Theory of transition from the tunneling regime to point contact in scanning tunneling microscopy

S. Ciraci and E. Tekman

Department of Physics, Bilkent University, Bilkent, 06533 Ankara, Turkey

(Received 21 June 1989; revised manuscript received 21 September 1989)

We analyze the transition from the tunneling regime to point contact in scanning tunneling microscopy. The variation of conductance as a function of tip-sample separation is sample and tip specific. Tunneling occurs through an effective barrier even if the potential barrier collapses. Subsequent to the collapse of the effective barrier the point contact is initiated leading to ballistic transport. The ballistic conductance through uniformly increasing contact area exhibits neither sharp quantized steps nor pronounced quantum oscillations. The observed oscillations are explained by the irregular enlargement of the contact area.

The current theory of scanning tunneling microscopy[1] (STM) is based on the first-order time-dependent perturbation theory and predicts that the tunneling conductance is proportional to the local density of states of the bare sample surface $\rho(r_0, E_F)$, evaluated at the center of the tip and at the Fermi level.[2] This theory has been used with reasonable success in STM operating in the usual conditions, where the tip and sample are widely separated. This is the independent-electrode regime. Earlier, the authors[3] identified two other regimes occurring at small bias voltage V_b, in which the tip-sample interaction is significant, and thus the electrodes are no longer independent. These were *electronic-contact* and *point-contact* regimes. In the first one, the transport takes place via tunneling through the strongly disturbed electronic states of the electrodes. At relatively smaller tip-sample distance d (from ion-core to ion-core), the potential barrier ϕ (i.e., the saddle-point value of the electronic potential energy above E_F) collapses and eventually point contact is initiated leading to the ballistic transport. In this paper we present a theoretical analysis of these two regimes, which is crucial for the interpretation of STM operating at small tip-sample distance. Major issues we are investigating are how the transition from tunneling to ballistic conduction takes place and whether the conductance is quantized in the ballistic regime.

As the tip approaches the sample, the potential barrier is lowered and the tip-sample interaction gradually increases.[4] The electronic charge is rearranged and the ions are displaced to attain the lowest total energy at the preset tip position. Owing to the significant overlap at small tip-sample distance in the electronic contact regime, the tip and sample states are combined to yield site-specific localized (or resonance) states[3] with a net binding interaction.[5] Moreover, due to the tip-sample interaction the parallel wave vector $k_\parallel$ which is normally conserved in the independent electrode regime, is no longer a good quantum number.[6] For example, in contrast to the case of STM of graphite in the independent electrode regime, the apparent barrier should not rise because of transversal momentum. In the presence of a significant tip-sample interaction, the effect of the local perturbation of electronic states can be sought in the expression of tunneling current[7]

$$I \propto \int_{E_F}^{E_F + eV} D_t(E) D_s(E - eV) T(E, V_b) dE . \qquad (1)$$

Apparently, the density of states of the tip and sample (D_t and D_s, respectively), as well as the transmission function $T(E, V_b)$ are modified by the local tip-sample interaction and thus deviate from those corresponding to the bare tip and bare sample.

The tip-sample interaction at small d has been investigated extensively by the recent studies of Ciraci, Baratoff, and Batra.[5] Their results obtained from *ab initio* force and electronic-structure calculations provide evidence that the corrugation amplitude at constant tunnel current is slightly reduced by the tip-induced elastic deformation. The tip-induced local modifications of the electronic structure and effective barrier have a much stronger effect on the STM images.[5] Since the wave function is laterally confined[5,6,8] between the tip and sample, the energy ε_1 of the lowest propagating state may occur above E_F even if ϕ is collapsed. This forms an effective barrier $\phi_{\text{eff}} = \varepsilon_1 - E_F$. Such a situation is illustrated in Fig. 1 by the contour plots of the potential energy $V(r)$ and charge density $\rho(r)$ of the multiatom Al tip and Al sample calculated by the self-consistent (SCF) pseudopotential method.[5] Already at $d = 4.2$ Å, ϕ is collapsed and $V(r) - E_F \simeq -1.9$ eV, whereas the effective barrier ϕ_{eff} is significant and estimated to be ~ 2.3 eV. On the other hand, the charge density is enhanced in the dividing plane [in the region where $V(x, y, z = d/2) < E_F$]. Therefore, in spite of significant tip-induced modifications of electronic structure in the electronic contact regime, tunneling still occurs across an effective barrier ϕ_{eff}, but the conductance is no longer proportional to $\rho(r_0, E_F)$ of the bare sample.

The perforation of the potential barrier (i.e., $\phi \rightarrow 0$) sets in even before the outermost tip atom enters in the strongest attraction of the sample.[5] However, the point contact is initiated only if the energy of the lowest state quantized in the orifice (i.e., first-subband energy) ε_1, dips below E_F and becomes occupied (i.e., ϕ_{eff} collapses). This way a new channel of conduction is opened, and consequently the character of the conductance undergoes a qualitative change. The gradual collapse of the barrier
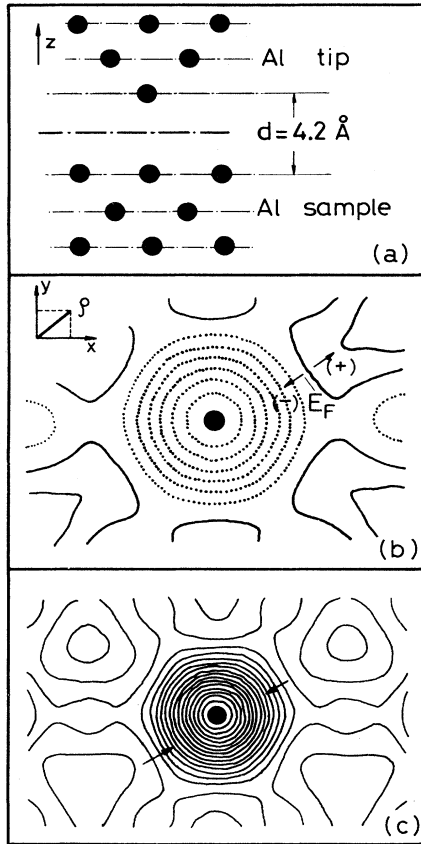
© 1989 The American Physical Society

FIG. 1. Contour plots of the Al tip and the Al sample with $d \simeq 4.2$ Å. (a) Atomic arrangement. (b) SCF potential $V(\mathbf{r})$ in a horizontal plane bisecting d indicated by dashed-dotted line in (a). Dotted contours indicate $V(\mathbf{r}) < E_F$, and contour spacing is 20 mRy. (c) Total charge density in the same plane. Contour spacings are $2 \times 10^{-4} \, e/(\text{a.u.})^3$.

and its effect on the current have been illustrated experimentally by Gimzewski and Möller[9] in studies of the transition from tunneling to point contact between the Ir tip and Ag sample. Their $\log I(d)$ plot at constant bias voltage shows that the current I increases first exponentially with decreasing d. This implies a tunneling behavior. The discontinuity observed at small d was attributed to the mechanical instability (or hysteric deformation).[10] The recorded values of the conductance just after the discontinuity was only $\sim 80\%$ of the unit of quantum conductance,[11] $(2e^2/h)$. Upon further approach of the tip, I continues to increase and exhibits an oscillatory behavior.[9]

Lang[12] simulated the point contact realized in the above experiment by two jellium electrodes, one of them having an Na atom attached to the jellium edge and thus representing a single atom tip. He found that the conductance G_t saturates at the value $\eta 2e^2/h$ and forms a plateau when d is in the range of the distance from the Na core to the positive background edge of the tip electrode. The value of η is only 0.4 for Na, and is found to depend on the identity of the tip. Within a tight-binding approximation and using the nonequilibrium Green's-function

method, Ferrer, Martin-Rodero, and Flores[13] also found that G_t reaches $\lesssim 2e^2/h$ at the smallest d. Apparently, it is not clear so far whether the transition from tunneling to ballistic regime takes place at small d, and how a plateau can occur prior to the plastic deformation. Moreover, in both calculations[12,13] the mechanism related to the experimentally observed increase of G_t following a discontinuity is not included. Gimzewski and Möller[9] give an estimate for the dimension of the contact area, which lies in the range of λ_F. If so, the observed transport beyond the discontinuity has to be associated with the quantum ballistic transport.[14,15]

In what follows we will show that after the point of discontinuity the conductance G normally increases as the tip continues to be pushed towards the sample, and its observed behavior[9] is compatible with the ballistic transport. However, the initiation of the point contact and the variation of the conductance depend on the material of electrodes and the shape of the tip, as well. Since the length of the orifice between the tip and surface is small ($< \lambda_F$), the quantization with sharp steps in G does not take place. Therefore, the pronounced large-period oscillations superposed on the $\log I(d)$ curve[9] are possibly related to the irregular evolution of the contact.

We model the tip-sample system by the two jellium electrodes separated by a potential barrier depending on d. Only at the point of the contact is an orifice formed. Since the SCF calculations of a sharp tip facing the metal sample surface predict a parabolic potential in the gap,[5] the potential of the orifice for a given d is represented by

$$V(d,\rho,z) = \phi_m(d,z)$$
$$+ \alpha(d)\rho^2 \theta(z + d_0/2)\theta(z - d + d_0/2) , \qquad (2)$$

with $\phi_m(d,z)$ calculated from the jellium model.[16] The maximum of $\phi_m(d,z) - E_F$ is $\zeta(d)$, and $\zeta = \phi$ if $\zeta > 0$. d_0 is twice the distance between the first atomic plane and the jellium edge. Subsequent to the collapse of the potential barrier the diameter of the orifice at E_F, $\rho_m = \sqrt{-\zeta/\alpha}$, increases as d decreases. Note that the material- and tip-specific features of the contract are represented by $\alpha(d)$. For the particular contact system treated here $\alpha(d)$ is obtained by using the diameter of the point contact given by the experiment[9] and by scaling those calculated[5] for Al as a function of d, and also by using electronic parameters of Ag for jellium electrodes. The calculated values of α and ρ_m are given in Table I.

In order to calculate the conductance we consider the current carrying states. These are the three-dimensional (3D) plane waves in the electrodes and quantized transversal momentum states in the orifice. The wave functions of these orifice states consist of products of 2D isotropic harmonic-oscillator solutions and 1D solutions

TABLE I. α and ρ_m as functions $d - d_0$. Values of ρ_m are given only for $\zeta(d) < 0$.

$d - d_0$ (Å)	0.0	1.0	2.0	3.0	4.0	5.0	6.0
α (eV/Å^2)	2.11	0.94	0.48	0.27	0.16	0.10	0.07
ρ_m (Å)	1.62	1.55	1.02	$\cdots$	$\cdots$	$\cdots$	$\cdots$

corresponding to the potential $\phi_m(d,z)$. The coefficients of these states are obtained from the boundary conditions at the edges of the orifice ($z = -d_0/2$ and $d - d_0/2$). Finally, we use the conductance expression derived earlier[17] within the linear-response theory:

$$G = \frac{e^2}{\pi h} \int_{FS} \frac{d\mathbf{k}_\parallel}{k_Z(\mathbf{k}_\parallel)} \{\tilde{\Theta}^\dagger(\mathbf{k}_\parallel)\tilde{\Gamma}_R\tilde{\Theta}(\mathbf{k}_\parallel) - \tilde{\Delta}^\dagger(\mathbf{k}_\parallel)\tilde{\Gamma}_R\tilde{\Delta}(\mathbf{k}_\parallel)$$
$$+ 2\,\mathrm{Im}[\tilde{\Theta}^\dagger(\mathbf{k}_\parallel)\tilde{\Gamma}_I\tilde{\Delta}(\mathbf{k}_\parallel)]\} , \qquad (3)$$

where FS denotes the Fermi surface and $\tilde{\Theta}$ and $\tilde{\Delta}$ are the coefficients of the right- and left-going orifice states, respectively, for an incident plane wave with wave vector $\mathbf{k} = \mathbf{k}_\parallel + k_z\hat{z}$. Since the tunneling conductance of the evanescent waves are taken into account by the third term, the above expression of G has a wide range of application and thus is appropriate for our study covering *ballistic* as well as *tunneling* regime. The ballistic transport of electrons through a quasi-1D constriction was studied by using a similar approach,[17] and for long constrictions ($d \geq 5\lambda_F$) the conductance was found to be quantized in units of $2e^2/h$ with resonance structure superimposed on the plateaus. However, the tapering at the entrance, impurity, and roughness in the constriction prevent the conductance from exact quantization but yields plateaus at relatively lower values. In the present case, the orifice in the potential barrier is a tubelike constriction through which the electrons pass from one electrode to the other. Since the diameter of the constriction is in the range of λ_F, the electrons have a quantum regime different from that in the 3D electrodes. Their mean free path is larger than the length of the orifice, so the energy dissipation in the constriction is neglected. We believe that our model is realistic and comprises the essential ingredients of the point contact. It can be extended to treat a hornlike entrance to the orifice and the nonuniformities in $V(d,\rho,z)$ by using a z dependent α. In this case a transfer-matrix method is used for multiple boundary matching.[17]

Figure 2(a) presents the results of our calculations. In agreement with previous calculations,[12,13] the conductance associated with a uniform orifice set up by a single atom at the vertex of the tip has a value less than $2e^2/h$. Since the length of the orifice is finite and in the range of the internuclear distance a_0 (i.e., sum of the atomic radii), this result implies that the energy of the first subband ε_1 is still above E_F, (i.e., $\phi_{\mathrm{eff}} > 0$), and hence the conductance is dominated by tunneling. Apart from the contact system given in Table I we also analyzed the conductance by using $\alpha(d)$ corresponding to different tips. We found that depending on the shape of the tip the form of $G(d)$ may differ from Fig. 2(a) even *qualitatively*. For example, if ρ_m is allowed to be less than the atomic radius of Ag, G decreases passing through a maximum ($\sim 2e^2/h$). This form is reminiscent of $G(d)$ obtained by Ferrer *et al.*[13] On the other hand, $G(d)$ may reach a plateau[12] before the point of discontinuity if α is very small. This suggests that ϕ_{eff} collapses prior to the hysteric deformation, and the value of G at the plateau may be smaller than $2e^2/h$ owing to the scattering by the ions in the orifice. Certain contacts may have several subbands close to E_F, each contrib-
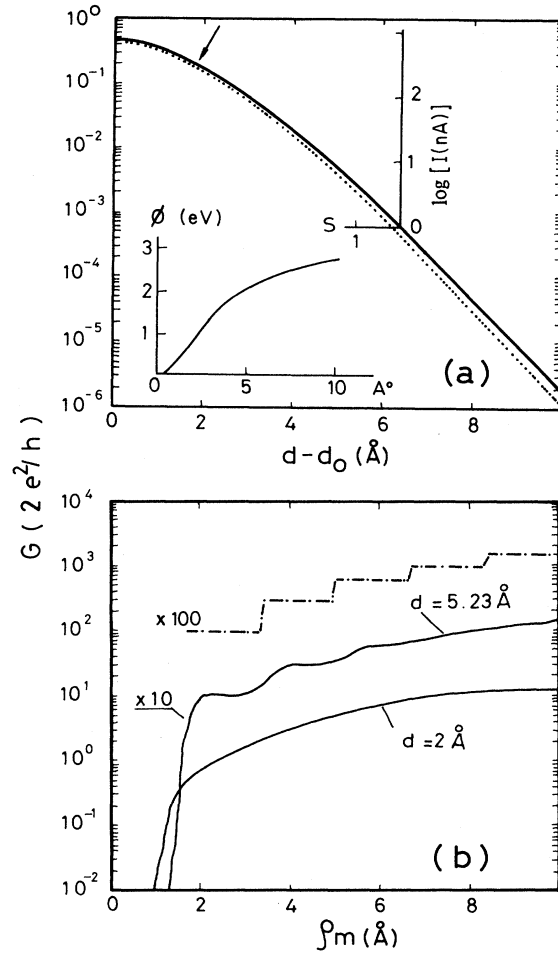

FIG. 2. (a) G_t vs d curve (or log$I$ vs tip displacement for bias $V_b = 20$ mV) calculated for two Ag electrodes. The dotted curve is the contribution of the first subband. The distance where the first mechanical instability was observed in the experiment is marked by the arrow. Inset shows the apparent barrier height calculated from the current. (b) G vs ρ_m of the Ag electrodes calculated for $d = 2$ Å ($\lesssim a_0$), $d = \lambda_F$, and for the exact quantization (dashed-dotted line).

uting to tunneling. In this case, plateaus do not occur but $G(d)$ increases almost linearly. For a flat tip (with several atoms at the apex) the contact area is large and hence G rises above $2e^2/h$ already before the tip-induced plastic deformation sets in. Of course, all these arguments are based on the assumption that there is neither an oxide layer nor a flake between the tip and sample, which influence the variation of $G(d)$ in an essential manner.

According to the experiment[9] it appears that the ballistic regime starts subsequent to the structural instability occurring at $d > a_0$. Under normal circumstances, if one continues to push the tip further, d saturates at $\simeq a_0$, but the contact area expands with an enhanced plastic deformation followed by the adhesion of several nearby atoms. The actual form and size of contact after the point of mechanical instability is uncertain and depends on several

parameters (such as the detailed atomic structure at the apex of the tip, intraelectrode and interelectrode interaction energies, etc.). For example, the tip atoms in the repulsive force range are displaced in such a way that the aperture of the orifice may be reduced incidentally causing G to decrease temporarily in the course of tip approach. Nevertheless, the value of conductance with respect to the displacement of the tip can be related to the radius of the orifice ρ_m, which normally increases with the continuing plastic deformation. Depending on the aperture of the orifice several subband states which are quantized in the orifice can be occupied. Each subband occupied contributes to G by $2e^2n/h$ (n is the degeneracy) and thus yields a step structure. The perfect quantization with sharp steps occurs if $d \gg \lambda_F$ and $\phi_m(d,z) = 0$, however. We extend our model to include the plastic deformation region, where a significant hysteresis is observed in the excursion of the tip.[9] Keeping $\phi_m(d,z) = 0$, we calculate the conductance as a function of $\rho_m = (E_F/a)^{1/2}$ for the fixed length of the orifice. Clearly this will be the continuation of the $G(d)$ shown in Fig. 2(a) after the point of discontinuity.

The variation of G with the radius of the contact in the plastic deformation region is presented in Fig. 2(b). Since $\phi_m(d,z) = 0$, the first channel is already opened at $\rho_m \simeq 1.5$ Å and the value of G rises by $\sim 2e^2/h$. Additional channels are opened sequentially as ρ_m increases. Pronounced quantum oscillations (or smeared-out step structure) of the $G(\rho_m)$ curve are apparent for $d = \lambda_F$, and are compared by the sharp steps corresponding to the exact quantization. However, for $d \simeq a_0$, weak oscillations are washed out in the logarithmic scale and hence $G(\rho_m)$ becomes almost featureless. It appears that the point contact through a short constriction prevents G from exact quantization or from having pronounced quantum oscillations, and the ballistic transport takes place in the "quantum" Sharvin regime.[14,17] Earlier, assuming an infinite potential well in the contact, the observed oscillations of the $\log I(d)$ curve were attributed to the quantization of the conductance.[15] The present results suggest that the pronounced quantum oscillations would appear in the $\log I(d)$ curve only if the length of the orifice were large ($\geq \lambda_F$). This situation would appear for the tip having a very unusual shape. These oscillations possibly originate from the irregular motion of the atoms as the tip is uniformly pushed towards the sample causing irregular enlargement of the contact area. Also the atoms of a blunt tip may undergo sequential contacts each time opening a new orifice and leading to abrupt changes in the current. Both cases give rise to the variation of $\log I(d)$ as observed experimentally.[9]

In conclusion, the character of transport and the variation of conductance as a function of d are not generic, but strongly depend on the properties of the tip and the sample. In particular, the appearance of a plateau before the point of discontinuity occurs only at certain conditions.

[1] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibell, Phys. Rev. Lett. **49**, 57 (1982).

[2] J. Tersoff and D. R. Hamann, Phys. Rev. Lett. **50**, 1998 (1983).

[3] E. Tekman and S. Ciraci, Phys. Scr. **38**, 486 (1988); Phys. Rev. B **40**, 10286 (1989).

[4] S. Ciraci and I. P. Batra, Phys. Rev. B **36**, 6194 (1987).

[5] S. Ciraci, A. Baratoff, and I. P. Batra, Phys. Rev. B (to be published); and (unpublished).

[6] S. Ciraci, in *Basic Concepts and Applications of Scanning Tunneling Microscopy and Related Techniques,* edited by H. Rohrer, N. Garcia, and J. Behm (Kluwer Academic, Amsterdam, in press).

[7] A. Selloni, P. Carnevalli, E. Tosatti, and C. D. Chen, Phys. Rev. B **31**, 2602 (1984).

[8] N. D. Lang, Phys. Rev. B **38**, 10395 (1988).

[9] J. K. Gimzewski and R. Möller, Phys. Rev. B **36**, 1284 (1987).

[10] J. B. Pethica and W. C. Oliver, Phys. Scr. **19A**, 61 (1987).

[11] R. Landauer, IBM J. Res. Dev. **1**, 233 (1957); Z. Phys. B **68**, 27 (1988).

[12] N. D. Lang, Phys. Rev. B **36**, 8173 (1987).

[13] J. Ferrer, A. Martin-Rodero, and F. Flores, Phys. Rev. B **38**, 10113 (1988).

[14] Yu. V. Sharvin, Zh. Eksp. Teor. Fiz. **48**, 984 (1965) [Sov. Phys. JETP **21**, 655 (1965)].

[15] N. Garcia (unpublished).

[16] J. Ferrante and J. R. Smith, Surf. Sci. **38**, 77 (1973).

[17] E. Tekman and S. Ciraci, Phys. Rev. B **39**, 8772 (1989), and references therein; **40**, 8559 (1989).