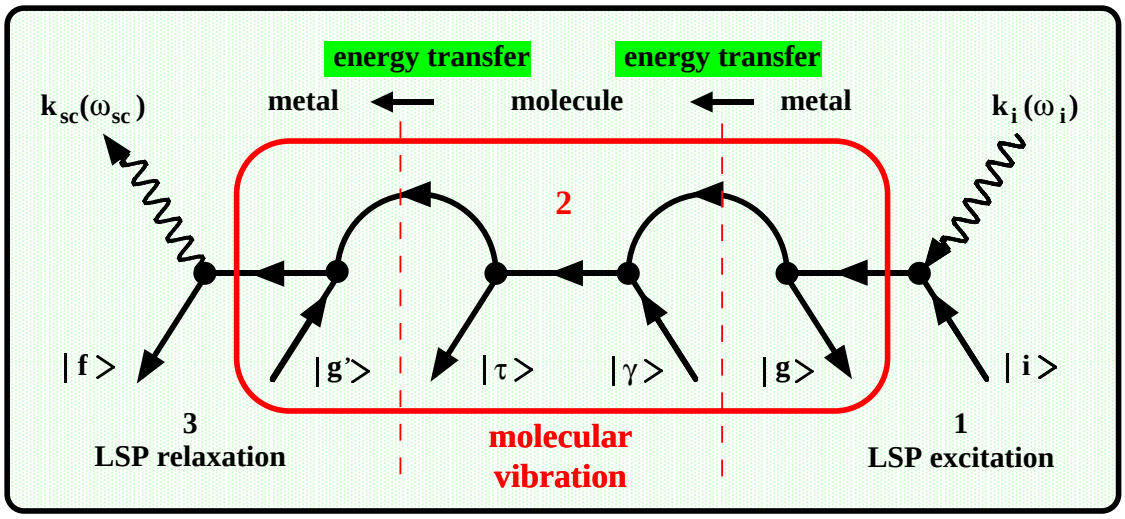
2. FUNDAMENTALS



2

2.1 Raman spectroscopy

A first understanding of the Raman effect may be obtained considering a stream of photons (*i.e.*, a beam of light) passing through a gaseous sample. Photons may collide with molecules and, in the case of an elastic collision, be deflected with unchanged energy. This process is called the Rayleigh (elastic) scattering. There is also a smaller probability that some energy will be exchanged during the collision, in which case an inelastic scattering occours, denoted as the Raman effect. The scattered photons may exhibit a lower energy, $\Delta E - h\nu_v$, if the molecule gains a quantum of energy $h\nu_v$ corresponding to one of its transitions, either rotational or vibrational. If the molecule is already in an excited vibrational (rotational) state, it can return to the ground state after the collision and the scattered photon has then an higher energy, $\Delta E + h\nu_v$.

Let us consider an electromagnetic field $\vec{E}$ associated with the radiation of frequency ν:

$$\vec{E} = \vec{E_0} \sin 2\pi\nu t. \tag{1}$$

This will generate an induced dipole moment $\vec{\mu}$ in the molecule:

$$\vec{\mu} = \alpha\vec{E} = \alpha\vec{E_0} \sin 2\pi\nu t \tag{2}$$

where α is the polarizability of the molecule, which is a second rank symmetric tensor. We see that the dipole moment will oscillate at the frequency ν and in doing so it will emit radiation at that same frequency, *i.e.*, this concept describes

the Rayleigh scattering. If the molecule is allowed to vibrate at the frequency ν_v of one of its normal modes, the atoms and electron cloud will be displaced during the oscillation, thus the polarizability will change periodically:

$$\alpha = \alpha_0 + \beta \sin 2\pi\nu_v t, \tag{3}$$

where α_0 is the static polarizability leading to the Rayleigh scattering. Then:

$$\vec{\mu} = (\vec{\alpha}_0 + \beta \sin 2\pi\nu_v t)\, \vec{E}_0 \sin 2\pi\nu t, \tag{4}$$

which can be decomposed into three terms:

$$\vec{\mu} = \underbrace{\alpha_0 \vec{E}_0 \sin 2\pi\nu t}_{\text{Rayleigh}} - \frac{1}{2}\beta\vec{E}_0\Big[\underbrace{\cos 2\pi(\nu+\nu_v)t}_{\text{Anti-Stokes}} - \underbrace{\cos 2\pi(\nu-\nu_v)t}_{\text{Stokes}}\Big]. \tag{5}$$

Our simple formula describes an induced dipole that emits three different types of radiation: Rayleigh scattering (first term, unchanged energy), Anti-Stokes Raman light (second term, higher energy) and Stokes Raman light (third term, lower energy). Incidentally, we may note that, if the transition does not cause a change in the polarizability $\beta \propto \frac{\delta\alpha}{\delta q_v} = 0$, no Raman light will be emitted. For a more rigorous description of the scattering processes, not only the molecular energy levels must be considered quantized but also the electromagnetic field (second quantization). Both the Rayleigh and the Raman scattering can be then derived in a more elegant way. The exhaustive (and rather lengthy) procedure is found in classic text books such as "The quantum theory of light" by Loudon [12]. Here we need to add just a few more considerations at an atomic level: the Raman event is not exclusively a vibrational process: it is better described as a vibronic one in which vibrational as well as electronic states are involved. To illustrate this point we refer to Fig. 2.1, where the molecule is represented as a three-level system: S_0 is the electronic ground state, split into two vibrational sub-levels, and S_1 its first excited state.

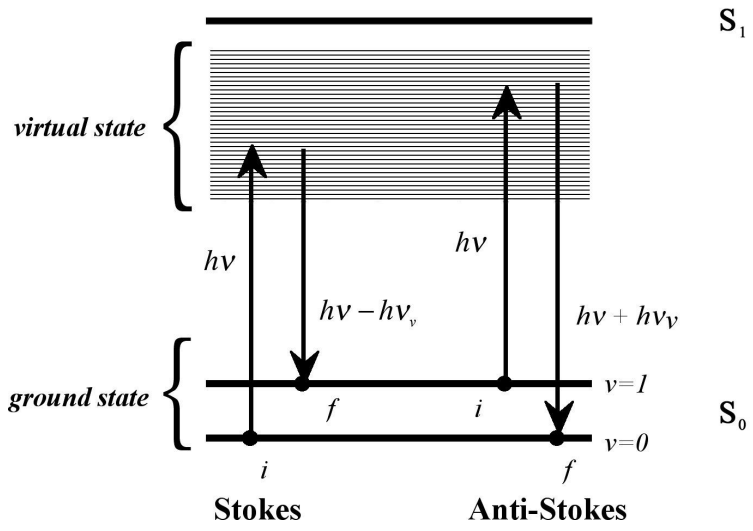


Figure 2.1: Scheme for the energy levels involved in the Raman process.

An electron in the ground state is excited into a so called "virtual" electronic state, actually still an S_1 state but not as well defined and existing only for the short time of the photon-molecule interaction. This excitation is governed by the Heisenberg uncertainty principle. The larger the energy mismatch ($\Delta E_{S_0-S_1} - h\nu$), the more simultaneous excitation and de-excitation processes must occur. Under such conditions, the excitation may be defined as virtual. In turn, the smaller the energy mismatch, the more resonant is the Raman process. Upon returning to the ground state, the electron arrives at a final (vibrational) state that may lie higher or lower in energy than the initial one. The energy available for the photon emission gives rise to the Stokes and anti-Stokes Raman lines, respectively.

2.2 Surface enhanced Raman spectroscopy

The main drawback of Raman spectroscopy is the extremely low cross section ($\frac{d\sigma}{d\Omega} \sim 10^{-28}$ cm^{-2} sr^{-1}) of the Raman event compared with the much higher prob-

ability of other optical processes such as the Rayleigh scattering or fluorescence. Thus, powerful exciting sources, very sensitive detectors and a large number of scattering molecules are needed in order to record a Raman spectrum.

The last constraint would make Raman spectroscopy rather unsuitable for surface studies, where usually one has to deal with a small amount of material, down to the monolayer level, or even less. In contrast to this expectation, some twenty years ago, at first Fleischmann [10, 13] and then Van Duyne *et al.* [11] reported huge Raman intensities observed for certain adsorbates at rough silver electrodes. Van Duyne concluded that the signal arises due to a tremendous enhancement of the Raman cross section for the adsorbed molecules. A suitable choice of the metal (copper, silver and gold being the most used) and a roughening of the surface turned out to be the key factors for the enhancement.

Since then, this effect is known as the surface enhanced Raman scattering (SERS) and has been widely studied [14–18]. Today, it is commonly accepted that it arises mainly due to two different mechanisms operating at the same time. In the literature, these are usually named the *electromagnetic enhancement* and the *chemical enhancement*; together, they may provide enhancement factors between 10^6 and 10^8; yet, the exact partition between the two is still a matter of debate.

2.2.1 Surface plasmon polaritons

The so-called electromagnetic effect involves the excitation of surface plasmon polaritons (SPP) in the substrate metal. Since these modes will play a major role in all the forthcoming discussion, it is worthwhile introducing them here.

As long the free electron model is consistent, a metal behaves like a plasma of an electron gas, the electrons being the mobile charge carriers, and the dielectric function may be expressed as:

$$\varepsilon(\omega) = 1 - \frac{Ne^2}{\varepsilon_0 m_e \omega^2} = 1 - \frac{\omega_{BP}^2}{\omega^2}, \tag{6}$$

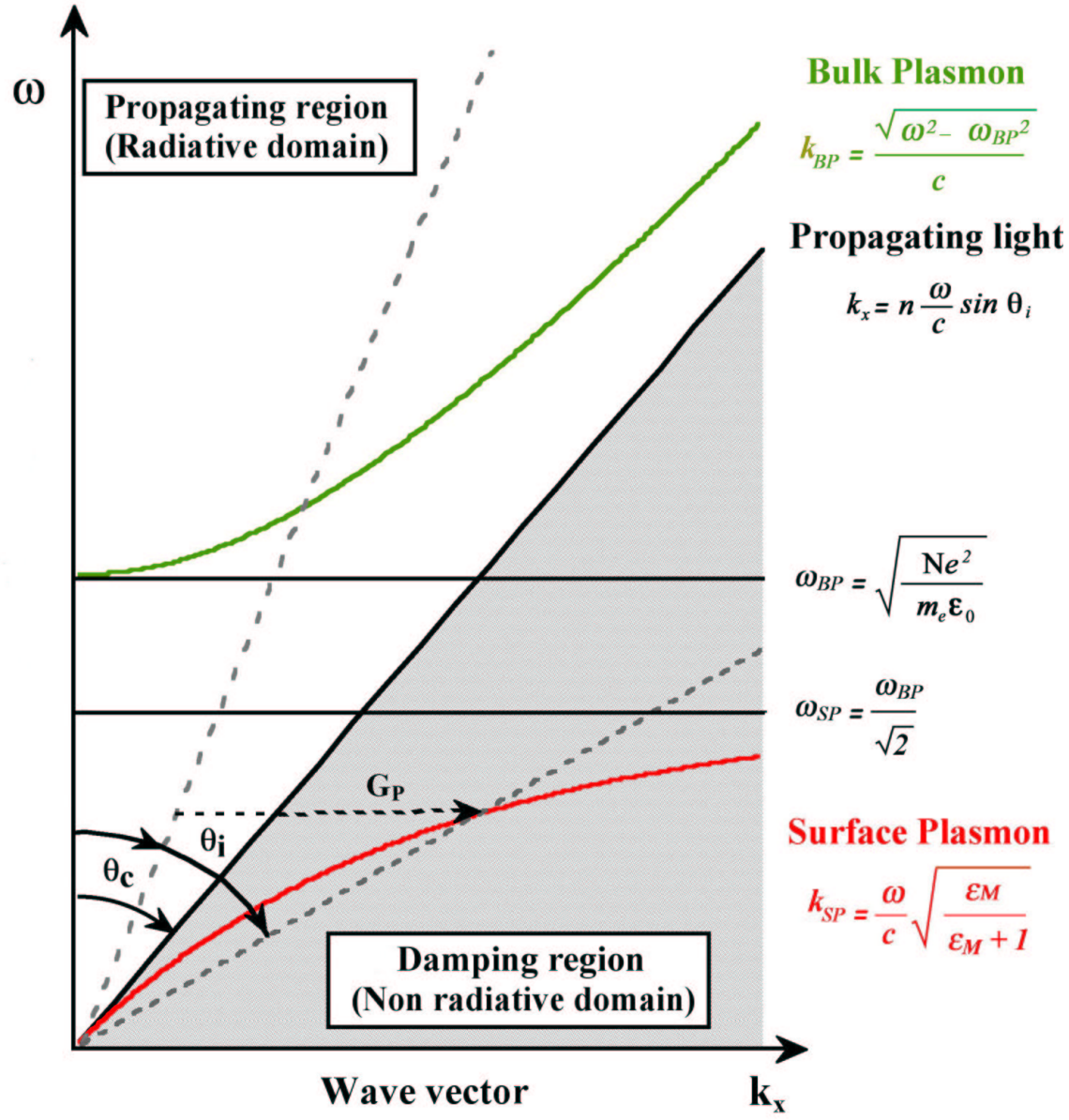


Figure 2.2: Dispersion curves for bulk and surface plasmons. Excitation of bulk plasmon in the metal is allowed at the frequency defined by the crossing point between the dispersion curve and the dotted line (angle of incidence $\vartheta_i < \vartheta_c$). Excitation of surface plasmon is normally forbidden (the required angle of incidence is larger than the critical angle ϑ_c).

$$\omega_{BP}^2 = \frac{Ne^2}{\varepsilon_0 m_e} \tag{7}$$

is defined as the (bulk) plasma frequency for the metal; ω is the frequency of light, e and m_e are the electron charge and mass, ε_0 the permittivity of free space and N is the electron density of the metal.

Using the general relation for the optical dispersion of an electromagnetic wave in an isotropic medium:

$$\varepsilon(\omega, k)\,\omega^2 = c^2 k^2 \tag{8}$$

the dispersion curve for the bulk plasma is obtained:

$$\omega^2 - \omega_{BP}^2 = c^2 k_{BP}^2. \tag{9}$$

The oscillating electron cloud of a metal behaves as a collection of coupled harmonic oscillators with a characteristic resonance frequency given by (7), where the plasmon polariton is the quasi particle of the excited mode.

This assumes so far a bulk (volume) plasma. Next, an air/metal phase boundary system with the metal surface at $z = 0$ and $\varepsilon(\omega) = \varepsilon_M(\omega)$ for $z < 0$ is considered. Solutions for the Laplace equation that fulfill the boundary conditions for the tangential components of the electric field and for the normal component of the electronic displacement exhibit the form:

$$E_z = E_{0z} \exp\left[i\left(kx - \omega t\right)\right] \exp\left(-k_z z\right), \tag{10}$$

the dispersion relationship reads as follows:

$$k_{sp} = \frac{\omega}{c}\sqrt{\frac{\varepsilon_M}{\varepsilon_M + 1}}. \tag{11}$$

The limiting frequency for the oscillating charges at the surface, the surface plasma frequency, is defined by:

$$\omega_{sp} = \frac{\omega_{BP}}{\sqrt{2}}. \tag{12}$$

The wave described in (10) is an evanescent wave, *i.e.*, it is a wave propagating along the interface and decaying exponentially in intensity as it penetrates the less dense medium. This is also seen in Fig. 2.2, where the dispersion curve of the surface plasma is confined below the propagating wave domain. It means that volume plasmons may be excited by any incident light with $\vartheta_i < \vartheta_c$ (the critical angle), if $k_x(\omega, \vartheta_i) = k_{BP}(\omega)$ (the red dotted line intersects the dispersion curve) but not surface plasmons, for which $k_x(\omega, \vartheta_i) < k_{sp}$. This is why they are usually described as radiative and non radiative processes, respectively.

Two different approaches are suitable to study surface plasmons by optical means. One is to work under the total internal reflection condition, so that an evanescent wave at the boundary is available for excitation. In this so-called Kretschmann-Raether configuration [19, 20] the reflectivity at a given wavelength of a thin metal film (usually Ag or Au) deposited on a prism with sufficiently high refractive index is measured as a function of the incident angle of the monochromatic light. The reflectivity follows the Fresnel formula: for p-polarisation and $\vartheta_i > \vartheta_c$, $R \simeq 100\%$ (total internal reflection) except at the angle at which the surface plasmons are excited and a clear dip in the reflectivity is seen. The angular location, width and depth of the absorption peak provides information about the flatness and thickness of the metal film but also on the dielectric constants of the metal and adjacent dieletric.

In the foregone discussion a planar surface was considered to derive the surface plasmon dispersion curves. Instead of enlarging the value of k_x with a prism, a sub-wavelength regular modulation of the surface structure may also provide the missing momentum to satisfy the law of momentum conservation:

$$k_x + G_p = k_{sp}. \tag{13}$$

where G_p is the reciprocal of the grating constant (such a possibility is often called the umklapp process [21]), so that also surface plasmon modes may

be excited and radiate. A number of different values for G_P may be provided by a rough surface, since in a first approximation, all the randomly distributed structures, as hills, bumps and crevices, may be described as a superposition of gratings. However, for a substantially rough surface, the plasmon resonances cannot be properly described by (11), which is valid only for flat surfaces or for the limit of superimposed small perturbations.

More suitable concepts are then introduced to better describe the electronic behavior and the corresponding electromagnetic field distribution in the proximity of subwavelength structures. Various models [22–30] have been proposed using spheres, ellipsoids, cylindrical rods or even arrays of particles or fractal structures to model the surface roughness. Later on we will discuss in more detail a particular model (a sphere placed close to a plane). Here we report briefly a first conclusion that may be drawn from all these theoretical efforts.

As said before, the surface plasmons are electromagnetic waves propagating along a flat surface. Any structure in their way acts as an hindrance or scattering center that can convert the delocalized mode into a localized one (LSP), the electromagnetic wave being confined into subwavelength structures. The local field strength, amplitude and the resonant condition associated with the charge oscillations are all governed by the size and geometry of each structure as well as by their spatial arrangement(s).

2.2.2 Electromagnetic enhancement

A diagram may be useful to visualize the electronic processes taking place at the activated, *i.e.*, roughened metal surface, due to illumination with a monochromatic light source (Fig. 1.3). The above mentioned LSP are excited by the absorption of a photon with energy $\hbar\omega_i$; the metal particle is now found in an excited state and will relax to the ground state (relaxation time, τ, in the order of 10^{-15} s) with the annihilation of the LSP mode and the creation of a second photon. This process is also a Rayleigh scattering with the surface structure acting as the scatterer of light: the frequency of the emitted photon remains the same as for the incident

photon $\omega_{sc} = \omega_i$, but the other components of the wave vector may have changed, $k_{sc} \neq k_i$.

A different optical process can arise if the surface is covered by an adsorbate [31]: the energy stored in the system (metal/molecule) may be transferred *via* dipole-dipole interaction to the molecule. Vibrational modes may be activated before a further energy transfer occurs, this time from the molecule back to the metal. Now the scattered photon has a different energy and the difference in energy is the quantity gained or lost by the molecule by its vibration. In that sense, we may speak of a Raman process (mediated by the metal surface).

As reported in the diagram, the whole process may be ideally divided into three resonant steps; it must be clear that the LSP take part twice (but with different energies): they are twice excited and twice de-excited. Compared with the normal resonance Raman process, where the molecule interacts directly with the incident electromagnetic field, the field experienced by the adsorbed molecule is the local field enhanced by the induced oscillations of the electrons in the metal particle. Since the cross section for the Raman process scales with the fourth power of the electromagnetic field enhancement at the molecule location [12], a moderately enhanced field gives an extremely enhanced Raman intensity: for example, a ten fold enhancement of the local field increases the Raman intensity by four orders of magnitude.

A rough surface is not the only substrate that supports SERS. Another excellent enhancing system is a colloidal solution [32–40]: a suspension of gold or silver particles, with dimensions between 20 nm and 200 nm, can yield huge Raman intensity for the adsorbate. Colloids exhibit a few straightforward advantages: they are easily made with some control of the size, size distribution and aggregation state; since the suspended system of particles resembles an array of isolated spheres (or dimers), the theoretical description for the field enhancement is greatly simplified: dye molecules adsorbed at Ag colloids in aqueous solution may be considered as a model system for SERS investigation, enabling an easier comparison between theory and experiment.

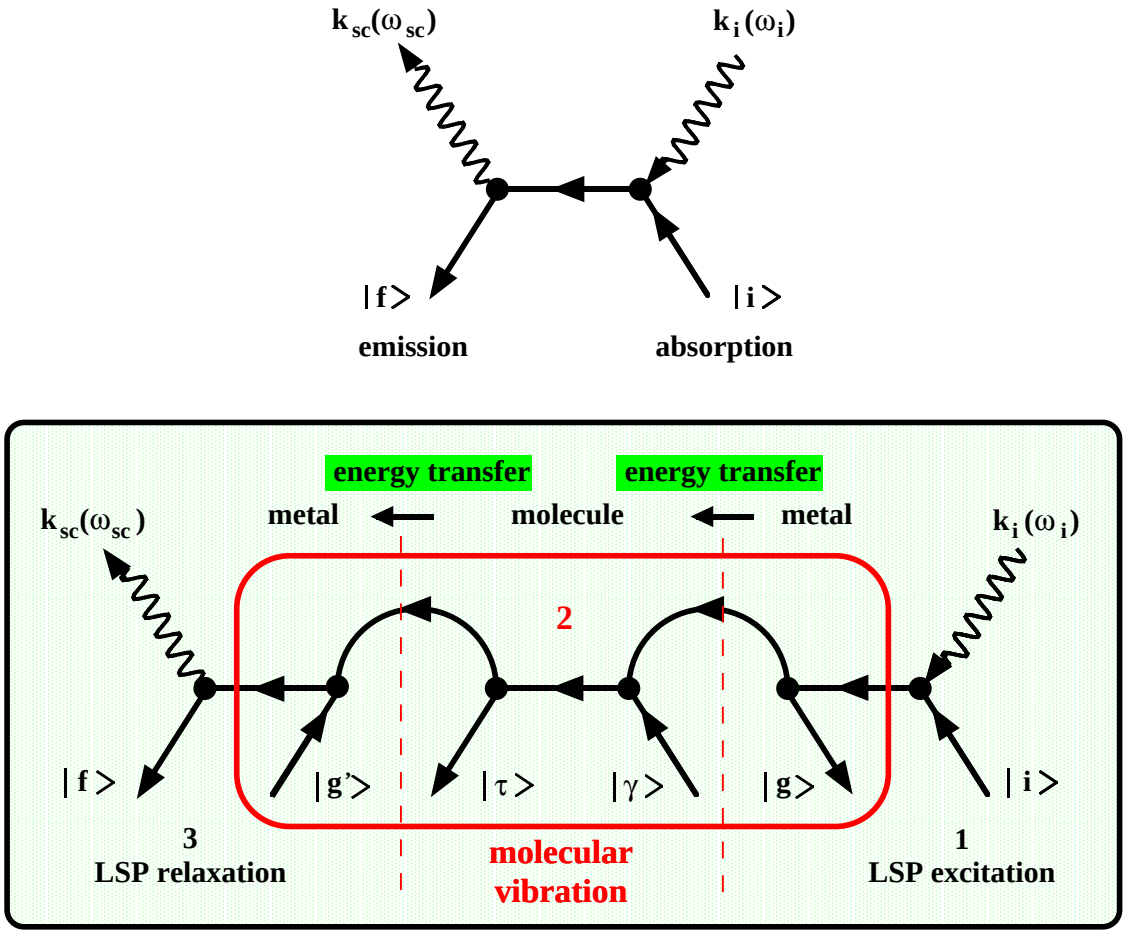


Figure 2.3: Above, Rayleigh scattering by a metal surface; below, Raman scattering enhanced by a metal surface (the surface is considered to be covered by a layer of molecules). To a first approximation, the total energy states are given by the product of the uncorrelated states of the molecule and the metal: $\mid i >= \mid g > \mid \tau >$, $\mid f >= \mid g' > \mid \gamma >$; $\mid \tau >$ and $\mid \gamma >$ are different vibrational states of the molecule, $\mid g >$ is the ground state for the metal.

2.2.3 Chemical enhancement

For species chemisorbed at metal surfaces there is a partial overlap of the molecular and metal orbitals. This can lead to the creation of a new set of energy states that are available for the Raman process. A simple mechanism, the "charge transfer process" [41, 42], has been proposed and experimentally confirmed. Consider an energy level diagram (Fig. 2.4) with the metal on the left and the molecule on the right. In the first step the incident radiation $\hbar\omega_i$ excites an electron within the metal conduction band to an empty state in the half filled band. This charge will be transfered from the metal to the molecule. E_{CT} is the electronic state given by the mixing of a molecule non-bonding state and an appropriate metal state. Clearly, E_{CT} exists only when the molecule is adsorbed at the surface. Once again, the whole process can be considered as a scattering process, involving an excited electron-hole pair. If the molecule is left in a different vibrational state, the e^--h^+ pair and, thus, the photon leaving our system has a (slightly) different energy. The final step is the radiative annihilation of the e^--h^+ pair[2].

If the metal acts as an electrode in an electrochemical environment, the position of the Fermi level E_F (relative to the zero energy at infinite) may be shifted at will by changing the applied voltage U. The optimal excitation energy for the overall resonance process is thus defined by the following:

$$\hbar\omega_i = E_{CT} - [E_F(0) + eU]. \tag{14}$$

Indeed, relation (14) has been confirmed by the observation of different (potential dependent) Raman intensity maxima for different excitation laser frequencies [42].

A final consideration about the two enhancement mechanisms discussed above: it is common to refer to the electromagnetic enhancement as a long range effect and to the chemical enhancement as a short range effect. Indeed, the latter is

[2]If the two discussed enhancing effects act cooperatively, the scattering of photons, as depicted by the CT mechanism, has to be replaced by the inelastic scattering of the LSP *via* CT process.

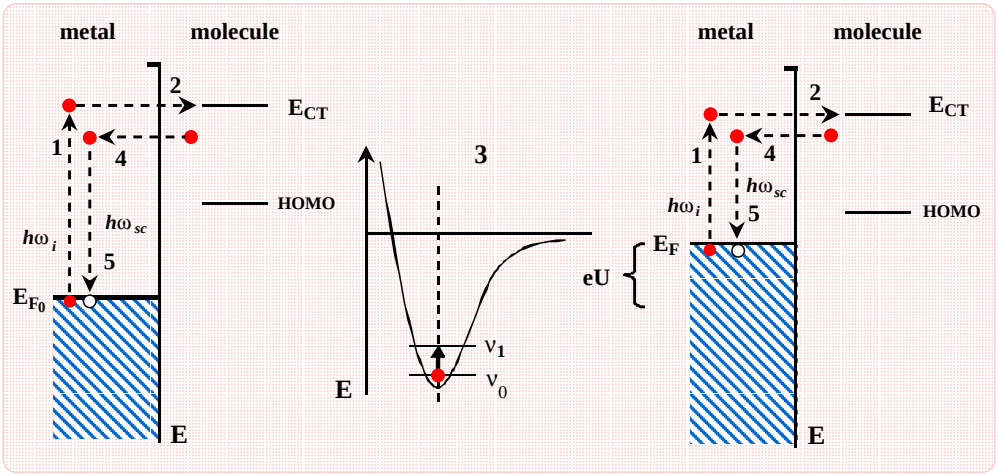


Figure 2.4: Energy diagram for a charge transfer process at the metal/adsorbed molecule system. On the right and left, two different positions for the Fermi level are shown; in the middle, the Morse potential for the state E_{CT} is schematically reported.

mainly effective for the first layer of the adsorbed species and drops quickly to zero for subsequent layers. On the other hand, the electromagnetic enhancement decreases more slowly with increasing distance from the metal surface.

2.3 Tip enhanced Raman spectroscopy

Since the first, actually fortuitous, discovery of a surface enhancement of the Raman scattering, SERS has grown as a valuable technique for the identification and study of adsorbed species in the surface science field. More than 3000 papers dealing with the subject appeared in less than 30 years. It also provided a significant input for theoretical investigations, that, as we will see, are rooted in the near field optics, a topic that recently attracted considerable interest for the development of new optical techniques [43–48].

Nonetheless, SERS still suffers from a few, but severe, limitations: first of

all, only those metals which support an easy excitation of the LSP modes in the visible region provide large enhancement for the Raman process. Silver, gold and copper are the best choices: silver shows the highest enhancement in the whole visible spectrum, while for copper excitation wavelengths above 520 nm are needed; other metals, such as tungsten, have the excitation energy in the UV, or have the plasma resonances heavily damped, such as platinum. One way to extend Raman spectroscopy to other metals was proposed by Weaver and the Tian group [49,50]: a typical SERS active substrate is prepared and then electrochemically covered with a few layers of a second metal (Pt, Pd). This approach profits from the long range nature of the electromagnetic enhancement: LSP are excited in the substrate metal and since the enhanced field may extend over several nanometers, reaching and going beyond the deposited layers, molecules adsorbed at the SERS inactive metal are probed.

A second limiting drawback is represented by the essential need to have the metal surface heavily roughened. A few attempts [51–56] have been made to study the Raman scattering from atomically smooth surfaces but the obtained spectra (if any) were rather weak with most of the enhancement provided by only the chemical effect. Finally, due to the random distribution of the surface roughness, even considering a uniform coverage of the effective area, two different spots will never give the same enhancement, thus making a quantitative analysis rather delicate. Moreover, within the laser focus, only a subgroup of metal structures are in resonance with a particular laser frequency. Changing the laser frequency, but illuminating the same spot, another subgroup of metal structures will be in resonance. Indeed, it has been suggested, and we refer mainly to the work of Moskovits [57–59], that a surface enhanced spectrum is most probably provided by only a few characteristic (and rather peculiar) adsorption sites, denoted as *hot spots*. They are in an optimal resonance condition with the used exciting source, whereas the vast majority of the illuminated area remains inactive or provide only a minor contribution (at that frequency). This specific property of SERS arises from the fact that the more localized the LSP are, the higher the field enhancement. SERS from a few small structures can dominate over the SERS from the rest of the illuminated surface.

The *hot spots* concept and subsequent non-uniform field distribution will be one of the guiding lines in our project.

As seen so far, the rough surface acts for the adsorbates both as a physical support and an enhancer of the Raman intensity. Most of the above mentioned limitations could be circumvented by using an external field enhancing unit. Consider a metallic object of sub-micrometer dimensions placed a few $\mathring{A}$ above an adsorbate supported by a flat metal slab (Fig. 2.5). This simple geometry actually describes the junction formed in a scanning tunneling microscope (STM) with the piezo elements controlling, at atomic resolution, the height and lateral movement of the scanning tip, in our case made out of silver or gold, and the metal slab to close the electric circuit and to enable the current feedback loop. By the illumination of such a cavity, plasmon modes will be excited that are no more related to the surface roughness but are localized below the tip, *i.e.*, at the gap formed by the tip apex and the surface. This approach should still bring a substantial Raman enhancement since the field associated with the external body should extend over several nanometers, thus reaching the underlying surface if the

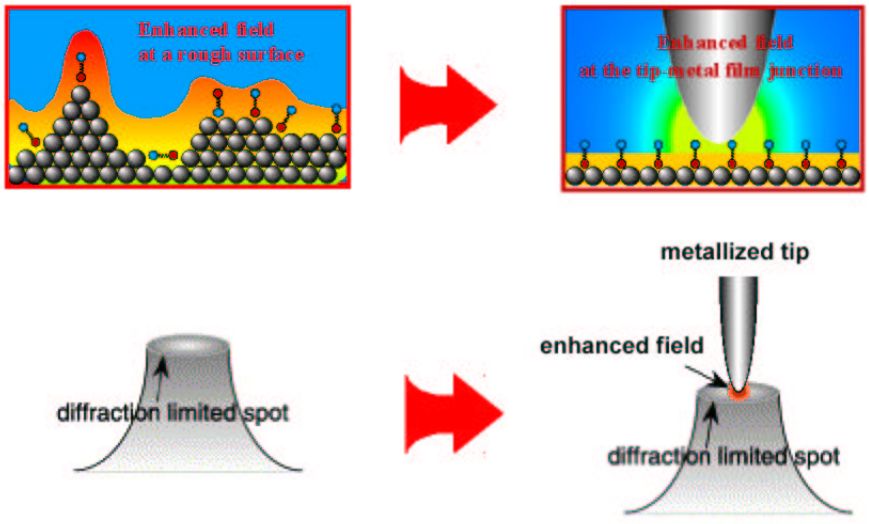


Figure 2.5: Schematic comparison between Surface and Tip Enhanced Raman Spectroscopy.

tip is at tunneling distance. In a figurative way, the tip may be considered as playing the role of a single bump where the enhancing field is extremely localized, in a sense acting as a *single hot spot*, inverted upside down over the surface. Since it is mainly the tip which provides the electromagnetic enhancement, the supporting metal may have a smooth surface and can still provide the chemical component of the enhancement.

Such an improvement, *i.e.,* the availability of a vibrational spectroscopy with high spatial resolution, would be extremely valuable for the study of catalytic processes. A rough surface represents a non-uniform system with a vast number of different adsorption sites, each one with its own vibrational frequency for the adsorbed molecule. The SER spectrum is then the sum of all this information (the band width is a measure of such an inhomogeneous broadening) and only in a few cases it is possible to clearly identify the adsorption geometry. On the other hand, a smooth surface with a precise crystallographic orientation offers well-defined adsorption sites and is usually regarded as a model system. Moreover, Raman studies could be extended to other metals: only the choice for the tip material is restricted, but virtually any conductive surface may be chosen as substrate: platinum, for example, or also a semi-conductor.

There are more features that make a Tip Enhanced Raman Spectroscopy (TERS) so appealing. First of all the possibility to have a local (in the sub-wavelength domain) non-invasive analytical tool: as far the illumination is concerned, Raman spectroscopy and SERS may be regarded as far field optical techniques, where the laser spot size and subsequent lateral resolution cannot be lower than half the wavelength of the excitation beam, due to the diffraction limit of light. In contrast, the extension of the tip induced enhanced field depends mainly on the size of the tip itself. A number of theoretical investigations point out that the highest fields are expected for tips with a small radius, between 10 nm and 200 nm. To a first approximation, we may consider that only the area directly below the tip, shadowed by it, will be influenced by the external field. Such an area is by far smaller than the average diffraction limited laser spot (0.5-2 μm). Consider now also the following: even if only a single monolayer covers the rough

surface, still the number of molecules found in the laser spot, and thus probed, is large, in the order of millions or more; the size of this ensemble will be drastically reduced in the tip configuration since only a small fraction of all the illuminated molecules are actually exposed to the near field of the tip. Finally, the number of probed molecules depends on the tip apex size [60]. For this reason our approach has prospects for being developed into a *single molecule spectroscopy*.

So far no further use for the scanning tunneling unit has been mentioned: our initial need is to control the height of the metal tip in order to place it in close optical, but not physical, contact with the surface; *i.e.*, it forms with the substrate a suitable cavity for LSP confinement. There is obviously the possibility to relate the surface topography recorded with the tunneling microscope with the unique information provided by a vibrational technique and thus to compensate the lack of "chemical sensitivity" quite often ascribed to this otherwise powerful tool.

After this general overview, we address here the main lines to solve theoretically the physical problem of the metal tip illuminated by a monochromatic light. Absorption and scattering of light by small particles is the source for many different and often remarkable optical phenomena; an excellent guide for the theory behind these is the book by Bohren and Huffman [61]. It is worth emphasizing that this class of problems has been studied for at least more than one century. For example, the classic Mie theory, which provides exact solutions for the problem of a sphere in an electromagnetic field, dates back to 1908.

The full electrodynamic problem based on the Maxwell equations may be reduced to the solution of the Laplace equation for the electric potential; it is sufficient to consider that our microstructures have dimensions much smaller than the wavelength of light. Plasmon resonances are usually found in the domain of visible frequency. Therefore, we may neglect, to a first approximation, the magnetic fields and any retardation effects, *i.e.*, we reduce the problem to solving the equations in the electrostatic limit.

Since the ideal geometry of the system (a tip or a sphere placed above a plane) possesses an axis of symmetry, cylindrical coordinates (r, θ, z) [62] are appropriate. Solutions for the Laplace equation then have the form:

$$\psi(r,\theta,z) = R(r)\,Q(\theta)\,Z(z) = J_v(\rho)\exp(\pm i\nu\theta)\exp(\pm ikz) \tag{15}$$

where J_v are the Bessel functions of the first kind and order v and $\rho = kr\sin\theta_i$; they build an orthogonal and complete set of solutions for the corresponding Bessel differential equation.

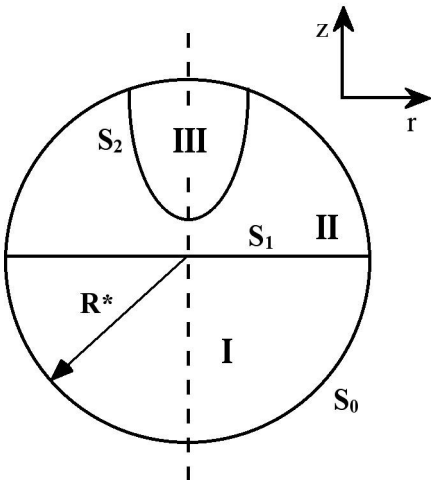


Figure 2.6: Model geometry used for the theoretical calculation by Dickmann *et al.* [63]. Domain III is the actual tip, different shapes for the surface S_2 may be chosen, such as a spherical, an ellipsoidal or hyberbolic surface. S_1 is the flat metal surface and S_0 a surface encompassing the three domains.

At first, we consider the interaction of the incident wave with the metal slab, rewriting the Fresnel formulae: for $z > 0$, that is, above the metal surface, the radial component of the electric field, averaged for all possible values taken by θ, is:

$$E_r = iE_0\cos\theta_i J_1(kr\sin\theta_i) \left[\exp(-ikz\cos\theta_i) - \frac{\varepsilon_M\cos\theta_i - \sqrt{\varepsilon_M - \sin^2\theta_i}}{\varepsilon_M\cos\theta_i + \sqrt{\varepsilon_M - \sin^2\theta_i}}\exp(ikz\cos\theta_i)\right], \tag{16}$$

for the vertical component, along the axis of symmetry z:

$$E_t = E_0 \sin\theta_i J_0\left(kr\sin\theta_i\right) \left[\exp\left(-ikz\cos\theta_i\right) + \frac{\varepsilon_M\cos\theta_i - \sqrt{\varepsilon_M - \sin^2\theta_i}}{\varepsilon_M\cos\theta_i + \sqrt{\varepsilon_M - \sin^2\theta_i}}\exp\left(ikz\cos\theta_i\right)\right]. \quad (17)$$

As usual ε_M is the dielectric constant of the metal, $k = \frac{2\pi}{\lambda}$, λ the light wavelength and θ_i the incident angle; inside the brackets two terms are reported, one for the incident and the other for the reflected wave, the opposite propagation directions are taken into account by the signs.

Next, the diffraction of these fields by the tip is considered. The electric potential (15), satisfying the scalar wave equation, is used as the generating function for vector cylindrical (spherical) harmonics into which the incident field may be expanded. A secondary field, the scattered one, inside and outside the small object, is then obtained by imposing the boundary conditions.

In order to have numerical results, a finite domain for integration has to be defined: in our case a sphere with radius R* encompassing the tip apex. This full domain is divided in three sub-domains: the sample domain (I), the free space (II) and the tip (III). The boundary conditions are then repeated for each interface (S_0, S_1, S_2):

$$\psi_1 = \psi_2 \qquad \varepsilon_1\frac{\delta\psi_1}{\delta n} = -\varepsilon_2\frac{\delta\psi_2}{\delta n} \quad (18)$$

the subscripts 1, 2 identify the two sides of a given interface.

The tip surface (S_2) is given by:

$$z = r_0\cot\varphi\left(\left(1+\frac{r}{r_0}\right)^{\frac{1}{n}} - 1\right) + z_{tip} \quad (19)$$

and its location above the surface (always considered at $z = 0$) is described by the constant z_{tip}. Two parameters are used to define the geometry of the tip: φ

is the aperture angle of the cone and r_0 the curvature radius of the apex; n is a parameter that allows one to change the form of the tip over a wide range, for example: for $n = 1$, the tip is just a cone, for $n = 2$, the apex is a sphere with radius r_0, for $n > 2$ the apex gets flattened.

Dickmann *et al.* published in a series of recent papers [63–67] the results of these calculations performed for a number of different tip materials. The most promising system for enhancing the Raman scattering is, as expected, a silver tip over a gold sample, which we have tested experimentally. In Fig. 2.7, the field increment or field enhancement factor (FEF), defined as the ratio between the total field and the incident field, is plotted *versus* the wavelength of the incident light.

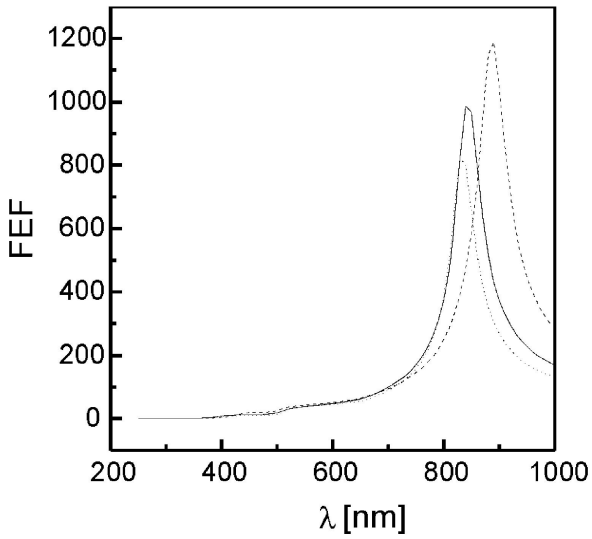


Figure 2.7: Field enhancement at the tip apex *versus* the wavelength of the incident light for a silver tip over a gold surface. Three different tip shapes are studied with curvature radius (from left to right) $r_0 = 10$ nm, 20 nm and 30 nm. Other parameters are $\varphi = 10°$, $\theta_i = 70°$ (lateral illumination), $n = 6$, $z_{tip} = 2$ nm. $FEF = \frac{E_{tot}}{E_0}$ [63].

A few remarks should be underlined at this point. First of all, the calculated field enhancement is extremely large. Four orders of magnitude for the electromag-

netic field would give 10^{12} for the Raman process. If true, this would easily enable *single molecule* studies. It is also seen that, as expected, the tip shape affects the position for the main resonance and its magnitude. As a general rule it is possible to predict that the sharpest tip gives the highest enhancement: as the value for r_0 increases, the enhancement due to the confinement of the surface plasmon modes below the tip becomes less condensed and shifts to lower wavelengths.

The spatial distribution of the enhanced field is reproduced in Fig. 2.8. As mentioned previously, the magnitude of the enhancement factor is mainly influenced by the curvature of the tip, with the sharpest apex giving the strongest and most localized field. As the tip becomes broader and flattened the extension of the field increases and the enhancement weakens. Such a behavior is sketched as a three-dimensional plot in Fig. 2.9 (same tip parameters as in Fig. 2.8b). Clearly, the field distribution is not homogenous but peaks along the axis of symmetry. Adsorbed molecules found further away, but still shadowed by the tip, will contribute less to the Raman signal due to the weaker field. Considering tips with different radii, there is then proportionality between the number of scattering molecules probed by the tip, the population, and their (enhanced) cross section. The strongest field, the one due to small tip radius, is available only for an extremely small number of molecules, since it extends laterally for a short range and drops quickly to zero, while more molecule will be involved in the near field process with the (weaker but more extended) field provided by a larger tip.

The main limitation in these calculations is the complete neglect of any size effect, which must be taken in account when the dimension of the tip is comparable or even smaller than the mean free path length of conduction electrons in the bulk metals. The damping constant used in the free electron term of the dielectric function is increased because of additional collisions with the boundary of the particle. The net effect is to increase the width and lower the peak height for plasmon resonances.

Concluding this section, we point out that many parameters may be changed in the calculation (and in the experiments) in order to acquire the highest enhance-

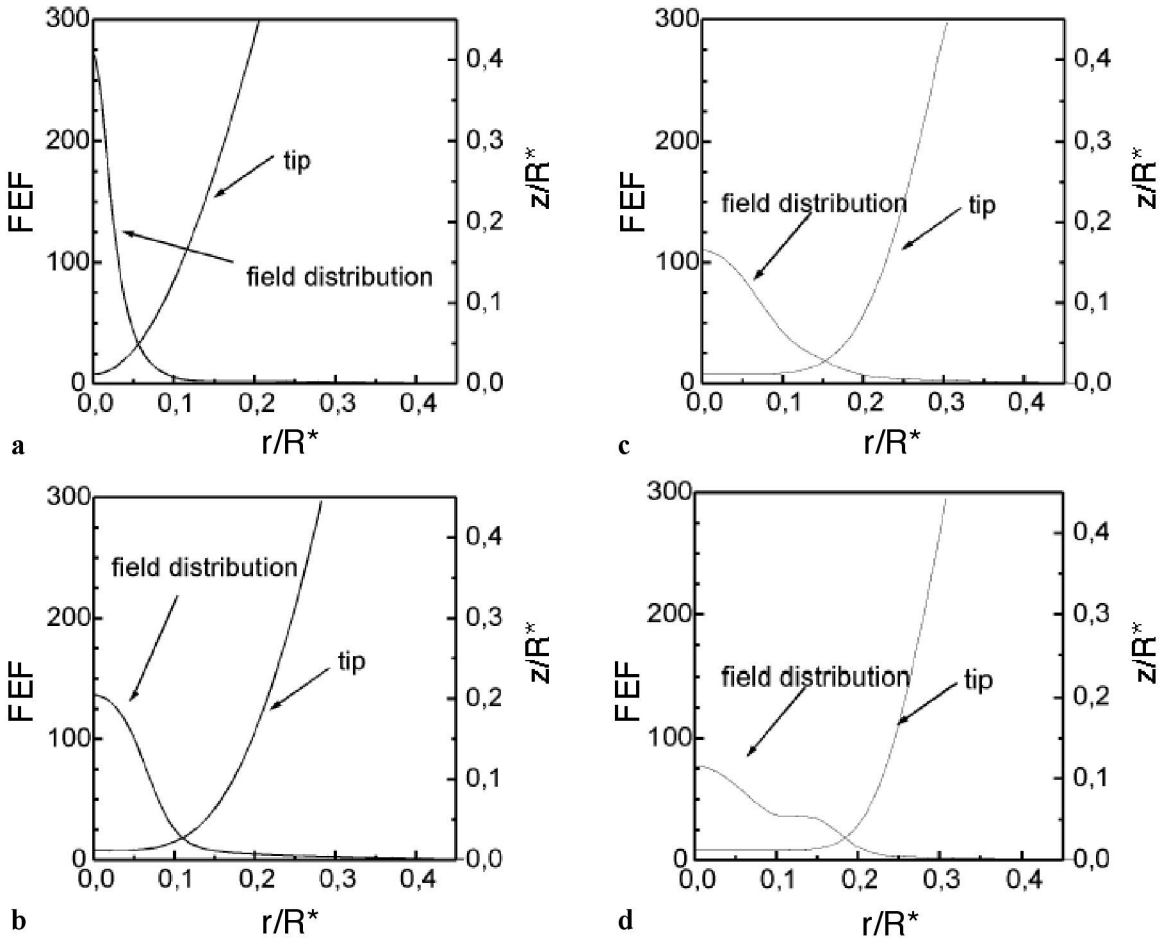


Figure 2.8: Field distribution along the tip profile; due to the axial symmetry only half of the tip and corresponding field is shown. A silver tip over a gold surface illuminated by light with $\lambda = 532$ nm has been considered. Parameters used in the calculations to define the tip geometry are as follows: $\varphi = 10°$, $r_0 = 20$ nm, $R* = 80nm$ $\theta_i = 70°$, $z_{tip} = 2$ nm, $n = 2$ (a), 4 (b), 6 (c), 9 (d) [63].

ment: undoubtedly, the size and the shape of the tip and its dielectric constant play a major role. Also the dielectric constant of the underlying sample material, the distance between tip and surface, the angle of incidence of the light, its polarization and the excitation wavelength are crucial factors.

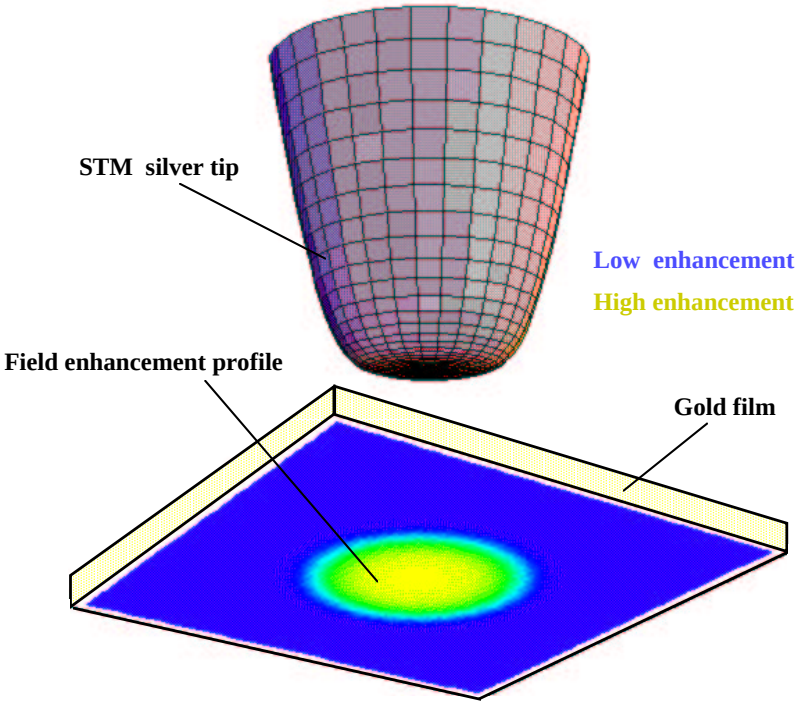


Figure 2.9: Contour plot of the (enhanced) field distribution at the tip sample junction.

2.4 Scanning tunneling microscopy

The tunneling probability, that is the non-zero probability for a particle to be found in a classically forbidden region is perhaps the most striking prediction obtained by applying quantum mechanics to simple physical systems (such as the unidimensional motion of a charge in a potential well or the harmonic oscillator, see Fig. 2.10 and caption).

Such a probability decays exponentially beyond the potential barrier, located at $z = 0$:

$$w = \mid \psi(0) \mid^2 \exp\left(-2d\frac{\sqrt{2m(U-E)}}{\hbar}\right) \tag{20}$$

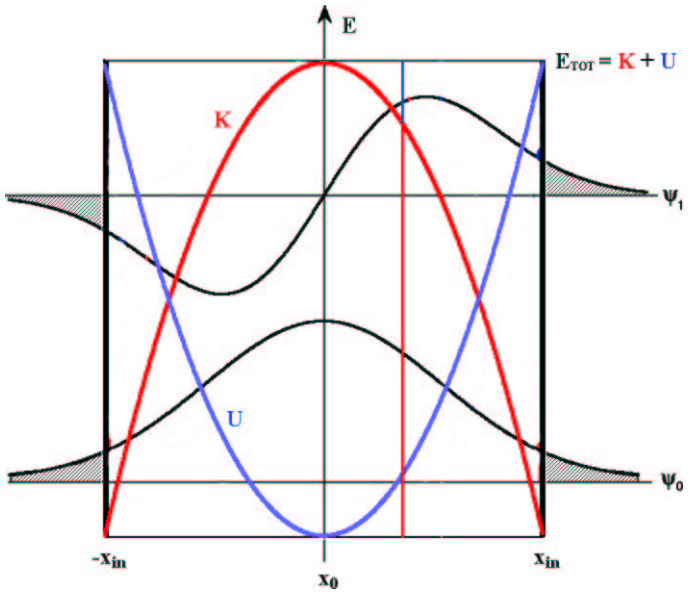


Figure 2.10: In the classical treatment of the harmonic oscillation, the motion is confined between the values $\pm x_{in}$. K is the kinetic energy of the particle, U its potential energy. In a conservative system the sum of the two, the total energy, E_{TOT}, is constant. The oscillator may possess any value of E_{TOT}, corresponding to different initial position, x_{in}. In contrast, only a well defined set of energy values is allowed by quantum mechanics (in the picture two are shown, with the corresponding wavefunctions). Also, the possibility of finding the particle beyond $\pm x_{in}$ (in classical terms, with a negative kinetic energy) is taken in account.

where $\psi(0)$ is the (non-zero) value of the solution at the barrier, E the energy of the particle, U the (finite) height of the barrier, d its width and m the particle mass.

Considering a metal-insulator-metal junction, electrons may jump, tunnel from one metal to the other, through a vacuum gap. A net current may be obtained when applying a bias voltage to the junction, thus changing the respective position of the Fermi levels of the two metals. Still considering the simple case of a unidimensional tunneling, the slope of the natural logarithm of the current plotted against the tunneling distance is proportional to the (apparent) tunnel barrier height $\phi_{app} = U - E$. The success of resolving the atomic structures of conductive surfaces with the STM is based on the exponential distance dependence of the

tunneling current [69].

In the usual topographic studies (performed in the constant current mode), the tunneling bias is kept much below ϕ_{app} (that is a few tens of millivolts) and the tunneling current is set at a fixed value (e.g.: 1 nA). A metal tip is laterally scanned over the surface by piezo-elements. *Via* an internal electrical loop the z-elongation of the piezos (onto wich the tip is mounted) is adjusted so to have always the set current between tip and sample. The voltage(s) applied are recorded and processed to produce an "image" of the surface.

2.5 Light emission from tunneling junctions

The theoretical calculations of the electromagnetic field enhancement occurring in the close proximity of an illuminated metal tip point out very clearly that photons in a well defined frequency range are needed to excite resonantly the localized plasmon modes (Fig. 2.7). The fabrication of tips closely resembling the geometrical models proved to be rather difficult. Tips with axial symmetry, narrow aperture cone and spherical apex with radii below 50 nm are hardly obtained by the usual techniques such as etching or thin film evaporation on commercially available STM or AFM tips. As shown by SEM images [68], the tips are usually composed of a layer of non spherical particles attached to the initial tip body. Other effects, not taken in account in the simulations, may also be present, as energy transfer between adjacent structures or as damping or shifting of the resonances due to the coupling with the metal body. Choosing the correct wavelength to drive the resonant excitation of the LSP is of highest importance to maximize the enhancement of the Raman signal. The theoretical predictions may hold for ideal tip shapes but, most likely, not as well for real tips. Hence, experimental knowledge of the energies of the plasmon resonances supported by the real tips is needed. Information of this type can be obtained by recording the emission spectrum of a given tip/metal substrate.

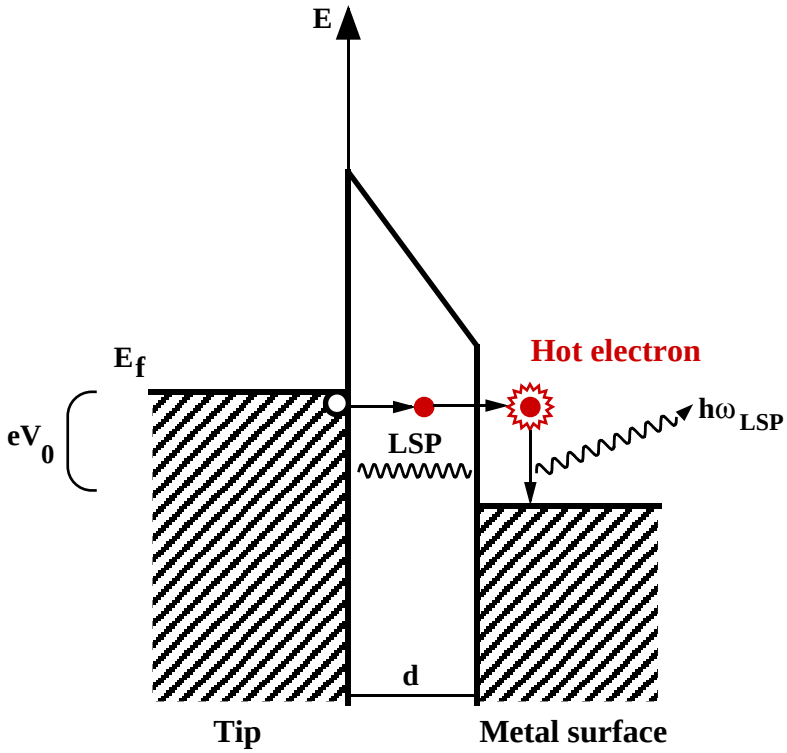


Figure 2.11: Energy diagram for a metal-insulator-metal tunneling junction. Localized plasmon modes confined in the gap may be excited either during the tunneling or by hot electrons. In both cases light is emitted from the junction due to the radiative decay of the LSP.

The LSP modes may be excited not only by incoming photons upon illumination (as in TERS and SERS) but also by inelastically tunneling electrons [70–84]. Let us consider the metal-insulator-metal (MIM) junction formed by the STM tip and the metal film, when it is biased at high voltages ($V_0 > 1.5$ V). Electrons may tunnel inelastically through the barrier losing a fraction or all of their kinetic energy and emit light. For an event involving a single electron the highest frequency of the emitted light is:

$$\hbar\omega_{\max} = eV_0. \tag{21}$$

This picture is not yet complete, since tunneling electrons cannot directly generate photons. There must be an intermediate step to satisfy both the energy and

momentum conservation laws: the excitation of surface plasmon modes localized at the junction. These are excited through the interaction with electrons during the tunneling and then decay by generating (far-field) photons. This means that the frequency of the emitted light is related to the excitation energy of these modes. A slightly different mechanism may also account for the light emission in which the electrons travel through the gap from one metal to reappear as hot electrons in the second metal, where they relax to the Fermi level, again with the excitation of LSP and their final radiative decay. The two mechanisms are schematically depicted in Fig. 2.11.

A number of theories have been developed to link the geometry of the resonant cavity with the energies of the supported modes [26, 85–89]. The main lines to deduce the energies for the LSP are similar to those illustrated in Sec. 2.3. Starting from the Maxwell equations and neglecting magnetic fields and retardation effects, an electric potential, the solution for the Poisson equation must be found which satisfies the boundary conditions. The plasmon driving source, which has to be included in the Poisson equation, is the tunneling charge fluctation [90]. For junction materials, which behave as free electron metals, the dielectric constant may be expressed as a function containing the bulk plasma frequency[3] as a parameter (eq. 6). The approximated solution for a perfectly conducting particle has the form:

$$\omega_n = \omega_{BP} \left[\frac{\tanh\left(n+\frac{1}{2}\right)\beta_0}{\varepsilon + \tanh\left(n+\frac{1}{2}\right)\beta_0} \right]^{1/2} . \tag{22}$$

$\beta_0 = \operatorname{arccosh}\left(1+\frac{d}{R}\right)$ describes, in bispherical coordinates, the surface of a sphere with radius R centered at a distance $(R+d)$ above the surface (Fig. 2.12). If the radius is much bigger than the gap d, β_0 may be approximated to $\left(\frac{2d}{R}\right)^{1/2}$. For such a geometry, the cavity size is defined by the distance from the bottom

[3] In the case that the junction is formed by two different metals ($\omega_{BP1} \neq \omega_{BP2}$), an average plasma frequency should be used: $\omega_{12} = \left[\frac{2\omega_{BP1}^2\omega_{BP2}^2}{\omega_{BP1}^2+\omega_{BP2}^2}\right]^{1/2}$.

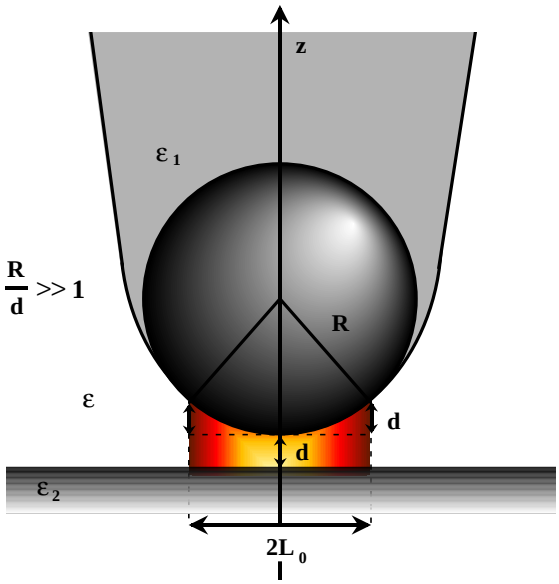


Figure 2.12: Model geometry for the tunneling junction used by R.W. Rendell and D.J. Scalapino [88]. Different geometries are also studied in [86, 88, 89].

of the particle to the point where it has curved up to the distance d, so that $L \simeq \sqrt{2Rd}$. Eq. (2.22) may also be written as (the dielectric constant in the gap, ε, is taken to be equal to unity):

$$\omega_n = \omega_{BP} \left[\frac{\tanh(q_n d)}{1 + \tanh(q_n d)} \right]^{1/2} . \tag{23}$$

q_n are the wavevectors associated with the resonant modes. They are related to the mode localization lengths by:

$$L_n = \frac{(2Rd)^{1/2}}{2n+1} = q_n^{-1} \tag{24}$$

By expanding the argument of tanh to the first order, an easy rule of thumb may be obtained:

$$\omega_n = \omega_{BP} \left(n + \frac{1}{2} \right)^{1/2} \left(\frac{2d}{R} \right)^{1/4} = \omega_{sp} \left(n + \frac{1}{2} \right)^{1/2} \left(\frac{8d}{R} \right)^{1/4} , \tag{25}$$

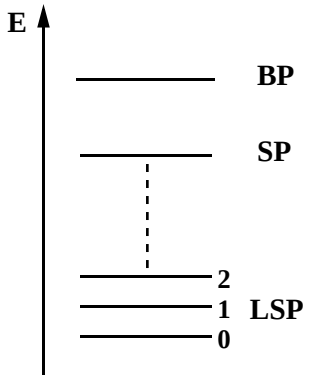


Figure 2.13: The energy sequence (not to scale) of all the plasmon excitations possibly supported by a metal. The initials are for: bulk plasmon, surface plasmon and localized surface plasmon modes. $\omega_{sp}^{Ag} = 3.8$ eV, $\omega_{sp}^{Au} = 2.5$ eV [69].

For the lowest mode, the nodeless one fitting inside the cavity:

$$\omega_0 = \omega_{sp}\left(\frac{2d}{R}\right)^{1/4}, \qquad L_0 = \sqrt{2Rd}. \tag{26}$$

A set of discrete resonant frequencies appears below the limiting surface plasmon frequency[4] (Fig. 2.13). From higher to lower energies the plasmons states become more and more localized, starting from an excitation that is spread all over the metal body to modes that are confined in small regions well below the diffraction limit of light.

Eq. 25 relates in a very straightforward way the geometrical parameters of the cavity and the properties of the material with the LSP energies. Its obvious limit is that modes even beyond ω_{sp} may be obtained. The more rigorous Eq. 22 should be used for higher n or for energies close to ω_{SP}. Both formulae predict that the resonances are shifted to higher energy (blue shift) if the particle size decreases. Figuratively, higher frequency modes are needed to fit into the smaller micro-cavity. On the other hand, for $d \to 0$ the LSP clearly disappear and for $d \to \infty$ the surface plasmon frequency is obtained once again. In a sense, when

[4]Once again it is stressed that while the delocalized surface plasmons are non radiative excitations below the bulk plasma frequency freely propagating along the smooth metal-insulator interface, any irregularity on the surface can scatter them into new radiative states (the LSP).

the tip (the sphere) is too far from the surface its influence is no longer relevant: the planar symmetry of the system is retained and the surface states are not disturbed.

The probability of creating a photon with an inelastically tunneling electron has been experimentally evaluated in vacuum to be in the order of 10^{-5} [71,74–76]. To compensate for this small value, light emission studies integrating over all frequencies and thus mainly focusing on the topography of metal surfaces were thoroughly performed. Photon emission maps are obtained while scanning the STM tip and continuously measuring the total intensity of the emitted light. The maps are then related to the simultaneously recorded topography. It is possible to appreciate from which structures more light is generated, *i.e.*, at which positions of the tip the LSP become highly localized and easily excited. The electromagnetic interaction between tip and sample is so sensitive to the local curvature that an atomic grating such as the (1x2) reconstructed Au(110) surface may also be resolved in the intensity of the emitted light [79]. Luminescence due to the LSP excitation is not restricted to junctions formed by gold or silver. Light has been detected also with tungsten tips and semi-conductor samples [91–93]. In the former case the resonances are more shifted to higher energies, closer to ω_{BP}^{W}; in the latter case the dominant process is the recombination of the electron with a hole and the frequency of the emitted light corresponds to the energy gap of the semi-conductor. It has recently been suggested that tunneling electrons may also occupy empty electronic states of molecules found in the gap, thus opening new avenues for (single molecule) fluorescence spectroscopy induced by the STM [94–97].

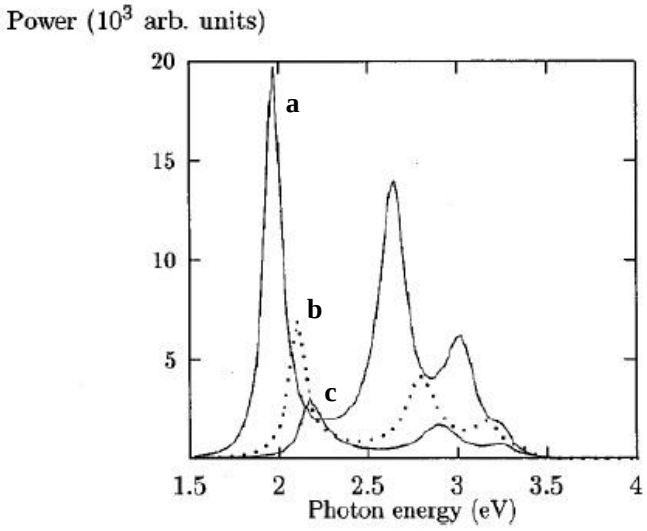


Figure 2.14: Calculated emission spectrum [89] of a silver tip modeled as a sphere with radius $R = 30$ nm above a silver slab for different gap distances: a) 0.5 nm, b) 0.75 nm and c) 1 nm. The resonances shift to lower energies for smaller gaps.

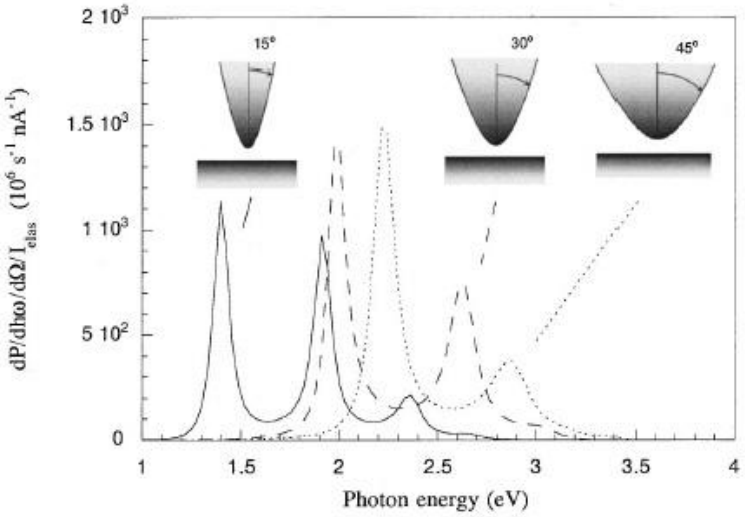


Figure 2.15: In the model by R. Berndt *et al.* [87] a hyperbolic function is used for the tip shape. The resonances and the energy spacing between them is in this case highly influenced by the aperture angle of the tip.