Review

STM-theory: Image potential, chemistry and surface relaxation

Jose Manuel Blanco \textsuperscript{a,b}, Fernando Flores \textsuperscript{a}, Rubén Pérez \textsuperscript{a,*,}\textsuperscript{a}

\textsuperscript{a} Facultad de Ciencias, Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, Modulo C-V, E-28049, Spain
\textsuperscript{b} European Space Astronomy Centre, ESAC-ESA, Ingeniería y Servicios Aeroespaciales, SA (INSA), Spain

Abstract

A critical review of the different methods used nowadays for calculating tunneling currents and STM-images is presented with a special emphasis on the role played by the interface image potential and the interaction between the tip and the sample at short distances. After presenting the most commonly used approaches to this problem, we discuss in full detail how the image potential modifies critically the interface tip–sample barrier and how neglecting this effect underestimates the tunneling currents by several orders of magnitude. Although interface non-local image potential effects are difficult to introduce in a plane-wave Density Functional approach, we show how a Green's function Density Functional formalism based on a local-orbital basis set allows us to introduce those image effects with a good accuracy. The effect of the interaction between the tip and the sample is illustrated for an Al-tip approaching an Al surface; and the role of the electronegative atoms adsorbed on the tip is discussed considering the O/Pd(111) interface and the effect of having an O-atom adsorbed on the tip apex. Finally, by analyzing the Si(112)–Ga interface we also show how the Green's function Density Functional approach based on a local orbital basis can also be reliably used to analyze surface stoichiometries.

© 2006 Elsevier Ltd. All rights reserved.

PACS: 68.37.Ef; 71.15.Mb; 73.40.Gk

Keywords: Scanning Tunneling Microscopy; Density Functional Theory; Green’s functions; Electronic transport; Image potential; Local orbital basis; Tip-sample interaction

* Corresponding author. Tel.: +34 914974906.
E-mail address: ruben.perez@uam.es (R. Pérez).
1. Introduction

The invention of the Scanning Tunneling Microscope [1,2] introduced a new and revolutionary tool in Surface Science, and started a development that is leading worldwide towards the new field of nanotechnology. In spite of the 24 years lapsed since that invention, we do not have yet a method reliable enough to yield tunneling currents and STM-images with sufficient accuracy. This is partly due to the lack of a full control of the tip structure [3–6], which has obscured the difficulties for understanding the theoretical challenges to solve in this area.

Typically, simple models (like the one proposed by Tersoff and Hamann – see Section 2) have been used to interpret STM-images. With the development of DFT-codes [7–9] a conventional and more accurate approach has been based on the calculation of the independent tip and sample electronic structures, and the coupling of both sides of the interface using Bardeen’s approach and an extended basis formalism (see Section 2). As we shall discuss in Section 3, this approach presents a fundamental limitation in its conventional applications: the neglect of the interface image potential [10]. As shown in Section 3, this approximation in the calculation of the tunneling currents introduces an important
underestimation of their values, and a drastic reduction of the tip sample distance if one tries to calculate those currents fitting this length.

Alternatively, methods based on a local basis formalism can also be used to calculate the tip and sample structure. In this approach, the tip–sample interaction is obtained calculating the hopping integrals between the orbitals of both sides. As we will discuss in Sections 3 and 4, this approximation, if properly used, can incorporate image potential effects.

In Section 4 we shall also discuss the role played in the tunneling currents by the tip–sample interaction that, at small tip–sample distances, can introduce important atomic relaxations and a change of the surface corrugation. Tip–sample chemical interaction can also be important [11]; in particular, tips contaminated with very electronegative atoms can show inverted STM-images [12,11]: in Section 5, we shall discuss an example, Pd(111)–O, where an O-atom chemisorbed on the apex of the tip introduces those effects. In this case, we shall also discuss the role played by the d-orbitals in the STM-images, an effect that can be found in other transition metal surfaces [13,14]. Finally, we shall present a case, Ga/Si(112)-6 \cdot 1, where the local orbital basis formalism has been successfully used to the determination of a complicated surface structure not having, a priori, a well known stoichiometry.

Although this critical review focuses on discussing theoretically the important issues of image potential, chemistry and tip-induced surface relaxation effects, the reader is referred to other works where a more comprehensive discussion of STM is presented [15–21].

2. Tunneling currents. General formulation

Typically, tunneling currents in STM are analyzed splitting the system into two parts: the tip (T) and the sample (S); then, some theoretical scheme has to be used to obtain the coupling between both sides, T and S, of the interface. In this section, we discuss some typical approaches adapted to the use of either a plane-wave or a local-orbital basis set.

2.1. Extended basis formalism

We first consider the case of an extended basis formalism in which wavefunctions are described in a continuum way, say as \( \psi_T(\vec{r}) \) or \( \psi_S(\vec{r}) \) for the tip and the sample. Fig. 1a shows the geometry of our system and Fig. 1b its energy diagram [18].

2.1.1. Bardeen’s approach

In this case a surface, \( \sigma \) defines the separation between T and S [22]. Then, as showed in Fig. 1c and d, one defines the potentials \( U_S(\vec{r}) \) and \( U_T(\vec{r}) \) such that the total potential is \( U_S(\vec{r}) + U_T(\vec{r}) \); with this geometry the functions \( \psi_T(\vec{r}) \) and \( \psi_S(\vec{r}) \) are defined by (using \( k \) or \( k' \) as the quantum numbers of the eigenstates)

\[
-\frac{1}{2} \nabla^2 \psi_k + U_S \psi_k = \epsilon_k \psi_k
\]

\[
-\frac{1}{2} \nabla^2 \psi_{k'} + U_T \psi_{k'} = \epsilon_{k'} \psi_{k'}
\]

(atomic units, \( \hbar = e = m = 1 \), are used along this paper unless specifically mentioned).

In this approach, \( \psi_k \) and \( \psi_{k'} \) are the wavefunctions of the uncoupled sample and tip, and the tunneling currents are obtained by analyzing how these wavefunctions are modified by the coupling between both sides. In the coupled system,
in such a way that, in lowest order perturbation theory, we expect to have

\[ \varphi_s \approx \varphi_k + \sum_{k'} C_{kk'} \varphi_{k'} \]

\[ \varphi_{s'} \approx \varphi_{k'} + \sum_k C_k \varphi_k \]

which means that the coupling between S and T makes the S(T)-electrons to tunnel into the T (S)-side.

The coefficients \( C_{kk'} \) and \( C_k \) in Eqs. (4) and (5) can be calculated using perturbation theory. In a first step, we need \( \tilde{T}_{kk'} = \int_\sigma \varphi_k \varphi_{k'} (-1/2 \nabla^2 + U_S + U_T) \varphi_{k'} d^3r \) the coupling between wavefunctions \( \varphi_k \) and \( \varphi_{k'} \). This is given by [23]:

\[ \tilde{T}_{kk'} = \epsilon_s S_{kk'} + T_{kk'} \]

where

\[ S_{kk'} = \int \varphi_k \varphi_{k'} d^3r \quad \text{and} \quad T_{kk'} = -\frac{1}{2} \int_\sigma (\varphi_{k'} \nabla \varphi_k - \varphi_k \nabla \varphi_{k'}) d^3\sigma \]

\( \sigma \) being the surface of separation between S and T (see Fig. 1a). In calculating Eq. (6), it has been assumed that \( \epsilon_s = \epsilon_k = \epsilon_{k'} \); this is the case of interest because electrons tunnel between both sides at constant energy.

Then, perturbation theory (taking into account that \( \varphi_k \) and \( \varphi_{k'} \) are not orthogonal) yields for \( \varphi_s \) and \( \varphi_{s'} \) [24]:

\[ \varphi_s \approx \phi_k + \sum_{k'} \frac{T_{kk'} \phi_{k'}}{\epsilon_k - \epsilon_{k'} - i\eta} \]

\[ \varphi_{s'} \approx \phi_{k'} + \sum_k \frac{T_{kk'} \phi_k}{\epsilon_{k'} - \epsilon_k - i\eta} \]
where \( \phi_k = \sum_k (S^{-1/2})_{kk'} \phi_{k'} \) and \( \phi_{k'} = \sum_k (S^{-1/2})_{kk'} \phi_{k'} \) are the Löwdin’s orthogonal basis and \( k'' \equiv (k,k') \). In the lowest order, \( \phi_k \approx \phi_k - 1/2 \sum_{k'} S_{kk'} \phi_{k'} \) and \( \phi_{k'} \approx \phi_{k'} - 1/2 \sum_{k''} S_{kk''} \phi_{k''} \).

Eqs. (8) and (9) can be used to calculate the current between \( S \) and \( T \) associated with the states, \( \phi_k \) and \( \phi_{k'} \). For example, \( \phi_2 \) yields the following density current \( \overrightarrow{j}_2(\overrightarrow{r}) \) [25–27]:

\[
\overrightarrow{j}_2(\overrightarrow{r}) = \frac{1}{2i} \left\{ \phi_2^* \nabla \phi_2 - \phi_2 \nabla \phi_2^* \right\} \approx \pi \sum_{k'} \left\{ \phi_k(\overrightarrow{r})T_{kk'} \nabla \phi_{k'}(\overrightarrow{r}) - \nabla \phi_k(\overrightarrow{r})T_{kk'} \phi_{k'}(\overrightarrow{r}) \right\}
\]

The current \( I(k \rightarrow k') \) is obtained by integrating \( \overrightarrow{j}_2(\overrightarrow{r}) \) across the separation surface:

\[
I(k \rightarrow k') = \int \overrightarrow{j}_2(\overrightarrow{r}) \cdot d\overrightarrow{s} = 2\pi \sum_{k'} |T_{kk'}|^2 \delta(\epsilon_k - \epsilon_{k'})
\]

The total current is given by (the factor \( e/h \) is explicitly included in the tunneling current):

\[
I_{\text{tunnel}}(S \rightarrow T) = 2\pi e/h \sum_{kk'} |T_{kk'}|^2 [f_S(\epsilon) - f_T(\epsilon')] \delta(\epsilon_k - \epsilon_{k'})
\]

where \( f \) is the Fermi distribution function. Eq. (12) can be written as a function of the density of states, \( \rho_k(\epsilon) \) and \( \rho_{k'}(\epsilon) \), for the sample and tip, introducing the quantum numbers \( k \) and \( k' \), such that \( k \equiv (k,\epsilon) \) and \( k' \equiv (k',\epsilon) \):

\[
I_{\text{tunnel}}(S \rightarrow T) = 4\pi e/h \int d\epsilon [f_S(\epsilon) - f_T(\epsilon)] \int \frac{d^2\kappa}{(2\pi)^2} \frac{d^2\kappa'}{(2\pi)^2} \rho_k(\epsilon)T_{kk'}(\epsilon)\rho_{k'}(\epsilon)T_{kk'}(\epsilon)
\]

where the factor \( 4\pi e/h \) includes the spin degeneracy.

Eq. (13) yields, in principle, an exact solution (in the lowest order perturbation theory) to the problem of calculating the tunneling currents between the tip and the sample. We should keep in mind, however, the main assumptions made to calculate Eq. (13):

(i) As shown in Fig. 1, \( \phi_k \) and \( \phi_{k'} \) are obtained from the potentials \( U_S \) and \( U_T \), respectively. In practice, these potentials are not known, unless a fully self-consistent potential including electron correlations effects is calculated for the total, \( S + T \), system. The reason for this is that, typically, image potential effects modify substantially the interface potential (see below) and, accordingly, the wavefunctions \( \phi_k \) and \( \phi_{k'} \) should keep information of this effect.

In many DF-calculations, the tunneling currents are obtained using only the short-range surface potential for each independent surface. As we will discuss below, this approximation underestimate drastically the tunneling currents.

(ii) On the other hand, Eq. (13) is calculated using lowest order perturbation theory. This means that Eq. (13) is only valid for long distances (typically for distances between the atoms of the tip and sample larger that 4–5, see below). One might argue that, in this limit, using the short-range potential for each independent surface (tip and sample) is reasonable; the problem is that in this limit, the image potential is still playing an important role modifying substantially the tunneling currents (see below).

### 2.1.2. Tersoff–Hamman’s approximation

In this approximation [28], the tip is assumed to have only one active s-like orbital in the apex atom (see Fig. 2). Then, the tip wavefunction goes like

\[
\phi_{k'}(\overrightarrow{r}) = \alpha_{k'} \phi^T_{s}(\overrightarrow{r}) + \cdots
\]

\[ \text{Eq. (14)} \]
where $\phi^{T}_{s}(\vec{r})$ is the apex s-orbital and $\alpha_s(k')$ a parameter depending on the quantum number $k'$.

Assuming also that the sample wavefunction near the tip apex has a slow variation in $\vec{r}$ (determined by $W$, the effective barrier for the tunneling electrons, see Fig. 2) we find (see Eq. (6))

$$T_{kk'} = \int \phi_k(\vec{r}) \left( -\frac{1}{2} \nabla^2 + U_T + U_S \right) \phi_{k'}(\vec{r}) d^3r - \epsilon_k S_{kk'}$$

$$\approx (\epsilon_s - \epsilon_x) \phi_x(\vec{R}_0) \alpha_s(k') \int \phi_{k}(\vec{r}) d^3r$$

(15)

where $\vec{R}_0$ is the coordinate of the apex atom and $\epsilon_s$ the s-orbital level.

In Eq. (15), $\epsilon_s - \epsilon_x$ and $\alpha_s(k') \int \phi_{k}(\vec{r}) d^3r$ are constants independent from the tip position; this shows that $T_{kk'} \sim \phi_k(\vec{R}_0)$ and

$$|T_{kk'}|^2 \sim |\phi_k(\vec{R}_0)|^2$$

(16)

Then, the tunneling current is given by

$$I_{tunnel}(S \rightarrow T) \sim \sum_k |\phi_k(\vec{R}_0)|^2 = \int_0^{\epsilon_V} \rho(\vec{R}_0, \epsilon) d\epsilon$$

(17)

where $\rho(\vec{R}_0, \epsilon)$ is the local density of states of the sample at the tip apex position. In Eq. (17), the energy origin is the tip Fermi level and $\epsilon_V$ the sample Fermi energy (we have also assumed zero temperature).

Eq. (17) is very appealing because it relates the tunneling intensity to the sample local density of states at the position of the tip active atom. Moreover, from Eq. (17) we find:

$$\frac{dI_{tunnel}}{dV} \sim \rho(\vec{R}_0, \epsilon_V)$$

(18)

This result has been frequently used to interpret STM-images; given a particular surface, one can use conventional plane-wave DF-codes to calculate $\rho(\vec{R}_0, \epsilon)$ which is directly related, in Eq. (18), to $dI_{tunnel}/dV$.

This is a very simple result that can be easily applied calculating the sample electron local density of states. This density of states far from the surface is small, however, and calculating it numerically is not an easy task; this explains why in these Tersoff–Hamman model calculations the tip–sample distance is rather small, typically 3 Å, smaller than the experimental distances of 5–8 Å. The main limitation of this model comes from its independence of the tip geometry and the tip electronic structure; in particular, tip and sample...
d-electrons have been found to play an important role in some metal STM-images (see below [29]).

2.1.3. Chen’s approximation

Tersoff–Hamman’s approach can be generalized to other active orbitals in the tip apex [29]. For an active orbital having p-symmetry the approximations made in Eq. (15) have to be changed; in particular, \(
\frac{\partial}{\partial x_i} \varphi_k(\mathbf{R}_0) \)

is zero and one has to go to a higher order term in \( \varphi_k(\mathbf{R}) \) around \( \mathbf{R}_0 \). This yields:

\[
T_{kk'} \approx (\epsilon_p - \epsilon_s)\alpha_p(k') \int \mathbf{r} \varphi_p d^3r \cdot \nabla \varphi_k(\mathbf{R}_0)
\]

showing that

\[
|T_{kk'}|^2 \sim \left| \frac{\partial}{\partial x_i} \varphi_k(\mathbf{R}_0) \right|^2
\]

Then

\[
I_{\text{tunnel}} \sim \int_{0}^{eV} \sum_{k} \left| \frac{\partial}{\partial x_i} \varphi_{\kappa,\epsilon}(\mathbf{R}_0) \right|^2 d\epsilon
\]

where we have used the quantum numbers \((\kappa, \epsilon) \equiv k \) (see above).

For an active d-orbital one can get straightforwardly the following equations:

\[
T_{kk'} \sim \frac{\partial^2 \varphi_k(\mathbf{R}_0)}{\partial R_x \partial R_y} \quad (\text{d}_{xy}-\text{orbital})
\]

\[
T_{kk'} \sim \left\{ \frac{2}{3} \frac{\partial^2}{\partial R_z^2} - \frac{1}{3} \left( \frac{\partial^2}{\partial R_x^2} + \frac{\partial^2}{\partial R_y^2} \right) \right\} \varphi_k(\mathbf{R}_0) \quad (\text{d}_{z^2}-\text{orbital})
\]

that can be used as in Eq. (21) to calculate the tunneling currents.

The main problem with this approach is that one needs to know in advance that there is only one active orbital because, otherwise, it is not so easy to combine appropriately the contributions coming from several active orbitals.

2.2. Local basis formalism

In this approach [30–38], a local orbital basis, \( \varphi_\mu(\mathbf{r}) \), is used to describe the electron states of the system \((i\varepsilon) \) represents the atom \( i \) and the orbital \( \varepsilon \); from now on we write \( \mu = (i, \varepsilon) \)). There are DFT-codes (like Fireball [39–41] and SIESTA [42]) that use that basis set and yield the electron states, \( \varphi_\mu(\mathbf{r}) \), as \( \sum_v c_\mu(v) \phi_i(\mathbf{r}) \), \( \phi_i(\mathbf{r}) \) being the orthogonal Löwdin’s wavefunction \((\phi_\mu = \sum_\nu (5^{-1/2})(\nu\mu)\phi_\nu(\mathbf{r}))\).

In this local basis formalism, we can describe the sample and the tip hamiltonians as follows (see Fig. 3):

\[
\hat{H}_S = \sum_\nu \{ \epsilon_\nu \hat{n}_\nu + T_{\nu\mu} c_\mu^+ c_\nu \} \quad (24)
\]

\[
\hat{H}_T = \sum_{\nu'} \{ \epsilon_{\nu'} \hat{n}_{\nu'} + T_{\nu'\mu'} c_{\nu'}^+ c_{\mu'} \} \quad (25)
\]
where the operators $c^+_v$, $c^+_m$ and $c^+_v$, $c^+_m$ are associated with the wavefunctions $\psi_v(\vec{r})$, $\psi_m(\vec{r})$.

In this formalism, the tunneling currents are associated with the tip–sample coupling Hamiltonian

$$\hat{H}_1 = \sum_{vv'} \left[ T_{vv'} \hat{c}_v^+ \hat{c}_{v'} + T_{v'v} \hat{c}_{v'}^+ \hat{c}_v \right]$$

(26)

defined by the hopping elements, $T_{vv'}$, that play the role $T_{kk'}$ do for the extended basis formalism.

### 2.2.1. Tunneling currents. Green’s function formalism

Assuming Hamiltonians $H_S$, $H_T$ and $\hat{H}_1$ to be known, we will discuss in this section how to calculate the tunneling currents of the problem. We stress that, in principle, we can calculate $H_T$ and $H_S$ from first-principles DF-methods based on a local orbital basis; in Section 2.2.4 we will discuss how to obtain $H_I$ using different approximations.

The analysis of the tunneling currents for weak coupling, $T_{vv'}$, is similar to the one presented in Section 2.1.1. In particular, Eq. (13) can be written in the local basis used here as follows:

$$I_{\text{tunnel}}(S \rightarrow T) = \frac{4\pi e}{\hbar} \int d\epsilon \left[ f_S(\epsilon) - f_T(\epsilon) \right] \cdot \text{Trace}\{ \hat{T}_{TS} \hat{\rho}_{SS} \hat{T}_{ST} \hat{\rho}_{TT} \}$$

(27)

In this equation $\hat{T}_{TS} \hat{\rho}_{SS} \hat{T}_{ST} \hat{\rho}_{TT}$ replaces the term $T_{v'\alpha} \rho_{\alpha'}(\epsilon) T_{\alpha'k'} \rho_{k'}(\epsilon)$ in Eq. (13), and Trace replaces the integrations $\int d^2\kappa(2\pi)^2 d^2\kappa'(2\pi)^2$. Notice also that in Eq. (27), $\hat{T}$ and $\hat{\rho}$ are matrices, $\hat{T}_{TS}$ representing $T_{vv'}$ and $\hat{\rho}$ the density of states matrix of either the independent tip or sample, $\rho_{vv'}(\epsilon)$ or $\rho_{\alpha'\alpha}(\epsilon)$. In this local approach, the calculation of the tunneling currents is reduced to including the active sites for which the tip–sample hopping are not negligible (see Fig. 3).

Eq. (27) yields the tunneling currents in the lowest order perturbation theory. Going beyond this approximation can be achieved by considering all the multiple processes shown in Fig. 4 [43,44]. Adding all the diagrams in this figure leads to the effective hopping matrix $\hat{T}_{TS}^{\text{eff}}$. 

---

**Fig. 3.** Geometry for the local basis formalism. Active sites (in black) are shown.
The point to realize is that Eq. (31) is also valid in an tip–sample currents for any geometrical configuration within a local orbital formalism. Then, the tunneling current to all orders in \( b \) is given by

\[
I_{\text{tunnel}}(S \rightarrow T) = \frac{4\pi e}{\hbar} \int \text{d}\epsilon [f_S(\epsilon) - f_T(\epsilon)] \cdot \text{Trace} \{ \hat{T}^{\text{eff}, R}_{\text{TS}} \hat{G}^{\text{eff}, A}_{\text{ST}} \hat{\rho}_{\text{TT}} \}
\]  

(30)

where both the retarded and advanced hopping matrices, \( \hat{T}^{\text{eff}, R}_{\text{TS}} \) and \( \hat{T}^{\text{eff}, A}_{\text{TS}} \), have been introduced (see Ref. [43] for a more formal discussion of Eq. (30), based on a Keldysh–Green’s function approach). Eq. (30) can also be written as follows:

\[
I_{\text{tunnel}}(S \rightarrow T) = \frac{4\pi e}{\hbar} \int \text{d}\epsilon [f_S(\epsilon) - f_T(\epsilon)] \cdot \text{Trace} \{ \hat{D}^R_{\text{TT}} \hat{T}_{\text{TS}} \hat{G}^{\text{eff}, A}_{\text{ST}} \hat{D}^A_{\text{TT}} \hat{T}_{\text{ST}} \hat{\rho}_{\text{TT}} \}
\]  

(31)

where

\[
\hat{D}^R_{\text{TT}} = \{ 1 - \hat{T}_{\text{TS}} \hat{G}^{(0)}_{\text{SS}} \hat{T}_{\text{ST}} \hat{G}^{(0)}_{\text{TT}} \}^{-1}
\]  

(32)

(other equivalent equations can be written using the identity \( \hat{D}^R_{\text{TT}} \hat{T}_{\text{ST}} \hat{G}^{(0)}_{\text{TT}} \hat{T}_{\text{TS}} \hat{G}^{(0)}_{\text{SS}} \hat{T}_{\text{ST}} \hat{\rho}_{\text{TT}} \) with \( \hat{G}^{(0)}_{\text{SS}} = \{ 1 - \hat{T}_{\text{TS}} \hat{G}^{(0)}_{\text{SS}} \hat{T}_{\text{ST}} \hat{G}^{(0)}_{\text{TT}} \}^{-1} \)).

Eq. (31) (or its lowest order approximation, Eq. (27)) is the central object yielding the tip–sample currents for any geometrical configuration within a local orbital formalism.

Before closing this section, it is worth discussing how Eq. (31) can be generalized to the case of a plane-wave formalism. The point to realize is that Eq. (31) is also valid in an \( \vec{r} \)-representation (see Eq. (27)); then, to all orders in perturbation theory:

\[
I_{\text{tunnel}}(S \rightarrow T) = \frac{4\pi e}{\hbar} \int \text{d}\epsilon [f_S(\epsilon) - f_T(\epsilon)] \cdot \sum_{k',k} \left\{ T_{k'k} \hat{D}^R_{kk} \rho_k(\epsilon) \hat{T}_{k'k'} \hat{D}^A_{k'k'} \rho_{k'}(\epsilon) \right\}
\]  

(33)

The new factors in this equation are \( \hat{D}^R_{kk} \) and \( \hat{D}^A_{k'k'} \), given by

\[
\hat{D}_{kk} = \{ 1 - \hat{G}^{(0)}_{kk} \hat{T}_{kk} \hat{G}^{(0)}_{kk} \hat{T}_{kk} \}^{-1}
\]  

(34)

For calculating \( I_{\text{tunnel}} \) to all orders in \( \hat{T} \) one should obtain first \( \hat{D}_{kk} \) and then use Eq. (33) (a task which has not yet been applied to any practical case).

2.2.2. Scattering theory

This approach is closely related to the Green’s function formalism discussed above, and it can be applied to the calculation of the tunneling currents using either an extended or a
local orbital basis for the description of the electron wavefunction at the interface. In the scattering theory, electrons propagate along different channels in the tip and sample [21,33]. In the tip (sample) one defines the incoming and outgoing wavefunctions, $i_{i}(S, T)$ and $i_{j}(S, T)$ (and $f_{i}(S, T)$ and $f_{j}(S, T)$), in such a way that the tunneling across the interface is described by the transmission, $T$, and reflection, $R$, matrixes defined by equations:

$$c_{f}^{(S),O} = \sum_{i} T_{fi}c_{i}^{(T),1} + \sum_{i'} R_{f'i}c_{i'}^{(S),1}$$

$$c_{j}^{(T),O} = \sum_{i} T_{ji}c_{i}^{(T),1} + \sum_{i'} R_{ji'}c_{i'}^{(S),1}$$

where the amplitudes, $c_{f}^{(S),O}$ and $c_{j}^{(T),O}$ of the outgoing wavefunction for the sample and the tip are linearity related to the amplitude of the incoming ones, $c_{i}^{(S),1}$ and $c_{i}^{(T),1}$. The transmission and reflection probabilities between two channels (say $i \leftarrow j'$ and $i' \leftarrow j$) are given by $|T_{fi}|^2$ and $|R_{f'i'}|^2$, respectively, and the total probability of having one electron scattered to the outgoing channel $j'(j)$ is given by $\sum_{i}|T_{fi}|^2 + \sum_{i'}|R_{f'i'}|^2$ (or $\sum_{i}|T_{ji}|^2 + \sum_{i'}|R_{ji'}|^2$). One can define the interface conductance (from tip to sample) as [45,46]:

$$G = \frac{2e^2}{\hbar} \sum_{i,j} |T_{fi}|^2$$

where we consider all the different transmission processes between the incoming (tip) and the outgoing (sample) channels, summing all the initial and final states. Eq. (37) can be written as follows:

$$G = \frac{2e^2}{\hbar} \text{Trace}(\hat{T} \hat{T}^\dagger)$$

(38)

this equation defining the interface conductance (and the tunneling currents) as a function of the transmission matrix, $T_{fi}$. For a finite bias, the tunneling current is given by

$$I = \int_{0}^{eV} \frac{2e}{\hbar} \text{Trace}(\hat{T} \hat{T}^\dagger) d\epsilon$$

(39)

where the energy integration is extended to the window (0,eV) defined by the applied bias, $V$. Notice that $T_{fi} = T_{f'i'}$ [46], in such a way that the interface conductance is symmetric: $G(T \rightarrow S) = G(S \rightarrow T)$, as it should.

We can establish contact between Landauer’s Eqs. (38) and (31) by defining

$$\hat{T} = 2\pi \rho_{SS}^{1/2} \hat{T}_{ST} \hat{D}_{TT}^{1/2}$$

(40)

Then, using the cyclic property of the Trace, we can write:

$$\frac{dI}{dV} = \frac{2e^2}{\hbar} \text{Trace}\left\{ 2\pi \rho_{SS}^{1/2} \hat{T}_{ST} \hat{D}_{TT}^{1/2} \right\} = \frac{2e^2}{\hbar} \text{Trace}(\hat{T} \hat{T}^\dagger)$$

(41)

The scattering matrix, or the transmission matrix $\hat{T}$, has been evaluated using either a transfer matrix technique [47,48] or Green’s function techniques [33]. Eqs. (41) and (31) show how closely related the scattering theory and our Green’s function formalism are. There is, however, one important point to stress: in the scattering theory one has to evaluate the incoming and outgoing wavefunctions of the tip and sample; although this is a simple task for the bulk channels (wavefunctions of the infinite crystal), things are more
difficult to evaluate when surface states appear in the sample and the tip. The great advantage of our Eq. (31) is that the contribution of bulk and surface channels is automatically included in the calculation of the Green’s functions and the $D^K, D^K$-matrixes.

2.2.3. Hückel approximation vs Tersoff–Hamman’s and Chen’s approximation

Our discussion of the tunneling currents in Section 2.2.1 can be straightforwardly applied if we have a prescription for calculating $T_{vv'}$. In Hückel approximation:

$$T_{vv'} = \frac{\kappa}{2} (\epsilon_v + \epsilon_{v'}) S_{vv'}$$

(42)

Assume we have only an active orbital in the tip apex, say $v' = v'_0$. Then

$$T_{vv'} \sim S_{v v'_0}$$

(43)

and

$$\frac{dI_{\text{tunnel}}}{dV} \sim \sum_{v_1 v_2} j_{v'_0 v'_0} S_{v_0 v'_1} j_{v_1 v_2} S_{v_2 v'_0}$$

(44)

For a very localized tip orbital having s-like symmetry:

$$S_{v'_0} = \int \varphi_{v'_0}(\vec{r}) \varphi_{s,v_0}(\vec{r}) d^3r \approx \varphi_{v'_0}(\vec{R}_0) \int \varphi_{v'_0}(\vec{r}) d^3r$$

(45)

This shows that the tunnel current depends on $\vec{R}_0$ as follows:

$$\frac{dI_{\text{tunnel}}}{dV} \sim \sum_{v_1 v_2} \varphi_{v_1}(\vec{R}_0) j_{v_1 v_2} \varphi_{v_2}(\vec{R}_0) = \rho(\vec{R}_0, \epsilon)$$

(46)

the sample density of states at the point $\vec{R}_0$, in agreement with Tersoff–Hamman [28]. On the other hand, for a p-like orbital ($i \equiv x, y, z$):

$$S_{v'_i} \approx \frac{d \varphi_{v'_0}(\vec{R}_0)}{dr_i} \int r_i \varphi_{x_i, v'_0}(\vec{r}) d^3r$$

(47)

then, the tunneling current is given by

$$\frac{dI}{dV} \sim \sum_{v_1 v_2} \frac{d \varphi_{v_1}(\vec{R}_0)}{dr_i} \rho_{v_1 v_2} \frac{d \varphi_{v_2}(\vec{R}_0)}{dr_i}$$

(48)

in agreement with Chen [18].

2.2.4. Calculation of the hopping elements $T_{vv'}$

Going beyond Hückel’s approximation implies having a more accurate prescription for obtaining $T_{vv'}$. We first discuss how to proceed in the dimer approximation shown in Fig. 5. In this case, we assume to have orbitals $v$ and $v'$ associated with the atoms $A$ and $A'$, creating short-range muffin-tin potentials $V_v(\vec{r})$ and $V_{v'}(\vec{r})$ respectively. As in Section 2.1.1 we first calculate $T_{vv'}$:

$$\widetilde{T}_{vv'} = \int \psi_v H \psi_{v'} d^3r = \int_{\Omega} \psi_v \left( -\frac{1}{2} \nabla^2 + V_v(\vec{r}) + V_{v'}(\vec{r}) \right) \psi_{v'} d^3r$$

(49)
If the total volume, $\Omega$ is divided into subspaces $\Omega_\nu$ and $\Omega_{\nu'}$ ($\Omega = \Omega_\nu + \Omega_{\nu'}$), with $V_\nu \in \Omega_\nu$ and $V_{\nu'} \in \Omega_{\nu'}$, Eq. (49) can be shown to yield [23,49]:

$$\tilde{T}_{\nu\nu'} = -\frac{1}{2} \int_{\sigma} \left( \psi_\nu \nabla \bar{\psi}_{\nu'} - \psi_{\nu'} \nabla \bar{\psi}_\nu \right) d\sigma + E_\nu \int_{\Omega_\nu} \psi_\nu \psi_{\nu'} d^3r + E_{\nu'} \int_{\Omega_{\nu'}} \psi_{\nu'} \psi_\nu d^3r$$

(50)

where $E_\nu = \langle \psi_\nu | \hat{H} | \psi_\nu \rangle$ and $E_{\nu'} = \langle \psi_{\nu'} | \hat{H} | \psi_{\nu'} \rangle$. Choosing $\Omega_\nu$, $\Omega_{\nu'}$, and the separation surface such that $\int_{\Omega_\nu} \psi_\nu \psi_{\nu'} d^3r = \int_{\Omega_{\nu'}} \psi_{\nu'} \psi_\nu d^3r$, one finds:

$$\tilde{T}_{\nu\nu'} = T_{\nu\nu'}^B + \frac{1}{2} S_{\nu\nu'} (E_\nu + E_{\nu'})$$

(51)

where

$$T_{\nu\nu'}^B = -\frac{1}{2} \int_{\sigma} \left( \psi_\nu \nabla \bar{\psi}_{\nu'} - \psi_{\nu'} \nabla \bar{\psi}_\nu \right) d\sigma$$

(52)

and $S_{\nu\nu'} = \int \psi_\nu \psi_{\nu'} d^3r$ is the overlap between both orbitals. Due to this overlap, $S_{\nu\nu'}$, the effective hopping between the Löwdin’s orbitals, $\phi_\nu$ and $\phi_{\nu'}$, associated with $\psi_\nu$ and $\psi_{\nu'}$ is given, to the smallest order in $S_{\nu\nu'}$, by $\langle \psi_{\nu'} | \hat{H} | \psi_\nu \rangle - \frac{1}{2} S_{\nu\nu'} \langle \psi_\nu | \hat{H} | \psi_{\nu'} \rangle$, which Eq. (51) shows to be, $T_{\nu\nu'}^B$.

We conclude that, in the dimer approximation, for short-range potentials and for the limit $S_{\nu\nu'} \to 0$, the hopping integral between Löwdin’s orthogonal orbitals, $T_{\nu\nu'}$, is given by the Bardeen’s tunneling current between orbitals $\psi_\nu$ and $\psi_{\nu'}$ (Eq. (52)).

Going beyond short-range potentials and any value of the overlap can be done as follows [49]. Consider the dimer approximation with the following hamiltonian:

$$\hat{H} = -\frac{1}{2} \nabla^2 + U_\nu(\vec{r}) + U_{\nu'}(\vec{r})$$

(53)

Assume that orbital $\psi_\nu$ and $\psi_{\nu'}$ satisfies the following equations:

$$\left( -\frac{1}{2} \nabla^2 + U_\nu \right) |\psi_\nu\rangle = E_\nu |\psi_\nu\rangle$$

(54)

$$\left( -\frac{1}{2} \nabla^2 + U_{\nu'} \right) |\psi_{\nu'}\rangle = E_{\nu'} |\psi_{\nu'}\rangle$$

(55)

Define $S_{\nu\nu'} = \langle \psi_\nu | \psi_{\nu'} \rangle$ and introduce Löwdin’s orbitals:

$$\phi_\nu = \lambda \psi_\nu + \mu \psi_{\nu'}$$

(56)

$$\phi_{\nu'} = \mu \psi_\nu + \lambda \psi_{\nu'}$$

(57)
where \( \lambda = (1/\sqrt{1+S} + 1/\sqrt{1+S})/2 \) and \( \mu = (1/\sqrt{1+S} - 1/\sqrt{1+S})/2 \). Then, Hamiltonian (53) can be written as follows, in a second quantization language:

\[
\hat{H} = \sum_{\sigma} (E_{\sigma} \hat{n}_{\sigma} + E_{\sigma}^* \hat{n}_{\sigma}^*) + T_{\sigma\sigma'} \sum_{\sigma} (c_{\sigma\sigma'}^+ c_{\sigma\sigma} + c_{\sigma\sigma}^+ c_{\sigma\sigma'})
\]  

(58)

where

\[
E_{\sigma} = \langle \phi_{\sigma} | \hat{H} | \phi_{\sigma} \rangle = \epsilon_{\sigma} - \frac{1}{2} \left[ 1 - \left( 1 - S_{\sigma\sigma}^2 \right)^{1/2} \right] (\epsilon_{\sigma} - \epsilon_{\sigma'}) - S_{\sigma\sigma} T_{\sigma\sigma'}
\]

(59)

\[
T_{\sigma\sigma'} = -\frac{1}{1 - S_{\sigma\sigma'}^2} \frac{S_{\sigma\sigma'}}{2} (\epsilon_{\sigma} + \epsilon_{\sigma'}) + \frac{1}{1 - S_{\sigma\sigma'}^2} \tilde{T}_{\sigma\sigma'}
\]

(60)

and \( \tilde{T}_{\sigma\sigma'} = \langle \psi_{\sigma} | \hat{H} | \psi_{\sigma'} \rangle \) is given by Eq. (51) for short-range potentials. In general, for long range potentials, \( T_{\sigma\sigma'} \) can be approximated by [23]:

\[
\tilde{T}_{\sigma\sigma'} = \gamma T_{\sigma\sigma'} + \frac{1}{2} S_{\sigma\sigma'} (\epsilon_{\sigma} + \epsilon_{\sigma'})
\]

(61)

where \( \gamma \) is a constant around 1.2–1.5. This yields:

\[
T_{\sigma\sigma'} \approx \frac{\gamma T_{\sigma\sigma'}^8}{(1 - S_{\sigma\sigma'}^2)}
\]

(62)

which generalizes \( T_{\sigma\sigma'}^8 \) to cases of long-range potentials and large overlaps.

Eq. (62) is only valid for one effective orbital per atom. In general, more than one orbital per atom has to be used, and the simple approach developed above cannot be followed. In this general case one is forced to use numerical methods to calculate the effective hopping, \( T_{\sigma\sigma'} \), between these orbitals, even using a dimer approximation. Basically, one defines first the Löwdin’s orthogonal basis: \( \phi_\mu = \sum_{\mu'} (S^{-1/2})_{\mu\mu'} \psi_{\mu'} (\mu \equiv v, v') \) and in a second step, calculates \( T_{\sigma\sigma'} \), as

\[
\int \phi_\nu \left( -\frac{1}{2} \nabla^2 + U_\nu (\vec{r}) + U_{\nu'} (\vec{r}) \right) \phi_{\nu'} d^3r
\]

(63)

where \( U_\nu (\vec{r}) \) and \( U_{\nu'} (\vec{r}) \) are the DF-atomic potentials associated with orbitals \( \psi_\nu \) and \( \psi_{\nu'} \).

Fig. 6. Hopping integrals (eV) between different orbitals for a Pd-dimer.
Fig. 6 shows $T_{\nu\nu'}$ for different orbitals of a Pd–Pd dimer. The hopping elements are shown between distances of 2.5 Å and 12 Å. A discussion of the behaviour of these hoppings integrals for long distances will be given below.

Up to this point, we have shown how to calculate $T_{\nu\nu'}$ in a dimer approximation. We shall also discuss below, Section 4, that this yields accurate tunneling currents for long and intermediate tip–sample distances. For short distances, however, a better approach is needed, as seen in Section 4. This is provided by a cluster approximation; in this case, instead of the dimer, a cluster embedding the atoms of the orbitals $\nu$ and $\nu'$ has to be used. Obviously, this is more demanding than a simple dimer, and first-principles codes have to be used to obtain $T_{\nu\nu'}$.

3. Image potential

Image potential plays a crucial role in the interface tip–sample barrier [10,50], modifying substantially the tunneling currents. We first discuss general concepts in a planar geometry, a case well adapted for systems where an extended basis formalism is used. In a second step we discuss the image potential from the point of view of a local basis formalism and analyze how to introduce its effects in the tunneling integrals, $T_{\nu\nu'}$.

3.1. Planar geometry. Extended basis

We start by considering the simple 1D case of Fig. 7, with a square barrier between the tip and sample. In this introductory model, electrons at the Fermi level see a constant barrier $W_0$, with a width $s$. The tunneling current of this case depends on $s$ as follows:

$$I_t = I_0 e^{-2\sqrt{2W}}$$

Obviously,

$$W = \frac{1}{8} \left[ \frac{d(\ln I)}{ds} \right]^2$$

an equation that can be used, in general, as the definition of the apparent barrier height $\phi_{ap}$ between tip and sample [51,52].

In a more general 3D case (see Fig. 8), the interface potential presents three regions, the two surfaces of the tip and sample respectively and the proper interface region. Around the
surfaces, for two metals with the same workfunction, the potential is controlled by Hartree, exchange and short-range correlation interactions; at the interface region, the potential is dominated by the image potential, a non-local correlation effect [53]. It is convenient to start analyzing the image potential effects by considering a simple model, similar to the one presented in Fig. 7 but assuming that the barrier height, \( W \), depends on \( s \) as follows [10]:

\[
W = W_0 - \frac{\ln 2}{s}
\]

(66)

In this equation \(-\ln 2/s\) is an effective “image potential” mean barrier height, that modifies the tunneling current between the two surfaces in the following way:

\[
I_t = I_0 e^{-2s\sqrt{2(W_0 - \ln 2)/s}}
\]

(67)

if we assume that \( I_0 \) does not depend on the distance, \( s \) [54]. For long tip–sample distances, \( \ln 2/s \ll W_0 \), we can write:

\[
-2s\sqrt{2\left(W_0 - \frac{\ln 2}{s}\right)} \approx 2s\sqrt{2W_0} - \sqrt{2}\frac{\ln 2}{\sqrt{W_0}}
\]

(68)

in such a way that

\[
I_t = I_0 e^{-2s\sqrt{2W_0}\sqrt{\ln 2/W_0}}
\]

(69)

This equation shows that, at very long tip–sample distances, the main effect of the image potential is to increase the tunneling current by a constant factor \( e^{\ln 2/\sqrt{W_0}} \). Notice that the apparent barrier, Eq. (65), is still \( W_0 \), which indicates that, even if the tunneling current is increased a lot, the apparent barrier is still the same of a system having the “ideal” barrier, \( W_0 \), with no image effects included.
It is worth looking at the order of magnitude of the factor $e^{\ln 2 \sqrt{2/W_0}}$, for $W_0 \approx 4.5$ eV (0.16 a.u.), its value is around 10, indicating the great importance that image effects have on the tunneling current.

More sophisticated calculations of the image potential effects on the tunneling currents for two planar surfaces have been performed by Pitarke et al. [51]. Fig. 9 presents the case of two metals with workfunctions of 4.8 eV; in this figure, the apparent barrier height (Eq. (65)), the maximum barrier height and the “mean” barrier height are presented as a function of the tip–sample distance (this is practically the distance between the “jellium” edges of both metals; see Fig. 8). The mean barrier height is defined by the equation:

$$\bar{W} = \frac{1}{\Delta s} \int_{s_1}^{s_2} W(z)dz$$

(70)

where $\Delta s = s_2 - s_1$ is the distance between the classical turning points of the $W(z)$ potential.

Several things should be commented about Fig. 9. First, notice that the apparent barrier height is larger than the maximum barrier height, and both are larger than the mean barrier height. The maximum and the mean barrier heights are zero, for $s \approx 1.7 \text{ Å}$; the apparent barrier height being zero, however, for $s \approx 0$. On the other hand, the apparent barrier height is close to $W_0$ for $s > 6 \text{ Å}$, in agreement with the results of the simplified model discussed above [55,52,56].

Fig. 10 shows $\ln(I(s)/I(0))$ for the same model of Fig. 9 ($I(0)$ is the tip–sample current for $s = 0$). Notice that $\ln(I(s)/I(0))$ presents a maximum at $s = 0$, while it tends to $-2s\sqrt{2W_0}$ for $s \to \infty$. In the same figure, we also show $\ln(I(s)/I(0))$, the straight line $-2s\sqrt{2W_0}$, for the simple case of Fig. 8, when image effects are neglected and a constant square barrier model is used at the interface. For the sake of completeness, the tunneling current calculated using, for each surface, the Lang–Kohn short-range potential [57] given by a DFT–LDA approximation for a jellium model is also shown. This simple model presents also a maximum at $s = 0$, and the same limit (the straight line $-2s\sqrt{2W_0}$) for $s \to \infty$.

---

**Fig. 9.** Apparent barrier height (full line), maximum barrier height an mean barrier height as a function of the tip–sample distance, for the model of Fig. 8 (taken from Pitarke et al.).
Fig. 10 summarizes, in a very descriptive way, our previous discussion about the effect of the image potential in the tunneling currents. For a constant square barrier, the straight line \(-2s\sqrt{W_0}\) defines the \(\ln(I(s)/I(0))\) behaviour. Short-range potentials, like the one defined by the DFT–LDA Lang–Kohn solution for the jellium model, increases a little the tunneling currents in the way shown in Fig. 10, with the largest increase (roughly a factor of three) in the tunneling currents for \(s \to 1\).

Image potential effects are dramatic when we compare the tunneling currents of the case presented in Fig. 9 with the two previous cases (see Fig. 10): at very long distances \((s \to \infty)\), the tunneling current is increased by the image potential by three orders of magnitude with respect to both the constant square barrier model \((2 \times 10^3\) larger) and the DFT–LDA short-range surface potential model \((0.7 \times 10^3\) larger). For a typical distance of 5 Å, we find an increase of the tunneling currents by a factor of 80 (with respect to the square barrier) or 26 (DFT–LDA surface potential model). These numbers show the extreme importance that the image potential has in the tunneling currents. One could argue, however, that for a planar surface and a rather sharp tip, image potential effects should be smaller. Although more sophisticated calculations, similar to the previous ones but considering a sharp tip, are needed to validate our conclusion, we should comment that tips are assumed typically to have a pyramidal geometry and that, in this case, image effects are still important. On the other hand, calculations based on the local orbital basis approach (see Sections 3.2 and 4) suggest that a tip having a pyramidal geometry creates image potential effects similar to the ones discussed above for two planar surfaces.

We conclude that an extended basis calculation using Bardeen’s approach, based on the solution of uncoupled surfaces (sample and tip) with short-range potentials (image potential effects neglected) is bound to underestimate the tunneling currents by several orders of magnitude.
3.2. Local orbital basis approach

Before analyzing explicitly how to introduce image potential effects in the local orbital approach, it is convenient, as a first step, to discuss in this formalism how tunneling currents depend on the tip–sample distance.

The starting point is the equation:
\[
\frac{dI}{dV} = \frac{4\pi e^2}{h} \text{Trace}\{\hat{D}^R_{TT} \hat{T}_{TS} \rho_{SS} \hat{T}_{ST} \hat{D}^R_{TT} \rho_{TT}\}
\]

(71)

with \(D_{TT}\) given by Eq. (32). Assume to have only one active orbital, \(\nu_0\), in the tip (in the apex atom). Then, \(D_{TT}\) is a number given by
\[
\hat{D}^R_{TT} = 1 \left\{ 1 - G^{R(A)}_{\nu_0\nu_0} \left[ \sum_{\nu, \mu} T_{\nu_0\nu} G^{R(A)}_{\nu_\mu} T_{\mu_0\mu} \right] \right\}
\]

(72)

and Eq. (72) reads:
\[
\frac{dI}{dV} = \frac{4\pi e^2}{h} \frac{\sum_{\nu, \mu} T_{\nu_0\nu} \rho_{\nu_0\nu} T_{\mu_0\mu} \rho_{\mu_0\mu}}{1 - G^{R}_{\nu_0\nu_0} \left[ \sum_{\nu, \mu} T_{\nu_0\nu} G^{R}_{\nu_\mu} T_{\mu_0\mu} \right]^2}
\]

(73)

Assuming \(G^{R}_{\nu_0\nu_0} \approx -\pi \rho_{\nu_0\nu_0}\) and \(G^{R}_{\nu_\mu} \approx -\pi \rho_{\nu_\mu}\), we obtain the following equation [30]:
\[
\frac{dI}{dV} = \frac{8e^2}{h} \frac{x}{(1 + x)^2}
\]

(74)

where \(x = \pi^2 \sum_{\nu, \mu} T_{\nu_0\nu} \rho_{\nu_0\nu} T_{\mu_0\mu} \rho_{\mu_0\mu}\). This equation shows the general behaviour of \(dI/dV\) for very long tip–sample distances, \(x \to 0\) and \(dI/dV \to 8e/hx\); in this limit \(x \sim \pi^2 \rho_{\nu_0\nu_0} T_{\nu_0\nu_0}^2\) (\(\rho_{\nu_\mu}\) being an effective density of states of the active orbitals in the sample), and \(dI/dV\) should decrease as \(T_{\nu_0\nu_0}^2\) (an exponential behaviour, \(e^{-2x\sqrt{2W_0}}\) (see below)). For short tip–sample distances, \(dI/dV\) presents a maximum for \(x = 1\); in this case \(dI/dV = 2e/h\) the conductance quantum associated with one channel (in our case, the channel for the active tip orbital). This discussion shows that \(\ln(dI/dV)\) (Eq. (74)) presents qualitatively a behaviour similar to the one found in Fig. 10 for two planar surfaces: in our local-orbital basis case, the maximum in \(\ln(dI/dV)\) appears for the first contact (\(s = 0\) in Fig. 9) between the tip and the sample [58]; if the tip is pushed against the sample, more channels can be opened when more that one atom of the tip make contact with the sample [59,60], the differential conductance increasing with the number of channels. On the other hand, at long tip sample distances, \(\ln(dI/dV)\) behaves as \((\text{constant} - 2s\sqrt{2W_0})\) as in Fig. 10. This discussion shows that the local orbital approach allows us to describe appropriately the qualitative behaviour of the tunneling currents for short tip–sample distances. We should stress that a similar approach has been used to analyze the behaviour of metallic nanowires submitted to a stretched deformation [61]; the results of these references show that this local orbital approach yields a rather accurate description of the conductance between two metals in the region close to their contacts (more about this point will be discussed in Section 4, for an Al/Al contact).

We discuss now the limit of long tip–sample distances in order to see which is the appropriate long distance behaviour for \(T_{\nu_\mu}\). Fig. 11 shows the system we are going to analyze: it consists of two planar surfaces with atoms distributed periodically in the planes.
We assume the two planes to be the last layers of two s-like metals located at a distance \(d\) (the distance between the atomic planes). At very long distances, electrons tunnel from one s-like metal to the other with probabilities proportional to \(e^{-\frac{d}{\sqrt{2W_0}}}\) \((W_0\) being the mean workfunction of the metals) so that the probability amplitude for that process is \(e^{-\frac{d}{\sqrt{2W_0}}}\). If we consider how electrons tunnel in our local basis description, we should think of the two wavefunctions going like \(P_m \psi_m(\vec{r})\) and \(P_{m'0} \psi_{m'0}(\vec{r})\) (this means considering the \(\vec{k} = 0\) states). Then, the tunneling probability amplitude between both wavefunctions would be proportional to \(P_{m0}^* T_{mm'} \), where we fix \(m\) and sum upon the orbitals, \(m'\), of the other plane. What we need is a hopping, \(T_{mm'}\), satisfying the condition:

\[
\sum_{m'} T_{mm'} \sim e^{-d \sqrt{2W_0}}.
\]

(75)

At long distances we expect \(T_{mm'}\) to behave as \(f(d) \times e^{-d \sqrt{2W_0}}\); we are going to determine \(f(d)\) by establishing the condition:

\[
\sum_{m'} f(d') e^{-d' \sqrt{2W_0}} \sim e^{-d \sqrt{2W_0}}
\]

(76)

for \(d \to \infty\). In the limit \(d \to \infty\), we can calculate \(\sum_{m'}\) as follows:

\[
\sum_{m'} f(d') e^{-d' \sqrt{2W_0}} \approx f(d) \sum_{m'} e^{-\left(d + \frac{z^2}{d} \right) \sqrt{2W_0}} \approx f(d) e^{-d \sqrt{2W_0}} \frac{d}{\sqrt{2W_0}}
\]

(77)

Then, from Eq. (76) we deduce that \(f(d) \sim \frac{1}{d}\) for s-like orbitals and that \(T_{mm'}\)

\[
T_{mm'}(s-s \text{ orbital}) \sim \frac{1}{d} e^{-d \sqrt{2W_0}}
\]

(78)

Following Harrison [62], we extend Eq. (78), to any orbital geometry and write:

\[
T_{mm'} \sim \frac{1}{d} e^{-d \sqrt{2W_0}}
\]

(79)
where \( \alpha = l + l' + 1 \), \( l \) and \( l' \) being the angular momentum quantum number associated with the orbitals \( \nu \) and \( \nu' \).

This long distance limit and the calculation for short distances discussed in Section 2.2 (Fig. 6) can be used to obtain an interpolative description of \( T_{\nu\nu'} \) in the dimer approximation; this yields a qualitative behaviour of the tunneling currents in agreement with our results of Fig. 11. Fig. 12 shows this procedure: As \( T_{\nu\nu'} \sim \frac{1}{d} e^{-d/\sqrt{2W_0}} \) is known except for a constant, we displace the long distance behaviour of \( \ln T_{\nu\nu'} \) until having a smooth match of both limits (Fig. 12 shows the case of an s–p hopping for a Pd–Pd dimer). We claim that the hoppings calculated with this procedure and Eq. (31) takes into account the main effects associated with the interface image potential as discussed above. Although a more complete discussion of this point will be deferred until Section 4, we already mention that a calculation along these lines for an Al pyramidal tip approaching an Al surface yields the tunneling currents shown (red squares)\(^1\) in Fig. 10. The agreement between these results and the ones calculated for the two planar surfaces (with work functions appropriate for the Al/Al case) are striking and gives a strong support to the claim that our local orbital basis approach combined with our prescription for the calculation of the hopping integrals provides a very accurate way of introducing image potential effects.

### 3.3. Finite voltage effects

So far we have considered the low bias case, where only the states close to the Fermi level can tunnel through the barrier. For these states the barrier height is defined, in the large distance regime, by the metal work function, and we expect an exponential decay \( e^{-2d\sqrt{2W_0}} \) for the tunneling current. (Notice that the constant factor associated with image effects that controls the absolute value of the current (see Eq. (69)) is not considered in this discussion.)

---

\(^1\) For interpretation of the references in color in Fig. 10, the reader is referred to the web version of this article.
The analysis is more complicated for the large bias case, where an energy dependence of the hoppings has to be included. In simple terms, this dependence comes from the fact that the lower lying energy states have to overcome a larger barrier that reduces the transmission probability. Furthermore, the shape of the barrier changes with the applied bias: upper lying states cross a narrower barrier and their tunneling probability is enhanced.

Starting from the WBK expression for the tunneling probability, we can discuss these bias effects with a simple model (similar to that of Simmons [54]). We assume a triangular barrier that we approximate by an effective square barrier with height $u = \phi + \frac{1}{2}V$ (see Fig. 13). The effective barrier for an electron with energy $E$ is then $\phi - E$, and the hoppings are given by

$$T \propto \frac{1}{d\phi} e^{-d\sqrt{2(\phi - E)}}$$

and the tunneling probability is

$$P \propto e^{-2d\sqrt{2(\phi - E)}}$$

As the work function is usually large compared with the bias, we can approximate the WBK expression for the tunneling probability by

$$P \propto e^{-2d\sqrt{2(\phi - E)}} \approx e^{-2d/2\phi(E-V/2)}$$

This defines an energy window [25,63] for the integration [15],

$$\Delta E = \frac{\sqrt{2\phi}}{2d}$$

that weights more the tip states close to the Fermi level (or those of the sample for $V < 0$). Thus we can include the energy dependence of the hoppings including this exponential factor as a weighting factor in the energy integration:

$$I = \frac{4\pi e}{\hbar} \int_0^{eV} dE Tr \left[ \rho_{SS}(E) D_{SS}(E) t_{ST} \rho_{TT}(E - eV) D_{TT}^*(E - eV) t_{TS} \right] \cdot e^{-V/2}$$

Now, only the electrons with energy $E = \frac{1}{2}V$ have the same transmission probability that we obtained in the previous subsection for the electrons close to the Fermi level.

![Triangular barrier model for the finite bias situation and its approximation by an effective square barrier. Notice that the value of the potential, usually small compared to the work function, has been exaggerated in the figure for the sake of clarity.](image-url)
4. From tunneling to the contact regime: first-principles STM simulations

The discussion in the previous sections shows that reliable calculations for STM currents require a theoretical approach capable to handle at the same time a sophisticated description of the electronic structure of the system and a rigorous treatment of the tunneling process. In this section, we show how to combine standard ab initio plane wave pseudopotential methods for the determination of the atomic relaxations induced in the tip and sample by their mutual interaction, with a calculation of non-perturbative STM currents. The connection between the structural and transport calculations is done through a fast local-orbital DFT technique (Fireball2003), that efficiently maps the electronic hamiltonian determined from the PW calculation, and can be naturally linked with the transport formalism, that is expressed in terms of localized orbitals. The ability of the Fireball code to reproduce, with a properly chosen minimal basis, the results of fully converged first-principles calculations has been extensively checked in a number of systems against fully converged plane-wave DFT calculations (see, in particular, our recent study of the mechanical properties and electrical conductance during the stretching and final breaking of Al nanocontacts). Notice that, although we find the combination with the Fireball code particularly efficient, the formulation for the STM simulations presented here can be easily implemented with any other local orbital code. Closely related approaches, based on the combination of ab initio calculations on local-orbital basis and Keldysh–Green’s function method, have been recently applied to study transport properties in nanocontacts.

Using this approach, a detailed study of the currents from the tunneling to the contact regime has been performed, using both an exact calculation of the hopping elements, $T_{\epsilon\mu}$ (with a cluster that includes the whole tip–sample system) and the simple dimer approximation discussed in the previous sections. These calculations provide a stringent test, particularly for the intermediate and long distance range, that confirms the validity of our approach for a quantitative description of the currents, including image potential effects. In the low tunneling resistance regime, these calculations offer a way to explore the interplay between the current increase due to the atomic relaxations and the saturation effects (due to multiple scattering processes) in the value of the current.

We have focused our discussion on Al(111) surfaces, where extremely large corrugations have been reported, and tentatively attributed to large tip-induced surface relaxations. These results for an Al(111) surface and an Al tip (a four-atom Al tip mimics the basic characteristics of the low-coordinated metallic cluster resulting from sample contamination that can be expected to be formed at the tip apex) indicate that the presence of both significant atomic relaxations and saturation effects in the current are associated with the onset of a chemical interaction between the tip apex and the surface atoms, and become relevant in a similar distance range.

4.1. Modelling the tip–sample contact

The initial geometry of the system considered for the calculation of the tip and sample tunneling currents (and also for the determination of the atomic relaxations due to their mutual interaction as discussed below) is shown in Fig. 14: an Al(111) slab with the correct FCC stacking and a tip, represented by a pyramid with four Al atoms, attached to one of the (111) surfaces. In this calculation a $3 \times 3$ periodicity in the direction parallel to the
surface, large enough to decouple the different tips from each other, is introduced. A supercell approach, imposing also periodic boundary conditions in the direction perpendicular to the surface, is used. In this way, the tip interacts with the free surface of the slab (see Fig. 14).

When relaxations induced by the tip–sample interaction are considered, the system is allowed to relax towards its configuration of minimum energy after each step. In this relaxation, the atoms located in the central three layers of the slab shown in Fig. 14 remain fixed. The tip–sample distance is defined as the distance between the central fixed layers in two neighbouring slabs (see Fig. 14). An arbitrary constant is subtracted in order to make the tip–sample distance equal to the normal distance between the tip apex and the topmost surface atoms when the tip and sample are far apart.

In all these current calculations a slab with 11 layers have been considered, keeping the positions from the total energy calculations for the five layer slab and including in between the third and fourth layers six more (111) planes with the correct stacking [64].

4.2. Describing the ideal contact (without relaxations) and validating the dimer approximation. Image potential effects

The main critical point in the simulation of the STM-images is related to the calculation of $H_{ts}$. During normal STM operation, the average tip–surface distance is assumed to be, depending on the voltage and current conditions, in the range between 4 and 8 Å. It is clear that, in the large distance regime, we cannot use in order to calculate the tip–sample hopping elements the strictly localized pseudoatomic orbitals (they are exactly zero for distances larger than the cutoff radius $r_c$) that are needed to have an accurate and efficient
description of the electronic bulk properties, as typical cutoff radii are smaller than the tip–
surface distance (our optimal minimal basis for Al includes s and p orbitals with a cutoff
radii \( R_c = 6.4 \text{ a.u.} \) \[64\]). These hoppings have been determined from the calculation of a
dimer formed by the corresponding tip and sample atom with the Fireball code. In this
calculation pseudoatomic orbitals – generated with a cuttoff radius larger than 20 a.u.,
that reproduce the behaviour of the real free-atom orbitals and, thus, are able to describe
properly the overlap of the tip and sample wavefunctions – have been used. For tip–sam-
ple distances beyond the cuttof radius, the correct distance behaviour for \( T_{nm} \) has been
enforced using the interpolative approach described in Section 3.2.

Notice that there is no inconsistency in the fact that different basis are used for the cal-
culations of the different terms in the conductance (see Eq. (71)). The formulation outlined
in the previous section explicitly separates the contributions coming from the “inter-
atomic” matrix elements (those between an atom in the tip and another atom in the sam-
ple) that are represented by \( \tilde{T}_{ua} \) and the “intraatomic” ones that are described through the
density of states matrices \( \hat{\rho}_t \) and \( \hat{\rho}_s \). With this separation one can concentrate the effort in
the proper calculation of each term. In particular, the charge density matrices have to be
calculated from the corresponding tip and sample “intraatomic” hamiltonians, that
include the structural and electronic effects induced by the mutual interaction between
tip and sample at close distances and are properly described by the pseudoatomic basis
with the finite cutoff. For large distances, where those effects can be neglected, this
approach simplifies significantly the calculation of STM currents as it only requires, apart
from the very quick dimer calculation for the hoppings as a function of the distance, two
independent (smaller) calculations for tip and sample.

As STM-experiments on metal surfaces are usually done with small bias voltages (less
than a 100 meV), the conductance provides a good characterization of the system: Com-
parison with experiments at different bias conditions are easy, as the theoretical current
can be simply estimated as the product of the conductance and the bias voltage. On the
other hand, the values of conductance, particularly when written in terms of its quantum
unit \( G_0 = 2e^2/h \), provide insight into the relation between the structure of the system and
its electronic transport properties. The calculated values for the conductance as a function
of the tip–surface distance for the top and fcc symmetry positions are shown in Fig. 15.
The curves display both the exponential decay and the saturation expected for the large
and short tip–sample distance regimes. A revealing quantitative assessment can be made
when plotting the corresponding normalized logarithmic conductance \( \ln(I(s)/I(0)) \) that
we introduce in Section 3 and comparing it with the exact results for the planar case, as
shown in Fig. 10 (we have assigned the contact \( s = 0 \) for a distance of 2.42 Å, that corre-
sponds to an average of the fcc and top cases). Apart from the expected coincidence in the
linear behaviour for the large distance regime (as similar work functions are used in both
calculations), there is a remarkable agreement for the whole distance range, that confirms
the ability of our approach to describe quantitatively the evolution of the barrier (and thus
the tunneling currents) due to the different effects (short-range exchange and correlation,
image potentials) that are relevant in the different distance regimes.

It is interesting to compare these results, in the pure tunneling regime where relaxations
do not play an important role, with other calculations where the tip and surface electronic
properties have been determined by ab initio methods, like those presented by Hofer et al.
\[11\] for a W tip on a Au(111) surface and Corbel et al. \[73\] for a Cu tip on a Cu(001)
surface. In the case of W on Au(111), they have used a simple Bardeen approach where
the wavefunctions from independent tip and sample calculations are used to calculate the tunneling currents. The conductance at the largest distance considered in that calculation (5 Å) can be estimated from the currents (see Fig. 3 in [11] and the bias considered, \( \Delta V / C_0 = 0.1 \) V). This gives a value of \( 0.3 \times 10^3 G_0 \), that is roughly three orders of magnitude smaller than our result for the Al tip on Al(111) surface for that same distance. Given the results in Fig. 10, we attribute this large discrepancy to the Bardeen’s model used in Ref. [11]: image potential effects reduce considerably the tunneling barrier and increase by two or three orders of magnitude the tunneling current.

The approach for the current calculations by Corbel et al. for a Cu tip on Cu(001) [73] is more similar to the work presented here, as they used scattering theory to calculate the transmission coefficient, by means of Green’s function techniques. The self-consistent hamiltonian is determined on a localized orbital basis set of Slater-type orbitals using the ADF-BAND program. In this case, the conductance values (e.g., 0.01 \( G_0 \) for \( d = 5.6 \) Å, see Fig. 5 in [73]) are a factor of 10 smaller than our results for an Al tip on Al(111). These differences in the absolute value of the conductance could be related to the different nature of the atoms involved, if we recall the atomic-like nature of the tip apex and the significant contribution of extended p states in the Al atom, as opposed to the s states that are dominant in the case of Cu. The results for the behaviour of the conductance in the long distance tunneling regime also agree with those of Corbel et al. We observe an exponential decay with a slope value, \( d \log_{10} \sigma / dz \), of \(-0.82\) (a reduction by a factor of 10 in 1.2 Å), close to the one found in experiments and calculations for other metallic surfaces (\(-0.9\) – 1)).

Although the use of the dimer calculation for the determination of the tip–sample hoppings adds flexibility and a significant speed-up to the STM calculation, it is based on an approximation that has to be validated. The hoppings calculated in this way include the dominant orthogonalization effects of the corresponding two atoms but not the orthogonalization and electrostatic potential contributions coming from the rest of the atoms in
the tip and sample. For the short distance range, where one can expect those effects to be particularly important, the influence of those contributions in the STM current has been checked. For that, the conductances obtained with the dimer approximation and those calculated using the hopping elements that appear in the final self-consistent hamiltonian (with $R_c = 6.4$ a.u.) for the whole tip–sample system have been compared (see Fig. 15).

These calculations show that, for distances smaller than 4.75 Å, although absolute values of the conductance differ due to the different cutoff radii for the orbitals used in both cases [73], the two approaches produce for the different symmetry positions, conductance vs distance curves that have the same shape, leading in both cases to the correct distance behaviour of the corrugation and quantitative differences in the estimated corrugation smaller than 2.5%. Larger corrugations correspond to the calculation with the full hamiltonian (0.29 Å vs 0.22 Å for the maximum value, as discussed in Section 4.3), where the orbitals with shorter range are considered. This increase in the corrugation with the use of more localized basis sets have also been described in STM simulations with semiempirical matrix elements [33]. For larger distances, orthogonalization effects are negligible and the different decay behaviour in the conductance is related to the different extension of the orbitals considered in each calculation (extended free-atom-like orbitals for the dimer calculation and localized pseudoatomic orbitals with cutoff radius equal to 6.4 a.u. in the hoppings in the final self-consistent hamiltonian).

4.3. Tip–sample relaxations and STM currents and corrugations

The relaxed structures of the tip and sample due to their mutual interaction have been determined using CASTEP [74], a standard first-principles plane wave total energy code (For calculation details and convergence criteria, see the original publication [75]). The nature and strength of the short-range tip–sample interaction for the tip approaching the surface in three different symmetry positions (top, hollow fcc and hollow hcp) can be characterized through the analysis of the total energies, the tip–sample forces (see figures 3 and 4 in the original publication [75]), and the atomic relaxations induced (monitored through the apex–surface distance in Fig. 17).

Marked differences between the top and the two hollow positions can be correlated with the onset of significant tip–sample forces and relaxations induced in tip and sample, that are particularly large in the range 4.25–4.50 Å, where the apex–sample distance decreases by 1.17 Å (see Fig. 17) leading to the formation of a strong covalent bond between the two atoms. This distance reduction is mainly due to the upwards displacement, by approximately 0.9 Å, of the Al atom right below the tip. Notice that, before that point, the relaxations induced in both the surface adatom and the tip apex are quite symmetric and of the order of 0.1 Å. For distances smaller than 4 Å, the surface atom is pushed back towards the surface, keeping the bond with the apex atom, that also relaxes back to the initial tip configuration. These atomic displacements lead to a slow decrease of the apex–surface distance, which is close to the Al bulk n.n. distance (2.80 Å). The two hollow positions present a different behaviour. Relaxations on the surface atoms are modest (below 0.05 Å at 4.5 Å), reaching a maximum value of 0.2 Å for a tip–surface distance of 3.3 Å, while the position of the apex atom changes 0.4 Å at that same distance.

Similar results have been found in other studies of the interaction of tips and metallic surfaces. Regarding the atomic relaxations, the behaviour mentioned here (Fig. 17) for the top position agrees quite well (in the absolute value of the relaxations of the apex and the
surface atom, and in the relevant tip–sample distance range) with the results of Hofer et al. [11] for a four-atom W tip and a five-layer slab of Au(111) planes. However, the even larger tip and sample relaxations that they have reported for the hollow position are not seen in Fig. 17.

The curves in Fig. 16 already reflect the role of the tip and sample relaxations, as the conductance of the top and hollow positions become markedly different in the distance range where those relaxations start to be significant for the top position. However, there are other effects, including multiple scattering effects, that have an important role in determining the final value of the conductance. Fig. 18 helps to quantify the contribution of these different factors comparing the final curves for the top and fcc positions with those where either the multiple scattering or the induced relaxations are not included. Multiple scattering effects start to play a role for distances below 5 Å, but they become particularly relevant for the distance range where relaxations take place [11,76]: If they are not included both the values of the conductance and the differences between different positions in the surface are grossly overestimated. Notice, in particular, that these conductance values would lead to an overestimation of the surface corrugation calculated theoretically. On the other hand, the proper calculation (including multiple scattering) for an ideal tip and sample (where we keep their structure when they are far apart for the whole distance range) shows that relaxations play the crucial role in the differences observed between the top and hollow positions.

Finally, we consider the influence on the corrugation of the STM-images of the tip and sample relaxations induced by their mutual interaction. The corrugation, defined as the height difference between the top and hollow positions in a constant current image, can
be extracted from the corresponding current vs distance curves for these two sites (see Fig. 16). The evolution of the corrugation as a function of the tip–sample distance is displayed in Fig. 19. These results show that, in the high resistance tunneling regime (tip–sample distances larger than 4.75 Å), the corrugation has a simple exponential behaviour, decreasing by a factor of 5 in 1 Å. Below that distance, the tip and sample relaxations

Fig. 17. Apex–surface distance as a function of the tip–sample distance for different symmetry positions in the Al(111) surface. The straight line shows the apex–surface atom distance if no relaxations of tip and sample were allowed. The faster change in the top case is mainly due to the upwards displacement of the surface atom below the tip.

Fig. 18. Conductance (in units of $2e^2/h$) as a function of the tip–sample distance for the top and fcc positions, showing the effects of multiple scattering and atomic relaxations.
produce a complicated behaviour in the corrugation. The important tip and sample relaxations induced in the top position for distances smaller than 4.75 Å produce a significant change in the exponential behaviour of the corrugation, that increases from 0.05 to 0.175 Å when the tip–sample distance is reduced from 4.5 to 4.25 Å. For smaller distances, the corrugation tends to saturate and reaches a maximum of 0.225 Å around 4.13 Å.

According to these results, corrugations of a few hundredths of Å are characteristic of metallic surfaces in the high resistance regime (where no significant relaxations are taking place). Similar values have been reported in other theoretical calculations [73]. Larger corrugations, of the order of few tenths of Å, are associated with the atomic relaxations induced by the tip–sample interaction. Corrugations of the order of 0.2 Å have been measured experimentally in certain close-packed fcc surfaces [72,71] and have also been attributed to tip and sample relaxations in the case of a W tip on an Au(111) surface [11]. Notice that, although the corrugation values shown in Fig. 19 for the relevant distance regime are of the same order (0.2 Å) as those found in Ref. [11], the calculations are significantly different: we speculate that the underestimation of the hoppings due to the Bardeen approximation and the overestimation of the current due to the neglect of multiple scattering effects compensate in order to give currents and corrugations closer to our well-converged results.

In the particular case of Al(111) [71], while most of the corrugation amplitudes measured lie in the range 0.1–0.4 Å, depending on tunneling conditions, corrugations as large as 0.8 Å have been also reported. These high-resolution conditions are very unstable and the tip often undergoes irreversible changes. From the calculations of Ref. [64] one can deduce that those instabilities should appear for distances close to the maximum of the force \( d \approx 3.5–4.25 \) Å. It is clear that relaxations and forces due to the tip–sample interaction depend significantly on the tip structure. While one cannot explain the large corrugations with the tip model presented here, we cannot rule out that a more realistic description of the tip state after the experimental treatment, with a loosely bound atom
or small metallic cluster at the apex, could lead to a stable operation with a significant enhancement of the corrugation. These soft modes of the tip apex have been also suggested to explain the large energy dissipation observed during some frequency-modulation atomic force microscopy experiments [77].

5. STM-images for adsorbates on metal and semiconductor surfaces

Sections 2–4 have shown that a rigorous calculation of the tip–sample tunneling currents is not an easy task. For a general formulation (including short tip–sample distances) one needs to introduce electron multiple scattering processes and image potential effects. Our discussion of Sections 3 and 4 shows that those two effects can be accurately included in a local orbital basis formalism, using Eq. (31), a local orbital DF-code for calculating the density of states and Green’s functions matrixes of the tip and the sample, as well as a dimer approximation for calculating the electron hopping elements, \( T_{TS} \), between the tip and sample orbitals.

This section is devoted to discussing a couple of examples, Pd(111)–O [75] and Ga/Si(111)-6 \( \times \) 1 [78], where this approach has provided a very efficient way of understanding the role played by the electron orbitals in the STM-image of adsorbates on semiconductors or metals and of yielding an accurate description of the surface stoichiometry.

5.1. Oxygen on Pd(111): contrast image dependence on tip structure and d-orbitals

This case provides an interesting example of the role played by different electron orbitals in the tunneling process [79,80]. In this section, we follow Ref. [75] and discuss some STM-experiments for the STM-images of atomic oxygen adsorbed on Pd(111), which forms a \( 2 \times 2 \)-periodic structure. We should show that a theoretical analysis of these images, based on the discussion presented above, allows us to understand the contribution to the contrast image of various tips and surface orbitals, which is crucial for a correct interpretation of STM-images.

Fig. 20 summarizes the results of two different experiments addressing the same O/Pd(111)-\( 2 \times 2 \) system. In both experiments the Pd crystal was cleaned by cycles of ion-sputtering and annealing. In the first case the atomic oxygen layer was prepared by thermal dissociation of a layer of adsorbed molecules from 60 K to 160 K, which produced patches or islands with internal \( (2 \times 2) \) ordering. Large, ordered areas of the same structure were produced in the other set of experiments by exposure to molecular oxygen gas at room temperature. In both cases STM-images were acquired at low temperature. The image of Fig. 20a was obtained with a Pt tip and low bias (\( V = -0.059 \) V and \( I = 16.9 \) nA), while the data of Fig. 20b correspond to an Ir tip and a larger bias (\( V = -1.4 \) V and \( I = 0.7 \) nA).

The crystallographic orientation of the surface was determined from images showing the atomic corrugation in two consecutive terraces separated by monoatomic steps. In this manner, the low coverage image presented in Fig. 20a, where both the oxygen and surrounding metal atoms are imaged, can be interpreted in terms of oxygen atoms adsorbed on hollow fcc sites and producing a depression. These same interpretation can be applied to the denser arrays shown in Fig. 20b where the oxygen atoms are responsible for the dark areas in the image. However, apart from this basic feature, common to most of the experimental images (cases where O atoms are imaged as protrusions are also observed [81] and will be discussed below), the shape of the depression associated to the oxygen
atom in both topographs is substantially different: triangular in Fig. 20a, and circular in Fig. 20b. The differences in the contrast structure of the two images are manifested more quantitatively in the line scans shown as insets in Fig. 20, taken along the same crystallographic direction. As can be seen, the maxima in the images (brighter spots) are associated with either top or hollow hcp positions, showing that the naive identification of protrusions with atomic positions is not always justified. Relevant to the present discussion is the shift of the contrast maximum from the top site, near the center of the scan line in (a) at low bias, to the hcp hollow site in the case of the high bias image shown in (b).

Fig. 20. STM-images for O adsorbed on Pd(111) forming a $2 \times 2$ superstructure. (a) Low coverage limit measured with a Pt tip, $V = -0.059$ V and $I = 16.0$ nA; the underlying metal surface is also shown. (b) Monolayer coverage measured with a Ir-tip, $V = -1.4$ V, $I = 0.7$ nA.
The tunneling current flowing through the tip–sample system (in the limit of small $T$, as corresponds to the tip–sample distances of the experiments, see below) can be written as follows (see Eq. (31), for $D_{TT}^A \simeq D_{TT}^A \simeq 1$):

$$I = \sum_{i,j} I_{ij}$$

$$= \sum_{i,j} \frac{4\pi e}{\hbar} \int_{-\infty}^{\infty} \mathrm{d} \epsilon [f_S(\epsilon) - f_T(\epsilon)] \langle \hat{\rho}_{SS} \rangle_{ij}(\omega) \sum_{p,q} \left[ \langle \hat{T}_{ST} \rangle_{ip}(\hat{\rho}_{ST})_{pq}(\omega - eV)(\hat{T}_{ST})_{qi} \right]$$

(85)

where $V$ is the bias voltage. $I_{ij}$ represents the tunneling current associated with the sample $(ij)$-bond; for $i = j$, $I_{ii}$ represents the current through the $i$-orbital. Similarly, $I = \sum_{p,q} I_{pq}$, $I_{pq}$ being the current associated with the $(pq)$-bond.

The relaxed structure and electronic properties of the isolated tip and sample have been determined using CASTEP\cite{74}. In the O/Pd(111)-2 × 2 relaxed geometry, the three Pd atoms close to the O have in-plane and upwards movements of 0.05 Å, the fourth Pd atom on the unit cell moves downwards 0.02 Å (with respect to their positions in the clean surface) and the O stands 1.12 Å above the surface, in agreement with previous calculations\cite{82}.

The tip apex was simulated by a tetrahedral four-atom cluster coupled to the (111) slab of the corresponding bulk material (Pt or Ir). Tip chemical composition was simulated with either a pure Pd tip or by replacing the apex Pd atom by O. As mentioned above, Fireball was used in the relaxed structure to calculate the density of states matrixes. The hopping elements, $(T_{TS})_{pq}$, have been calculated using the dimer approximation.

Fig. 21a shows the calculated O/Pd(111)-2 × 2 constant current STM-images for a tetrahedral Pt tip, a negative bias of 60 meV, and a tip–sample current of 58 nA (giving a contact resistance of $\sim 10^6$ Ω, in agreement with Fig. 20a conditions). The inset of Fig. 21a shows the corrugation along the relevant scan line. The good agreement between the calculations and the experimental data in Fig. 20a is evident. The shape of the STM-image, which depends critically on the relative corrugation heights of the top and hcp positions along the scan line, is also correctly predicted. The main result is that when the corrugation height of the top position is higher than that of the hcp site a triangular

Fig. 21. Calculated O/Pd(111)-2 × 2 STM-images for (a) Pt-tip, $V = -0.060$ V and $I = 58$ nA; (b) Ir-tip, $V = -1.0$ V and $I = 158$ nA.
shape is obtained, as in Fig. 21a; when both points have the same corrugation the shape of the oxygen depression becomes circular (STM-images calculated for a Ir tip, $V = -1.0$ V (positive) and $I = 158$ nA, Fig. 21b). Finally, when the corrugation at the hcp site is higher than that over the top site, the triangular shape re-appears but inverted relative to the one of Fig. 21a [83].

From these observations one can analyze all the O/Pd(1 1 1)-(2 × 2) STM-images characterized by a depression at the position of the O atoms and the triangular or circular shapes around them in terms of the current profiles along the scan line.

Fig. 22a shows different components of the tunneling current calculated along the scan line for a constant tip–sample distance of 6.45 Å defined as the distance between the centers of the apex atom and the last Pd(1 1 1) layer and corresponds to the top position of the inset of Fig. 21a. The magnitude of the total current is basically given by the current through Pd–Pd orbitals. The role of the O orbitals is basically to create a large depression in the STM current on top of the O atoms. The bonding between the Pd-p$_z$ and the O-p$_z$ orbitals yields the largest contribution to this image depression. On the other hand, as mentioned above, the shape of the image around the O is determined by the relative intensities for the top and hcp positions. Fig. 22a shows that the currents through the O orbitals are the same for both sites. The currents through the Pd (d–s or d–p)-bonds, however, are responsible for the shape, round or triangular of the O depression, indicating that the symmetry of the STM-image depends crucially on the Pd d density of states. This finding also explains the dependence of STM-images with the tip–sample distance: for large distances the contribution of the Pd-d orbitals to the tunneling current decreases more rapidly than the contribution from the more extended s and p orbitals and tend to make the Pd (d–s or d–p) orbital contributions to the current negligible. This explains why the shape of the STM-image around O for large distances is a circle; notice that, although the Ir-case of Fig. 22b has a contact resistance os $3 \times 10^6$ Ω (with $d \approx 6.5$ Å), the experiment of Fig. 20b has a contact resistance of $2 \times 10^9$ Ω. This shows that $d$ is around 9 Å in this case and explains the circular shape of the depression associated with the O-image.

Fig. 22. Different components of the total tunneling current along the top-top scan line (see Fig. 21) for (a) $d = 6.45$ Å, $V = -0.060$ V and Pt-tip; (b) $d = 6.45$Å, $V = -1.0$ V and Ir-tip.
A full understanding of the STM-image dependence on the applied bias (Fig. 21b) can be obtained by analyzing the current components, $I_{ij}$, as shown in Fig. 22b. Comparing Fig. 22a and b we see that the main effect of the applied bias is to change the shape of the current profile associated with the Pd (d–p or d–s) bonds: this component is now very flat, yielding a total STM current with similar values at the top and hcp sites, this fact adding a new argument for the circular shape of the O-depression of Fig. 20b. As discussed below, this behaviour is independent of the tip metal atom, Pt or Ir.

Finally, we discuss how the STM-images depend on the chemical nature of the tip. First of all, it was found that different metal atoms at the tip apex do not change appreciably the simulated images. Indeed, nearly identical tunneling current with Pd and Pt tips are obtained, as shown in Fig. 23a. Non-metallic tips, specifically oxygen terminated tips were constructed by replacing the Pd atom at the tip apex by an O atom. For the metal terminated tips the tunneling current is mainly associated with the apex atom (Fig. 23a) and in particular, with its $p_z$ orbital. This explains why different metal atoms yield similar images and why tips of the same metal, but having different geometries, may have quite different tunneling currents due to changes in the $p_z$-density of states. However, oxygen contamination of the tip introduces significant qualitative changes (see Fig. 23b): oxygen atoms are now imaged as bright protrusions located on top of the atom \[81\].

These changes can be easily understood in terms of the $I_{pq}$ current components: while for the metal tip most of the current is associated with the apex atom, for the oxygen terminated tip the apex atom contributes a very small fraction, and most of the current goes through the three metal atoms on the layer above the oxygen, a result in agreement with previous approximate calculations using the scattering theory formalism and the Extended Debye–Huckel approximation \[84\].

In conclusion, based on the approach published in this paper, it has been shown that the contrast and symmetry of STM-images of O/Pd(111)-(2×2) that are found experimentally to vary from a triangular to a circular depression, or even to a protrusion when the tip is contaminated with O, can be explained by the relative contributions of short
range Pd-d orbitals and long range Pd-s and Pd-p orbitals. The former are accessed at short tip–surface distances and low bias voltages. The latter are dominant at large separations and/or larger bias voltages. The results also show that for a clean tip, the sample oxygen-p$_z$ orbital is mainly responsible of the image depression at the oxygen position. All these results conclusively prove that the shape and symmetry of the image (triangular or circular), depend crucially on the contribution to the tunneling current of the Pd-d orbitals that can be reduced or increased by changing the tip–sample distance or the applied bias. Additionally, it has also been shown that the protrusion found on top of oxygen for an oxygen contaminated tip is mainly associated with the tip Pd–O bonds.

5.2. Self-assembly of metallic wires on vicinal surfaces: Ga/Si(112)-6×1

Metallic one-dimensional wires have attracted a lot of attention. In particular, deposition of metal atoms on vicinal Si surfaces is expected to form “atomic wire arrays” because incoming atoms are expected to adsorb preferentially at the high reactive step edges \[85,86\]. Ga on Si(112) represents a typical example of those possible structures and, in this section, we discuss how the theoretical approach discussed in this paper has been used to analyze the STM-images to obtain the surface structure and the stoichiometry of that system \[87\].

Fig. 24 shows a large scale empty state STM-image of the Si(112)-6×1–Ga surface; this 6×1 (or 5×1) periodicity appears to release the compressive strain along [110]. Parallel atomic lines along [1 1 0] are intersected by dark vacancy lines. The vacancy lines are not straight; as discussed below, this is due to the coexistence of 6×1 and 5×1 units in the surface \[88\], and to the presence of intrinsic fluctuations in the vacancy lines. Fig. 25a shows an atomic resolution empty state STM-image of a 6×1 Ga/Si(112) surface. The brightest atomic lines in Fig. 25a are the same atomic lines as imaged by Baski et al. \[88\]. But due to better STM resolution, an atomic line lying in between the brighter lines is also observed (see also Fig. 24). These two parallel atomic lines appear to form a zig-zag pattern, resulting in a zig-zag asymmetry in the vacancy line. In the filled state image,
A relatively bright protrusion is the dominant feature. Registry aligned dual bias images show that this bright protrusion is located in the vacancy lines. Bright protrusions in different unit cells are connected by fuzzy lines along the $\frac{1}{2}10/C221/138$ direction. These STM-images imply that the step-edge decorated structural model [87–89] of the Si(112)-6 $\cdot$ Ga surface needs to be revisited because the vacancy lines in that model have mirror plane symmetry (i.e. mirror plane perpendicular to $\frac{1}{2}10$). We thus have to consider alternative atomic arrangements and allow for the possibility of a different Ga/Si composition ratio of the surface layer.

Extensive DFT and STM-images were performed to identify the precise structure of the Ga/Si(112) surface. New candidate structural models corresponding to Ga coverages ranging from 5 to 11 Ga atoms per $6 \times 1$ unit cell were explored using Fireball. In total, more than 40 new structures were fully relaxed and their STM-images calculated. Fig. 26 shows the surface structure giving the best agreement with the experimental STM-images. As shown in Fig. 26 the Ga atoms in this surface form two parallel rows: a row of
‘step-edge’-Ga atoms adsorbed at the (111)-like step and a second row of Ga atoms adsorbed at the terraces. The two Ga-rows form a zig-zag pattern. Each Ga-row contains 5 Ga atoms per 6 × 1 unit cell, i.e. there is a Ga-vacancy in each row. Vacancies in different Ga-rows line up and form vacancy lines [90]. The creation of Ga-vacancies (and concomitant breaking of the Ga backbonds) is compensated by the formation of a Si–Ga dimer at the step-edge and a neighboring Si–Si dimer in the terrace row; both dimer bonds are parallel to the [110] direction. In this zig-zag structure all the partially-filled Si dangling bonds are replaced by empty Ga dangling bonds, and the surface is semiconducting. Evidently, the system’s desire to passivate all dangling bonds precludes the formation of metallic single-atom wires at the step-edges. Vacancy-line creation is driven by the need to relieve the compressive strain along [110] [89].

Fig. 25c and d show simulated STM-images for the zig-zag model with a top view of the structural model superimposed. The agreement between the experimental and theoretical STM-images is excellent. It shows that the two atomic lines seen in the empty state STM-images are the step-edge and the terrace Ga lines; the zig-zag asymmetry in the vacancy lines observed experimentally is neatly reproduced by the simulated STM-images. In the filled state image, fuzzy lines with a bright protrusion inside the vacancy line are observed, also in agreement with the experimental images. The bright protrusion originates from the Si–Ga dimer in the vacancy lines, while the fuzzy lines are due to Si–Ga bonds on the (111)-like terrace, close to the step-edge.
An independent analysis of the 40 different surface structures as a function of their surface energies \( F = E_{\text{total}} - \mu_{\text{Si}}N_{\text{Si}} - \mu_{\text{Ga}}N_{\text{Ga}} \) confirms that the structure of Fig. 26 is the most stable (see Refs. [78,91]).

It is also interesting to notice that the current structural model of Ga/Si(112) also implies the possibility of intrinsic fluctuations, i.e., fluctuations within each \( 6 \times 1 \) or \( 5 \times 1 \) domain. For example, a closer scrutiny of the empty state STM-image of Fig. 25a reveals that in this pure \( 6 \times 1 \) region the number of step-edge Ga atoms between the vacancy lines fluctuates between four and six while the number of terrace Ga atoms is always five. The fluctuations associated with the step-edge atoms are related to the two possible orientations of the Si–Ga dimers (in Fig. 25(a) there are a total of 9 Si–Ga dimers with the orientation shown in Fig. 25, and 6 Si–Ga dimers with the opposite orientation). This twofold degeneracy is a direct consequence of the \([1 \bar{1} 0]\) mirror plane symmetry of the vicinal substrate, broken by the Si–Ga dimers. No such fluctuations can be discerned from the filled-state images since the observed bright protrusions appear in the middle of the Si–Ga dimer (see Fig. 25d).

In conclusion, the structure and fluctuating disorder of the quasi-1D atom wire arrays that appear for the deposition of Ga on Si(111) can be determined by means of a careful study of the filled and empty state STM-images of the system. This evidence is confirmed by total energy minimization calculations and RBS-experiments [91]. All together, these results are a convincing evidence of the accurate approach that for calculating the STM-images is provided by the method discussed in this paper.

6. Conclusions

A critical review of the different methods used nowadays for the calculation of the tunneling currents and STM-images has been presented in this paper.

In Section 2, we discuss Bardeen’s approach adapted to an extended basis formalism, as well as a Green’s function method and the scattering theory based on a local basis description of the electron states. In all these approaches, the tip and sample electronic properties are calculated independently in a first step (which we assume to have been done using a DF-method) and, in a second step, the tip and sample interaction is switched on and the tunneling currents are calculated by a suitable approach. Although, in principle, the methods discussed in Section 2 provide an exact calculation of the tip–sample tunneling currents, the particular implementation of each case may present approximations that limit the accuracy of the calculated currents. In Section 3 we discuss how the image potential plays a critical role in the interface tip–sample barrier, and how neglecting its effect in the tip–sample coupling can underestimate the tunneling current by several orders of magnitude. We have shown, however, in Sections 3 and 4, how a Green’s function formalism based on a local orbital basis approach can introduce those image effects with a good accuracy. The main approximation in this method is, however, the calculation of the hopping parameters between the orbitals of the tip and the sample, interaction that is obtained using a dimer approximation. Although this approximation can be improved using a cluster (embedding the dimer of each case), the calculations shown in Sections 4 and 5 indicate that, that dimer-approach yields a good description of the tunneling currents. Regarding the scattering theory approach, in Section 2 we have shown it to be equivalent to the Green’s function formalism; there is, however, a limitation in that approach: the difficulty of introducing the corresponding incoming and outgoing wave functions of localized
states (like surface states). The advantage of the Green’s function approach is that all the different contributions from either the bulk or the surface are included automatically in the formalism.

Finally, the results shown in Sections 4 and 5 for different systems (Al(1 1 1)-surface, Pd(1 1 1)–O and Si(1 1 2)–Ga-6·1), calculated with the Green’s function formalism, and including tip and surface relaxations, as well as tip sample chemical interaction show the validity and good accuracy that this method has for obtaining the tunneling currents and the STM-images.

Acknowledgements

This work has been supported by the DGI-MCyT (Spain) under contracts MAT2001-0665, MAT2002-01534, MAT2005-01298 and NAN-2004-09183-C10-07, and the Comunidad Autonoma de Madrid under contract S-0501/MAT/0303.

References