Scanning Probe Microscopy

Lydia Alvarez¹ and J. M. Siqueiros²

¹ Instituto de Ingeniería, Universidad Autónoma de Baja California, Blvd. Benito Juárez esq. Calle de la Normal, Col. Insurgentes Este, CP 21280, Mexicali, B.C. México
² Centro de Nanociencias y Nanotecnología, Universidad Nacional Autónoma de México, Km. 107 Carretera Tijuana-Ensenada, Apartado Postal 356, 22800 Ensenada B.C. México.

Almost thirty years ago, microscopy experienced a paradigm shift. Microscopy had been classically regarded as a method to expand our vision power and, for centuries, had maintained the same methodology: the radiation scattered by an object was collected by a detector and placed at an appreciable distance from it. Although the change of electromagnetic waves to wavelike electrons had resulted in a large increment in resolution, it was not until recently that microscopy as an expansion of our sense of touch has begun to be considered. The idea had been explored as early as eighty years ago by Synge, who discussed it within the context of electromagnetic waves. It was understood that the limits of optical microscopy had already been reached, but Synge pointed out that the Rayleigh-Abbe diffraction limit was only valid if light was collected in the far-field. By the approximation of a small aperture to the source, we can collect light before it experiences diffraction and obtain a resolution as good as the aperture is small. From the beginning, two major difficulties were apparent: that a small aperture would only transmit an infinitesimal amount of light and that there was no known technique to hold the aperture at a distance short enough to avoid diffraction effects. The second problem was solved when the way to manipulate piezoelectric materials was learned. However, it was not an aperture as Synge predicted but a metal probe which was held in close proximity to a sample. It was not the intention to collect light before it diffracted but to collect electrons by virtue of the tunneling quantum effect. Nevertheless, the result was similar to that predicted by Synge, a new technique of microscopy with resolution limited only by the size of the probe: Scanning Tunneling Microscopy. It did not take much time for useful variations to be found: two metals are needed to have quantum tunneling but any two materials experience forces when they are brought close enough to each other, which is the physical principle behind Atomic Force Microscopy. The proposal of Synge was also successfully tried a little later, leading to Scanning Near-Field Optical Microscopy in its different configurations. All these techniques share characteristics and are grouped under the label Scanning Probe Microscopy. Such common characteristics include the piezoelectric control of the position of the probe, the construction of an image pixel by pixel with the help of suitable software and, most important, the collection of information based on an interaction that can be only detected at nanometric distances. Scanning Probe Microscopy is one of the oldest nanotechnologies and will be the support for most of the nanotechnologies that will be developed in the future. In this work we intend to give a general description of this group of technologies, discussing separately their different approaches, their relationships and their prospects for the future.

Keywords scanning; probe

1. Origins of microscopy

The historical origins of microscopy are hazy, but evidence exists that in classical Greek and Roman times, globes filled with water or even pieces of glass were used to improve vision and concentrate sunlight. However, the first notice of the use of spectacles is found in the Florence of the 13th century. High quality single lens microscopes were made by Anton van Leeuwenhoek in the 17th century [1].

The compound microscope was invented in the 16th century by Zacharias Jansen, but the lenses in his design, required to be skillfully fabricated because in this approach any imperfection was doubled. The difficulty of fabricating lenses that were good enough turned the market to the single lens microscope of Leeuwenhoek for many years. However, as the quality of the lenses improved, Jansen’s design proved superior, being able in theory to reach any desired magnification.

A modern optical microscope consists of three different stages, each consisting of the combination of several lenses. The first stage is the illumination system consisting of a condenser lens and an aperture which limit and focus light into the plane of the object being observed. The objective, an optical system with a relatively short focal length and therefore a large magnification, creates a virtual image of the object close to the focus of the next optical system, the eyepiece. The eyepiece further magnifies the image and places it in the infinite for the eye to observe it comfortably.

2. Resolution limit

A complex sample is made of many spatial frequencies, the periods of which are called pitches. Each frequency generates a set of orders of diffraction, as those shown in Fig. 1a. Abbe’s theory of image formation [2] establishes that in order to recover any given frequency we must be able to collect at least order one. Order one is diffracted at an angle \( \beta = \sin^{-1}\left(\frac{\lambda}{p}\right) \). For this order to be collected it should be at least within the half acceptance...
angle $\alpha = \sin^{-1}(NA)$, where NA means numerical aperture. If we assume that the minimum feature is half the minimum pitch, we can obtain the resolution limit for a microscope using coherent illumination as

$$R = \frac{\lambda}{2NA}$$

(1)

On the other hand, if we assume an infinitely thin feature we can approximate it by an impulse function, which is converted into an Airy’s disk by effect of the circular aperture stops in an optical system. By assuming incoherent illumination, the Airy disks of neighboring features will not interfere and we can observe the points just before their images overlap and before they are seen as a single point, as illustrated in Fig. 1b.

![Diagram of microscope principle](image)

**Fig. 1** (a) Conventional microscopy. (b) Rayleigh criterion

The Rayleigh criterion establishes that the resolution limit is reached when the maximum of one function coincides with the minimum of the other, giving the resolution limit for the case incoherent illumination of

$$R = \frac{0.61\lambda}{NA}$$

(2)

This result is of the same order of magnitude as that obtained by Abbe. It has been pointed out that the Rayleigh criterion is somewhat arbitrary and that it seems that we could resolve the features even if they are closer if we could detect the difference in intensity from maxima to minima. This is true, but if the light detector has a linear response, the increase is not actually significant.

Resolution beyond the Abbe-Rayleigh limit is known generically as superresolution. To accomplish this, some assumptions in the previous model should be challenged. We can place the detector so close to the sample as for it to be able to detect evanescent orders. We can also use a strongly nonlinear detector, which will enhance our ability to distinguish between maxima and minima, even if the evanescent orders are lost [3].

### 3. Looking for an alternative probe radiation

Without surpassing Abbe-Rayleigh limit the only way to increase resolution is to use the shortest wavelengths. For systems that work with visible light, the shorter wavelength possible will give a resolution of about 200 nm. If we go out of the visible spectrum and enter the ultraviolet range, we will find that these shorter electromagnetic waves interact with matter in a much different way than visible light and that they require a completely different technology.

Ultraviolet microscopy was not developed until 1905 by Kholer and Rohr [4], but the equipment was so expensive and the gain in resolution so small that their instrument never had much market. Nevertheless, the high intensity of their
illumination led them to discover the self-fluorescence of some samples, which were later the basis for the fluorescence microscope.

Even more complicated to use are X-rays for which the same characteristic that makes them interesting—passing through matter—makes them difficult to focus and therefore to find an application in microscopy. Sources of radiation intense enough were also a problem before synchrotron radiation was available. Nevertheless, the first commercial X-ray microscope was patented [5] by Sterling Newberry, an employee at General Electric in 1954. Today, a host of different approaches are found in the most important research laboratories of the world.

A line of research much more successful was the use of electrons as illuminating radiation. In 1928 Louis De Broglie hypothesized that if light has particle-like properties electrons should have wave-like properties and his guesswork was soon proved true. Technology for manipulating electron beams by means of electric and magnetic fields was already available and was used to create what are called electronic lenses. Ernst Ruska [6] was the developer of the first successful electron microscope for which he received half a Nobel Prize in 1986.

The approach of Ruska is nowadays referred to as Transmission Electron Microscope (TEM). In spite of its technology being completely different from that of the optical microscope, the truth is that each subsystem in one of them has a counterpart in the other and the schematic in Fig. 1a is still valid. The velocity of the electrons in the beam is associated with a wavelength in the same order of magnitude as that of X-rays, allowing much better resolutions than those obtained in an optical microscope. However, a thin electron beam will only be able to illuminate a small spot on the sample, which gives the TEM a field of view comparatively smaller than that of an optical microscope.

4. Forming the image by pieces

In microscopes like those represented in Fig. 1a, the image is formed at once in the detector, be it the human eye or a special plate sensitive to electrons. For this reason they are called conventional microscopes. The idea of collecting an image in fractions was used for the first time in the 19th century and was the best option to increase the field of view of an electron microscope. The approach of repeating the same imaging process in a series of spots on the sample until covering the desired area is called scanning. The application of scanning to electron microscopy led to the invention of the Scanning Transmission Electron Microscope (STEM) by von Ardenne in 1928 [6].

![Fig. 2](a) Scanning microscopy (b) Nyquist sampling (c) Subsampling

In a scanning microscope the detector only needs a single scalar measurement associated to each point, and if several orders are collected their individuality is not maintained. The arrangement can work perfectly collecting only the zero order. This invalidates both the Abbe analysis and the Rayleigh criterion, for which it is said that scanning microscopes are not diffraction limited.
Scanning microscopes are better analyzed using the Nyquist-sampling theorem. If we are sure that our sample does not contain any spatial frequency over $k_{\text{max}}$, we can reconstruct it from the data in the sample if the sampling interval is at least

$$s = \frac{\pi}{k_{\text{max}}} \quad (3)$$

This is the condition shown in Fig. 2a, where the repetitions of the Fourier transform of the sample originated by the sampling process just touch each other. If the sampling interval is larger, the repetitions of the Fourier transform will start overlapping and we cannot collect but a distorted signal that will not give us reliable information. This situation is shown in Fig. 2b.

Actually, the spatial frequencies contained in an unknown sample are probably infinite, requiring a sampling interval of zero for perfect reconstruction. Since this is not possible, there will always be overlapping of the repetitions of Fourier transforms and some level of distortion that will be difficult to estimate. Filtering of the signal upon sampling can help but there is no perfect filter, and some residual high frequency signals will remain, causing a distortion known as aliasing. A brute force strategy to improve the resolution of images is to use a sampling interval as small as possible.

Nevertheless, there are also physical limits to the reduction of the sampling interval. For example, in an electron microscope it is not possible for the intervals to be smaller than the area of interaction of the electron beam, which is always larger than the width of the beam. Therefore, we can take the width of the beam as the ultimate physical limit to the resolution of a scanning electron microscope.

The STEM of Ardenne was soon replaced by an extremely marketable offspring: the Scanning Electron Microscope (SEM). In a SEM, electrons do not go through the sample but are reflected from it, either by backscattering or secondary generation. This allows for the imaging of a wider variety of samples. The secondary electrons are perfectly suited to modulate the cathode ray beam of a TV monitor and form an image in real time. The X-rays generated by the sudden stopping of the electrons can be used to do in-situ chemical analysis of the sample [7].

5. Touching microscopes

Conventional microscopes enhance our power of sight, but sight is not always the best way to perceive a shape. Sight gives a 2D image and so does microscopy. However, even with all its disadvantages, touch has the ability to create 3D images. Microscopes that enhance our power of touch are also able to do that.

In the study of surfaces it was common to use a stylus to scratch a thin film to check its adhesion [8]. The systematic use of such a stylus together with a scanning system was soon found to be a convenient way to record the profile of a sample first in a chart recorder and then as digital data on a tape, as J. B. P. Williamson did in 1968 [9]. Profilometers, as described in Fig. 3b, evolved together with the electronic data processing technology and are still very popular.

Those early profilometers worked in the contact regime, where contact is understood as an approximation so close that a strong reaction force, most probably of electric and quantum origin, arises. The reaction force makes the travelling stylus follow what we call the topography of the sample. However, contact always carries the risk of alteration of the sample which leads to what are now called non-contact profilometers. In these instruments, a different interaction was used to make the stylus follow the topography.

In one of these schemes, the interaction used was field emission, which happens when an electric field large enough is applied to a metal. With the application of such voltage and the approximation between stylus and sample, a tunneling current could be established. If the stylus moves so as to maintain this current constant, it will follow the topography of the sample, same as if it made “physical” contact.

Such an instrument, called a Topographiner, was created in 1972 by Young, Ward and Scire [10]. The possibility of this instrument to operate with the stylus so close to the sample as to have direct quantum tunneling, not requiring a strong electric field, was demonstrated. However, images were not obtained because of difficulties in the electronic instrumentation. It was understood that resolution would be greatly improved with this mode of operation, being only limited by the radius of the stylus. However, an estimation of the physical limit of this radius was not available.

6. Scanning Tunneling Microscopy

In 1985, Binnig and Rohrer built an instrument notably similar to the Topographiner. In tune with other advances developed in that decade, this new instrument included computational acquisition of data and more sophisticated electronics [11]. These improvements allowed the new instrument to work with the quantum tunneling current, a feature of which the Topographiner was not capable. This instrument, represented in Fig. 3a, would have been regarded as a simple step forward in the long path to obtain better images if it not were for an amazing characteristic: it was able to reach atomic resolution.
It was predicted before that the limit of the resolution of any profilometer is the radius of the stylus tip and it happened that the stylus of this arrangement had only one atom at the tip. The breaking of the old paradigm establishing that atoms are impossible to see gave the Scanning Tunneling Microscope an instantaneous mediatic success, which was crowned with the other half of the Nobel Prize that was given to Ernst Ruska in 1986.

This series of events was the beginning of what can be called the Scanning Probe Microscopy Age. Contrary to the inventors of the Topographiner who had worked with severe budget limitations, now resources were available to attempt any possible variation of the original design. It was enough to think of a different property that could be probed with a stylus to conceive a completely new instrument.

In a profilometer the stylus moves in response to the reaction of the surface in a manner analogous to blind people using their hands to feel the shape of one object. Since the force used by the profilometer is the same that upon being constant defines what we call topography, the profilometer works by its own nature in a mode of constant force and renders the topography of the sample.

When an STM works analogous to a profilometer, the stylus moves up and down in order to keep constant the tunneling current. It is said that it is working in the mode of constant current, which is true but when we say that it is rendering the topography of the sample, this is only an approximation. Tunneling current is not the physical quantity that we use to define contact, so it could part from what we understand as topography. The clearest example is what happens on a flat surface when its composition changes suddenly from one element with low density of electron states to one with a high density. Obviously, the stylus will sense the increase in tunneling current and will move up, although there is no topographic feature corresponding to this movement.

In spite of these precisions, the mode of operation where the stylus moves to keep any parameter constant is known both as mode of constant signal and as topographic mode. This is the reason that images from SPMs require usually trained people for their correct interpretation. There are other difficulties associated with SPMs like the limited area that can be scanned and the difficulty in locating the spot of interest. These are the major reasons behind the fact that, despite their amazing resolutions, SPMs are still not able to replace conventional microscopy in most industrial applications.

7. Atomic Force Microscopy

The most obvious limitation of an STM is that it requires a conductive sample. In the search for an interaction available in any type of sample, force was the first candidate. However, it was known that profilometers did not have atomic resolution. The reason for this is that the vertical force applied to the stylus combined with a very small tip radius will make furrows on the sample before it is able to image it.
To solve this problem a much smaller stylus was placed at the end of a cantilever beam. However, this changed the up and down movement of the stylus to an up and down bending of the cantilever, which required a different scheme for its detection. In 1986, the research team responsible of STM placed a metallic stylus behind the back of a cantilever in such a way that any bending was translated into variations of the detected tunneling current which could be used to construct the topographic image of the sample, as shown in Fig. 4a. This was the birth of Atomic Force Microscopy (AFM) [12].

In spite of some general similarity, the forces detected by an AFM are not exactly the same as those detected by a profilometer. These forces are not those that we use to define contact and, like in an STM in relation to electron density, using an AFM requires careful interpretation so as not to confuse topography with force distribution.

In this shortcoming can be converted into an advantage by changing the mode of operation. If when the cantilever bends a compensating force is provided so that the signal is kept constant, the magnitude of the compensating force can be registered for each spot and used to make a map of the distribution of forces over the sample. The creation of such maps, which are not images in the conventional sense, is a unique characteristic of SPMs.

Today’s AFMs have found more practical alternatives to the tunneling current for the detection of the cantilever bending. The approach of focusing a laser on the back of the cantilever which is reflected toward a four-quadrant photodetector is found in several commercial instruments, as shown in Fig. 4b. Some of these commercial instruments include an option to make the cantilever vibrate at its resonant frequency [13]. The imaging can be accomplished by measuring the frequency shift caused by the forces on the surface, as shown in Fig. 4c. The use of this approach makes it unnecessary for the cantilever tip to be too close to the sample —contact mode, although it could still make contact with the sample intermittently —tapping mode.

8. Scanning Near-field Optical Microscopy

Since the development of electron and scanning electron microscopy, optical microscopy was somewhat neglected. It was considered that this technology had already reached its physical limit and, although it was a very well established tool, few could have expected some improvement. This was not the case with Synge [14], who in 1928 proposed a scheme to do optical microscopy overpassing the diffraction limit. He proposed to make light pass through a nanometric aperture and use it to illuminate a sample placed at a nanometric distance of it. The light passing through the sample should be focused using a conventional microscope into a photodetector, which converted it into an electrical current signal. The sample should be scanned in nanometric increments as to construct the whole image.
Synge identified correctly the most important difficulties with this approach: the fabrication of a nanometric hole in an opaque plate, the need of a flat sample, the required nanometric control of positions and the low intensity of the light transmitted through such a small aperture. By 1984, all of these difficulties had been solved and Dieter Pohl decided to build what Synge had proposed more than fifty years before.

What Pohl did was to press a metal-coated polished pipette against a sample as to create a small aperture at its apex and used it to illuminate a sample as flat as it was possible in his time. A piezoelectric system similar to that of STM was used to control the position of the illumination source and sweep it to scan the sample. An objective lens was used to focus the transmitted light into a state-of-the-art photomultiplier which could amplify the weak signal thus obtained.

The result, shown in Fig. 5a, was called first Optical Stethoscope and later Scanning Near-Field Optical Microscope (SNOM) [15].

The Synge/Pohl solution is not the only SPM that works with visible light. An arrangement similar to a profilometer and an STM is obtained by illuminating the back of a sample at the appropriate angle to reach the condition of total reflection. In this condition the sample could only support evanescent waves that are converted into propagating waves when total reflection is frustrated by an optical fiber, as shown in Fig. 3c. The analogy with the STM is close enough for the instrument to be called a Photon Scanning Tunneling Microscope (PSTM) [16].

PSTM does not use an aperture, and this approach was pushed further ahead with the development of the apertureless SNOM (aSNOM). In this instrument light is focused just in the apex of an AFM or STM probe, which will diffract a signal containing the fine structure on the sample, combined with light diffracted by other parts of the cantilever which forms the background noise as shown in Fig. 5b. Sophisticated electronics are required to extract the useful part in such a signal, but the arrangement can potentially reach resolutions similar to those of other SPMs [17]. However, no optical SPM has yet reached atomic resolution and the best of them are still in the order of tens of nanometers [18].

9. Conclusions

It is true that the invention of the SNOM was made possible by the previous developments of the STM and AFM, and that PSTM is nothing but another sibling of this family. However, the history of SNOM dates back and is in line with conventional optical microscopes. Because of this, the explanation of why SNOM was able to overpass the Rayleigh-Abbe diffraction limit became an issue, what did not happen when STM and AFM were invented.

It was believed that overpassing the diffraction limit was equivalent to violating Heisenberg’s uncertainty principle but Vigoreaux dispelled this misconception [19]. Evanescent waves are only evanescent in a direction normal to the surface while in the direction parallel to the surface they are propagating and are associated with very large spatial

Fig. 5 (a) Pohl’s Scanning Near-field Optical Microscope. (b) Apertureless Scanning Near-field Optical Microscope
frequencies. Large spatial frequencies correspond to large momentum uncertainties allowing for a position uncertainty almost as small as we could desire.

Evanescent modes each have a different rate of evanescence and the distance between probe and sample determines how many of them can be accessible and can, therefore, serve as the low pass filter required to satisfy the Nyquist criterion. Tunneling current and atomic force are also physical quantities that decay strongly with distance, but their analysis cannot be carried out in the framework of classical physics. It is possible that finding the connection between the classical approach to evanescent waves and the quantum approach for the other phenomena will provide insight about these nanoscale physical phenomena.

Scanning Probe Microscopes were second in the Top Ten of Advances in Materials Science of the last 50 years, as classified by Materials Today. It is notable that SPMs are only after the International Technology Roadmap for Semiconductors (ITRS), which in spite of their importance is not really a scientific discovery. SPMs are not only important as new amazing microscopy techniques, but for opening the possibility to probe the nanoworld, which no doubt will be hand in hand with any development in nanotechnology in the foreseeable future [20].

Acknowledgment One of us (LA) thanks to the support of the grant DGIP UABC 2399

References