

**Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology**

Seeing Atoms – the Scanning Tunneling Microscope

In 1983, Binnig and Rohrer and their co-workers¹ reported a new type of microscope called the *scanning tunneling microscope* which permitted the imaging of features on the atomic scale. This exciting discovery has opened an entirely new field, generally referred to as scanning probe microscopy, in which many different physical mechanisms are exploited to gain views of matter at atomic dimensions. In this section, we explore the basic physics behind the scanning tunneling microscope.

In the previous section, we noted (Eq. 8.21) that the probability of tunneling through a rectangular barrier depends exponentially on the width of the barrier. Specifically, we noted that the transmission coefficient $T(E)$ has the form

$$T(E) \propto e^{-\alpha L} \tag{1}$$

where L is the barrier width and where

$$\alpha = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \tag{2}$$

for an electron of energy E impinging on a barrier of height V_0 . This exponential dependence of tunneling probability on distance creates the opportunity to image atomic-scale features.

Figure 1 shows to the left a highly schematic and exaggerated sketch of an atomically rough sharp tip brought into proximity with an atomically rough surface of a sample below. Without apology for the coarse artwork, one can recognize that there is clearly a point of closest approach between the tip

¹G Binnig, H. Rohrer, C. Gerber and E. Weibel, *Phys. Rev. Lett.*, **50**, 120 (1983); See also their Nobel Laureate Address for 1986 at nobelprize.org

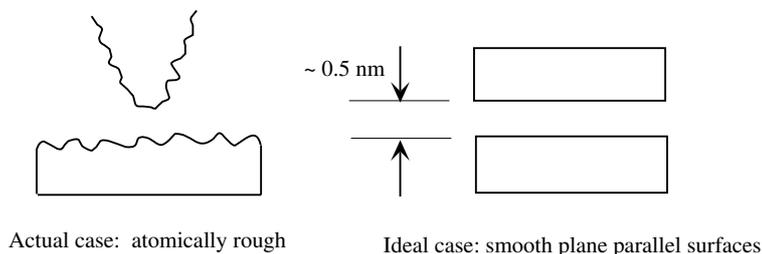


Figure 1: A highly schematic view, on the left, of an atomically rough tip in proximity to an atomically rough surface. On the right is shown an idealized one-dimensional equivalent with two perfectly flat parallel surfaces: something that cannot be realized in practice but which can be analyzed using the methods developed so far.

and the sample. Further, because of the exponential dependence of tunneling probability on distance, if there is to be tunneling between the two materials, most of the tunneling will occur at the point of closest approach. As the tip is moved laterally across the surface, this distance will vary and, hence, so will the tunneling current. In practice, as discussed below, an instrument is designed to maintain the amount of tunneling by moving the tip up and down as it is moved laterally. Either the variation in tunneling with lateral displacement or the vertical motion required to maintain a constant amount of tunneling can be used to image the atomic features of the sample.

We cannot analyze the complex three-dimensional structure. Instead, we shall use a one-dimensional model, as shown on the right of Fig. 1, with two perfectly smooth planar parallel surfaces separated by a small gap on the order of 0.5 nm. We now need to connect this physical picture to the tunneling calculation of the previous section.

0.1 A First Look at Metals

Figure 2 shows a schematic representation of the electron states in two identical metals separated by an air gap. The metal on the left might be the idealized tunneling tip while the right-metal might be the sample. The horizontal axis is position, while the vertical axis is electron energy. Our picture of a metal is that its electron states are those of particles in a very large box. These states form a quasi-continuum. The states are filled from lowest

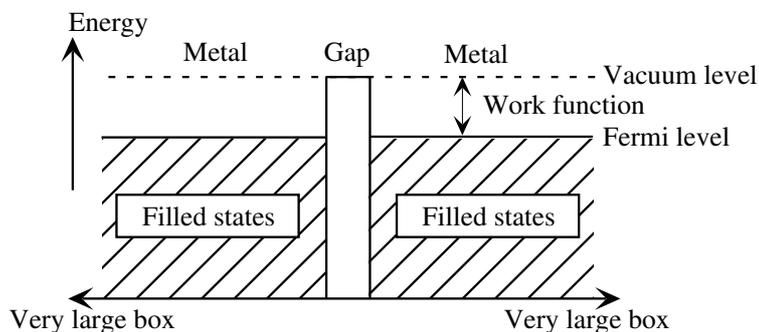


Figure 2: Schematic energy-level diagram for two identical metals separated by an air gap.

to highest following the *Pauli exclusion principle*, which states that each allowed electron state can be occupied by at most one electron. The number of filled states is determined by the electron density required by the metal valency. Thus the states are filled to some specific energy, called the *Fermi level*, but with a small transition zone of width $\approx 3k_B T$ around the Fermi level, where k_B is Boltzmann's constant (1.38×10^{-23} Joules/Kelvin) and T is the absolute temperature in Kelvins. In this transition zone, some of the states beneath the Fermi level are empty and some of the states above the Fermi energy are full.

The Fermi level is located below the *vacuum level* by an amount called the *work function*. The vacuum level is the energy that an electron would require to escape from the metal.² The work function, which is in the range 2-5 eV for typical metals, is a measure of how tightly the electrons are bound in the metal.

Looking again at Fig. 2, we notice that if there is to be tunneling, it will have to involve the states in the transition zone around the Fermi level because, by the Pauli exclusion principle, tunneling can occur only when an electron in an occupied state in one material can tunnel to an unoccupied state in the other material. This is, in fact, the case we studied in our one-dimensional tunneling analysis. An electron entering from the left and impinging on a barrier was not restricted by occupancy considerations from

²More precisely, the work function is the energy that an electron would require to escape from the metal in the absence of electrostatic interactions with the remaining electrons and atoms in the metal.

tunneling across the barrier to the other side. Figure 3 illustrates this situation. The horizontal arrows indicate tunneling from filled states on one material to empty states in the other material within a narrow transition zone about the Fermi level. In equilibrium, the amount of tunneling from left to right will exactly match that from right to left, so there is no net current in the structure.

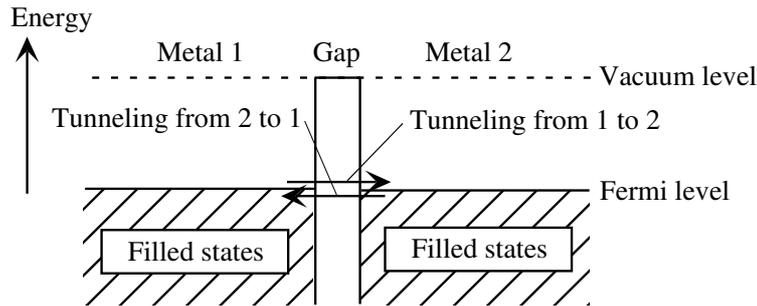


Figure 3: The sketch of Figure 2 showing equal and opposite tunneling between the two metals from energies near the Fermi level.

If we now apply a small voltage V between the two metals, the effect of this is to raise the Fermi level of one of the metals relative to the other, as shown in Fig. 4.

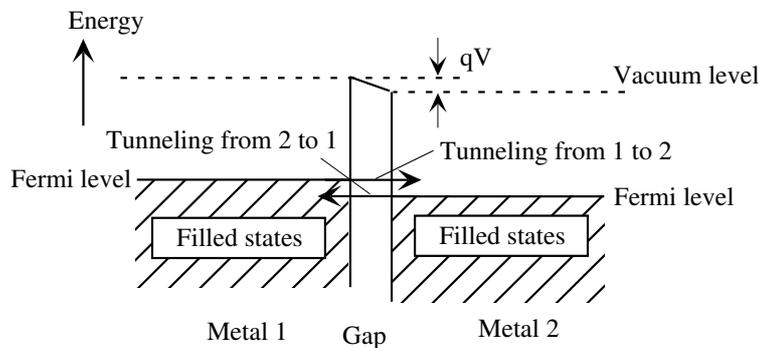


Figure 4: Tunneling between the two metals when a voltage V is applied.

In this non-equilibrium situation, the applied voltage shifts all of the electron states of one of the metals by an amount qV . The result is that there is now a significant unbalance in the occupancy. The left-hand metal,

with its higher Fermi level, has many occupied states at the same energy as, now, *unoccupied* states in the right-hand metal, but the corresponding states in the right-hand metal are now at an energy well below the Fermi level of the right-hand metal and these states are filled. Therefore, while there can be tunneling from left to right, from occupied to unoccupied, there cannot be much tunneling from right to left because the final states are heavily occupied. This unbalance results in a net electron current from the left to the right. This current can be measured with a picoammeter and used as the basis of an analytical instrument.

The net tunneling current, in addition to depending on the details of the quantum states in the vicinity of the Fermi level, carries the same exponential dependence on gap width L as the basic tunneling probability. To show how sensitive this can be, consider a current (at some fixed applied voltage) of the form

$$I = I_0 e^{-\alpha L} \quad (3)$$

where I_0 is a constant that depends on the specific metals and on the applied voltage. As a measure of sensitivity of this current to the tunnel distance, we calculate the following:

$$\frac{\Delta I}{I} = -\alpha L \frac{\Delta L}{L} \quad (4)$$

The factor αL , which is the ratio of the fractional change in current to the fractional change in gap width, is called the gauge factor of the tunneling phenomenon. For a typical value of $V_0 - E$ of 3eV and a gap spacing of 0.5 nm, this gauge factor is 6.6. This means that a 10% change in current, usually a readily measurable change, corresponds to a change of gap by only 1.6%. That is, L must change by only 8 picometers or 0.08 Angstroms – far less than a single atomic dimension – to product a 10% change in current. This sensitivity to positional variation is incredible!

To convert this positional sensitivity into a scanning instrument, a sharp tip is mounted on a frame that can be scanned horizontally above a sample while simultaneously being controlled vertically so that the tip-to-sample distance can be continuously varied to maintain constant current. In this configuration, the required vertical motion of the tip as a function of the lateral position forms an image that matches the atomic-scale topography of the surface.