# Local probe microscopy: contact and manipulation

The invention of scanning tunnelling microscopy was a breakthrough for the nanosciences. This technique makes it possible not only to observe atoms but also to move them one by one and assemble them in a planned way. Atomic force microscopy, which is derived from it, has proved to be particularly well-suited to exploring the nanoworld. Here again, the probe used, in special conditions, offers a nanotool for the assembly of artificial surface nanostructures.
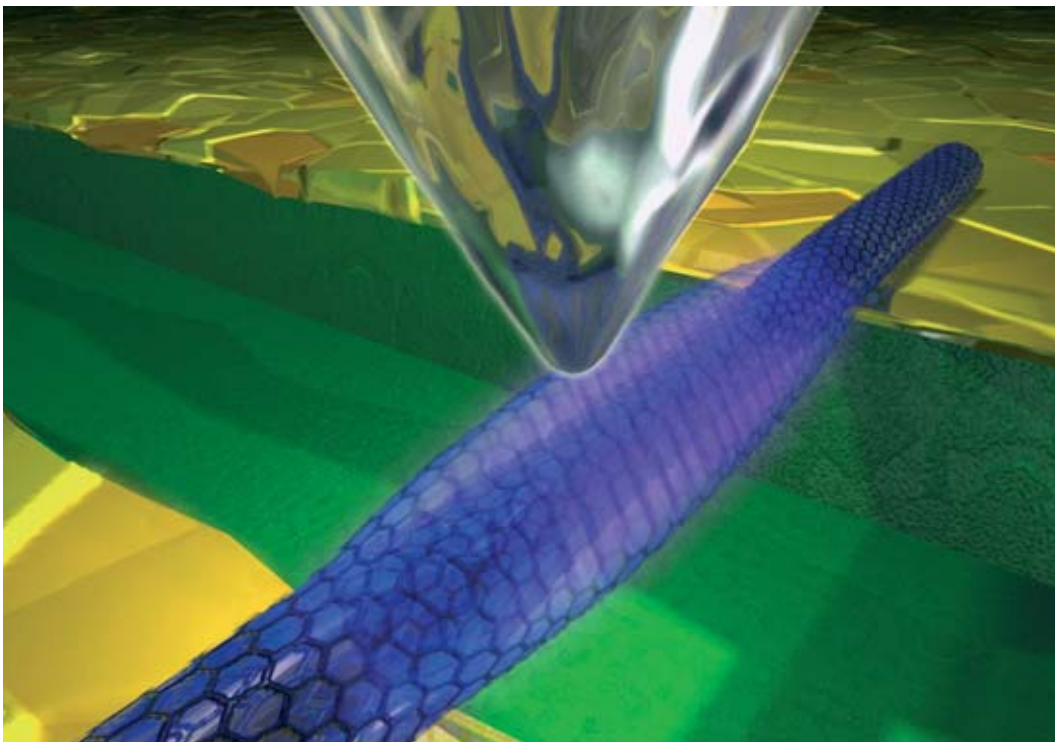
Tremani/Molecular Biophysics Group/Delft University of Technology

Artist's impression of the STM probing of a suspended vibrating carbon nanotube.

Conventional electronic microscopy (see *Conventional electron microscopy: transmission and scanning*), although very effective in certain applications, is often ill-suited to characterizing or building nano-objects. For example, accurate measurement of the relief of a sample using scanning electron microscopy (SEM) requires complex processing of the signal. Although chemical information can be extracted from an analysis of the **photons** generated by the impact of the **electron** beam on the sample, SEM is inappropriate in many systems that are of interest in the **nanosciences** and in **nanotechnology**. Its lateral **resolution**, which is directly related to the beam diameter is at best only a few **nanometres**[1]. For example, the structure of self-assembled **monolayers** cannot be approached for two reasons. First, the high energy density deposited by the electron beam induces some damages to the monolayer. Second, the molecules are generally much smaller in size than the beam diameter. In transmission electron microscopy (TEM), atomic resolution is obtained, but it corresponds to the projection onto a screen of columns of atoms parallel to the beam.

A major advance in the development of the nanosciences came with the invention of the **scanning tunnelling** microscope in 1981 at IBM's Zurich laboratories by Gerd Binnig and Heinrich Rohrer. Using such a microscope, these researchers succeeded in mapping, at the atomic scale, first a gold surface and later a silicon surface. This invention jointly earned them the Nobel Prize in Physics 1986, which they shared with Ernst Ruska, the inventor of the conventional electron microscope in the 1930s. Also in 1986, Binnig proposed a new instrument he called the atomic force microscope (AFM), derived from the STM. This device, which works with very weak forces ($\sim 10^{-9}$ N), makes it possible to examine the surface of all types of solid, and is particularly well-suited to exploring the nanoworld. Both STM and AFM can characterize local surface properties, unlike the diffraction techniques, which provide information on symmetry and mean distances between atoms on typically millimetre-scale objects of interest present in the analysis area.

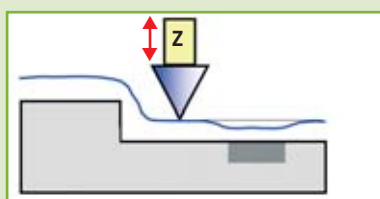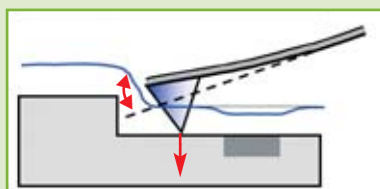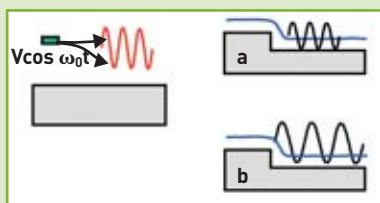(1) A beam of less than 1 nm is claimed on a Hitachi prototype)

Figure 1.
Top, path of the average position of the probe of a local probe microscope along a surface while a system adjusts Z so that the tip-surface interaction remains constant. When the surface is homogeneous, this path shows the topography. When a chemical heterogeneity appears (grey area) that modifies this interaction, the path is modified.
Centre, principle of AFM in contact mode. The tip is in repulsive contact with the surface and exerts a force perpendicular to the surface, causing a deviation matching the spring stiffness force. The green arrow indicates this gap between the position away from the surface (dashed line) and the contact. The path of the tip (blue) corresponds to the topography, except above a heterogeneity where the tip-surface interaction changes.
Bottom, principle of vibrating AFM. Away from the surface, the cantilever vibrates at its resonance frequency f_0 with an amplitude A_0. When the point interacts with the surface, the resonance of the cantilever is modified. The average position of the probe can be adjusted in two ways: a) the amplitude is reduced by a set amount at a set frequency (amplitude modulation), b) the frequency is adjusted by a set amount at constant amplitude (frequency modulation).

An astonishingly simple principle

The principle of local probe microscopy is based on the movement of a probe along the surface, like an electron beam in SEM. The probe has a fine solid tip, sharpened so that its end is composed of only a few atoms. It interacts with the surface when placed in close proximity to it. The interaction intensity depends strictly on the tip-surface distance. Consequently, if this interaction is kept constant, then the scanning of the probe will follow the relief of the sample surface. However, when physical or chemical surface heterogeneities modify locally the tip-surface interaction, the profile departs from the relief (Figure 1). In practice, all these microscopes share many features (Figure 2). The probe is attached to a set of
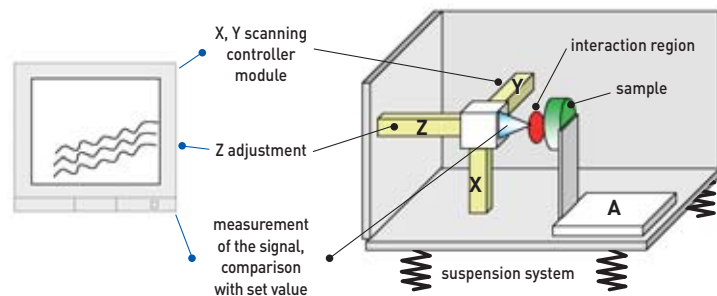

Figure 2.
Schematic view of a local probe microscope composed of three parts. The computer provides the man-machine interface and records and visualizes the data. The electronic controller runs the mechanical parts and regulates the potential voltage V_z so that the signal remains equal to the set value. Lastly, the mechanical parts must be well protected from the vibrations in the laboratory. The probe is attached to the ceramics, which displace it in a maximum volume measuring 100 x 100 x 10 mm³. The whole setup typically takes up a volume of 10 x 10 x 10 cm³. The system for approaching the sample must be extremely stable.

piezoelectric ceramics. These ceramics expand or contract according to the sign of the voltage applied, and will move through distances as small as a picometre (10^{-12} m). An electronic device continuously measures the tip-surface interaction, and, by means of a feed back system, continuously adjusts the voltage applied to the ceramic Z so that the interaction remains equal to a value set by the operator. In these conditions, the tip or its mean position remains at a constant distance from the sample surface. By moving the probe along the surface by means of the piezoelectric ceramics X and Y by point-to-point scanning, we obtain a series of iso-intensity profiles for the tip-surface interaction. The juxtaposition of these profiles, visualized on a computer screen, forms a pattern that corresponds to $V_z = f(V_x, V_y)$ where $V_{x,y,z}$ are the voltages applied to the ceramics X, Y and Z. This pattern generally represents the topography of the surface.

To obtain quality topographical images regardless of the nature of the tip-surface interaction chosen, a major difficulty has to be overcome: the elimination of all vibration. Any variation in the tip-surface distance due to extraneous vibration triggers a response in the servo-loop that produces spurious "relief" in the image. Elimination of these vibrations becomes very difficult when working at the atomic scale with measured relief amplitudes as low as a few picometres. Two systems have been devised to suppress vibration: suspensions to dampen all the vibrations from the laboratory, and very high stability of the mechanical systems that displace the sample towards the probe during the initial approach. These systems for millimetre-scale displacement are used to move the surface within range of the probe, because the maximum elongation of ceramic Z is 10 µm.

The extraordinary development of scanning tunnelling microscopy

Since its invention, scanning tunnelling microscopy has undergone an extraordinary development, owing to its simplicity and the quality of the information it yields (atomic resolution has been reached in many systems).

In a scanning tunnelling microscope (STM), the tip-surface interaction is measured by the number of electrons circulating by the tunnel effect between the metal probe and the conducting surface. This tunnel current, due to the quantum nature of the electrons (Box G) appears only when the probe-surface distance is no more than a nanometre, and when a low bias is applied to the junction ($V_t < 3V$). In a highly simplified description of this pair of electrodes, which are almost touching, the intensity of the tunnel current (I_t) varies exponentially with the tip-sample distance[2]. Typically, a current with an intensity of the order of a nanoampere (10^{-9} A) flows between the tip and the surface. It is this exponential fall-off in the current with increasing distance that accounts for the extremely fine spatial resolution

(2) A shown by the equation: $I_t = I_0 \exp(-2\kappa d)$, where d is the distance between the tip and the sample with κ constant and equal to ~ 0.1 nm⁻¹, and I_0 depends on V and the nature of the materials of the tip and the surface.

achieved with STM. The tunnel current increases tenfold when the tip approaches the surface by 0.1 nm! Another consequence of this dependence is that most of the electrons travel via the atom in the tip that is closest to the surface. This means that the useful size of the probe is in the range of a single atom.

After preparation under ultra-high vacuum the surface of a single copper crystal displays plane areas called terraces, separated by steps of monoatomic height (0.8 nm) (Figure 3). In this case, the measured amplitude of the relief due to the atoms is 10 pm (0.01 nm). STM also makes it possible to study surfaces immersed in a liquid. When the liquid is electrically conducting, protection to insulate the body of the tip from the liquid is necessary. Obviously, the tip end has to remain exposed to allow the electrons to tunnel. *In situ* studies in electrochemical cells show the growth, atom by atom, of a metal deposit on a gold surface. In the case of insulating liquids such as some hydrocarbons, the tip is immersed in the liquid and explores the liquid-solid interface with no special precautions being needed. For example, researchers at the CEA have worked with a dilute solution of fullerenes (C_{60}) in tetradecane ($C_{14}H_{30}$). An organized molecular layer of C_{60}, which partly covers a gold surface, forms spontaneously at the interface (Figure 4). The STM images of this layer show that the molecules of C_{60} form protrusions set 1 nm apart. In addition to imaging of the surface structure, scanning tunnelling spectroscopy (STS) can be used to measure, for example, the tunnel current as a function of the polarizing voltage applied between the tip and the substrate, in order to plot graphs of $I = f(V)$ for a set tip-sample distance. These measurements provide information on surface electron density with very high spatial resolution.

Power and versatility of atomic force microscopy

Atomic force microscopy, based on the measurement of force (*F*) or force gradient ($\partial F/\partial z$) between a probe and a surface has proved to be an extremely powerful and versatile technique. Unlike STM, which is limited to conducting samples, it allows the surface exploration of all types of solids.

The principle is straightforward. The free end of a thin cantilever of silicon or silicon nitride (typically L = 300 μm, l = 40 μm; e = 5μm), which acts as a spring, has a pointed tip that forms the probe. When the tip interacts with the surface, the force that develops bends the cantilever relative to its position away from the surface. This bend is often measured by the deviation of a fine laser beam. Knowing the amplitude of this deviation and the stiffness of the spring, it is easy to evaluate the force exerted between the tip and the surface. For a conducting tip and the surface of a non-magnetic metal, several forces contribute to the tip-surface interaction when the distance between tip and sample is shortened. Far from the surface, **forces** of attraction called **Van der Waals** forces predominate, and as the tip comes into contact with the surface, forces of repulsion take over. The forces in play between the tip and the surface thus vary nonmonotonously with distance. This variation makes it difficult to regulate the position of the
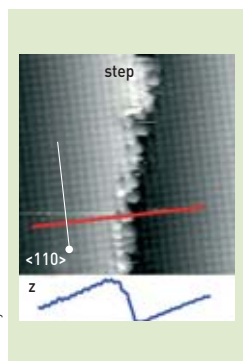

Figure 3.
STM image of a clean copper surface (001) under ultra-high vacuum at ambient temperature (area 6.3 nm x 6.3 nm, the z-axis is graduated in shades of grey). The surface is formed of plane terraces where the atoms, 0.26 nm apart, form a square lattice. The fourth-order symmetry of the atom lattice is clearly visible on these terraces. A cross-section along the red line shows the relief. The atomic monolayer step has a height of 0.18 nm while the relief associated with the atoms is 0.01 nm. The arrangement of the atoms at the step edge is ill-defined for two reasons: the imaging conditions are locally disturbed because several atoms on the tip end conduct the current, and the copper atoms move at ambient temperature. Note that STM is capable of positioning two macroscopic objects (the tip and the sample) with an accuracy of up to 10^{-12}m !
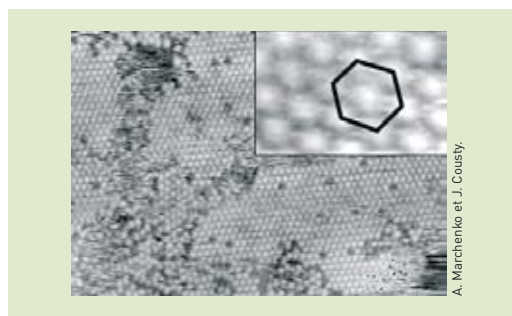

Figure 4.
STM image obtained at the interface between a liquid and a gold surface (area 50 x 36 nm²). The liquid contains molecules of C_{60} (fullerene), which are adsorbed and self-organized on the metal surface. Islands are formed with a compact hexagonal pattern of molecules 1 nm apart, as shown in the magnified view. Vacancies and defects in the molecular array are clearly visible. This molecular structure is comparable to that obtained when the same molecules are deposited in vacuum.

probe relative to the surface because certain force intensities correspond to two different equilibrium distances. Another difficulty is that there are forces that interact at long distances: the sides of the probe thus contribute to the forces detected by the spring. This is one of the reasons why it is more difficult to obtain atomic resolution with AFM than with STM. AFM can be used with different operating modes. The most common include the *contact mode* (where the tip touches the surface in the classical sense), and two *vibrating modes*, where the cantilever vibrates close to the surface at a frequency near its mechanical resonance.

Contact mode
The principle of this operating mode is shown in Figure 1 (centre), page 92. The tip is in repulsive contact with the surface. The intensity of the force applied by the tip on the surface, which is assumed to be rigid, depends on the stiffness of the spring and its deviation relative to its rest position. When the tip scans the surface at constant deviation of the spring the tip path reproduces the relief of the surface. However, the contact and the scanning movement cause friction. Experience has shown that a local variation of this force interferes with the mea-

Coupling of a scanning tunnelling microscope and an atomic microscope at CEA Grenoble, dedicated to research in organic and molecular electronics, mainly on the self-organization of molecular systems.
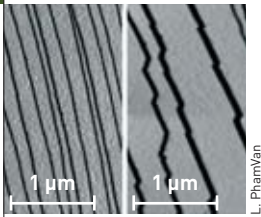
1 µm   1 µm

L. PhamVan

Figure 5.
AFM image in contact mode of the surface of a single alumina crystal showing its morphological evolution after cooling by annealing in air. After annealing at 1200 °C, the surface shows rectilinear steps (h ~ 1 nm) and terraces. Annealing at 1500°C for 4 h causes the formation of facets on the steps with a serrated profile.

surement of the spring deviation. Consequently, the AFM image modified by the local variations in the friction force can differ from the real relief on the sample. When measurements are made in air, the force that is actually exerted between the tip and the surface also depends on the ambient humidity. The hydrated film on the surface forms a meniscus, which introduces an additional force due to capillarity. Despite the limits of the technique, AFM in contact mode yields hitherto unattainable information, *e.g.*, the surface topography of dielectric materials, which are difficult to examine with conventional electron microscopy because the electron impact causes electrical charges to build up on them. AFM observations in air show that the surface of a crystal of aluminium oxide also displays terraces and steps. When annealed in air at 1500 °C, the density and shape of the steps evolve (Figure 5).

Vibrating modes
AFM operating in vibrating mode can overcome some of limits of the contact mode, in particular the difficulties arising from friction forces. In vibrating mode, the cantilever is mounted on a piezoelectric ceramic that is excited by applying an electric potential V[3]. The cantilever then oscillates at its mechanical resonance frequency (f_0) with an amplitude A_0 when the tip is away from the surface. When it interacts with the surface, the amplitude and frequency of the cantilever resonance are modified. There are thus two operating modes; one in which the amplitude is attenuated by a set amount, and another in which the frequency is shifted (Figure 1, bottom).
In the "amplitude modulation" mode (a), the average position of the tip relative to the surface is

■ (3) Where $V = V_0 \sin (2\pi f_0 t + \varphi)$

adjusted with the piezoelectric ceramic Z so that the amplitude of vibration of the cantilever is reduced by a set amount relative to A_0 at a set frequency. This operating mode is very robust when implemented in air, and yields reproducible results, because the hydrated film coating the samples interferes less than in the contact mode. For high attenuation values, the tip touches the surface during each cycle, which provides a good resolution in the z axis comparable to that obtained in the contact mode. This mode, sometimes called *tapping mode*, is widely used for studies in air and on fragile samples (*e.g.*, organic layers or biological objects). In addition, the simultaneous recording of the conditions of excitation of the cantilever yields information on the local variation of the tip-surface interaction caused by chemical heterogeneities. Thus both topographical and local chemical information can be obtained for a sample.
In the "frequency modulation" mode of cantilever vibration (b), the average position of the tip relative to the surface is adjusted with the piezoelectric ceramic Z so that the frequency shift relative to f_0 remains constant at a set amplitude. It was using this mode in a vacuum that atomic resolution was attained on various materials such as silicon, ionic crystals and alumina. A detailed analysis of the movement of the tip shows that the frequency shift is proportional to the local force gradient for low vibration amplitudes. However, the origin of the contrast in these high resolution images is still subject to debate (Figure 6). The AFM and STM obviously show the same atomic structure, but there are slight differences in contrast between the two, resulting from the different nature of the interactions in play between the tip and the surface.
In addition, the force or force gradient used to map the surfaces can result from different types of tip-

G The tunnel effect, a quantum phenomenon

Quantum physics predicts unexpected behaviour that defies ordinary intuition. The tunnel effect is an example. Take the case of a marble that rolls over a bump. Classical physics predicts that unless the marble has enough kinetic energy it will not reach the top of the bump, and will roll back towards its starting point. In quantum physics, a particle (proton, electron) can get past the bump even if its initial energy is insufficient, by "tunnelling" through. The tunnel effect makes it possible for two protons to overcome their mutual electrical repulsion at lower relative velocities than those predicted by classical calculations.
Tunnel effect microscopy is based on the fact that there is a finite probability that a particle with energy lower than the height of a potential barrier (the bump)

can still jump over it. The particles are electrons travelling through the space between two electrodes. These electrodes are a fine metal tip terminating in a single atom, and the metal or semiconductor surface of the sample. In classical physics a solid surface is considered as a well-defined boundary with electrons confined inside the solid. By contrast, in quantum physics each electron has wave properties that make its location uncertain. It can be visualized as an electron cloud located close to the surface. The density of this cloud falls off exponentially with increasing distance from the solid surface. There is thus a certain probability that an electron will be located "outside" the solid at a given time. When the fine metal tip is brought near the surface at a distance of less than a nanometre, the wave function associated with the electron is non-null on the other side of the potential barrier and so electrons can travel from the surface to the tip, and *vice versa*, by the tunnel effect. The potential barrier crossed by the electron is called the tunnel barrier. When a low potential is applied between the tip and the surface, a tunnel current can be detected. The tip and the surface being studied together form a local tunnel junction. The tunnel effect is also at work in Josephson junctions where a direct current can flow through a narrow discontinuity between two superconductors.
In a transistor, an unwanted tunnel effect can appear when the insulator or grid is very thin (nanometre scale). Conversely, the effect is put to use in novel devices such as Schottky barrier tunnel transistors and carbon nanotube assemblies.

surface interaction, which can provide further information after appropriate processing of the data. For example, with magnetic or dielectric materials that have trapped electrical charges, additional forces come into play between the surface and a specially-designed tip. By using specific techniques to separate the different contributions made by these additional forces, it is possible to map the magnetic or electric properties of the sample simultaneously with simple surface topography. This adaptability of AFM to new research applications continues to drive efforts in instrumentation development.

Local probe microscopy as a nanotool

The tip used in a local probe microscope, which is remote-controlled by the operator, can be used to modify a surface. The technique consists in first examining the surface (initial state), and then modifying it using a special procedure, and finally examining the resulting modifications. Several procedures have been used to sculpt a surface at the nanometre or even atomic scale. For example the tip can move atoms randomly **adsorbed** on the surface and arrange them in an artificial nanostructure. By means of a potential pulse the tip can decompose organometallic molecules to form a local deposit or extract atoms from the surface. Other modes of action have been explored that allow modifications to be made to a sample with varying degrees of control. However, these procedures are time-consuming because the point has to scan each nano-object to be made.
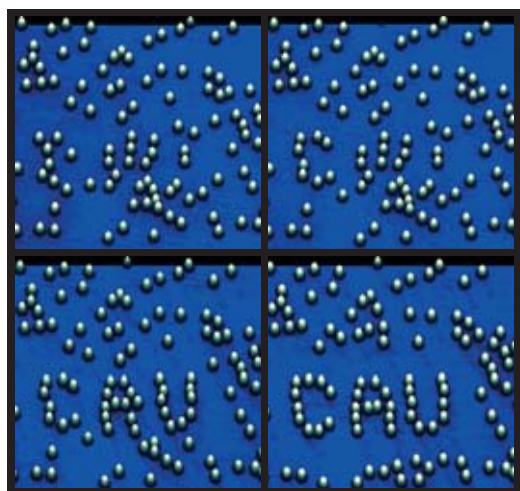
Series of STM images showing some steps in the ordering of manganese atoms on a silver surface cooled to 4.6 K, on which manganese atoms have been deposited (area 32 x 32 nm²). In Image 1, the white dots are atoms of manganese (Mn) distributed randomly on the surface. In the other images the operator uses the STM tip in two ways: 1) as a probe, to *image* the relief, and 2) as a tool to move certain atoms. This function is obtained by carefully bringing the tip closer to the surface than in imaging mode. In this way the interaction between an adsorbed atom and the tip increases, and the atom becomes attached to the tip and moves with it when it is displaced. The Mn atom thus 'slides' along the surface. Once it arrives at the chosen site, the tip is moved away from the surface, leaving the atom in place. This very delicate sequence must be repeated for every atom. Finally, the letters CAU (for Christian-Albrecht Universität) were drawn using a few dozen atoms in about one hour.

(Taken from Fig. 1b in Lantz et al., *Science 291* ; 2580–2583 (2001) with the author's permission © 2005 AAAS)
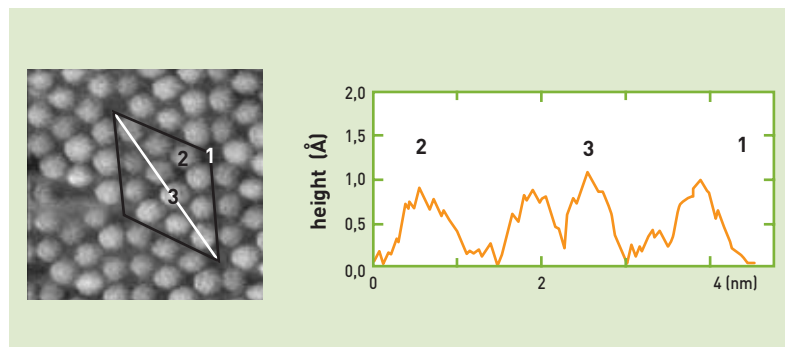
Figure 6.
Atomic resolution image of a surface (111) of silicon by AFM in frequency modulation mode obtained in a vacuum at ambient temperature. Each bump corresponds to a silicon atom. These surface atoms reorganize into a new large-lattice array because of the dangling bonds due to the presence of the surface. This lattice (black line) matches that obtained by STM. The cross-section along the white line shows the relief. However, there are differences in the relative heights of the atoms measured in the STM and AFM images. .

Further expected developments

The nanosciences and nanotechnology need to be able to carry out structural and physical-chemical investigations in the range between atomic and micrometre scales. In near-field microscopy, STM remains the prime method for the atomic-scale study of the surfaces of conducting samples or ultra-thin insulating layer deposited on a conducting substrate.
AFM microscopy has proved to be very versatile and yields a wide variety of types of information. In addition to the surface topography of all types of materials, magnetic, electrostatic and chemical force variations have already been successfully mapped. AFM should develop further; new probes that eliminate the optical detection of the spring deviation are currently being tested.
The principle of operation of local probe microscopy sets limits to its use as a tool for fabricating nano-objects or nanostructures. The tip, the nanotool, moves slowly along the surface, which entails a very long "milling" time. Various solutions have been put forward to reduce this time. For example, a comb consisting of several hundred individually actuated cantilevers allows parallel operation of nanotools, thereby reducing total processing time.

> **Jacques Cousty***
Materials Science Division
CEA Saclay Centre
(*) Assisted by Laurent PhamVan,
Christophe Lubin, François Thoyer
and Aldo Vittiglio

A From the macroscopic to the nanoworld, and vice versa...

In order to gain a better idea of the size of microscopic and nanoscopic* objects, it is useful to make comparisons, usually by aligning different scales, *i.e.* matching the natural world, from molecules to man, to engineered or fabricated objects (Figure). Hence, comparing the "artificial" with the "natural" shows that artificially-produced **nanoparticles** are in fact smaller than red blood cells.

Another advantage of juxtaposing the two is that it provides a good illustration of the two main ways of developing nanoscale systems or objects: ***top-down*** and ***bottom-up***. In fact, there are two ways

* From the Greek *nano meaning* "very small", which is also used as a prefix meaning a billionth (10^{-9}) of a unit. In fact, the **nanometre** (1 nm = 10^{-9} metres , or a billionth of a metre), is the master unit for nanosciences and nanotechnologies.

Artechnique

300-mm silicon wafer produced by the Crolles2 Alliance, an illustration of current capabilities using top-down microelectronics.

into the nanoworld: molecular manufacturing, involving the control of single **atoms** and the building from the ground up, and extreme miniaturization, generating progressively smaller systems.

Top-down technology is based on the artificial, using macroscopic materials that we chip away using our hands and our tools: for decades now, electronics has been applied using **silicon** as a substrate and what are called "**wafers**" as workpieces. In fact, microelectronics is also where the "top-down" synthesis approach gets its name from. However, we have reached a stage where, over and above simply adapting the miniaturization of silicon, we also
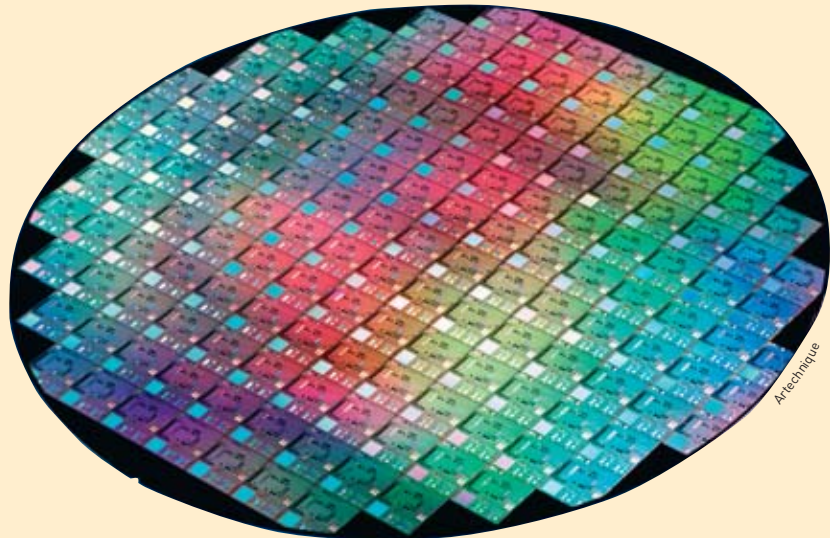
have to take on or use certain physical phenomena, particularly from **quantum** physics, that operate when working at the nanoscale.

The bottom-up approach can get around these physical limits and also cut manufacturing costs, which it does by using component **self-assembly**. This is the approach that follows nature by assembling molecules to create **proteins**, which are a series of amino acids that the super-molecules, *i.e.* **nucleic acids** (**DNA**, **RNA**), are able to produce within cells to form functional structures that can reproduce in more complex patterns. Bottom-up synthesis aims at structuring the material using

"building blocks", including atoms themselves, as is the case with living objects in nature. Nanoelectronics seeks to follow this assembly approach to make functional structures at lower manufacturing cost.

The **nanosciences** can be defined as the body of research into the physical, chemical or biological properties of nano-objects, how to manufacture them, and how they self-assemble by auto-organisazation.

Nanotechnologies cover all the methods that can be used to work at molecular scale to reorganize matter into objects and materials, even progressing to the macroscopic scale.

A (next)

natural world

bottom-up

molecule
a few Å

DNA ,
3.4 nm

virus,
0.1 µm

red blood cell,
5 µm

a hair, 50 µm
(diameter)

pollen grain,
10 µm to 20 µm

flea,
1 mm

ant,
1 cm

butterfly,
5 cm

man,
2 m

0.1 nm	1 nm	10 nm	100 nm	1 µm	10 µm	100 µm	1 mm	1 cm	10 cm	1 m

nanoworld

10^{-10} m	10^{-9} m	10^{-8} m	10^{-7} m	10^{-6} m	10^{-5} m	10^{-4} m	10^{-3} m	10^{-2} m	10^{-1} m	

atome,
0.1-1 nm

quantum dot,
5 nm

nanoparticle,
10 nm

nanotransistor,
20 nm

"Cooper"
transistor,
1 µm

integrated circuit
interconnects,
1 to 10 µm

microsystem,
10 to 100 µm

card chip,
1 cm

mobile phone,
10 cm

private vehicle,
2 m

top-down

artificial world

B A guide to quantum physics

Quantum physics (historically known as quantum mechanics) covers a set of physical laws that apply at microscopic scale. While fundamentally different from the majority of laws that appear to apply at our own scale, the laws of quantum physics nevertheless underpin the general basis of physics at all scales. That said, on the macroscopic scale, quantum physics in action appears to behave particularly strangely, except for a certain number of phenomena that were already curious, such as superconductivity or superfluidity, which in fact can only explained by the laws of quantum physics. Furthermore, the transition from the validating the paradoxes of quantum physics to the laws of classical physics, which we find easier to comprehend, can be explained in a very general way, as will be mentioned later.

Quantum physics gets its name from the fundamental characteristics of quantum objects: characteristics such as the angular momentum (spin) of discrete or discontinuous particles called quanta, which can only take values multiplied by an elementary quantum. There is also a quantum of action (product of a unit of energy multiplied by time) called Planck's cons-
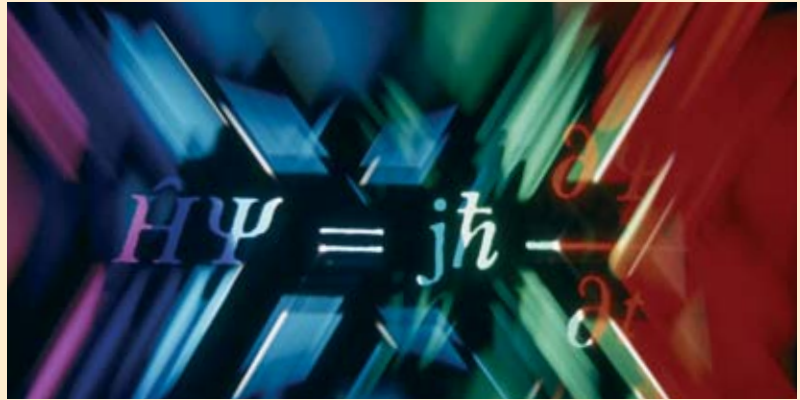
An "artist's impression" of the Schrödinger equation.

D. Sarraute/CEA

tant (symbolized as h) which has a value of 6.626×10^{-34} joule·second.

While classical physics separates *waves* from *particles*, quantum physics somehow covers both these concepts in a third group, which goes beyond the simple wave-particle duality that Louis de Broglie imagined. When we attempt to comprehend it, it sometimes seems closer to waves, and sometimes to particles. A quantum object cannot be separated from how it is observed, and has no fixed attributes. This applies equally to a particle - which in no way can be likened to a tiny little bead following some kind of trajectory - of light (photon)

or matter (electron, proton, neutron, atom, etc.).

This is the underlying feature behind the Heisenberg uncertainty principle, which is another cornerstone of quantum physics. According to this principle (which is more *indeterminacy* than *uncertainty*), the position and the velocity of a particle cannot be measured *simultaneously* at a given point in time. Measurement remains possible, but can never be more accurate than *h*, Planck's constant. Given that these approximations have no intrinsically real value outside the observation process, this simultaneous determination of both position and velocity becomes simply impossible.

B (next)

At any moment in time, the quantum object presents the characteristic of *superposing* several states, in the same way that one wave can be the *sum* of several others. In quantum theory, the amplitude of a wave (like the peak, for example) is equal to a **probability amplitude** (or probability wave), a complex number-valued function associated with each of the possible sates of a system thus described as quantum. Mathematically speaking, a physical state in this kind of system is represented by a **state vector**, a function that can be added to others *via* superposition. In other words, the sum of two possible state vectors of a system is *also* a possible state vector of that system. Also, the product of two vector spaces is also the sum of the vector products, which indicates **entanglement**: as a state vector is generally spread through space, the notion of local objects no longer holds true. For a pair of entangled particles, *i.e.* particles created together or having already interacted, that is, described by the *product* and not the *sum* of the two individual state vectors, the fate of each particle is linked - entangled - with the other, regardless of the distance between the two. This characteristic, also called *quantum state entanglement*, has staggering consequences, even before considering the potential applications, such as quantum cryptography or - why not? - teleportation. From this point on, the ability to predict the behaviour of a quantum system is reduced to probabilistic or statistical predictability. It is as if the quantum object is some kind of "juxtaposition of possibilities". Until it has been measured, the measurable size that supposedly quantifies the physical property under study is not strictly defined. Yet as soon as this measurement process is launched, it destroys the **quantum superposition** through the "collapse of the wave-packet" described by Werner Heisenberg in 1927. All the properties of a quantum system can be deduced from the equation that Erwin Schrödinger put forward the previous year. Solving the **Schrödinger equation** made it possible to determine the energy of a system as well as the **wave function**, a notion that tends to be replaced by the probability amplitude.

According to another cornerstone principle of quantum physics, the **Pauli exclusion principle**, two identical half-spin ions (**fermions**, particularly electrons) cannot simultaneously share the same position, spin and velocity (within the limits imposed by the uncertainty principle), *i.e.* share the same *quantum state*. **Bosons** (especially photons) do not follow this principle, and can exist in the same quantum state.

The coexistence of **superposition states** is what lends **coherence** to a quantum system. This means that the theory of **quantum decoherence** is able to explain why macroscopic objects, atoms and other particles, present "classical" behaviour whereas microscopic objects show quantum behaviour. Far more influence is exerted by the "environment" (air, background radiation, etc.) than an advanced measurement device, as the environment radically removes all *superposition of states* at this scale. The larger the system considered, the more it is coupled to a large number of degrees of freedom in the environment, which means the less "chance" (to stick with a probabilistic logic) it has of maintaining any degree of quantum coherence.

TO FIND OUT MORE:

Étienne Klein, *Petit voyage dans le monde des quanta*, Champs, Flammarion, 2004.

C Molecular beam epitaxy

Quantum wells are grown using Molecular Beam Epitaxy (from the Greek *taxi*, meaning order, and *epi*, meaning over), or MBE. The principle of this physical deposition technique, which was first developed for growing III-V semiconductor crystals, is based on the evaporation of ultra-pure elements of the component to be grown, in a furnace under ultra-high vacuum (where the pressure can be as low as $5 \cdot 10^{-11}$ mbar) in order to create a pure, pollution-free surface. One or more thermal beams of atoms or molecules react on the surface of a single-crystal wafer placed on a substrate kept at high temperature (several hundred °C), which serves as a lattice for the formation of a film called epitaxial film. It thus becomes possible to stack ultra-thin layers that measure a millionth of a millimetre each, *i.e.* composed of only a few atom planes.

The elements are evaporated or sublimated from an ultra-pure source placed in an effusion cell (or Knudsen cell; an enclosure where a molecular flux moves from a region with a given pressure to another region of lower pressure) heated by the Joule effect. A range of structural and analytical probes can monitor film growth *in situ* in real time, particularly using surface quality analysis and grazing angle phase transitions by LEED (*Low energy electron diffraction*) or RHEED (*Reflection high-energy electron diffraction*). Various spectroscopic methods are also used, including Auger electron spectroscopy, secondary ion mass spectrometry (SIMS), X-ray photoelectron spectrometry (XPS) or ultraviolet photoelectron spectrometry (UPS).
As *ultra-high-vacuum* technology has progressed, molecular beam epitaxy has branched out to be applied beyond III-V semiconductors to embrace metals and insulators. In fact, the vacuum in the growth chamber, whose design changes depending on the properties of the matter intended to be deposited, has to be better than 10^{-11} mbar in order to grow an ultra-pure film of exceptional crystal quality at relatively low substrate temperatures. This value corresponds to the vacuum quality when the growth chamber is at rest. Arsenides, for example, grow at a residual vacuum of around 10^{-8} mbar as soon as the arsenic cell has reached its set growth temperature.
The pumping necessary to achieve these performance levels draws on several techniques using ion pumps, cryopumping, titanium sublimation pumping, diffusion pumps or turbo-molecular pumps. The main impurities (H_2, H_2O, CO and CO_2) can present partial pressures of lower than 10^{-13} mbar.

D The transistor, fundamental component of integrated circuits

The first transistor was made in germanium by John Bardeen and Walter H. Brattain, in December 1947. The year after, along with William B. Shockley at Bell Laboratories, they developed the bipolar transistor and the associated theory. During the 1950s, transistors were made with silicon (Si), which to this day remains the most widely-used semiconductor due to the exceptional quality of the interface created by silicon and silicon oxide (SiO_2), which serves as an insulator. In 1958, Jack Kilby invented the integrated circuit by manufacturing 5 components on the same substrate. The 1970s saw the advent of the first microprocessor, produced by Intel and incorporating 2,250 transistors, and the first memory. The complexity of integrated circuits has grown exponentially (doubling every 2 to 3 years according to "Moore's law") as transistors continue to become increasingly miniaturized.

The transistor, a name derived from transfer and resistor, is a fundamental component of microelectronic integrated circuits, and is set to remain so with the necessary changes at the nanoelectronics scale: also well-suited to amplification, among other functions, it performs one essential basic function which is to open or close a current as required, like a switching device (Figure). Its basic working principle therefore applies directly to processing binary code (0, the current is blocked, 1 it goes through) in logic circuits (inverters, gates, adders, and memory cells).

The transistor, which is based on the transport of electrons in a solid and not in a vacuum, as in the electron tubes of the old triodes, comprises three electrodes (anode, cathode and gate), two of which serve as an electron reservoir: the source, which acts as the emitter filament of an electron tube, the drain, which acts as the collector plate, with the gate as "controller". These elements work differently in the two main types of transistor used today: bipolar junction transistors, which came first, and field effect transistors (FET).

Bipolar transistors use two types of charge carriers, electrons (negative charge) and holes (positive charge), and are comprised of identically doped (p or n) semiconductor substrate parts
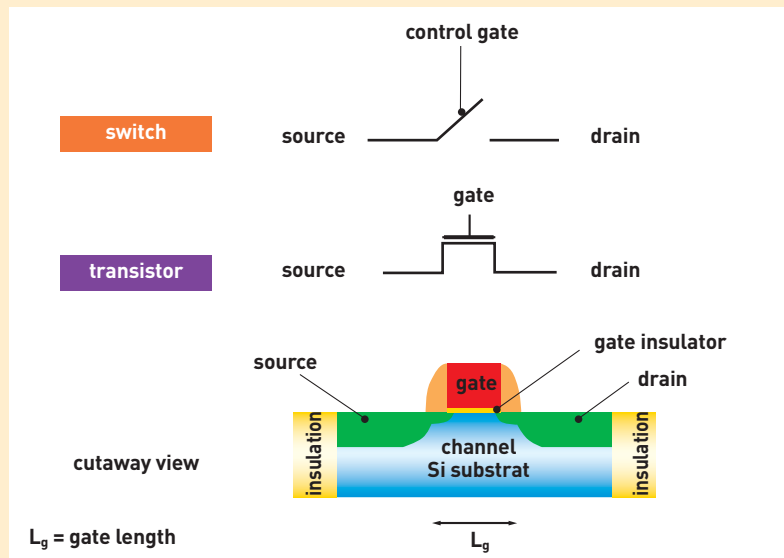

Figure.
A MOS transistor is a switching device for controlling the passage of an electric current from the source (S) to the drain (D) via a gate (G) that is electrically insulated from the conducting channel. The silicon substrate is marked B for Bulk.

separated by a thin layer of inversely-doped semiconductor. By assembling two semiconductors of opposite types (a p-n junction), the current can be made to pass through in only one direction. Bipolar transistors, whether n-p-n type or p-n-p type, are all basically current amplifier controlled by a gate current[1]: thus, in an n-p-n transistor, the voltage applied to the p part controls the flow of current between the two n regions. Logic circuits that use bipolar transistors, which are called TTL (for transistor-transistor logic), consume more energy than field effect transistors which present a zero gate current in off-state and are voltage-controlled.

Field effect transistors, most commonly of **MOS** (metal oxide semiconductor) type, are used in the majority of today's **CMOS** (C for complementary) logic circuits[2]. Two n-type regions are created on a p-type silicon crystal by doping the surface. These two regions, also called drain and source, are thus separated by a very narrow p-type space called the **channel**. The effect of a positive current on the control electrode, naturally called the **gate**, positioned over the semiconductor forces the holes to the surface, where they attract the few mobile electrons of the semiconductor. This forms a conducting channel between source and drain (Figure). When a negative voltage is applied to the gate, which is electrically insulated by an oxide layer, the electrons are forced out of the channel. As the positive voltage increases, the channel resistance decreases, letting progressively more current through.

In an integrated circuit, transistors together with the other components (**diodes**, condensers, resistances) are initially incorporated into a "chip" with more or less complex functions. The circuit is built by "sandwiching" layer upon layer of conducting materials and insulators formed by **lithography** (Box E, *Lithography, the key to miniaturization*). By far the most classic application of this is the microprocessor at the heart of our computers, which contains several hundred million transistors (whose size has been reduced 10,000-fold since the 1960s), soon a billion. This has led to industrial manufacturers splitting the core of the processors into several subunits working in parallel!

The very first transistor.
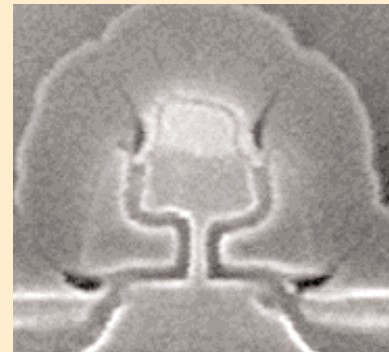
8 nanometre transistor developed by the Crolles2 Alliance bringing together STMicroelectronics, Philips and Freescale Semiconductor.

(1) This category includes **Schottky transistors** or **Schottky barrier transistors** which are field effect transistors with a metal/semiconductor control gate that, while more complex, gives improved charge-carrier mobility and response times.

(2) Giving **MOSFET** transistor (for Metal Oxide Semiconductor Field Effect Transistor).

E Lithography, the key to miniaturization

Optical lithography (photolithography) is a major application in the particle-matter interaction, and constitutes the classical process for fabricating integrated circuits. It is a key step in defining circuit patterns, and remains a barrier to any future development. Since resolution, at the outset, appears to be directly proportional to wavelength, feature-size first progressed by a step-wise shortening of the wavelength λ of the radiation used.

The operation works via a reduction lens system, by the *exposure* of a photoresist film to energy particles, from the ultraviolet (UV) photons currently used through to X photons, ions, and finally electrons, all through a mask template carrying a pattern of the desired circuit. The aim of all this is to transfer this pattern onto a stack of insulating or conducting layers that make up the mask. These layers will have been deposited previously (the *layering* stage) on a wafer of semiconductor material, generally silicon. After this process, the resin dissolves under exposure to the air (*development*). The exposed parts of the initial layer can then be etched selectively, then the resin is lifted away chemically before deposition of the following layer. This lithography step can take place over twenty times during the fabrication of an integrated circuit (Figure).

In the 1980s, the microelectronics industry used mercury lamps delivering near-UV (g, h and i lines) through quartz optics, with an emission line of 436 nanometres (nm). This system was able to etch structures to a feature-size of 3 microns (µm). This system was used through to the mid-90s, when it was replaced by excimer lasers emitting **far-UV** light (KrF, krypton fluoride at 248 nm, then ArF, argon fluoride at 193 nm, with the photons thus created generating several electronvolts) that were able to reach a resolution of 110 nm, pushed to under 90 nm with new processes.

In the 1980s, the CEA's Electronics and Information Technology Laboratory (Leti) pioneered the application of lasers in lithography and the fabrication of integrated circuits using excimer lasers, and even the most advanced integrated circuit production still uses these sources.

Photolithography section in ultra-clean facilities at the STMicroelectronics unit in Crolles (Isère).

The next step for high-volume production was expected to be the F_2 laser (λ = 157 nm), but this lithography technology has to all intents and purposes been abandoned due to complications involved in producing optics in CaF_2, which is transparent at this wavelength. While the shortening of wavelengths in exposure tools has been the driving factor behind the strong resolution gain already achieved, two other factors have nevertheless played key roles. The first was the development of polymer-lattice photoresists with low absorbance at the wavelengths used, implementing progressively more innovative input energy reflection/emission systems. The second was enhanced optics reducing diffraction interference (better surface quality, increase in numerical aperture).

Over the years, the increasing complexity of the optical systems has led to resolutions actually *below* the source wavelength. This development could not continue without a major technological breakthrough, a huge step forward in wavelength. For generations of integrated circuits with a lowest resolution of between 80 and 50 nm (the next "node" being at 65 nm), various different approaches are competing to offer particle projection at ever-shorter wavelengths. They use either "soft" X-rays at extreme ultraviolet wavelength (around 10 nm), "hard" X-rays at wavelengths below 1 nm, ions or electrons.

The step crossing below the 50 nm barrier will lead towards low-electron-energy (10 eV)-enabled nanolithography with technology solutions such as the scanning tunnelling microscope and molecular beam epitaxy (Box C) for producing "superlattices".
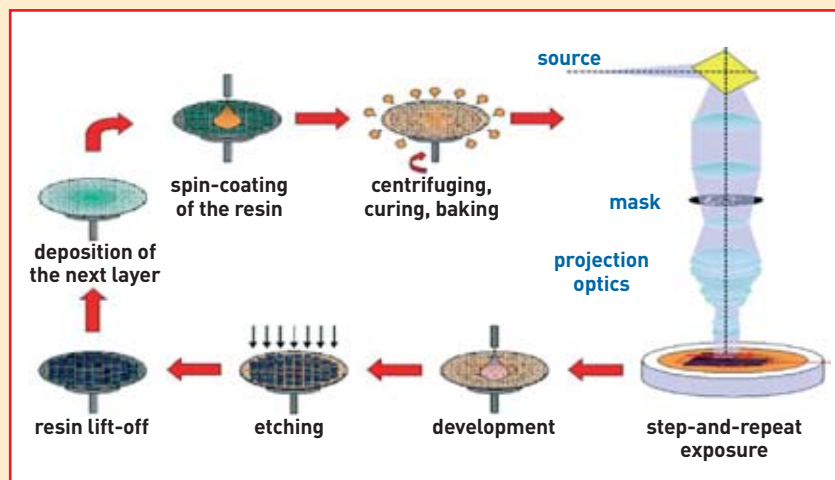

Figure. The various phases in the lithography process are designed to carve features out of the layers of conducting or insulating materials making up an integrated circuit. The sequences of the operation are laying of a photoresist, then projecting the pattern on a mask using a reduction optics system, which is followed by dissolution of the resin that is exposed to the light beam (development). The exposed parts of the initial layer can then be etched selectively, then the resin is lifted away before deposition of the following layer.

G The tunnel effect, a quantum phenomenon

Quantum physics predicts unexpected behaviour that defies ordinary intuition. The **tunnel effect** is an example. Take the case of a marble that rolls over a bump. Classical physics predicts that unless the marble has enough kinetic energy it will not reach the top of the bump, and will roll back towards its starting point. In quantum physics, a particle (**proton, electron**) can get past the bump even if its initial energy is insufficient, by "tunnelling" through. The tunnel effect makes it possible for two protons to overcome their mutual electrical repulsion at lower relative velocities than those predicted by classical calculations.

Tunnel effect microscopy is based on the fact that there is a finite probability that a particle with energy lower than the height of a potential barrier (the bump) can still jump over it. The particles are electrons travelling through the space between two **electrodes**. These electrodes are a fine metal tip terminating in a single **atom**, and the metal or **semiconductor** surface of the sample. In classical physics a solid surface is considered as a well-defined boundary with electrons confined inside the solid. By contrast, in quantum physics each electron has wave properties that make its location uncertain. It can be visualized as an electron cloud located close to the surface. The density of this cloud falls off exponentially with increasing distance from the solid surface. There is thus a certain probability that an electron will be located "outside" the solid at a given time. When the fine metal tip is brought near the surface at a distance of less than a **nanometre**, the **wave function** associated with the electron is non-null on the other side of the potential barrier and so electrons can travel from the surface to the tip, and *vice versa*, by the tunnel effect. The potential barrier crossed by the electron is called the **tunnel barrier**. When a low potential is applied between the tip and the surface, a **tunnel current** can be detected. The tip and the surface being studied together form a local **tunnel junction**. The tunnel effect is also at work in **Josephson junctions** where a direct current can flow through a narrow discontinuity between two **superconductors**.

In a **transistor**, an unwanted tunnel effect can appear when the insulator or **grid** is very thin (nanometre scale). Conversely, the effect is put to use in novel devices such as **Schottky barrier tunnel transistors** and **carbon nanotube** assemblies.