

A matrix method for treating the coupling between an electron and a surface plasmon: a dynamical image potential in model tunnelling junctions

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1998 J. Phys.: Condens. Matter 10 3697

(<http://iopscience.iop.org/0953-8984/10/17/004>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.209

The article was downloaded on 14/05/2010 at 13:02

Please note that [terms and conditions apply](#).

A matrix method for treating the coupling between an electron and a surface plasmon: a dynamical image potential in model tunnelling junctions

H Ness and A J Fisher

Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Received 13 November 1997, in final form 22 January 1998

Abstract. In this paper, we present a method for treating in a general manner the coupling between a single electron and boson fields represented by harmonic oscillators. The (general) solution of the corresponding many-body Schrödinger equation in real space is obtained by means of a propagation matrix method without using any *ansatz* for the many-body wave functions. We study the particular case of coupling between electrons and surface plasmons in tunnelling junctions. The electron is coupled inside the tunnelling barrier to a few surface plasmon modes. We present results for the dynamical effective potential felt by an electron tunnelling in model one-dimensional tunnelling junctions as well as for the various traversal times. As expected, significant differences from the corresponding static image potential are obtained when the tunnelling times τ are shorter than the characteristic response time of the surface charge (i.e. the inverse of the surface plasmon frequency ω). Examples of such dynamical effective potentials are given for various typical tunnelling conditions in the presence and absence of an applied bias voltage. The behaviour of the potential is studied versus the tunnelling times τ . It is also shown that the apparent barrier height that can be deduced from experiments does not always contain useful information about the dynamics of the coupled electron–plasmon system. On the other hand, the absolute values of the current in the tunnelling junctions are strongly dependent on the characteristic parameters of this dynamics.

1. Introduction

For a given geometry and tip–sample separation, the actual conductance of a scanning tunnelling microscope (STM) junction (or in general any tunnelling junction) depends crucially on the potential felt by the electrons. A substantial contribution to this potential depends on the extent to which the tunnel junction is polarized by the electron’s field. As is well known, if tunnelling is sufficiently slow in comparison to the characteristic response time of the electrodes (i.e. the inverse of the corresponding surface plasmon frequency), the electron will feel the classical ‘image force’ potential. If the tunnelling is very rapid, however, the redistribution of the charge around the junction cannot occur sufficiently rapidly to build up the image charge. These dynamical effects can alter drastically the shape of the effective potential felt by the tunnelling electrons and therefore also the values of the tunnelling current.

To our knowledge, the first experimental evidence of the image potential effects on the conductance in a STM junction were pointed out by Binnig *et al* [1]. The authors used a static model for the image potential to interpret their data. This model cannot determine

in which regime the dynamical effects on the potential are important. Furthermore, it is not obvious that the observed modifications of the STM conductance are mainly due to the effects of the image potential. They may also arise from the strong interaction between the electronic states of the tip and the sample when their separations are small.

More recently the interaction between tunnelling electrons and localized plasmon modes has been investigated by Berndt *et al* [2, 3]. These plasmon modes are induced by the proximity of the tip to the surface and are spatially confined. Their energy distribution can be observed via photon emission from the STM junction. In another class of experiments, evidence of dynamical effects in the image potential has been found in the conductance of semiconductor heterojunctions, especially in the case of tunnelling through thin and low tunnelling barriers [4].

On the other hand, the dynamical aspects of the image potential have also been studied theoretically by several authors within different approaches and approximations. Different classical treatments have been proposed; for example, Ray and Mahan determined self-consistently the image potential with the finite velocity of the electron treated as a classical particle [5]. However, for a more realistic description, quantum mechanical treatments of the dynamic image potential are necessary. Different types of approach have been used. They can essentially be classified into three groups: perturbation theory for the electron-plasmon coupling terms [6], the self-energy formalism and the path integral formalism. Within the self-energy formalism, Inkson derived analytical expressions of the Green's functions and the screened Coulomb potential for non-tunnelling electrons [7]. Green's functions and the resulting energy-dependent image potential have been determined by non-self-consistent calculations in the local and non-local limit for the Green's function [8–10], by self-consistent calculations in the local limit [11, 10] and by fully self-consistent calculations [12, 13]. To our knowledge, the first important contribution to a path integral formulation of the dynamic image potentials was given by Young [14]. In their spirit, all of the path integral methods are similar to the approach adopted by Caldeira and Leggett for calculating the tunnelling probability of a particle coupled to a heat bath [15, 16]. The same formalism has been used by Persson and Baratoff to study inelastic tunnelling and the corresponding dynamic image potential [17, 18]. The formalism has been corrected by Sebastian and Doyen to account for the correct boundary conditions and the fact that it is not possible to transfer arbitrary amounts of energy to the plasmons [19, 20]. Finally, Klipa and Šunjić recently generalized the previous path integral formalisms to go beyond the high-plasmon-frequency limit [21].

Despite the existence of this literature on the dynamical image potential, there is still a need to develop a general method to treat the coupling of electrons with harmonic modes (for instance the surface plasmon modes). The aim of this paper is to point out that this class of problems can be effectively solved by the method of Bonča and Trugman [22] whenever the interaction with the harmonic modes is localized in some region. The method should not rely on the analytical knowledge of the essential quantities like the wave functions or the Green's functions. This means that it should work, in principle, for any shape of the nominal static potential barrier and for situations not limited just to the tunnelling regime. We do not want to introduce any *ansatz* for the form of the many-body wave functions that will assume any particular dynamical correlation between the electron position and displacements of the harmonic modes. It should be possible to include virtual as well as real excitations of the harmonic modes. And finally, the model should enable us to separate the different contributions (elastic and inelastic processes) to the current in the region of space considered, which cannot be done with the existing path integral models.

For this purpose, we develop a matrix method to solve the corresponding many-body

Schrödinger equation on a real-space mesh where a single electron interacts with harmonic modes ‘located’ in the region of space of interest. At first sight, our model has some resemblance to the models used by Puri and Schaich [23] or Tagliacozzo and Tosatti [24]. But the present model generalizes the former studies by considering a more realistic description for the harmonic modes and for their coupling matrix elements to the electron in the junction considered.

The paper is organized as follows. In section 2 we describe the type of Hamiltonian used to treat the coupling between the electron and the surface plasmon. We also give the expressions for the surface plasmon frequencies and the matrix elements describing the coupling to the electron. We outline the method for solving the many-body Schrödinger equation (section 2.2). Details of the propagation matrix method that we use are given in the appendix. We derive the expression for the local, energy-dependent, effective dynamical potential (section 2.4). In sections 3.1 and 3.2, results for typical tunnelling conditions are presented for the partial and total transmission amplitudes and currents. The dependence of the dynamical effective potential upon the plasmon frequencies, the tunnelling energies and the applied bias are considered in section 3.3. A quantitative study of the corresponding tunnelling times is performed in section 3.4. We discuss the variation of the apparent barrier height for different tunnelling conditions in section 3.5. Finally in section 4, we summarize the most significant results that we have obtained, and discuss improvements and other potential applications of the method.

2. Model for the tunnelling barrier

2.1. The electron–surface plasmon coupling

We consider a model case where an electron tunnels through a static barrier V between two planar metallic surfaces. The metallic electrodes are described by a local scalar dielectric function. Within the tunnelling barrier ($0 \leq z \leq L$), the electron is coupled only to the surface plasmon (SP) modes. Each mode, of frequency ω_ν and coupling matrix element Γ_ν , is characterized by a composite index $\nu = (\mathbf{q}, \alpha)$ where \mathbf{q} is a vector parallel to the electrode surfaces and $\alpha = (\pm)$ represents the even and odd SP modes. These modes correspond to the in-phase and out-of-phase oscillation of the charge on the two electrode surfaces. The expressions for the frequencies and the coupling elements are (in atomic units) [25, 21]

$$\omega_\nu = \omega_{\mathbf{q},\alpha} = \omega_s \sqrt{1 + \alpha e^{-qL}} \quad (1)$$

and

$$\Gamma_\nu = \Gamma_{\mathbf{q},\alpha}(z) = \sqrt{\frac{2\pi\omega_s^2}{Aq\omega_{\mathbf{q},\alpha}}} \frac{e^{q(z-L)} + \alpha e^{-qz}}{2} \quad (2)$$

where A is the surface area of the unit cell and ω_s is the surface plasmon frequency, related to the bulk plasmon frequency ω_p by the usual relation ($\omega_s = \omega_p/\sqrt{2}$).

The Hamiltonian for the model consists of three terms: a purely electronic term (the kinetic energy operator and the static barrier V), the energy of the SP modes and the term describing the coupling between the electron and the potential generated by the surface charge oscillations (SP modes). The coupling to the SP modes is taken to be linear in the coordinates of the plasmons. The translational invariance parallel to the electrode surfaces allows us to write a one-dimensional Hamiltonian for the model system when only the motion of the electron perpendicular to the electrode surfaces is considered. Furthermore, on discretization of real space, the electron kinetic energy operator $-\frac{1}{2}\nabla_z^2 f(z)$ becomes

$-\frac{1}{2}(f(z_{j+1}) - 2f(z_j) + f(z_{j-1}))/\Delta^2$ (Δ being the grid spacing in the z -direction). Then, the corresponding Hamiltonian is written (in atomic units) as

$$H = \sum_j \epsilon_j c_j^\dagger c_j + \sum_{j,k} \beta_{j,k} (c_j^\dagger c_k + \text{HC}) + \sum_v \omega_v b_v^\dagger b_v + \sum_{j,v} \Gamma_v(z_j) c_j^\dagger c_j (b_v + b_v^\dagger) \quad (3)$$

where b_v^\dagger and b_v (c_j^\dagger and c_j) are the creation and annihilation operators of the SP modes (of the electron at site j), $\beta_{j,k} = -(1/2\Delta^2)\delta_{k,j+1}$ and $\epsilon_j = V(z_j) + 1/\Delta^2$. The basis vectors when a single electron is present are of the form $|j, n_{v_1}, n_{v_2}, \dots\rangle \equiv |j, \{n_v\}\rangle$, where each harmonic mode v contains $n_v = 0, 1, 2, \dots$ quanta. The main difficulty is that of determining the many-body wave functions of equation (3) inside the barrier without making any *ansatz* about the form of such wave functions, such as a separation between the harmonic part and the real-space part of the wave functions [26, 27]. In the present model, we want to retain fully the dynamical correlation between the tunnelling electron and the plasmon displacements.

2.2. A way to solve the Schrödinger equation

The Schrödinger equation for the Hamiltonian equation (3) can be solved for a given energy E following the procedure of Bonča and Trugman [22]. In this approach, the many-body problem is mapped exactly onto a one-body problem with many channels. The Schrödinger equation is then solved for the one-body problem. The problem is not solved as a standard eigenvalue–eigenvector system, since the eigenvalue E is known in advance. Essentially any problem can be solved with this method when a single electron tunnels and when the many-body interactions are limited to a finite region of space. These conditions are fulfilled for the model junction where the electron is coupled to the SP modes only inside the tunnelling barrier.

The different channels correspond to the incoming and backscattered waves on the left-hand electrode and the outgoing (transmitted) waves on the right-hand electrode. Each channel is characterized by a set of occupation numbers $\{n_v\}$ for the harmonic modes v . We consider only one incident electron, in a plane wave coming from the left, in the elastic channel. These are the physically relevant boundary conditions, since we normally consider the regime for which $k_B T \ll \hbar\omega_p$ and no plasmons are present before the electron arrives. In the elastic channels (left-hand and right-hand electrodes), all of the harmonic modes are in the ground state, i.e. $n_v = 0$ for all modes v (also denoted the $\{n_v\} = \{0\}$ channels) in contrast to the inelastic channels for which some modes are excited. Then the solution of the Schrödinger equation for the energy E is obtained by expanding the wave functions $\Psi(z_j, E)$ at a site j inside the barrier as

$$\Psi(z_j, E) \equiv \Psi_j(E) = \sum_v a_{j,\{n_v\}} |j, \{n_v\}\rangle \quad (4)$$

where the coefficients $a_{j,\{n_v\}}$ are complex numbers. Note that $j = 0$ (M) corresponds to the left-hand (right-hand) side of the barrier i.e. $z = 0$ ($L = M\Delta$).

Outside the barrier, no electron–SP couplings are considered and only the first three terms in the Hamiltonian equation (3) remain. The solutions of the corresponding Schrödinger equations in the left-hand and right-hand electrodes give the dispersion relations for the wave vectors:

$$E = \epsilon_{L,R} + 2\beta \cos(k_{L,R}^{(0)} \Delta) \quad \text{for the elastic channels} \quad (5)$$

and

$$E - \sum_v n_v \omega_v = \epsilon_{L,R} + 2\beta \cos(k_{L,R}^{(n_v)} \Delta) \quad \text{for the inelastic channels.} \quad (6)$$

We have assumed that the static potentials inside the left-hand and right-hand electrodes (defining $\epsilon_{L,R}$) are constant. This will be the case inside metals because of their characteristic short screening lengths.

With the above boundary conditions, the incoming and outgoing channel wave functions are described by

$$a_{j,\{n_v\}}^L = \begin{cases} e^{ik_L^{(0)} \Delta j} + r_{\{0\}} e^{-ik_L^{(0)} \Delta j} & \text{(elastic channel)} \\ r_{\{n_v\}} e^{-ik_L^{(n_v)} \Delta j} & \text{(inelastic channels)} \end{cases}$$

for the left-hand electrode (i.e. $j \leq -1$) and

$$a_{j,\{n_v\}}^R = t_{\{n_v\}} e^{ik_R^{(n_v)} \Delta j} \quad (\text{all of the channels}) \quad (7)$$

for the right-hand (i.e. $j \geq M+1$) electrode. The complex coefficients $r_{\{n_v\}}$ and $t_{\{n_v\}}$ are the reflection and transmission coefficients for each channel. Note that when the energy lies outside the band energies defined by equation (5) and equation (6), the corresponding channel is not propagating. Instead, we take exponentially decaying wave functions of the form $\exp(q_L^{(n_v)} \Delta j)$ and $\exp(-q_R^{(n_v)} \Delta j)$ for the left-hand and right-hand electrodes respectively.

On solving the Schrödinger equation for site $j = -1$ (site $j = M+1$), the reflection (transmission) coefficients are related to the wave-function coefficients $a_{j=0,\{n_v\}}$ ($a_{j=M,\{n_v\}}$) at the barrier edges. The relations are

$$r_{\{0\}} = -1 + a_{0,\{0\}} \quad \text{and} \quad r_{\{n_v\}} = a_{0,\{n_v\}} \quad (8)$$

and

$$t_{\{n_v\}} = e^{-ik_R^{(n_v)} \Delta M} a_{M,\{n_v\}} \quad (9)$$

(in the case of non-propagating channels, the complex exponential is replaced by the corresponding decaying exponential).

Taking into account the boundary conditions, the problem becomes equivalent to solving a complex linear system of the form $\mathbf{M}\mathbf{a} = \mathbf{d}$ where the components of vector \mathbf{a} are the $a_{j,\{n_v\}}$ coefficients and \mathbf{d} is proportional to the imaginary part of the incident wave function in the left-hand elastic channel. The matrix \mathbf{M} is by definition a sparse matrix, so algorithms for large sparse systems can be used. However, the solution of this linear system becomes numerically unfeasible when one considers a large number of sites inside the barrier and a large number of different harmonic modes ν (with a lot of different possible occupation numbers n_ν). This numerical problem can be substantially reduced by considering the propagation of the wave-function coefficients through the barrier. Introducing the vectors \mathbf{a}_j whose components are the wave-function coefficients on the different sites j , simpler tight-binding matrix equations can be derived from equation (3). These coupled equations can be solved in principle by a propagation matrix method [28], by a scattering matrix method [29] or by a recursion ratio matrix method [30]. In appendix A, we present a propagation matrix method for solving the problem. All of the results presented in this paper have been calculated with this method.

2.3. Reduced parameter space

The size of the parameter space (i.e. the size of the set $\{n_\nu\}$ of the occupation numbers) can be reduced further under certain conditions. When the voltage between the two electrodes is small (a few volts as in standard experimental conditions), there is only virtual excitation of the SP because of the high value of the SP frequency [19]. Then the essential physics of the problem will be given by the states with the lowest occupation numbers. This situation

corresponds to the weak-coupling limit considered by Puri and Schaich [23] for which the extent of the virtual excitation of the plasmon modes is bounded. Their method is similar to but less general than the method that we use in the present work.

When the energy E is large, it is possible to have real excitations of the plasmon modes which will correspond to propagating channels for the excited modes in the right-hand electrode. However, one can no longer speak about electron tunnelling in such conditions, so this regime is not relevant for STM.

On the other hand, the summation over the $\nu = (\mathbf{q}, \alpha)$ index in equation (3) goes in principle to an infinite value for \mathbf{q} . The SP modes decay into single-particle excitations for large \mathbf{q} , so a cut-off wave vector q_c can be introduced in the summation.

Although the method described in sections 2.1 and 2.2 is valid for the most general situations, we consider in this paper a model case corresponding to a reduced parameter space, i.e. the electron inside the tunnelling barrier is interacting with only two possible SP modes. One value for the \mathbf{q} -vector is chosen and associated with the modes whose frequency is given by equation (1), and the coupling matrix elements are now chosen to be [31]

$$\Gamma_{q,\alpha}(z) = \sqrt{\frac{\omega_s^2}{\omega_{q,\alpha}}} \frac{e^{q(z-L)} + \alpha e^{-qz}}{2}. \quad (10)$$

We also drop for simplicity the normalization versus the unit surface area A . Then, the wave-function coefficients $a_{j,\{n_\nu\}}$ inside the barrier are of the form $a_{j,n,m}$ where n and m are the occupation numbers of the even $\omega_{q,(+)}$ and odd $\omega_{q,-()}$ SP mode respectively.

Note finally that it is possible to rederive equation (3) for a three-dimensional system. For planar interfaces, the use of cylindrical coordinates and a Bessel transform is useful for decomposing the wave functions and surface charge oscillation modes according to their symmetry about a cylindrical axis. This extension to the present model will be considered in a forthcoming paper. Furthermore, for a more realistic description of the STM junctions, the corresponding surface plasmon frequency and coupling matrix elements can be derived from classical electrostatics [32].

2.4. The effective potential

In order to study in detail the dynamical aspects of the image potential for the present model, an expression for the effective electron potential V_{eff} has to be determined. Within the present formalism a natural way of proceeding is to define V_{eff} as the local (energy-dependent) potential

$$V_{\text{eff}}(z, E) = \frac{\langle \Psi(z, E) | \hat{V}_{\text{eff}} | \Psi(z, E) \rangle_\nu}{\langle \Psi(z, E) | \Psi(z, E) \rangle_\nu} \quad (11)$$

where

$$\hat{V}_{\text{eff}} = V + \sum_\nu \omega_\nu b_\nu^\dagger b_\nu + \sum_\nu \Gamma_\nu (b_\nu + b_\nu^\dagger).$$

The inner product in equation (11) goes only over the plasmon modes; no integration over the electron z -coordinate is implied. Then $V_{\text{eff}}(z_j, E)$ can be written as

$$V_{\text{eff}}(z_j, E) = V(z_j) + V_{\text{e-SP}}(z_j) / \sum_\nu \sum_{\{n_\nu\}} |a_{j,\{n_\nu\}}|^2 \quad (12)$$

where

$$V_{\text{e-SP}}(z_j) = \sum_v \sum_{\{n_v\}} \left\{ n_v \omega_v |a_{j,\{n_v\}}|^2 + \Gamma_v(z_j) 2 \sum_{\{m_v\}} \text{Re} [a_{j,\{m_v\}}^* \langle j, \{m_v\} | b_v^\dagger | j, \{n_v\} \rangle a_{j,\{n_v\}}] \right\}. \quad (13)$$

This potential contains all of the dynamical effects of the response of the surface plasmon due to the presence of an electron in the tunnelling barrier. For the reduced parameter space introduced in section 2.3, $V_{\text{e-SP}}$ becomes

$$V_{\text{e-SP}}(z_j) = \sum_{n,m} (n\omega_{q,(+)} + m\omega_{q,(-)}) |a_{j,n,m}|^2 + 2\sqrt{n}\Gamma_{q,(+)}(z_j) \text{Re} [a_{j,n,m} a_{j,n-1,m}^*] + 2\sqrt{m}\Gamma_{q,(-)}(z_j) \text{Re} [a_{j,n,m} a_{j,n,m-1}^*]. \quad (14)$$

The normalization factor in equation (12) is simply $\sum_{n,m} |a_{j,n,m}|^2$.

3. Results

In this section, we present results for a model tunnelling junction where the static barrier V is a square barrier of height V_0 (zero bias) and where a uniform electric field is applied between the two electrodes (non-zero bias ΔV), i.e. $V(z) = V_0 + \Delta V(z/L)$. The partial and total transmission amplitudes and currents are considered both when only virtual excitations of the SP modes are possible, and also when actual excitations occur. Then we present the effect of the dynamics on the effective electron potential for a given junction geometry. The relation between these dynamical effects and the tunnelling times of the electron is also examined. As a parameter that can be deduced from the experiments, we study the evolution of the apparent barrier height of the junction as a function of the electrode separation. Finally, in connection with experimental conditions, we discuss the possibilities of observing the dynamical effects in electron tunnelling.

3.1. Transmission amplitudes

In order to give a qualitative understanding of the difference of behaviour between purely elastic tunnelling through square barrier and inelastic tunnelling due to the coupling of the electron with the SP modes, figure 1 represents the transmission amplitudes $T_{nm} = |t_{nm}|^2$ for the reduced parameter space. We recall that n and m are the occupation numbers for the even $\omega_{q,(+)}$ and odd $\omega_{q,-(-)}$ modes respectively. Calculations have been done for two different maximum occupation numbers $n_{\text{max}} = 1$ and 2. The overall shape of the transmission amplitude T_{00} for the elastic channel is rather similar to the corresponding result obtained with no coupling to the SP modes inside the barrier. An exponential behaviour of the transmission amplitude is obtained as expected in the tunnelling regime ($E/V_0 < 1$) and oscillations characteristic of resonance effects occur for energies above the static barrier height V_0 . However, in the propagating regime ($E > V_0$), T_{00} is reduced compared to its values in the absence of coupling to the SP modes because some amplitudes are distributed not only in the elastic channels but in the inelastic channels as well. It should be noted that in any case the elastic channel makes the dominant contribution to the transmission in both tunnelling and propagating regimes. For $E < V_0$, the results are independent of the maximum occupation number used in the calculations (compare (a) and (b) in figure 1), but there are differences in the results at higher energies $E > V_0$. Indeed, the larger the energy, the more excitations can be created in the inelastic channels. The number of modes included should therefore be chosen with great care if one wishes to study the propagating

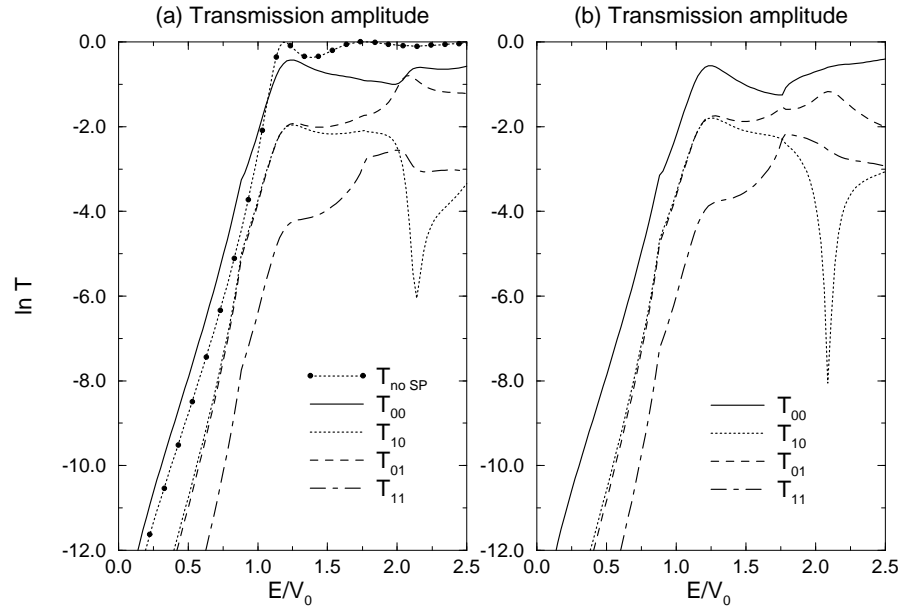


Figure 1. Transmission amplitudes T_{nm} (logarithmic scale) versus the electron energy E calculated for the following set of parameters characterizing the tunnelling junction: $V_0 = 0.4$, $L = 8.0$, $\omega_p = 0.5$ and $q = 0.90$. n and m are the occupation numbers for the even $\omega_{q,+}$ and odd $\omega_{q,-}$ modes respectively. Calculations have been performed for a maximum occupation per mode of (a) $n_{\max} = 1$ and (b) $n_{\max} = 2$. The transmission amplitudes are presented only for the channels with the lowest occupation numbers: $n, m = (0, 0)$, —; $(1, 0)$, ·····; $(0, 1)$, ---; and $(1, 1)$, — · —. The dotted line with filled circles in (a) represents the transmission amplitude obtained for a simple square barrier (i.e. when no electron-SP coupling inside the barrier is considered).

regime in detail; however, as previously mentioned, we concentrate in the present paper just on the tunnelling regime and do not discuss the behaviour of the transmission amplitudes for energies above the barrier height.

Note that, within the tunnelling regime, the values of the transmission amplitude T_{00} for the elastic channel are *larger* than the corresponding values obtained without the SP coupling inside the barrier. This behaviour is a signature of the lowering of the tunnelling barrier due to the interaction between the electron and the SP modes. This effect is also present in the tunnelling current as we will show below.

3.2. Tunnelling currents

In the present model, the current leaving through the outgoing channels in the right-hand electrode (or the reflected waves in the left-hand electrode channels) is obtained using

$$J_{k \rightarrow l}^{\{n_v\}} = 2 \operatorname{Im} [a_{k,\{n_v\}}^* \beta_{k,l} a_{l,\{n_v\}}]. \quad (15)$$

Since only hopping to adjacent sites is considered in equation (3), the current in each channel of the right-hand electrode (similar expressions can be obtained for the left-hand electrode) is found to be proportional to the transmission amplitudes:

$$J_R^{\{n_v\}} = 2\beta \sin(k_R^{\{n_v\}} \Delta) |t_{\{n_v\}}|^2 = 2\beta \sin(k_R^{\{n_v\}} \Delta) |a_{M,\{n_v\}}|^2. \quad (16)$$

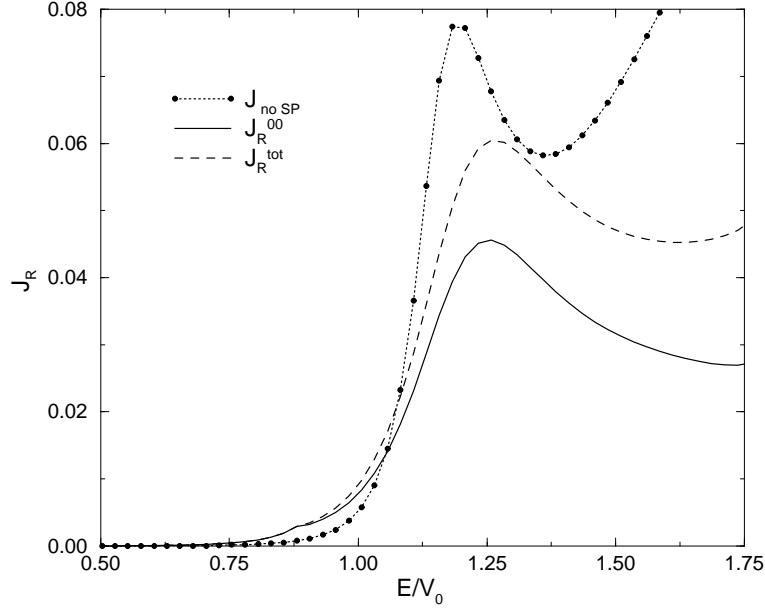


Figure 2. Currents going out to the right versus the energy E for the same set of parameters as in figure 1. The calculations have been performed for a maximum occupation per mode of $n_{\max} = 2$. The current J_R^{00} passing through the right-hand elastic channel is shown (—) as well as the total current leaving the junction on the right $J_R^{\text{tot}} = \sum_{n,m} J_R^{nm}$ (---). The dotted line with filled circles represents the current obtained for a simple square barrier without electron–plasmon coupling.

From equation (A12), it is obvious that the non-propagating channels (the corresponding wave functions have an exponentially decaying behaviour outside the barrier) do not carry any current, as expected. The current leaving the junction on the right versus the energy E is shown in figure 2. We have checked that within the range of parameters that we have studied, the main contribution to the total outgoing current J_R^{tot} always comes from the current J_R^{00} passing through the elastic channel. Non-zero current exists in the inelastic channels only above the energy threshold corresponding the real excitation of the SP modes. For example, the currents J_R^{10} , J_R^{01} and J_R^{11} are non-zero only for $E > \omega_{q,(+)}$, $\omega_{q,-}$ and $\omega_{q,(+)} + \omega_{q,-}$ respectively.

Note again that in the tunnelling regime, which is our main focus of interest, the total current J_R^{tot} and the partial current J_R^{00} are always larger than the current calculated without the coupling to the SP modes. As mentioned previously, this is a manifestation of the lowering of the tunnelling barrier felt by the electron because of their interaction with the SP modes. In this regime, the electron–SP interaction exists only via virtual excitation of those mode since the outgoing channels with $n_v \neq 0$ are not propagating. A detailed study of the modification of the tunnelling barrier due to the electron–SP mode interaction is presented in the next section.

3.3. The dynamical effective potential

The dynamical effects in the effective potential $V_{\text{eff}}(z, E)$ given by equations (12) and (14) are studied for various conditions. For a given geometry of the tunnelling junction,

$V_{\text{eff}}(z, E)$ is calculated for different tunnelling energies E and different values of the plasmon frequency $\omega_p = \omega_s \sqrt{2}$. These parameters vary the effective time taken by the electron to tunnel through the barrier and the speed at which the SP modes respond to the presence of an electron inside the barrier. The results for the effective potential are also compared to the static effective potential $V_{\text{eff}}^{\text{stat}}$ defined as

$$V_{\text{eff}}^{\text{stat}}(z) = V(z) - \sum_{q,\alpha} \frac{|\Gamma_{q,\alpha}(z)|^2}{\omega_{q,\alpha}}. \quad (17)$$

$V_{\text{eff}}^{\text{stat}}$ is the local, non-energy-dependent, potential felt by an electron interacting with the two SP modes and held fixed at position z inside the barrier. This situation corresponds to the semiclassical limit where the electron is treated as a classical point charge. No dynamical effects are included in this potential.

By performing explicitly the summation over q and α in equation (17), the exact classical image potential V_{im} is obtained [21, 33] as

$$V_{\text{im}}(z) = V(z) - \frac{1}{4L} [2\Psi(1) - \Psi(z/L) - \Psi(1 - z/L)] \quad (18)$$

where Ψ is the Psi (digamma function). The potential V_{im} has the classical divergences at both electrode surfaces; in particular, it can be shown that in the limit of large electrode separation, (i.e., where the electron ‘interacts’ only with one surface), the classical image potential takes the usual form $V_{\text{im}} = V(z) - 1/4z$. This can be done by taking the limit $L \rightarrow \infty$ in equation (18), or more easily by performing the summation in equation (17) with the corresponding asymptotic values for $\omega_{q,\alpha}$ and $\Gamma_{q,\alpha}$ defined by equation (1) and equation (2) respectively. In the case of the reduced parameter space, i.e., when only one value of q is retained, $V_{\text{eff}}^{\text{stat}}$ will be different from V_{im} since the summation over q is not complete. Therefore $V_{\text{eff}}^{\text{stat}}$ has no divergences at the surface of the electrodes but rather reflects the exponential behaviour of $\Gamma_{q,\alpha}(z)$ for the chosen value of q as can be seen from equation (17). However, $V_{\text{eff}}^{\text{stat}}$ will differ in principle from the dynamical effective potential $V_{\text{eff}}(z, E)$ because by definition $V_{\text{eff}}^{\text{stat}}$ does not contain any information about the dynamics of the coupled electron–SP system.

In figure 3 we show the variation of the dynamical effective potential for different values of the plasmon frequency ω_p . As expected, for a given tunnelling energy, the larger the plasmon frequency ω_p is, the closer $V_{\text{eff}}(z, E)$ is to the static effective potential. Indeed, for large ω_p -values, the dynamics of the SP modes is fast enough to build the corresponding classical image charge of the tunnelling electron. In the other extreme case where the plasmon frequency is very small, the SP dynamics is so slow that the tunnelling electron can cross the barrier without ‘feeling’ any modification of the potential. In that case, $V_{\text{eff}}(z, E)$ is almost equal to the nominal static barrier $V(z)$.

The asymmetry with respect to the middle of the tunnelling barrier that can be observed in the dynamical effective potential $V_{\text{eff}}(z, E)$ is a characteristic of a non-classical treatment for the electron (by contrast the static effective potential $V_{\text{eff}}^{\text{stat}}(z)$ is symmetric with respect to the position $z = L/2$, as also is the effective potential if the electron is taken as propagating with a constant velocity v). This behaviour is a signature of the intrinsic asymmetry of the tunnelling wave functions (for both elastic and inelastic channels) inside the barrier. These effects have already been pointed out by Šunjić and co-workers [9] and also observed in the results of Jonson and co-workers [12].

Figure 4 shows the modifications of $V_{\text{eff}}(z, E)$ when the tunnelling energy E varies. When the plasmon frequency is fixed, the dynamical effective potential becomes closer to the static effective potential $V_{\text{eff}}^{\text{stat}}(z)$ for tunnelling energies that approach the top of the

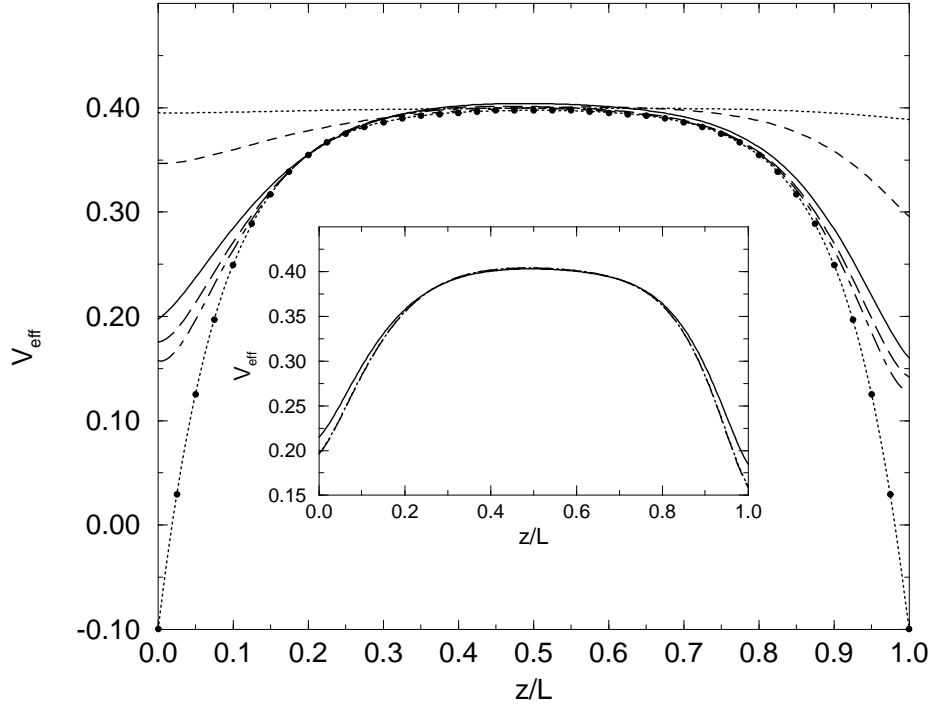


Figure 3. The effective potential $V_{\text{eff}}(z, E)$ for different values of the plasmon frequency: $\omega_p = 0.01$ (\cdots), $\omega_p = 0.1$ ($---$), $\omega_p = 0.5$ ($—$), $\omega_p = 1.0$ ($---$), $\omega_p = 1.5$ ($- \cdot -$). The square-barrier height is $V_0 = 0.4$, the tunnelling energy $E = 0.25$, the barrier length $L = 5.0$ and $q = 1.00$. The static effective potential $V_{\text{eff}}^{\text{stat}}(z)$ is also shown (the dotted line with filled circles). Note that for a given junction geometry, $V_{\text{eff}}^{\text{stat}}(z)$ is independent of the plasmon frequency ω_p . The inset shows the convergence of the results versus the maximum number of quanta considered in each mode (for $\omega_p = 0.5$ and the same values of the other parameters): $n_{\text{max}} = 1$ ($—$), $n_{\text{max}} = 2$ (\cdots), $n_{\text{max}} = 3$ ($---$). Convergence is reached for $n_{\omega}^{\text{max}} = 2$; this value is used for all of the results shown in the present work.

nominal static barrier V_0 . This behaviour is fully consistent with the modifications of the tunnelling time of the electron with respect to the tunnelling energy (see section 3.4).

There exist different definitions for the tunnelling time inside a barrier and we devote the next section to this problem. However, even if all of these times have different values for a given barrier and tunnelling energy, the same trends are observed as far as their dependence on E is concerned. The tunnelling times always increase when the tunnelling energy becomes closer to the top of the barrier; this can be interpreted as meaning that the electron tunnels more slowly as the energy E approaches the top of the barrier. Therefore, for slow electrons, the SP dynamics is fast enough to build almost the corresponding classical image charge as illustrated in figure 4, while for fast electrons it is not. Similar results are also obtained (for a given plasmon frequency ω_p and a given tunnelling energy E) by changing the barrier length L . Indeed for decreasing values of L , the tunnelling time becomes shorter (as might be expected) and therefore the dynamical effective potential strongly deviates from the corresponding static effective potential.

In figure 5, we present the modification of $V_{\text{eff}}(z, E)$ due to a negative bias ΔV applied to the right-hand electrode. The same trends are observed. The applied bias reduces the barrier height; therefore the ‘velocity’ of the tunnelling electron decreases and $V_{\text{eff}}(z, E)$

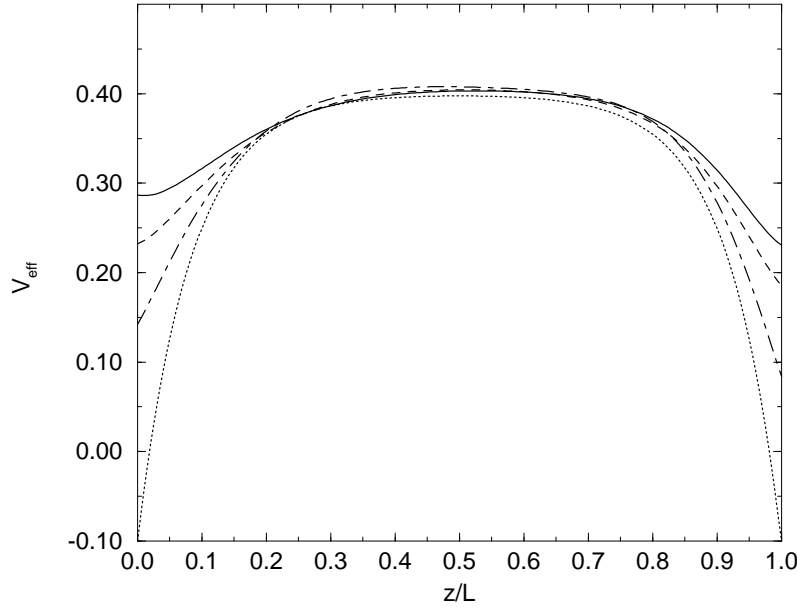


Figure 4. The effective potential $V_{\text{eff}}(z, E)$ for different values of the tunnelling energy: $E = 0.05$ (—), $E = 0.20$ (---), $E = 0.35$ (— · —), The square-barrier height is $V_0 = 0.4$, the plasmon frequency $\omega_p = 0.4$, the barrier length $L = 6.0$ and $q = 1.00$. The static effective potential $V_{\text{eff}}^{\text{stat}}(z)$ is represented by the dotted line.

becomes closer to $V_{\text{eff}}^{\text{stat}}$. This behaviour is particularly apparent on the right of the tunnelling barrier (i.e. $z \approx L$). It should, however, be noticed that the modifications of V_{eff} are almost negligible on the left-hand side of the barrier (i.e. $z \approx 0$), where the effects of the applied bias on the nominal static barrier are not important (and especially when the bias is not too large). The departure of V_{eff} from $V_{\text{eff}}^{\text{stat}}$ that can be observed when $\Delta V > E$ is due to the fact that the wave function of the elastic channel is no longer behaving like tunnelling wave functions for $0 \leq z \leq L$. The elastic channel wave function becomes oscillatory for $z \geq L(E - V_0)/\Delta V$. However, for the case presented in figure 5, the energies corresponding to the inelastic channels are still below the static barrier. And the results are still converged for $n_{\text{max}} = 2$. However, in other cases (for example, that of a larger value for the bias or that of a smaller barrier length), it might be necessary to enlarge the parameter space ($n_{\text{max}} > 2$) to obtain the converged results.

Finally, in a recent paper [34], we studied and compared the values of $V_{\text{eff}}(z, E)$ obtained by a path integral technique [21] to those obtained with the present matrix method. We obtained the same trends for the dependence of $V_{\text{eff}}(z, E)$ on the plasmon frequency and/or the tunnelling energy. However, the values obtained by the path integral technique are systematically larger than those obtained with the matrix method. We attribute this behaviour to the semiclassical approximation used to describe the electron motion in the path integral formalism.

3.4. Tunnelling times

In order to perform a more quantitative study of the concept of ‘slow’- or ‘fast’-tunnelling electrons introduced in the previous section, we calculate the values of different tunnelling

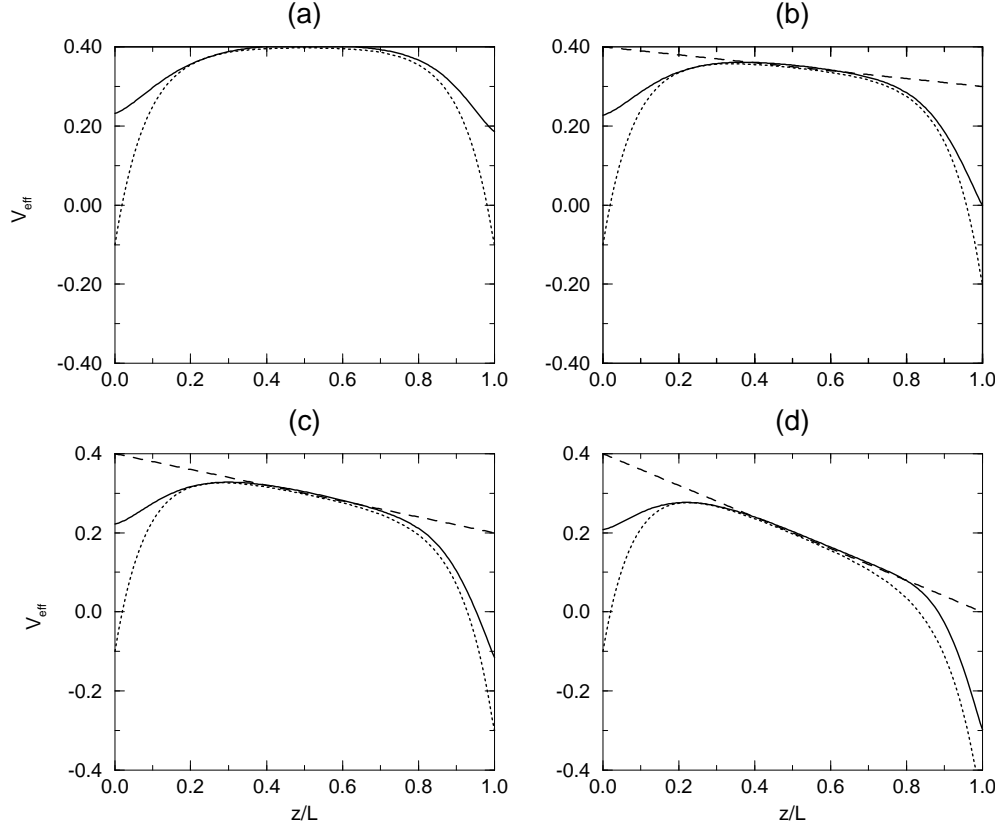


Figure 5. The effective potential $V_{\text{eff}}(z, E)$ (—) for different values of a negative bias ΔV applied to the right-hand electrode. (a) $\Delta V = 0.0$, (b) $\Delta V = -0.1$, (c) $\Delta V = -0.2$, (d) $\Delta V = -0.4$. The nominal square-barrier height is $V_0 = 0.4$, the barrier length $L = 6.0$, the tunnelling energy $E = 0.20$, the plasmon frequency $\omega_p = 0.4$ and $q = 1.00$. The static effective potential $V_{\text{eff}}^{\text{stat}}(z)$ is represented by the dotted line and the static barrier $V(z) = V_0 + \Delta V(z/L)$ by the dashed line.

times. There are several ways to define the time taken for a particle to complete a tunnelling process. One method involves the various ‘phase times’, which are related to the mean time taken by a tunnelling wave-packet [35]. However, it is a characteristic of tunnelling problems that the transmission depends exponentially on the energy, and this prevents one from carrying out a proper limiting process as the wave-packet is made broader and its spectral range is reduced. It has therefore been argued [36] that it is more physical to consider the interaction of the tunnelling particle with some local perturbation, present only in the barrier, and to use the effect that the perturbation has as a measure of the duration of the particle’s interaction with this perturbation. This approach leads to the Büttiker–Landauer time τ^{BL} [37, 38] which we consider below. Yet another possibility is derived from evaluating the tunnelling semiclassically, in which case the amplitude is dominated by complex ‘paths’ corresponding to classical motion of the particle in the inverted potential (i.e., to motion in which $E - V$ is everywhere replaced by $V - E$) [18, 20, 21]. One then uses the time associated with this classical motion as the tunnelling time. Another expression of the traversal time for a tunnelling electron has been proposed by Rudberg and

Jonson [12]. This time is obtained from a quantity interpreted as the speed of the particle and appears in the expression of the modified image potential obtained by Jonson from a self-energy formalism.

Table 1. Values for the tunnelling time proposed by Büttiker and Landauer through the barrier defined by the effective potential $V_{\text{eff}}(z, E)$. τ_0^{BL} is the tunnelling time corresponding to the static barrier $V(z) = V_0 + \Delta V(z/L)$. The other parameters are $V_0 = 0.4$, $L = 6.0$, $\Delta V = -0.2$, $\omega_p = 0.4$ and $q = 1.0$. Atomic units are used.

E	τ_0^{BL} No bias	τ^{BL} No bias	τ_0^{BL} ΔV	τ^{BL} ΔV
0.01	8.05	8.64	8.79	9.60
0.05	8.33	8.99	9.20	10.18
0.10	8.74	9.51	9.82	10.92
0.15	9.23	10.16	10.62	11.47
0.20	9.85	11.01	11.70	12.29
0.25	10.65	12.20	—	—
0.30	11.73	13.41	—	—

In this section, we consider only the tunnelling times proposed by Büttiker and Landauer and by Rudberg and Jonson. The Büttiker and Landauer tunnelling time is defined as

$$\tau^{\text{BL}} = \left| \frac{d}{dV} \ln t(E, \bar{V}) \right| \quad (19)$$

where t is the complex transmission coefficient and \bar{V} is the average height of the barrier. In the present calculations, we use the transmission coefficient $t_{\{0\}}$ of the outgoing elastic channel and the derivative with respect to the barrier height is obtained by varying the height V_0 of the nominal static barrier. The results obtained for a given set of parameters are listed in table 1.

Table 2. Values for the tunnelling time proposed by Rudberg and Jonson through the effective potential $V_{\text{eff}}(z, E)$. τ_0^Ψ is the tunnelling time corresponding to the static barrier $V(z) = V_0 + \Delta V(z/L)$. The other parameters used are the same as in table 1.

E	τ_0^Ψ No bias	τ^Ψ No bias	τ_0^Ψ ΔV	τ^Ψ ΔV
0.05	8.46	8.91	8.59	8.96
0.10	8.55	8.98	9.18	9.44
0.15	8.91	9.38	10.01	10.11
0.20	9.49	10.04	11.22	11.20
0.25	10.35	11.10	—	—
0.30	11.69	12.75	—	—

The tunnelling time proposed by Rudberg and Jonson is defined as

$$\tau^\Psi = \int_0^L \frac{dz}{v_\Psi(z)} \quad (20)$$

where v_Ψ is interpreted as the ‘velocity’ of the electron and defined by $v_\Psi(z) = |d \ln \Psi(z, E)/dz|$. Similarly to in the calculations of τ^{BL} , the ‘velocity’ v_Ψ is calculated from the values of the elastic channel wave-function coefficients $a_{j,\{0\}}$ inside the barrier. The results are shown in table 2.

In all situations (coupling to the SP modes or not, presence of a bias or not) and as expected, the tunnelling times τ^{BL} and τ^{Ψ} increase with increasing tunnelling energy E . Therefore as mentioned above, the electron tunnels faster when its energy lies well below the barrier height than when it is close to the barrier height. When the coupling to the SP modes is introduced, the tunnelling times increase systematically. This is a signature, as already seen, of the lowering of the dynamical effective potential compared to the nominal static barrier. However, the values of τ^{BL} and τ^{Ψ} are always different and, for almost all tunnelling energies, τ^{Ψ} appears smaller than τ^{BL} . The time differences between when the SP coupling is included and when it is not are also always larger for the Büttiker–Landauer time than for the Rudberg–Jonson time.

Note that the values of τ_0^{BL} ($\Delta V = 0$) are slightly different to those obtained from the analytical expression of $t(E, V_0)$ derived for a square barrier (see equation (4.5) in reference [38]). These differences are due the fact that the Büttiker–Landauer times shown in table 1 are obtained with our matrix method. In this method, the real-space discretization gives results equal to the exact analytical one only for infinitely small grid spacing. It should also be mentioned that from the definition of τ^{Ψ} , divergent results are obtained when $v_{\Psi}(z) \approx 0$ close to the right-hand electrode. This situation occurs when the tunnelling energy E is very small.

Although it is possible to calculate τ^{BL} and τ^{Ψ} numerically for (almost) any energy E when a bias is applied, we do not give results for $E > 0.2$ in tables 1 and 2. In this range of energy, one does not have a purely tunnelling behaviour for the wave functions inside the entire nominal barrier length ($0 \leq z \leq L$). In this case, the concept of a tunnelling time becomes rather misleading.

3.5. Apparent barrier height

The results that we presented above are for the dynamical effective potential and for theoretical entities which are not directly accessible from the experiments. However, understanding and knowledge of them are essential for an accurate calculation of the current in a tunnelling junction. In order to make an easier connection with experimental results, we studied the evolution of the apparent barrier height ϕ_{app} versus the barrier length. We define the apparent barrier height in a common manner [1, 30] (in atomic units):

$$\phi_{\text{app}} = \frac{1}{8} \left(\frac{d \ln J_R^{\text{tot}}}{dL} \right)^2 \quad (21)$$

where J_R^{tot} is the total current coming from all the right-hand outgoing channels.

The evolution of ϕ_{app} versus the barrier length L is shown in figure 6 for different plasmon frequencies. All of the curves show a similar behaviour: the apparent barrier height is constant for large barrier lengths and ϕ_{app} decreases continuously when the two electrodes are brought close together. As expected, the values of $\phi_{\text{app}}(L)$ are bounded by the two limiting cases which correspond to the results obtained for a static square barrier (the dynamics of the SP is not fast enough to respond to the presence of the tunnelling electron) and for the static image potential characterized by a large plasmon frequency (in the present case $\omega_p \geq 12.0$).

For large barrier lengths L , the behaviour of the wave functions close to the surface electrodes is no longer affected by an increase of L . Then, the apparent barrier height comes from the values of $V_{\text{eff}}(z, E)$ close to the middle of the barrier as can be seen in the insets in figure 6. The apparent barrier height is then equal to $V_0 - E - \Delta V/2$ (ΔV being the applied bias) and does not contain any detailed information about the electron–SP

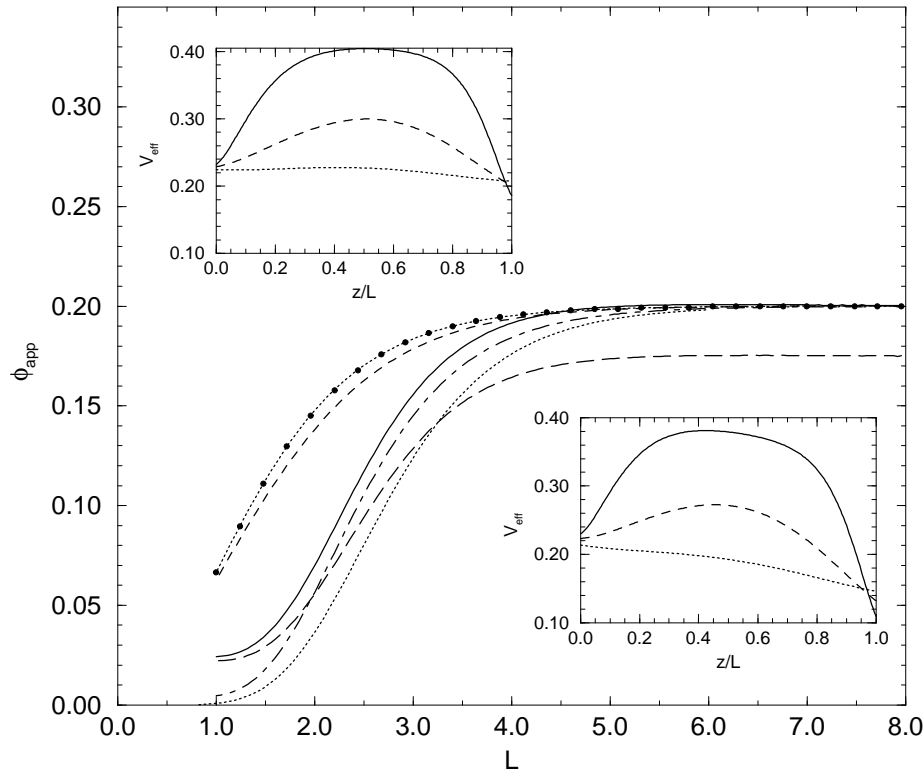


Figure 6. The apparent barrier height ϕ_{app} versus the barrier length L for $V_0 = 0.4$, $E = 0.20$ and $q = 1.00$. The calculations have been performed for different plasmon frequencies in the zero-bias limit: $\omega_p = 0.02$ (---), $\omega_p = 0.4$ (—), $\omega_p = 1.0$ (— · —), $\omega_p = 12.4$ (·····); and for a finite bias $\Delta V = -0.05$ and $\omega_p = 0.4$ (— · —). The dotted line with filled circles represents the apparent barrier height obtained for a simple square barrier (no electron–SP coupling). The upper inset represents the corresponding effective potential $V_{\text{eff}}(z, E)$ for three given barrier lengths $L = 6$ (—), $L = 2$ (---) and $L = 1$ (·····) in the zero-bias limit ($\omega_p = 0.4$). The lower inset shows $V_{\text{eff}}(z, E)$ for the same parameters as in the upper inset but with the finite applied bias $\Delta V = -0.05$.

coupling. However, the effects of the SP coupling affect the exponential behaviour of the wave functions mainly close to the electrode surfaces. Therefore the absolute values of the current, even for large separations L , are strongly dependent on the electron–SP coupling characteristics as can be seen in figure 7. For the set of parameters chosen, the current values vary according to the plasmon frequency up to more than four times the value of the current obtained in the absence of coupling to the surface plasmon modes. Figure 7 is a typical example showing the importance of the role of dynamics of the coupled electron–surface plasmon system on the current which is not always apparent from the evolution of the barrier height.

For intermediate and small L , the apparent barrier decreases continuously to zero which is, in principle, the signature of the collapse of the tunnelling barrier (i.e. the energy E of the electron is larger than the maximum of $V_{\text{eff}}(z, E)$ in between the two electrodes). However, it is interesting to note that even in the case of a simple square barrier, $\phi_{\text{app}}(L)$ goes to zero for decreasing L although the barrier height ($\phi = V_0 - E$) should remain constant for all electrode separations. This behaviour is due to the fact that the transmission amplitude

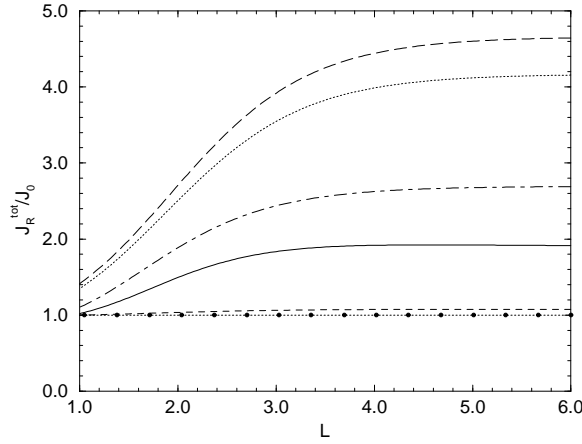


Figure 7. The ratio of the total current J_R^{tot} to the current J_0 obtained for a simple square barrier versus the barrier length L for $V_0 = 0.4$, $E = 0.20$ and $q = 1.00$. Calculations have been performed for different plasmon frequencies in the zero-bias limit: $\omega_p = 0.02$ (---), $\omega_p = 0.4$ (—), $\omega_p = 1.0$ (— · —), $\omega_p = 5.0$ (·····), $\omega_p = 12.4$ (— — —).

$T = |t|^2$ goes continuously to $T \rightarrow 1$ when $L \rightarrow 0$ (note that T is proportional to the current for a given tunnelling energy E). The corresponding expression for $\phi_{\text{app}}(L)$ can be easily obtained from equation (21) for the analytical solution of the square-barrier problem.

When the two electrodes are brought close together, the strength of the coupling between the tunnelling electron and the surface plasmon modes increases. The increase of the coupling is more important for the even (+) mode than for the odd (−) mode. Ultimately, the out-of-phase oscillation (odd mode) of the surface charges vanishes when $L \rightarrow 0$ which is the expected behaviour. The apparent barrier height decreases faster in the presence of coupling to the SP modes. This decrease is due to the two effects mentioned above: (i) the effective collapse of the potential barrier; and (ii) the increasing (and eventual saturation) of the transmission amplitude when the electrodes come close together. The detailed analysis of the evolution of $\phi_{\text{app}}(L)$ especially for small L -values is complex, as it results from a competition between the strengthening of the matrix coupling elements $\Gamma_{q,(\pm)}(z)$ and the shortening of the tunnelling times when L decreases. Note that we have also considered the evolution of the maximum barrier height $\phi_m(L) = \max_{z \in [0, L]} (V_{\text{eff}}(z, E) - E)$. A qualitatively similar behaviour is obtained for $\phi_m(L)$, i.e. the gradual collapse of the potential barrier when L decreases characterizes the decrease of $\phi_m(L)$. However, since the effects of the transmission amplitude are not included in the definition of ϕ_m , its modifications versus the plasmon frequency (for example) are not as important than those of ϕ_{app} shown in figure 6. Furthermore, as expected, $\phi_m(L)$ remains (almost) constant in situations where the SP dynamics is not fast enough to respond to the presence of the tunnelling electron inside the barrier.

Finally, it is also interesting to note the close resemblance of our results for $\phi_{\text{app}}(L)$ to those obtained by other methods which at first sight appear to contain quite different physics. For example, the barrier height between two metallic electrodes has been studied within the local density approximation for *jellium* surfaces in references [39, 30]. A similar evolution for the barrier height is observed when the electrodes are brought close together. However, the physics involved in this work is different from that in the present model, since the overlap of the charge density on the two sides of the tunnelling junction is treated self-consistently.

This overlap increases when the electrodes come close together and corresponds to the gradual collapse of the potential barrier inside the tunnelling gap. This barrier collapsing is, as mentioned above, is only partially taken into account (in the present results) via the electron–SP ‘induced’ potential. Although in the model proposed in this paper, the nominal static barrier $V(z)$ is not calculated self-consistently, the part added to the effective potential due to the electron–SP coupling includes the dynamical effects of the coupled system (i.e. the effective potential is energy dependent). This dependence is only taken into account in the work of references [39, 30] within the local density approximation. Further work to obtain a more accurate description of the static barrier $V(z)$ will be considered in the near future.

4. Conclusion

In this paper, we have presented a general approach for studying the coupling between a single electron and ‘boson fields’ which are represented by a set of harmonic oscillators. The approach can be used for any nominal static potential barrier, for different harmonic modes and does not include any local or semiclassical approximation. The method consists in solving the many-body Schrödinger equation on a real-space grid by mapping the many-body problem exactly onto a one-body problem with many incoming and outgoing channels. A propagation matrix technique is used to solve the corresponding one-body problem. We have studied in particular the coupling of electrons with surface plasmon modes in model tunnelling junctions. We mainly considered the effects of the dynamics of the coupled electron–plasmon system on the effective dynamical potential ‘felt’ by the tunnelling electron. The modifications of the electron effective potential compared to the static image potential have been interpreted in terms of competition between timescales (the electron tunnelling times and the characteristic response time of the surface plasmon, i.e. the inverse of the plasmon frequency). The most important deviations of the effective dynamical potential from the static image potential are obtained, as expected, for short tunnelling times (i.e. for small separation between the electrodes or for tunnelling energies well below the static barrier height) and for small plasmon frequencies. We have also shown that the apparent barrier height which is often derived from the experiments is not the best quantity to use in order to characterize the dynamical effects. Instead, the absolute values of the current are strongly dependent on the dynamics of the electron–plasmon coupled system. Improvements of the present study would be necessary to consider three-dimensional systems consisting of planar surfaces or a more realistic STM such as the tunnelling junction described in reference [32]. It would also be interesting to investigate (and possibly implement in the present model) the procedure proposed by Grillo *et al* [40] to transform delocalized plasmon modes into projected local plasmon modes. This procedure permits one in principle to reduce the dimensionality of the problem. This possibility is also promising as regards studying the tip-induced plasmon modes observed in photon emission experiments by Berndt *et al* [2]. Finally, it should be noted that the model used in this paper is a general model by definition. Therefore it is possible in principle to study other kinds of coupling between electrons and harmonic modes.

Acknowledgment

We are grateful to the Engineering and Physical Sciences Research Council for support under grants GR/J67734 and GR/K80495.

Appendix A. The propagation matrix technique

In order to improve the numerical efficiency of the solution of the problem, it is convenient to consider the propagation of the wave-function coefficients $a_{j,\{n_v\}}$ through the barrier. For this, we introduce vectors \mathbf{a}_j whose components are the wave-function coefficients on sites j . The components are classified according to the occupation numbers n_v for each harmonic mode v . In general, for N_v different modes, there are $(n_{\max} + 1)^{N_v}$ components for each \mathbf{a}_j -vector if each harmonic mode has the same maximum occupation number n_{\max} . Then the propagation of the \mathbf{a}_j -vectors inside the barrier is given from equation (3) by the following tight-binding-like matrix equation:

$$\mathbf{M}_j \mathbf{a}_j = \beta \mathbf{a}_{j+1} + \beta \mathbf{a}_{j-1}. \quad (\text{A1})$$

The matrices \mathbf{M}_j are sparse matrices whose diagonal elements are $E - \epsilon_j - \sum_v n_v \omega_v$ for the different occupation numbers of each mode. The off-diagonal elements of \mathbf{M}_j are proportional to the electron-SP coupling elements (i.e. $\sqrt{1 + n_v} \Gamma_v(z_j)$ for the upper off-diagonal and $\sqrt{n_v} \Gamma_v(z_j)$ for the lower off-diagonal elements). For example, in the case of the two odd and even modes $\omega_{q,-}$ and $\omega_{q,+}$, the matrices \mathbf{M}_j are of the form (with $n_{\max} = 1$)

$$\begin{bmatrix} E - \epsilon_j & \sqrt{1} \Gamma_{q,+}(z_j) & \sqrt{1} \Gamma_{q,-}(z_j) & 0 \\ \sqrt{1} \Gamma_{q,+}(z_j) & E - \epsilon_j - \omega_{q,+} & 0 & \sqrt{1} \Gamma_{q,-}(z_j) \\ \sqrt{1} \Gamma_{q,-}(z_j) & 0 & E - \epsilon_j - \omega_{q,-} & \sqrt{1} \Gamma_{q,+}(z_j) \\ 0 & \sqrt{1} \Gamma_{q,-}(z_j) & \sqrt{1} \Gamma_{q,+}(z_j) & E - \epsilon_j - \omega_{q,+} - \omega_{q,-} \end{bmatrix}. \quad (\text{A2})$$

The aim of the method is to solve equation (A1) inside the barrier knowing the boundary conditions at $j = 0$ and $j = M$. The solution of equation (A1) on the left-hand ($j \leq -1$) and on the right-hand ($j \geq M$) sides of the barrier (where the electron-SP coupling does not exist but where the inelastic effects are taken into account) gives the dispersion relation for the $k_{L,R}^{(n_v)}$ numbers in the different channels.

Equation (A1) can be rewritten in the form of a propagation matrix equation [28] (forward propagation):

$$\begin{bmatrix} \mathbf{a}_{j+1} \\ \mathbf{a}_j \end{bmatrix} = \begin{bmatrix} \beta^{-1} \mathbf{M}_j & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a}_j \\ \mathbf{a}_{j-1} \end{bmatrix}. \quad (\text{A3})$$

It is then possible to link the wave-function coefficients from one side of the barrier to those from the other side. For example in the case of backward propagation, we write

$$\begin{aligned} \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} &= \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{a}_M \\ \mathbf{a}_{M+1} \end{bmatrix} \\ &= \begin{bmatrix} \beta^{-1} \mathbf{M}_1 & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_2 & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_3 & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \cdots \\ &\cdots \begin{bmatrix} \beta^{-1} \mathbf{M}_{M-1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_M & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a}_M \\ \mathbf{a}_{M+1} \end{bmatrix}. \end{aligned} \quad (\text{A4})$$

Then the Schrödinger equation can be solved for a particular j -site, for example the site $j = 0$ of the barrier:

$$\beta^{-1} \mathbf{M}_0 \mathbf{a}_0 = \mathbf{a}_1 + \mathbf{a}_{-1} \quad (\text{A5})$$

knowing that the wave-function coefficients \mathbf{a}_j on sites on the left of the barrier ($j \leq -1$) are related to the coefficients \mathbf{a}_0 by

$$\mathbf{a}_j \equiv \begin{bmatrix} e^{ik_L^{(0)} \Delta j} + r_{\{0\}} e^{-ik_L^{(0)} \Delta j} \\ \left(r_{\{n_v\}} \begin{Bmatrix} e^{-ik_L^{(n_v)} \Delta j} \\ e^{q_L^{(n_v)} \Delta j} \end{Bmatrix} \right) \end{bmatrix} \quad \text{for } j \leq -1. \quad (\text{A6})$$

($\{n_v\} = \{0\}$) represents the ground state for all of the harmonic modes, i.e. the occupation number for each mode equals zero; this state is also called the elastic channel as opposed to the inelastic channels for which a modes v can be ‘excited’, i.e. $n_v \neq 0$.) The complex reflection coefficients $r_{\{n_v\}}$ given by the boundary conditions are

$$r_{\{0\}} = -1 + a_{j=0, \{0\}} \quad \text{and} \quad r_{\{n_v\}} = a_{j=0, \{n_v\}}. \quad (\text{A7})$$

Using equation (A4) to relate \mathbf{a}_0 to \mathbf{a}_M and \mathbf{a}_M to \mathbf{a}_1 and equation (A6) for $j = -1$, one can transform equation (A5) into the following linear system:

$$(\beta^{-1} \mathbf{M}_0 - \mathbf{E}^L - \mathbf{X}_B \mathbf{Y}_B^{-1}) \mathbf{a}_0 = -2i \sin(k_L^{(0)} \Delta) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (\text{A8})$$

where

$$\mathbf{X}_B = \mathbf{B}_{21} + \mathbf{B}_{22} \mathbf{E}^R \quad \text{and} \quad \mathbf{Y}_B = \mathbf{B}_{11} + \mathbf{B}_{12} \mathbf{E}^R. \quad (\text{A9})$$

The matrices $\mathbf{E}^{L,R}$ are diagonal matrices whose elements are

$$\begin{Bmatrix} e^{ik_{L,R}^{(n_v)} \Delta} \\ e^{-q_{L,R}^{(n_v)} \Delta} \end{Bmatrix}.$$

These elements are arranged the same manner with respect to the different SP modes and occupation numbers as in the wave-function coefficient vectors \mathbf{a}_j .

The solution of the linear system equation (A8) gives the values of the wave-function coefficients at the particular site $j = 0$ and this solution can be propagated through the barrier using equation (A3) and equation (A6) for \mathbf{a}_{-1} to give the complete solution for the wave functions inside the entire barrier.

There are other ways to solve the problem. For example, the forward-propagation equivalent of equation (A4) is

$$\begin{bmatrix} \mathbf{a}_M \\ \mathbf{a}_{M-1} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_{-1} \end{bmatrix} \quad (\text{A10})$$

where

$$\begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} = \begin{bmatrix} \beta^{-1} \mathbf{M}_{M-1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_{M-2} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_{M-3} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \cdots \\ \cdots \begin{bmatrix} \beta^{-1} \mathbf{M}_1 & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta^{-1} \mathbf{M}_0 & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}. \quad (\text{A11})$$

Knowing that the wave-function coefficients on the right-hand side of the barrier are given by

$$\mathbf{a}_j \equiv \begin{bmatrix} t_{\{n_v\}} \begin{Bmatrix} e^{ik_R^{(n_v)} \Delta j} \\ e^{-q_R^{(n_v)} \Delta j} \end{Bmatrix} \end{bmatrix} \quad \text{for } j \geq M+1 \quad (\text{A12})$$

with

$$t_{\{n_v\}} = \left\{ \begin{array}{c} e^{-ik_R^{[n_v]} \Delta M} \\ e^{i q_R^{[n_v]} \Delta M} \end{array} \right\} a_{j=M, \{n_v\}} \quad (\text{A13})$$

(the same expression for the elastic and inelastic channels is obtained for the wave-function coefficients $a_{j \geq M}$ and the complex transmission coefficient $t_{\{n_v\}}$ due to the boundary conditions), the Schrödinger equation for the site $j = M$ is transformed into the following linear system:

$$(\beta^{-1} \mathbf{M}_M - \mathbf{E}^R - \mathbf{X}_F \mathbf{Y}_F^{-1}) \mathbf{a}_M = 2i \sin(k_L^{(0)} \Delta) (\mathbf{X}_F \mathbf{Y}_F^{-1} \mathbf{F}_{12} - \mathbf{F}_{22}) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (\text{A14})$$

where

$$\mathbf{X}_F = \mathbf{F}_{21} + \mathbf{F}_{22} \mathbf{E}^L \quad \text{and} \quad \mathbf{Y}_F = \mathbf{F}_{11} + \mathbf{F}_{12} \mathbf{E}^L. \quad (\text{A15})$$

The inversion of the matrices \mathbf{Y} can be avoided, by using the forward-propagation matrices \mathbf{F} to relate \mathbf{a}_0 to \mathbf{a}_M and the backward-propagation matrices \mathbf{B} to relate \mathbf{a}_M to \mathbf{a}_0 . In these conditions, the linear system to be solved becomes

$$(\beta^{-1} \mathbf{M}_M - \mathbf{E}^R - \mathbf{X}_F \mathbf{Y}_B) \mathbf{a}_M = -2i \sin(k_L^{(0)} \Delta) \mathbf{F}_{22} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (\text{A16})$$

In equation (A16) the matrix inversion of \mathbf{Y} is avoided; however, the simultaneous knowledge of both the backward-propagation \mathbf{B} and forward-propagation \mathbf{F} matrices is needed. In certain cases like for a symmetric (with respect to the middle $z = L/2$ of the barrier) static potential barrier $V(z)$, these backward-propagation and forward-propagation matrices are equivalent. This is not the case when a bias is applied to the nominal square barrier.

The solution of equation (A14) or equation (A16) gives \mathbf{a}_M . This solution can be back-propagated through the barrier to give the complete solution for the wave functions inside the entire barrier. This procedure is in principle strictly equivalent to the forward propagation of the solution \mathbf{a}_0 of equation (A8).

We have compared the results obtained by the present propagation matrix method to the exact analytic solution for the case of a square barrier (with no coupling to the SP modes) and to the results obtained by solving the original large linear system described in the main text. The propagation matrix method gives the same results as the analytic solution in the limit of very small grid spacing and the same results as for the large linear system. But the computational times for the propagation matrix method are reduced by at least an order of magnitude compared to the solution of the large linear system. We also found that the propagation matrix method ‘propagates’ errors due to the numerical finite precision. In some cases, especially for wave-function coefficients that have very small values, the propagation matrix method can lead to incorrect results close to the side of the barrier where the propagation finishes. These cases correspond to large barrier lengths and tunnelling energies well below the barrier height. We found that working with quadruple-precision routines suppresses these problems of instabilities of the algorithm. We have not found any such problems in the systems presented in this paper.

References

- [1] Binnig G, García N, Rohrer H, Soler J M and Flores F 1984 *Phys. Rev. B* **30** 4816
- [2] Berndt R, Gimzewski J K and Johansson P 1991 *Phys. Rev. Lett.* **67** 3796
- [3] Berndt R, Gimzewski J K and Johansson P 1993 *Phys. Rev. Lett.* **71** 3493
- [4] Guéret P, Marclay E and Meier H 1984 *Appl. Phys. Lett.* **53** 1617
- [5] Ray R and Mahan G D 1972 *Phys. Lett.* **42A** 301
- [6] Wu J W and Mahan G D 1983 *Phys. Rev. B* **28** 4839
- [7] Inkson J C 1973 *J. Phys. F: Met. Phys.* **3** 2143
- [8] Jonson M 1980 *Solid State Commun.* **33** 743
- [9] Tran Thoai D B and Šunjić M 1991 *Solid State Commun.* **77** 955
- [10] Šunjić M and Marušić L 1992 *Solid State Commun.* **84** 123
- [11] Šunjić M and Marušić L 1991 *Phys. Rev. B* **44** 9092
- [12] Rudberg B G R and Jonson M 1991 *Phys. Rev. B* **43** 9358
- [13] Marušić L and Šunjić M 1993 *Solid State Commun.* **88** 781
- [14] Young R A 1983 *Solid State Commun.* **45** 263
- [15] Caldeira A O and Leggett A J 1981 *Phys. Rev. Lett.* **46** 211
- [16] Caldeira A O and Leggett A J 1983 *Ann. Phys., NY* **149** 374
- [17] Persson B N J 1988 *Phys. Scr.* **38** 282
- [18] Persson B N J and Baratoff A 1988 *Phys. Rev. B* **38** 9616
- [19] Sebastian K L and Doyen G 1993 *Phys. Rev. B* **47** 7634
- [20] Sebastian K L and Doyen G 1993 *J. Chem.* **99** 6677
- [21] Klipa N and Šunjić 1995 *Phys. Rev. B* **52** 12408
- [22] Bonča J and Trugman S A 1995 *Phys. Rev. Lett.* **75** 2566
- [23] Puri A and Schaich W L 1983 *Phys. Rev. B* **28** 1781
- [24] Tagliacozzo A and Tosatti E 1988 *Phys. Scr.* **38** 301
- [25] Šunjić M and Lucas A A 1971 *Phys. Rev. B* **3** 719
- [26] Lee T D, Low F E and Pines D 1953 *Phys. Rev.* **90** 297
- [27] Evans E and Mills D L 1972 *Phys. Rev. B* **8** 4004
- [28] Sautet P and Joachim C 1988 *Phys. Rev. B* **38** 12238
- [29] Ko D Y K and Inkson J C 1988 *Phys. Rev. B* **38** 9945
- [30] Hirose K and Tsukada M 1995 *Phys. Rev. B* **51** 5278
- [31] In principle, the summation $\sum_{q,\alpha} \rightarrow \int d^2q/(2\pi)^2 = \int q dq d\theta/(2\pi)^2$ suppresses the possible singularity in $1/q$ that can be obtained in the expression for the electron effective potential. In a self-energy or path integral formalism, the dynamical image potential is always found to be related to the square of the coupling matrix elements Γ_v . This is why we do not consider the factor $\sqrt{2\pi/q}$ in equation (10).
- [32] Šestović D, Marušić L and Šunjić M 1997 *Phys. Rev. B* **55** 1741
- [33] Lenac Z and Šunjić M 1976 *Nuovo Cimento B* **33** 681
- [34] Ness H and Fisher A J 1998 *Appl. Phys. A* **66** at press
- [35] Hauge E H and Støvneng J A 1989 *Rev. Mod. Phys.* **61** 917
- [36] Landauer R and Martin Th 1994 *Rev. Mod. Phys.* **66** 217
- [37] Büttiker M and Landauer R 1982 *Phys. Rev. Lett.* **49** 1739
- [38] Büttiker M and Landauer R 1985 *Phys. Scr.* **32** 429
- [39] Ferrante J and Smith J R 1985 *Phys. Rev. B* **31** 3427
- [40] Grillo M E, Castro G R and Doyen G 1992 *J. Phys.: Condens. Matter* **4** 5103