LS Scienza dei Materiali - a.a. 2006/07

Fisica delle Nanotecnologie – part 5.2

Version 5a, Nov 2006

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: SNOM e litografie

20/11/2006 – 14.30-16.30 – room T1

Outlook

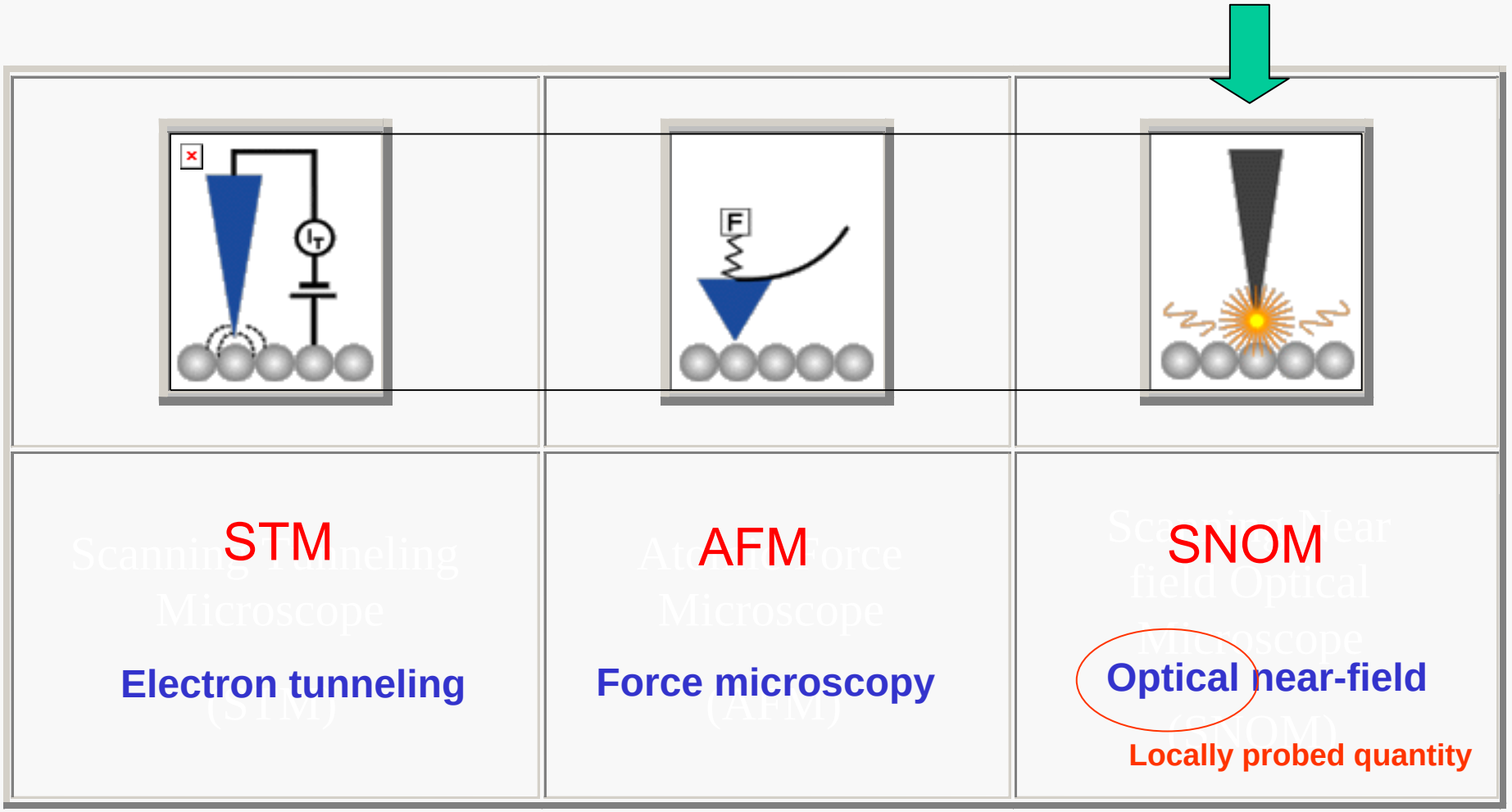
1. The mother of all SPMs: Scanning Tunneling Microscopy (STM): mechanisms and instruments to investigate local electronic properties

2. SPM based on probing mechanical forces:
   A. Atomic Force Microscopy (AFM);
   B. Variants (lateral, electrostatic, magnetic forces,...)

3. Sub-diffraction properties of electromagnetic waves:
   A. Scanning Near Field Optical Microscopy (SNOM);
   B. SNOM-based “nanoscopies”

4. “Lithographies” (better denoted as “nanomanipulations”) associated with:
   A. AFM;
   B. STM;
   C. SNOM

3. Scanning Near Field Optical Microscopy (SNOM)

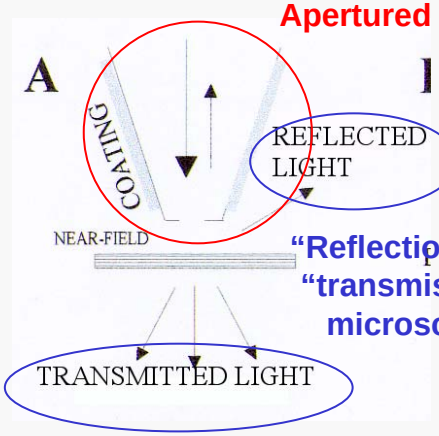


*SNOM holds the **unique** ability to analyze optical properties with sub-diffraction space resolution*

Introducing near-field and SNOM

Main motivation: extending optical microscopy (and sepctroscopy) analyses into the nanoworld (which implies to overcome the diffraction limit!!)

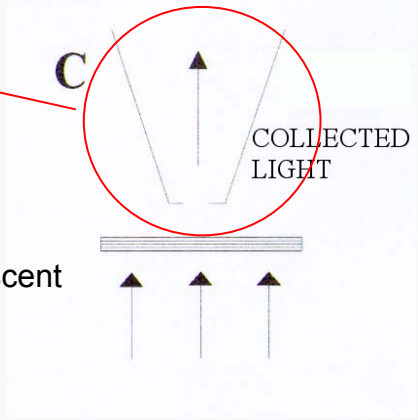
Optical near-field: an e.m. field with frequency in the optical range and a **non-propagating** nature
Near-field is the probed quantity in SNOM, and can be exploited in many different configurations



Apertured probe

Apertured tapered optical fibers;
Hollow cantilevers;
Etc.

"Reflection" and "transmission" microscopy



(Also for electroluminescent samples)

EMISSION MODE
(the near-field interacts with the sample and the result of the interaction is collected and analyzed in the far field)

COLLECTION MODE
(the near-field produced by conventional irradiation is collected by the near-field probe)

Other configurations can be involved, e.g., apertureless, "photon tunneling", etc.

SNOM probes (tapered fibers) I

Different configurations for SNOM exist

Here we will mention mostly **aperture-SNOM**, which often exploits tapered optical fibers as probes

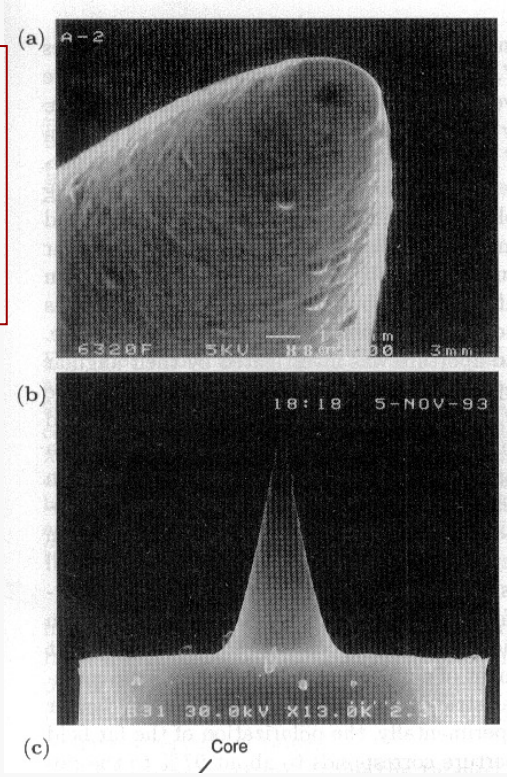
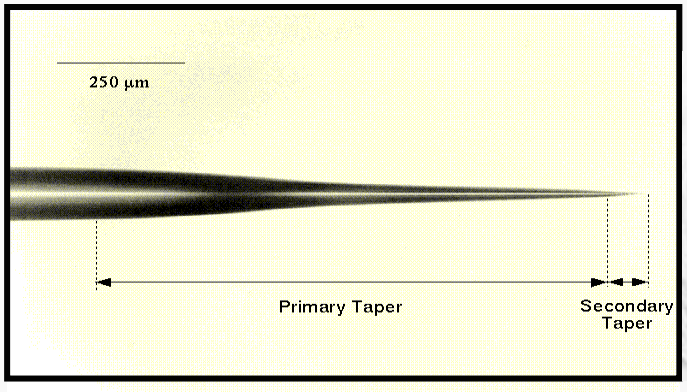
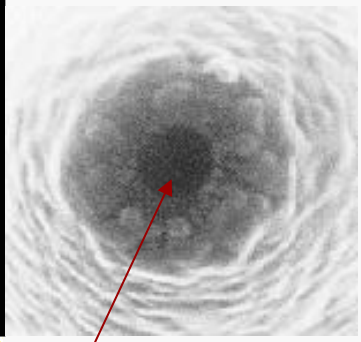


Fig. 7.7a–d. SEM micrographs of SNOM probes: (**a**) Tip of an aperture probe consisting of a thermally pulled tip of a quartz monomode fiber coated with aluminum. (By Courtesy of Sunney Xie). (**b**) Etched fiber tip according to Ohtsu [7.23]. The tip is fabricated by wet etching of a monomode quartz fiber. The thickness of the fiber coating is strongly reduced at the end of the fiber and a sharply pointed tip sticking out from the end is formed from the core. (**c**) Aperture probe fabricated on the basis of an etched tip, as shown in (**b**). The etched tip is coated with gold which is removed from the apex of the tip by a lithographic process such that a small aperture is formed with the tip sticking out [7.117]. (**d**) Tetrahedral tip. The tetrahedral tip consists of a glass fragment which is coated with metal. By courtesy of, R. Reichelt, Institute of medical Physics and Biophysics, University of Münster

Most common SNOM probe: tapered optical fiber, with metallization and apical aperture a<<λ (*aperture-SNOM*)





Typical aperture diameter: 50-100 nm

Note: metal layer (typ Cr, Ni) can absorb radiation → power entering the fiber cannot exceed the mW range!

Note: probe "throughput" (i.e., ratio between output/input power) is quite low for fiber probes, ~ 1/10000, but near field **intensity** can be large enough

SNOM probes (tapered fibers) II

Conventional probe fabrication

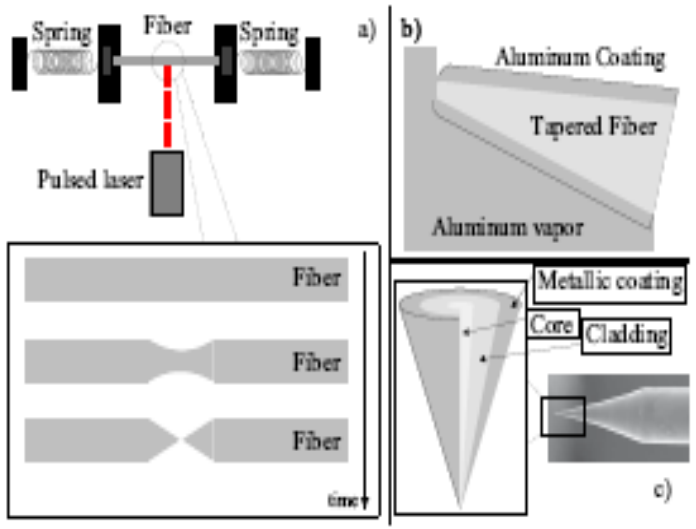


Figure 1.1: a) Heat and pull, one of the procedure to taper optical fibres for SNOM applications. b) Metallization of the SNOM fibre tip at steep angle to leave a sub-wavelength aperture at the end. c) SEM image of the resulting tip and sketch of its interior.

An alternative technique

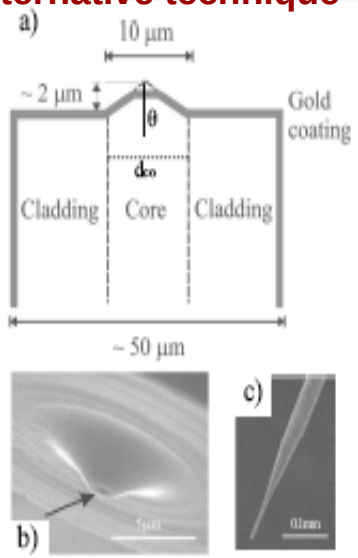


Figure 3.13: (a) Schematic of a SNOM fibre probe produced by selective chemical etching (SCE). The protruding cone is formed due to a slower etching rate of the core with respect to the cladding. After gold metallization (indicated by light gray lines), the tip is punched against a hard surface, producing a flattened apex with a sub-wavelength aperture at the center (black arrow), as evidenced by the SEM micrograph in (b).

An alternative probe

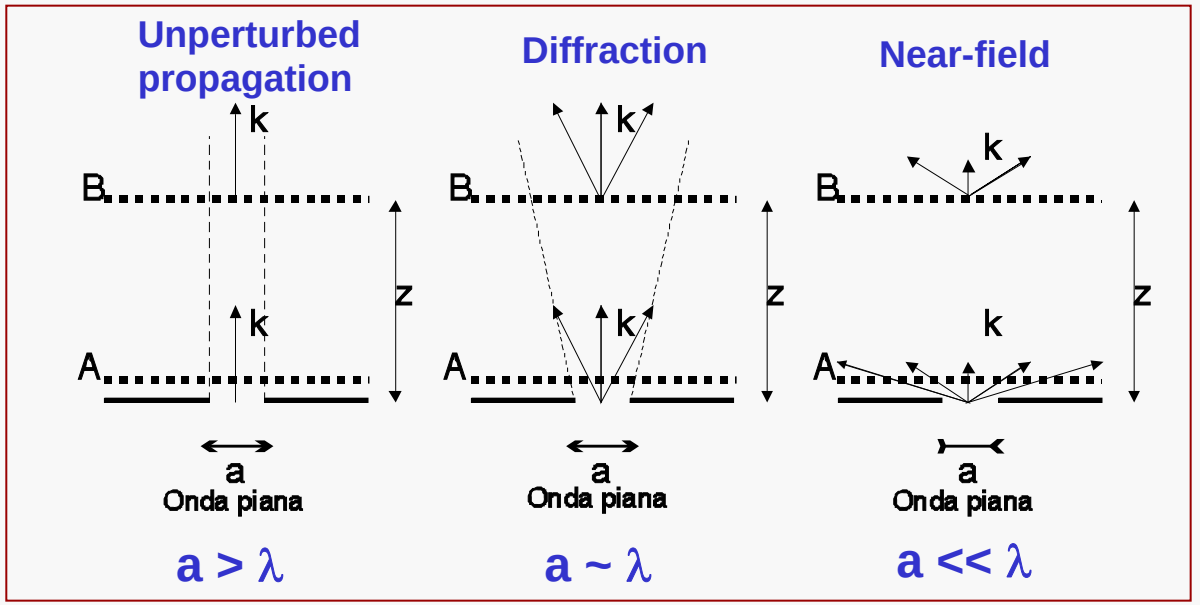
Figure 5.1: Back of an hollow cantilever. The square is the back aperture of the pyramidal hole produced by selective chemical etching. The inset shows a light spot coupled directly into the cantilever's hole by means of a microscope objective.

Many probes are available, including hollow cantilevers (similar to those for AFM, but with a pyramidal aperture)

Materiale tratto da Antonio Ambrosio
PhD Thesis Applied Physics, Pisa, 2005

Optical Near Field and Fourier optics

Ideal case (e.g., studied in the 20's by Synge and reworked in the 40's by Bethe): radiation sent onto a conductive plane with a subwavelength circular or elliptic aperture



In terms of Fourier optics, the subwavelength aperture produces radiation with extremely high spatial frequencies (transverse wavevectors)
→ space resolution no longer limited by diffraction

When aperture diameter is much smaller than the wavelength, far-field (propagating) intensity gets negligible compared to near-field (**non-propagating**) intensity

The subwavelength aperture acts as a hi-pass filter for the spatial frequencies

Reminders on evanescent waves

From Maxwell's equations, we know that the component of the electric field tangent to the interface of two dielectrics must be continuous across that interface. For a plane wave moving from one dielectric to another,

$$k_i \, \mathrm{Sin}[\theta_i] = k_t \, \mathrm{Sin}[\theta_t]$$

where k is the wave vector, theta is the angle between the wave propagation and the interface normal, and subscripts i and t stand for incident and transmitted wave fronts, respectively.

The frequency of the wave is identical on either side of the interface, so we have Snell's law:

$$n_i \, \mathrm{Sin}[\theta_i] = n_t \, \mathrm{Sin}[\theta_t]$$

with n the index of refraction of each of the media.

At incoming angles equal to and above the critical angle

$$\theta_c = \mathrm{ArcSin}\left[\frac{n_t}{n_i}\right]$$

The reflectance is 1, the transmittance is 0, and all energy is reflected back to the incoming side of the interface. However, as electric fields impinge on our uncharged dielectric interface, the boundary conditions due to Maxwell's equations and the conservation of momentum demand that there be a matching field on the far side. The component of k_t parallel to the interface is still equal to the component of k_i to obey the boundary conditions.

We have an incoming plane wave

$$E_i[r, t] = \mathcal{E}_i \, e^{(i \, (k_{ix} \, x + k_{iz} \, z - \omega \, t))}$$

With amplitude $\mathcal{E}i$, wave vector k and frequency w. The z direction is normal to the plan of interface of two dielectrics, and the x direction is chosen so that the wave vector lies entirely in the xz plane.

Likewise, the transmitted wave is

$$E_t[r, t] = \frac{1}{2} E_{0t} \, e^{(i \, (k_t \cdot r - \omega \, t))}$$

where

$$k_{tz} = k_t \, \mathrm{Cos}[\theta_t] = \pm k_t \sqrt{1 - \mathrm{Sin}^2[\theta_t]}$$

$$= \pm k_t \sqrt{1 - \frac{\mathrm{Sin}^2[\theta_i]}{(n_t / n_i)^2}}$$

The second line is thanks to Snell's law, and is an imaginary quantity for $\theta_i > \theta_c$.

Thus, the transmitted wave can be rewritten:

$$E_t[r, t] = \frac{1}{2} E_{0t} \, e^{(i \, (k_t \cdot r - \omega \, t))}$$

d is the positive, real, evanescent penetration depth.

$$d = \frac{\lambda}{2 \pi \sqrt{n_1^2 \, \mathrm{Sin}^2[\theta_i] - n_t^2}}$$

The positive exponential d z is unphysical if the second medium is semi-infinite.

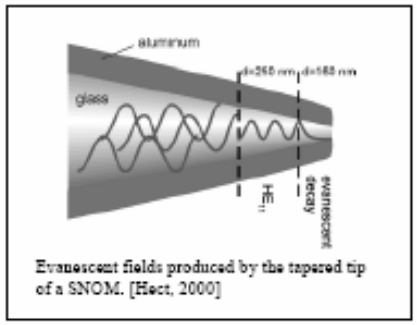
Suggested reading: Hecht et al., J. Chem Phys. 112, 7761 (2000).

Physical situations exist where *k* gets an imaginary character (e.g., evanescent waves)

Light propagation into a SNOM tapered fiber can be shown to lead to evanescent fields

Tom Hunt
http://www.physics.harvard.edu/~tomhunt/pubs/evanescent.pdf

A whole slew of scanning near field optical microscopes (SNOMs) have been developed by researchers over the past 15 years. These microscopes smash the diffraction limit of far field microscopes, potentially achieving resolution an order of magnitude better than a standard confocal microscope [Hect 2000]. One such SNOM is the scanning tunneling optical microscope. This microscope uses a sharp glass tip to locally frustrate total internal reflection below a surface, indirectly imaging features on that surface at high spatial resolution. Aperture based SNOMs are more common and more practical. [Hect 2000] They produce an evanescent field by forcing light through a small aperture (see figure). The evanescent field locally illuminates the sample. Once free of the aperture, the field is no longer evanescent, and it expands in the far field to be picked up by a detector. To achieve high resolution, the aperture must be small, and close to the sample surface so the field is tightly confined when it interacts with the sample. Full analysis of the field-sample interaction of a SNOM is a difficult or impossible undertaking, but the data produced can yield important and detailed information about a sample surface.



Evanescent fields produced by the tapered tip of a SNOM. [Hect, 2000]

Optical Near Field and diffraction

Diffraction and the Heisenberg's microscope

The resolving power of a conventional microscope, i.e. the distance (Δx) between two object points which a microscope can just resolve, depends on the numerical aperture (NA) of the objective and the wavelength (λ) of the light used. The relation between these quantities is fixed by the *Rayleigh's criterion* that sets:

$$\Delta x = 0.61 \frac{\lambda}{NA} \qquad (1.1)$$

The whole limits discussed above can be regarded in terms of quantum mechanics. In fact, applying the Heisenberg's uncertain principle to the components (x_i) of a photon's position and to those of the linear momentum (p_i) of the photon:

$$\Delta x_i \cdot |p_i| \geq h \qquad (1.2)$$

where i is an index indicating the projections along x, y or z axes. Each component of the linear momentum of the photon is related to the corresponding components (k_i) of the light wavevector ($\vec{k}$) by $p_i = \hbar k_i$. The relation 1.2 may then be written as:

$$\Delta x_i \geq \frac{1}{|k_i|} = \frac{1}{|n_i|} \frac{\lambda}{2\pi} \qquad (1.3)$$

This formula fixes the physical limit for the linear dimensions of a focused beam as well as the achievable optical resolution. The possible values of k_i are limited by the mathematical condition being between each vector and its components:

$$|k| = \sqrt{k_x^2 + k_y^2 + k_z^2} \qquad (1.4)$$

Classical optics and microscopy employ free propagating waves for which all the components k_i are real. In this case $k_i^2 = |k_i|^2$ and the relation 1.3 limits the best resolution achievable to values no much smaller than $\lambda/2$.

Materiale tratto da Antonio Ambrosio
PhD Thesis Applied Physics, Pisa, 2005

Propagating waves:
k_i are real and $|k_x| \leq k = 2\pi/\lambda$

⇩

Heisenberg's principle: $k_x \geq 2\pi / \Delta x$

$\Delta x \geq \lambda$
(actual parameters give the Abbe's limit)

... but ...

In **non-propagating (e.g., evanescent)** waves:
k_i can be imaginary, and, e.g.: $|k_x| \geq k = 2\pi/\lambda$

⇩

The Heisenberg's principle is no longer ruling the ultimate resolution!

Sub-diffraction space resolution associated with the non-propagating character of the field

Optical Near Field and Maxwell eqs. I

Typical approach to radiative problems is based on manipulations of Maxwell's eqs exploiting potentials (scalar and vector)

Vector potential

Possiamo infatti esprimere il campo $\vec{B}$ in termini del potenziale vettore $A(\vec{r})$, con:

$$\vec{\nabla} \times \vec{A} = \vec{B} \tag{5}$$

dove $\vec{A} = \vec{A}(\vec{r}, t)$, con che l'equazione $\vec{\nabla} \cdot \vec{B} = 0$ é automaticamente soddisfatta.

Scalar potential (in the Lorentz gauge)

$$\vec{\nabla} \cdot \vec{A} + \epsilon\mu \frac{\partial V}{\partial t} = 0$$

$$\vec{\nabla} V = -(\vec{E} + \frac{\partial \vec{A}}{\partial t}) \tag{6}$$

Potentials and charge, currents density

$$\vec{A}(\vec{r}, t) = \frac{\mu}{4\pi} \int_\tau \frac{\vec{J}(\vec{r'}, t - \Delta r/v)}{\Delta r} \, d\tau' \tag{12}$$

$$V(\vec{r}, t) = \frac{1}{4\pi\epsilon} \int_\tau \frac{\rho(\vec{r'}, t - \Delta r/v)}{\Delta r} \, d\tau' \tag{13}$$

dove $\vec{r}$ é il vettore tracciato a partire dall'origine al punto P dove si calcola i campi; $\vec{r'}$ é il vettore tracciato a partire dall'origine al generico punto dove é presente l'elemento di carica $\rho d\tau$ o l'elemento di corrente $\vec{J} d\tau$; $\Delta r = |\vec{r} - \vec{r'}|$ il vettore che va dal punto dove sono i generici elementi di carica/corrente al punto P; $v = 1/\sqrt{\epsilon\mu}$ é la velocitá di propagazione delle onde elettromagnetiche nel mezzo.

In the far-field (retardation effects!):

e possono essere trascurati. Quindi, a grandi distanze dal sistema di cariche, i potenziali diventano:

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi r} \int_\tau \vec{J}(\vec{r'}, t - r/c + \frac{\vec{r'} \cdot \vec{n}}{c}) \, d\tau' \tag{19}$$

$$V(\vec{r}, t) = \frac{1}{4\pi\epsilon_0 r} \int_\tau \rho(\vec{r'}, t - r/c + \frac{\vec{r'} \cdot \vec{n}}{c}) \, d\tau' \tag{20}$$

dove si é ipotizzato di essere nel vuoto ($\epsilon = \epsilon_0$, $\mu = \mu_0$) Il termine $\vec{r'} \cdot \vec{n}/c$ tra gli argomenti degli integrandi indica quanto l'onda elettromagnetica proveniente dalle parti piú distanti del sistema che irradia é ritardata rispetto all'onda proveniente dalle parti vicine. In altre parole, il termine $\vec{r'} \cdot \vec{n}/c$ determina il tempo che l'onda elettromagnetica impiega ad attraversare il sistema. Se la velocitá delle cariche é v, in tale tempo esse si saranno spostate di $v(\vec{r'} \cdot \vec{n})/c$. Il ritardo interno al sistema sará trascurabile quando tale distanza sará piccola rispetto alle dimensioni del sistema, r'. Quindi la condizione é: $v(\vec{r'} \cdot \vec{n})/c \ll r'$ o, equivalentemente: $v \ll c$. In tal caso le cariche non avranno il tempo di cambiare apprezzabilmente le loro posizioni durante il tempo che l'onda impiega ad attraversare il sistema.

Da Chiara Roda
www.df.unipi.it/~roda/fisica2/postscript/potenziali.ps

Trascurando il ritardo interno al sistema, l'espressione del potenziale vettore diventa:

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi r} \int_\tau \vec{J}(\vec{r'}, t - r/c) \, d\tau' \tag{21}$$

Sostituiamo in questa espressione a $\vec{J}$ il prodotto $\rho\vec{v}$, dove ora la densitá di carica ρ é calcolata al tempo $t - r/c$. Si ottiene quindi:

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi r} \int_\tau \rho(\vec{r'}, t - r/c)\vec{v} \, d\tau'$$

Notiamo ora che, se avessimo a che fare con una singola carica puntiforme e, l'integrale nell'ultima espressione scritta sarebbe semplicemente $e\vec{v}$. Per un sistema di cariche esso sará:

$$\sum_i e_i \vec{v}_{i\,(t-r/c)}$$

dove la somma va calcolata al tempo ritardato $t - r/c$. Ora, il momento di dipolo di un sistema é definito come:

$$\vec{p} = \sum e_i \vec{r'}_i$$

per cui é:

$$\vec{\dot{p}} = \sum_i e_i \frac{d\vec{r'}_i}{dt}$$

ed infine, il potenziale vettore diviene:

$$\vec{A} = \frac{\mu_0}{4\pi r} \vec{\dot{p}}_{(t-r/c)}$$

Retardation (inside the emitter) is neglected, i.e., emitter is considered pointlike

Abbiamo in definitiva ottenuto delle espressioni per i potenziali vettore e scalare relativi ad un dipolo variabile nel tempo. Le espressioni ottenute sono:

$$V = \frac{1}{4\pi\epsilon_0 r^3} [\vec{p} + \frac{r}{c}\vec{\dot{p}}] \cdot \vec{r}_{\,(t-r/c)} \tag{22}$$

$$\vec{A} = \frac{\mu_0}{4\pi r} \vec{\dot{p}}_{(t-r/c)} \tag{23}$$

Retarted potential methods allow to derive potential wavefunctions

Optical Near Field and Maxwell eqs. II

Explicit field solution in the far-field

Suggested reading: Jackson, Classical Electrodynamics

Le equazioni di Maxwell nel vuoto sono in genere risolte assumendo che la distanza del punto $\mathbf{r}$ in cui si vogliono calcolare i campi $\mathbf{E}(\mathbf{r})$ e $\mathbf{B}(\mathbf{r})$ dalla sorgente sia molto maggiore delle dimensioni della stessa. In realtà [21] le soluzioni *complete* delle equazioni per i campi generati da un dipolo elettrico $\mathbf{p}$ posto nell'origine sono:

$$\mathbf{E}(\mathbf{r}) = k^2(\mathbf{n}\times\mathbf{p})\times\mathbf{n}\frac{e^{ikr}}{r} + [3\mathbf{n}(\mathbf{n}\cdot\mathbf{p}) - \mathbf{p}]\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)e^{ikr} \qquad (27)$$

$$\mathbf{B}(\mathbf{r}) = k^2(\mathbf{n}\times\mathbf{p})\frac{e^{ikr}}{r}\left(1 - \frac{1}{ikr}\right) \qquad (28)$$

Dove $\mathbf{n}$ è il versore di $\mathbf{r}$ e k è il modulo del vettore d'onda $\mathbf{k}$, con $k = \omega/c = 2\pi/\lambda$. Sia $\mathbf{E}$ che $\mathbf{B}$ sono dati dalla somma di due contributi con un diverso andamento in r. Nel limite di campo lontano, con $kr \gg 1$, si ritrovano le usuali soluzioni

$$\mathbf{E}_{far}(\mathbf{r}) \approx k^2(\mathbf{n}\times\mathbf{p})\times\mathbf{n}\frac{e^{ikr}}{r} \qquad (29)$$

$$\mathbf{B}_{far}(\mathbf{r}) \approx k^2(\mathbf{n}\times\mathbf{p})\frac{e^{ikr}}{r} \qquad (30)$$

per le quali vale $\mathbf{E} = \mathbf{B}\times\mathbf{n}$. Queste soluzioni rappresentano onde sferiche la cui ampiezza scala come $1/r$.

Explicit field solution in the near-field

Nel limite opposto in cui $kr \ll 1$ sviluppando e^{ikr} si ottengono i campi

$$\mathbf{E}_{near}(\mathbf{r}) \approx [3\mathbf{n}(\mathbf{n}\cdot\mathbf{p}) - \mathbf{p}]\frac{1}{r^3} \qquad (31)$$

$$\mathbf{B}_{near}(\mathbf{r}) \approx ik(\mathbf{n}\times\mathbf{p})\frac{1}{r^2} \qquad (32)$$

i quali sono nella stessa forma di quelli prodotti da un dipolo statico o lentamente variabile nel tempo, come ci si aspetta nel limite di $kr \to 0$.

Static dipole means non propagating (the oscillating temporal behavior is indeed preserved)!

Per comodità consideriamo separatamente i campi nel caso $\mathbf{r} \perp \mathbf{p}$ e nel caso $\mathbf{r} \parallel \mathbf{p}$. Definendo il parametro adimensionale $x = kr$, le (27) e (28) si riscrivono

$$\mathbf{E}_\perp(x) = \frac{e^{ix}}{x}k^3\mathbf{p}\left(1 - \frac{1-ix}{x^2}\right) \qquad (33)$$

$$\mathbf{B}_\perp(x) = i\frac{e^{ix}}{x}(\mathbf{n}\times\mathbf{p})\frac{1-ix}{x^2} \qquad (34)$$

$$\mathbf{E}_\parallel(x) = 2\frac{e^{ix}}{x}k^3\mathbf{p}\frac{1-ix}{x^2} \qquad (35)$$

$$\mathbf{B}_\parallel(x) = 0 \qquad (36)$$

Nel primo caso $\mathbf{E}$ e $\mathbf{B}$ sono entrambi trasversi, mentre nel secondo caso il campo elettrico ha una componente longitudinale non nulla dovuta al contributo del campo prossimo. Tale componente tende a zero per $x \to \infty$. Le ampiezze dei campi $\mathbf{E}_\perp(x)$, $\mathbf{B}_\perp(x)$ e $\mathbf{E}_\parallel$ sono rispettivamente

$$|\mathbf{E}_\perp(x)| = k^3\frac{p}{x}\sqrt{1 - \frac{1}{x^2} + \frac{1}{x^4}} \qquad (37)$$

$$|\mathbf{B}_\perp(x)| = k^3\frac{p}{x}\sqrt{1 + \frac{1}{x^2}} \qquad (38)$$

$$|\mathbf{E}_\parallel(x)| = \frac{2k^3}{x^2}p\sqrt{1 + \frac{1}{x^2}} \qquad (39)$$

Come si può vedere il campo $\mathbf{E}_\parallel$ diventa dominante per x piccoli. Questo significa che per valori di r molto più piccoli di λ gli effetti del campo prossimo diventano importanti. Questi risultati, benché ricavati nel caso particolare di un dipolo elettrico nell'origine, sono molto generali e almeno qualitativamente descrivono l'emissione in campo prossimo di ogni tipo di sorgente.

The field close to the emitter (the near field) holds unique features

Nicola Paradiso, Tesi di Laurea in Fisica, Pisa 2005

Electric/magnetic dipoles in the near-field

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0} k^2 (\mathbf{n}\times\mathbf{p})\times\mathbf{n}\frac{e^{ikr}}{r} + [3\mathbf{n}(\mathbf{n}\cdot\mathbf{p}) - \mathbf{p}]\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)e^{ikr}$$

$$\mathbf{H} = \frac{ck^2}{4\pi}(\mathbf{n}\times\mathbf{p})\frac{e^{ikr}}{r}\left(1 - \frac{1}{ikr}\right)$$

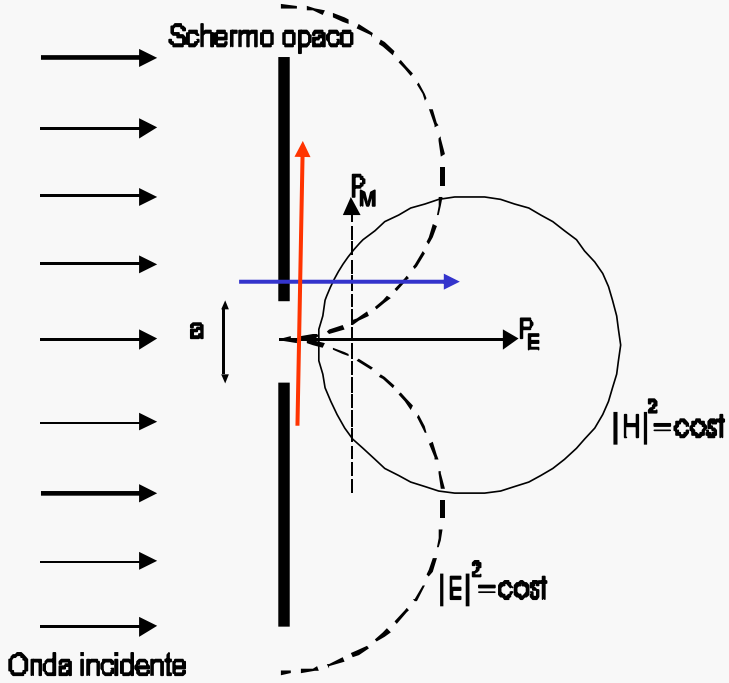
Behavior of electric and magnetic fields

Per $kr \ll 1$: **Near-field condition**

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}[3\mathbf{n}(\mathbf{n}\cdot\mathbf{p}) - \mathbf{p}]\frac{1}{r^3}$$

$$\mathbf{H} = \frac{i\omega}{4\pi}(\mathbf{n}\times\mathbf{p})\frac{1}{r^2}$$

Equivalente all'emissione di un dipolo statico.



Near-field space distribution of electric and magnetic fields in the ideal case can be approximated with two mutually orthogonal dipoles (electric and magnetic, respectively)

Materiale tratto dal seminario di Stefano Tirelli, 2006

Optical Near Field and Maxwell eqs. III

Bouwkamp, following Bethe's ideas, carefully considered the boundary conditions

IV. BOUWKAMP'S SOLUTION

50's

Bouwkamp's approach [7] in solving the diffraction problem consists in first deriving the solution for a disk and then using Babinet's principle to obtain the magnetic currents for the case of the aperture. [2]

Electric currents are induced on the disk due to the interaction with the incident fields. Once such currents are known, the vector potential can be obtained through:

$$\vec{A} = \frac{\mu_0}{4\pi}\int \vec{I}\,\frac{\exp(ikR)}{R}\,da\,, \qquad (29)$$

which in the Lorentz gauge gives the scattered fields:

$$\vec{H}^s = \frac{1}{\mu_0}\nabla\times\vec{A} \qquad \vec{E}^s = i\omega\vec{A} - \frac{c}{ik}\nabla\left(\nabla\cdot\vec{A}\right). \qquad (30)$$

Using the appropriate boundary conditions for the problem:

$$E_x^s = -1,\ E_y^s = 0,\ H_z^s = 0 \quad (x^2+y^2 \le a) \qquad (31)$$

combined with eqs. *(29)* and *(30)* a system of integro-differential equations is obtained for the fields in the aperture:

$$\frac{\partial A_x}{\partial y} = \frac{\partial A_y}{\partial y}$$

$$\frac{\partial^2 A_x}{\partial x^2} + \frac{\partial^2 A_x}{\partial y^2} + k^2 A_x = ik$$

$$\frac{\partial^2 A_y}{\partial x^2} + \frac{\partial^2 A_y}{\partial y^2} + k^2 A_y = 0$$

and

$$A_x = \frac{\mu_0}{4\pi}\int I_x\,\frac{\exp(ikR)}{R}\,da \qquad A_y = \frac{\mu_0}{4\pi}\int I_y\,\frac{\exp(ikR)}{R}\,da \qquad (32)$$

This system of equations cannot be solved analytically. In order to obtain a solution, it is necessary to use the symmetry of the problem and impose the additional boundary conditions at the edge of the disk.

Since a plane wave at normal incidence is being considered, the problem has axial symmetry. This together with the boundary conditions at the edge of disk implies that the behavior of the electric current density must satisfy:

$$\frac{1}{c}I_x = \frac{A(\rho)+B(\rho)\cos 2\varphi}{\sqrt{1-\rho^2}}, \quad \frac{1}{c}I_y = \frac{B(\rho)}{\pi^2\sqrt{1-\rho^2}}\sin 2\varphi \qquad (33)$$

from which the following condition can be obtained:

$$I_y = -\frac{1}{2}\frac{\partial I_x}{\partial\varphi} = \frac{1}{2}\left(y\frac{\partial I_x}{\partial x} - x\frac{\partial I_x}{\partial y}\right) \qquad (34)$$

A similar condition can be obtained for the vector potential. These relations allow eqs. *(32)* to be reduced to an integral equation containing only one variable (I_x or I_y), so that through a series expansion the coefficients A and B of eqs. *(33)* can be obtained.

Through the use of Babinet's principle the electric currents will act as magnetic current densities, which to first order are given by:

$$K_y = \frac{8ik}{3\pi}\frac{2a^2-x^2-2y^2}{\sqrt{a^2-x^2-y^2}}, \quad K_x = -\frac{8ik}{3\pi}\frac{xy}{\sqrt{a^2-x^2-y^2}} \qquad (35)$$

The fields in the aperture can now be obtained:

$$E_x = -\frac{4ik}{3\pi}\frac{2a^2-x^2-2y^2}{\sqrt{a^2-x^2-y^2}} \qquad H_x = 0$$

$$E_y = -\frac{4ik}{3\pi}\frac{xy}{\sqrt{a^2-x^2-y^2}} \qquad H_y = \frac{1}{\mu_0 c} \qquad (36)$$

$$E_z = 0 \qquad H_z = -\frac{4}{\mu_0 c\pi}\frac{y}{\sqrt{a^2-x^2-y^2}}$$

These equations can now be directly compared to Bethe's results [eqs. *(28)*]. As can be seen, the expressions for the magnetic fields are the same, which is due to the fact that to first order there is no contribution from the magnetic current density to the magnetic fields.

A contour plot of both Bethe's and Bouwkamp's results can be seen in *Fig.2* and *Fig. 3*, respectively. In these plots, the lighter regions represent higher intensity.

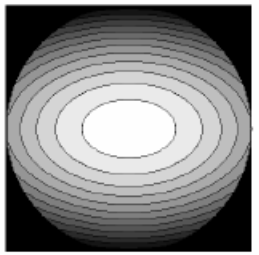


Fig. 2. Total electric field $|E|^2$ in the aperture obtained by Bethe.

Fig. 3. Total electric field $|E|^2$ in the aperture obtained by Bouwkamp.

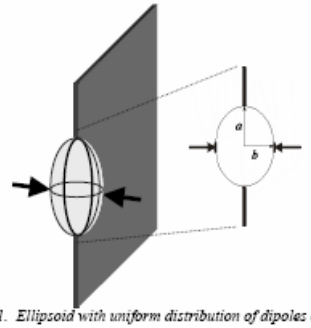


Fig. 1. Ellipsoid with uniform distribution of dipoles oriented along the plane of the aperture. The charge distribution in the hole is obtained by making the length of axis b very small.

Diffraction by a Small Circular Aperture
Contributed by Alberto M. Marino and Giovanni Piredda

How to calculate the actual near-field distribution

3.2 THEORETICAL MODELING OF NEAR-FIELD NANOSCOPIC INTERACTIONS

Although Maxwell's equations provide a general description of electromagnetic phenomena, their analytical solutions are limited to relatively simple cases and rigorous treatment of nanoscale optical interactions presents numerous challenges. The various ways of approaching the theory of near-field optics can be classified according to the following considerations (Courjon, 2003):

- The physical model of the light beam
- The space chosen to carry out the modeling (i.e., the direct space or Fourier space modeling)
- Global or nonglobal way of treating the problem (e.g., performing separate calculation for the field in the sample and then computing the capacity of the tip to collect the field)

Among several methods used for electromagnetic field calculation, one can distinguish techniques derived from the rigorous theory of gratings like the differential method (Courjon, 2003) and the Reciprocal-Space Perturbative Method (RSPM), as well as techniques that operate in direct space like the Finite-Difference Time-Domain Method (FDTD) and the Direct-Space Integral Equation Method (DSIEM).

In general, analytical solutions can provide a good theoretical understanding of simple problems, while a purely numerical approach (like that of the FTDT method) can be applied to complex structures. A compromise between a purely analytical and a purely numerical approach is the multiple multipole (MMP) model (Girard and Dereux, 1996). With the MMP model, the system being simulated is divided into homogeneous domains having well-defined dielectric properties. Within individual domains, enumerated by the index i, the electromagnetic field $f^{(i)}(\mathbf{r}, \omega_0)$ is expanded as a linear combination of basis functions

$$f^{(i)}(\mathbf{r}, \omega_0) \approx \sum_j A_j^{(i)} f_j(\mathbf{r}, \omega_0) \tag{3.1}$$

where the basis functions $f_j(\mathbf{r}, \omega_0)$ are the analytical solutions for the field within a homogeneous domain. These basic functions satisfy the eigenwave equation for the eigenvalue q_j (analogous to the equation in Table 2.1):

$$-\nabla \times \nabla \times f_j(\mathbf{r}, \omega_0) + q_j^2 f_j(\mathbf{r}, \omega_0) = 0 \tag{3.2}$$

MMP can use many different sets of basis fields, but fields of multipole character are considered the most useful. The parameters $A_j^{(i)}$ are obtained by numerical matching of the boundary conditions on the interfaces between the domains.

As an example of the use of this technique for investigations of nonlinear optical processes in the near field, we show here investigations of second harmonic generation in a noncentrosymmetric nanocrystal exposed to fundamental light from a near-field scanning tip (Jiang et al., 2000).

One notes that a consequence of nonlinear optical interaction in the near-field is that the phase-matching conditions do not need to be fulfilled because the domains are much smaller than the coherence length. Starting from Maxwell's equations, the electric fields of the fundamental and the second harmonic (SH) wave can be shown to satisfy the nonlinear coupled vector wave equations

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, \omega_0) - \frac{\omega_0^2}{c^2} \varepsilon(\mathbf{r}, \omega_0)\mathbf{E}(\mathbf{r}, \omega_0) = 4\pi \frac{\omega_0^2}{c^2} \mathbf{P}^2(\mathbf{r}, \omega_0) \tag{3.3}$$

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, 2\omega_0) - \frac{4\omega_0^2}{c^2} \varepsilon(\mathbf{r}, 2\omega_0)\mathbf{E}(\mathbf{r}, 2\omega_0) = 4\pi \frac{4\omega_0^2}{c^2} \mathbf{P}^2(\mathbf{r}, 2\omega_0) \tag{3.4}$$

where $\varepsilon(\mathbf{r}, \omega_0)$ and $\varepsilon(\mathbf{r}, 2\omega_0)$ are linear dielectric functions for the fundamental and the SH waves, respectively.

The propagation constant k_z along the z direction is

$$k_z = (\mathbf{k}^2 - \mathbf{k}_{||}^2)^{1/2} = k_0(1 - n_1^2 \sin^2 \theta)^{1/2} \tag{3.5}$$

where $k_0 = 2\pi/\lambda$, λ is the wavelength of illumination light in free space; n_1 is the refractive index of the tip, and θ is the incident angle. If $1 - n_1^2 \sin^2 \theta > 0$ (i.e., k_z is real), the waves will propagate with constant amplitude between the probe and the sample, which corresponds to the "allowed light" in the sample. In the areas where k_z is imaginary, the waves will decay exponentially within distances comparable to the wavelength, thus such waves have evanescent character and produce the "forbidden light" in the sample. From the electrical field distribution of the fundamental wave calculated with the MMP method, we can obtain the electrical field distribution of the SH wave and the different contributions of "allowed light" and "forbidden light."

Da P.Prasad, Nanophotonics (Wiley, 2004)

Detailed description of the near field produced by actual probes requires sophisticated numerical methods and accounting for multipole distributions

Actual calculations: an example

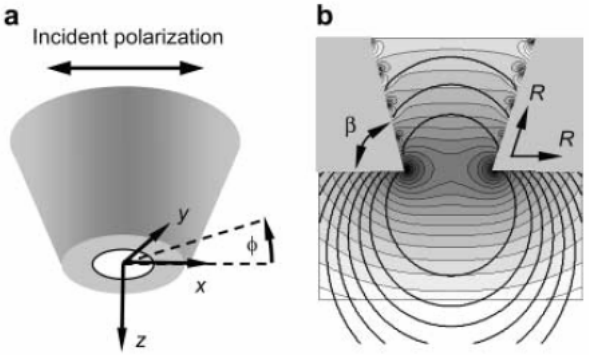


Fig. 2 – (a) 3D view and (b) (x, z) cross-section of an optical fiber tip with the associated electric-field lines and a grey-scale logarithmic intensity map. The polarization of the incident light propagating in the fiber core (z direction) is oriented along the x-axis. The electric field is generated by polarization charges in a perfect metal coating modelled by a surface charge distribution on the lateral and inner surface of the metal coating surrounding the tapered fiber (aperture diameter $2a$). The distance from the aperture rim R and the cone angle β are indicated in the figure. The electric-field lines in the immediate vicinity of the metal surface show artefacts mainly due to the numerical discretization procedure, which become, however, insignificant at larger distances.

In a tapered optical fiber near-field distribution can be simulated by considering nanosized emitters

simulation

Bethe case

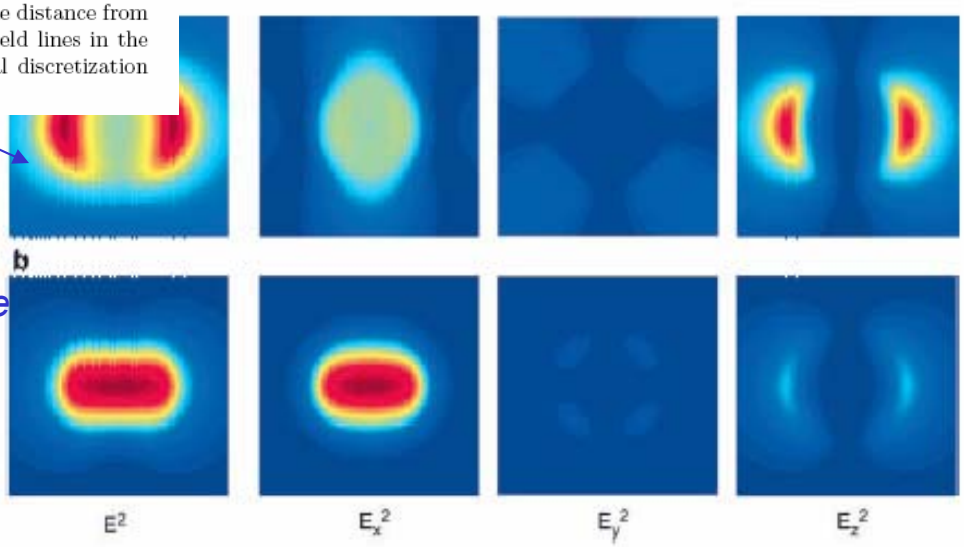


Fig. 3 – Intensity maps for the electric field (E^2 and components $E_x{}^2$, $E_y{}^2$, and $E_z{}^2$) at a distance of $z = 0.3a$ from the tip aperture of radius a created by (a) a surface charge distribution as illustrated in fig. 2, and (b) a planar Bethe aperture (Bethe-Bouwkamp solution). All images show an area of $4a$ by $4a$, and the images in each row share the same color bar. The two images on the left correspond to those that would be obtained using a pointlike, scalar detector for the electric field (*e.g.* an idealized, infinitely small fluorescent nanosphere). The images of the last three columns correspond to fluorescence images predicted for pointlike vector detectors for the electric field (*e.g.* single, fluorescent molecules) oriented along the x, y, and z direction, respectively.

Actual space distribution of the near-field can differ from the ideal case

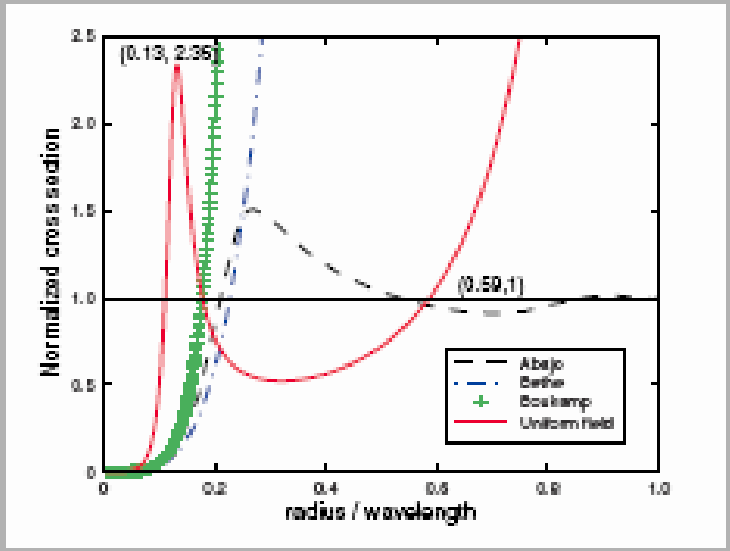
EUROPHYSICS LETTERS 1 April 2004
Europhys. Lett., **66** (1), pp. 41–47 (2004)
DOI: 10.1209/epl/i2003-10138-7

The optical near-field of an aperture tip

A. Drezet, M. J. Nasse, S. Huant and J. C. Woehl(*)

Actual calculations: another example

Abstract: Diffraction of normally incident light by a subwavelength circular aperture is calculated analytically. The aperture is opened on a perfectly conducting planar screen with infinitesimal thickness. In our model, the aperture is replaced by uniform magnetic currents and charges. The model allows one to obtain the normalized cross section for the aperture radius up to half of the wavelength, which exceeds the 0.2 wavelength limit of the Bethe-Bouwkamp's dipole model [1–3]. Also, in addition to reproducing the $(ka)^4$ dependence, which is characteristic of the dipole mode, our uniform field model explains the transmission enhancement obtained in Abajo's numerical simulation [4].



Normalized transmission cross section as a function of the normalized radius

In any case, near-field distribution gets a maximum value for apertures ~ 0.1 λ

Light diffraction by a subwavelength circular aperture

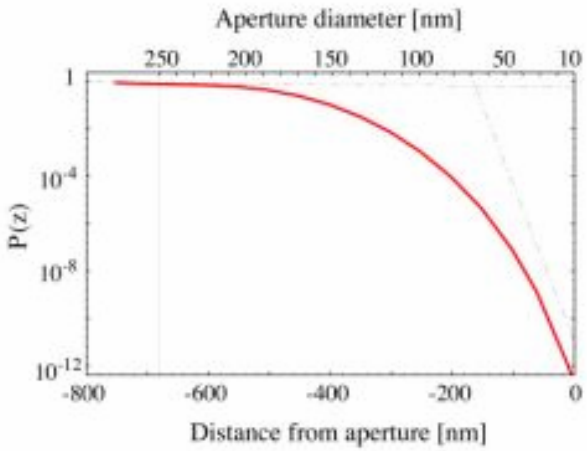
Che-Wei Chang, A.K. Sarychev, and V.M. Shalaev*

Laser Phys. Lett. 2, No. 7, 351–355 (2005) / DOI 10.1002/lapl.200510006 Laser Physics Letters 351

Near-field throughput

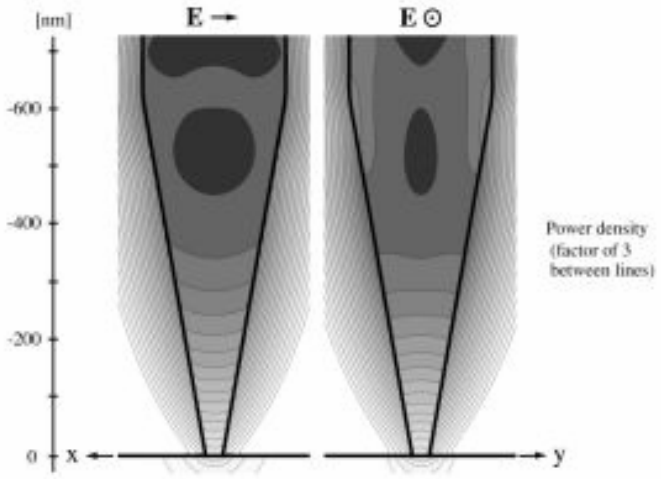
However, the amount of light that can be transmitted by a small aperture poses a limit on how small it can be made before nothing gets though. To a degree this can be lived with, as more optical power can be generated, but the cutoff is so severe that it cannot be made smaller. As the figures illustrate, this is not a subtle extinction.

Little power is coupled into the near-field!



L. Novotny and D. W. Pohl, in *Photons and Local Probes*, NATO ASI Series E, p.21-33, Kluwer Academic, 1995.

When the aperture is 100 nm, the cutoff is down four orders of magnitude, and when it reaches 50 nm, only one part in 10^8 makes it through. Furthermore, the input power cannot be increased arbitrarily because 1/3 of the power is absorbed in the coating. Increasing the input power above approximately 10mW will destroy the coating. This severly limits the signal-to-noise ratio of small apertures, and is the reason our group uses another approach.



L. Novotny and D. W. Pohl, in *Photons and Local Probes*, NATO ASI Series E, p.21-33, Kluwer Academic, 1995.

http://xray.optics.rochester.edu/workgroups/novotny/snom.html

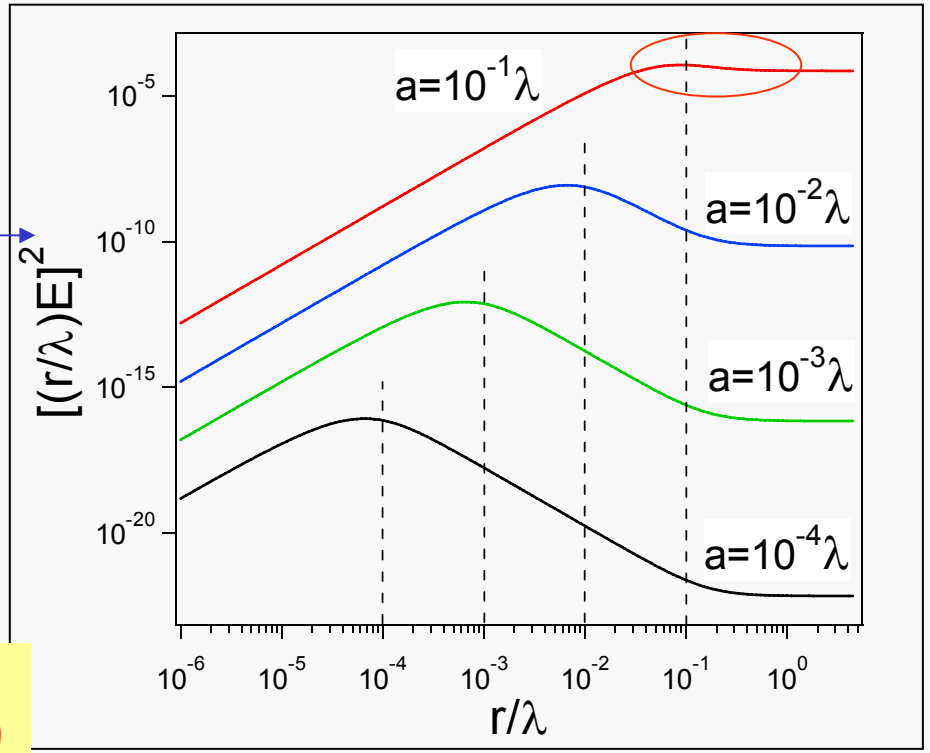
Near-field extinction

In the presence of a non-propagating (evanescent) wave, field undergoes a fast extinction as a function of the distance from the aperture (*z*)

$$I(0,0,z) \cong \left(\frac{8ka}{3\pi}\right)^2 e^{\frac{-3\pi z}{4a}}$$

(ideal case, after Bethe)

"Normalized" field intensity as a function of the distance **(*r*)** from the aperture (estimation)



In practice:
a ~ 50 nm (~ λ/10) and the field drops rapidly to zero for *r* > *a* (i.e., for *r*/λ ~ 0.1)

The space distribution of the near-field (confined within a distance ~ *a*) gives SNOM its sub-diffraction space resolution

How to concern a sample with the near-field

A distinctive feature of near-field is that its amplitude rapidly drops to zero within a range ~*a*

⇩

During the scan, the probe tip must be kept "close" to the surface (typ at a distance < 10 nm)

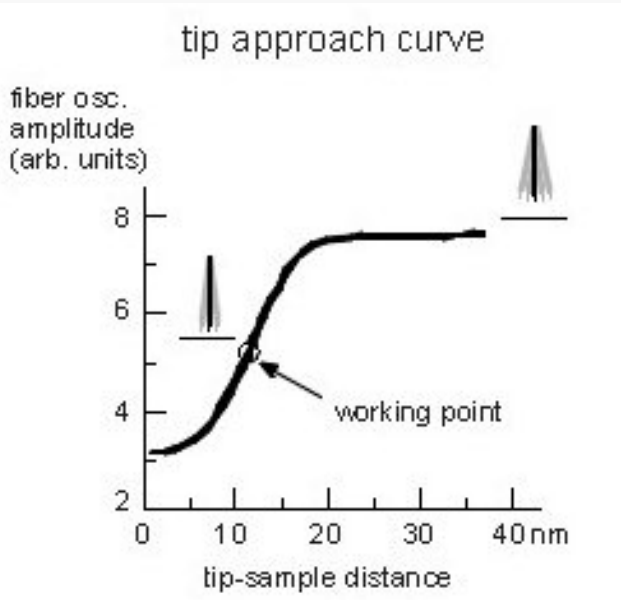
"Constant gap" operation is strictly required for the SNOM images to be reliable

⇩

A method to continuously monitor tip/sample distance is needed
A feedback acting on the vertical piezo displacement is used (as, e.g., in non-contact AFM)

⇩

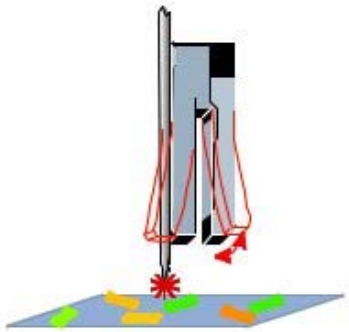
A topography image is simultaneously acquired during each scan, with a lateral resolution depending on the probe size, typ in the 100 nm range

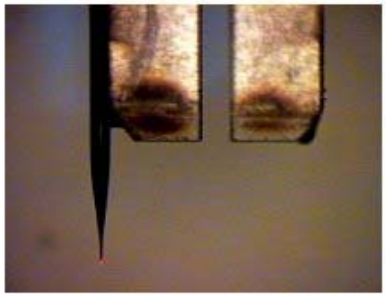


If the tip is kept in **longitudinal** oscillation, the oscillation amplitude depends on the distance due to shear-forces (mostly associated with friction of the air layers between tip and surface)

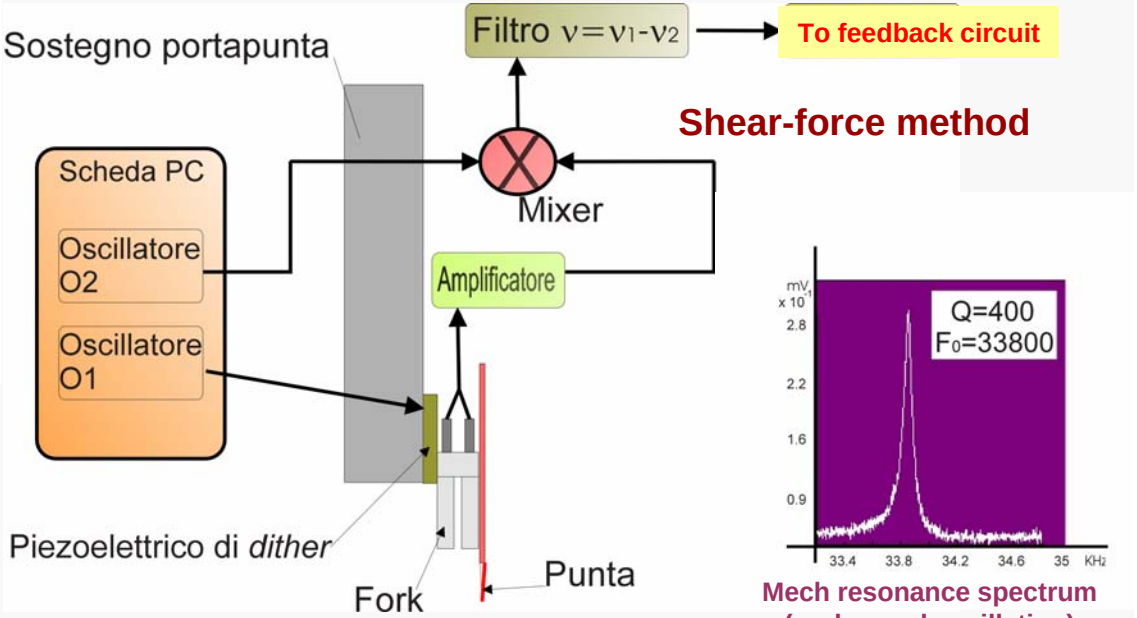
Notes:
Oscillation amplitude must be small (typ ~ 1 nm) to prevent resolution loss
Oscillation at resonance frequency is required to get maximum sensitivity

Shear-force method





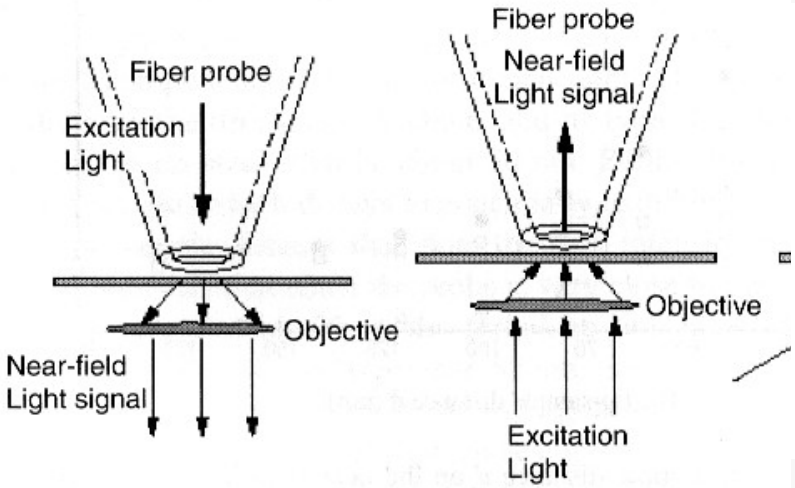
Fiber tip is glued onto a *tuning fork* (a quartz diapason acting as a mechanical oscillator and showing capabilities to measure the oscillation amplitude)



Mech resonance spectrum (undamped oscillation)

- A dithering piezoelectric transducer keeps the probe tip in oscillation along a direction **parallel** to the surface
- Oscillation amplitude is monitored by a *tuning fork*
- When the distance gets smaller (typ., below 10 nm), the oscillation is **damped** (and phase is changed) due to **shear-forces** involving many effects (e.g., viscous interaction of the air layer between tip and sample)
- Similar to AFM in tapping mode, but for the oscillation direction, the relevant distance and the involved mechanisms

Operation modes for SNOM



Illumination mode **Collection mode**

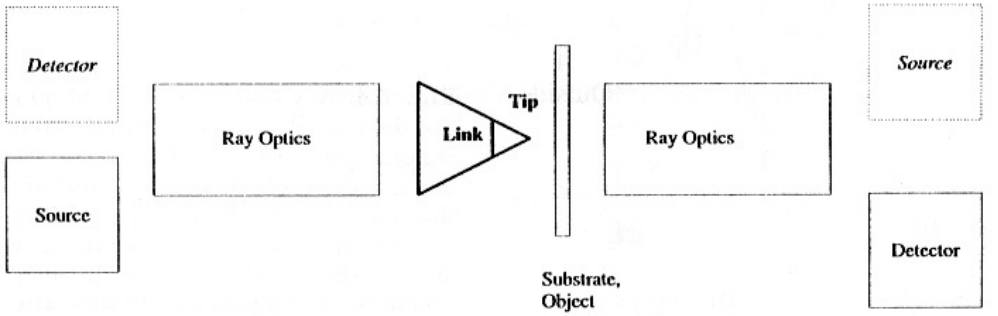


Fig. 7.3. The characteristic components of a SNOM. In collection modes the position of the detector and of the source is interchanged with respect to their position in the illumination modes

In **illumination mode** (the most common) the surface is concerned by the near field and the resulting scattered light is collected "in the far field" (either in transmission or reflection)

In **collection mode** the surface is illuminated by a propagating (conventional) field and the resulting scattered light is collected in the near field by the probe

Sub-diffraction space resolution is due to the non-propagating character of the near-field (typ resolution comparable to the aperture size, i.e., tens of nanometers)

Near-field "conversion" to far-field

7.2.1 Ray Optics of a SNOM

SNOM techniques differ mainly in the types of probes which are used and also by their ray optical components, i.e., their optical scheme, which is useful for a classification of most types of SNOM concepts. As shown schematically in Fig. 7.4, three different regions where the rays propagate can be distinguished: I) the body of the probe, II) the outside and III) the substrate of the object. In general, regions I and III will have a higher refractive index than the outside region II. Different angular domains of rays propagating in regions I and III exist, which can be distinguished by the criterion of total reflection of a ray falling on its boundary or of it being partially refracted into the outside II. Thus, in the case of a transparent substrate III, we distinguish between the angular domain III_1 of angles ε with $-\varepsilon_t < \varepsilon < \varepsilon_t$, where ε_t is the critical angle of total reflection ($\varepsilon_t = 41.5°$ for glass of refractive index 1.5) and the angular domain III_2 with $90° > \varepsilon > \varepsilon_t$ or $-90° < \varepsilon < -\varepsilon_t$, which is sometimes called the range of forbidden light. Rays of the domain III_2 are totally reflected in the substrate, whereas rays of domain III_1 are partially refracted into the outside II. Also within the body of the tip I two different domains may be distinguished (Fig. 7.4). This figure only shows the case of a rectangular wedge, a two-dimensional analog of the three-dimensional body of the tip. For such a wedge, with a refractive index $n = 1.5$, rays entering at an angle within the angular domain ($-3.5° < \varepsilon < +3.5°$, region I_2) will be totally reflected back into a reflected ray of the same angle ε. Rays entering the wedge at different angles will also be reflected into the same angle and be partially refracted into the outside II of the wedge. This situation also applies, if the wedge is coated with a partially transparent metal film, as is typical for SNOM probes. Similar considerations also apply for a three-dimensional tip.

In summary, in many cases it is possible to distinguish in regions I and III between angular domains I_1 and III_1 where total reflection of the rays occurs into the same domain and the domains I_2 and III_2 from where light is partially refracted to the outside II.

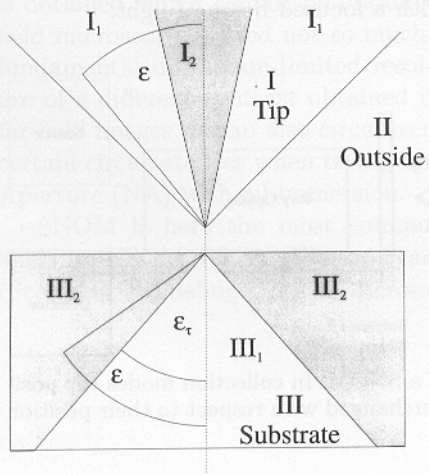


Fig. 7.4. Ray optics of a SNOM. One can distinguish between three different regions where the rays propagate; the body of the probe I, the outside II and the substrate III of the object. Different angular domains of rays propagating in regions III and I at an angle ε can be distinguished by the criterion of a ray falling on its boundary being totally reflected into the same domain (III_2, I_2) or being partially refracted into the outside II (III_1, I_1)

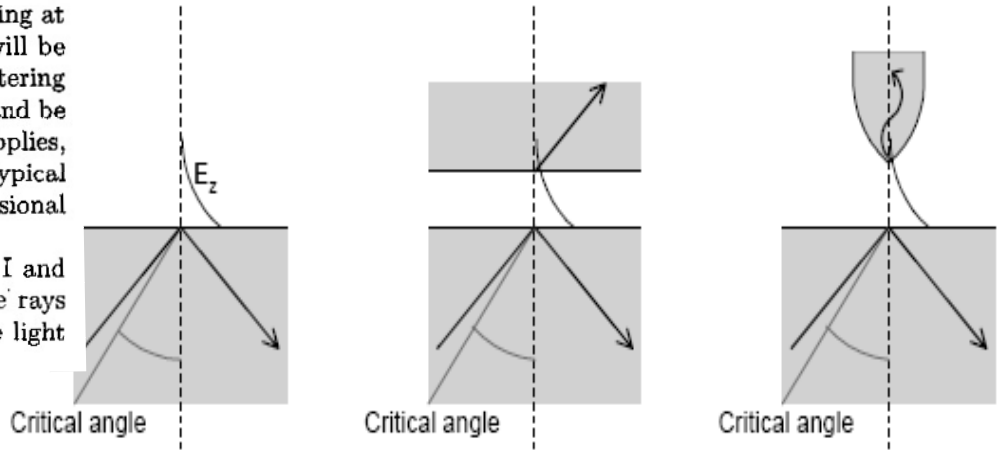


Figure 9: the left panel shows the penetration of an electromagnetic field in the less dense medium when total internal reflection occurs. The middle and the right panel schematically envision the energy flow between two dense media through an air gap (frustrated total internal reflection).

SNOM probe can "convert" near-field into far-field (especially relevant in collection mode)

http://www.chem.vu.nl/~sneppen/literaturereport.pdf

Da Wiesendanger Ed., Scanning Probe Microscopies (Springer, 1998)

A very few words on apertureless SNOM

3.5 APERTURELESS NEAR-FIELD SPECTROSCOPY AND MICROSCOPY

As mentioned in Section 3.3, an emerging approach is the apertureless near-field spectroscopy and microscopy (Novotny et al., 1998; Sanchez et al., 1999; Bouhelier et al., 2003). The use of an aperture such as a tapered fiber opening poses a number of experimental limitations. Some of these are:

- Low light throughput due to the small fiber aperture and the finite skin depth (light penetration) into the aluminum metal coating around the tapered fiber.
- Absorption of light in the metal coating; this can produce significant heating that can create a problem in imaging, particularly of biological samples.
- Pulse broadening in the fiber, when using short pulses for nonlinear optical studies. Also, the fiber tip may be damaged by the high peak intensity as already discussed in Section 3.3.

The apertureless approach overcomes these limitations, at the same time providing a significantly improved resolution. It has been demonstrated by Novotny, Xie, and co-workers (Sanchez et al., 1999; Hartschuh et al., 2003; Bouhelier et al., 2003) that optical images and spectra of nanodomains ≤25 nm can be obtained using the aperturel m.

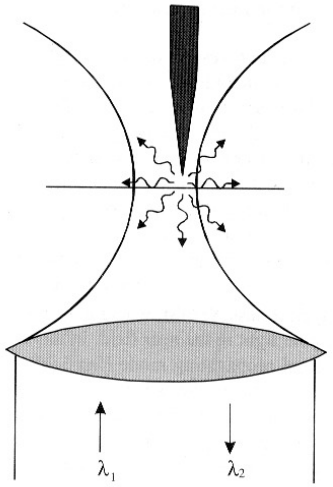


Figure 3.22. Metallic tip enhancing the local field by interacting with the focused beam at λ_1. The optical response at another wavelength λ_2 is collected by the same objective lens.

The two approaches used for apertureless NSOM are:

1. Scattering type, which involves nanoscopic localization and field enhancement of the electromagnetic radiation by scattering of the light from a metallic nanostructure. An example is provided by Figure 3.2 where the light is scattered by a sharp metallic tip. Scattering and field localization can also be produced by a metallic nanoparticle within nanometers of distance from the sample surface. The localization and enhancement of electromagnetic field by plasmon coupling to a metallic nanoparticle is discussed in Chapter 5 under "Plasmonics." This principle of obtaining nanoscopic resolution using scattering from a metallic nanoparticle also forms the basis of "plasmonic printing," discussed in Chapter 11 on "Nanolithography".
2. Field-enhancing apertureless NSOM, where a metallic tip is used to enhance the field of an incident light in the near field. In this case, the light is incident on the tip as a normal propagating mode (far-field). The strongly enhanced electric field at the metal tip produces nanoscopic localization of optical excitation. This approach offers simplicity and versatility of using light by just focusing on the metallic tip through a high-numerical-aperture lens. Hence it is described here in detail, with examples of some recent studies utilizing this approach.

A nanoparticle, or a nanosized tip, irradiated by a propagating field, acts as a quasi-pointlike source of the near field

Advantages of apertureless sources

Apertureless SNOM

In addition to methods based upon optical aperture probes, it is also possible to excite near-field phenomena in the absence of an aperture by using a metal tip as the probe. There has recently been a growth of interest in these so-called aperture-less SNOM techniques, in the hope of delivering superior resolution to that provided by aperture probes. They rely upon the fact that when a nanoscale metallic asperity is held in close proximity to a surface and illuminated with polarised light, the electric field associated with the excitation may be significantly enhanced in a small region directly beneath the tip.[5] The magnitude of the enhancement in the field strength may be very large—several orders of magnitude under optimal conditions. The phenomenon has become known rather evocatively as "the lightning-rod effect". The field enhance-ment is associated with scattering from the tip, and consists mainly of non-propagating (evanescent) components. The advantages of apertureless approaches are that they deliver significantly enhanced resolution compared to techniques based on aperture probes and, moreover, tips are significantly easier to fabricate than optical fibre probes. Fig. 3 shows the extent of the resolution improvement that is possible in principle by comparing the spatial variation in the electric field

Advantages:
Better space resolution (below 10 nm??)
No, or negligible, throughput limitations

Disadvantages:
Cumbersome operation
Stray light (in the far field)

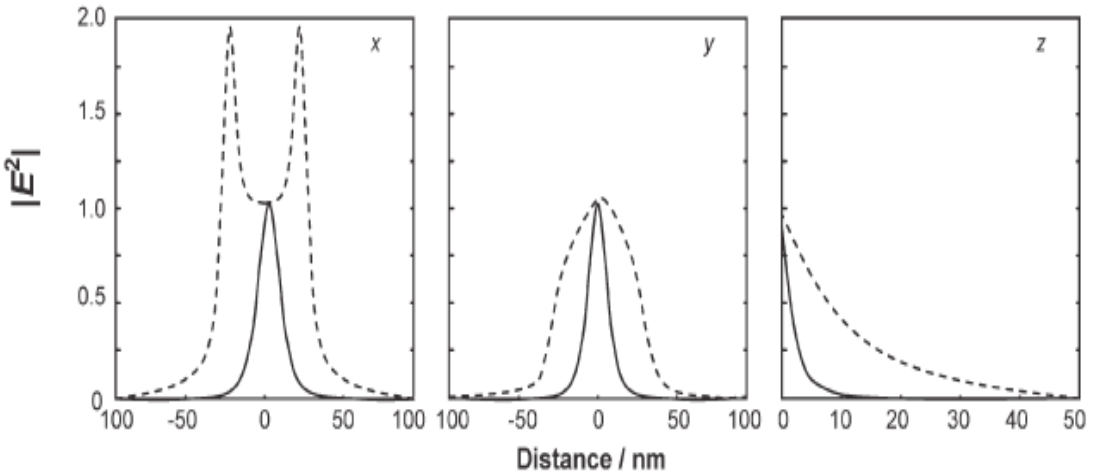


Fig. 3 Decay of the electric field intensity $|E^2|$ along the x, y and z directions (z is parallel to the probe axis) for an aperture probe (dashed line) and a silver plasmon probe excited aperturelessly (solid line). (Reproduced with permission from ref. 5. Copyright 1995 Elsevier.)

3.B What can be measured by SNOM

The effects of the **local interaction between the sample surface** (i.e., a layer with thickness comparable to the near field range) **and the near field** photons are recorded

They can be regarded as analogous (but for the sub-diffraction resolution and the surface origin) of conventional **optical transmission and/or reflection** measurements (depending whether the sample is transparent or opaque)

Non propagating behavior of the exciting near field can however play a role (for instance, specific polarization can give access to otherwise forbidden transitions, …)

Examples

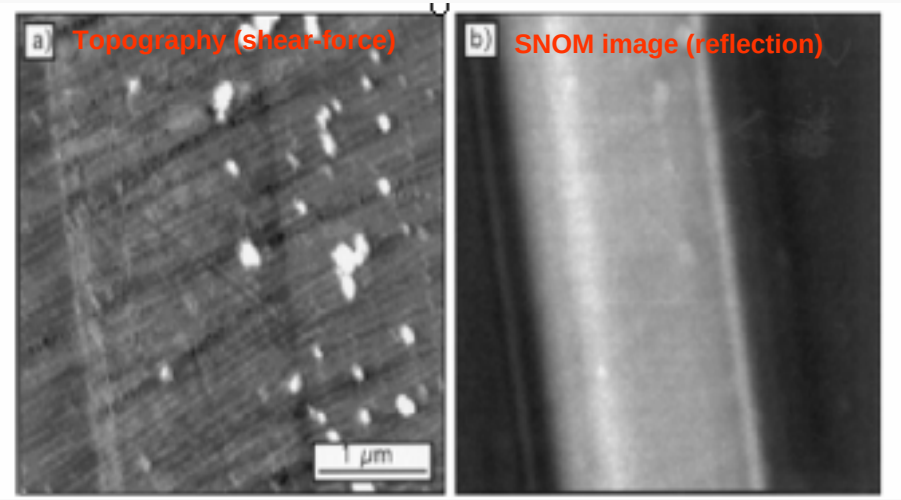
- ✓ **Local variations of the "refractive index"** can be derived by analyzing the scattered radiation
- ✓ In case of emitting (photoluminescent) samples, fluorescence can be excited by the near field, and **photoluminescence** maps can be acquired
- ✓ By implementing a polarization control system *(see later on)*, **optical activity** of the sample (e.g., dichroism, birifringence) can be analyzed at the sub-diffraction level
- ✓ More advanced spectroscopy (e.g., Raman) can be carried out

Examples

Collection mode can be used to map emission of, e.g., electroluminescent devices
Also, evanescent radiation, e.g., stemming from a waveguide surface, can be mapped

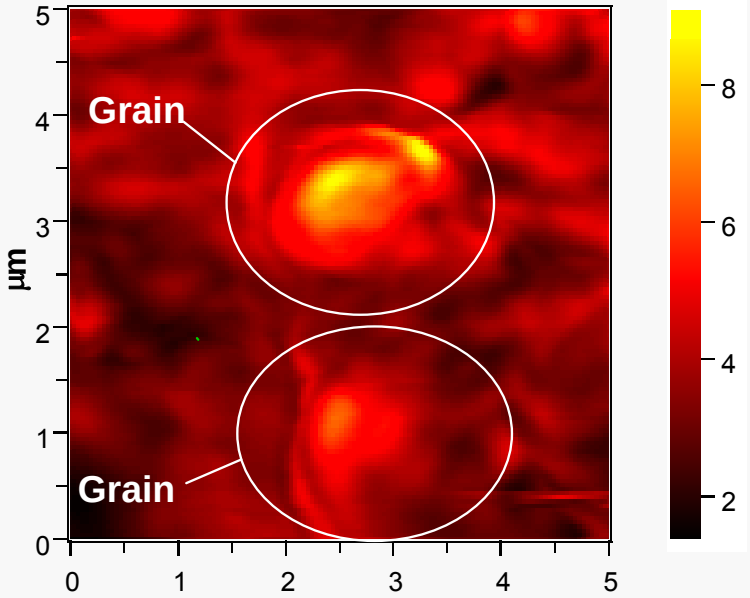
In addition, the tip/sample distance control, being based on a feedback system, allows acquisition of topography maps simultaneously with every SNOM scan (with a space resolution in the tens nm range)
Morphological and optical information acquired and compared at a glance!

Examples of SNOM in reflection mode I



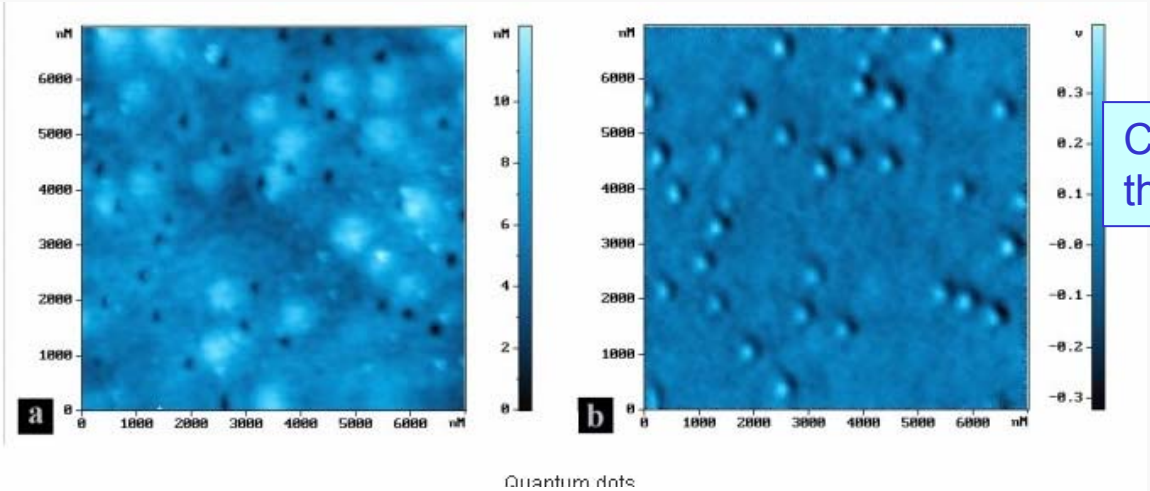
NiFe stripes embedded in an alumina matrix

YBCO/YSZ/NiFe multilayer deposited by PLD



Compositional differences (e.g., two-phase materials or structural fluctuations) associated with variations in the optical properties (i.e., refractive index) are evidenced in SNOM with a sub-diffraction space resolution

Examples of SNOM in reflection mode II



Quantum dots

Shear Force (topography) (a) and reflection (b) images of In-Ga quantum dots made with the use of He-Cd 442 nm laser.

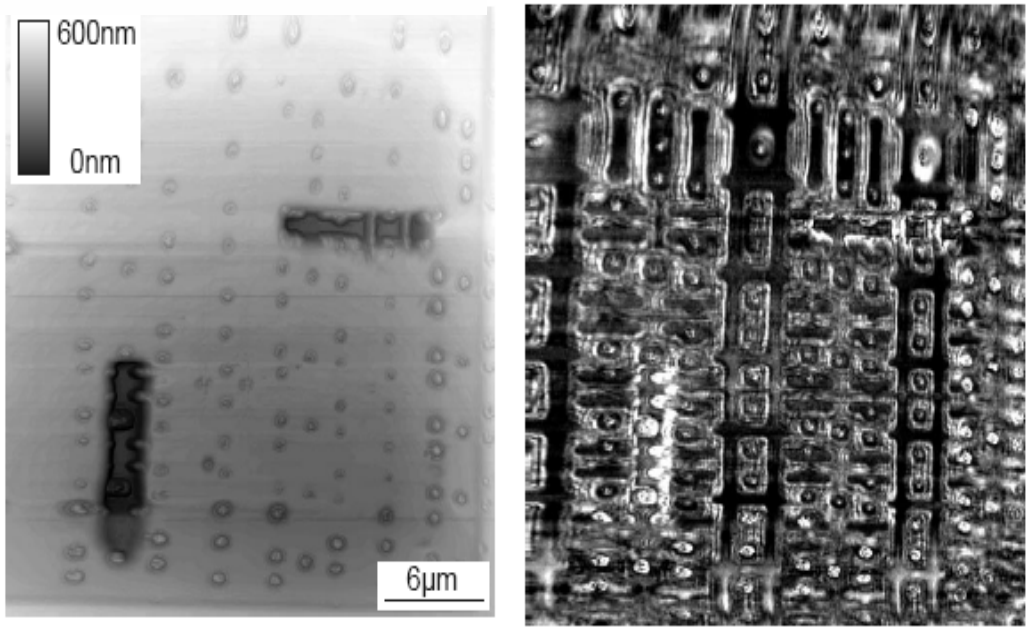
Images courtesy of Igor Dushkin, NT-MDT.

Scan size: 7x7 μm

Source MDT-file: download (515.06 Kb)

Contrast enhancement due to the use of optical signals

Buried structures can be detected (when using illumination light at a wavelength transmitted by the upper layers of the device)



g. 2. Topography (left) and oblique reflection (right) micrographs of aluminium structures buried underneath interlayer dielectrics.

Examples of SNOM photoluminescence

Saiki et al. (1998) and Matsuda et al. (2001) conducted room temperature photoluminescence study on a single quantum dot from InGaAs quantum dots grown on a GaAs substrate. Their result is shown in Figure 3.12. Because of the spectral resolution obtained by sampling only a single quantum dot (no inhomogeneous broadening), they were able to observe, at an appropriate excitation density, emission not only from the lowest level (subband) of the conduction band but also from higher levels. (See Chapter 4 for a description of these bands.) They were able to study the homogeneous line width, determined by the dephasing time of excitation (see Chapter 6 for a description of dephasing time), as a function of the interlevel spacing energy. They found that the line width was larger for a smaller-size quantum dot for which the interlevel spacing is larger. (This is predicted by a simple particle in a box model as the length of the box becomes smaller, see Chapters 2 and 4.)

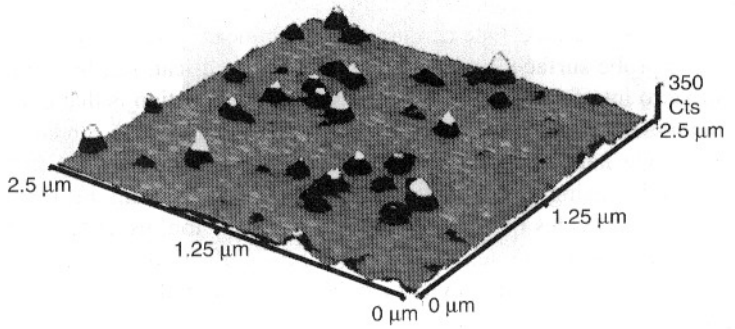


Figure 3.13. Fluorescence NSOM images of single molecules. From Professor D. Higgins and Professor P. Barbara, unpublished results.

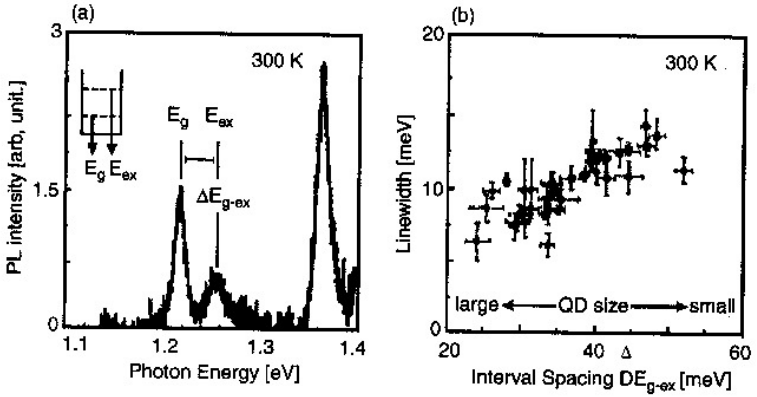


Figure 3.12. Photoluminescence spectrum of single QD at room temperature (a), and dependence of the homogeneous linewidth of ground-state emission on interval spacing, which is closely related to size of Qd's (b). From Saiki and Narita (2002), reproduced with permission.

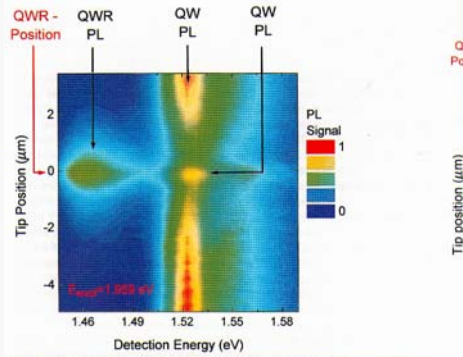


Fig. 7.23. Spatially resolved near-field photoluminescence spectrum of the quantum wire heterostructure. The spectrum was recorded for spatially resolved excitation of the sample at 1.959 eV. The tip was scanned along the lateral direction perpendicular to the wire. The PL intensity (in arbitrary units) is plotted as a function of tip position and detection energy. The color red corresponds to high, while purple corresponds to low PL intensity. The quantum wire emission is centered around 1.46 eV. Note that, in addition to the flat quantum well luminescence at 1.522 eV, a further, slightly blue shifted, sidewall quantum well emission is resolved

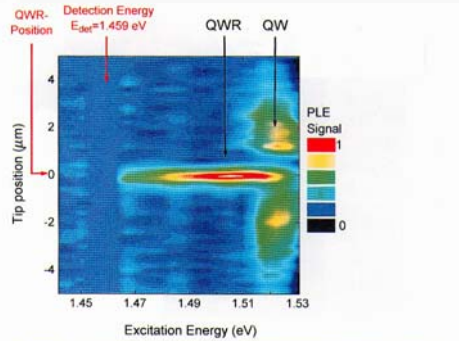


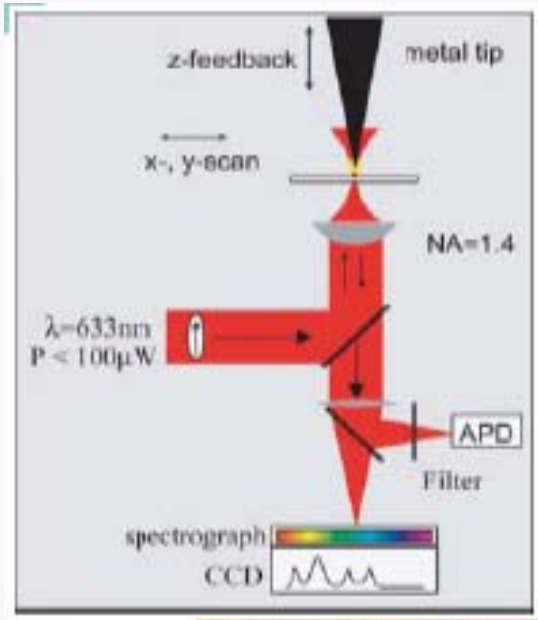
Fig. 7.24. Spatially resolved near-field photoluminescence excitation spectrum. The PLE signal is plotted as a function of tip position and excitation energy. The scan direction is perpendicular to the wire. The photoluminescence was detected at 1.46 eV. The color red corresponds to high and purple to low intensity

Single, isolated nanostructures, nanoparticles, or emitting molecules can be analyzed in spectral terms

Da Wiesendanger Ed., Scanning Probe Microscopies (Springer, 1998)

SERS and TERS microscopies

Light scattering from a nanoparticle or a tip (apertureless SNOM) is used to locally enhance the field distribution allowing reliable acquisition of (weak) Raman signals



La SERS (surface enhanced raman scattering), abbinata a una sonda di dimensioni nanometriche, permette lo studio di spettri vibrazionali localizzati.

Inoltre, il segnale Raman è molto debole (x-section inferiori di 10^14 rispetto alla fluorescenza) e necessita di tecniche particolari per essere rilevato. Sin dal 1974 si usa la SERS, amplificazione dell'effetto Raman dovuto a risonanze plasmoniche e "lightning rod effect" dovuto alla geometria delle superfici nanostrutturate. La TERS (tip enhanched...) permette di combinare questi effetti all'amplificazione del campo che si ha in prossimità di una punta NOM (stimato almeno di un fattore 4).

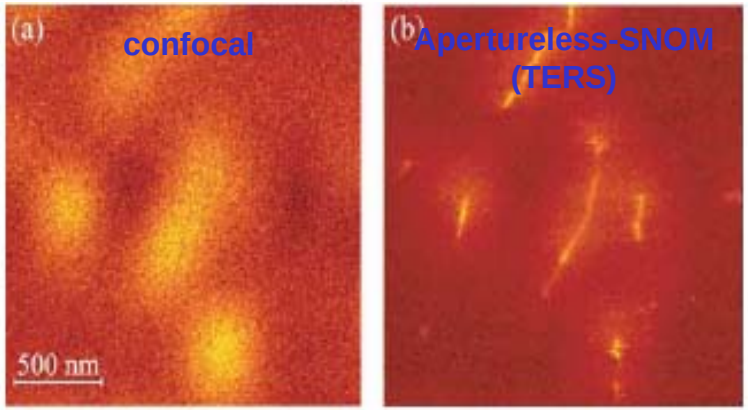
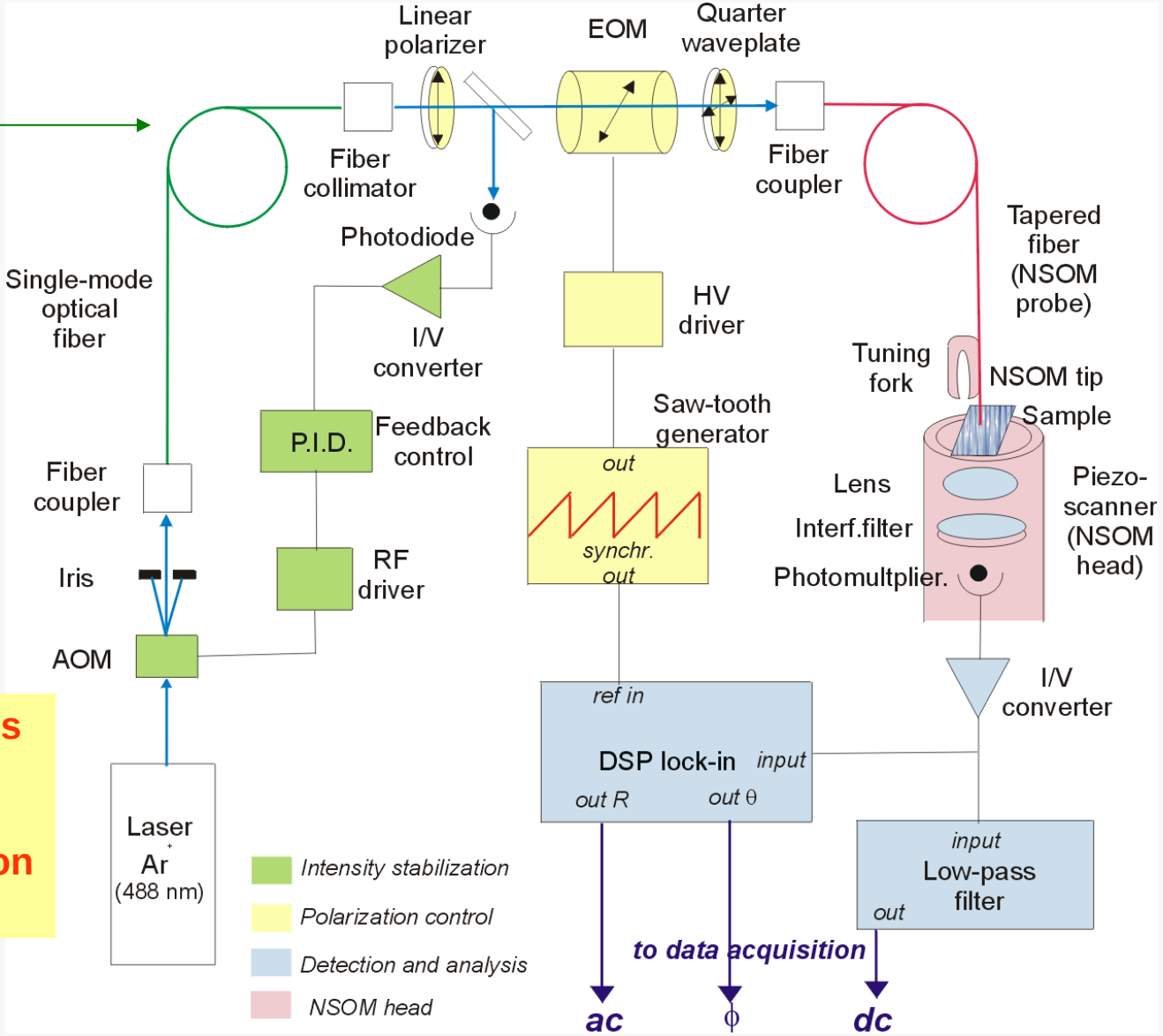


Fig. 3 Raman scattering images of SWNTs. (a) Confocal Raman scattering image integrated over the G' Raman band (~2600 cm^{-1}). (b) Raman scattering image taken over the same scan area with a sharp Au tip placed in the laser focus. The integration time per image pixel is ~10 ms (256 x 256 pixels).

Spectroscopy for advanced characterization of materials can be carried out at the very local scale

SNOM with polarization modulation (PM-SNOM) I

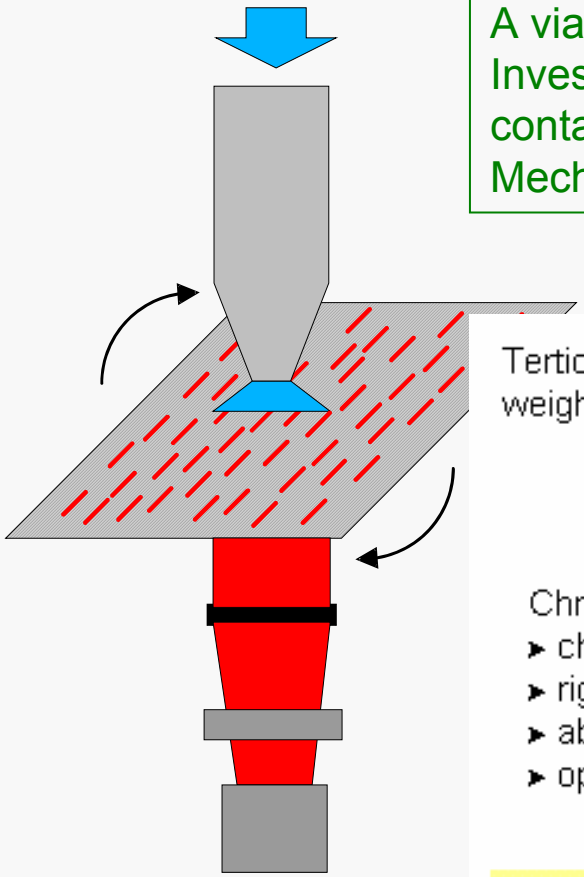
A system is used to control the polarization of the laser light entering the fiber probe: polarization is linear but continuously rotating (at a frequency typ. in the kHz range)



Optical activity of samples (e.g., dichroism, birifringence) can be anlyzed with sub-diffraction space resolution

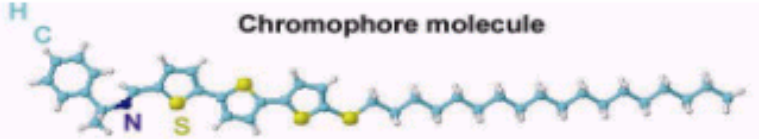
Michele Alderighi, Tesi di Laurea in Scienza dei Materiali, Pisa 2003

PM-SNOM II



A viable example of application:
Investigation of host-guest systems (high density PET matrix containing a dispersion of molecular chromophore)
Mechanical stretching of the polymer film leads to ***linear dichroism***

Tertiophene-like chromophore dispersed (3% wt) in a ultra-high molecular weight polyethilene (UHMWP)

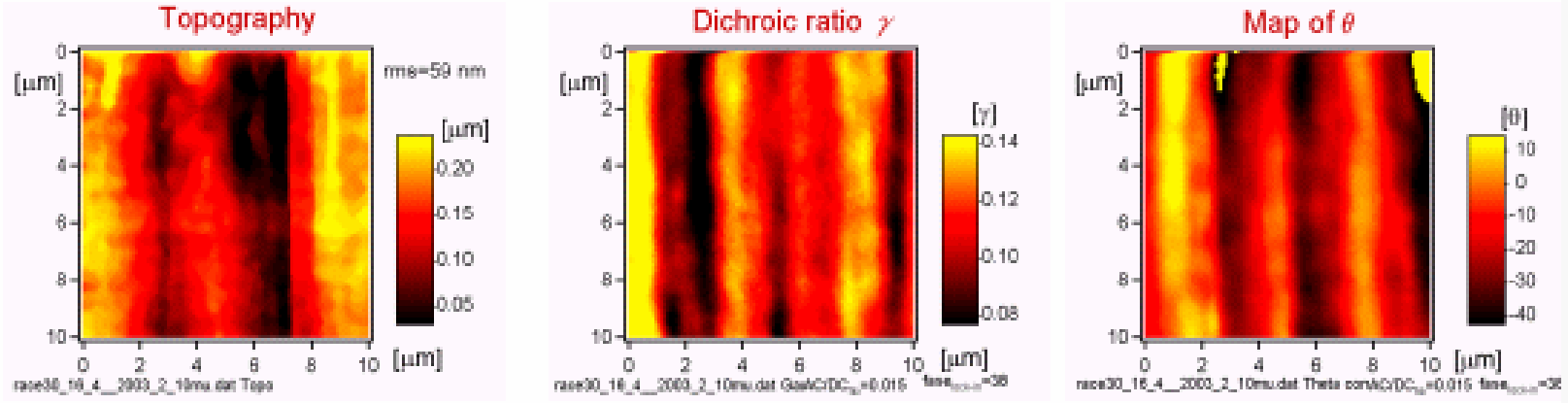


Chromophore choice:
- chemical compatibility with the host;
- rigid structure;
- absorption in the blue (peaked around 400 nm);
- optical activity (right- and left-handed isomers available).

Fabrication process (at *Dipartimento di Chimica e Chimica Industriale, Università di Pisa, group of Prof. Ciardelli and Prof. Ruggeri*):
- cast (to obtain pristine films) followed by
- high temperature drawing (to obtain films with stretched polymer chains)

PM-SNOM III

Sample: L-R (50:50) stretched (drawing ratio 30), 10x10 µm scan



Topographical variations due to local changes in polymer chain stretching

Local optical activity strongly affected by (host) stretching

Chromophore molecules follow host molecule alignment
Macroscopic dichroism (γ~0.1 in this sample) due to the combined effect of elongated islands with locally inhomogeneous optical activity

Information on the sample properties (and suggestions to improve the fabrication process) can be found which are masked in conventional (macroscopic) polarimetry

4. “Lithography” with SPMs

“As usual”, (scanning probe) microscopy is “accompanied” by (scanning probe) “lithography”

In this case, the term “lithography” can be better replaced by “**manipulation**”, since position and/or local properties of the matter can be modified, i.e., manipulated, by SPMs

Indeed, as we will see with several examples, matter can be **manipulated at the nanometer or sub-nanometer level** by using SPMs in order to fabricate nanostructures or prototypal nanodevices (*the latter we will see in the following)*

This kind of “lithography” is obviously *serial*, due to the scanning nature of SPMs, and not suited for industrial environment, though attempts are being made to get some *parallel* character (e.g., by using several probes in parallel)

Wide variety of methods can be envisioned: we will mention here in the following a few of them, based on AFM, STM and SNOM

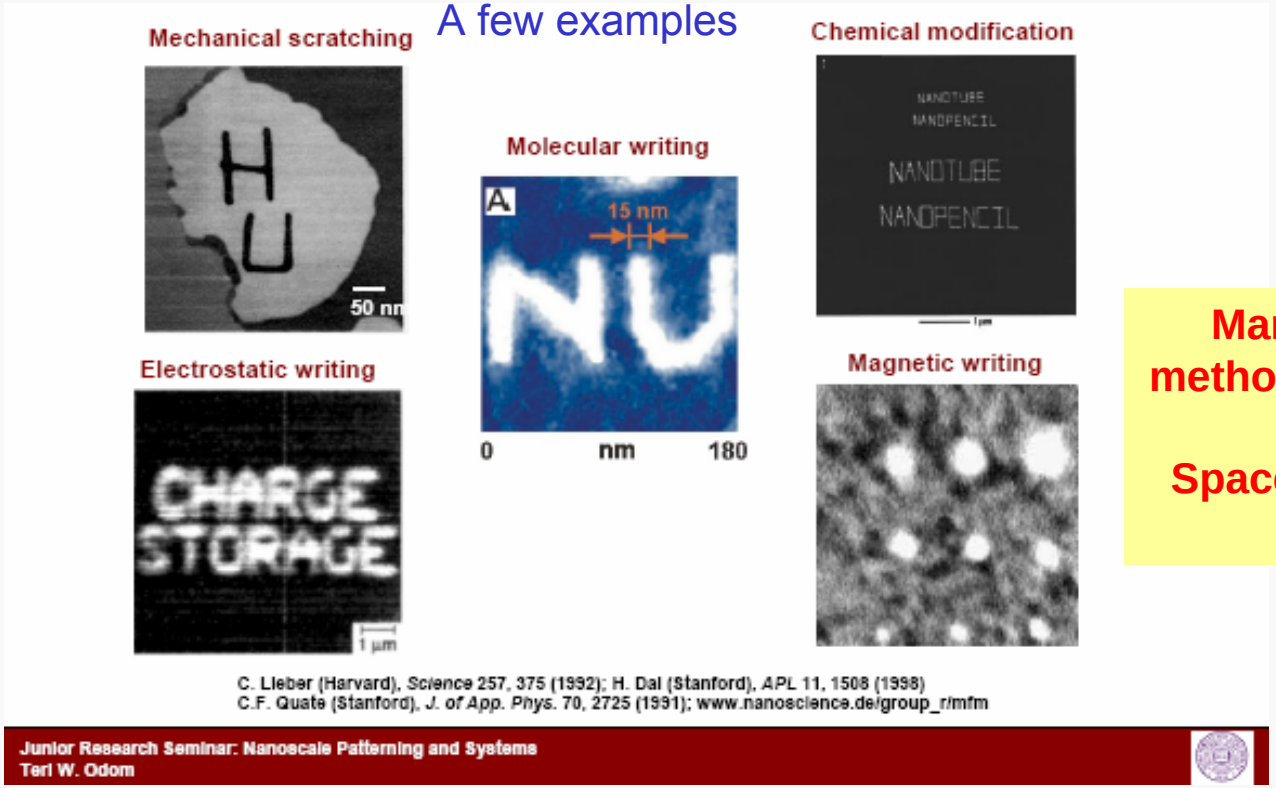
Their relevance is in fabrication, in surface modifications, in nanowriting for **data storage**

Examples of surface nanomodifications

Roughly speaking, we can for instance:

- ✓ Modify the surface, e.g. by mechanical scratching (nanoindenation) or by electrical means (STM-based);
- ✓ Induce “chemical” modifications on the surface (mostly STM-based), including resists impression (lithography);
- ✓ Manipulate single atoms or nanoparticles on the surface (AFM and STM-based);
- ✓ “Write” patterns on a material surface for data storage purposes (mostly SNOM-based)

A few examples



C. Lieber (Harvard), *Science* 257, 375 (1992); H. Dai (Stanford), *APL* 11, 1508 (1998)
C.F. Quate (Stanford), *J. of App. Phys.* 70, 2725 (1991); www.nanoscience.de/group_r/mfm

Many different tools and methods have been proposed!

Space resolution is typically excellent

4.A. Nanoindentation and AFM

Introduction to nanoindentation

http://www.nanoindentation.cornell.edu/

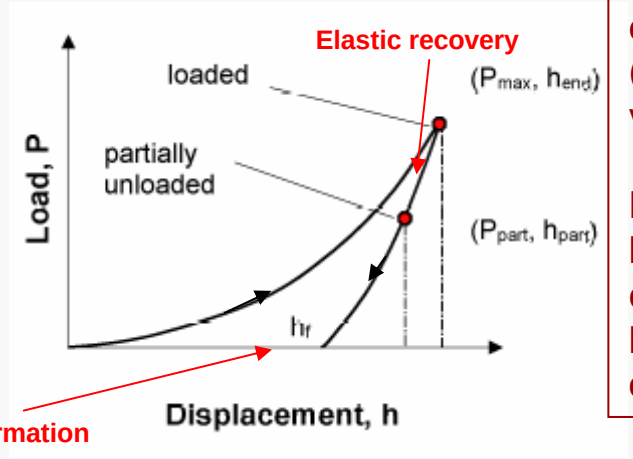
Indentation tests are perhaps the most commonly applied means of testing the mechanical properties of materials. In such a test, a hard tip, typically a diamond, is pressed into the sample with a known load. After some time, the load is removed. The area of the residual indentation in the sample is measured and the hardness, H, is defined as the maximum load, P, divided by the residual indentation area, A_r , or

$H = P/A_r$

The idea of nanoindentation arose from the realization that an indentation test is an excellent way to measure very small volumes of materials. In principle, if a very sharp tip is used, the contact area between the sample and the tip, and thus the volume of material that is tested, can be made arbitrarily small. The only problem is determining the indentation area. It is easy to make an indentation that is so small that it is difficult to see without a powerful microscope.

To solve this problem depth sensing indentation methods were developed. In this method, the load and displacement of the indenter are recorded during the indentation process and these data are analyzed to obtain the contact area, and thereby mechanical properties, without having to see the indentations.

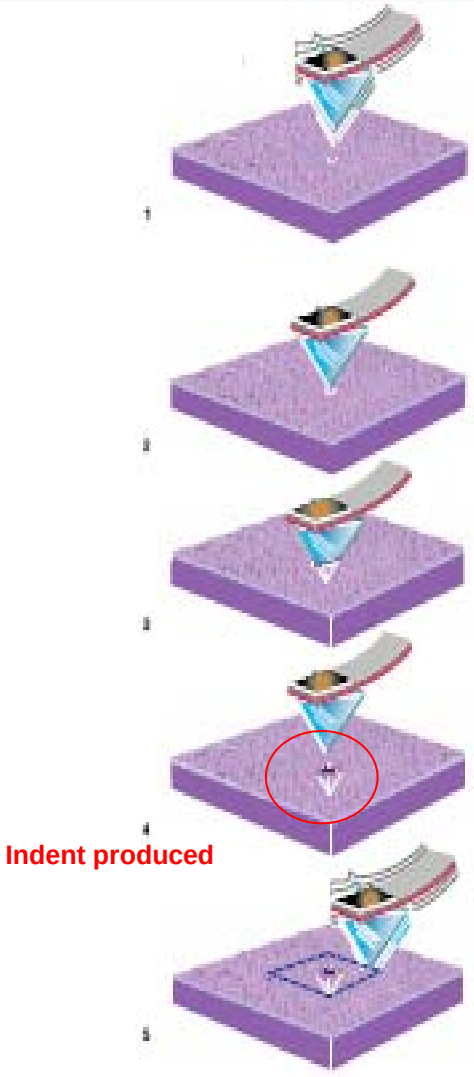
Nanoindentation refers to depth-sensing indentation testing in the submicrometer range and has been made possible by the development of 1) machines that can make such tiny indentations while recording load and displacement with very high accuracy and precision, and 2) analysis models by which the load displacement data can be interpreted to obtain hardness, modulus, and other mechanical properties.



(Nano)indentation is a common technique to ascertain elastic/plastic behavior of the materials
(if carried out with a load modulation, also surface viscoelasticity can be analyzed)

Data pertaining to the elastic modulus and to the plastic behavior (e.g., shear modulus) can be attained and comparison with macroscopic results (e.g., Vickers hardness, Rockwell,…) may lead to interesting insights on the microscopic nature of surfaces and nanostructures

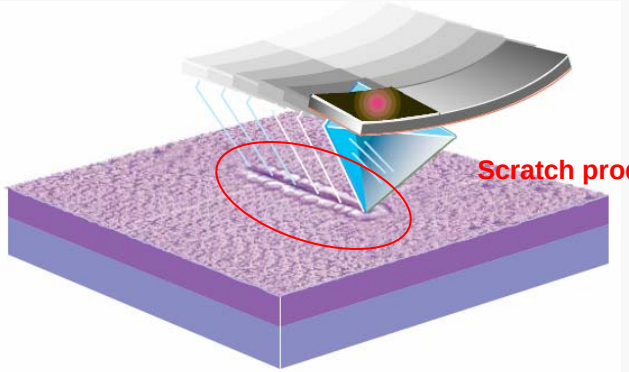
Surface modifications and scratching



225-2 Nanoindentation sequence: (1) engagement in TappingMode; (2-4) indentation; (5) imaging in TappingMode.

AFM tip can be used to produce a controlled indentation (note: typ force is in the nN range, but indentation area can be in the nm^2 range, hence pressures in the GPa range can be easily achieved – remember that max elastic modulus is ~ 1 TPa for diamond; shear moduli are typ in the 10 MPa range)

(Nano)scratches can be also obtained by scanning the tip (that involves nanotrobology issues)



225-3 Nanoscratching consists of moving a diamond tip through material at a specified depth and with a measured force.

Mechanical forces (pressures) in tip/surface interaction can be used to *permanently* modify the surface (nanoscribing on plastic materials)

http://www.veeco.com

More details on nanoindentation

Polymers can be plastically deformed

If *thermoplastics* are used (and tip is heated), deformations can be reversed

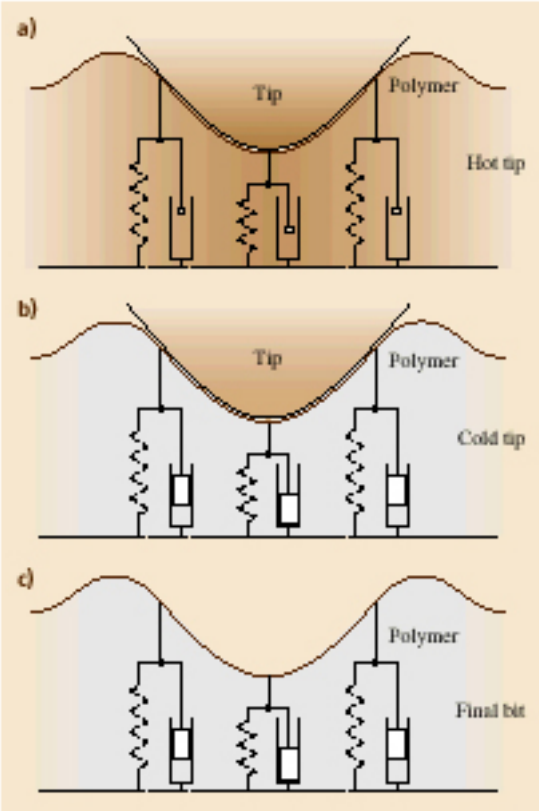


Fig. 31.18a–c Viscoelastic model of indentation writing. (**a**) The hot tip heats a small volume of polymer material to more than Θ_g. The shear modulus of the polymer drops drastically from GPa to MPa, which in turn allows the tip to indent the polymer. In response, elastic stress (represented as compression springs) builds up in the polymer. In addition, viscous forces (represented as pistons) associated with the relaxation time for the local deformation of molecular segments limit the indentation speed. (**b**) At the end of the writing process, the temperature is quenched on a microsecond time scale to room temperature: The stressed configuration of the polymer is frozen-in (represented by the locked pistons). (**c**) The final indentation corresponds to a metastable configuration. The original unstressed flat state of the polymer can be recovered by heating the indentation volume to more than Θ_g, which unlocks the compressed springs (after [31.15])

31.7 Polymer Medium

The polymer storage medium plays a crucial role in millipede-like thermomechanical storage systems. The thin-film multilayer structure with PMMA as active layer (see Fig. 31.2) is not the only possible choice, considering the almost unlimited range of polymer materials available. The ideal medium should be easily deformable for writing, yet indentations should be stable against tip wear and thermal degradation. Finally, one would also like to be able to erase and rewrite data repeatedly. In order to be able to address all important aspects properly, some understanding of the basic physical mechanism of thermomechanical writing and erasing is required.

31.7.1 Writing Mechanism

In a *gedanken* experiment we visualize writing of an indentation as the motion of a rigid body (the tip) in a viscous medium (the polymer melt). Let us initially assume that the polymer, i. e., PMMA, behaves like a simple liquid after it has been heated above the glass-transition temperature in a small volume around the tip. As viscous drag forces must not exceed the loading force applied to the tip during indentation, we can estimate an upper bound for the viscosity ζ of the polymer melt using Stokes's equation:

$$F = 6\pi\zeta\varrho v\,. \qquad (31.1)$$

In actual indentation formation, the tip loading force is on the order of $F = 50\,\mathrm{nN}$ and the radius of curvature at the apex of the tip is typically $\varrho = 20\,\mathrm{nm}$. Assuming a depth of the indentation of, say, $h = 50\,\mathrm{nm}$ and a heat pulse of $\tau_h = 10\,\mu\mathrm{s}$ duration, the mean velocity during indentation formation is on the order of $v = h/\tau_h = 5\,\mathrm{mm/s}$. Note that thermal relaxation times are on the order of microseconds [31.20, 21] and, hence, the heating time can be equated to the time it takes to form an indentation. With these parameters we obtain $\zeta < 25\,\mathrm{Pa\,s}$, whereas typical values for the shear viscosity of PMMA are at least seven orders of magnitude larger even at temperatures well above the glass-transition point [31.39].

This apparent contradiction can be resolved by considering that polymer properties are strongly dependent on the time scale of observation. At time scales on the order of 1 ms and below, entanglement motion is in effect frozen in and the PMMA molecules form a relatively static network. Deformation of the PMMA now proceeds by means of uncorrelated deformations of short molecular segments, rather than by a flow mechanism involving the coordinated motion of entire molecular chains. The price one has to pay is that elastic stress builds up in the molecular network as a result of the deformation (the polymer is in a so-called rubbery state). On the other hand, corresponding relaxation times are orders of magnitude smaller, giving rise to an effective viscosity at millipede time scales on the order of $10\,\mathrm{Pa\,s}$ [31.39], as required by our simple argument (see (31.1)). Note that, unlike normal viscosity, this high-frequency viscosity is basically independent of the detailed molecular structure of the PMMA, i. e., chain length, tacticity, polydispersity, etc. In fact, we can even expect that similar high-frequency viscous properties can be found in a large class of other polymer materials, which makes thermomechanical writing a rather robust process in terms of material selection.

Examples of indentation

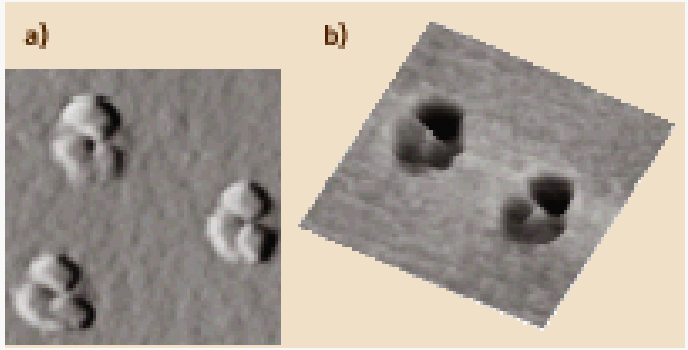


Fig. 31.20a,b Topographic image of individual indentations. (**a**) The region around the actual indentations clearly shows the threefold symmetry of the tip, here a three-sided pyramid. (**b**) The indentations themselves exhibit sharp edges, as can be seen from the inverted 3-D image. Image size is $2 \times 2\,\mu m^2$ (from [31.15] © 2002 IEEE)

Besides tribological and nanomechanical applications, nanoindentation can be envisioned as a nanofabrication tool or a data storage method

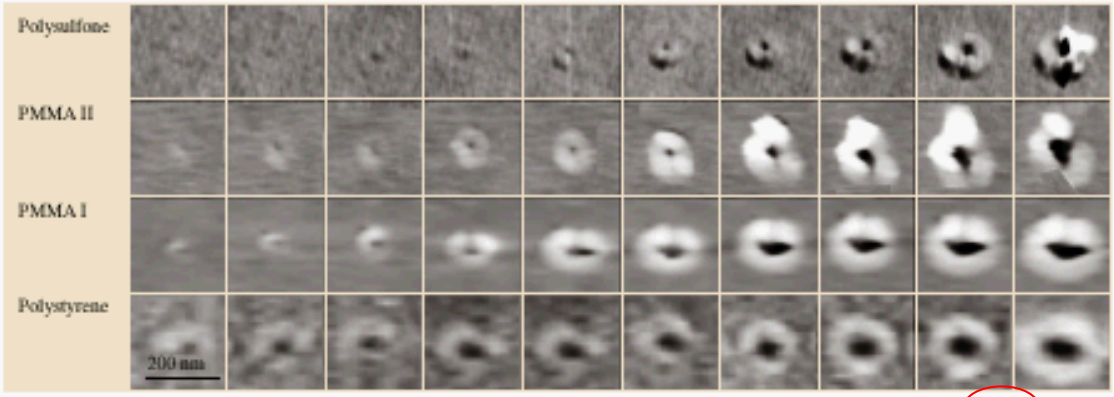


Fig. 31.21 Written indentations for different polymer materials. The heating pulse length was 10 µs, the load about 10 nN. The grey scale is the same for all images. The heater temperatures for the indentation on the left-hand side are 445, 400, 365, and 275 °C for the polymers Polysulfone, PMMA II (anionically polymerized PMMA, $M \approx 26$ k), PMMA I (Polymer Standard Service, Germany, $M \approx 500$ k), and Polystyrene, respectively. The temperature increase between events on the horizontal axis is 14, 22, 20, and 9 °C, respectively (from [31.15] © 2002 IEEE)

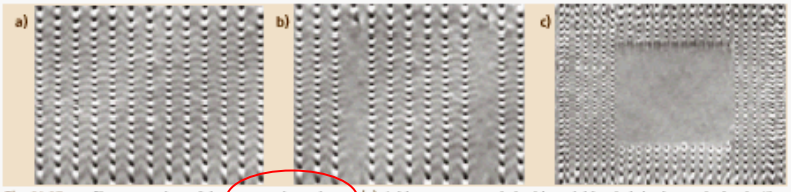


Fig. 31.25a–c Demonstration of the new erasing scheme. (**a**) A bit pattern recorded with variable pitch in the vertical axis (fast scan axis) and constant pitch in the horizontal direction (slow scan axis) was prepared. (**b**) Then two of the lines were erased by decreasing the pitch in the vertical direction by a factor of three, showing that the erasing scheme works for individual lines. One can also erase entire fields of indentations without destroying indentations at the edges of the fields. This is demonstrated in (**c**), where a field has been erased from an indentation field similar to the one shown in (**a**). The distance between the lines is 70 nm (from [31.15] © 2002 IEEE)

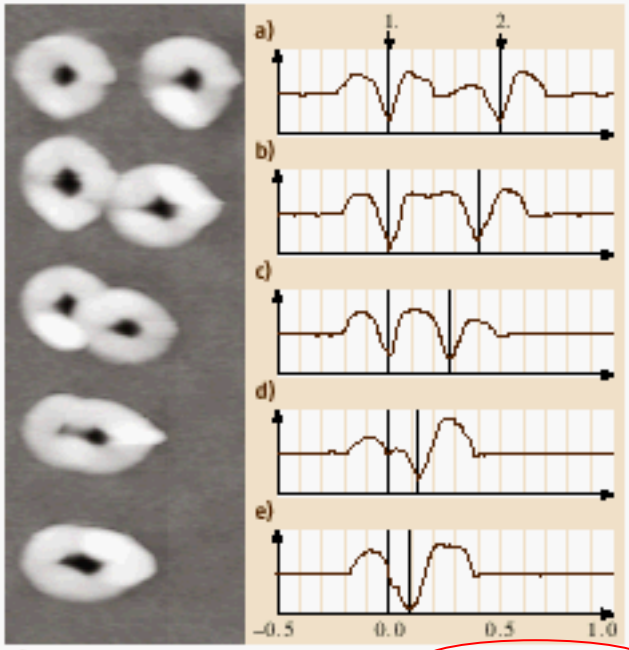


Fig. 31.24a–e Indentations in a PMMA film at several distances. The depth of the indentations is ~ 15 nm, roughly the same as the thickness of the PMMA layer. The indentations on the left-hand side were written first, then a second series of indentations was made with decreasing distance from the first series going from (**a**) to (**e**) (after [31.15])

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

Another example of nanoindentation and scratch

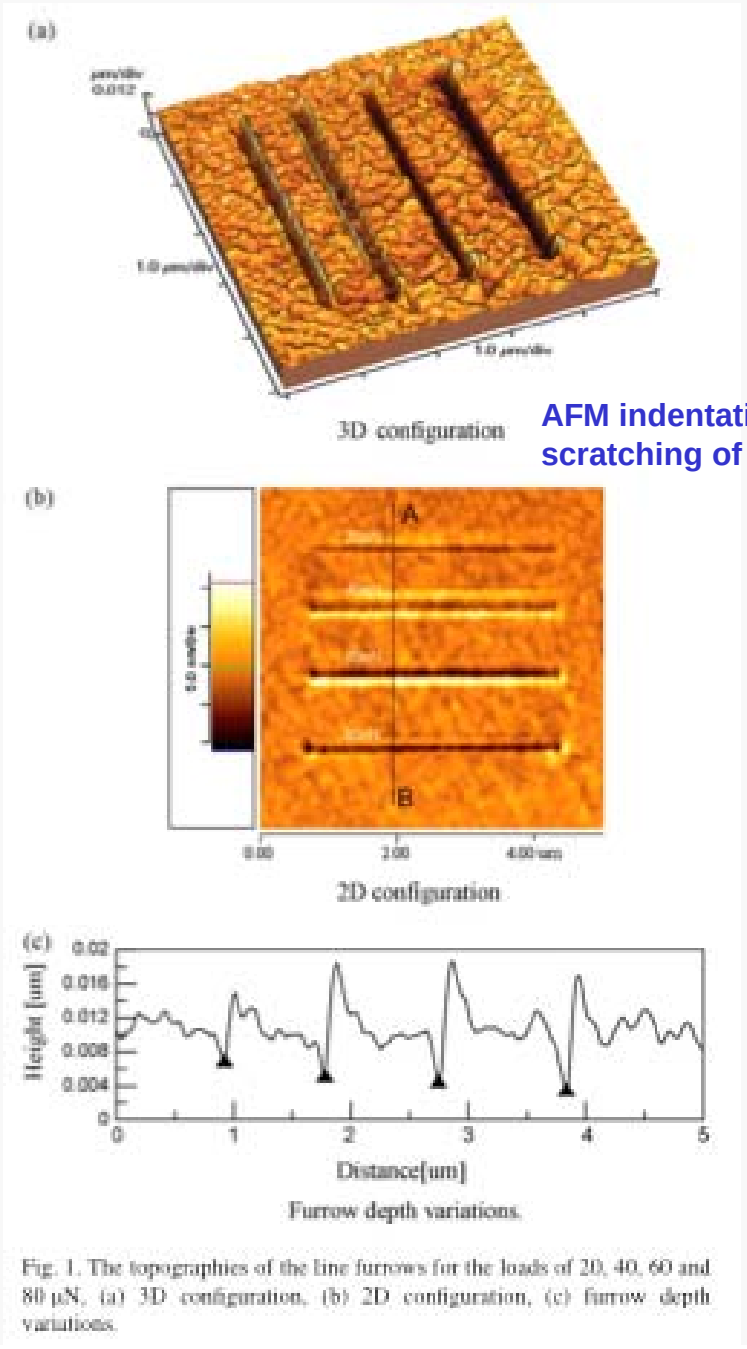


Fig. 1. The topographies of the line furrows for the loads of 20, 40, 60 and 80 μN. (a) 3D configuration, (b) 2D configuration, (c) furrow depth variations.

AFM indentation and scratching of polycarbonate

The highly controlled operating conditions of AFM (load, speed, tip position, etc.) opens the way to nanostructuring of (plastic) surfaces

Table 1
The working parameters of the scribing experiments

Properties	
Resonant frequency	325 kHz
Force constant	40 Nm^{-1}
Cantilever length	125 μm
Cantilever thickness	4 μm
Cantilever width	35 μm
Tip material	Silicon
Tip radius	10 nm
Tip height	15 μm
Workpiece material	Polycarbonate
Workpiece roughness, Ra	1.1 nm

Te-Hua Fang et al., Microelectr. J. 36, 55 (2005)

4.B. Nanomanipulation by SPM

The excellent space resolution offered by SPM can be exploited also to:
- Manipulate (spatially) nanoparticles on a surface;
- Produce very local (chemical) modifications of the surface;
- Induce chemical reactions on the surface involving additional “reactants”

5 Manipulation of Atoms and Molecules

The scanning probe microscopes do not only have the ability to image individual atoms. The interaction needed for imaging the surfaces can also be used to manipulate individual adatoms, molecules, or the surface structure itself on the atomic scale. Indeed a large number of works concentrated on the manipulation of individual atoms and in the following novel nanostructures were built. Here we briefly show the work by Eigler and coworkers [44], [45] as examples, followed by a more subtle tip-induced manipulation of atoms, and the tip-induced migration of defects by tip-induced excitement of defects [46], [47]. More recently the group of Rieder could even perform full chemical reactions with single molecules [48]. Three different manipulation modes can by distinguished: the lateral and the vertical manipulation as well as the tunnel current induced changes. The combination of all three modes enables to achieve tip controlled chemical reactions.

The same instrument can be used to produce and assess the manipulation

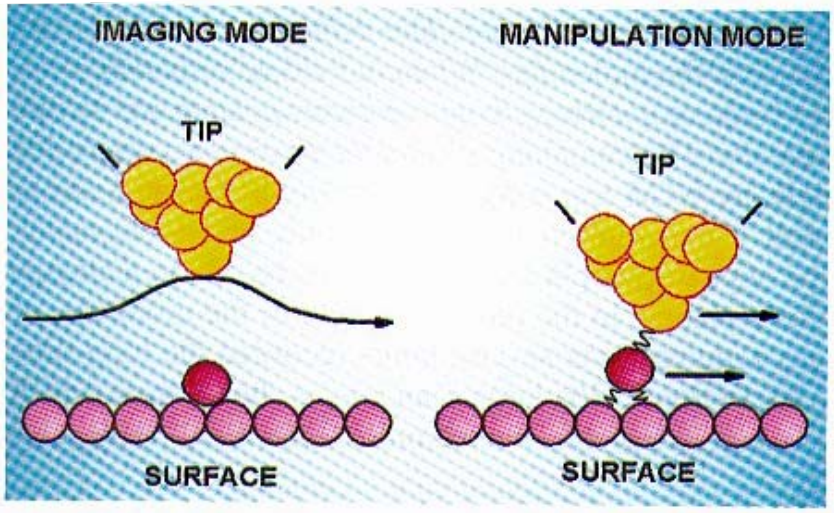


Figure 32: In the STM imaging mode the tunnel current is kept constant and the cantilever is raised. For manipulation the tip is lowered above an atom dragging it to the desired position, lifting the cantilever losing interaction with the atom [44].

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

STM and nanofabrication

In the STM device, the tip is so close to the target as to make the electron current highly spatially confined - it is the key to the extremely high spatial resolution of STM images. The key element in surface modifications is also the STM tip. To affect these modifications one utilizes a variety of tip-sample interactions, including attractive and repulsive forces, electric fields, and the effect of highly spatially confined electron currents. The small distance between tip and sample, which is about one nanometer, causes electrons to tunnel to (or from) a region on the sample that is approximately one nanometer in diameter, with an even smaller major distribution area. Thus, the surface fabrication produced by STM must be performed on the nanometer scale, i.e., STM can nanofabricate. Later, we will see that it is also possible to manipulate a single atom or molecule adsorbed on the surface with STM.

Since the invention of STM in 1981, as a nanofabrication tool it has been used in direct surface identation, electron-beam-induced deposition, etching, single-atom manipulation, and so on. All of these techniques have a wide-spread application potential. First, it is possible to reduce the linewidth of large-scale integrated circuits from the micrometer scale to the nanometer scale by lithography, beam-induced deposition and etching, which is one of the goals of high technology. In most cases the resulting feature with dimensions on the order of hundred nm [9.1], but features with dimensions of a few nm have also been achieved [9.2-4]. An exciting possibility will be to use the tip to "operate" on biomolecules such as DNA and proteins. The electronic properties of devices may be dominated by quantum-size effects when their size is reduced to the nanometer or atomic scale. By STM and other techniques, it is possible to discover new phenomena, design new devices and fabricate them. Next, STM can be utilized to repair masks and integrated circuits. The surface topographies can be imaged in situ during the surface fabrication process by STM, which makes it possible to discover defects in masks and circuits, to repair them by surface deposition and etching, and then to examine the final results by STM. Lastly, using the STM as a tool, the essential research on the growth, migration and diffusion of clusters on surfaces, and the interactions between small particles or between substrates and particles can be performed in order to manipulate clusters or atoms on purpose.

Lens-focused electron beams, ion beams and X-rays can also be employed in nanofabrication. Although the STM seems unlikely to become competitive in some areas of nanofabrication such as wafer-scale resist patterning, it has its own characteristics. First, an STM can work in either the tunneling mode or the field-emission mode. When working in the latter, a low applied voltage (higher than a few volts) can produce a strong enough electric field to make electrons emit from the tip over the barrier, because the distance between tip and sample is very small. These emitting electrons with a certain current and energy, will not diverge greatly because of the small separation which results in a nanometer beam diamter on the substrate surface. Unlike conventional high-energy electron lithography, the low-energy STM beam reduces the problems associated with electron backscattering and the generation of secondary electrons. A resolution of about 10 nm, and exposure rates comparable to those of conventional electron lithography have been achieved. Secondly, by moving the tip to contact the sample, the STM tip can also produce local contact forces and electrostatic forces in a small region on the sample surface to create indentations directly. Lastly, at present, STM is the only instrument that can provide a nm-sized beam of very low energy electrons ($0 \div 20$ eV). The importance of electrons with low energy is obvious when it is considered that many of the processes such as migration, bond breaking, chemical reactions that would be interesting to control, have activation energies less than 10 eV per atom which require a low-energy beam.

Nanomanipulation by STM: application of suitable voltages or voltage pulses can be used to "modify" (i.e., to damage) matter through field-emission from the tip

Resist-assisted lithography with STM I

9.3 Nanolithography on Resist Films

Direct writing of nanometer-scale features on organic materials opens the door to using film as a resist for the etching mask and the possibility of etching in air [9.30].

The resists can directly be exposed with conventionally focused electron beams to create various structures. The e-beam resist materials frequently utilized in conventional lithography are PolyMethylMethAcrylate (PMMA) and polydiacetylene with urethane substitutuents (P4BCMU). They are typically sensitive to low-energy (<20eV) electrons. The primary beam of a conventional e-beam lithography system has a considerably higher energy, so exposure of the resist occurs through interactions with secondary and backscattered electrons produced by the primary beam; a resist is consequently exposed over an area which is significantly larger than the primary beam spot size. The STM, even working in the field-emission mode, can supply focused electron beams with low energy which can interact with resists directly. Because the tip can be held within a few nanometers from the sample, which leads to an effective beam spot size on the sample of the order of the tip-sample separation, and the tip can transversely scan over the surface controlled precisely by a computer, STM can easily be employed in lithography for writing directly on the resist surface. The degradation in resolution because of interactions between the resist material and secondary electrons in conventional lithography can be overcome in STM lithography, which makes it possible to obtain more precise structures. Using STM it is also possible to make a thorough investigation of the exposure mechanism by controlling the bias voltage precisely (i.e., the energy of electrons with which the resists are exposed) in a certain time interval under constant current. In addition, STM generally ought to work in the field-emission mode because the electrons must have enough energy to induce a chemical reaction in resists (i.e., to expose resists). In this mode, a linear dependence of the tip-sample separation on the bias is expected in the absence of geometric effects, which make the widths of the features increase with the bias voltage.

In order to be successfully exposed with an STM, the resist film coated on conducting substrates such as Si, GaAs, Au, graphite and so on must be extremely thin, on the order of a few tens of nanometers, for two reasons. First, the low-energy electrons must be able to completely penetrate the film in order to properly expose it and to prevent excessive charging of the surface. Second, if the film thickness is greater than the gap between the tip and the conducting substrate, the tip will penetrate and damage the resist film. Experimentally, the thickest film that can be used is V nm, where V is the bias voltage in volts. There are different kinds of materials which can be used as resist films in STM lithography, including polymers such as PMMA and P4BCMU which were often employed in conventional e-beam lithography, metal halides such as GaF_2 and AlF_3, etc. The resist can be applied to the substrates by evaporation or deposited from a Langmuir-Blodgett (LB) film balance and spin coating.

The local electric field produced by the STM tip behaves similarly to the electron beam in EBL

⇩

Lithography can be accomplished only in a serial manner, but resolution can be excellent (without the need for complex EBL setups)

Resist-assisted lithography with STM II

By applying a liftoff process, McCord and Pease[7] used the STM to pattern a 20 nm thick polymethylmethacrylate (PMMA) film that acts as a mask for the subsequent deposition or etching process. As shown in Fig. 2(a), a 22 nm wide and 12 nm thick gold–palladium (Au–Pd) line was deposited on a Si substrate. The positive PMMA resist was exposed to a 100 pA; −20 V (negative polarity at tip) beam at a writing speed of 1 μm/s. In addition to PMMA, several other polymer resists have also been adopted by STM for patterning structures with a linewidth below 50 nm, including the Shipley Advanced Lithography 601 negative resist[8] and self-assembled monolayers (SAMs) of molecules.[9] The latter, after exposure and ligation of the catalyst, has also been used as a template for electroless plating, thus, producing a metallic etch mask for pattern transfer with a feature size of 15 nm.[10]

Both polarities can be used in resist exposures. McCord and Pease[7] have tried a positive bias (the bias voltage with a positive polarity at the STM tip), so that the electrons can be field emitted from the sample up through the resist without scattering inelastically and can quickly gain enough energy to cause exposure. As shown in Fig. 2(b), a 2.5 kΩ thin-film resistor device having less than 100 nm width at its narrowest region was fabricated using a patterned PMMA resist exposed by a positive polarity. The thin-film device was fabricated by the liftoff technique having a 13.5 nm thickness of Au–Pd. In addition to increasing the pattern resolution by avoiding the problems associated with backscattering and secondary electrons, the other advantage of reversing polarity is that the beam current instability caused by electron induced desorption (or ionization) of atoms and molecules can be reduced, especially at relatively high voltages.

STML in an ambient condition can be used for modifications of metallic and ceramic materials. As mentioned earlier, Dagata *et al.*[5] used STM to induce local oxidation on a hydrogen-passivated Si substrate in air, in which the field-

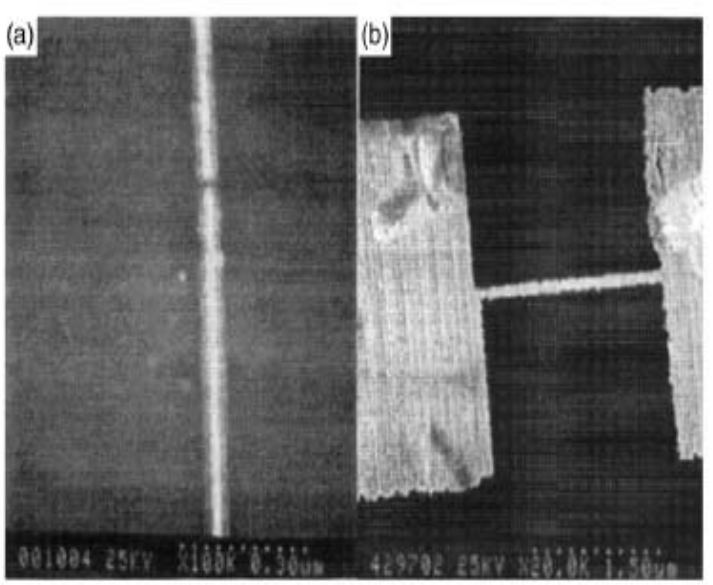


FIG. 2. Nanostructures fabricated by liftoff using STML patterning of PMMA resist subsequently with Au–Pd deposition: (a) Au–Pd line with 12 nm thickness and 22 nm width adopting 100 pA beam with 20 V (negative polarity at tip) and 1 μm/s writing speed in resist patterning; (b) 13.5 nm thick Au–Pd thin-film resistor having less than 100 nm width at narrowest region using tip-positive polarity pulse for patterning (after Ref. 7).

Conventional technology (e.g., liftoff) can be combined with STM lithography

Nanofabrication by scanning probe microscope lithography: A review

Ampere A. Tseng[a)]
Department of Mechanical and Aerospace Engineering, Arizona State University, Tempe, Arizona 85287-6106

Andrea Notargiacomo[b)]
Department of Physics, Roma TRE University, Rome 00146, Italy

T. P. Chen
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

877 J. Vac. Sci. Technol. B 23(3), May/Jun 2005

Nanofabrication by STM

9.4 Nanofabrication in Solution and in Gaseous Environments

Electron-beam-induced etching and deposition is a way of writing patterns on substrates in solution and in gaseous environments. The basic idea for etching and deposition is very simple. The focused beam is used to supply energy to decompose chemicals in a localized region. The decomposition products can include a metallic species to be deposited on a surface, or a corrosive species intended to participate in an etching reaction resulting in locally etched structures on a surface. The substrates used include Si, GaAs, graphite and metals. The STM system for operation in solution should supply the necessary solution. To minimize unwanted Faradic leakage current, the tip must be treated using some special methods such as coating wax up to the extreme end of the tip (Chap.4). Deposition and etching can also be induced with the STM tip as an electrochemical electrode to drive a localized Faradic current of ions. A gaseous environment can be utilized by introducing organometallic gas with a pressure of several Pa into a vacuum chamber with a base pressure of $10^{-5} \div 10^{-6}$ Pa. For different deposition metals, the introduced gas is different. They include DMCd, $W(CO)_6$, WF_6, and organometallic chemical of Au. Three possible mechanisms have been assumed to account for the dissociation of organometallic molecules [9.35]. (i) Electrons tunneling inelastically between the tip and sample break apart gas molecules adsorbed on the surface of the substrate. (ii) Current traveling between the tip and sample can locally heat the surface of the substrate enough to cause pyrolytic dissociation of adsorbed gas molecules. (iii) High fields between the tip and sample break down the gas creating a microscopic plasma between the tip and sample which then deposits the metal atoms on the surface. All of these will break the chemical bonds from the energy of electrons traveling between the tip and sample, thus the STM ought to be operated in the field-emission mode.

Electric fields locally applied by STM can induce local modifications, decomposition, chemical reactions, structure growing, ...)

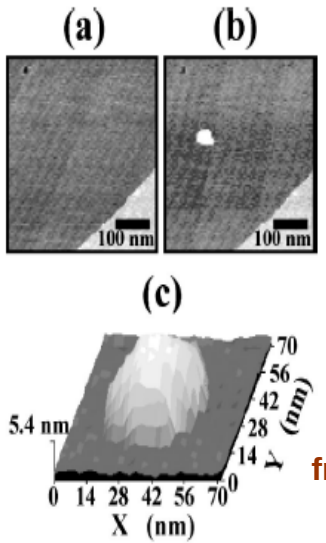


FIG. 2. Deposit from Me_2Au-tfac on HOPG substrate at a tip voltage of −6 V, current of −300 nA and a pulse time of 0.5 ms (ten times); (a) before, (b) after deposition, (c) landscape plot of a close-up of (b).

STM-induced deposition from precursors *in ambient atmosphere*

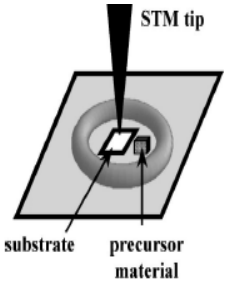


FIG. 1. Schematic of the setup for the STM CVD deposition in air. The substrate together with the solid precursor material is placed in a trough (viton ring) confining the vapor of the precursor molecules.

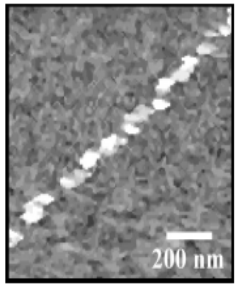


FIG. 5. Row of single dots deposited from $CpPtMe_3$ on ITO at a tip voltage of −6 V, current of −50 nA and a pulse time of 15 μs (single pulse).

Low energy electron beam decomposition of metalorganic precursors with a scanning tunneling microscope at ambient atmosphere

H. Brückl[a)]
University of Bielefeld, 33615 Bielefeld, Germany and Institute of Solid State and Materials Research, IFF, 01177 Dresden, Germany
J. Kretz
Institute of Solid State and Materials Research, IFF, 01177 Dresden, Germany
H. W. Koops
Deutsche Telekom AG, 64295 Darmstadt, Germany
G. Reiss
University of Bielefeld, 33615 Bielefeld, Germany

1350 J. Vac. Sci. Technol. B 17(4), Jul/Aug 1999

STM nanomodifications I

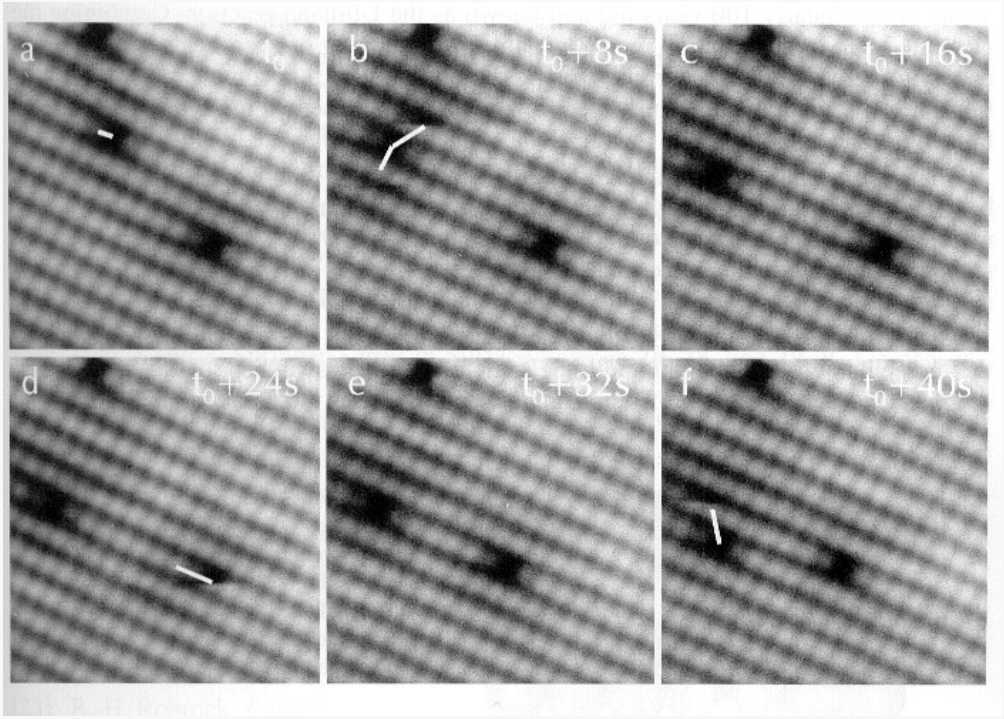
5.3 Effects induced by the Tunnel Current

It is also possible to excite atoms by the tip-sample interactions. Figure 42 shows a set of consecutive STM images acquired with 8 s time interval. The images show that the defects change their lattice positions [58], [59]. The tip can excite defects by several physical mechanism. The case shown here is based on a field-induced migration, due to the strong electrostatic field penetrating into the semiconductor. Defects can, however, also be excited by tunnelling of minority carriers into defect states followed by a charge carrier recombination with electron-phonon coupling [47]. There are surely even further mechanisms which may possibly excite atoms on the surface. Which of those will take place depends sensitively on the measurement conditions.

Finally, in Figure 43 the cutting of a carbon nano tube with an AFM is presented [60]. Earlier experiments controlling the length of carbon nano tubes were carried out using a STM [61]. At first the 600 nm² area is scanned then the AFM cantilever is lowered at the positions marked in the left image and for cutting a voltage pulse of –6 V is applied. The image on the right hand side shows the carbon nano tube after cutting.

Defects can be “excited” (or created) by local tunneling currents

Figure 42: Migration of three phosphorus vacancies on the GaP(110) surface. The change of the lattice positions of the vacancies is induced by the tip of the scanning tunnelling microscope. In this particular case the jumps are field-induced.



Field-induced migration of a single P vacancy in a GaP surface

STM nanomodifications II

5.4 Complex Chemical reactions with the STM

In 1904, Ullmann et al. heated iodobenzene with copper powder as catalyst and discovered the formation of biphenyl with high purity [62]. This aromatic ring coupling mechanism is now nearly 100 years old and known as the Ullmann reaction. Combining the presented STM manipulation methods, namely the moving of adsorbents and the influence via increased tunnel current, it is possible to control this complex chemical reactions at low temperatures step by step. S. H. Hla et al. presented the synthesis of to one biphenyl molecule out of two iodobenzene on a copper surface at 20 K [48].

The synthesis consist of three different steps. First two iodobenzene (C_6H_5I) have to be dissociated into phenyl (C_6H_5) and iodine (Figure 44a and b). Secondly the two phenyl rings have to be located one to another (Figure 44d and e) and finally in the third step, through tunnelling electrons the two phenyl rings are associated to biphenyl (Figure 44e).

To abstract the iodine from the iodobenzene the STM tip is positioned right above the molecule at a fixed height and a the sample voltage is switched to 1.5 V for several seconds. The energy transfer from a single electron causes the breaking of the C-I bond Figure 45a – c, [62]. As the bond energies of the C-H and C-C bonds are two and three times higher than the C-I bond, it is not possible to break them with a single electron process at this voltage. After preparing to phenyl reactants and moving away the iodine, the left phenyl Figure 45c is brought close to the other one by lateral manipulation using the tip adsorbate forces Figure 45d. Though the two phenyls are close together they do not join at 20 K. The two phenyls can easily be separated again by lateral manipulation. Both phenyls are still bond to the Cu step edge via their $\sigma_{C\text{-}Cu}$ bonds. Figure 46 shows a model where the phenyl is lying with its ð ring on the terrace while one of its C atoms is pointing towards the step edge and σ-bonding to a Cu atom. The final reaction step to associate the two phenyls to biphenyl is done by positioning the tip right above the centre of the phenyl couple and increasing the current drastically. The successful chemical association can be proved by pulling the synthesized molecule by its front end with the STM tip [62].

STM-controlled electrochemical reactions

Field-induced iodobenzene dissociation and diphenyl formation

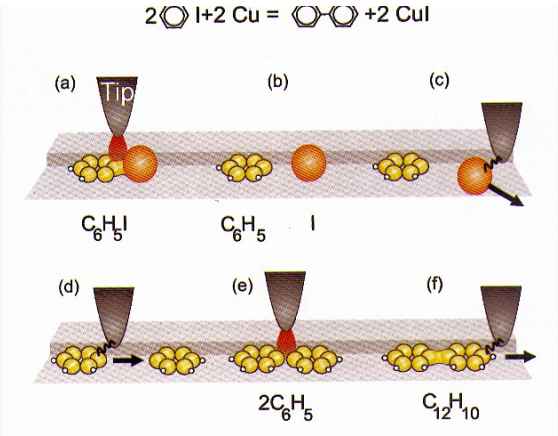


Figure 44: Schematic presentation of the tip-induced Ullmann reaction.
(a), (b) Electron-induced abstraction of the iodine from the iodobenzene.
(c) Pulling the iodine atom to a terrace site.
(d) Bringing together to two phenyl molecules by lateral manipulation and
(e) electron-induced chemical association to biphenyl.
(f) Pulling the synthesized molecule by its front end to prove the association.

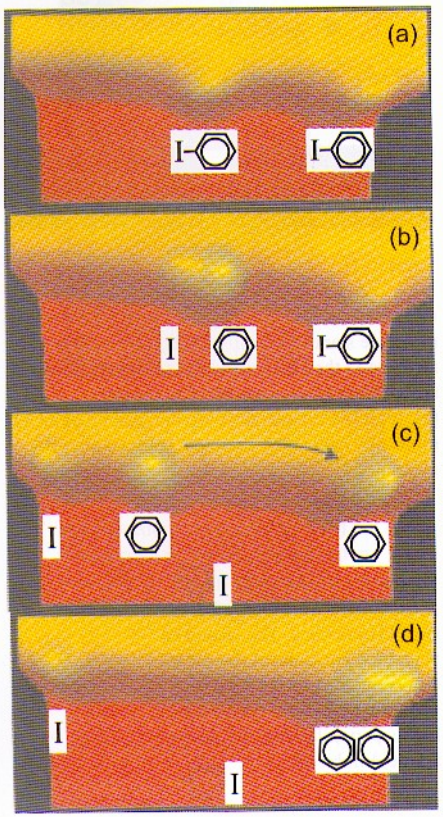


Figure 45: STM image of the Ullmann synthesis induced by the tip.
(a) Two iodobenzene molecules are absorbed at a Cu(111) step edge. Introducing a voltage pulse through the tip abstracts the iodine from the phenyl molecules
(b) (the left molecule).
(c) By lateral manipulation the molecules are further separated and
(d) the phenyl molecules are moved together to prepare for their association. (scan area 7 × 3 nm²) [48]

“Vertical” manipulation

5.2 Vertical Manipulation

In the vertical manipulation process, the adparticles are transferred from the sample surface to the tip apex and vice versa [51], [52]. The first experiments on vertical STM manipulation were carried out by Eigler “picking up” Xe atoms [51]. The group of Rieder showed that transferring a Xe atom to the tip apex leads to markedly improved resolution [53]. The single Xe atom obviously “sharpens” the tip.

In Figure 41(a) [54], [55] a schematic presentation of the *pick up* process of a Co molecule from Cu(111) is shown. It is well known that CO molecules stand upright on a Cu(111) surface [56] with the carbon atom bonding to the copper atoms. Due to occasional contact between the tip and the surface some copper atoms are transferred to the tip apex. During the transfer of the CO molecule to the tip, the molecule must consequently rotate. A reliable procedure for transferring the CO to the tip and back to the surface requires ramping of the tunneling voltage and the simultaneous decrease of the tip-surface distance. Figure 41b and c show that scanning with a CO molecule on the tip apex leads to a clear chemical contrast. Figure 41b is scanned with a clean metal tip and all adsorbents appear as depressions. After the transfer of the CO molecule to the tip apex (indicated with an white arrow) and rescanning the area, Figure 41c shows that all CO molecules changed their appearance to protrusions. Only the oxygen atom in the upper left part of Figure 41b and c retains its appearance.

In [57] it has been described, how to combine the potential of single atom manipulation of STM and single atom sensitivity of an atom probe mass spectrum to realize an ultimate technique for surface science. The System used by Shimizu et al. consists of an STM, an atom probe, load lock chambers and a mechanism to transfer tip and sample. The tip can be transferred reversibly between the STM and the atom probe stages. To investigate the pick-up of Si atoms during manipulation, a clean Si surface was approached with a clean tungsten tip applying a bias of + 2 V and 0.3 nA at the sample. After manipulation the tip was transferred from the STM to the atom probe. The atom probe analysis showed the formation of to different layers on top of the tip apex. The top most layer was WSi_2 and the next layer was W_5Si_3, finally the clean tungsten surface appeared. Using this combination of an STM and an atom probe could prove that the tips during manipulation do not only adsorb atoms but furthermore depending on the conditions alloys can be formed.

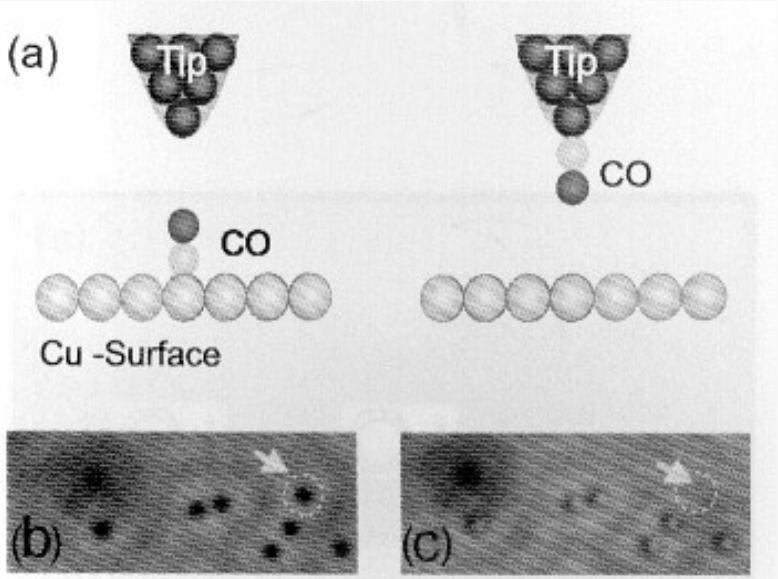


Figure 41:
(a) A sketch of the picking up procedure of CO molecules on Cu(111) surfaces. Notice that the CO molecule stands upright with the carbon atom attached to the surface and has to switch its orientation when being transferred to the tip.
(b), (c) STM images showing the pick up of a CO molecule. Notice the chemical contrast after the pick up [54].

Molecules or atoms can be “trapped” or “released” by the STM tip by simply varying the bias voltage

“Lateral” manipulation of single atoms I

Eigler and Schweizer first used an STM tip to position individual atoms on a surface by moving Xe atoms on Ni(110) at 4K [23, 24]. To move a Xe atom, the tip-sample separation was decreased by increasing the setpoint current (10–60 nA), resulting in a stronger tip-sample interaction. The tip with an accompanying Xe atom was then moved to its new destination with the feedback activated. To release the Xe atom, the tunnel current was reduced to imaging conditions (<1 nA). In addition to noble gases, a number of metal adsorbate atoms have been manipulated using low temperature STM (LTSTM). For example, Fe adatoms have been arranged on Cu(111) in order to construct a circular corral having a radius of 7 nm [25]. Due to strong scattering between the surface state electrons and Fe adatoms, the electrons become confined within the corral and produce an electron density distribution consistent with that for an electron trapped in a round 2D box [26]. Fig. 2 shows STM images of Fe atoms arranged on Cu(111) to form a variety of quantum “corrals,” where the ripple features indicate variations in the surface electronic density [27]. In another example, Fig. 3 shows the atom-by-atom manipulation of Ag atoms on Ag(111) to form a triangular corral [28]. Although all of the previous experiments have been performed with LTSTM, it is possible to perform limited lateral manipulation of atoms at room temperature (RT), one example being the movement of Br atoms on a Cu(001) surface [29]. However, the creation of highly ordered structures such as quantum corrals requires the controlled environment of low temperatures.

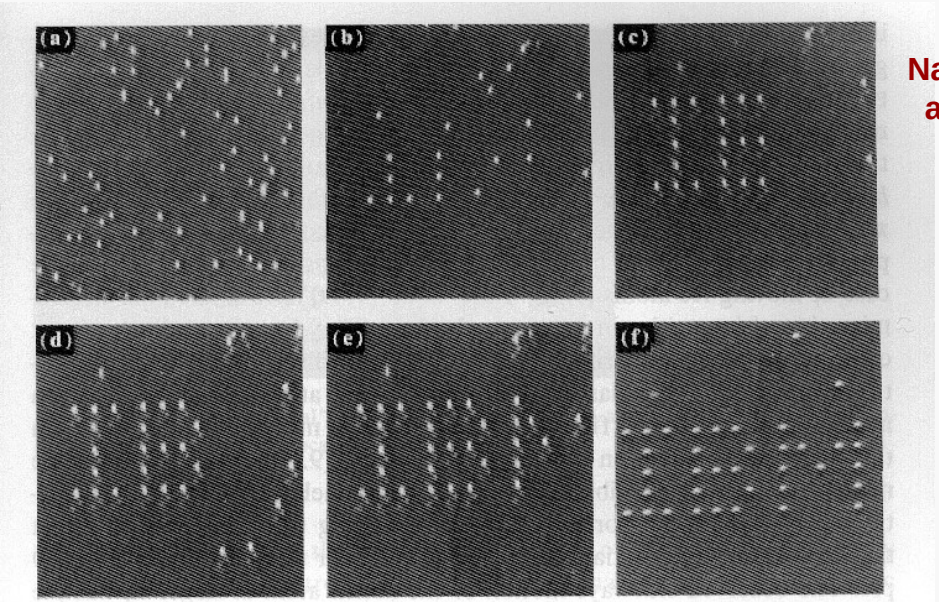
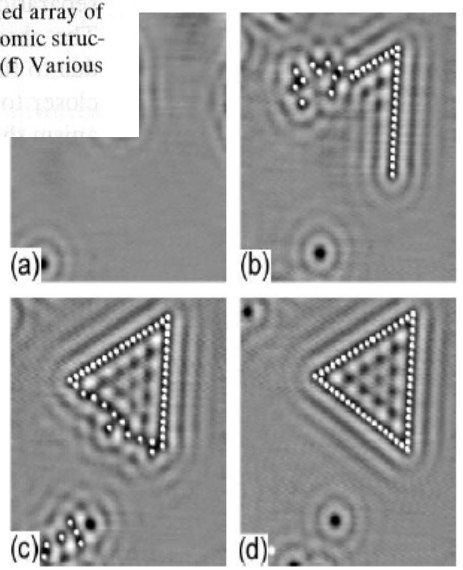


Fig. 9.10a-f. A sequence of STM images taken during the built-up of a patterned array of the letters I, B and M constructed of xenon atoms on the Ni (110) surface. The atomic structure of the nickel surface is not resolved. (**a**) The surface after xenon dosing. (**b**)-(**f**) Various stages during the construction. Each letter is 5 nm from top to bottom [9.40]

Nanomanipulation of Xe atoms on a Ni surface

Nanomanipulation of Ag atoms on a Ag surface



Typically, LowTemperature STM are used (e.g., to prevent atom diffusion)

To be published in “Advanced Semiconductor and Organic Nano-techniques, Part 3 (edited by Morkoc),” Academic Press (2002)

Suggested reading:

Fabrication of Nanoscale Structures using STM and AFM

Alison A. Baski

“Lateral” manipulation of single atoms II

5.1 Lateral Manipulation

In the lateral manipulation mode a particle on the sample surface is moved along the surface to the desired location without losing contact to the surface. The motion can be obtained either by dragging or pushing. Figure 31a to d show the build up of a quantum corral by manipulating individual Fe atoms on a Cu(111) surface at 4 K [44]. The final structure of 15 nm diameter consists of 48 Fe atoms. At this stage one may ask how such a fine manipulation can be achieved. The procedure is as follows: Fe atoms are evaporated onto a Cu(111) surface cooled to 4 K. The surface containing statistically distributed Fe atoms is then examined by STM. Normally, no atoms are displaced, but if the distance between tunnelling tip and a Fe atom is reduced, then the tunnelling tip exercises an attractive force on the Fe atom and the Fe atom can be dragged by the tunnelling tip to the desired location on the surface [44], [45]. Once the desired location is reached, the tip is retracted. Increasing the distance between the tip and the sample reduces the tip-Fe atom interactions and, hence, the Fe atom remains at its new position (Figure 32). Figure 31e shows that in this way a whole circle of iron atoms can be built up. The artificially build nanostructure shown in Figure 31 confines the electrons of the two-dimensional surface electron gas on Cu(111). Therefore, as soon as the circle is complete, the electrons are scattered in the circle and form standing electron waves due to quantum mechanics. Figure 31e thus illustrates the wave nature of the electrons.

In a quantum well, due to the quantization of the electron states, not only standing electron waves but also discrete energy values of the electrons are expected. Consequently, increased electron densities should occur at specific energies. As already described, the density of states can be approximately calculated from experimentally measured current-voltage characteristics by calculating $(dI/dV)/(I/V)$. For metals, however, I/V is generally constant and the sample density of states is therefore proportional to dI/dV. The variation of dI/dV as a function of voltage reflects the density of states variation as a function of energy. Figure 33 shows the density of states thus obtained for three different surface positions. At the centre of the circle, as expected, peaked energy levels occur (curve a), whereas outside the circle no structure in the density of states is measurable (curve c). If the density of states is measured at a distance of 0.9 nm from the circle centre, even more energy levels occur as shown by the arrows in curve b.

This example of spectroscopic measurements and of the spatial distribution of the electron waves in a potential well provides a particularly illustrative picture of quantum mechanics. The construction of different quantum structures by an atom-by-atom manipulation approach using scanning probe microscopes nowadays allows a new look into the quantum world and a direct spatial measurement of the electron waves.

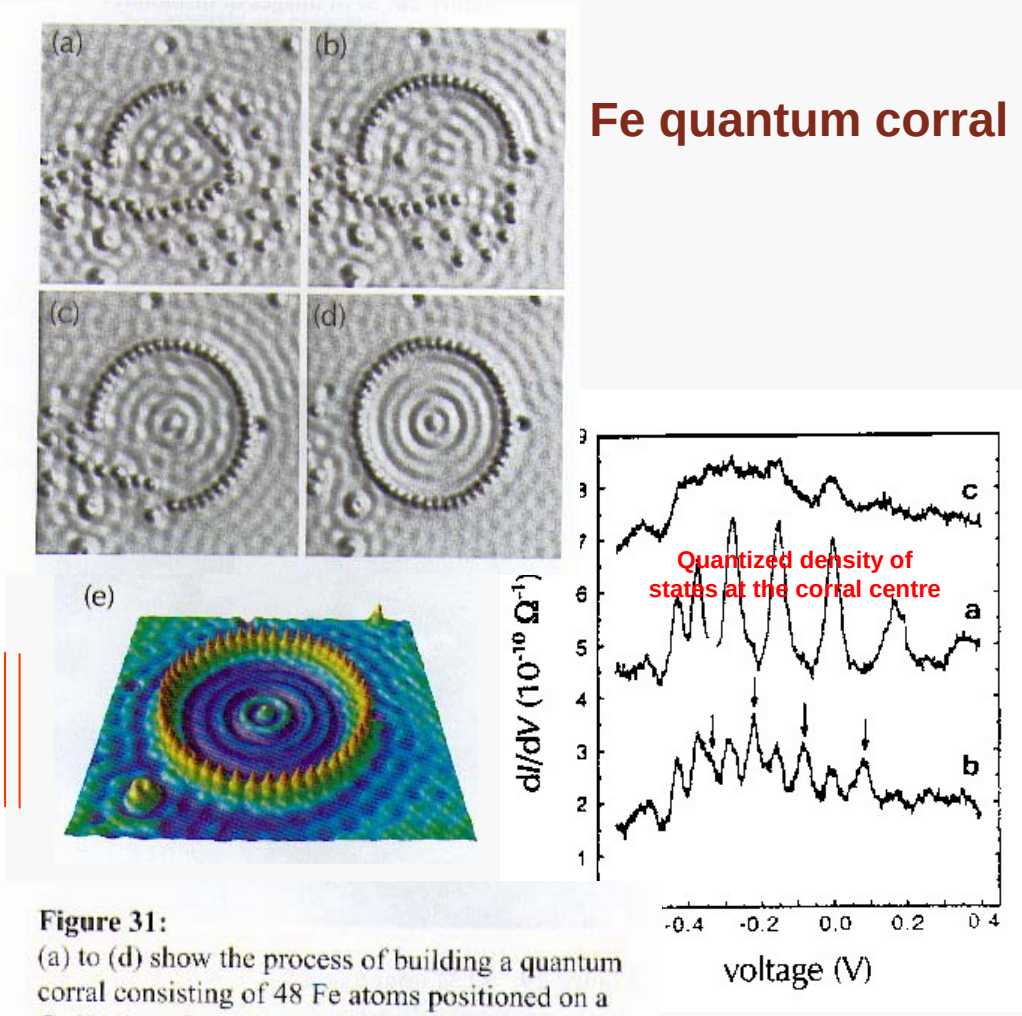


Figure 31:
(a) to (d) show the process of building a quantum corral consisting of 48 Fe atoms positioned on a Cu(111) surface. The resulting structure and the standing waves induced by the quantum confinement of surface electrons in the structure is visible in the three-dimensional view
(e) of the quantum corral [44].

Fe quantum corral

Artificial quantum structures can be produced by placing in close proximity different atoms (physisorbed) on a surface

“Lateral” manipulation and surface reconstruction

Instead of evaporating foreign atoms onto a copper surface it is also possible to reconstruct the substrate surface itself, which is more difficult due to the higher coordination number and binding energy of the atoms located in the surface or in steps [49]. The experiments shown here are carried out on Cu(211) substrates at 30 – 40K. In Figure 34 a sphere model of the copper surface is shown, whereby the atoms are shaded darker the deeper they lie. Lateral manipulation of single Cu atoms parallel and perpendicular to step edges is presented in Figure 35 [49]. A measure for the minimum force necessary to move a copper atom is the tunnel resistance which displays the distance between tip and sample. The tunnel resistance used for motion along a step edge was approx. 700 kΩ and ~500 kΩ for moving them over a step edge. Figure 36a – c demonstrate that it is even possible to “dig out” single copper atoms from even higher coordinated sites. The single Cu atom (Figure 36a) is used as a marker. Figure 36b, c show the drag out of single Cu atoms leading to a corresponding vacancies in the initial site of the atoms.

Furthermore instead of moving single atoms, the lateral manipulation technique is also capable to move entire molecules. Gimzweski et al. deposited hexa-*tert*-butyl decacyclene (HB-DC) molecules onto a Cu(100) surface [50]. The decacyclene core of the HB-DC is equipped with six bulky *t*-butyl-legs (Figure 37). At monolayer coverage, the molecules are immobile, forming a two dimensional van der Waals crystal (Figure 38). Separated HB-DC molecules on a Cu(100) surface are extremely mobile, making it impossible to get STM images with atomic resolution.

For this reason a coverage of just less than one monolayer was chosen and STM images resemble those of the immobilized 2-D lattice at full monolayer coverage. However, there are some random voids. In this layer the molecules can be at sites with different symmetry with respect to the surrounding molecules (Figure 39). Molecules at sites of lower symmetry rotate at speeds higher than the scan rate used for imaging and therefore appear as torus Figure 40a. The molecules at the higher symmetry sites are observed as six-lobed images, proving that they are immobile Figure 40b. Gimzweski used the lateral manipulation to drag a rotating HB-DC molecule from a low symmetry site into a higher symmetry site and the six lobes of the immobilized molecule was again clearly observed.

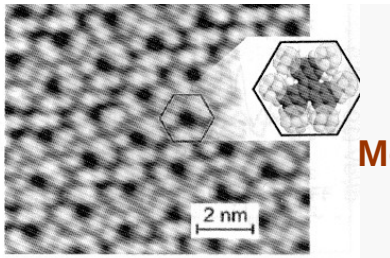


Figure 38: STM image of an Cu(100) surface after exposure to a full monolayer coverage of HB-DC molecules at room temperature. Image area is 11.4 nm by 11.4 nm [50].

Molecular manipulation

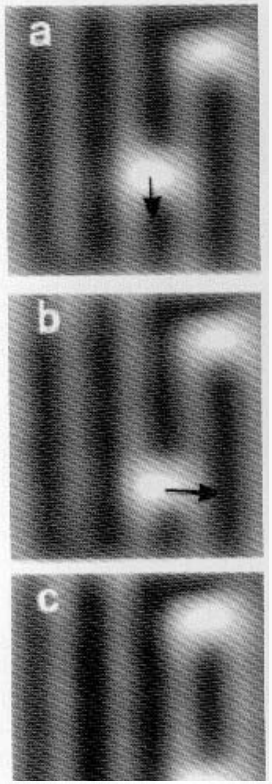


Figure 35: STM image showing the lateral manipulation of a single Cu atom (a) parallel and (b) perpendicular to a step. The motion is indicated by an arrow. In the upper part of the image a single copper atom serves as marker. The processes involved correspond to the motion shown in Figure 34a and b [49].

Surface reconstruction: practically, atoms belonging to the surface can be displaced from one site to another

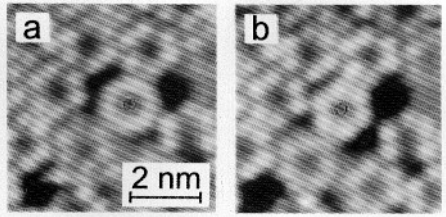


Figure 40: STM images of a Cu(100) surface after exposure to a coverage just below one complete monolayer of HB-DC molecules at room temperature. In (a) the molecule is imaged as a torus and is in a location where it is not in phase with the overall 2D molecular overlayer. The molecule is rotating. (b) The same molecule is translated by 0.26 nm and imaged as a six-lodeb structure in registry with the surrounding molecules. Image area is 5.75 nm by 5.75 nm.

4.C Resist-assisted nanolithography with SNOM

SNOM (in emission mode) represents a very local scanning source of light, hence allowing for a (serial) implementation of optical lithography

SNOM scanning head (Product of NT-MDT). We use a mechanical shutter to make a light pulses. We can control the shutter directly from Solver SNOM software (Lithography option) through scanning head controller.

In our experiments we used a thin layer of positive photoresist on flat Si substrate. After probe approaching we made a shear force image of sample surface to ensure that it is rather flat and free from artifacts. Then we moved probe over surface by especial marker in software or by programming it behavior. Software allows as to draw raster (pcx files) or vector images (lines, circles and alphabet signs). During probe movements a laser pulses were made. The pulses length was of 0.1-5 seconds. The total power reached probe aperture and hence the sample surface was 1-10 mW. Drawing immediately followed by shear force scanning.

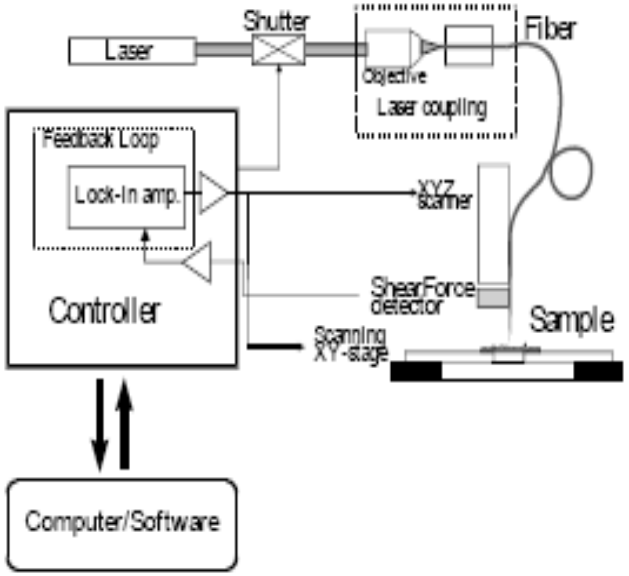


Figure 1. Experimental setup.

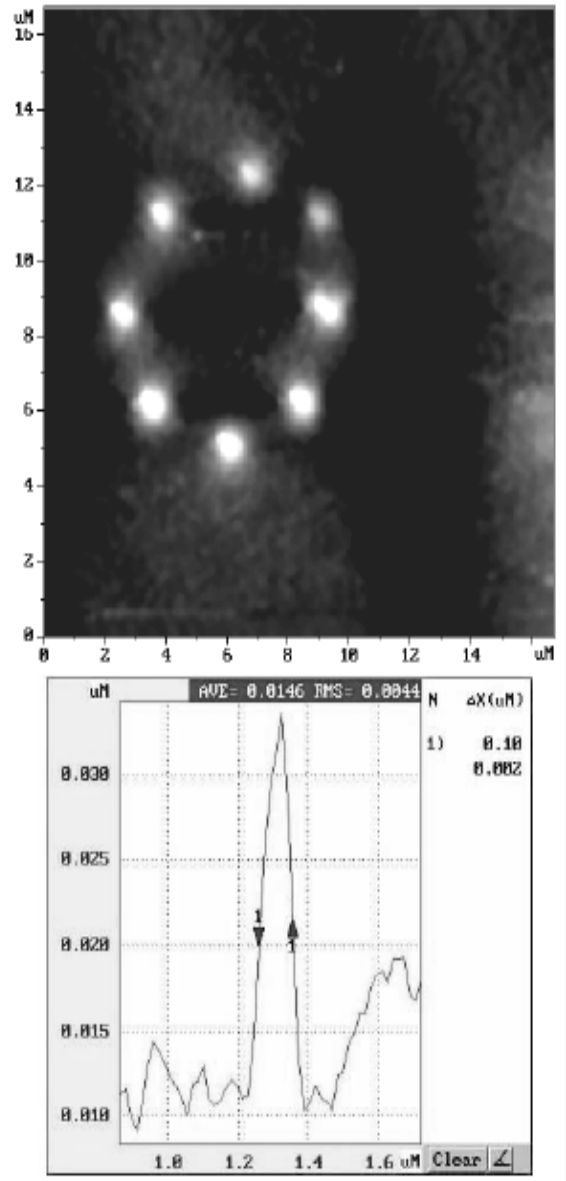


Figure 2. Dots on a positive photoresist and a crossection of one dot.

www.ntmdt.ru

SNOM “nanowriting” I

However, SNOM capabilities seem particularly appealing in the area of “nanowriting” (of photosensitive materials) for extreme data storage applications (*serial reversible* process)

5 Photoaddressable Polymers

5.1 Introduction

Polymers are perfect materials for a diversity of applications. Due to their macromolecular architecture there is an enormous variety to modify and control the synthesis of polymers. Thus, they can be tailored to even quite difficult demands. Since a whole industry deals with the processing of polymers, efficient production lines have been developed for almost every polymer. Not only the molecular composition, but also the cost of polymers has been optimized. Now these materials can be considered even for highly sophisticated applications, as optical and holographic data storage.

Looking at the development of optical media, it is obvious that polymeric materials play a major role for the compact disc (CD): Important for the success of the CD-ROM and CD-AUDIO was the development of a specific polymer – polycarbonate – which serves as substrate material, carrying the digital information in the form of tiny pits. In the ROM the information is imprinted during production and cannot be erased.

In photopolymers the optical information is stored by changing the optical properties of the material by a light-induced polymerization or a ring opening reaction. These materials turned out to be the most suitable holographic media for write-once-read-many (WORM) applications [66]. Since especially the optical requirements for holographic data storage are very stringent, there are some drawbacks that hinder the introduction of those polymers into mass production. So far, the shrinkage during illumination is the most severe problem for photopolymers.

As far as rewritable holographic storage is concerned, photorefractive crystals like iron-doped lithium niobate ($LiNbO_3$) are mostly used for rewritable (R/W) laboratory demonstrations [67]. In the following chapter, a polymeric class of materials, the so-called photoaddressable polymers [68] - [72] and the results of holographic measurements with these materials will be presented. These polymers show no shrinkage effects and react in a reversible way to the illumination of light and are therefore able to meet the requirements for R/W holographic data storage.

In principle, all materials that react to light with a change of specific properties can be described as photoaddressable polymers. Since the change of material properties is the basic condition for storing information in a material, we have to deal with a reversible modification in order to fulfill the requirements of an R/W material. In the following section the name photoaddressable polymer (PAP) is used as a synonym for a class of polymers with azo dyes as reversible antennae for the incident light. Those PAPs are basically azobenzene-containing side-group polymers. The aim of this chapter is to give an overview of the optical and photophysical properties of those polymers.

Basic idea:
To modify the optical properties of a photoaddressable material (i.e., to **write**) on a very local scale through a light-controlled mechanism and to **read** by the same method

… but …

Suitable materials are needed

Purely light-driven optical modifications are accessible (without any morphology change?)

⇩

Fast and reversible writing

Photopolymers I

5.2 Photochemistry of Azobenzene

The photophysical reactions of azobenzene have been well studied. It can be demonstrated that due to controlled light-induced reactions of azo chromophores, properties of the whole system incorporating the dye can be modified. Those properties include for example viscosity, solubility, mechanical parameters, bioactivity, and optical constants [73]. Additionally azobenzene molecules can be used as a probe for their molecular surroundings. In this way a detailed study of polymeric parameters can be performed by monitoring the photochemical behavior of the azo dyes. The success of azo dyes in all these applications can be explained by detailed knowledge of their light-induced photoreactions: Azobenzene chromophores exist in two isomeric states. The rodlike long shaped trans form and the bent cis configuration. The isomerization can be induced by light in both directions, from trans to cis and from cis to trans, whereas the cis-isomer can also undergo a thermal back relaxation to the thermodynamically more stable trans-isomer (Figure 21).

A look at a typical absorption spectrum (Figure 22) of unsubstituted azo chromophores (Figure 21) shows two absorption bands: The so-called $\pi\pi^*$ band with a maximum absorbance at $\lambda_{max} \sim 360$ nm and the so-called $n\pi^*$ band with $\lambda_{max} \sim 460$ nm. The spectral position of the absorption maximum of the $\pi\pi^*$ band can be shifted by chemical modification of the dye, namely by the replacement of donor/acceptor substituents. Those substituents do not influence the spectral position of the $n\pi^*$ band. The isomerization cycles have a distinct influence on absorption and optical index, which can be used for data recording.

The absorption bands define the laser wavelength needed: Azo-dyes undergo their isomerization cycles by illumination with light of wavelengths from the UV to the green/yellow range of the optical spectrum.

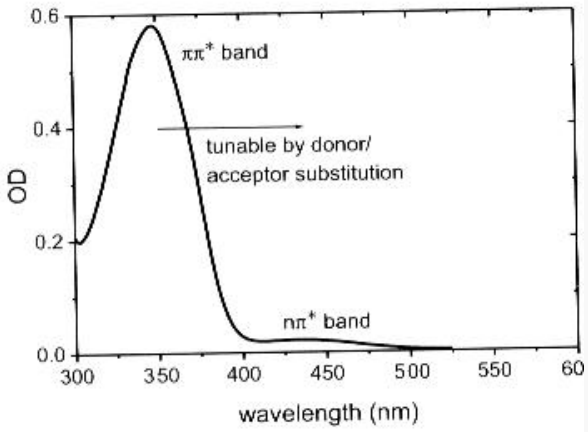


Figure 22: Absorption spectrum of unsubstituted azobenzene in solution.

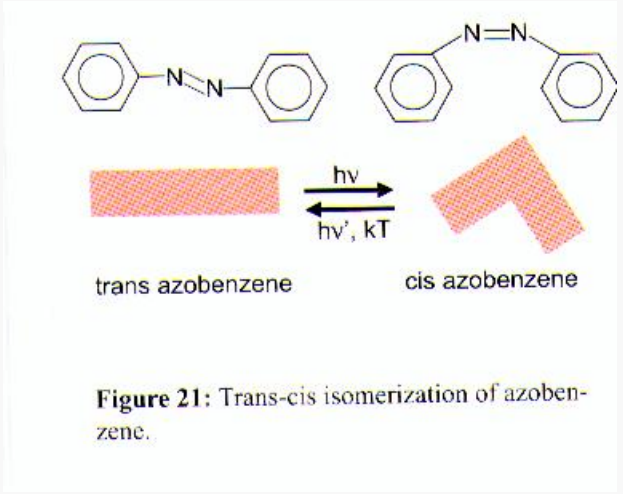


Figure 21: Trans-cis isomerization of azobenzene.

5.3 Azobenzene-Containing Polymers

There are three ways to incorporate chromophores into a polymer:

- In guest-host-systems chromophore guests are doped into a polymeric host. In these systems the chromophore concentration cannot exceed a specific value, because highly concentrated chromophores tend to phase separation and crystallization.
- This disadvantage can be avoided by attaching the chromophores to the backbone as side chains.
- Or the chromophores are fixed as part of the main chain.

The first investigations of the isomerization kinetics of azo chromophores in polymers were performed in 1972 [74] and mainly focused on spectroscopic measurements [75] – [78]. Those experiments can be summarized by the experimental observation that the isomerization of azo dyes in a polymeric environment is possible, even at temperatures below the glass transition temperature. This result applies to side chain as well as to main chain polymers.

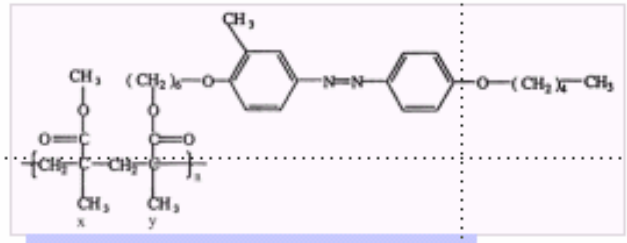
Since the isomerization results in a change of the molecular shape of the azo chromophores, as described above, there is a specific demand for free volume to enable this reaction[4]. In solution this free volume condition is fulfilled. In polymers, however, there might be some sterical constraints due to the inhomogeneous free volume distribution. For some chromophores there are local environments with large free volume, where the isomerization can be performed in a similar way as in solution. But there are also configurations where the free volume is not sufficient for the isomerization of the dyes. Therefore strong deviations of the photochemical reactions in polymers occur, if compared to the reactions in a solution.

However, experiments performed in polymeric systems demonstrate that azo dyes are able to undergo isomerization reactions in this environment.

Photopolymers II

Photosensitive polymer sample

Polymethacrylate (PMA) network, modified in the four position with an azobenzene mesogenic unit (3-methyl-4'-pentyloxy) connected to the main chain by an hexamethylene spacer (PMA4).



PMA4
chemical structure
homopolymer ($x = 0$)
co-polymer ($x \neq 0$)

Samples preparation

- **highly concentrated solution of PMA4 powder dissolved in chlorobenzene, stirred at room temperature for a few hours**
- **thermal treatment before impression**
- **film from single drops of solution on Corning 4079 glass substrate and spinning at ~ 2000 - 10000 rpm.**

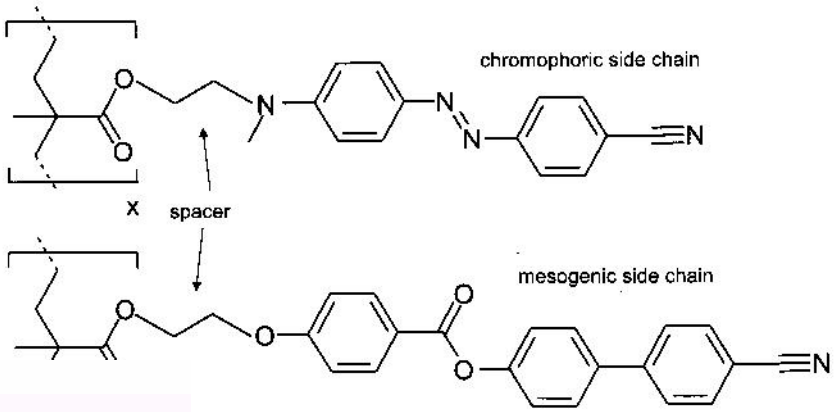
- **T_g = 294 K.**
- **Above T_g, nematic phase up to T_c ~ 353 K.**
- **Conformational transition at 320 K.**
- **Film thickness ~ 100 - 200 nm**

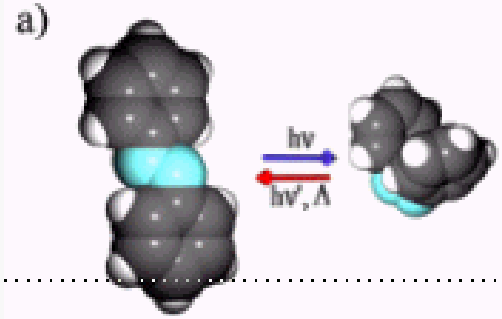
5.5 Photoaddressable Polymers for Optical Storage

The motivation of the development of polymers following the scheme of Figure 23 can be summarized as follows:

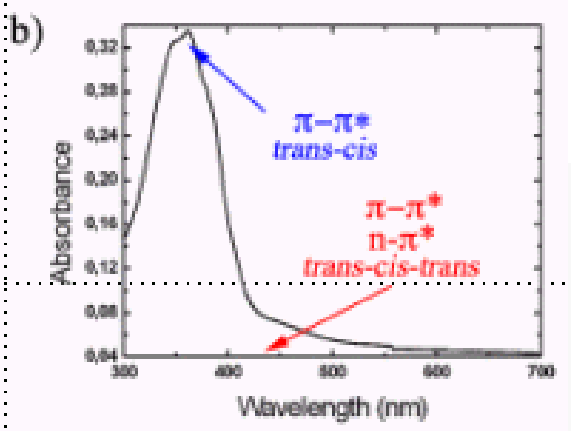
The dye acts as an antenna for the incident light. In order to optimize the response to light, adequate dye systems must be used. Azo dyes are the perfect chromophores for this purpose because of their well-known photochemistry.

It is the task of the mesogenic groups to stabilize and to amplify the reorientation of the chromophores. The physical reason for this can be explained by the intrinsic tendency of the mesogens to spontaneously organize in domains.





Trans-cis photoisomerization of the azobenzene moiety.



Absorption spectrum

Photopolymers can be engineered to improve their properties (e.g., to limit the topography changes upon irradiation)

SNOM “nanowriting” II

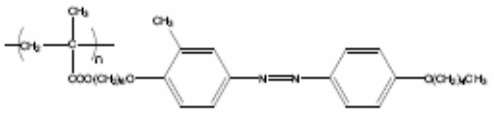


Figure 4.2: Schematic of the azo-polymer we used for near-field lithography (*PMA*4). *PMA*4 is polymethacrylate (*PMA*) containing 3-methyl-4′-pentyloxy-azobenzene unit connected at the 4-position of the main polymer chain via six hexamethylene spacers.

Near-field illumination has already been used for pure optical lithography on *PMA*4 samples [132]. This technique has provided (see also next section) sub-wavelength resolution. The best resolution achieved in this process until now is of 120 nm [132]. The application of near-field microscopy (SNOM) in pure-optical lithography on azo-based polymers leads to estimated potential storage density of at least 1 Gbit/cm^2.

Material engineering

the spacer length is adjusted to obtain both stability of the azo-molecules configurations in the polymer and a liquid crystaline phase, particularly appreciated for holographic high density data storage (see next section). As a result, the glass transition temperature of the resulting compounds ranges into the interval 30°C-100°C.

The rate of thermal isomerization from *cis* to *trans* configuration is negligible with respect to the rate of photo-activated transitions. In order to achieve light-driven complete *trans-cis-trans* isomerization cycles, research is oriented towards materials presenting *trans-to-cis* absorption band superimpose to *cis-to-trans* absorption band. Following the classification introduced by Rau [117], these materials are the so called *pseudostilbenes*, i.e., azobenzene-based polymers containing electron-donor and electron-acceptor substituents connected to the azobenzene groups. *Pseudostilbenes* have a $\pi-\pi^\star$ band practically superimposed to $n-\pi^\star$ band allowing light-induced excitation from *trans-to-cis* and *cis-to-trans* using the same light source.

SNOM writing is based on a (reversible) modification of optical properties following azo isomerization (remember: trans and cis have a different geometry and exhibit different birifringence properties)

Write mechanism

Since the pioneering work by Todorov et al. [121], azobenzene containing polymers have been largely studied as materials for high density data storage. Due to the rod-like shape of the azobenzene *trans* isomer, in the light driven isomerization, the transition momentum of *trans-cis* isomerization has its maximum when the molecule is oriented along the light polarization direction. Thus, complete *trans-cis-trans* isomerization cycles driven by illumination with linearly polarized light yields to a statistical orientation of the azo-moieties perpendicularly to the light polarization direction.

The orientation of azo-moieties perpendicularly to the light polarization direction results into the variation of the refractive index components inside the polymeric compound as has been studied using attenuated total reflection (ATR) [42]. As a consequence, a modification of the properties of azo-containing polymers occurs under illumination. This phenomenon has suggested azo-based polymers as material for holographic data storage.

Birefringence patterns are written well below the glass transition temperature of the polymeric compound and are stable for months. Nevertheless the written data may be erased by heating the sample above its glass transition temperature (T_g) or by exposure to a single circularly polarized light beam that destroys the order among the azo-molecules [125]. The writing process is not energetically

Antonio Ambrosio, PhD Thesis Appl. Phys., Pisa 2005 (unpublished)

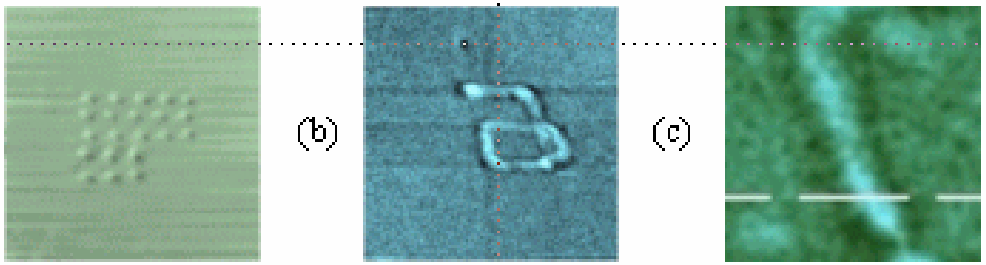
Examples of nanowriting (topographical)

SNOM topography images of the optically nanostructured PMA4 thin film

(a) homopolymer (exposure time = 1 s/dot, image size = 5 x 5 μm²)

(b) 30/70 copolymer (sensor speed = 20 nm/s, size = 2 x 2 μm²)

(c) 30/70 copolymer (speed = 50 nm/s, size = 300 x 300 nm²)



180 nm width

75 nm width

45 nm width

[S. Patanè, A. Arena, M. Allegrini, L. Andreozzi, M. Faetti, and M. Giordano, Opt. Commun. 210, 37 (2002)]

Topographical nanowriting (due to local photoinduced mass migration)

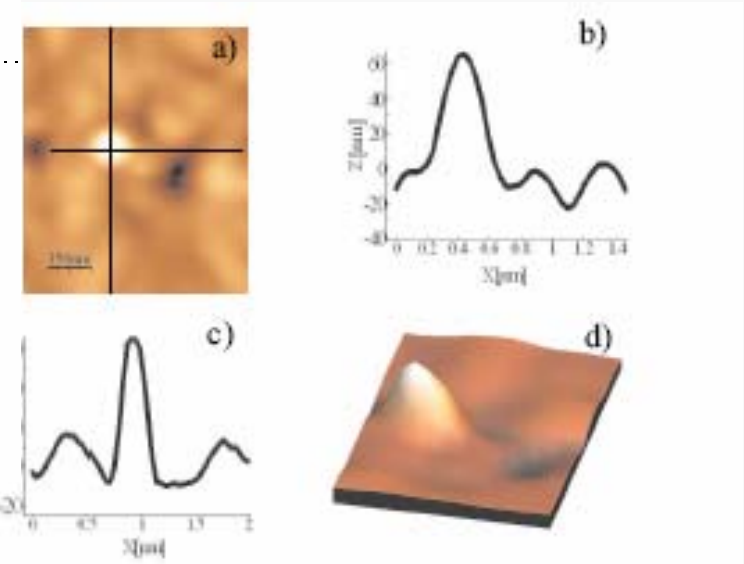
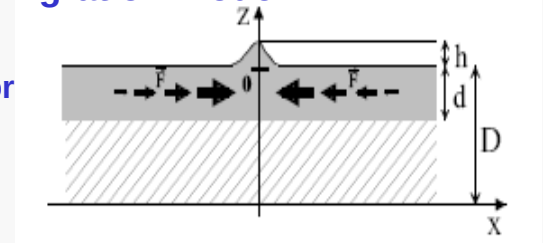


Figure 4.6: Typical dot drawn in our experiment on the surface of *PMA/PMA4* copolymer films. The dot height is of about 65 nm while the dot diameter is of ~ 250 nm (FWHM)

Mass-migration model

Material is "attracted" (or "repulsed") by the light gradient: slow process!



Antonio Ambrosio, PhD Thesis Appl. Phys., Pisa 2005 (unpublished)

Examples of nanowriting (optical)

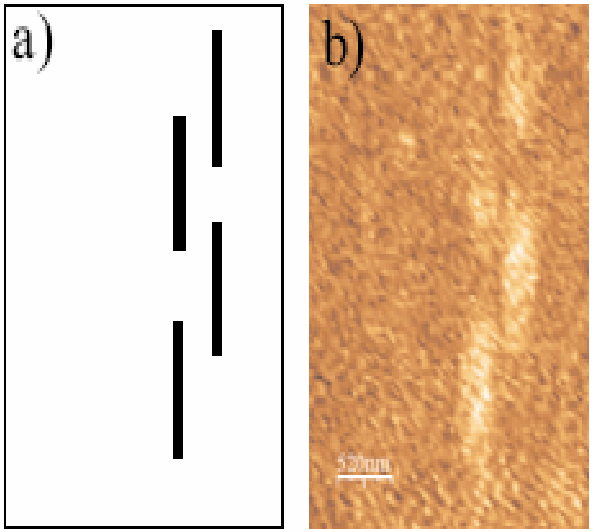
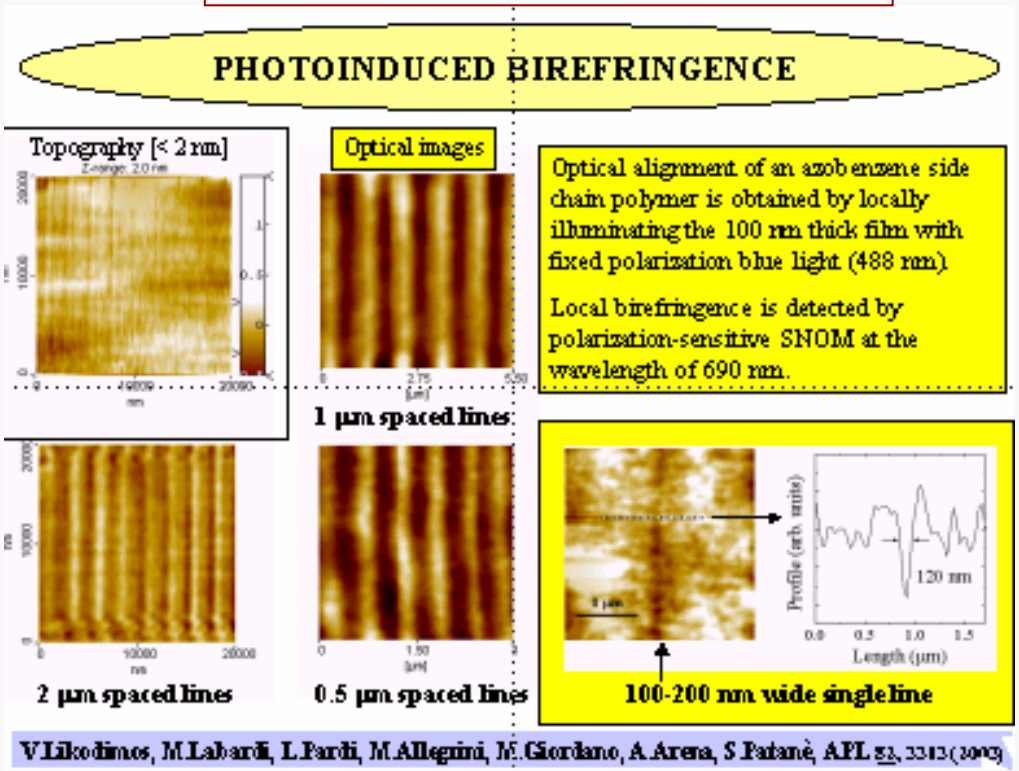


Figure 4.4: a) Representation of the birefringence pattern realized on the free surface of a 30/70 *PMA/PMA4* copolymer film. b) SNOM image of the pattern really realized. In order to read the written information we have used the high-contrast mechanism resulting from polarization modulation that we have implemented in our SNOM setup.

PM-SNOM is used for reading

Promises exist for a ultra-high dense writing of information (estimated above 10 Gbit/cm^2)

Optical writing result in the formation of a pattern of optically-active regions (i.e., in such regions azo-groups birifringence is used to modify the transmission of a polarized light)



Conclusions

✓ Exploitation of a probe able to “produce” (or “detect”) an optical near-field allows to overcome limitations due to diffraction in the space resolution of conventional microscopy

✓ Spectroscopy methods (including advanced ones) can be transferred to the local scale thanks to SNOM

✓ The excellent space control offered by SPM can be exploited also for fabrication (nanomanipulation) purposes, leading to techniques with a poor applicative potential (slow, complicated), but with enormous capabilities in a bottoms-up context, compatible also with “soft” (typ organic) matter