Anomalous current-voltage characteristics of tunnel junctions

Yu. V. Nazarov

Landau Institute of Theoretical Physics, USSR Academy of Sciences

(Submitted 4 July 1988)
Zh. Eksp. Teor. Fiz. **95**, 975–984 (March 1989)

A method is developed that permits the influence of electromagnetic interaction on the rate of one-electron tunneling to be taken into account outside the scope of perturbation theory. Such an account is possible because of the semiclassical character of the carrier motion and the weakly varying electromagnetic field. The method is used to calculate the zero-bias anomalies of the current-voltage characteristic and to solve problems involving the Coulomb blockade of the tunneling; a number of new results are obtained.

Two classes among the diverse phenomena in tunnel junctions cannot be explained in the framework of the one-particle scheme and are due to interelectron interactions. The first class comprises the zero-bias anomalies.[1] These are small deviations from Ohm's law ($\delta R/R \propto V^{1/2}$ or $\ln V$) observed frequently in the current-voltage characteristics (IVC) of a tunnel junction at voltages $V < 10$ mV. In Ref. 2 and in preceding papers[3,4] these effects are attributed to the influence of interelectron Coulomb interaction on the carrier energy density.

The second class of phenomena, actively investigated of late, is connected with the Coulomb blockade of the tunneling (see Ref. 5). This blockade was indirectly observed also in an experiment[6] under conditions when the tunneling between the metallic banks was via conducting granules imbedded in an insulating interlayer. In this case the need for charging the granule on passage of an electron and the corresponding increase of the Coulomb energy of the system lead to suppression of the tunneling at a potential difference $\sim e/C$, where C is the capacitance of the granule. Experiments were reported in which this phenomenon was observed also without granules, in which case C pertains to the capacitor formed by the banks of the junction. A well developed theory exists for this phenomenon, makes use of the concepts of network theory, and is therefore phenomenological. However, in view of the advances in the experimental techniques, the need increases for a microscopic corroboration of the employed approach.

Since both phenomenon classes are of the same nature, it is natural to treat them from a unified point of view. The problem of electron tunneling in the presence of electromagnetic interaction can be divided into two, by considering separately electron motion in a specified external field and the fluctuations of this field. Note that the phenomena of interest to us take place at low contact potential differences, $eV \ll \varepsilon_F$, and at low temperatures, $T \lesssim eV$. Under these conditions the tunneling rate is determined by the behavior of the system over relatively long times, of order $(eV)^{-1}$, so that account need be taken of only the low-frequency part of the electromagnetic field. The field fluctuations can then be regarded as quadratic, and the carrier motion as semiclassical. This permits a departure from perturbation theory and consideration of large deviations of the IVC from Ohm's law.

The plan of the article is the following: The method employed is described in Sec. 1. In Secs. 2 and 3 the zero-bias anomalies are calculated for a simple symmetric geometry, and a comparison is made with the available theoretical and experimental results. The connection between the Coulomb blockade and the zero-bias anomalies is considered in Sec. 4, and the corrections to the results of the phenomenological theory are determined. In Sec. 5 is considered a new type of low-voltage IVC anomaly that takes place in long junctions, and the possibilities of observing this effect in experiment are determined.

1. We use the tunnel-Hamiltonian method, which we modify to take the electromagnetic interaction into account. The total Hamiltonian of the problem consists of the following parts:

$$\hat{H} = \hat{H}_R + \hat{H}_L + \hat{H}_{Imp} + \hat{H}_T + \hat{H}_{ph}.$$

Here $\hat{H}_{R,L}$ describe the carriers in the two banks of the junction and the interaction of these carriers with the electromagnetic field; it is convenient to use the gauge of this field in the form $\varphi = 0$,

$$\hat{H}_R = \int_{x\subset R} d^3x\, \psi^+(\mathbf{x})[\varepsilon(\hat{\mathbf{p}}) - \mu + eV/2]\psi(\mathbf{x}),$$

$$\hat{H}_L = \int_{x\subset L} d^3x\, \psi^+(\mathbf{x})[\varepsilon(\hat{\mathbf{p}}) - \mu - eV/2]\psi(\mathbf{x}),$$

$$\hat{p}^\alpha = -i\frac{\partial}{\partial x^\alpha} - e\hat{A}^\alpha(x)/c,$$

V is the potential difference across the junction. The Hamiltonian $H_{\rm Imp}$ describes carrier scattering by the impurities and the boundaries of the metals. The tunnel term H_T describes the transitions of the carriers through the insulating interlayer:

$$H_T = \int_{x\subset R, x'\subset L} d^3x\, d^3x'\, T(\mathbf{x},\mathbf{x}')\psi^+(\mathbf{x})\psi(\mathbf{x}') + \text{c.c.}$$

at $A = 0$. The corresponding expression for the current operator is

$$I = i\int_{x\subset R, x'\subset L} d^3x\, d^3x'\, T(\mathbf{x},\mathbf{x}')\psi^+(\mathbf{x})\psi(\mathbf{x}') + \text{c.c.}$$

Finally, the term $H_{\rm ph}$ is the free-photon operator. In this scheme it is necessary to take into account the modification of the tunnel amplitude by the electromagnetic field. To determine this modification it is necessary to use gauge invariance premises as well as the facts that a) the electron tunnels along the most probable trajectory, and b) as stated in the Introduction, only the low-frequency part of the electromagnetic field is important. This makes it possible to write

 0038-5646/89/030561-06$04.00 © 1989 American Institute of Physics

$$T(\mathbf{x},\mathbf{x}'|A)=T(\mathbf{x},\mathbf{x}')\exp\Big(-ie\int_{\mathbf{x}}^{\mathbf{x}'} dz\,\mathbf{A}(z)/c\Big), \qquad (1)$$

where the integral is taken along the most probable trajectory—a straight line joining the points $\mathbf{x}$ and $\mathbf{x}'$.

Our purpose is to calculate the current in second-order perturbation theory in the tunnel Hamiltonian, as well as in the context of the standard approach. The expression of the current for arbitrary V can be obtained by analytically continuing to the complex plane the Matsubara current Green's function determined at $\omega_n = 2\pi Tn$, as well as H_T taken at $V=0$:

$$I=\mathrm{Im}\Big\{\int_{-\beta}^{+\beta} d\tau \langle\langle \hat{I}(0)\hat{H}_T(\tau)\rangle\rangle e^{i\omega_n\tau}|_{i\omega_n\to eV+i0}\Big\}.$$

This Green's function can be written in standard fashion (see, e.g., Ref. 7) in the form of a functional integral over the Fermi fields of the electrons and over the Bose fields of the photons. Taking the Gaussian integral over the Fermi fields, we express the sought quantity in terms of the Green's function of electrons in an external electromagnetic field:

$$I=\mathrm{Re}\Big\{\int_{-\beta}^{+\beta} d\tau\, e^{i\omega_n\tau}\int_{\substack{\mathbf{x},\mathbf{x}_1\subset R;\\ \mathbf{x}',\mathbf{x}_1'\subset L}} d^3x\, d^3x_1\, d^3x'\, d^3x_1'$$
$$\times\langle T^*(\mathbf{x},\mathbf{x}'|A)T(\mathbf{x}_1,\mathbf{x}_1'|A)G^R(\mathbf{x},\tau_1;\mathbf{x}_1,\tau_1+\tau|A)$$
$$\times G^L(\mathbf{x}_1',\tau+\tau_1;\mathbf{x}',\tau_1|A)\rangle_A|_{i\omega_n\to eV+i0}\Big\}. \qquad (2)$$

The angle brackets denote averaging over the electromagnetic-field configurations

$$\langle X(A)\rangle=\int\prod_{x,\tau} d\mathbf{A}(x,\tau)X(A)\exp(-S(A))\Big/\int\prod_{x,\tau} d\mathbf{A}(x,\tau)\exp(-S(A)).$$

The field fluctuations can be regarded as Gaussian,

$$S(A)=\frac{T}{2}\sum_{\omega_n}\int d^3x\, d^3x'\, A^\alpha_{\omega_n}(\mathbf{x})D^{\alpha\beta}(\mathbf{x},\mathbf{x}')A^\beta_{-\omega_n}(\mathbf{x}'),$$

where $D^{\alpha\beta}_\omega(\mathbf{x},\mathbf{x}')$ is the Green's function of a photon in the medium. Equation (2) was symmetrized with respect to R and L; it is implied thereby that the banks are made of the same material.

We proceed now to the Wigner representation $G(\mathbf{x},\mathbf{x}')\to G(\mathbf{x},\mathbf{p})$ for the Green's function of the electrons, and integrate with respect to the variable $\xi=\varepsilon(p)-\mu$. These transformations allow us to write a clear expression for I:

$$I=\mathrm{Re}\Big\{\int_{-\beta}^{+\beta} d\tau\exp(i\omega_n\tau)\int d^2x\, d^2x'\, d^2n\, d^2n'\, W(\mathbf{n},\mathbf{x};\mathbf{n}',\mathbf{x}')$$
$$\times\langle\tilde{G}^R(\mathbf{x},\mathbf{n};\tau_1,\tau_1+\tau|A)$$
$$\times\tilde{G}^L(\mathbf{x}',\mathbf{n}',\tau_1+\tau,\tau_1|A)F(A)\rangle_A|_{i\omega_n\to eV+i0}\Big\}. \qquad (3)$$

Here $\tilde{G}$ is the electron Green's function integrated with respect to ξ, n parametrizes the Fermi surface of the metal, x and x' have values on the surfaces of the right and left banks, and $W(\mathbf{n},\mathbf{x};\mathbf{n}',\mathbf{x})$ has the meaning of the probability of carrier tunneling from point $\mathbf{x}$ with momentum $\mathbf{p}(n)$ to point $\mathbf{x}'$ with momentum $\mathbf{p}(n')$. The factor $F(A)$ stems from the modification (1) of the tunneling amplitude

$$F(A)=\exp\Big\{-ie\int_{x}^{x'} dz[\mathbf{A}(z,\tau_1+\tau)-\mathbf{A}(z,\tau)]\Big\}.$$

In the approximation we need it is easy to obtain an explicit expression for $\tilde{G}$:

$$\tilde{G}(\mathbf{x},n;\tau_1,\tau_2|A)=\tilde{G}_0(\tau_1-\tau_2)\exp\Big\{-iT\sum_{\omega_n}[\exp(i\omega_n\tau_1)-\exp(i\omega_n\tau_2)]\int j^\alpha_{\omega_n}(\mathbf{x},\mathbf{n};\mathbf{x}_1)A^\alpha_{-\omega_n}(x_1)d^3x_1\Big\},$$
$$\Big(\omega_n+\frac{\partial}{\partial x_\beta}v^\beta(n)+\mathrm{sign}\,\omega_n\int w(n,n')d^2n'\Big)j^\alpha_{\omega_n}(\mathbf{x},n;\mathbf{x}_1)$$
$$-\mathrm{sign}\,\omega_n\int w(n',n)j^\alpha_{\omega_n}(n')d^2n'=ev^\alpha(n)\delta^3(\mathbf{x}-\mathbf{x}_1). \qquad (4)$$

Here

$$\tilde{G}_0(\tau_1-\tau_2)=\sum_{\varepsilon_n}\exp(-i\varepsilon_n\tau)\,\mathrm{sign}\,\varepsilon_n,\qquad \varepsilon_n=(2n+1)\pi T,$$

and the probabilities $w(\mathbf{n},\mathbf{n}')$ of scattering by the impurities have been introduced. The derivation and discussion of the limits of validity of Eq. (4) for our problem can be found in the Appendix. The path integral with respect to A in Eq. (3) has a Gaussian form and can be evaluated. It is natural to assume that the tunneling takes place between nearest points of the metal–insulator surfaces; this can be written in the symbolic form $W(\mathbf{n},\mathbf{x};\mathbf{n}',\mathbf{x}')=w(\mathbf{n},\mathbf{n}';\mathbf{x})\delta(\mathbf{x}-\mathbf{x}')$. From (3) and (4) we obtain

$$I(V)\propto\mathrm{Im}\Big\{T^2\int_{-\beta}^{+\beta} d\tau\frac{\sin^2(\omega_n\tau/2)}{\sin^2(\pi T\tau)}\int d^2n\, d^2n'\, d^2x\,\tilde{w}(\mathbf{n},\mathbf{n}',\mathbf{x})$$
$$\times\exp[-S(\tau,\mathbf{n},\mathbf{n}',\mathbf{x})]|_{i\omega_n\to eV+i0}\Big\},$$
$$S(\tau)=T\sum_{\omega_n}S(\omega_n)\sin^2(\omega_n\tau/2),\qquad S(\omega_n,\mathbf{n},\mathbf{n}',\mathbf{x}) \qquad (5)$$
$$=\int d^3x_1\, d^3x_2\, j^\alpha_{-\omega_n}(\mathbf{x},\mathbf{x}_1,\mathbf{n},\mathbf{n}')D^{\alpha\beta}_{\omega_n}(x_1x_2)j^\beta_{\omega_n}(x,x_2,n,n'),$$
$$\mathbf{j}_{\omega_n}(\mathbf{x},\mathbf{x}_1,\mathbf{n},\mathbf{n}')=\mathbf{j}^R_{\omega_n}(\mathbf{x},\mathbf{x}_1,n)-\mathbf{j}^L_{\omega_n}(\mathbf{x},\mathbf{x}_1,n')+\mathbf{j}^{RL}_{\omega_n}(\mathbf{x},\mathbf{x}_1)$$

[$\mathbf{j}^{RL}_{\omega n}$ is determined here from $F(A)$], and these relations do solve our problem. The matter has now been reduced to solution of a kinetic equation for the quantities $j^{R,L}$ and of the material part of the Maxwell equations, and to solution of the electrodynamics equation for $D^{\alpha\beta}_\omega$.

If we consider now, as will be done below, space–time scales for which the kinetic equation can be replaced by the diffusion equation, and assume the junction to be homogeneous, we can neglect the dependences of $S(\tau,n,n',x)$ on n and n' and of $\tilde{w}(n,n',x)$ on x. The result is a simpler equation for the current:

$$I(V)\propto\mathrm{Im}\Big\{T^2\int_{-\beta}^{+\beta} d\tau\frac{\sin^2(\omega_n\tau/2)}{\sin^2(\pi T\tau)}\exp(-S(\tau))|_{i\omega_n\to eV+i0}\Big\}. \qquad (5a)$$

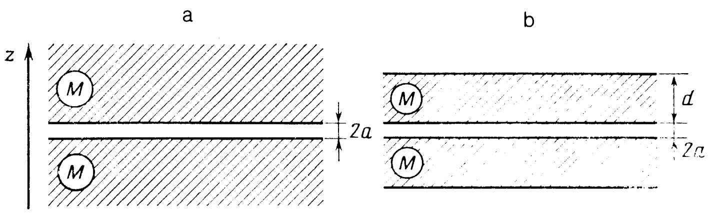

FIG. 1.

The coefficient in (5a) must be determined as follows: since we are interested here in low-voltage IVC anomalies, we separate in $S(\tau)$ the part responsible for these anomalies. Subtracting this part from $S(\tau)$ we find that $I\propto V$ as $V\to\infty$. This enables us to express the coefficient in terms of R_0, which is the junction resistance in the absence of anomalies.

We note in conclusion that the described method is similar to the procedure used to exclude infrared divergences in zero-spin quantum electrodynamics.[7]

2. We consider now an actual junction geometry (Fig. 1a). Two metallic half-spaces are separated by an insulating liner of thickness $2a$, and the z axis is perpendicular to the liner. We determine the influence exerted in the IVC by field fluctuations of space and time scales much larger than the electron momentum relaxation time $\tau_{\rm Imp}$ and the mean free path l. This allows us to change from the kinetic equation to the diffusion equation. For the isotropic part of $j^{R,L}$ we have

$$\left(|\omega_n|+D\mathbf{q}^2-D\frac{\partial^2}{\partial z^2}\right)\begin{Bmatrix} j^z \\ j^{x,y}\end{Bmatrix} = -D\begin{Bmatrix}\dfrac{\partial}{\partial z}\\ -iq^{x,y}\end{Bmatrix}\delta(z-z')\,\mathrm{sign}\,\omega_n.$$

Here D is the carrier diffusion coefficient and $\mathbf{q}$ is a wave vector lying in the liner plane. From this and from (1) we get for $\mathbf{j}$

$$e^{-1}j^z(0,z,\mathbf{q})=\theta(a-|z|)+\theta(|z|-a)\exp(-|z|\delta),$$
$$e^{-1}j^{x,y}(0,z,\mathbf{q})=-\mathrm{sign}\,z[\theta(|z|-a)\exp(-|z|\delta)i\mathbf{q}^{x,y}/\delta],$$
$$\delta^2=q^2+|\omega_n|/D.$$

It is convenient to represent $S(\omega_n)$ in the form

$$e\int d^3x\, j^\alpha_{-\omega}(x)A_\omega^\alpha(x)/2c,$$

and for A^α we have the electrodynamics equation

$$(\omega_n^2/c^2+\mathbf{rot\,rot})\mathbf{A}=4\pi[j^\alpha(\mathbf{A})+j_\omega^\alpha(x)]/c,\quad j^\alpha(\mathbf{A})=0,\quad |z|<a. \tag{6}$$

For materials with good metallic conductivity the Debye length r_D and the resistivity σ^{-1} are much smaller than the spatiotemporal scales of interest to us. This permits the use of the electroneutrality condition, which has in this case the form $\mathrm{div}\mathbf{A}=0$ in the metal; with this, $j^\alpha(A)=\sigma A^\alpha$. Solving (6) and calculating $S(\tau)$ we obtain

$$S(\tau)=8e^2a\int dq\,qT\sum_{\omega_n}\frac{\sin^2(\omega_n\tau/2)}{|\omega_n|(|\omega_n|+4\pi\sigma a\delta_1^{-1}q^2)}\times\left(1-\frac{q^2}{\delta_1(|\omega_n|/D+q^2)^{1/2}}\right),$$
$$\delta_1^2=q^2+4\pi\sigma|\omega_n|/c^2. \tag{7}$$

We call attention to the denominators. One of them describes carrier diffusion. The second corresponds to propagation of electromagnetic perturbations in the space near the liner. In real time the excitation of these perturbations can be visualized as follows: the tunneling act produces on the surfaces, near a certain point in the layer $\sim r_D$, excess charge densities of opposite sign on the different banks. The electromagnetic field produced by these charges excites conduction currents in the metal, and this causes the charge spot to spread out like $r\propto t^{3/4}$. As $\tau\to\infty$ we have $S(\tau)=\mathrm{const}+\tau^{-1/2}$. This corresponds to appearance of an increment $\propto V^{3/2}$ to the tunnel current. Note that $S(\tau)$ tends to a finite limit as $\tau\to\infty$, remaining always small. This allows us to expand the exponential in (5) in terms of this quantity and retain only the first term. The answer depends substantially on the ratio of the carrier diffusion coefficient and the magnetic-field diffusion coefficient $D^*=(4\pi\sigma)^{-1}c^2$, which depends in turn on the purity of the metal. In the "pure" limit ($k_Fl\gg c/e^2$, impurity density $\ll 1\%$) we have $D^*\ll D$, and in the "dirty" limit $D^*\gg D$. We ultimately obtain in these limiting cases

$$R_0\delta I/V=\begin{cases}\dfrac{4e^2}{3\pi^2\sigma}(eV/2D)^{1/2}\left[1+\left(\dfrac{D}{16D^*}\right)^{1/2}\ln(V_1/V)\right], & D^*\gg D,\\[2ex] \dfrac{2e^2}{3\pi c}(eV/2\pi\sigma)^{1/2}(\ln V_1/V), & D^*\ll D\end{cases}$$
$$V_1=4\pi\sigma ac/eD,\quad V_1\gg V. \tag{8}$$

For a given V, the anomalous part of the resistance decreases with increase of the bank-material purity:

$$R_0\partial I/V\propto l^{-3/2}\ \text{for}\ k_Fl\ll c/e^2,$$
$$R_0\partial I/V\propto l^{-1/2}\ \text{for}\ k_Fl\gg c/e^2.$$

The size of this correction makes it possible to determine it reliably using the characteristic dependence on V in the 1–10 mV range.

The zero-bias anomalies were calculated in Refs. 2 and 4 under the assumption that the tunneling rate was determined only by the density of states of the carriers in the junction banks. This assumption, which is undoubtedly correct within the scope of the one-particle scheme, does not take correct account of interaction effects (in the scheme described here this corresponds to a splitting of the mean values

$$\langle G^RG^L\rangle_A\to\langle G^R\rangle_A\langle G^L\rangle_A$$

and to neglect of the retardation). In the dirty limit $D^*\gg D$ the results (8) agree with the conclusions of Refs. 2 and 4, apart from a numerical factor and a small logarithmic term. In the opposite limiting case the difference between the results is more substantial: the dependence of the correction on

σ and on the diffusion coefficient changes. There are unfortunately no experimental data for this limiting case. We note only that in the experiment of Ref. 8, where one of the banks was pure and the other dirty, the observed dependence of δI and σ was intermediate between the pure and dirty cases.

We note in conclusion that the density-of-states correction investigated in Refs. 2 and 4 can be easily obtained within the framework of the scheme employed here by using the relation

$$\nu(\varepsilon, x) = \int_{-\beta}^{+\beta} d\tau\, e^{i\varepsilon_n\tau} \int d^2\mathbf{n} \langle \check{G}(\mathbf{x}, \mathbf{n}; \tau, \tau_1+\tau|A)\rangle_A \Bigg|_{i\omega_n \to eV+i0}$$

and the diffusion equation. No account is taken in this case, however, of the Fermi-liquid effects allowed for in Refs. 2 and 4. In all other respects the results are the same.

3. Let us complicate the geometry somewhat (Fig. 1b). The banks are now films of thickness $d \gg l$, separated by an insulating interlayer. Equations (6) must be solved with allowance for this circumstance. It turns out then that $\check{f}(0,\mathbf{x}) = 0$ on the outer boundary of the films and, accurate to $\sim|\omega|/\sigma$, the electromagnetic field does not leave the films. The frequency and wave-vector regions of interest are $\omega \ll D/d^2$, D^*/d^2 and $dq \ll 1$. In this limit we get for $S(\tau)$

$$S(\tau) = 8e^2 a \int dq\, qT \sum_{\omega_n} \sin^2(\omega_n\tau/2) \times \frac{1}{|\omega|+4\pi\sigma a\, dq^2} \frac{1}{|\omega_n|+Dq^2}. \quad (9)$$

Just as in (7), the first denominator describes here the propagation of electromagnetic excitations, and the second the carrier diffusion. The spreading of the spot is slower here than in (7), so that the degree of divergence of $S(\tau)$ increases. For long times $S(\tau) \propto \ln \tau$. This means formally that at sufficiently low V and T the IVC distortions become large, but this takes place in the region

$$V, T \lessdot \exp(-1/e^2 R_\square), \quad R_\square = (\sigma d)^{-1}.$$

This region cannot be reached in practice, since $e^2 R_\square \ll 1$ in the range in which our analysis is valid. We consider here therefore a small correction to the tunnel current, expanding the exponential in powers of $S(\tau)$. We have

$$R_0 \delta I/V = -\delta R/R_0 = \frac{e^2}{\pi^2 \sigma d} \ln\left(\frac{da}{r_D^2}\right) \ln(eV/\omega_{\text{cut}}). \quad (10)$$

The cutoff frequency ω_{cut} is here the lowest of the following three frequencies:

$$\omega_1 \sim 4\pi\sigma a/d, \quad \omega_2 \sim c^2/d^2 4\pi\sigma, \quad \omega_3 \sim 1/\tau_{Imp}.$$

This agrees in order of magnitude with the results of Refs. 3 and 4, the difference occurring in the logarithmic factors. It is possible that this is just the deviation noted in Ref. 4 from the experimental data.

4. We consider now the description of the Coulomb blockade of the tunneling within the framework of the scheme employed. We take into account the finite longitudinal dimensions of the films (Fig. 2a), assuming that the film linear dimension L is much larger than its thickness d. As noted in the preceding section, the electromagnetic field excited by the tunneling does not go outside the films. This

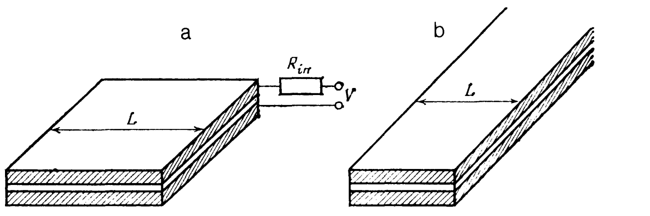

FIG. 2.

permits a trivial allowance for the finite dimensions in the case, e.g., of a square film, by replacing in (9) the integration with respect to $\mathbf{q}$ by summation over the discrete values $q_x = \pi n L^{-1}$ and $q_y = \pi m L^{-1}$:

$$S(\tau; x, y) = 16\pi e^2 a L^{-2} \sum_{\mathbf{q}} T \sum_{\omega_n} \frac{\sin^2(\omega_n\tau/2)\sin^2(q^x x)\sin^2(q^y y)}{(|\omega_n|+4\pi\sigma\, da\mathbf{q}^2)(|\omega_n|+D\mathbf{q}^2)}, \quad (11)$$

In this case $\exp(-S(\tau))$ in (5) is replaced by

$$\int_0^L dx \int_0^L dy \exp[-S(\tau; x, y)]/L^2.$$

It is now appropriate to state how the potential difference between the films is maintained. In the literature (e.g., in Ref. 9), there is some confusion due to the literal application of the concepts of network theory. It can be understood, for example from Ref. 10, that Coulomb blockade can be observed for dc voltage if the internal resistance R_{in} of the voltage source satisfies the condition $e^{-2} \ll R_{\text{in}} \ll R_0$, which we assume. The decisive contribution to (9) at long times will then be made by the zeroth harmonic; in this case $S(\tau) = e^2\tau/2C$, where C is the capacitance of the capacitor whose electrodes are the films. For the current we have

$$R_0 I(V) = \begin{cases} 0, & |V| < e/2C \\ V - \text{sign}\, Ve/2C, & |V| > e/2C \end{cases}$$

i.e., the tunneling is totally suppressed at $|V| < e/2C$. This result was first obtained in Ref. 6. Here we calculate the correction for the remaining harmonics in (9). Expanding the exponential in terms of the contribution of the nonzero harmonics in (5) and integrating, we obtain for $T = 0$

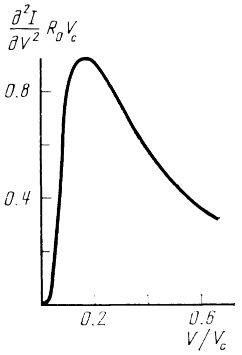

FIG. 3. Second derivative of the current with respect to voltage vs the voltage, for the case of a long tunnel junction.

$$\delta R/R_0=\begin{cases}\dfrac{e^2}{\pi\sigma d}\sum_q (\mathbf{q}^2L^2)^{-1}\ln\left(1+4\left(V-\dfrac{e}{2C}\right)^2\Big/D^2\mathbf{q}^4\right), & |V|>e/2C\\ 0, & |V|<e/2C\end{cases}$$

This correction to the resistance is small, $R_\square e^2$, and the characteristic scale of its variation is of the order D/L^2, the reciprocal of the time of carrier diffusion over a distance of the order of the film dimensions. For $V-e/2C \gg D/L^2$ we obtain $\delta R \propto \ln V$.

At still higher voltages $V \sim \sigma daL^{-2}$, i.e., V approximately equal to the reciprocal time of propagation of the electromagnetic excitation over the entire film, relation (12) is replaced by (10).

The corrections to the results of the "phenomenological" theory of the Coulomb blockade are thus small. It should be noted that this holds true only if the films have metallic conductivity and if the geometry ensures a long ($\gg C/e^2$) discharge time of the capacitor made up by the films.

5. Let us examine the result of lengthening the film in one direction (Fig. 2b). It is clear from (11) that once the film resistance reaches the quantum limit e^{-2} all the harmonics with wave vectors along the larger axis of the film become significant. For the infinite strip shown in Fig. 2b we have

$$S(\omega_n)=8e^2aL^{-1}\int\frac{dq}{(|\omega_n|+4\pi\sigma adq^2)(|\omega_n|+Dq^2)}.$$

For $T=0$ we have $S(\tau)=(eV_c\tau)^{1/2}$, where $V_c=e^3(\tilde{R}/\pi\tilde{C})$; $\tilde{R}$ and $\tilde{C}$ are the resistance and capacitance per unit length of the strip. For the second derivative of the current with respect to voltage we can obtain the analytic expression

$$\frac{\partial^2 I}{\partial V^2}R_0V_c=\frac{\exp(-V_c/4V)}{2[\pi(V/V_c)^3]^{1/2}}.$$

The shape of the IVC is shown in Fig. 3. The divergence of $S(\tau)$ as $\tau\to\infty$ is weaker than in the case of the Coulomb blockade, therefore the tunneling is not fully blocked anywhere, even at $T=0$, although it is substantially suppressed at $V\ll V_c$. The value of V_c in terms of the geometric dimensions and the mean free path can be estimated at $V_c\sim ea(k_F^2L^2dl)^{-1}$. For a, $l\sim 10$ Å and l, $d\sim 100$ Å we obtain $V_c\sim 10^{-6}$ V. The phenomenon will be observed at $T<eV_c\sim 0.1$ K.

The author takes pleasure in thanking K. K. Likharev and D. V. Averin for calling his attention to the questions discussed and for a number of helpful discussions.

APPENDIX

Let us examine relation (4). The electron Green's function integrated with respect to ξ satisfies the equation

$$\left\{-\left(\frac{\partial}{\partial\tau_1}+\frac{\partial}{\partial\tau_2}\right)+iv^\alpha(\mathbf{n})\frac{\partial}{\partial x^\alpha}-ev^\alpha[A^\alpha(\tau_1)-A^\alpha(\tau_2)]/c\right\}\times\tilde{G}(\tau_1,\tau_2;\mathbf{x},\mathbf{n})=\mathrm{St}\{\tilde{G}\},\tag{13}$$

$$\mathrm{St}\{\tilde{G}\}=\int_0^\beta d\tau_3(\Sigma(\tau_1,\tau_3;\mathbf{n},\mathbf{x})\tilde{G}(\tau_3,\tau_2;\mathbf{n},\mathbf{x})-\tilde{G}(\tau_1,\tau_3;\mathbf{n},\mathbf{x})\Sigma(\tau_3,\tau_2;\mathbf{n},\mathbf{x})).$$

Here St stands for scattering by impurities and Σ is the self-energy part given in the Born approximation by

$$\Sigma(\mathbf{n})=\int d^2n'\,w(\mathbf{n},\mathbf{n}')\tilde{G}(\mathbf{n}').$$

The conditions for the validity of (13) are: $eA/c\ll p_F$ and the characteristic spatial and "temporal" scales of the field variations must be r, $t\gg p_F^{-1}$, ε_F^{-1}.

We consider first the scales r, $t\ll 1$, and $\tau_{\rm Imp}$. In this case we can neglect the collision term and Eq. (13) can be easily integrated:

$$\tilde{G}(\tau_1,\tau_2;\mathbf{n},\mathbf{x})=\tilde{G}_0(\tau_1-\tau_2)\exp[i\varphi(\tau_1,\mathbf{n},\mathbf{x})-i\varphi(\tau_2,\mathbf{n},\mathbf{x})],$$
$$\left(\frac{\partial}{\partial\tau}-iv^\alpha(\mathbf{n})\frac{\partial}{\partial x^\alpha}\right)\varphi(\tau,\mathbf{n},\mathbf{x})=iev^\alpha(\mathbf{n})A^\alpha(\tau,x).\tag{14}$$

This result agrees with (4) if impurities are neglected. We now modify (14) in such a way that the first correction with respect to $\mathbf{A}$ to the Green's function $\tilde{G}$ is correctly described by (14). We obtain then the relation (4).

We now substitute $\tilde{G}$ in the form (14) in (13). We get

$$\int d^2n'\,w(\mathbf{n},\mathbf{n}')T\sum_{\omega_n}\operatorname{sign}\omega_n[\exp(i\tau_1\omega_n)-\exp(i\tau_2\omega_n)]\times[\varphi_{\omega_n}(\mathbf{n})-\varphi_{\omega_n}(\mathbf{n}')]$$
$$=\int d^2n'\,w(\mathbf{n},\mathbf{n}')\int_0^\beta d\tau_3\tilde{G}(\tau_1-\tau_3)\tilde{G}_0(\tau_3-\tau_2)\times\tilde{G}_0^{-1}(\tau_1-\tau_2)(\exp\{i[\varphi(\mathbf{n}',\tau_1)-\varphi(\mathbf{n},\tau_1)+\varphi(\mathbf{n},\tau_3)-\varphi(\mathbf{n}',\tau_3)]\}-\exp\{i[\varphi(\mathbf{n},\tau_2)-\varphi(\mathbf{n}',\tau_2)+\varphi(\mathbf{n}',\tau_3)-\varphi(\mathbf{n},\tau_3)]\}).$$

For simplicity, we have taken Σ here in the Born approximation. This equality is valid in the limit $\varphi(\mathbf{n}')-\varphi(\mathbf{n})\ll 1$. The isotropic part φ can also be of order of unity and larger.

This situation can be realized when r, $t\ll l$, $\tau_{\rm Imp}$. The isotropic part of the Green's function is then much larger than the anisotropic, and the diffusion approximation used in the problems treated in the text is valid.

Relation (4), which is valid for $r,t\ll l,\tau_{\rm Imp}$ for arbitrary fields $\mathbf{A}$, is thus valid over large scales if $eAl/c\ll 1$. This is equivalent to the condition that the electron distribution be weakly distorted by low-frequency fluctuation-type electromagnetic perturbations, and is well satisfied in metals.

What is the physical meaning of retaining in (4) the terms nonlinear in $\mathbf{A}$? After averaging over the configurations of the field $\mathbf{A}$, the expressions for the Green's function should describe the processes of emission and absorption of electromagnetic photons and exchange of virtual photons as the electrons move. Inclusion in (4) of terms of finite order in $\mathbf{A}$ means allowance for processes in which only a finite number of photons participate. The use of expression (4) for $\tilde{G}$ corresponds to allowance for processes with an arbitrarily large number of photons, and under the assumption that each act of emission, absorption, and exchange is independent, which corresponds to a semiclassical treatment of interaction with electromagnetic excitations.

[1]B. L. Altshuler, in: *Electron–electron Interactions in Disordered Systems*, A. L. Efros, ed., North-Holland, 1985, p. 190.
[2]B. L. Al'tshuler and A. G. Aronov, Zh. Eksp. Teor. Fiz. **77**, 2028 (1979) [Sov. Phys. JETP **50**, 968 (1979)].
[3]B. L. Altshuler and P. A. Lee, Phys. Rev. Lett. **44**, 1288 (1980).
[4]B. L. Al'tshuler, A. G. Aronov, and A. Yu. Zyuzin, Zh. Eksp. Teor. Fiz. **86**, 709 (1984) [Sov. Phys. JETP **59**, 415 (1984).
[5]K. K. Likharev, IBM J. Res. Develop. **32**, 144 (1988).
[6]H. R. Zeller and I. A. Giaever, Phys. Rev. **181**, 789 (1969).
[7]N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Interscience, 1951.
[8]A. Dynes and B. Garno, Phys. Rev. Lett. **46**, 137 (1981).
[9]K. Mullen, E. Ben-Jacob, and R. Jaklevic, Phys. Rev. B37, 104 (1988).
[10]D. V. Averin and K. K. Likharev, Zh. Eksp. Teor. Fiz. **90**, 733 (1986) [Sov. Pys. JETP **63**, 427 (1986)].

Translated by J. G. Adashko