Three-dimensional approach to scanning tunneling spectroscopy and application to Shockley states

Three-dimensional approach to scanning tunneling spectroscopy and application to Shockley states

F Donati[1]**, S Piccoli, C E Bottani and M Passoni**

Dipartimento di Energia and NEMAS—Center for NanoEngineered MAterials and Surfaces, CNISM, Politecnico di Milano, Via Ponzio 34/3, I-20133 Milan, Italy
E-mail: fabio1.donati@mail.polimi.it

Abstract. The problem of the interpretation of scanning tunneling spectroscopy (STS) data is analytically solved using a three-dimensional (3D) transfer Hamiltonian approach. We present an analytical model capable of including both the electronic structure of the sample and the symmetry of the tip states (s, p_z, d_{z^2}, ...) and we discuss the role of these 3D aspects in tunneling. We applied this model to the case of Shockley states. This system, allowing a full analytical treatment, led us to a detailed simulation and comprehension of the tunneling process. A procedure for the recovery of the sample local density of states from STS measurements is then proposed and applied to both the simulated and the experimental STS data of Shockley states. Comparing this approach with other methods proposed in the literature, the importance of considering the 3D aspects in treating and interpreting STS data is demonstrated.

Contents

---

[1] Author to whom any correspondence should be addressed.

IOP Institute of Physics ⦿ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

1. Introduction

The investigation of the local, surface-projected electron density of states (DOS) at the nanometric scale and the possibility of accessing both the occupied and the unoccupied states are two of the most attractive features of scanning tunneling spectroscopy (STS) [1]. Over the last few years, different approaches have been followed to solve the challenging problem of interpreting the raw STS quantities and connecting them to the physical properties of surfaces.

The tunneling current is the result of a non-equilibrium process taking place on a surface in the presence of a tip (with a given electronic structure), and its general relationship with the surface DOS is indeed quite complicated. In principle, making use of the theory of the non-equilibrium Green's function [2], a full computational approach can be developed to simulate the process [3]. Various attempts for more simplified treatments exist. One of the most successful approaches consists in describing the tunneling in the one-dimensional (1D) Wentzel–Kramer–Brillouin (WKB) approximation. From the earliest works [4], it has been clear that both the tunneling current I and, in a more direct way, the differential conductivity dI/dV provide information on the electron DOS of the surface. However, such quantities generally contain other spurious contributions due both to the presence of the tip and to the tunneling process itself. As a consequence, the identification of dI/dV as a function of the applied bias V with the surface DOS is valid only in a short-energy range around the Fermi level. Stroscio *et al* [5] proposed to normalize the differential conductivity to the total conductivity $(dI/dV)/(I/V)$ to qualitatively remove the tunneling process contribution. Ukraintsev [6] proposed a more consistent normalization procedure based on the analytical evaluation of the 1D-WKB transmission coefficient of the tunneling barrier. A more refined treatment was provided by Koslowski *et al* [7], who proposed an analytical normalization procedure based on a suitable combination of dI/dV and I, to quantitatively recover the surface DOS. The possibility of recovering the sample DOS by solving iteratively the corresponding Volterra equation of the second kind in the Neumann scheme was also suggested [7]. Wagner *et al* [8] directly approached the same problem numerically inverting the integral relation between the sample DOS and I (which is a Volterra equation of the first kind) and applied it to the case of organic molecule spectra, neglecting possible tip effects. Passoni *et al* [9] evaluated the effectiveness of these methods on experimental data and considered the effect of the presence of a non-constant tip DOS in dI/dV spectra. The possibility of removing such tip effects has also been demonstrated in principle, exploiting iterative solution methods [10, 11]. Finally, Ziegler *et al* [12] proposed a method to treat and interpret the spectra taken by varying the tip–sample distance in the so-called constant current mode. One of the most attractive features of the 1D-WKB approach is that it 'bypasses' the non-equilibrium situation, connecting in a simple and direct way measured quantities with physical properties. On the other hand, several important aspects cannot be treated within the 1D-WKB approximation. In particular, 3D effects have been included in the WKB framework considering the problem of tunneling through a 3D extended tunneling barrier only within the zero bias approximation [13], while the contributions of a non-vanishing applied bias, of electron states located in different regions of the sample Brillouin zone and of the symmetry of the tip states cannot be easily included. As a consequence, a 3D-WKB theory suitable for STS is not available yet.

In a different framework, the theory of tunneling can be treated within the so-called transfer Hamiltonian (TH) approach, as was originally done by Bardeen [14]. This approach, within its limits of validity [15], opens the possibility for an analytical treatment of the 3D effects.

From a historical point of view, these aspects were included in scanning tunneling microscopy (STM) descriptions first by Tersoff and Hamann [16] and then by Chen [17–19], in the low bias approximation, with the purpose of explaining the origin of atomic resolution in STM. The extension of this model to finite bias has been recently proposed with the aim of providing a solid theoretical framework for the interpretation of STS measurements [20].

We followed the possibilities offered by this approach to present a complete 3D analytical model of tunneling and a specific STS normalization procedure, applied in particular to the physical system consisting of Shockley states with a parabolic dispersion in the presence of a tip state with well-defined angular symmetry. In section 2, we present the analytical description of the tunneling process in a 3D-TH approach, providing an expression for I and dI/dV that is able to take into account both the band structure of the sample surface states and the tip state angular symmetry. In section 3, we consider the case of an ideal Shockley surface state as a sample in the STS measurement, showing 3D-related features in simulated tunneling current. The presented analytical formulation is exploited in section 4, where we propose a normalization procedure for the recovery of the sample local density of states (LDOS). The normalization method is applied to the simulated Shockley state spectra and compared with existing procedures [5–7]. The reliability of the method is also evaluated by testing sensitivity to the model parameters (e.g. estimated tip–sample distance and tip symmetry) and eventually applying the method to experimental STS spectra taken on Cu(111). The concluding remarks are presented in section 5.

2. Theory

We start by summarizing the formulation of the tunneling problem in the TH approximation. Under the quite general assumptions of independent electrons for both the sample and the tip and of weak coupling between the two physical systems, we can define the LDOS $\rho_{S(T)}(E) = \sum_\gamma \delta(E - E_\gamma)$ and the LDOS $L_{S(T)}(E, \mathbf{r}) = \sum_\gamma |\psi_\gamma(\mathbf{r})|^2 \delta(E - E_\gamma)$, where E_γ and ψ_γ are the γth eigenvalue and eigenfunction of the unperturbed single-electron Hamiltonian, respectively. In the following, we will use the notation $\gamma \to \mu$ for the tip states and $\gamma \to \nu$ for the sample states. In the 1D treatment, a difference between LDOS and DOS does not explicitly appear and these two quantities are used almost equivalently. On the other hand, they have to be distinguished whenever we consider a real 3D system. One of the most attractive features of the 1D-WKB approximation is that it leads to a formulation of the tunneling current in the form of a convolution integral of tip and sample DOSs, weighted by the WKB tunneling transmission coefficient [21]. Such an expression is achieved after several simplifying assumptions and it is not valid in the general case, as we will show in the following. The TH formulation for the tunneling current in the zero-temperature, elastic regime [14, 22], using dimensionless units ($E/W_F \to E$, $eV/W_F \to V$ and $z\sqrt{2m_e W_F}/\hbar \to z$, where V is the applied voltage, W_F is the effective work function and e, m_e are the modulus of electron charge and mass, respectively), is given by

$$I = \int_0^V \sum_{\mu,\nu} |M_{\nu,\mu}|^2 \delta(E - V - E_\mu)\delta(E - E_\nu)\, dE. \tag{1}$$

Here, the probability of a transition between the sample state with the energy $E_\nu = E$ and the tip state with the energy $E_\mu = E - V$ is described by means of the Fermi golden rule through the matrix elements $M_{\nu,\mu}$, which depend on the corresponding wave functions ψ_μ, ψ_ν [14].

In the general situation, the evaluation of $M_{\nu,\mu}$ in (1) does not lead to a simple connection between the current and the sample DOS or LDOS [15, 20]. If we expand the tip wave functions into spherical harmonics [18], it is possible to write these matrix elements in the general form [20]

$$M_{\nu,\mu} = \hat{O}_\mu(E, V)\psi_\nu(\mathbf{r_0}, E, V), \tag{2}$$

where $\mathbf{r_0}$ is the center of curvature of the tip apex and $\hat{O}_\mu(E, V)$ is a differential operator that acts on the spatial coordinates of the sample wave functions ψ_ν and depends on both the tip properties and the applied finite bias. If we assume that

(a) the sample wave functions are the solutions of a Schrödinger equation that is separable into Cartesian coordinates (x, y for the in-plane parallel ($\parallel$) directions, which are associated with an eigenvalue $E_\parallel$, and z for the out-of-plane orthogonal ($\perp$) direction, which is associated with an eigenvalue $E_z = E - E_\parallel$).

(b) The potential barrier between the two systems can be well approximated by a (bias-dependent) effective rectangular barrier of height equal to $W_F + V/2$;

we can express the sample wave function in the vacuum region in the form

$$\psi_\nu(\mathbf{r}, E, V) = A_{\nu_\parallel}(x, y; E, V)\chi_{\nu_\perp}(E_z, V, z)$$
$$= A_{\nu_\parallel}(x, y; E, V)C_{\nu_\perp}(E_z, V)\,e^{-k_s(E_z, V)z}, \tag{3}$$

where $A_{\nu_\parallel}(x, y)$ and $\chi_{\nu_\perp}(E_z, V, z)$ are the eigenfunctions for the parallel and perpendicular equations which depend on the quantum numbers $\nu_\perp$ and $\nu_\parallel$, respectively; $k_S = \sqrt{1 + V/2 - E_z}$ is the decay factor for the sample state in vacuum and $C_{\nu_\perp}(E_z, V)$ is a normalization parameter. Although for some special cases such a parameter could be calculated, its dependence on V is in general not known and some assumptions are therefore needed. Here, we neglect the quantity $\partial|C_{\nu_\perp}|^2/\partial V$ and assume that the electric field affects only the out-of-plane component of the eigenfunction; we can write

$$\psi_\nu(\mathbf{r}, E, V) \approx \psi_\nu(\mathbf{r}_\parallel, z = 0, E, V = 0)\,e^{-k_s(E_z, V)z}. \tag{4}$$

Let us now consider, in the spherical harmonics expansion, the tip states s, p_z and d_{z^2} that are generally responsible for the main contribution to the tunneling current [18]. The matrix elements (2) corresponding to these states contain differential operators acting only on the z-coordinate and can be expressed in the functional form [20]

$$M_{\nu,\mu} = M_{\nu,l_m} = \alpha_{l_m}(E, E_z, V)\psi_\nu(\mathbf{r_0}, E, V), \tag{5}$$

where l and m are the quantum numbers related to the tip spherical harmonics expansion and $\alpha_{l_m}(E, E_z, V)$ depends on the particular tip state

$$\alpha_s = \frac{4\pi a_s}{k_T(E, V)}, \tag{6}$$

$$\alpha_{p_z} = \frac{4\pi a_{p_z}}{k_T(E, V)} \cdot \frac{k_S(E_z, V)}{k_T(E, V)}, \tag{7}$$

$$\alpha_{d_{z^2}} = \frac{4\pi a_{d_{z^2}}}{k_T(E, V)} \cdot \left[\frac{k_S^2(E_z, V)}{k_T^2(E, V)} - \frac{1}{3}\right]. \tag{8}$$

Here, $k_T(E, V) = \sqrt{1 + V/2 - E}$ is the decay constant for the tip wave functions and a_s, a_{p_z}, $a_{d_{z^2}}$ are constants related to the relative weight of the corresponding l_m state in the spherical harmonics expansion. Now, introducing the l_m-partial tip DOS ρ_{T,l_m}, obtained by summing over the tip states with the symmetry l_m, we can rewrite the tunneling current as a sum of contributions I_{l_m} coming from each tip state symmetry $l_m = s, p_z, d_{z^2}$:

$$I = \sum_{l_m} I_{l_m}$$

$$= \sum_{l_m} \int_0^V \rho_{T,l_m}(E - V) \sum_\nu \alpha_{l_m}^2 |\psi_\nu(\mathbf{r_0}, E, V)|^2 \delta(E_\nu - E)\, dE. \tag{9}$$

Equation (9) can be expressed in a more transparent form by introducing the partial sample LDOS $L_{S,\nu}(E, \mathbf{r_0}) = |\psi_\nu(\mathbf{r_0})|^2 \delta(E - E_\nu)$, so that

$$I = \sum_{l_m} \int_0^V \rho_{T,l_m}(E - V) \sum_\nu \alpha_{l_m}^2 L_{S,\nu}(E, V, \mathbf{r_0})\, dE. \tag{10}$$

In (10), the tunneling current is expressed as an integral including $L_{S,\nu}$ evaluated at the center of curvature of the tip apex. Usually, it is more interesting to consider this quantity evaluated at the sample surface. Making use of (4), it follows that

$$L_{S,\nu}(E, V, \mathbf{r_0}) \approx L_{S,\nu}(E, V = 0, \mathbf{r_\parallel}, z = 0)\, e^{-2k_s(E_z, V)z} \equiv L_{S,\nu}(E, \mathbf{r_\parallel})\, e^{-2k_s(E_z, V)z}, \tag{11}$$

identifying in this way the sample perturbed partial LDOS in vacuum as a product of an unperturbed partial LDOSs $L_{S,\nu}(E, \mathbf{r_\parallel})$ evaluated at the surface and a tunneling decay function that depends on the bias and the tip–sample distance. This allows us to write (10) as

$$I = \sum_{l_m} \int_0^V \rho_{T,l_m}(E - V) \sum_\nu \tau_{l_m}(E, E_z, V, z) L_{S,\nu}(E, \mathbf{r_\parallel})\, dE, \tag{12}$$

where we introduce the function

$$\tau_{l_m}(E, E_z, V, z) = \alpha_{l_m}^2(E, E_z, V)\, e^{-2k_s(E_z, V)z}. \tag{13}$$

Equation (12) shows a number of similarities to the common 1D-WKB formula. It relates the tunneling current to the tip DOS and sample unperturbed partial LDOSs, each weighted by a generalized transmission coefficient τ_{l_m}. However, even though (12) has the same structure as that of the 1D formulation of the tunneling current, τ_{l_m} here depends not only on the total energy E of the surface states and the tip–sample distance z, but also on the out-of-plane energy E_z and the symmetry of the tip apex l_m, in this way including the relevant 3D features of the tunneling process. Comparing with the 1D-WKB transmission coefficient [8], τ_{l_m} still depends exponentially on z and V. However, the exponential term contains only the E_z energy instead of E, which is the physical quantity contributing to the decaying behavior in vacuum of ψ_ν [16, 23]. Moreover, the 1D coefficient T does not contain the term $\alpha_{l_m}^2$, which arises naturally in this TH formulation and allows us to include the assumed symmetry of the orbitals at the tip apex.

Since the α_{l_m} coefficients are functions of V, E and E_z, τ_{l_m} always differs from the 1D-WKB transmission coefficient. It is therefore important to understand how the transmission coefficients τ_{l_m} depend on the relevant quantities, in particular on the out-of-plane sample energy E_z (see figure 1(a)). In figure 1(a), the total energy is fixed at a certain positive value, equal to
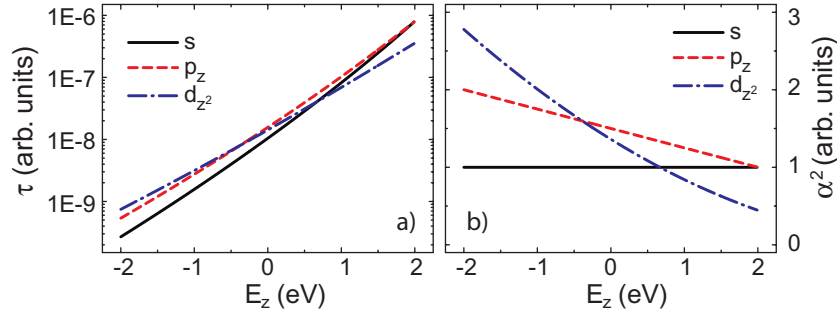
Figure 1. Plots of (a) τ_{l_m} and (b) $\alpha_{l_m}^2$ as a function of E_z ($l_m = s, p_z, d_{z^2}$; $z = 10$ Å; $V = 2$ V; $E = 2$ eV).

the applied voltage ($E = V$), and the behavior of τ_{l_m} as a function of E_z is shown. The out-of-plane energy E_z cannot overcome the total energy E and $E_z = E$ corresponds to a condition for which the in-plane motion energy is zero (i.e. the $\bar{\Gamma}$ point in the projected Brillouin zone). It can be observed that the transmission coefficient is overwhelmed by the exponential decay factor $e^{-2k_s z}$ that strongly damps the tunneling probability of the states with low E_z (the ones that lie far away from the $\bar{\Gamma}$ point) [23]. However, even though this effect is the dominant one, differences are visible when we compare the behavior of τ_{l_m} associated with different tip states. Looking at the functions $\alpha_{l_m}^2$ in figure 1(b), it is clear that for p_z and d_{z^2} states the tunneling probability is enhanced for surface states with lower E_z or in other words, for surface states far from the $\bar{\Gamma}$ point, the probability does not depend on E_z for s tip states. Furthermore, the transmission probability is always enhanced for the p_z states with respect to the s states, whereas the d_{z^2} states damp the contribution when approaching $\bar{\Gamma}$. Therefore, considering sample states with the same total energy, the tunneling is indeed dominated by those that are closer to the $\bar{\Gamma}$ point, but the symmetry of the tip state in general modulates the contributions of states belonging to different regions of the Brillouin zone. This aspect has to be properly considered when facing the issue of quantitatively extracting the sample LDOS from STS experiments.

We now consider the form of dI/dV, which represents the most relevant quantity in STS experiments, by directly differentiating (12):

$$\frac{dI}{dV} = \sum_{l_m} \frac{dI_{l_m}}{dV}$$

$$= \sum_{l_m} \rho_{T,l_m}(0) \sum_{\nu} \tau_{l_m}(E = V, E_z, V, z) L_{S,\nu}(E = V, \mathbf{r}_\parallel) + B_{t,l_m} + B_{s,l_m}, \quad (14)$$

where B_{t,l_m} and B_{s,l_m} are given by

$$B_{t,l_m} = - \int_0^V \frac{\partial \rho_{T,l_m}(E-V)}{\partial E} \sum_\nu \tau_{l_m}(E, E_z, V, z) L_{S,\nu}(E, \mathbf{r}_\parallel) \, dE. \quad (15)$$

$$B_{s,l_m} = \int_0^V \rho_{T,l_m}(E-V) \left[\frac{\partial}{\partial V} \sum_\nu \tau_{l_m}(E, E_z, V, z) L_{S,\nu}(E, \mathbf{r}_\parallel) \right] dE. \quad (16)$$

In (14), the first term contains the most direct information on the surface electron properties. It represents a sum of every sample surface state having a total energy $E = V$

weighted by the transmission coefficient τ_{l_m}, which accounts both for the specific decaying behavior in vacuum of the sample wave functions and for the coupling with the tip states. The term B_{t,l_m} (in (15)) depends on the derivative of ρ_T with respect to the energy and has to be considered only when the tip DOS is non-constant [9]. Experimentally, it is common to try to modify the tip apex electronic configuration by applying a strong voltage pulse during the measurement to achieve a sufficiently smooth tip DOS around the Fermi level. This *modus operandi* usually justifies the possibility of neglecting the contribution described in (15). In a recent work [9], different effects due to the presence of a non-constant tip DOS have been theoretically evaluated in the 1D framework and measured by experiments. Here, we want to focus mainly on the role of the symmetry of the tip apex orbitals in tunneling, and we will not further consider the effects due to its energetic structure, thus neglecting the term B_{t,l_m} in the following, leaving this for future work.

The term B_{s,l_m} (in (16)) involves a sum containing the derivative of the transmission coefficient τ_{l_m}, given by the following relations,

$$\frac{\partial \tau_{l_m}}{\partial V} = [\beta_{l_m,\nu}(E, V) + \gamma_\nu(E, V, z)]\bar{\tau}_{l_m}, \tag{17}$$

where

$$\beta_{s,\nu} = -\frac{1}{2k_T}, \quad \beta_{p_z,\nu} = \frac{1}{2k_S^2} - \frac{1}{k_T^2}, \quad \beta_{d_{z^2},\nu} = \frac{7 - 9k_S^2/k_T^2}{6k_S^2 - 2k_T^2}, \quad \gamma_\nu = -\frac{z}{2k_S}, \tag{18}$$

and consequently, by applying the mean value theorem [7] it is possible to write

$$B_{s,l_m} = \sum_\nu (\bar{\beta}_{l_m,\nu} + \bar{\gamma}_\nu) \int_0^V \rho_{T,l_m}(E - V)\tau_{l_m}(E, E_z, V, z) L_{S,\nu}(E, \mathbf{r}_\parallel) \, \mathrm{d}E$$

$$= \sum_\nu (\bar{\beta}_{l_m,\nu} + \bar{\gamma}_\nu) I_{l_m,\nu}. \tag{19}$$

The terms $\bar{\beta}_{l_m,\nu}$ and $\bar{\gamma}_\nu$ can be calculated for the case of constant tip and sample DOS by following the approaches proposed in the literature [7, 9]. The term B_{s,l_m} has therefore been decomposed into a weighted sum of partial currents $I_{l_m,\nu}$, which represent the contributions to the total current of each sample state ν, tunneling into the tip state with the symmetry l_m. It is a smooth term representing a *background* contribution to $\mathrm{d}I/\mathrm{d}V$, which has to be properly considered for a quantitative recovery of the surface LDOS, as will be discussed in section 4.

In (10)–(14), a simple dependence of I on the total sample LDOS is by no means evident. In order to obtain a very clear connection between the tunneling current and the sample LDOS, we introduce a further assumption:

$$\sum_\nu \tau_{l_m}(E, E_z, V, z) L_{S,\nu}(E, \mathbf{r}_\parallel) \approx \bar{\tau}_{l_m}(E, V, z) \sum_\nu L_S(E, \mathbf{r}_\parallel) = \bar{\tau}_{l_m} L_S(E, \mathbf{r}_\parallel). \tag{20}$$

The tunneling current I in (12) thus explicitly contains the term of interest $L_S(E, \mathbf{r}_\parallel)$:

$$I \approx \sum_{l_m} \int_0^V \rho_{T,l_m}(E - V)\bar{\tau}_{l_m}(E, V, z) L_S(E, \mathbf{r}_\parallel) \, \mathrm{d}E, \tag{21}$$

and consequently, (14) becomes

$$\frac{\mathrm{d}I}{\mathrm{d}V} = \sum_{l_m} \frac{\mathrm{d}I_{l_m}}{\mathrm{d}V}$$

$$= L_S(E = V, \mathbf{r}_\parallel) \sum_{l_m} \rho_{\mathrm{T},l_m}(0) \bar{\tau}_{l_m}(E = V, V, z) + \sum_{l_m} B_{s,l_m}, \tag{22}$$

where B_{s,l_m} in (16), under condition (20), becomes

$$B_{s,l_m} = \sum_\nu (\bar{\beta}_{l_m,\nu} + \bar{\gamma}_\nu) I_{l_m,\nu} \approx (\bar{\beta}_{l_m} + \bar{\gamma}) I_{l_m}. \tag{23}$$

Equation (22) now directly contains information on the sample LDOS L_S at the energy $E = V$. L_S is multiplied by the generalized transmission coefficient $\bar{\tau}_{l_m}$, evaluated at $E = V$ and superposed on the background term B_{s,l_m}.

The assumption expressed in (20) is thus pointing to the possibility of a simplified analytical treatment and it is worthwhile to understand the conditions under which it holds. Looking at (6)–(8) and (13), we see that (20) is rigorously correct only if all the states described in the system have the same constant out-of-plane energy E_z, which can occur in several cases of physical interest (see section 3). In the general case, the energy $E_z = E - E_\parallel$ is different for different sample states. Consequently, the contribution of each sample state to the tunneling current cannot be simplified by considering only the quantity $L_S(E)$, and (12) has to be evaluated considering the full band structure of the investigated system.

If the sample electronic structure is in some way known (e.g. by means of *ab initio* calculations or analytical modeling), it is possible to use (9) or (10) to simulate an STS measurement depending on the assumed symmetry of the tip apex orbitals. This approach will be presented in section 3. On the other hand, if one is interested in recovering the surface LDOS from experimental STS data, (21) and (22) can be used for this purpose, as will be shown in section 4.

3. Tunneling spectroscopy of Shockley states

To provide an example of physical interest, we consider the case of noble metal surfaces such as Cu(111) and Au(111), which are characterized by a surface state confined in a few monolayers from the surface termination [24]. The presence of a projected gap at $\bar{\Gamma}$ produces a weak interaction between the surface and the bulk states [25] and, in these conditions, the surface states are well described by an ideal 2D free-electron gas regarding the in-plane motion. The possibility of treating the electronic properties of this system using an analytical formulation makes it a good candidate for applying the theoretical approach presented in section 2 and simulating the results of an STS measurement. For this system, a parabolic relation between energy and momentum is expected; neglecting all the deviations that affect the real system, we express such a relation in the form

$$E = E_S + \frac{|k_\parallel|^2}{2m^*}, \tag{24}$$

where E_S is the bottom energy of the parabolic band (for Au and Cu, this value is estimated to be 460 and 430 meV below E_F [26, 27]), respectively, $k_\parallel$ is the momentum related to the in-plane motion and m^* is the electron effective mass of the state (0.23 and 0.39 m_e for
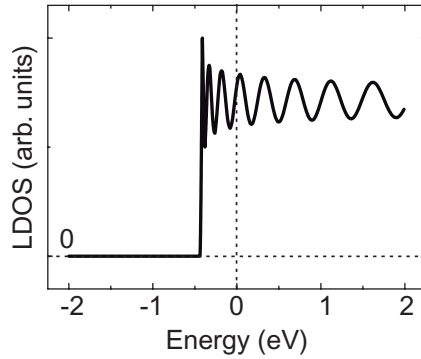
Figure 2. Simulated LDOS of a Shockley surface state with $E_S = 0.42\,\text{eV}$ as a function of E ($z = 0$, distance from the ideal step $x = 46\,\text{Å}$).

Au and Cu, respectively [26, 27]). It is worth noting that the out-of-plane component of the energy E_z does not depend on the momentum in this case, so that $E_z = E_S$ for every $k_\parallel$. For this system, the use of (20)–(21) is then fully justified. The DOS is obtained by integrating over the momentum space. For a full 2D translational symmetry, a step function with the discontinuity placed at $E = E_S$ is obtained. In STS, this step is usually visible as a jump in dI/dV signal for $V = E_S$ [25–27]. In the presence of a scattering center like point defects or steps, the translational symmetry is partially broken (scattering sites interact with the electrons of the surface and generate the well-known phenomenon of standing waves [28]) and the proper physical quantity to be considered is the LDOS instead of the DOS. The LDOS is modulated both in space and in energy, as can be observed routinely in low-temperature STM experiments.

In order to include such effects, we describe a model system of a 2D Shockley state, with an ideal step barrier along the x-direction. In this configuration, the eigenfunctions of the system are [29]

$$\psi_\nu = g_{k_x}(x) h_{k_y}(y) \chi(z), \tag{25}$$

where

$$g_{k_x}(x) = L^{-1/2} \sin(k_x x); \quad h_{k_y}(y) = L^{-1/2} e^{ik_y y}; \quad \chi(z) = C e^{-k_S z},$$

and $k_x^2 + k_y^2 = k_\parallel^2$. The corresponding LDOS is

$$L_S(E, x) = \frac{m^*}{4\pi} |C|^2 e^{-2k_S(E_S)z} \left[1 - J_0(2\sqrt{E - E_S}x) \right]. \tag{26}$$

The oscillations of the LDOS in energy are then described by the Bessel function J_0 (see figure 2).

We apply the tunneling theory described in section 2 to this system to perform a full analytical simulation of the tunneling current. Simulated I–V characteristics at typical tunneling parameters are shown in figure 3(a), considering the cases of s, p_z and d_{z^2} tip states. As the common behavior of all these curves, the absolute value of the tunneling current is larger in the negative bias region, whereas in the positive region it approaches a zero slope limit. Due to the integral nature of the current, we can observe only a very weak contribution coming from the oscillations of the LDOS. In figure 3(b), dI/dV curves are shown; we immediately recognize the two basic features of the model sample LDOS, namely the edge at the bias $V = E_S$ and
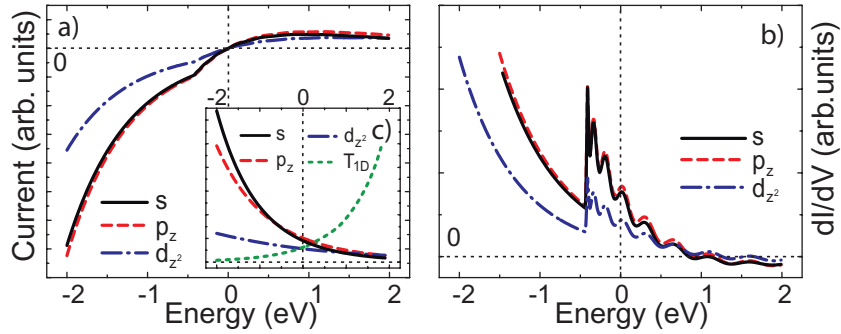
Figure 3. Simulated (a) I_{l_m} and (b) dI/dV_{l_m} for the system shown in figure 2 as a function of V ($l_m = s,\ p_z,\ d_{z^2};\ z = 7.36$ Å). (c) The transmission coefficient $\bar{\tau}_{l_m}(E, V)|_{E=V}$ as a function of V. The WKB transmission coefficient $T(E, V)|_{E=V}$ is shown for reference.

the oscillations in energy. The simulated dI/dVs agree well with typical experimental data on Shockley surfaces, such as that in figure 6 acquired on Cu(111), in particular regarding the monotonically decreasing behavior and the sharp onset. The maxima of these oscillations in the dI/dV curves are close to the L_S ones but a shift towards lower energy is present. In this sense, our simulation supports the well-known problems in extracting physical quantities of the sample directly from the dI/dV data, without further treatment. More precisely, we want to point out that the shape of the dI/dV is both *qualitatively* and *quantitatively* different from that of the sample LDOS. From a comparison of figures 2 and 3(b), we can see that (i) the differential conductivity shows a decreasing background distortion and it can even become negative after a certain value of the bias (figure 3(b)), and (ii) for negative biases, dI/dV continuously increases even where the LDOS evaluated for $E = V$ is zero. Furthermore, the behavior of dI/dV is affected in a non-obvious way by the symmetry of the tip apex orbitals.

In order to better understand the origin of these spurious contributions, we now interpret the shape of the dI/dV curves in the light of (22). Let us consider the decreasing behavior of the dI/dV signal. This effect can be usually ascribed to the energy behavior of the tunneling barrier and has been exhaustively explained in the framework of 1D-WKB theory [8, 9]. However, in the case of Shockley states, this effect is not expected simulating the tunneling using a 1D model [7], so that for this specific case its origin has to be ascribed to the 3D features included in the transmission coefficient. In order to elucidate this point, the quantity $\bar{\tau}_{l_m}(E, V, E_z, z)|_{E=V}$ is shown in figure 3(c) as a function of V for a constant E_z. We can see that in any case it shows completely different behavior compared with the 1D-WKB transmission coefficient [6]. For the Shockley states, the tunneling probability *decreases* when the applied bias *increases*. This is due to the fact that each state with a total energy E has a fixed out-of-plane energy E_z, and increasing the applied voltage only has the effect of increasing the barrier height, which determines the observed decrease in the transmission coefficient and the damping of the differential conductivity. On the other hand, for the general case of surface or surface-projected states, it is likely that a higher total energy E results in a higher out-of-plane energy E_z and a higher tunneling probability. Therefore, increasing the applied voltage not only increases the barrier height but also allows tunneling among states with a higher energy and tunneling probability; thus the resulting voltage behavior is the opposite.

Equation (22) is also helpful in understanding why we observe a non-vanishing $\mathrm{d}I/\mathrm{d}V$ below the edge of the Shockley state. If we consider the term B_{s,l_m} of this expression, we see that it is proportional to the tunneling current, which, being non-vanishing at negative biases, provides a finite contribution also in the energetic interval in which the surface has no states.

We now want to highlight the effect of the symmetry of the tip apex orbitals. From figure 3(b), we see that the s, p_z and d_{z^2} tip symmetries generate three different $\mathrm{d}I/\mathrm{d}V$ curves with different amplitudes. In particular, the s and p_z cases are very similar, whereas for d_{z^2} the amplitude of $\mathrm{d}I/\mathrm{d}V$ is reduced. The differences in their amplitude can be once again understood by looking at the behavior of $\bar{\tau}_{l_m}$ in figure 3(c). The effect of the tip apex orbital symmetry is contained in $\bar{\tau}_{l_m}$, which depends on the tunneling matrix elements. Since the coefficient $\tau_{d_{z^2}}$ is always smaller than the other two coefficients, this results in decreased values of tunneling current and differential conductivity for the d_{z^2} states, as visible in figure 3(b). To summarize the previous discussion, we observed that the main features of the Shockley state can be directly understood by looking at the $\mathrm{d}I/\mathrm{d}V$ curve. However, various spurious contributions from the tunneling process arise and they have to be properly removed for a quantitative recovery of the sample LDOS.

4. Recovery of the sample local density of states

Let us now discuss the crucial topic of how to extract quantitative information on the sample LDOS from the tunneling current. If the tunneling is dominated by one tip state symmetry in the energy range of interest ($I_{l_m} \equiv I$), an explicit expression for the sample LDOS in terms of experimental and modeling parameters can be obtained combining (22) and (23) for systems satisfying the conditions (20):

$$L_{\mathrm{S}}(E = V) = \left[\frac{\mathrm{d}I}{\mathrm{d}V} - (\bar{\beta}_{l_m} + \bar{\gamma})I \right] \cdot \frac{1}{\rho_{\mathrm{T},l_m}(0)\bar{\tau}_{l_m}(E = V, V, z)}. \tag{27}$$

Equation (27) is expressed in terms of the tunneling current I, in a similar way to that suggested by Koslowski *et al* in the 1D case [7]. The parameters $\bar{\beta}_{l_m}$ and $\bar{\gamma}$ are functions of V. Equation (27) represents an explicit expression to obtain the sample LDOS as a function of the measured quantities $\mathrm{d}I/\mathrm{d}V$ and I and simple analytical functions. In the general case, we may expect that more than one tip state contributes to the tunneling, so that the terms ρ_{T,l_m}, $\bar{\tau}_{l_m}$ and $\bar{\beta}_{l_m}$ should consequently be substituted by proper mean quantities. However, the assumption of one symmetry tunneling represents a simple and usually the best solution from a practical point of view.

We are therefore led to evaluate the effectiveness of (27) on the physical system discussed in section 3. For this purpose, a comparison of (27) with the normalization methods proposed in the literature by Stroscio *et al* [5], Ukraintsev [6] and Koslowski *et al* [7] will be presented below, followed by an analysis of the sensitivity to the modeling parameters (i.e. tip–sample distance z and tip orbital symmetry l_m). Then, (27) will be also applied to experimental data acquired regarding Cu(111). In these investigations, both the presented method and those from the literature have been applied to the I and $\mathrm{d}I/\mathrm{d}V$ curves simulated in the case of the s tip. Except normalization to total conductivity I/V, all these methods require an estimate of the tunneling parameters such as tip–sample distance and effective work-function [9]. For the sake of simplicity, here these parameters have been directly taken from the simulations. As far as the
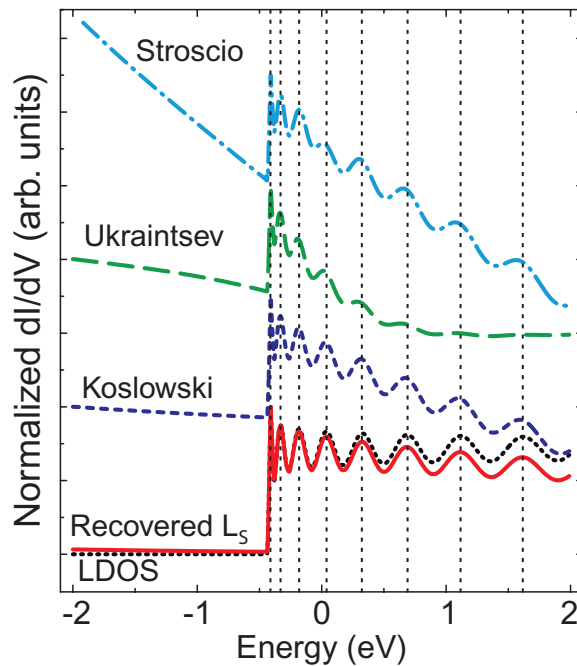
Figure 4. Comparison between the normalization methods obtained considering I_s and dI/dV_s represented in figure 3 and those obtained using the procedures of [5] (Stroscio *et al*), [6] (Ukraintsev), [7] (Koslowski *et al*) and (27) (recovered L_S). Sample LDOS is shown for reference. The maxima of the LDOS oscillations are indicated by vertical dashed lines.

Koslowski method is concerned, we directly use the 3D-WKB extended method (equation (13) of [7]), specifically suited for the Shockley state system.

The comparison is shown in figure 4. All the previous methods preserve an overall background distortion represented by a negative slope in all the recovered spectra. This distortion also affects the positions of the maxima of the oscillation in the spectra, producing a shift towards lower energy values. Moreover, some other shape distortions are visible in the spectra normalized using the total conductivity method and the Ukraintsev method. To explain such effects, we can observe that these procedures have been developed in a 1D framework, while this peculiar system presents important features that can be taken into account only in a 3D treatment. Even using Koslowski normalization, which is a 3D-WKB extension of the corresponding 1D-WKB method, it is not possible to completely remove the overall distortion, mainly because a proper inclusion of the tunneling matrix elements between the sample and the tip is missing. From figure 4, it is clear that the best result is achieved by applying (27).

The proposed method is also capable of properly removing the effects of different tip states. In figure 5(a), we show the results of applying (27) to the three dI/dV and I obtained by simulating the tunneling through the s, p_z and d_{z^2} tip states making use of the corresponding β, γ and τ functions. Even though, below the Shockley edge, the recovered curves show some small differences, all three normalized curves are coincident above this edge and in full agreement with the sample LDOS, thus showing the capability of the procedure to remove the influence of tip apex orbitals symmetry and achieve the correct result.
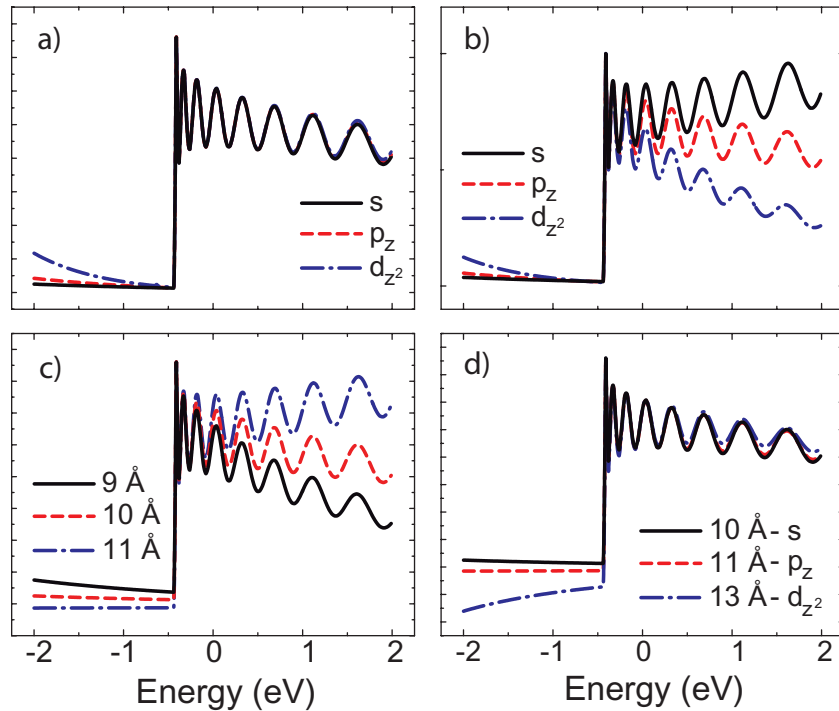
Figure 5. Recovered L_S obtained using (27) on: (a) simulated STS data ($l_m = s$, p_z, d_{z^2}, $z = 10\,\text{Å}$), assuming the correct tip state and tip–sample distance, (b) simulated STS data ($l_m = p_z$, $z = 10\,\text{Å}$), assuming $l_m = s$, p_z and d_{z^2} and the correct tip–sample distance, (c) simulated STS data ($l_m = s$, $z = 10\,\text{Å}$), assuming the correct tip state but $z = 9$, 10, 11 Å, (d) simulated STS data ($l_m = s$, $z = 10\,\text{Å}$), assuming different parameter sets ($l_m = s$ and $z = 10\,\text{Å}$; $l_m = p_z$ and $z = 11\,\text{Å}$; $l_m = d_{z^2}$ and $z = 13\,\text{Å}$).

It might appear that the applicability of this method relies on a detailed knowledge of the tip apex orbital symmetry and the tip–sample distance, which are often not achievable from a practical point of view. In order to evaluate possible errors introduced by wrong estimates of these experimental parameters, we consider the behavior of the proposed procedure when a 'wrong' estimated tip state is used, as shown in figure 5(b). Here, a recovered dI/dV curve related to a p_z state tip is shown, together with the application of (27) also assuming other tip symmetries. Some background distortion can be introduced if the tip apex orbital symmetry differs from that responsible for the tunneling process, but such an effect is anyhow clearly less pronounced when compared to that obtained using conventional 1D methods. In other words, it is better to normalize using (27) with a trial matrix element (possibly, even an incorrect one) instead of applying a method that does not consider the tip apex symmetry at all. We similarly consider the problem of a wrong estimation of the tip–sample distance. In figure 5(c), the dI/dV curve is simulated by considering an s state tip located 10 Å from the surface, and (27) was applied using three different estimated distances. Once again, the overall effect of using a wrong parameter is a background distortion, particularly visible in the change of slope of the recovered LDOS at voltages above the surface state onset. Since both the tip state and the tip–sample distance parameters act in a similar way on the background of the recovered spectra,
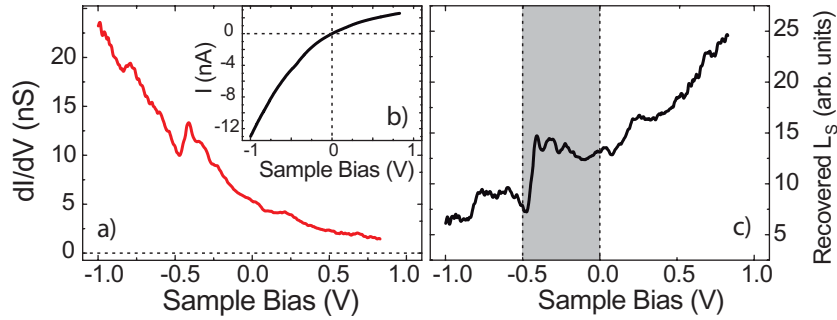
Figure 6. (a) Experimental dI/dV and (b) I curves taken using a Cr bulk tip on Cu(111). Set point: $I = 2\,\mathrm{nA}$, $V = 0.5\,\mathrm{V}$. Distance from the edge: 42 Å. (c) Recovered L_S obtained by using (27), assuming $l_m = s$ and $z = 10$ Å. Gray area corresponds to the reference region (see text for details).

it might be possible to obtain similar results by applying (27) using different combinations of estimated parameters. As an example of this, we show in figure 5(d) the results of applying (27) to the simulated STS curve with $l_m = s$ and $z = 10$ Å using both the correct parameters and two different combinations of estimated l_m and z, namely $l_m = p_z$, $z = 11$ Å and $l_m = d_{z^2}$, $z = 13$ Å. In all these three cases, the shape of the recovered LDOS is very close to that of the ideal surface LDOS (see figure 2 for a comparison), and only small errors are introduced because of using the improper set of parameters.

We now conclude this section by applying the procedure to experimental STS data acquired regarding the Cu(111) surface. Measurements were made at 7 K using an Omicron low-temperature STM on a conventionally prepared Cu(111) single crystal [27] and using a bulk Cr tip [30]. In figures 6(a) and (b) the dI/dV and I curves are presented. One can apply (27) in order to extract the measurement parameter l_m and z from the STS data. Since real measurements may contain several sources of spurious contributions not considered in this work (e.g. surface-projected bulk band states and non-constant tip DOS effects [9]), it is likely that complete agreement between recovered experimental and theoretical LDOSs could not be expected over the whole energy range of the experiments. Therefore, we focus our attention on the energy region close to the Shockley onset, which represents the most reproducible and reliable feature in STS measurements of this surface. We varied the set of parameters l_m and z until finding satisfactory agreement between theoretical and recovered LDOS in the range of -0.5–0 V. Figure 6(c) shows the result achieved using $l_m = s$ and $z = 10$ Å. In agreement with what was obtained on simulated spectra, a similar result can be also achieved using other sets of parameters, such as $l_m = p_z$, $z = 11$ Å and $l_m = d_{z^2}$, $z = 13$ Å. Although none of these values can be excluded *a priori*, we are here led to choose the s tip case as the most probable one, being the only one which corresponds to the typical z range of 5–10 Å according to the set point usually adopted in STS measurements for this class of surfaces [12, 31, 32]. Comparing figures 2 and 6(a), we see two effects that do not appear in our simulated spectra outside the reference range, namely a wide feature below the Shockley edge and an increasing background of the recovered LDOS curve as a function of the applied voltage above +0.1 V. These features are not due to our procedure since we found that they are stably detectable also employing other l_m values and varying z by ± 2 Å. Since no sample states, neither from the surface nor from the projected bulk band, are expected within the energy range of -0.8–0.45 eV [33],

we safely attributed the feature below the Shockley onset to a non-constant tip DOS effect [9]. The rising signal above Fermi energy suggests instead a different origin, probably connected with a change of the decaying behavior of the sample states in vacuum as a function of energy. It is worth recalling that Cu surface states are located within the surface projected bulk band gap below Fermi energy, whereas, with increasing momentum as a function of energy, they cross the bulk band edge just above the Fermi level [34], losing in this way its pure surface character and becoming instead a surface resonance [29]. As a consequence, the decaying behavior in vacuum is partially modified by the bulk contributions and the assumption that $E_z = E_S$ is not valid in this situation. This might therefore explain the different behavior of the recovered L_S as a function of energy with respect to pure Shockley states LDOS. It is remarkable that such a feature is particularly evident when STS data are treated by the proposed procedure, since (27) relies on the presence of a pure surface state and not on possible bulk band state or surface resonances, so that possible deviations from the ideal free-electron behavior are particularly emphasized.

The presented analysis suggests two possible ways of applying (27) to real situations. First, one can be interested in normalizing STS spectra from unknown surfaces using some reasonable trial parameters. In this case, we showed that it is possible to achieve a proper recovery of the LDOS by using different sets of parameters, which makes the procedure even more reliable. On the other hand, the proposed procedure could be applied to STS data from well-known surfaces (such as Shockley states) with the aim of extracting tip apex orbital symmetry and experimental parameters. This can be implemented by tentatively varying the parameters until the recovered LDOS (or alternatively a selected part of the recovered LDOS in a certain energy window) is very similar in shape to the ideal form of the surface LDOS. Since this result can be achieved through more than one combination of estimated parameters, a unique set could not be univocally found. However, it is also possible that some of the extracted parameter sets could be excluded if some values are not physically reasonable (e.g. too short/long tip–sample distances or unexpected tip apex symmetry for the used tip material), achieving in this way a unique result.

5. Conclusions

In this paper, we have presented a 3D-TH model of tunneling that is capable of including both the tip angular symmetry and the sample band structure. We found that the tunneling current can be represented as an energy convolution between the tip DOS and the sample LDOS weighted by an effective transmission coefficient τ. In contrast to what occurs in the 1D-WKB description, this coefficient depends on the 3D properties of the tip and the sample, in particular on the out-of-plane sample state energy (E_z) and the tip state angular symmetry (s, p_z or d_{z^2}). On the basis of these findings, we proposed a specific normalization procedure to properly recover the sample LDOS from an STS measurement. We considered the case of a Shockley state, which is analytically developed to also compare this normalization method with other relevant existing approaches. This analysis clearly shows the importance of including the 3D effects in the description of tunneling in order to better quantitatively recover the sample LDOS from experimental measurements.

The presented theory shows several advantages with respect to the other 1D-WKB-based models. First, the theory can be used to simulate the dI/dV curves starting from a theoretical sample electronic structure in order to reveal how an STS measurement should appear. This

IOP Institute of Physics **Φ** DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

would be of great help in overcoming the usual difficulties faced in comparing the *ab initio* simulations of surface electronic properties with STS data. With respect to WKB theories, the decaying behavior in vacuum of states belonging to different regions of the Brillouin zone can now be properly considered, together with effects introduced by the tip apex orbital symmetry. When the starting point is the experimental data, the proposed normalization procedure can be exploited to recover (i) the sample LDOS from STS measurements assuming a constant tip DOS or (ii) the tip apex orbital symmetry and electronic structure when STS is performed on well-known surfaces. In principle, the presented method for the recovery of tip and sample LDOS could also be exploited in connection with iterative methods, as proposed in the 1D-WKB framework [7, 10], e.g. by using a set of measurements taken at different tip–sample distances [11], providing also in this case a correct inclusion of the 3D effects.

Acknowledgments

References

 [1] Chen C J 2008 *Introduction to Scanning Tunneling Microscopy* 2nd edn (Oxford: Oxford University Press)
 [2] Datta S 1995 *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press)
 [3] Feuchtwang T E 1974 *Phys. Rev.* B **10** 4135
 [4] Selloni A, Carnevali P, Tosatti E and Chen C D 1985 *Phys. Rev.* B **31** 2602
 [5] Stroscio J A, Feenstra R M and Fein A P 1987 *Phys. Rev. Lett.* **58** 1668
 [6] Ukraintsev V A 1996 *Phys. Rev.* B **53** 11176
 [7] Koslowski B, Dietrich C, Tschetschetkin A and Ziemann P 2007 *Phys. Rev.* B **75** 035421
 [8] Wagner C, Franke R and Fritz T 2007 *Phys. Rev.* B **75** 235432
 [9] Passoni M, Donati F, Li Bassi A, Casari C S and Bottani C E 2009 *Phys. Rev.* B **79** 045404
[10] Wahl P, Diekhöner L, Schneider M A and Kern K 2008 *Rev. Sci. Instrum.* **79** 043104
[11] Koslowski B, Pfeiffer H and Ziemann P 2009 *Phys. Rev.* B **80** 165419
[12] Ziegler M, Néel M, Sperl A, Kröger J and Berndt R 2009 *Phys. Rev.* B **80** 125402
[13] Das B and Mahanty J 1987 *Phys. Rev.* B **36** 898
[14] Bardeen J 1961 *Phys. Rev. Lett.* **6** 57
[15] Feuchtwang T E and Cutler P H 1988 *Phys. Scr.* **38** 252
[16] Tersoff J and Hamann D R 1985 *Phys. Rev.* B **31** 805
[17] Chen C J 1990 *Phys. Rev. Lett.* **65** 448
[18] Chen C J 1990 *Phys. Rev.* B **42** 8841
[19] Chen C J 1992 *Phys. Rev. Lett.* **69** 1656
[20] Passoni M and Bottani C E 2007 *Phys. Rev.* B **76** 115404
[21] Appelbaum J A and Brinkman W F 1969 *Phys. Rev.* **186** 464
[22] Duke C B 1969 *Tunneling in Solids* (New York: Academic)
[23] Harrison W A 1961 *Phys. Rev.* **123** 85
[24] Lüth H 2001 *Solid Surfaces, Interfaces and Thin Films* 4th edn (Berlin: Springer)
[25] Kliewer J, Berndt R, Chulkov E V, Silkin V M, Echenique P M and Crampin S 2000 *Science* **288** 1399
[26] Schouteden K, Lievens P and Van Haesendonck C 2009 *Phys. Rev.* B **79** 195409

IOP Institute of Physics ⬧ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

[27] Rodary G, Sander D, Liu H, Zhao H, Niebergall L, Stepanyuk V S, Bruno P and Kirschner J 2007 *Phys. Rev. B* **75** 233412
[28] Crommie M F, Lutz C P and Eigler D M 1993 *Science* **262** 218
[29] Desjonqueres M C and Spanjaard D 1998 *Concepts in Surface Physics* 2nd edn (Berlin: Springer)
[30] Li Bassi A *et al* 2007 *Appl. Phys. Lett.* **91** 173120
[31] Li J, Schneider W-D and Berndt R 1997 *Phys. Rev. B* **56** 7656
[32] Kröger J, Limot L, Jensen H, Berndt R and Johansson P 2004 *Phys. Rev. B* **70** 033401
[33] Kevan S D and Gaylord R H 1987 *Phys. Rev. B* **36** 5809
[34] Hulbert S L, Johnson P D, Stoffel N G, Royer W A and Smith N V 1985 *Phys. Rev. B* **31** 6815