy of dipole  $\mathbf{d}$  placed in field  $\mathbf{E}$  is [1]:

$$W_D = -\mathbf{d}\mathbf{E} \quad (2)$$

and the electric field produced by the dipole  $\mathbf{d}$  is [1]:

$$\mathbf{E} = \frac{3(\mathbf{nd})\mathbf{n} - \mathbf{d}}{r^3}, \quad |\mathbf{E}| \sim \frac{1}{r^3} \quad (3)$$

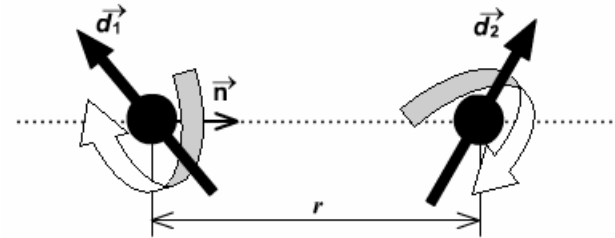
where  $\mathbf{n}$  – unit vector directed from the point at which the energy is determined to the dipole.

The orientational interaction (or the Casimir force) arises between two polar molecules each of which has the electric dipole moment. In accordance with (2), (3) the interaction energy of dipoles  $\mathbf{d}_1$  and  $\mathbf{d}_2$  separated by distance  $r$

$$W_D = \frac{\mathbf{d}_1 \mathbf{d}_2 - 3(\mathbf{d}_1 \mathbf{n})(\mathbf{d}_2 \mathbf{n})}{r^3} \sim \frac{1}{r^3} \quad (4)$$

depends sufficiently upon the molecules relative position. Here  $\mathbf{n}$  is the unit vector directed along the line between molecules.

In order to reach the potential minimum, dipoles tend to align along the common axis (Fig. 1). The thermal motion, however, breaks this order. To determine the "resulting" orientation potential  $W_{\text{orient}}$  one should average statistically interactions over all possible orientations of molecules pair. Notice that in accordance with the Gibbs distribution  $\exp(-W/kT)$ , which gives the probability of the system being in the state with energy  $W$  at temperature  $T$ , the energetically advantageous orientations are preferable. That is why despite the isotropy of possible mutual orientations, the average result will be nonzero.



Averaging with the use of the Gibbs distribution is performed in accordance with the following formula:

$$W_{\text{orient}} = \frac{\int W_D \exp\left(-\frac{W_D}{kT}\right) d\mathbf{v}}{\int \exp\left(-\frac{W_D}{kT}\right) d\mathbf{v}} \quad (5)$$

where, for the sake of normalization, the denominator is the statistical sum and  $\mathbf{v}$  is the integration parameter providing enumeration of all the system possible states (a pair of dipoles mutual orientations).

If  $W_D \ll kT$  the exponent can be approximated by the series expansion:

$$\exp\left(-\frac{W_D}{kT}\right) \approx 1 - \frac{W_D}{kT} \quad (6)$$

so the energy of orientation interaction is approximated as:

$$W_{\text{orient}} \sim \frac{\int W_D d\mathbf{v} + \int \frac{W_D^2}{kT} d\mathbf{v}}{\int d\mathbf{v} + \int W_D d\mathbf{v}} \quad (7)$$

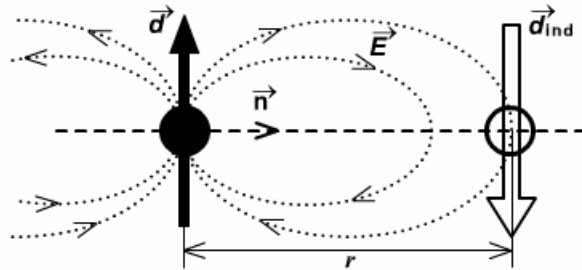
On performing integration it can be shown that  $\int W_D d\mathbf{v} = 0$ , thus,  $W_{\text{orient}} \sim W_D^2$ . Introducing constant  $A_1$  in accordance with (4), we finally have:

$$W_{\text{orient}} = \frac{\text{const}}{kT} \frac{1}{r^6} = -\frac{A_1}{r^6} \quad (8)$$



# A few words on van der Waals II

The **induction interaction** (or the **Debye force**) arises between polar and nonpolar molecules. Electric field  $\mathbf{E}$  generated by dipole  $\mathbf{d}_1$  polarizes the other molecule (Fig. 2). The induced moment calculated in the first order of the quantum perturbation theory is equal to  $\mathbf{d}_{\text{ind}} = \chi \mathbf{E}$  where  $\chi$  stands for the molecule polarizability.



Then, the potential of induction interaction is computed as follows:

$$W_{\text{ind}} = \mathbf{d}_{\text{ind}} \mathbf{E} = \chi \mathbf{E}^2 = \chi \frac{-3(\mathbf{n} \mathbf{d}_1)^2 + \mathbf{d}_1^2}{r^6} = -\frac{\chi \mathbf{d}_1^2}{2r^6} \sim \frac{1}{r^6} \quad (9)$$

Thus, this kind of interaction also "universally" depends on  $W = W_{\text{orient}} + W_{\text{ind}} + W_{\text{disp}} \sim 1/r^6$  though having the other reason and the other constant.

It should be noted that in liquids and solids the polarized molecule experiences the symmetric influence of many neighbor molecules, the induction interaction being strongly compensated by their action. The result is that the real induction interaction is estimated as:

$$W_{\text{ind}} \sim \frac{1}{r^n}, \quad n = 8 + 13 \quad (10)$$

Obviously, the force is determined by

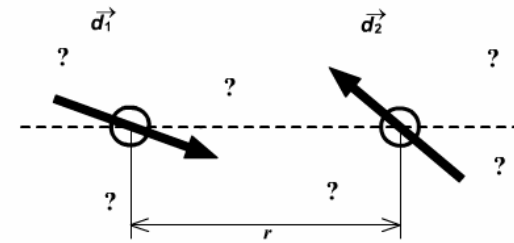
$$\mathbf{F} = -\text{grad} W_D, \quad W_{\text{disp}} = -\frac{A_3}{r^6} \quad (13)$$

Estimations of the Van der Waals attraction for AFM studies in the contact mode give:  
 $F_{\text{vdV}} \sim 10^{-8} + 10^{-9}$  N.

**Van der Waals interaction is based on different dipole/dipole interaction mechanisms, all leading to a  $r^{-6}$  behavior**

**This results into a  $r^{-m}$  force on the cantilever depending on specific tip shape and distance**

The **dispersion interaction** (or the **London force**) is a prevailing one because it involves nonpolar molecules as well. This third term in (1) is always presented that is why it is the major one.



3. Due to the quantum uncertainty, nonpolar molecules have "momentary" dipole moments, interaction between which is of the second order of smallness of the perturbation theory.

In a system of nonpolar molecules the electrons wave function  $\Psi$  is such that average values of dipole moments in any state  $n$  are equal to zero:  $\langle \psi_n | \mathbf{d}_{1,2} | \psi_n \rangle = 0$ . However, nondiagonal matrix elements  $\langle \psi_n | \mathbf{d}_{1,2} | \psi_m \rangle$  are nonzero. Moreover, the second quantum mechanical correction to the interaction energy calculated as is known [2] according to the formula below, is nonzero too:

In a system of nonpolar molecules the electrons wave function  $\Psi$  is such that average values of dipole moments in any state  $n$  are equal to zero:  $\langle \psi_n | \mathbf{d}_{1,2} | \psi_n \rangle = 0$ . However, nondiagonal matrix elements  $\langle \psi_n | \mathbf{d}_{1,2} | \psi_m \rangle$  are nonzero. Moreover, the second quantum mechanical correction to the interaction energy calculated as is known [2] according to the formula below, is nonzero too:

$$W_n^{(2)} = \sum_{n,m} \frac{|\langle \psi_n | W | \psi_m \rangle|^2}{\epsilon_n - \epsilon_m} \quad (11)$$

where perturbation  $W$  is given by (4),  $\epsilon_n, \epsilon_m$  – energies of the system of two molecules in arbitrary states  $n$  and  $m$ .

In a certain sense, "momentary" magnitudes of dipole moments (at zero average value) are nonzero and they interact (Fig. 3). In the second order of smallness the averaged magnitude of such "momentary" potential is not already vanished and namely this is the potential of dispersion interaction.

Correction (11) as is seen, is proportional to the square of perturbation  $W_D$ . From this it is clear that

$$W_{\text{disp}} \sim W_D^2, \quad W_{\text{disp}} = -\frac{A_3}{r^6} \quad (12)$$

Constant  $A_3 = \frac{3I_1 I_2}{2(I_1 + I_2)} \chi_1 \chi_2$  is called the Hamaker constant (here  $I_1, I_2$  – ionization potentials,  $\chi_1, \chi_2$  – molecules polarizability).

The classical interpretation of this interaction is as follows. The dipole moment of one molecule arisen from fluctuations, generates field which, in turn, polarizes the second molecule. The already nonzero field of the second molecule polarizes the first one. The potential of this peculiar system with a "positive feedback" is calculated similarly to the induction interaction.

# Not to forget: adhesion/capillary effects

Let us examine the effect of the surface tension on AFM measurements. [1] At the moment of a cantilever contact with a liquid film on a flat surface, the film surface reshapes producing the "neck". The water wets the cantilever surface (Fig. 1) because the water-cantilever contact (if it is hydrophilic) is energetically advantageous as compared to the water-air contact. Notice that in such cases the contact angle is always less than  $90^\circ$ .

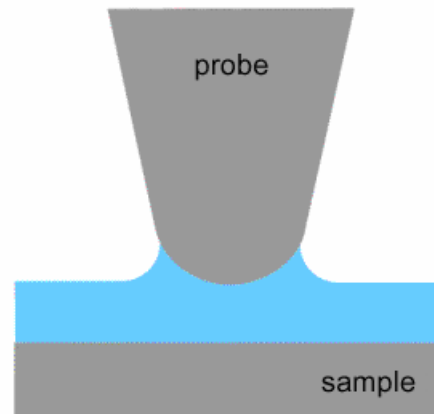


Fig. 1. "Neck" formation.

It is intuitively clear that the neck curved surface will tend to flatten that is available only at the expense of the cantilever pulling down. This means that the cantilever attracts to the sample.

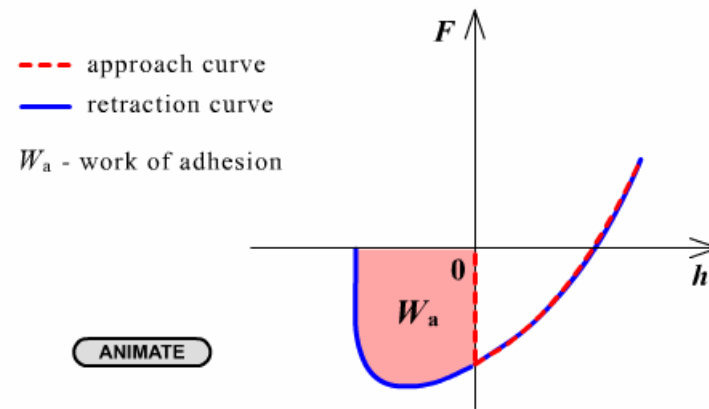
Calculation of this attraction force is a simple task. Let the tip curvature radius be much larger than other characteristic dimensions of the case. In Fig. 2 the following designations are introduced:  $D$  – tip-sample separation,  $d$  – "immersion depth",  $h$  – film thickness,  $\rho_1$  – lesser curvature radius of the liquid surface,  $\rho_2$  – tip-liquid contact area radius.

We will not concentrate attention on  $d$  value determination. For the estimation purpose we will use the maximum value of the capillary attraction force that takes place at  $D = 0$ . In this case the unknown parameter  $d$  vanishes:

$$F_{\text{cap}} = F_{\text{max}} = 4\pi R \sigma \cos \theta \quad (5)$$

Taking into account that cantilever radius  $R$  is  $10 \text{ nm}$ , water surface tension at  $20^\circ \text{C}$  is equal to  $0.073 \text{ N/m}$  and the contact angle is small i.e.  $\cos \theta$  is close to 1, we get the following estimation:  $F_{\text{cap}} \sim 10^{-8} - 10^{-9} \text{ N}$ . Thus, the capillary force by the order of magnitude is the same as the Van der Waals interaction and the electrostatic force.

During the cantilever approach-retraction cycle, the hysteresis arises. At the upward move the neck stays longer because the cantilever surface is already wetted and the liquid neck goes with the tip. As bonds break, the capillary force stops to act and the cantilever suddenly returns into its undeflected state.



**Adhesion effects can further affect the interaction (unless UHV operation is performed!)**

# Scanning Force Microscopy

## 3.2 The Operation Principle of Scanning Force Microscope

The main electronic components of the SFM are the same as for the STM, only the topography of the scanned surface is reconstructed by analysing the deflection of the tip at the end of a spring. Today, the interferometrical and optical lever method dominate commercial SFM apparatus. The most common method for detecting the deflection of cantilever is by measuring the position of a reflected laser-beam on a photosensitive detector. The principle of this optical lever method is presented in Figure 18 a. Without

cantilever displacement both quadrants of the photodiode (A and B) have the same irradiation  $P_A = P_B = P/2$  ( $P$  represents the total light intensity). The change of the irradiated area in the quadrants A and B is a linear function of the displacement

$$\delta \propto \Delta d = 2 \sin(\Theta) \cdot S_2 = 2\Theta \cdot S_2 = 3S_2 \cdot \delta / L \quad (10)$$

For small angles  $\sin(\Theta) \approx \Theta$  and  $\Theta$  may be evaluated from the relation  $\Theta = 3\delta / 2L$  (Figure 18b). For  $P_A$  and  $P_B$  one would get approximately  $P_A = P/2 \cdot (d + \Delta d) / 2$  and  $P_B = P/2 \cdot (d - \Delta d) / 2$ . Using the simple difference between  $P_A$  and  $P_B$  would lead to

$\Delta P = P \cdot 3S_2\delta / (Ld)$  but in this case one cannot distinguish between the displacement  $\delta$  of the cantilever and the variation in the laser power  $P$ . Hence the normalised difference is used, which is only dependent of  $\delta$ :

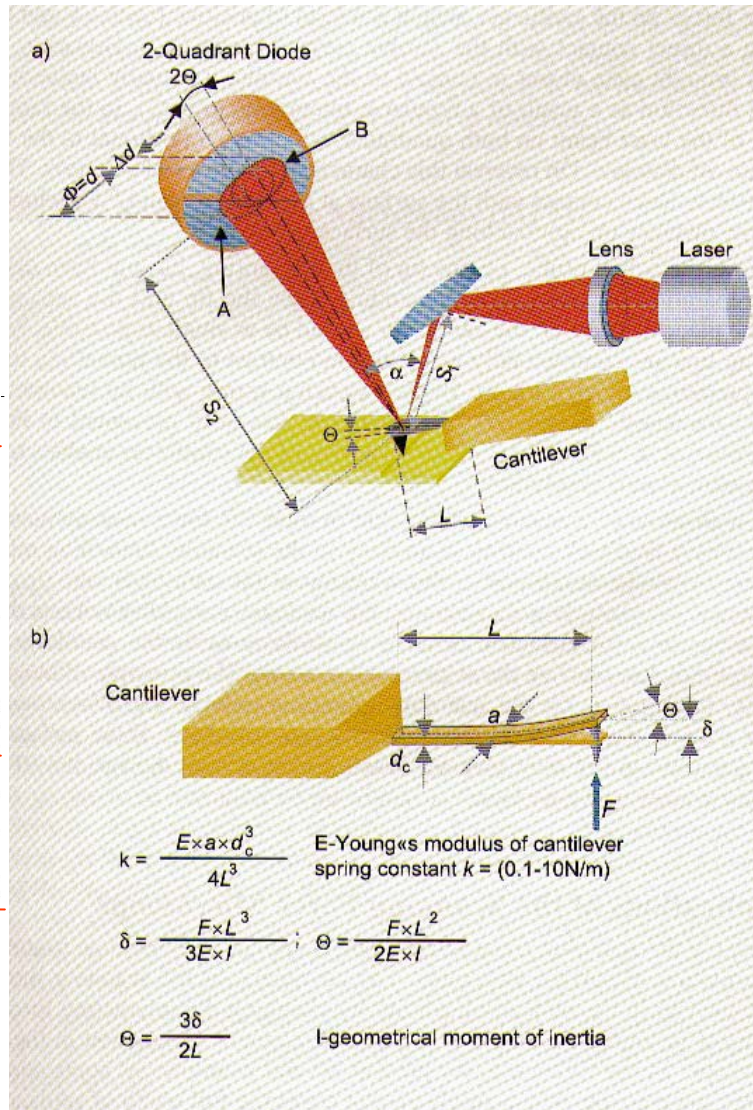
$$\frac{P_A - P_B}{P_A + P_B} = \delta \cdot \frac{3S_2}{Ld} \quad (11)$$

The "lever amplification"  $\Delta d / \delta = 3S_2 / L$  is about a factor of one thousand. On the basis of this kind of technique one is able to detect changes in the position of a cantilever of the order of 0.01 nm.

For large distances between the tip and the sample the bending of the cantilever by attractive forces is negligible. After the cantilever is brought closer to the surface of the sample (point "a" Figure 18c) the van der Waals forces induce a strong deflection of the cantilever and, simultaneously, the cantilever is moving towards the surface. This increases the forces on the cantilever, which is a kind of positive feedback and brings the cantilever to a direct contact with the sample surface (point "b"). However, when the cantilever is brought even closer in contact to the sample, it actually begins to bend in the opposite direction as a result of a repulsive interaction ("b-c"). In the range ("b-c") the position of the laser beam on both quadrants, which is proportional to the force, is a linear function of distance. On reversal this characteristic shows a hysteresis. This means that the cantilever loses contact with the surface at a distance (point "d") which is much larger than the distance on approaching the surface (point "a").

Up to now, the actual probe, i.e. the tip of the leaf spring, has not been discussed in detail. Its preparation is particularly demanding since the tip and the sensitive spring should be one piece. Moreover, the cantilever should be as small as possible. Nowadays, such scanning tips are commercially available (in contrast to the tunnelling tips, which you should prepare yourself). Figure 19 shows such a spring with tip (cantilever) made of Si. The characteristic parameters of a cantilever has been presented in Figure 18b. The spring constant  $k = Ead_c^3 / 4L^3 \sim 0.1 - 10 \text{ N/m}$  of the cantilever enables topographical analysis with atomic resolution.

For the realisation of a scanning force microscope, the force measurement must be supplemented by a feedback control, in analogy to the scanning tunnelling microscope. The controller keeps the amplitude of the vibration of the cantilever (the tip), and thus also the distance, constant. During scanning the feedback controller retracts the sample with the scanner of a piezoelectric ceramic or shifts towards the cantilever until the vibration amplitude has reached the setpoint value again. The principle of height regulation is exactly the same as for the scanning tunnelling microscope. The scanning force micrographs thus show areas of constant effective force constant. If the surface is chemically homogeneous and if only van der Waals forces act on the tip, the SFM image shows the topography of the surface.



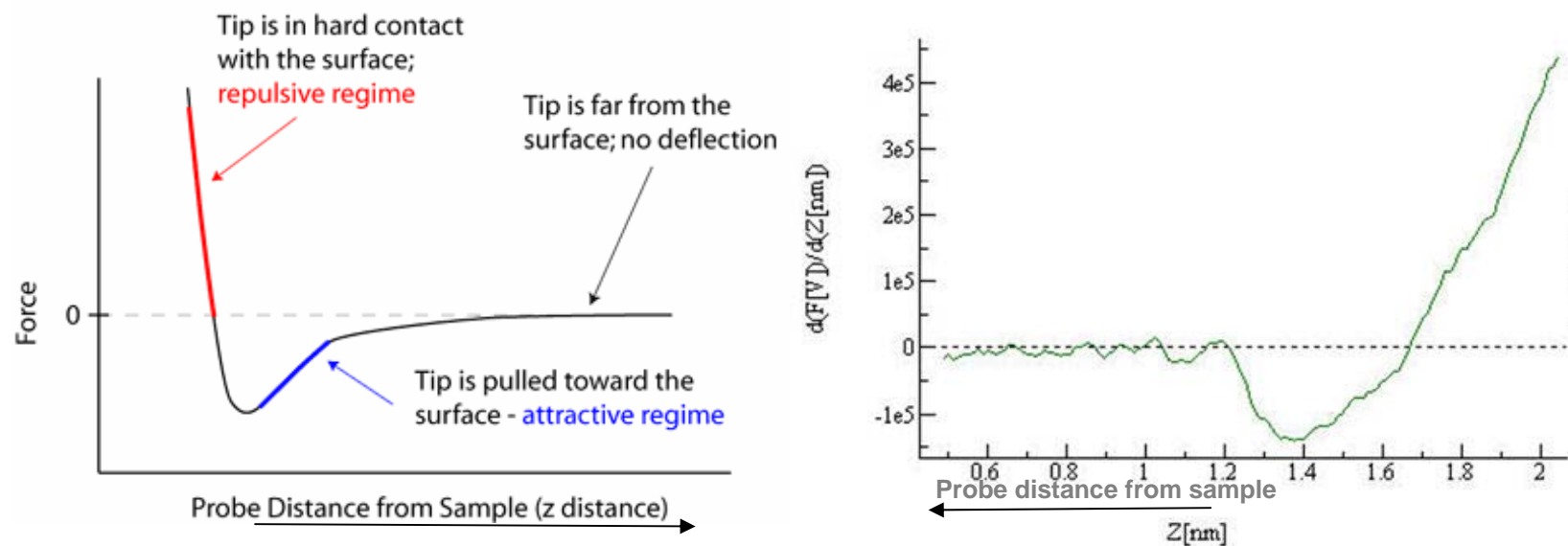
**Figure 18:** The amplification of the cantilever motion through the optical lever arm method. (a) Optical laser path in the standard AFM set-up. (b) Cantilever beam in bending. (c) Cantilever force as a function of the distance tip – sample distance.

**An optical lever method is used to detect the cantilever deflection**

## Force vs distance curves

In the approaching step, force (i.e., cantilever deflection) vs distance plots have a typical behavior

Force vs distance curves can be used to get local information on the mechanical properties of the surface (*force spectroscopy*)



**Note:** we are discussing of the **contact mode operation** and tip might penetrate into the sample (*as in nanoindentation – we will see later!*)

# Contact mode of operation

## 3.1 Theoretical Principles of the Scanning Force Microscope

As already mentioned above, van der Waals forces lead to an attractive interaction between the tip on the spring and the sample surface. Figure 15 shows schematically the van der Waals potential between two atoms. The potential can be described in a simpler classical picture as the interaction potential between the time dependent dipole moments of the two atoms. Although the centres of gravity of the electronic charge density and the charge of nucleus are exactly overlapping on a time average, the separation of the centres of gravity is spatially fluctuating in every moment. This produces statistical fluctuations of the atoms' dipole moments. The dipole moment of an atom can again induce a dipole moment in the neighbouring atom and the induced dipole moment acts back on the first atom. This creates a dipole-dipole interaction on basis of the fluctuating dipole moments. This interaction decreases with  $d^6$  in the case of small distances  $d$  (Lenard-Jones potential). At larger distances, the interaction potential decreases more rapidly ( $d^7$ ). This arises from the fact that the interaction between dipole moments occurs through the exchange of virtual photons. If the transit time of the virtual photon between atoms 1 and 2 is longer than the typical fluctuation time of the instantaneous dipole moment, the virtual photon weakens the interaction. This range of the van der Waals interaction is therefore called retarded, whereas that at short distances is unretarded.

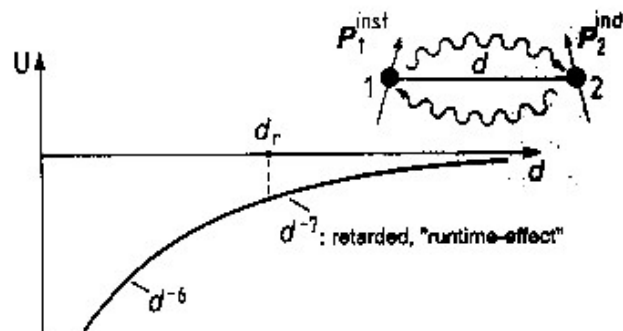


Figure 15: The van der Waals potential  $U$  between two atoms.  $d_r$  is the critical distance above which the transit time effects weaken the interaction [23].

**Contact mode is suitable for rather rigid surfaces**

The scanning force microscope is not based on the interaction of individual atom only. Both the sample and the tip are large in comparison to the distance. In order to obtain their interaction, all forces between the atoms of both bodies need to be integrated. The result of this is known for simple bodies and geometries. In all cases, the summation leads to a weaker decrease of the interaction. A single atom at distance  $d$  relative to a half-space leads to an interaction potential of

$$U = -\frac{C\pi\rho}{6} \cdot \frac{1}{d^3} \quad (7)$$

where  $C$  is the interaction constant of the van der Waals potential and  $\rho$  the density of the solid.  $C$  is basically determined by the electronic polarizabilities of the atoms in the half-space and of the single atom. If one has two spheres with radii  $R_1$  and  $R_2$  at distance  $d$  (distance between sphere surfaces) one obtains an interaction potential of

$$U = -\frac{AR_1R_2}{6(R_1 + R_2)} \cdot \frac{1}{d} \quad (8)$$

where  $A$  is the so-called Hamaker constant. It is materials specific and essentially contains the densities of the two bodies and the interaction constant  $C$  of the van der Waals potential. If a sphere with radius  $R$  has a distance  $d$  from a half-space, an interaction potential of

$$U = -\frac{AR}{6} \cdot \frac{1}{d} \quad \text{Realistic tip/surface potential} \quad (9)$$

is obtained from Eq. (8). This case describes the geometry in a scanning force microscope best and is most widely used. The distance dependence of the van der Waals potential thus obtained is used analogously to the distance dependence of the tunnel current in a scanning tunnelling microscope to achieve a high resolution of the scanning force microscope. However, since the distance dependence is much weaker, the sensitivity of the scanning force microscope is lower.

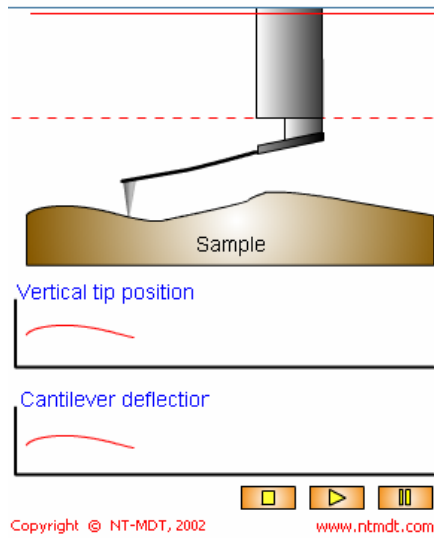
In the contact mode of operation, mechanical interaction leads to tip displacement, i.e., to cantilever deflection related to topography changes

As in STM (constant gap), typical operation foresees a **feedback** system, acting on the Z direction of the piezoscanner, which keeps constant the cantilever deflection during the scan

The “error signal” of the feedback system provides a **topography map (with a calibrated sub-nm space resolution)**

# Operation modes (AFM contact)

The most straightforward AFM operation mode involves “**contact**” (repulsive) forces  
 Constant height and **constant force** configurations are possible (the latter, the most common, is based on feedback)  
 Major drawback: surface degradation, especially with soft matter



In **Contact mode** of operation the cantilever deflection under scanning reflects repulsive force acting upon the tip.

Repulsion force  $F$  acting upon the tip is related to the cantilever deflection value  $x$  under Hooke's law:  $F = -kx$ , where  $k$  is cantilever spring constant. The spring constant value for different cantilevers usually vary from 0.01 to several N/m.

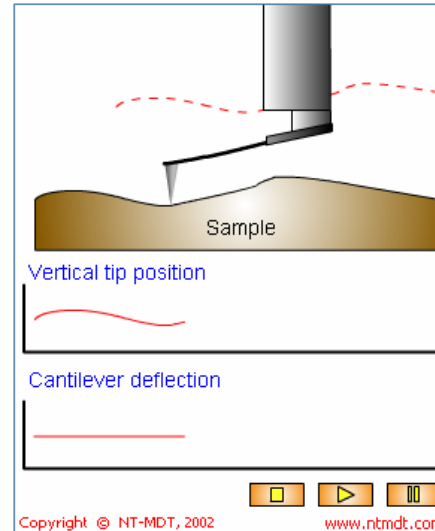
In our units the vertical cantilever deflection value is measured by means of the optical registration system and converted into electrical signal DFL. In contact mode the DFL signal is used as a parameter characterizing the interaction force between the tip and the surface. There is a linear relationship between the DFL value and the force. In Constant Height mode of operation the scanner of the microscope maintains fixed end of cantilever on the constant height value. So deflection of the cantilever under scanning reflects topography of sample under investigation.

Constant Height mode has some advantages and disadvantages.

Main advantage of Constant Height mode is high scanning speeds. It is restricted only by resonant frequency of the

cantilever.

Constant Height mode has also some disadvantages. Samples must be sufficiently smooth. When exploring soft samples (like polymers, biological samples, Langmuir-Blodgett films etc.) they can be destroyed by the scratching because the probe scanning tip is in direct contact with the surface. Thereunto under scanning soft samples with relatively high relief the pressure upon the surface varies, simultaneously varies local flexure of sample surface. As a result acquired topography of the sample can prove distorted. Possible existence of substantial capillary forces imposed by a liquid adsorption layer can decrease the resolution.



In **Contact mode** of operation the cantilever deflection under scanning reflects repulsive force acting upon the tip.

Repulsion force  $F$  acting upon the tip is related to the cantilever deflection value  $x$  under Hooke's law:  $F = -kx$ , where  $k$  is cantilever spring constant. The spring constant value for different cantilevers usually vary from 0.01 to several N/m.

In our units the vertical cantilever deflection value is measured by means of the optical registration system and converted into electrical signal DFL. In contact mode the DFL signal is used as a parameter characterizing the interaction force between the tip and the surface. There is a linear relationship between the DFL value and the force. In Constant Force mode of operation the deflection of the cantilever is maintained by the feedback circuitry on the preset value. So vertical displacement of the scanner under scanning reflects topography of sample under investigation.

Constant Force mode has some advantages and disadvantages.

Main advantage of Constant Force mode is possibility to measure with high resolution simultaneously with topography some other characteristics - [Friction Forces](#), [Spreading Resistance](#) etc.

Constant Force mode has also some disadvantages. Speed of scanning is restricted by the response time of feedback system. When exploring soft samples (like polymers, biological samples, Langmuir-Blodgett films etc.) they can be destroyed by the scratching because the probe scanning tip is in direct contact with the surface. Thereunto under scanning soft unhomogeneous samples the local flexure of sample surface varies. As a result acquired topography of the sample can prove distorted. Possible existence of substantial capillary forces imposed by a liquid adsorption layer can decrease the resolution.

**Locally probing the (repulsive) force allows for topography and morphology reconstruction**

# Non-contact modes of operation

The dynamic operation method of a scanning force microscope has proved to be particularly useful. In this method the nominal force constant of the van der Waals potential, i.e. the second derivative of the potential, is exploited. This can be measured by using a vibrating tip (Figure 16). If a tip vibrates at distance  $d$ , which is outside the interaction range of the van der Waals potential, then the vibration frequency and the amplitude are only determined by the spring constant  $k$  of the spring. This corresponds to a harmonic potential. When the tip comes into the interaction range of the van der Waals potential, the harmonic potential and the interaction potential are superimposed thus changing the vibration frequency and the amplitude of the spring.

This is described by modifying the spring constant  $k$  of the spring by an additional contribution  $f$  of the van der Waals potential. As a consequence, the vibration frequency is shifted to lower frequencies as shown in Figure 17.  $\omega_0$  is the resonance frequency without interaction and  $\Delta\omega$  the frequency shift to lower values. If an excitation frequency of the tip of  $\omega_m > \omega_0$  is selected and kept constant, the amplitude of the vibration decreases as the tip approaches the sample, since the interaction becomes increasingly stronger. Thus, the vibration amplitude also becomes a measure for the distance of the tip from the sample surface. If a spring with low damping  $Q^{-1}$  is selected, the resonance curve is steep and the ratio of the amplitude change for a given frequency shift becomes large.

In practice, small amplitudes (approx. 1 nm) in comparison to distance  $d$  are used to ensure the linearity of the amplitude signal. With a given measurement accuracy of 1 %, however, this means that the assembly must measure deflection changes of 0.01 nm, which is achieved most simply by a laser interferometer or optical lever method.

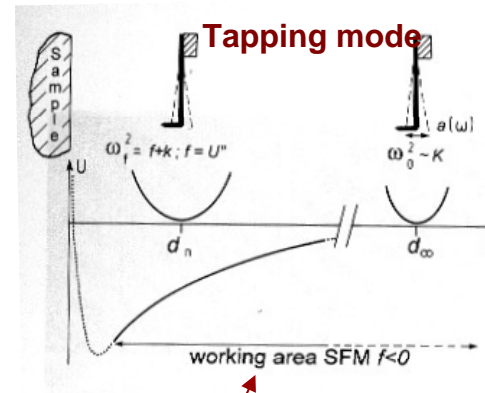


Figure 16: Schematic representation of the effect of the van der Waals interaction potential on the vibration frequency of the spring with tip. As the tip approaches the surface, the resonance frequency of the leaf spring is shifted. (from [23]).

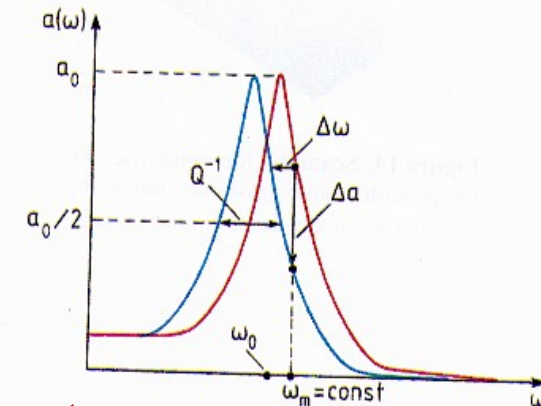


Figure 17: Resonance curves of the tip without and with interaction with a van der Waals potential. The interaction leads to a shift  $\Delta\omega$  of the resonance frequency with the consequence that the tip excited with the frequency  $\omega_m$  has a vibration amplitude  $a(\omega)$  attenuated by  $\Delta a$  [23].

In **non-contact (tapping) mode**, the tip/sample distance is continuously modulated thanks to a vibrating tip

**Tip vibration** is typically achieved by using a piezoelectric transducer fed by an oscillating voltage and mechanically coupled to the cantilever

Oscillation frequency is typically set around the mechanical resonance frequency of the system (cantilever+tip), i.e., hundreds of kHz

The vibration reflects in an oscillation of the position-sensitive detector (multiquadrant diode) and amplitude is monitored

Tip/sample interaction leads to a **damping** (and **phase shift**) of the recorded oscillation when the distance gets small

Suitably conditioned electronic signals are sent into the feedback system in order to stabilize the distance and to derive the topography map

**Non-contact modes suitable for “soft” surfaces**

**No sample preparation is needed!!**

# Effects on the cantilever resonance

Consider a cantilever oscillations when in addition to driving force ((1) in [chapter 2.2.3.3](#)), an external force  $F_{ts}(z)$  acts on it. The equation of motion in this case is written as

$$\ddot{z} + 2\delta\dot{z} + \omega_0^2 z = A_0 \cos \Omega t + F_{ts}(z)/m \quad (1)$$

In a general case the steady-state solution of equation (1) is the sum of harmonics with frequencies divisible by a driving force frequency  $\Omega$ :

$$z(t) = \sum_n A_n \cos(n\Omega t + \varphi_n) \quad (2)$$

In [chapter 2.2.3.4](#) we considered the particular case of equation (1) solution - **small oscillations** when the following condition is met

$$A_0 \ll \frac{m\omega_0^2}{\left\langle \frac{d^2 F_{ts}}{dz^2} \right\rangle} \quad (3)$$

where  $\omega_0$  – cantilever natural resonant frequency,  $\left\langle \frac{d^2 F_{ts}}{dz^2} \right\rangle$  – mean of the second derivative of the tip-sample interaction force (averaged with respect to oscillations amplitude).

In practice, condition (3) is seldom met. However, utilizing numerical methods, one can show that even under weak condition (4), the character of steady-state oscillations will only slightly differ from harmonic (a major contribution is made only by the first harmonic)

condition 
$$A_0 \leq \frac{m\omega_0^2}{\left\langle \frac{d^2 F_{ts}}{dz^2} \right\rangle} \quad (4)$$

solution 
$$z(t) \sim A \cos(\Omega t + \varphi) \quad (5)$$

In contrast to the case of small oscillations where the steady-state condition is entirely determined by system parameters, the motion in the considered case depends on the initial state. That is, depending on the initial position of the cantilever relative to the equilibrium position, the character of the steady-state oscillations will vary (Fig. 1).

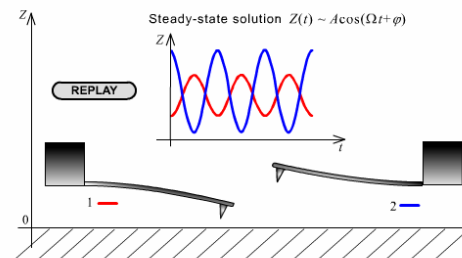


Fig. 1. Feasibility of several quasisteady-state solutions obtaining.

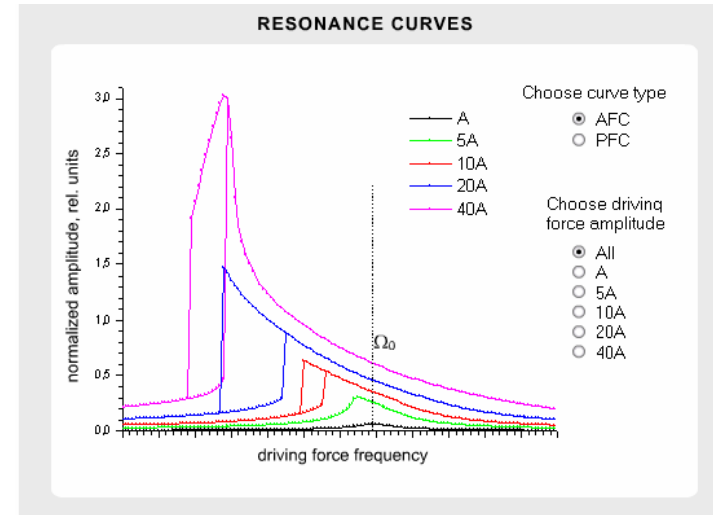
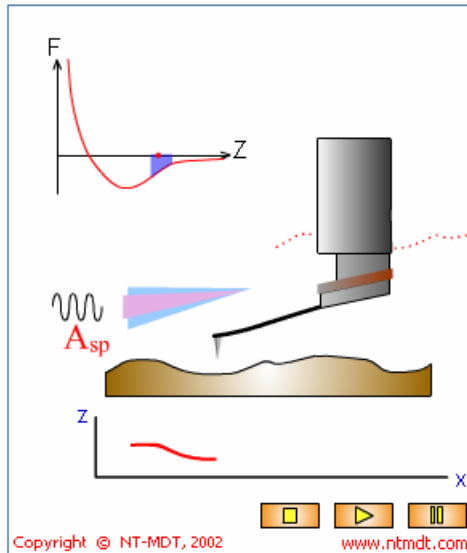


Fig. 2. Resonance curves in case of nonlinear oscillations.

**Cantilever oscillations are strongly dependent on “damping” (both in terms of amplitude and frequency/phase)**



### Non-Contact mode.



The Non-Contact AFM (NC AFM), invented in 1987 [1], offers unique advantages over other contemporary scanning probe techniques such as contact AFM and STM. The absence of repulsive forces (presenting in **Contact AFM**) in NC AFM permits its use in the imaging "soft" samples and, unlike the STM, the NC AFM does not require conducting samples.

The NC AFM works via the principle "amplitude modulation" detection. The corresponding detection scheme exploits the change in the amplitude,  $A$ , of the oscillation of a cantilever due to the interaction of a tip with a sample. To the first order, the working of the NC AFM can be understood in terms of a force-gradient model [1]. According to this model, in the limit of small  $A$ , a cantilever approaching a sample undergoes a shift,  $df$ , in its natural frequency,  $f_0$ , towards a new value given by

$$f_{\text{eff}} = f_0 (1 - F(z)/k_0)^{1/2}$$

where  $f_{\text{eff}}$  is the new, effective resonance frequency of the cantilever of nominal stiffness  $k_0$  in the presence of a force gradient  $F'(z)$  due to the sample. The quantity  $z$  represents an effective tip-sample separation while  $df = f_{\text{eff}} - f_0$  is typically negative, for the case of attractive forces.

If cantilever is initially forced to vibrate at a  $f_{\text{set}} > f_0$ , then the shift in the resonance spectrum of the cantilever towards lower frequencies will cause a decrease in the oscillation amplitude at  $f_{\text{set}}$  as the tip approaches the sample [1].

This change in  $A$  is used as the input to the NC-AFM feedback. To obtain a NC AFM image the user initially chooses a value  $A_{\text{set}}$  as the set-point such that  $A_{\text{set}} < A(f_{\text{set}})$  when the cantilever is far away from the sample. The NC AFM feedback then moves the cantilever closer to the sample until its instantaneous oscillation amplitude,  $A$ , drops to  $A_{\text{set}}$  at the user-defined driving frequency  $f_{\text{set}}$ . At this point the sample can be scanned in the  $x$ - $y$  plane with the feedback keeping  $A = A_{\text{set}} = \text{constant}$  in order to obtain a NC AFM image. The NC AFM feedback brings the cantilever closer (on average) to the sample if  $A_{\text{set}}$  is decreased at any point, and moves the cantilever farther away from the sample (on average) if  $A_{\text{set}}$  is increased. Overall, the implication of the above model is that the NC AFM image may be considered, in the limit of small  $A$ , to be a map of constant interaction-force gradient experienced by the tip due to the sample.

The non-contact mode has the advantage that the tip never makes contact with the sample and therefore cannot disturb or destroy the sample. This is particularly important in biological applications.

### References

1. J. Appl. Phys. 61, 4723 (1987).

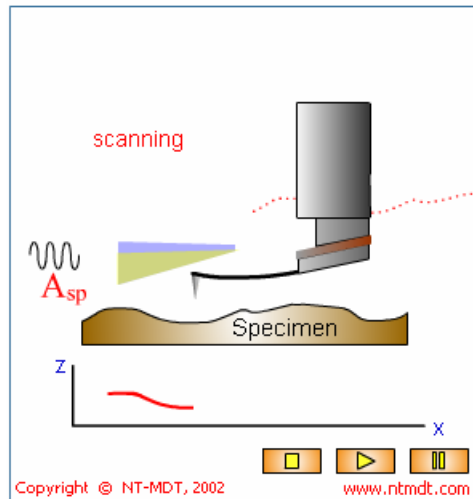
## True non-contact mode

- ✓ Vertical oscillation (tapping) is applied to the tip, at a frequency close to its resonance
- ✓ Resonance frequency variations are measured
- ✓ Feedback keeps the resonance frequency constant by changing the probe/sample distance, hence topography is reconstructed

**Non-contact modes allow for analysis of soft surfaces**

# Examples of other operation modes

Semicontact mode.



Usage of [Scanning Force Microscopy](#) with oscillating cantilever was firstly anticipated by Binnig [1]. Earlier experimental realizations of scanning with oscillated cantilever was realized in works [2, 3]. It was demonstrated influence of the force gradients on the cantilever frequency shift and possibility of non-contact scanning sample surface. It must be noted also that Durig studied frequency shift of oscillating cantilever under influence of STM tip [4]. In [2] was demonstrated also possibility of materials sensing under abrupt decreasing of cantilever oscillation amplitude. Possibility of scanning sample surface not only in attractive but also in repulsive forces was demonstrated in [4]. Relatively small shift of oscillating frequency with sensing repulsive forces means that contact of cantilever tip with sample surface under oscillation is not constant. Only during small part of oscillating period the tip "feels" contact repulsive force. Especially it concerns to oscillations with relatively high

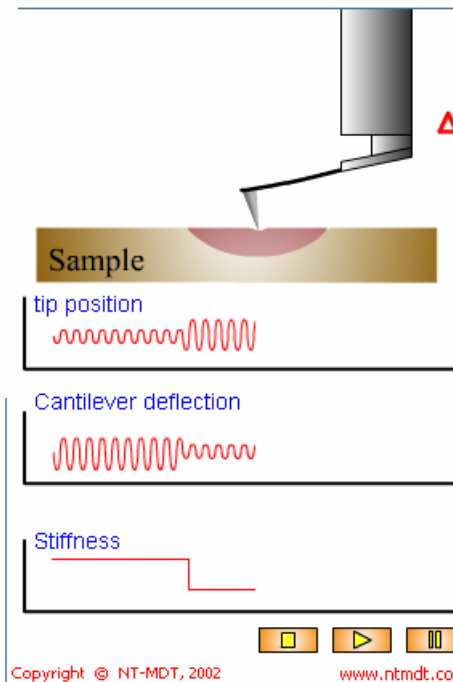
amplitudes. Scanning sample surface with cantilever oscillated in this manner is not non-contact, but intermittent contact. Corresponding mode of Scanning Force Microscope operation (Intermittent Contact mode) is in common practice.

The Intermittent Contact mode can be characterized by some advantages in comparison with [dc Contact mode](#). First of all, in this mode the force of pressure of the cantilever onto the surface is less, that allows to work with softer and easy to damage materials such as polymers and bioorganics. The semicontact mode is also more sensitive to the interaction with the surface that gives a possibility to investigate some characteristics of the surface - distribution of magnetic and electric domains, elasticity and viscosity of the surface.

Force modulation: an additional contrast mechanism related to material properties is found and exploited

Semicontact: the tip gets in "temporarily" contact thus combining advantages of contact and non-contact

Force Modulation mode.



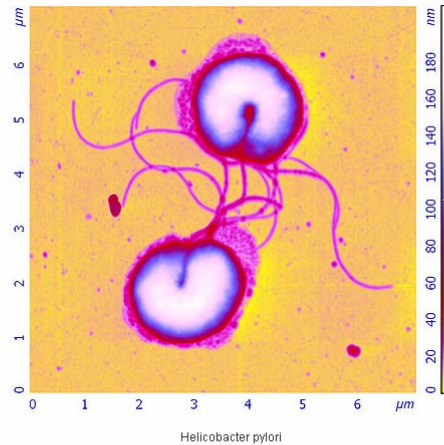
Under realization of Force Modulation mode (FM-mode) along with scanning of sample surface as in [Constant Force mode](#) (CFC-mode) the scanner (or the sample) executes a vertical periodic motion [1]. Under this periodic motion cantilever "feels out" the sample surface. At that the pressure of the probe tip on the sample surface does not remain constant but has periodic component, usually sinusoidal. In accordance with the local elasticity of the sample value of corresponding indentation will change under scanning. On the stiff areas of the sample surface depth of indentation will be smaller, and on the compliant areas - larger. Tracing of the sample surface relief height is conducted by the usage of the averaged cantilever deflection in the feedback circuit [2]. If values of the scanner vertical displacement  $Dz$ , the probe tip vertical displacement  $D$  and cantilever force constant  $\kappa_c$  are known, one can determine the local elasticity of the sample under investigation  $\kappa_s$

$$\kappa_s = \kappa_c \cdot (Dz/D - 1)$$

In turn with known value of the local elasticity one can to determine the modulus of elasticity of the sample. It can be done with usage of the calibrating measurements or with usage of the Hertzian model [3]. Force Modulation mode is widely used in polymers, semiconductors, biological, especially in

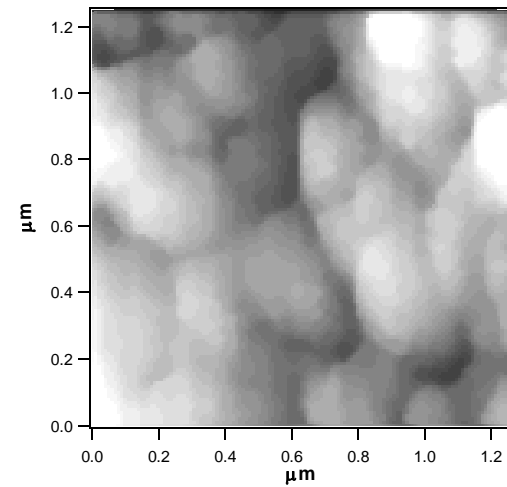
composite materials investigations.

# A very few examples of AFM images I

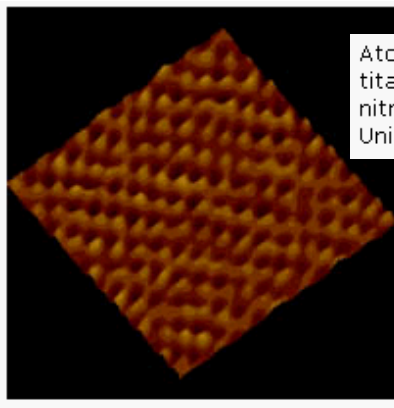


Mode: [Semicontact mode](#)  
 SPM Model: [Solver P47H-PRO](#)  
 Scan size: 7.2x7.2 μm  
 Source MDT-file: [download \(1.01 Mb\)](#)

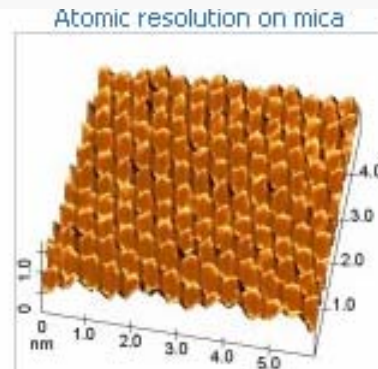
Conversion of two cells of bacterium *Helicobacter pylori* into coccoid forms. Polished silicone covered by polymer.  
 Image courtesy of Budashov I.A., Moscow State University, Institute of Biochemical Physics.  
 Sample courtesy of Momynaliev K.T., Scientific Research Institute of Physical-Chemical Medicine, Moscow.



YBCO/YSZ/Ni

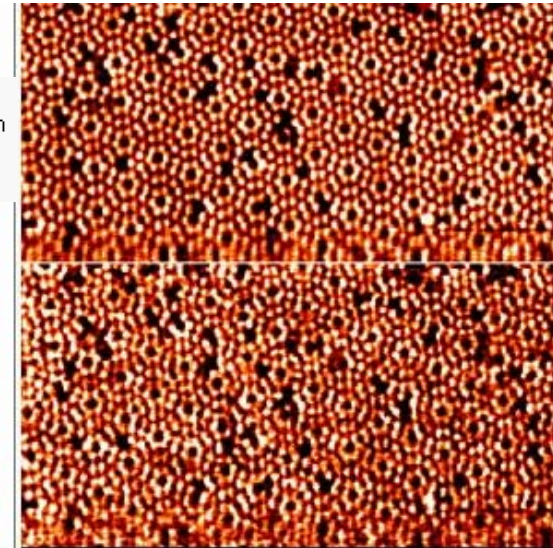


Atomic resolution image of the titanium oxide layer on top of a titanium substrate. Contact mode AFM in air, commercial silicon nitride cantilever. 5 nm scan courtesy P. Cacciafesta, University of Bristol, UK.



Muscovite is fairly common and is found in igneous, metamorphic and detrital sedimentary rocks. It has a layer-like structure of aluminum silicate sheets not strongly bonded, and they are held together by the  $K^+$  ions. For further reading on this topic see the following publication: Gelatin on Mica Surfaces, *J Phys. Chem.* **94**, 4611-4617.

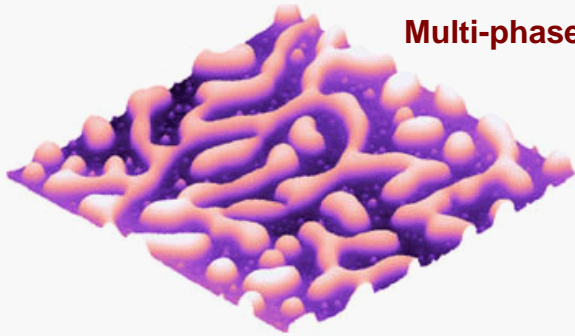
See <http://www.veeco.com>



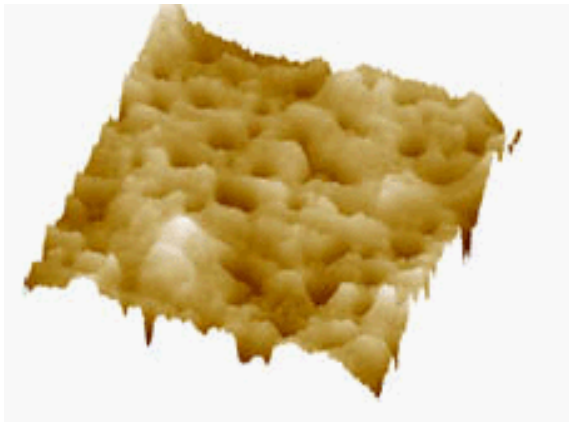
Multi-mode AFM operation : with simultaneous measurement of the topography in STM mode (upper image) with atomic resolution on Si(111)7x7 using a conductive cantilever, and of the atomic scale variation of the force, i.e. cantilever deflection (lower image).

## A very few examples of AFM images II

### Multi-phase systems



TappingMode AFM image of poly(styrene) and poly(methyl methacrylate) blend polymer film. The film was spin-cast on mica substrate from chloroform solution. The surface structure is resulted from the spinodal decomposition. The islands consist of a PMMS-rich phase while the surface matrix composes of a PS-rich phase. 3 $\mu$ m scan courtesy C. Ton-That, Robert Gordon University, U.K.



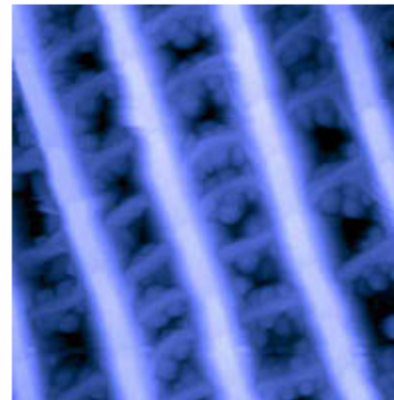
The sample is a strip of adhesive (3M Scotch tape) that has been peeled of a metal surface. The image shows small pits in the sticky surfaces of the adhesive. The image was acquired in TappingMode at frequency of 3 Hz and setpoint of 1.8 V. 2 $\mu$ m scan courtesy L. Scudiero, Washington State University, USA.

### DNA Molecules



A high resolution of the SPM is able to discern very subtle features such as these two linear dsDNA molecules overlapping each other. 155nm scan. Image courtesy of W. Blaine Stine at email address [stineb@stineb.pprd.abbott.com](mailto:stineb@stineb.pprd.abbott.com).

### A Butterfly Wing Imaged in TappingMode AFM

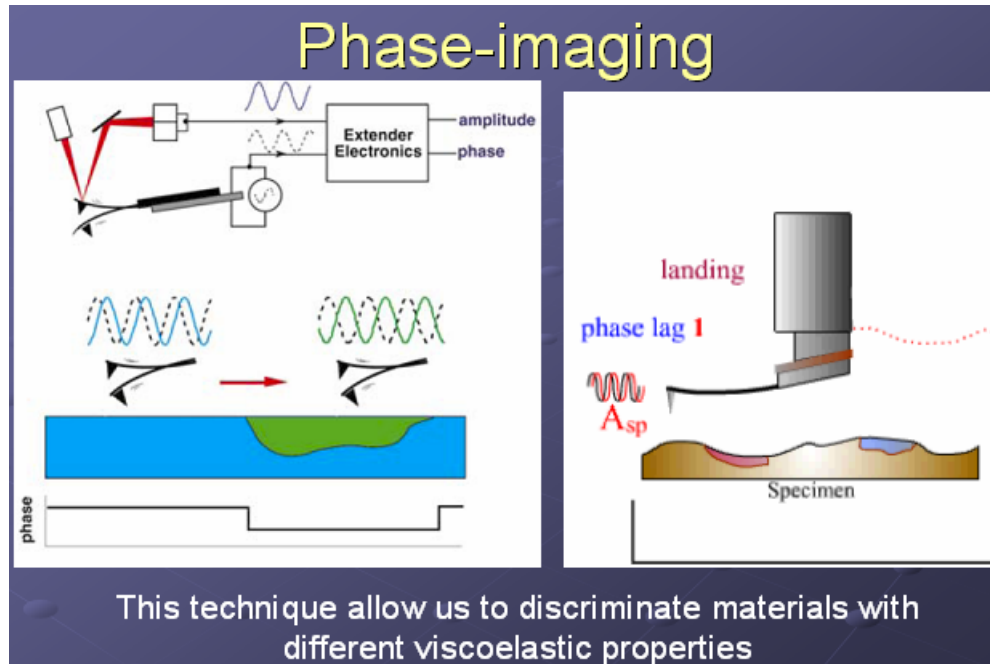


A butterfly (*anthocaris cardamines*) wing imaged in TappingMode AFM. The well-ordered 3-D lattice is an array of holes on the cuticle of the scales. The order of such a system leads to strong diffraction of favored wavelengths in certain directions. This scattering process and pigmentation causes the coloration of the butterfly. 6.25 $\mu$ m scan courtesy of J. Wuest, Museum d'Histoire Naturelle, Switzerland.

**Huge variety of AFM applications!!**

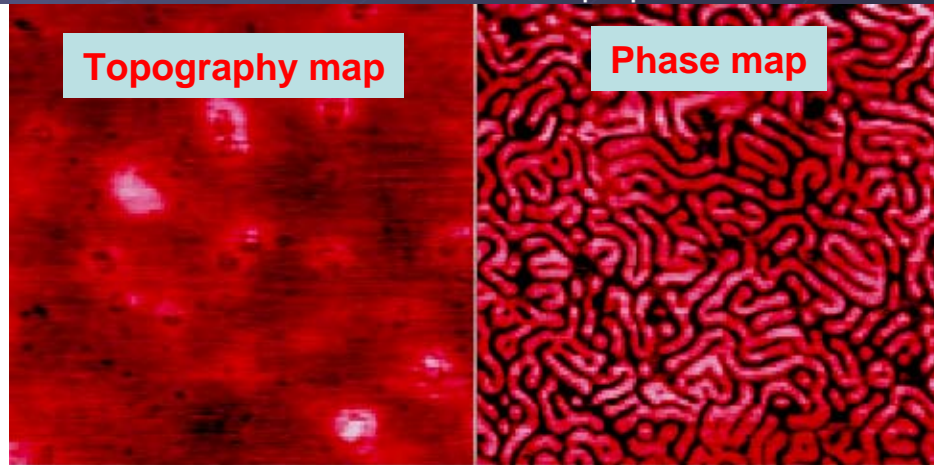
# Phase imaging techniques

Animations at [www.ntmdt.com](http://www.ntmdt.com)!



Dephasing between mechanical oscillation (e.g., the tapping oscillation) and the response of the surface (affecting the tip deflection) depends on the viscoelasticity of the surface (purely elastic vs Newton fluid)

Interpretation similar to a forced and damped mechanical oscillator



Phase imaging:  
- adds a contrast mechanism;  
- allows for local material analyses

Materiale tratto da seminario PhD di Michele Alderighi, 2005

## 2.B. Force microscopies derived from AFM

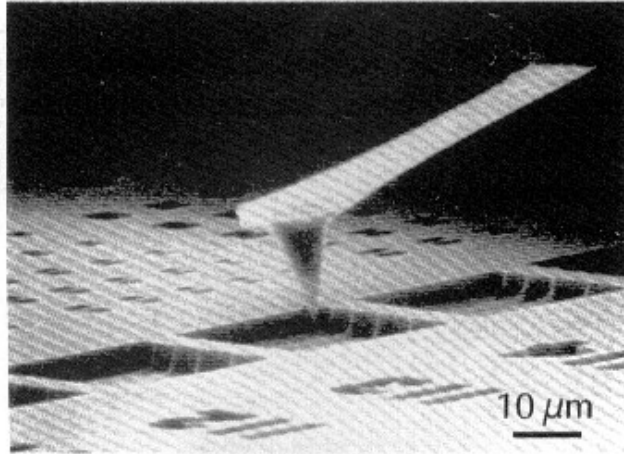


Figure 19: Scanning electron micrograph of a cantilever made of Si. [24].

We have seen how AFM, based on the occurrence of tip/surface van der Waals forces, can map the local topography of the sample

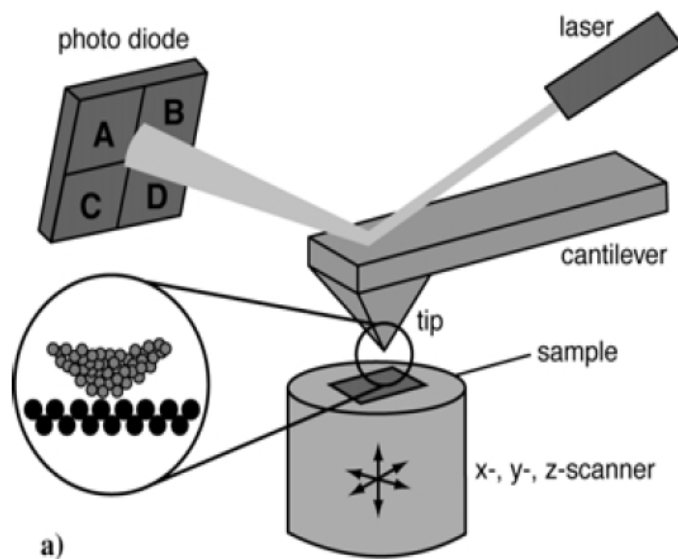
No sample preparation is needed, and the topography map is obtained with “absolute” calibration

The achievable space resolution can reach the **atomic level**, even though most common instruments are capable of a slightly smaller resolution (in the nm range, depending also on the sample properties!)

The close vicinity between tip and surface realized in AFM opens the way for probing physical **quantities other** than the van der Waals interaction force

For instance, tribological and material quantities can be measured (e.g., friction, viscoelasticity, Young modulus, etc.)  
With suitable tips (conductive, magnetic), static and quasi-static electromagnetic forces locally occurring at the sample surface can be analyzed

## Lateral Force Microscopy (LFM, SFFM)

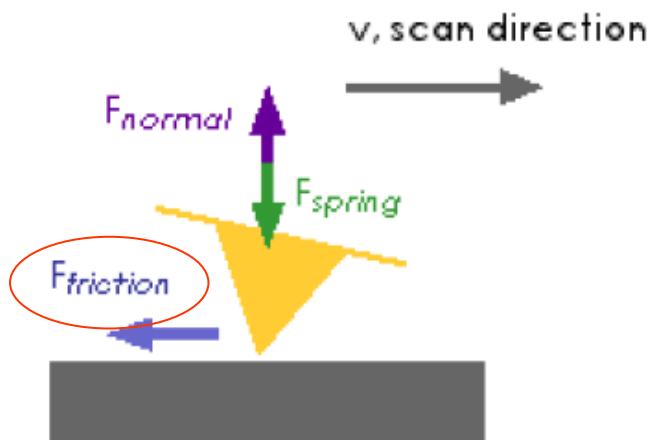


During the scan, the tip is continuously displaced with respect to the surface

**Friction forces** occur, resulting in a *twisting* of the cantilever

Cantilever twist can be recorded by a two-dimension position sensitive detector (i.e., a 4-quadrant photodetector)

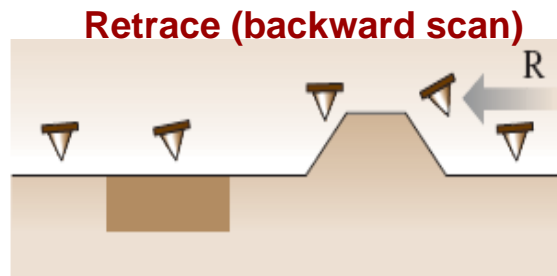
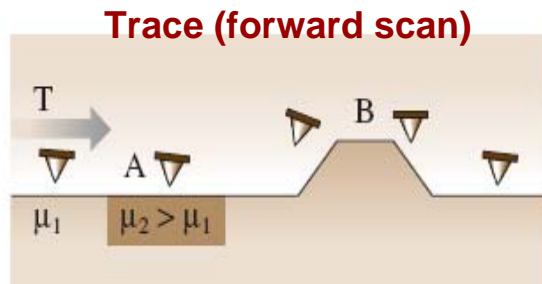
Friction effects can be corrected by the topographical artifacts by comparing forward and backward scans



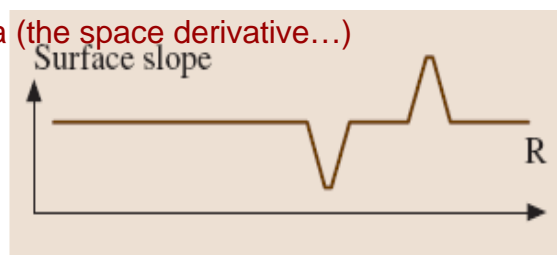
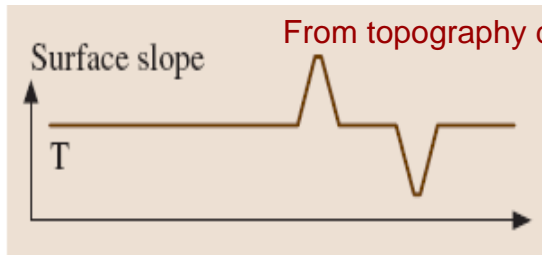
- $(A+B)-(C+D)$  = normal force (AFM signal)
- $(A+C)-(B+D)$  = lateral force (LFM signal)

Materiale tratto dal seminario di Cinzia Rotella, 2006

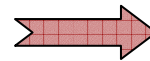
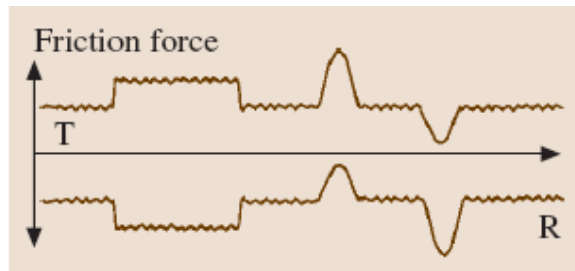
# Friction and topography: artifacts and genuine



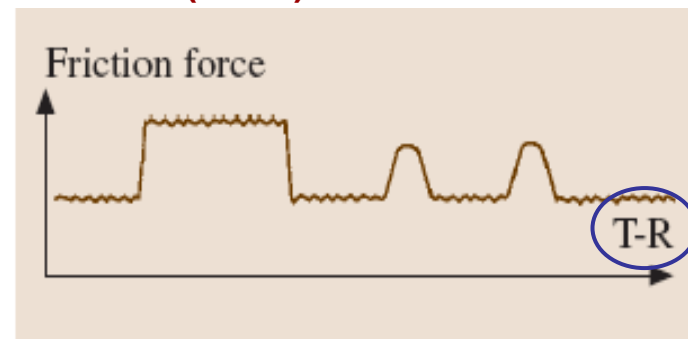
Friction force is **always** oriented against motion!



**Lateral force data**



**(Local) friction data**



Lateral force data are always convoluted with topography (slope), but genuine information of the local friction can be derived by comparing trace and retrace and considering simultaneously acquired topography data



## Examples of LFM/SFFM images

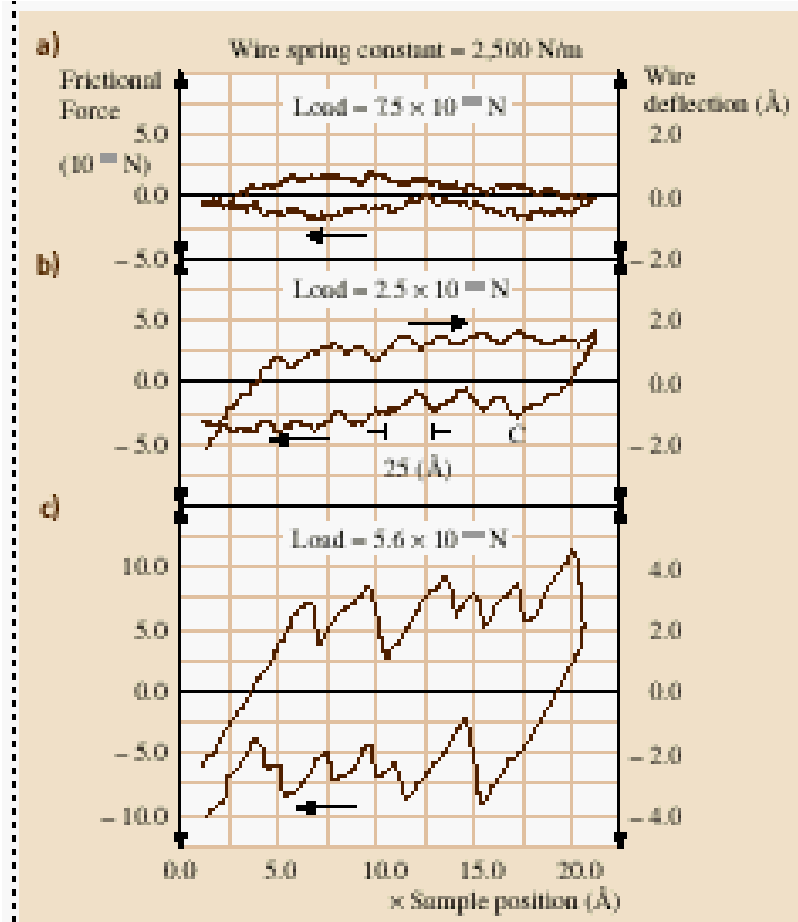


Fig. 20.15a-c Friction loops on graphite acquired with (a)  $F_N = 7.5 \mu\text{N}$ , (b)  $24 \mu\text{N}$  and (c)  $75 \mu\text{N}$ . (After [20.1])

- ✓ LFM/SFFM offers an additional contrast mechanism
- ✓ Possibility to discriminate different materials at the atom level
- ✓ Nanotribology investigations can be carried out

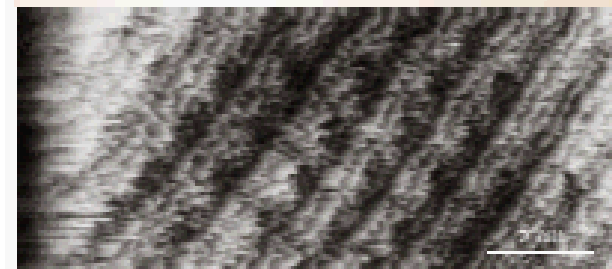


Fig. 20.19 (a) Topography and (b) friction image of Si(111)7x7 measured with a PTFE coated Si-tip. (After [20.29])

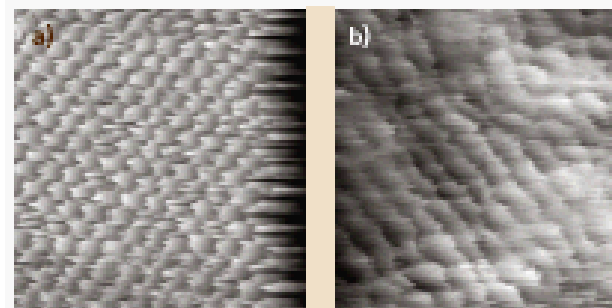
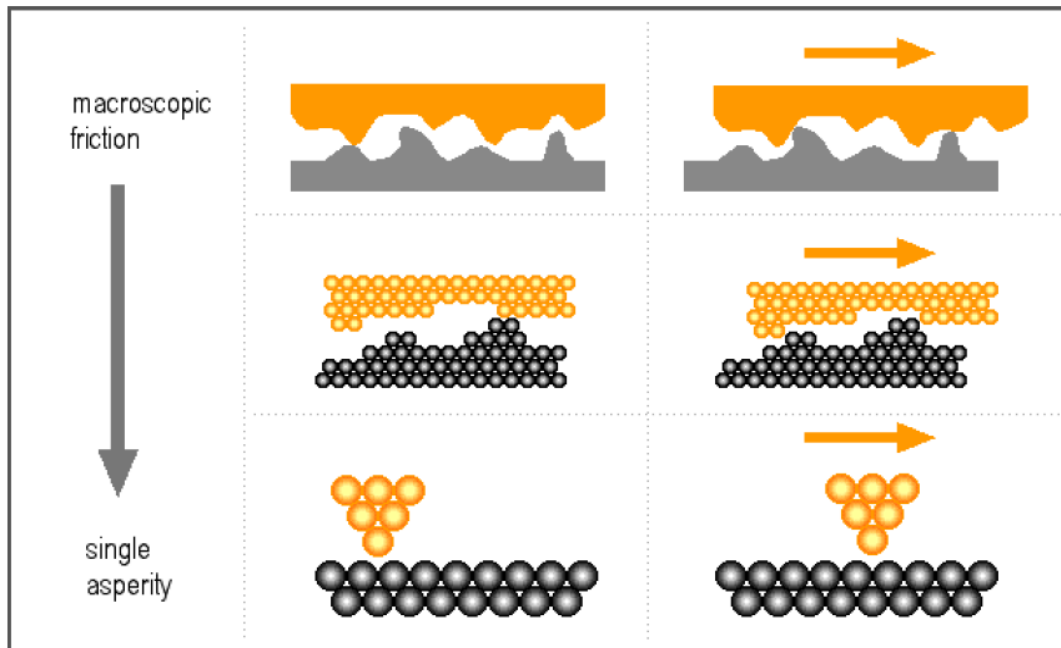


Fig. 20.19a,b Friction images of (a) Ca(111) and (b) Ca(100). Frame size: 3 nm. (After [20.34])

Da B. Bhushan, Handbook of nanotechnology (Springer, 2003)

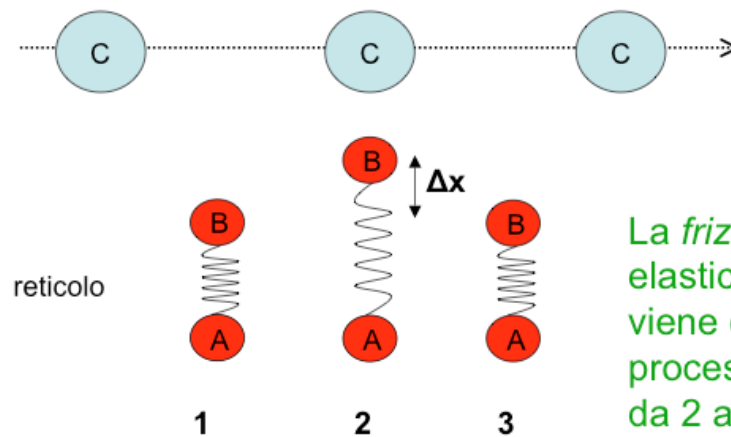
# A few words on nanotribology



Sliding motion of the AFM tip “in contact” with the surface turns out affected by “tribological mechanisms” at the atomic scale:

- **Adhesion;**
- **Ploughing;**
- **Deformation**

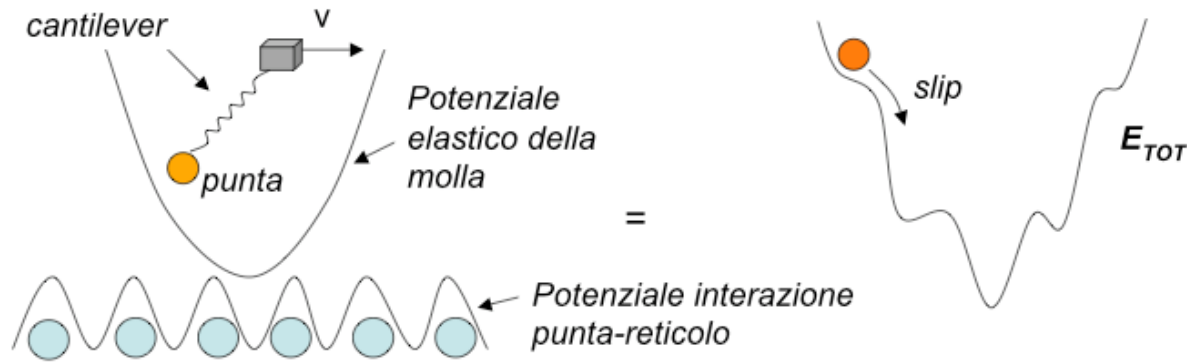
- Modello di Tomlinson (1929) :  
modello di attrito applicabile su scala atomica



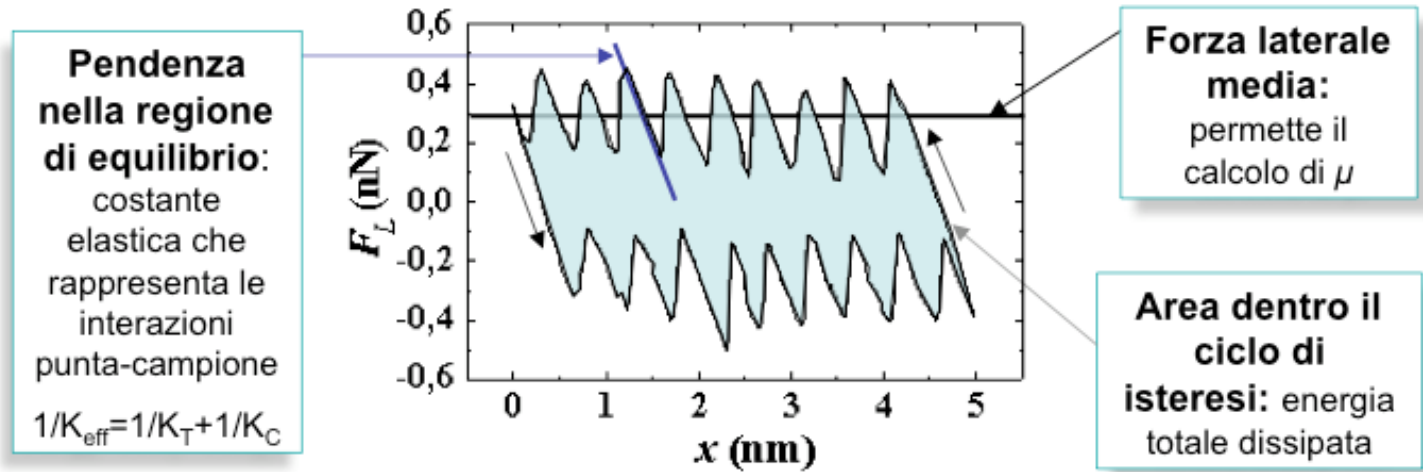
La *frizione* è l'energia elastica  $\frac{1}{2} K(\Delta x)^2$  che viene dissipata nel processo di rilassamento da 2 a 3

Friction models at the atomic scale must account for local tip /surface interaction

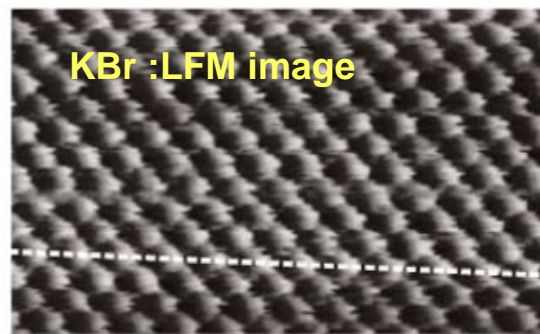
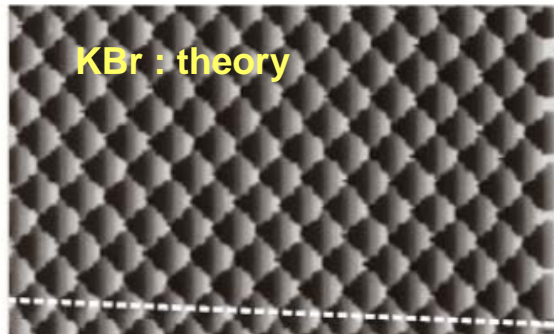
# Stick-slip mechanism during the scan



Interaction potential accounts for the periodic surface potential



Singola linea di scansione LFM di NaCl a  $v = 2.5$  nm/s



Detailed and quantitative info can be achieved at the atomic level

# Magnetic Force Microscopy (MFM)

## 3.3.2 Magnetic Scanning Force Microscopy (MFM)

If a magnetic tip is used in the scanning force microscope, magnetic structures can be imaged. Magnetic scanning force microscopy is of interest, in particular, for the investigation of magnetic storage media. In the most general case, the magnetic force between sample and tip is

$$F_{\text{mag}} = -\nabla \int_{\text{tip}} \mathbf{M}_{\text{tip}} \cdot \mathbf{H}_{\text{sample}} dV \quad (13)$$

or

$$F_{\text{mag}} = (\mathbf{m}_{\text{tip}} \nabla) \mathbf{B}_{\text{sample}} \quad (14)$$

where  $\mathbf{H}_{\text{sample}}$  and  $\mathbf{B}_{\text{sample}}$  are the magnetic stray field and the magnetic induction of the sample, respectively.  $\mathbf{M}_{\text{tip}}$  and  $\mathbf{m}_{\text{tip}}$  are the magnetisation and the magnetic moment of the tip, respectively. Since in most cases the exact magnetic structure of the tip is not known, a model tip magnetization must be assumed. In the simplest case, the tip is a spherically structured magnetic single domain with the magnetisation  $\mathbf{M}_{\text{tip}}$ . Of particular interest are the stray fields of magnetic storage media which consist of different domains. Since the important aspect in force microscopy is not the forces but the force gradient, a pronounced variation of the signal is found near the domain walls, but not inside a domain. This situation is sketched in Figure 21. The parameter of the two curves shown (solid and broken lines) is the ratio of the working distance  $d$  and the radius  $R$  of the magnetic domain of the tip.

Figure 22a shows an experimentally measured picture of four different oriented magnetic domains. Images b and c show the fine structure of a 180° domain. Alternating bright and dark contrasts can be seen. These contrast changes show that the domain wall consists of segments with different wall orientation. This example illustrates that magnetic SFM is well suited for imaging magnetic structures that are commonly used in today's storage media.

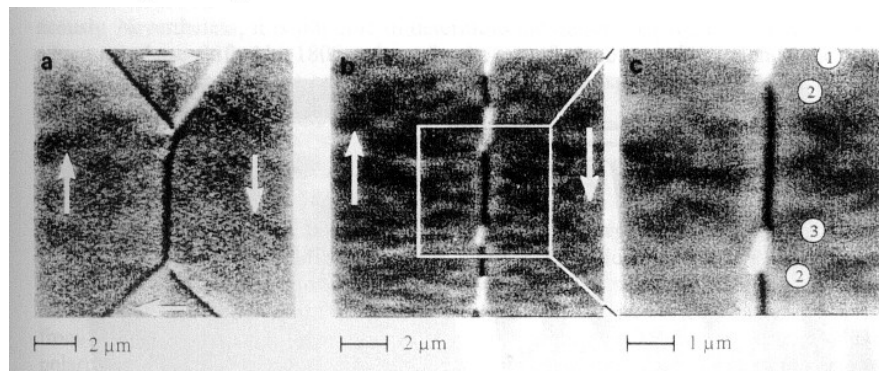


Figure 22: Magnetic SFM image of magnetic domains. (a) shows four domains of a Landau-Lifshitz structure in which the domain walls are the dark and bright lines. (b) and (c) show the fine structure of a 180° domain wall. The domain wall consists of segments with different wall orientation. Arrows denote the domain orientation. (after [26]).

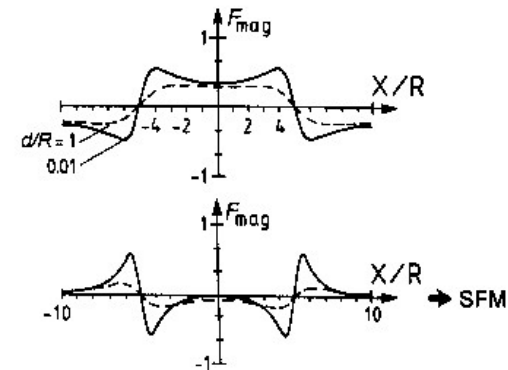
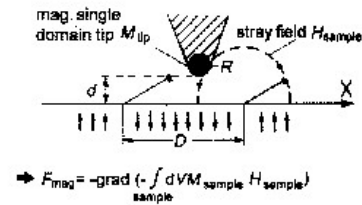


Figure 21: Principle of magnetic scanning force microscopy. On the left, the tip-sample configuration is shown and on the right the force and nominal force constant as a function of distance for this configuration. Two domain walls exist at position 5 (after [23]).

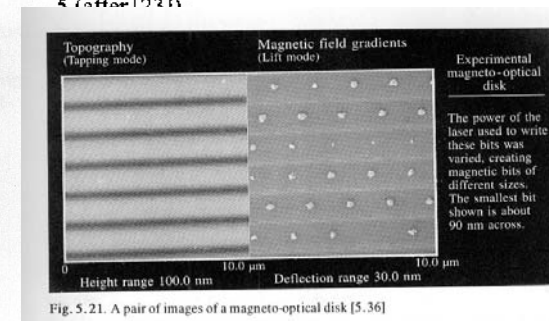


Fig. 5.21. A pair of images of a magneto-optical disk [5.36]

# Electrostatic Force Microscopy (EFM)

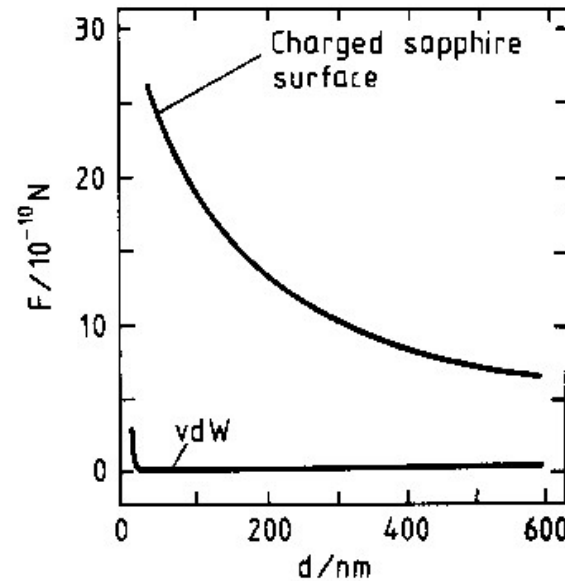
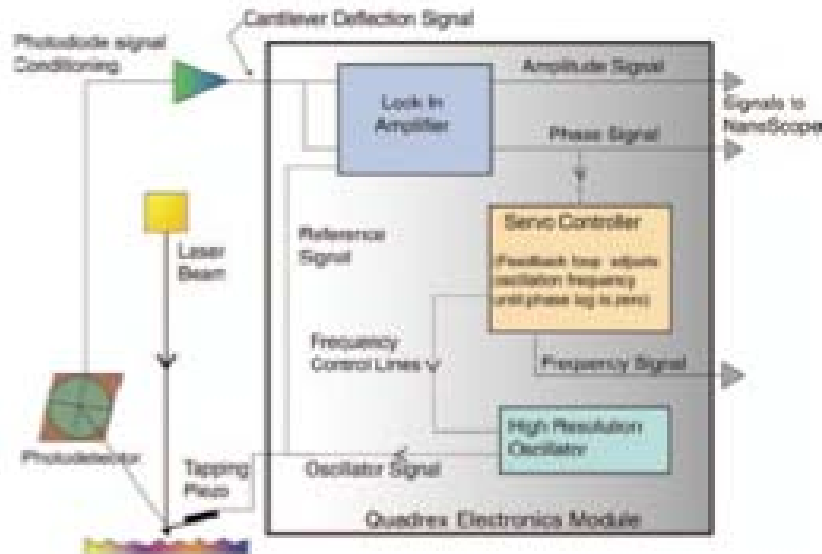


Figure 26: Comparison of the distance dependence of the electrical and van der Waals forces between a tip and a local electrical surface charge of  $2 \cdot 10^{-16} \text{ C}$  (after [23]).

Basic idea: application of a ddp between tip at sample in order to be sensitive to electric forces, thus to the space distribution of charges on the sample surface and to its electrostatic potential

Modulation/demodulation techniques used to get direct information on various surface/tip interaction features

**Excellent sensitivity to local charges**

Application to electronic devices (also in operating conditions)

Application to ferroelectric materials (also known as Piezoelectric Force Microscopy - PFM)

# Operating modes in EFM

Un EFM viene utilizzato principalmente per misurare il potenziale elettrostatico locale della superficie. A tale scopo alla punta viene applicata esternamente una tensione con una componente continua  $V_{dc}$  e una componente  $V_{ac}$  modulata alla frequenza  $\Omega$ . Il campione è invece collegato a massa. La forza elettrostatica tra punta e campione è data da [1]:

$$F = \frac{1}{2} \frac{dC}{dz} V^2$$

$C$  rappresenta l'accoppiamento capacitivo tra punta e campione mentre  $z$  è la variabile che indica la quota della punta. In  $V$  compaiono la tensione applicata ( $V_{dc} + V_{ac} \sin \Omega t$ ) una eventuale tensione indotta sulla punta da altri effetti secondari ( $V_{ind}$ ) e la tensione locale del campione ( $V_{cp}$ ).

$$V = (V_{cp} + V_{dc} + V_{ind}) + V_{ac} \sin(\Omega t)$$

La forza elettrostatica tra il campione e la punta può essere scomposta in tre componenti a frequenze diverse:

$$F = \frac{1}{2} \frac{dC}{dz} V^2 = \frac{1}{2} \frac{dC}{dz} [(V_{cp} + V_{dc} + V_{ind}) + V_{ac} \sin(\Omega t)]^2 = F_{dc} + F_{\Omega} + F_{2\Omega}$$

Esaminiamo ora in dettaglio le tre componenti.

1) Il primo termine della forza è costante nel tempo ed è pari a:

$$F_{dc} = \frac{1}{2} \frac{dC}{dz} [(V_{cp} + V_{dc} + V_{ind})^2 + \frac{1}{2} V_{ac}^2]$$

2) Il secondo termine invece dipende direttamente dalle quantità di interesse come  $dC/dz$  o  $(V_{cp} + V_{ind})$  ed è modulato alla frequenza  $\Omega$ :

$$F_{\Omega} = \frac{dC}{dz} (V_{cp} + V_{dc} + V_{ind}) V_{ac} \sin(\Omega t)$$

3) Il terzo termine è modulato alla frequenza  $2\Omega$  ed è molto importante perché dipende esclusivamente da  $V_{ac}$ , che è noto, e da  $dC/dz$ :

$$F_{2\Omega} = -\frac{1}{4} \frac{dC}{dz} V_{ac}^2 \cos(2\Omega t)$$

Materiale tratto dal seminario di Nicola Paradiso, 2006

## Operazione in modalità nano-Kelvin

Come si è detto precedentemente operando con un EFM occorre fornire una tensione con una componente continua e una modulata. Se con un circuito di reazione noi iniettiamo una tensione  $V_{dc}$  tale che il termine  $F_{dc} = 0$ , per cui  $V_{dc} = -(V_{cp} + V_{ind})$ , allora potremo misurare variazioni di tensione sulla superficie dovute a sia a  $V_{cp}$  che a  $V_{ind}$ .

## Impiego di un EFM per misure in dc

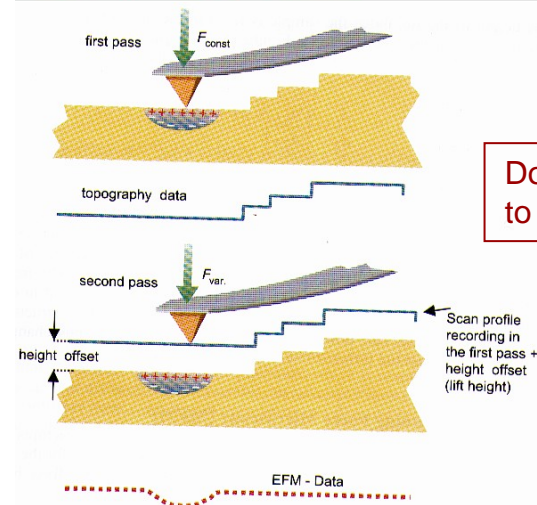
L'utilizzo dell'EFM è basato sull'assunto che la forza elettrostatica sia un effetto al secondo ordine sull'oscillazione meccanica [1]. Tali effetti vengono separati per mezzo di due amplificatori lock-in che permettono di estrarre il segnale modulato alla frequenza  $\Omega$  e quello alla frequenza  $2\Omega$ . E' conveniente scegliere la frequenza  $\Omega$  in modo da lavorare lontani dalla risonanza di oscillazione meccanica. Come si è detto, conviene inserire anche un circuito di reazione per poter misurare e mappare il potenziale elettrico superficiale. Queste operazioni richiedono che la distanza punta-campione sia stabile durante la misura. Esaminiamo brevemente due metodi che permettono di lavorare in tali condizioni.

## Modalità single-pass

Operando in aria, l'ampiezza dell'oscillazione meccanica della cantilever vicina alla risonanza è in genere debolmente dipendente da eventuali interazioni elettrostatiche. Se però localmente i valori di  $V_{cp}$  e di  $V_{ind}$  diventano molto intensi l'immagine topografica che si ricava può risultare disturbata. In modalità single-pass si utilizzano tensioni dc basse e si acquisiscono simultaneamente i dati sulla topografia (ottenuti facendo oscillare la sonda in tapping mode), i dati del segnale alla frequenza  $\Omega$  e alla frequenza  $2\Omega$ . Operando in modalità nano-Kelvin in pratica si sottrae la tensione  $V_{cp}$  che disturba l'immagine, mentre con il segnale alla frequenza  $2\Omega$  si ricava il valore dell'accoppiamento capacitivo locale.

## Modalità double-pass

Questa modalità prevede un primo passaggio volto ad acquisire il profilo topografico locale. La cantilever viene posta in oscillazione (tapping mode) senza che venga applicata alcuna tensione esterna. Alla fine della scansione della linea, il piezo si ritrae, allontana la sonda dal campione di una certa quantità fissata e riprende la scansione nel senso opposto, mantenendo fissa la quota basandosi sui dati topografici appena acquisiti. In questa fase la punta è relativamente più sensibile alle forze elettrostatiche in quanto forze a lungo range. Nello specifico qui viene rilevata la variazione di fase dovuta al gradiente di forza elettrostatica. Rispetto alla modalità single-pass questa modalità ha il vantaggio di essere più sensibile alle proprietà elettrostatiche del sistema. Lo svantaggio è rappresentato dalla perdita di risoluzione dovuta all'aver aumentato la quota di lavoro.



# Examples of EFM images

## Electronic devices

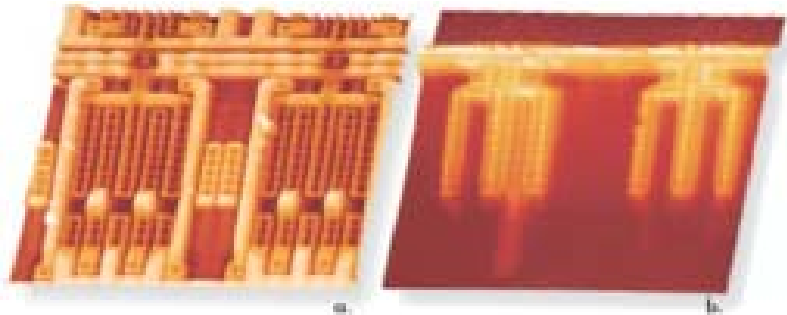


Figure 6. Topography (a) and EFM image (b) of a live packaged IC with passivation layer on. EFM image detects transistor in saturation. 80µm scans.

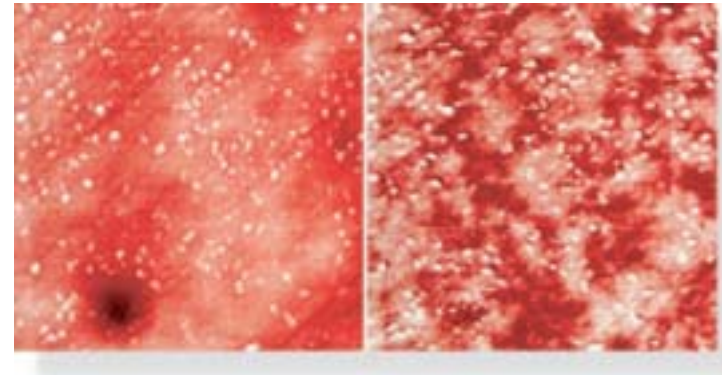
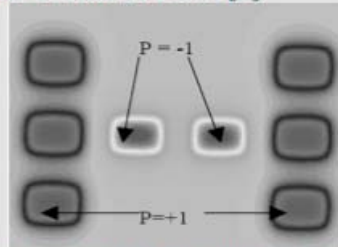


Figure 11. Topography (left) and phase detection EFM (right) images of the surface of a thick-film resistor (TFR). EFM image depicts the conductive RuO<sub>2</sub> network (dark) exposed at the surface. 56µm scans.

## Ferroelectric materials

Riportiamo a titolo di esempio l'immagine ricavata dalla scansione di un campione sulla cui superficie sono stati distribuiti dei domini ferroelectrici. L'area esaminata misura 2µm x 3µm, mentre i singoli domini misurano 500nm x 300nm [3].

*Domini ferroelectrici depositati su un substrato isolante. E' indicata la polarizzazione normalizzata del campo in corrispondenza dei domini. La distanza tip-sample è stata di 60nm. L'area di scansione misura 2µm x 3µm.*



Per materiali che non hanno una polarizzazione permanente vale l'analisi svolta precedentemente. La polarizzazione va calcolata auto-consistentemente; la risoluzione in ogni caso decresce con l'aumentare della separazione tip-sample, come messo in evidenza dalla seguente figura:

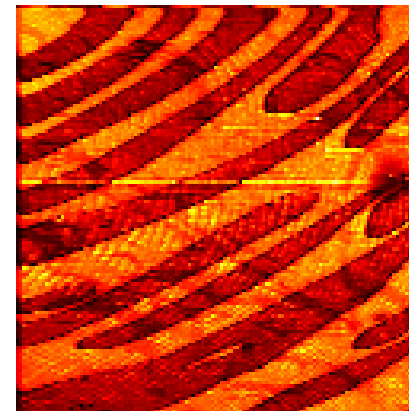


Image of ferroelectric domain structure in TGS Sample by Voltage Modulation Atomic Force Microscopy  
Scan size 27 µm x 27 µm

## Conclusions

- ✓ Scanning Probe techniques have been developed thanks to advances in material fabrication (atomic probes), electronics (and piezoelectric translators), and *methods of operation* (e.g., the role of feedback)
- ✓ Space resolution is excellent in SPMs (at the “atomic level” for some of them)
- ✓ Most important: physical quantities can be *measured* (in quantitative terms) with SPM, as topographical height, surface density of states, optical properties, ...
- ✓ STM is a powerful method to access local electronic properties of a surface (and, through a suitable feedback system, topography at the atomic level)
- ✓ AFM extends STM capabilities to *any kind* of material surfaces (not restricted to conductive as in STM)
- ✓ AFM “relatives” open the way to a wide variety of local measurements and imaging methods