Quantum fluctuations of the charge near the Coulomb-blockade threshold

D. S. Golubev and A. D. Zaikin
I. E. Tamm Department of Theoretical Physics, P. N. Lebedev Physics Institute, Leninski prospect 53, Moscow 117924, Russia
(Received 10 December 1993; revised manuscript received 1 April 1994)

In the vicinity of the Coulomb-blockade threshold, quantum fluctuations of the charge strongly influence the energy spectrum of small metallic junctions. We develop a nonperturbative calculation of the junction ground-state energy and evaluate the junction charge in the limit of a small junction conductance. Close to the Coulomb-blockade threshold this charge is effectively screened due to intensive virtual electron tunneling. We also describe level splitting close to the edges of the Brillouin zone and show that the electron tunneling rate can be substantially suppressed due to quantum fluctuations of the charge. As a result the junction current-voltage characteristic becomes non-Ohmic near the Coulomb gap.

I. INTRODUCTION

Charging effects in mesoscopic tunnel junctions have attracted much attention during the last several years (see, e.g., Refs. 1 and 2 for a review). A convenient experimental realization of these effects can be provided in the so-called electron box structure.[3] It consists of a small metallic island connected via a tunnel junction with a capacitance C_J to a lead electrode and via a gate capacitance C_g to a voltage source V_g. The charge of an island is quantized in units of an electron charge e. Therefore, the Coulomb energy of the system depends on the number of electrons on an island n and the externally applied charge $Q_x = C_gV_g$

$$E_0^{(0)}(n,Q_x) = \frac{(ne-Q_x)^2}{2C}, \tag{1}$$

$C = C_j + C_g$ is the total capacitance. At sufficiently low temperatures $T \ll E_c = e^2/2C$ and $ne - e/2 < Q_x < ne + e/2$ Coulomb interaction suppresses single electron tunneling across the junction.[1] Then the equilibrium number of electrons on an island is a step function of Q_x. Within the interval $ne - e/2 < Q_x < ne + e/2$ this number is equal to n. At $Q_x = ne + e/2$ an extra electron tunnels to an island and the picture repeats e periodically in Q_x.

Quantum fluctuations of the junction charge, or, in other words, virtual electron tunneling across the junction, can substantially modify this simple classical picture. Quantum effects become particularly important close to the points $Q_x = ne+e/2$ where the energies of the states n and $n+1$ are nearly equal. It was pointed out in Ref. 4 that quantum fluctuations of the charge lead to the effective renormalization of the junction capacitance and to the flattening of the band in the vicinity of the points $Q_x = \pm e/2$ (from now on we restrict our attention to the case $n = 0$ and $-e/2 < Q_x < e/2$). More recent analysis of the problem[5,6] demonstrated that in the limit of large tunneling conductance $\alpha_t \gg 1$ ($\alpha_t = R_q/R_t$, $R_q = \pi/2e^2$, R_t is the junction tunneling resistance) for small Q_x quantum fluctuations lead to the effective renormalization of the junction capacitance $C_{\rm eff} \sim C/\alpha_t^2 \exp(2\alpha_t)$, whereas for the values Q_x close to $e/2$ such fluctuations destroy the Coulomb blockade and the band becomes flat. As a result the expectation value for the junction charge operator becomes smaller than $|Q_x|$, i.e., an external charge turns out to be (partially) screened due to intensive virtual electron tunneling across the junction. A clear experimental indication of the presence of this effect in the electron box structure has been reported in Ref. 3.

In the limit of a small junction conductance $\alpha_t \ll 1$ one can proceed perturbatively in α_t and calculate the first order correction to the ground-state energy of the system $E_0(Q_x)$ (see, e.g., Refs. 2 and 4). This naive perturbation theory fails in the vicinity of the points $Q_x = \pm e/2$ because the first order correction to the expectation value of the charge operator

$$\langle Q\rangle = C\frac{dE_0}{dQ_x} \tag{2}$$

diverges logarithmically at $Q_x \to \pm e/2$

$$\langle Q\rangle = Q_x + \frac{\alpha_t e}{\pi^2}\ln\left(\frac{e/2-Q_x}{e/2+Q_x}\right). \tag{3}$$

For such values of Q_x higher order terms of the perturbation expansion become important. The physical reason for that is clear: for $Q_x \to e/2$ the energy states $Q_x^2/2C$ and $(Q_x - e)^2/2C$ are nearly degenerate. Therefore, strong tunneling between these two states occurs even for very small α_t and a naive perturbation theory in α_t cannot be sufficient.

Note that the origin of the logarithmic divergence (4) at $Q_x = \pm e/2$ is similar to that discussed in the familiar Kondo problem. This useful analogy has been established by Matveev[7] who calculated an expectation value for the charge operator (Q) making use of a renormalization group approach developed for the anisotropic Kondo problem. Note that the results of Ref. 7 have been obtained under certain assumptions about the particular form of the tunneling matrix element. Therefore, *a priori* it is not completely clear how far one can go with this analogy in the description of metallic tunnel junctions. In this paper, we develop a general analysis of the problem based on the microscopic effective action for a tunnel junction.[2,8] The results of our analysis expressed in terms of the junction conductance are insensitive to the details of the model describing electron tunneling across the junction.

In Sec. II we evaluate the higher order terms of a perturbation expansion in powers of α_t. Then we develop a diagram technique which allows one to sum up the perturbation series and to formulate the closed system of equations for the partition function and the self-energy of our problem. In Sec. III we calculate the energy spectrum of the tunnel junction in the limit of small α_t. We also evaluate the electron tunneling rate close to the Coulomb-blockade threshold and study the effect of quantum fluctuations of the charge on the current-voltage characteristic of a tunnel junction. A brief discussion of our results is presented in Sec. IV. Some details of a regularization procedure are outlined in the Appendix.

II. PERTURBATION THEORY AND DIAGRAM TECHNIQUE

Let us consider a tunnel junction between two normal metals with an externally controlled charge Q_x. As we already discussed, this physical situation can be easily realized, e.g., in the simple electron box configuration.[3] The ground partition function for this system has the form[2]

$$\mathcal{Z}(\beta, Q_x) = \sum_m \int d\varphi_0 \int_{\varphi_0}^{\varphi_0+4\pi m} D\varphi \times \exp\left(\frac{2\pi m i Q_x}{e} - S[\varphi]\right). \tag{4}$$

The term $S[\varphi]$ represents the action for a tunnel junction[2,8]

$$S[\varphi] = \int d\tau \frac{C}{2}\left(\frac{\dot\varphi}{2e}\right)^2 - \int d\tau \int d\tau' \alpha(\tau-\tau') \times \cos\left(\frac{\varphi(\tau)-\varphi(\tau')}{2}\right), \tag{5}$$

where

$$\alpha(\tau-\tau') = \frac{\alpha_t T^2}{\sin^2[\pi T(\tau-\tau')]}. \tag{6}$$

The phase $\varphi(\tau)$ is linked to the voltage across the junction by the relation $V(\tau) = \dot\varphi(\tau)/2e$. If we do not take electron tunneling into account (i.e., put $\alpha_t = 0$) the junction energy would be that of a capacitor $E_0^{(0)}(Q_x) = Q_x^2/2C$. However, for any nonzero α_t, electrons can tunnel across the junction and, therefore, the system "spends" part of the time in the charge states $Q \pm e$, $Q \pm 2e$, etc. As a result the value $\langle Q\rangle$ will deviate from Q_x and the junction ground-state energy $E_0(Q_x) = -T \ln \mathcal{Z}(Q_x)|_{T\to 0}$ will be renormalized. Proceeding perturbatively one can expand the expression (4) in powers of α_t and recover corrections to the charging energy (1) due to virtual electron tunneling. Then making use of (2) and keeping only the first two terms of a perturbation expansion in α_t we get

$$\langle Q\rangle = Q_x C/C_{\text{eff}}, \tag{7}$$

$$C/C_{\text{eff}} = 1 - 4g + 2g^2\Big(\frac{7\pi^2}{12} + \frac{29\ln 2}{9} - \frac{16}{3} - 4\ln 2 \ln 3 + 2\text{Li}_2(1/4)\Big) + \cdots \tag{8}$$

for $Q_x \ll e$ and

$$\langle Q\rangle = Q_x + ge\ln[a(Q_x)] + 2g^2 e(\{\ln[a(Q_x)]\}^2 + 3\ln[a(Q_x)]) + \cdots \tag{9}$$

for $a(Q_x) = 1/2 - Q_x/e \ll 1$. Here $\text{Li}_2(x)$ is the dilogarithm function and $g = \alpha_t/\pi^2$. We see that for small Q_x virtual electron tunneling yields a capacitance renormalization (8) whereas for the values Q_x close to $e/2$ the situation turns out to be more complicated because of the presence of logarithmically diverging terms in the perturbation expansion (9). It is straightforward to show that the third order correction to $\langle Q\rangle$ contains the main logarithmic contribution $4g^3 e[\ln a(Q_x)]^3$ as well as lower powers of $\ln a$. Analogously the nth order term of the perturbation expansion for $\langle Q\rangle$ should contain all powers of $\ln a$ from one to n. Below we shall develop a diagram technique which allows one to sum up our perturbation expansion in all orders in α_t and to evaluate the energy spectrum of our problem.

To proceed it is convenient to introduce the density matrix

$$\rho(\tau,\varphi) = \int_0^{\varphi} D\varphi \exp\{-S[\varphi]\}. \tag{10}$$

Making use of a Poissons resummation theorem after a trivial algebra one can rewrite the expression for $\mathcal{Z}(Q_x)$ (4) in the form (see also Ref. 2)

$$\mathcal{Z} = \frac{A}{4\pi}\sum Z(Q_x - en), \quad A = \int d\varphi, \tag{11}$$

where—in contrast to $\mathcal{Z}(Q_x)$—the function

$$Z(Q_x) = \int_{-\infty}^{\infty} \rho(\varphi)\exp(iQ_x\varphi/2e)d\varphi \tag{12}$$

is not e periodic in Q_x. Let us furthermore define the function $Z(\tau, Q_x)$ which coincides with $Z(Q_x)$ (12) at $\tau = \beta$. The function $Z(\tau, Q_x)$ will play the role of a propagator in a diagram technique developed below.

Expanding $Z(\tau, Q_x)$ in powers of α_t we get

$$Z(\tau, Q_x) = \exp\left(-\frac{Q_x^2}{2C}\tau\right) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^{\tau} \cdots \int_0^{\tau} d\tau_1 d\tau_1' \cdots d\tau_n d\tau_n' \alpha(\tau_1 - \tau_1') \cdots \alpha(\tau_n - \tau_n')$$
$$\times \exp\left(-\frac{1}{2C}\int_0^{\tau} dt \left\{ Q_x - e\sum_{j=1}^{n}[\theta(t-\tau_j) - \theta(t-\tau_j')]\right\}^2\right), \quad (13)$$

where $\theta(x)$ is the step function and $0 < \tau_j, \tau_j' < \tau$ for all j. Let us order the times τ_j, τ_j' and denote them according to their ordering: $\tau_1 < \tau_2 < \cdots < \tau_{2n-1} < \tau_{2n}$. Then we have

$$Z(\tau, Q_x) = \exp\left(-\frac{Q_x^2}{2C}\tau\right) + \sum_{n=1}^{\infty} \sum_{\epsilon_r = \pm 1} \sum_{\text{pairs}} \int_0^{\tau} d\tau_{2n} \int_0^{\tau_{2n}} d\tau_{2n-1} \cdots \int_0^{\tau_2} d\tau_1 \prod_{n=1}^{n} \alpha(\tau_{i_k} - \tau_{j_k})$$
$$\times \exp\left\{-\frac{1}{2C}\int_0^{\tau} dt \left[Q_x - e\sum_{r=1}^{n} \epsilon_r \theta(t-\tau_r)\right]^2\right\}. \quad (14)$$

The sum $\sum_{\epsilon_r=\pm 1}$ is taken over neutral charge configurations $\sum_r \epsilon_r = 0$ whereas the sum $\sum_{\text{pairs}}$ is taken over all possible pairs i_k and j_k with $\epsilon_{i_k} = 1$, $\epsilon_{j_k} = -1$. It is convenient to express the expansion (14) graphically (Fig. 1). Each term of this expansion corresponds to an even number of successive electron tunneling events (stepwise function of Fig. 1), logarithmic interaction between each pair of these events τ_{i_k} and τ_{j_k} introduced by the kernel $\alpha(\tau_{i_k} - \tau_{j_k})$ is depicted by a dashed line. We shall distinguish between connected and disconnected diagrams. Connected diagrams (e.g., the first diagram in Fig. 1) cannot be divided into two parts without the crossing of at least one dashed line. Disconnected diagrams (e.g., the second diagram in Fig. 1) have parts without dashed lines. The total charge corresponding to these parts is always equal to Q_x because connected blocks are neutral.

Analogously to the standard Feynman diagram technique we introduce the sum of all connected diagrams $\Sigma(\tau, Q_x)$ and then rewrite the diagram series in terms of a Dyson equation (Fig. 2)

$$Z(\tau, Q_x) = \exp(-Q_x^2\tau/2C) + \int_0^{\tau} d\tau_2 \int_0^{\tau_2} d\tau_1$$
$$\times Z(\tau - \tau_2)\Sigma(\tau_2 - \tau_1)\exp(-Q_x^2\tau_1/2C). \quad (15)$$

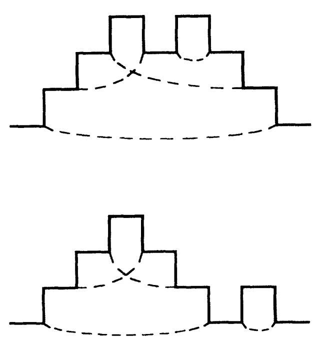

FIG. 1. Typical diagrams (one connected and one disconnected) which describe terms of a perturbation expansion for Z (14). Dashed lines correspond to logarithmic interaction between pairs of electron tunneling events.

Making use of the Laplace transformation we find

$$Z_p(Q_x) = \int_0^{\infty} d\tau \exp(-p\tau) Z(\tau, Q_x)$$
$$= 1/[p + Q_x^2/2C - \Sigma_p(Q_x)], \quad (16)$$

p is the Laplace variable. Below we shall use this equation to evaluate the energy spectrum of a tunnel junction.

The self-energy Σ in the Dyson Eqs. (15) and (16) can be represented in terms of the diagram series shown in Fig. 3. These series allow one to express Σ via $Z(\tau, Q_x)$ and to obtain the closed system of equations for Z. Equivalently one can rewrite the Dyson equation introducing the vertex part $\Gamma(p, Q_x; p', Q_x')$

$$\frac{1}{Z_p(Q_x)} = p + Q_x^2/2C + \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dp' \alpha(p - p')$$
$$\times \{\Gamma(p, Q_x; p', Q_x + e) Z_{p'}(Q_x + e)$$
$$+ \Gamma(p, Q_x; p', Q_x - e) Z_{p'}(Q_x - e)\} \quad (17)$$

and express Γ in terms of the diagram series presented in Fig. 4. Note that this diagram expansion is very similar to that discussed in the zero charge problem of quantum electrodynamics.[9] To recover this analogy one should identify solid and dashed lines, respectively, with electron and photon propagators. The only important difference between these two problems is that in contrast to Ref. 9 there is no "vacuum polarization" in our problem and the expression for the "photon propagator" $\alpha(\tau)$ is fixed.

III. ENERGY SPECTRUM AND TUNNELING RATE

The function $Z_p(Q_x)$ contains complete information about the energy spectrum of our problem. For $-e/2 <$

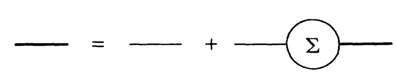

FIG. 2. Graphical representation of the Dyson Eq. (15).

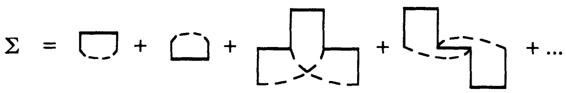

FIG. 3. Diagrammatic representation of a perturbation expansion for the self-energy Σ.

$Q_x < e/2$ this function defines the junction ground-state energy $E_0(Q_x) = -T \ln Z(Q_x)|_{T\to 0}$. From the inverse Laplace transformation

$$Z(\tau, Q_x) = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} dp Z_p(Q_x)\exp(p\tau) \qquad (18)$$

one can easily find $Z(\tau, Q_x) \sim \exp(p_0\tau)$ for $\tau \to \infty$. Here p_0 is the pole of $Z_p(Q_x)$ in the complex plane p with the largest value $\mathrm{Re}\, p_0$ (all poles of the function Z_p are in the half plane $\mathrm{Re}\, p < 0$). Therefore, the ground-state energy $E_0(Q_x)$ coincides with the smallest positive solution of the equation

$$Z_p^{-1}(Q_x) = 0 \qquad (19)$$

or, equivalently,

$$E_0(Q_x) = \frac{Q_x^2}{2C} - \Sigma_p(Q_x)|_{p=-E_0(Q_x)}. \qquad (20)$$

Making use of the analytic properties of the function $Z_p(Q_x)$ one can also obtain information about higher energy states. Indeed, provided $Z_p(Q_x)$ is an analytic function of the parameters p and Q_x the function $E_0(Q_x)$ is also an analytic function of Q_x. This allows one to recover the energy of excited states by means of an analytic continuation of the function $Z_p(Q_x)$ from the interval $|Q_x| < e/2$ to $|Q_x| > e/2$.

Before proceeding with the calculation of the energy spectrum let us note that the terms of our perturbation expansion for Z in powers of α_t formally diverge in the high frequency limit. These divergencies are induced by the behavior of the kernel $\alpha_t(\tau)$ (6) at $\tau \to 0$ and have no physical meaning. A standard way to treat this problem is to introduce a high frequency cutoff ω_c. Then one can solve Eq. (20) for $|p| \ll \omega_c$ and obtain the results which in general depend on the parameter ω_c. Therefore, this approach contains a certain ambiguity related to a particular choice of the high frequency cutoff procedure. Here we regularize our problem by making a shift of the effective action (5) by the φ-independent term

$$S[\varphi] \to S[\varphi] + \int d\tau \int d\tau' \alpha(\tau - \tau')\exp(-E_c|\tau - \tau'|). \qquad (21)$$

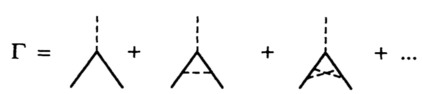

FIG. 4. Diagrammatic representation of a perturbation expansion for the vertex part Γ.

This procedure implies only an energy shift by a constant value and does not influence any physical quantity which we calculate below. Expanding (4) in powers of α_t with the aid of (21) it is straightforward to check (see the Appendix) that no divergent terms appear for $\tau - \tau' \to 0$ and our theory remains finite.

Making use of a regularization (21) one can rewrite the Dyson Eq. (17) in the form

$$\frac{1}{Z_p(Q_x)} = p + \frac{Q_x^2}{2C} - g\int_p^{+\infty} dp'(p'-p) \times[\Gamma(p, Q_x; p', Q_x - e) \times Z_{p'}(Q_x - e) + \Gamma(p, Q_x; p', Q_x + e) \times Z_{p'}(Q_x + e) - 2\Gamma(p, Q_x; p' + E_c, Q_x) \times Z_{p'+E_c}(Q_x)]. \qquad (22)$$

Equations (15) and (16) can be changed accordingly. To solve these equations in the limit of a small junction conductance we apply a perturbation theory in α_t and take into account only two first diagrams in the diagram series for the self-energy Σ (Fig. 3). This approximation corresponds to a summation of a subsequence of diagrams with noncrossing dashed lines. It is equivalent to the well-known approximation of main logarithms in the Kondo problem or in the zero charge problem. Analogously, only the first diagram in the series for the vertex part Γ (Fig. 4) has to be taken into account. Within the framework of this approximation the self-energy Σ and the vertex part Γ are equal to

$$\Sigma(\tau, Q_x) = \alpha(\tau)[Z(\tau, Q_x - e) + Z(\tau, Q_x + e) -2\exp(-E_c\tau)Z(\tau, Q_x)], \quad \Gamma = 1. \qquad (23)$$

Then with the aid of (23) one can rewrite the equation for Z (22) in the form

$$\frac{1}{Z_p(Q_x)} = p + \frac{Q_x^2}{2C} - g\int_p^{+\infty} dp'(p'-p)[Z_{p'}(Q_x - e) +Z_{p'}(Q_x + e) - 2Z_{p'+E_c}(Q_x)]. \qquad (24)$$

Taking the second derivative of (24) with respect to p we get

$$\frac{d^2}{dp^2}\frac{1}{Z_p(Q_x)} = -gZ_p(Q_x - e) - gZ_p(Q_x + e) +2gZ_{p+E_c}(Q_x). \qquad (25)$$

Our analysis can be simplified further for values Q_x close to $e/2$. In this limit diagrams describing tunneling between the charge states Q_x and $Q - e$ give the main contribution and one can reduce the problem to the effective two-state problem. Taking the Laplace parameter p close to the poles of the functions $Z_p(Q_x)$ and $Z_p(Q_x - e)$ for $|e/2 - Q_x| \ll e$ and $g \ll 1$ one can estimate the functions Z_p in the right-hand side of (23) as

$$Z_p^{-1}(Q_x - e) \sim \frac{(Q_x - e)^2}{2C} - \frac{Q_x^2}{2C} \equiv \Delta(Q_x),$$

$$Z_p^{-1}(Q_x + e) \sim Z_{p+E_c}^{-1}(Q_x) \sim E_c.$$

Then for $p + Q_x^2/2C \ll E_c$ and $\Delta(Q_x) \ll E_c$ Eq. (25) reduces to

$$\ddot{x}(p) = -g/y(p), \quad \ddot{y}(p) = -g/x(p), \tag{26}$$

where we define $x(p) = Z_p^{-1}(Q_x)$, $y(p) = Z_p^{-1}(Q_x - e)$ and denote the derivatives over p by overdots. To obtain initial conditions for these equations we rewrite the first of Eq. (26) in the form of an integral equation

$$x(p) = -g \int_p^0 dp'(p' - p)y^{-1}(p') + x_0 + \dot{x}_0 p \tag{27}$$

and compare it with (24) for $p + Q_x^2/2C \ll E_c$, i.e., in the vicinity of the pole of the function $Z_p(Q_x)$. As a result we arrive at the initial conditions:

$$x_0 = \frac{Q_x^2}{2C} - g \int_0^\infty dp' p' Z_{p'}(Q_x - e) - g \int_{-Q_x^2/2C}^\infty dp' p' [Z_{p'}(Q_x + e) - 2Z_{p'+E_c}(Q_x)], \tag{28}$$

$$\dot{x}_0 = 1 + g \int_0^\infty dp' Z_{p'}(Q_x - e) - g \int_{-Q_x^2/2C}^\infty dp' \times [Z_{p'}(Q_x + e) - 2Z_{p'+E_c}(Q_x)]. \tag{29}$$

Initial conditions for y_0 and $\dot{y}_0$ can be derived analogously. As integration over p' in (28) and (29) runs far from the poles of Z_p with a sufficient accuracy one can put $Z_p(Q_x) = (p + Q_x^2/2C)^{-1}$. Then for $Q_x \simeq e/2$ we immediately get

$$x_0 = Q_x^2/2C + gE_c[(1 - 7\ln 2)/4 + (1/4 + 5\ln 2)a(Q_x)], \tag{30}$$

$$y_0 = (Q_x - e)^2/2C + gE_c[(1 - 7\ln 2)/4 - (1/4 + 5\ln 2)a(Q_x)], \tag{31}$$

$$\dot{x}_0 + 3ga(Q_x) = \dot{y}_0 - 3ga(Q_x) = 1 + g\ln 2. \tag{32}$$

To solve Eqs. (26) with the initial conditions (30)–(32) it is convenient to introduce the function $u(p) = x(p)y(p)$. This function obeys the equation

$$\ddot{u} = 2\dot{x}_0\dot{y}_0 - 2g - 2g\ln(u/x_0y_0), \tag{33}$$

which has a solution

$$p = \int_{x_0y_0}^u du[(x_0\dot{y}_0 - \dot{x}_0y_0)^2 + 4\dot{x}_0\dot{y}_0 - 4gu\ln(u/x_0y_0)]^{-1/2}. \tag{34}$$

Combining this solution with Eq. (20) we arrive at the final result for the junction ground-state energy $E_0(Q_x)$

$$E_0(Q_x) = \frac{1}{2}\left(\frac{x_0y_0}{\dot{x}_0\dot{y}_0}\right)^{1/2} \int_0^1 dt \left(\frac{(x_0\dot{y}_0 - \dot{x}_0y_0)^2}{4x_0y_0\dot{x}_0\dot{y}_0} + t - \frac{gt\ln t}{\dot{x}_0\dot{y}_0}\right)^{-1/2}. \tag{35}$$

Within the accuracy of our calculation this expression yields

$$E_0(Q_x) = E_0(e/2) - E_c a(Q_x)\{(1 - gb)/L(Q_x) + g(b - 2)/[L(Q_x)]^2\}, \tag{36}$$

where $E_0(e/2) = E_c(1 - 8g\ln 2)/4$, $b = 1 + 4\ln 2$,

$$L(Q_x) = 1 - g\{\ln[a(Q_x)] + \ln[a_r(Q_x)]\} \tag{37}$$

and $a_r(Q_x) = [E_0(e/2) - E_0(Q_x)]/E_c$. Taking a derivative of (36) with respect to Q_x we get

$$\langle Q \rangle = (e/2)\{(1 - gb)/L(Q_x) + gb/[L(Q_x)]^2\}. \tag{38}$$

This dependence is presented on Fig. 5 for different values of g. In the main approximation the result (37) and (38) coincides with that obtained by Matveev[7] in the limit of a large number of channels. It is interesting to point out that the result (37) and (38) obtained by a summation of noncrossing two-state problem diagrams in some sense goes beyond the approximation of main logarithms. The reason for that lies in the fact that diagrams with crossing dashed lines give a small (of order g^2) contribution to the self-energy and, therefore, it is possible to recover the next order term in $g/(1 - 2g\ln a)$ in the expression for $\langle Q \rangle$. On the other hand, the approximation of noncrossing diagrams is obviously insufficient to reproduce the correct prefactor for the term $g^2 \ln a$ in the expansion (9) because diagrams describing tunneling to higher energy states also contribute to this term. This effect might lead to a small unimportant renormalization of g in Eqs. (36)–(38) which we do not consider here.

As we already discussed, the energy of excited states can be obtained by means of an analytic continuation of $E_0(Q_x)$ to the values $|Q_x| > e/2$. Here we are interested in the expression for the energy of the first excited state $E_1(Q_x)$ in the vicinity of the point $Q_x = e/2$. The analytic continuation procedure is straightforward and yields

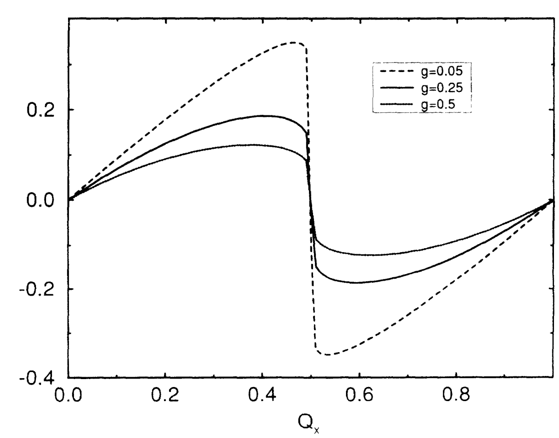

FIG. 5. Dependence of the expectation value for the charge operator $\langle Q \rangle = dE_0/dQ_x$ on the external charge Q_x calculated from the Eq. (35) for different values of g. Close to the Coulomb-blockade threshold $Q_x = e/2$ the value $\langle Q \rangle$ is defined by the Eqs. (37) and (38).

$$E_1(Q_x) = E_0(e/2) + E_c a(Q_x)\left(\frac{1-gb}{L(Q_x)-2gi\pi} + \frac{g(b-2)}{[L(Q_x)-2gi\pi]^2}\right). \quad (39)$$

Combining this result with Eq. (36) we immediately obtain the expression for the level splitting $\Delta_{10} = \mathrm{Re}E_1 - E_0$ in the vicinity of the point $Q_x = e/2$

$$\Delta_{10}(Q_x) = 2[E_0(e/2) - E_0(Q_x)]\{1 - 2\pi^2 g^2/[L(Q_x)]^2\}. \quad (40)$$

At $Q_x \to e/2$ in the main approximation this result reduces to a simple formula

$$\Delta_{10}(Q_x) = -E_c a(Q_x)/g \ln a(Q_x). \quad (41)$$

The behavior of the two lowest energy bands $E_0(Q_x)$ and $E_1(Q_x)$ in the vicinity of the Coulomb-blockade threshold $Q_x = e/2$ is depicted in Fig. 6 for different values of g. We see that for larger values of g both the slope of the bands and the value $\Delta_{10}(Q_x)$ become smaller but remain nonzero except for the point $Q_x = e/2$. We also point out an asymmetry of the bands which increases with increasing g.

The presence of an imaginary part in the expression for $E_1(Q_x)$ (39) indicates an instability of this state with respect to tunneling (decay) to the ground state with the energy $E_0(Q_x)$. For $Q_x < e/2$ this process corresponds to the tunneling of one electron from the charge state $Q_x - e$ to the state Q_x. The rate of such a tunneling process $\Gamma(Q_x)$ is defined by a well-known formula

$$\Gamma(Q_x) = 2\mathrm{Im}E(Q_x). \quad (42)$$

Not very close to the point $Q_x = e/2$, this rate was calculated perturbatively in α_t.[1,2] Our analysis allows one to find a nonperturbative expression for $\Gamma(Q_x)$ which includes the effect of strong quantum fluctuations in the vicinity of the point $Q_x = e/2$. Making use of (39) and (42) in the main approximation we get

$$\Gamma(Q_x) = \frac{2\pi g e}{C}\theta(Q_x - e/2)\frac{Q_x - e/2}{[1 - 2g\ln|a(Q_x)|]^2}. \quad (43)$$

This expression shows that intensive virtual electron tunneling close to the Coulomb-blockade threshold $Q_x \to e/2$ not only leads to the band flattening (36) and charge screening (38) but also reduces the single electron tunneling rate by a factor of order $[g\ln|a(Q_x)|]^{-2}$ below the standard perturbative value.[1,2] To calculate the current-voltage characteristic close to the point $Q_x = e/2$ it is necessary to combine both effects. Defining an externally applied voltage $V_x = Q_x/C$ we arrive at the expression for the current across the junction

$$I(V_x) = 2\langle Q(V_x)\rangle\Gamma(V_x). \quad (44)$$

In combination with the results (37), (38), and (43) at $T \to 0$ this expression yields

$$I(V_x) = \frac{V_x - e/2C}{R_t[1 - 2g\ln(CV_x/e - 1/2)]^3}, \quad (45)$$

for $V_x > e/2C$ and $I(V_x) = 0$ for $V_x < e/2C$. The result (45) demonstrates (see also Fig. 7) that the effect of quantum fluctuations on the current-voltage characteristic of a small tunnel junction is important for $g\ln(CV_x/e - 1/2) \gtrsim 1$ leading to a smearing of the jump in the classical conductance[1] $G(Q_x = e/2 + 0) - G(Q_x = e/2 - 0) = 1/R_t$ at $Q_x = e/2$.

At first sight the result (45) might look counter intuitive: it shows that at $V_x > e/2C$ the junction conductance decreases with increasing α_t. The reason for such a behavior lies in the effect of charge screening. Due to this effect close to the Coulomb-blockade threshold (a)

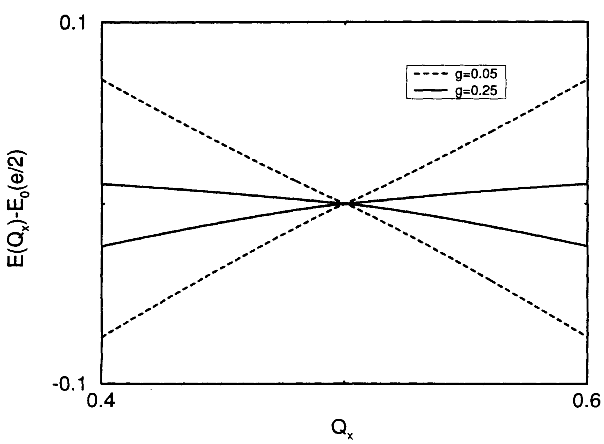

FIG. 6. The behavior of the two lowest bands $E_0(Q_x)$ (lower curves) and $E_1(Q_x)$ (upper curves) in the vicinity of the Coulomb-blockade threshold $Q_x = e/2$ for different values of g.

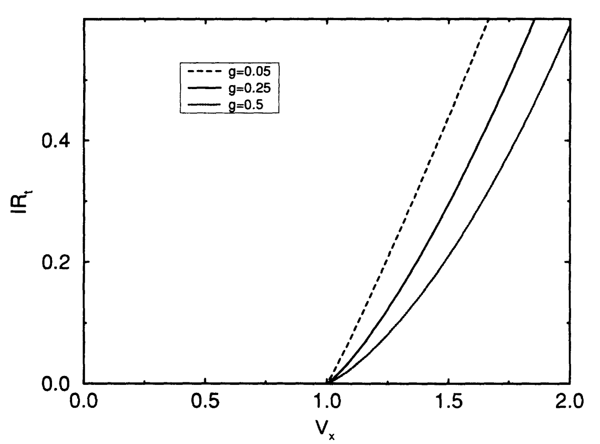

FIG. 7. The current-voltage characteristic of a small conductance tunnel junction (45) in the vicinity of the Coulomb-blockade threshold. At $T = 0$ the current I is zero below the Coulomb gap $V_x < e/2C$. Above this gap the effect of quantum fluctuations on the current is strong in the region $g\ln(CV_x/e - 1/2) > 1$ where I substantially deviates from its classical value (Ref. 1).

the junction "feels" a smaller value of the external charge than its bare value $Q_x = V_x C$ (and thus both the electron tunneling rate and the current across the junction decrease) and (b) each tunneling event changes the junction charge by the quantity $2\langle Q\rangle$ smaller than e. Both features become more pronounced for larger values of α_t leading to the dependence (45).

At the same time, the value α_t should be still small enough for the result (45) to remain valid. For relatively large α_t (typically for $\alpha_t \gtrsim 2$–3—see below) crossing diagrams become important and one might expect substantial changes in the system behavior.[5,6] Furthermore, even for small α_t the result (45) applies only for a physical situation of a fixed external charge across the junction. This situation can be achieved, e.g., for several junctions connected in series. To see the charging effects in a single junction one should essentially decouple this junction from a voltage source by a large external impedance (see, e.g., Refs. 1 and 2). The latter situation is better described by the current biased model. Making use of the results (36)–(43) and proceeding within the framework of a simple master equation analysis[1] we easily get for $I_x \ll e/R_t C$

$$V = \sqrt{\frac{\pi I_x R_t e}{2C}}\left(1 - \frac{2g}{1 - \ln(I_x R_t C/e)}\right). \qquad (46)$$

Here V is the average voltage across the junction and I_x is a constant current bias. In contrast to the voltage biased case (45) two quantum effects—suppression of the electron tunneling rate and electron charge screening—act in the "opposite directions" and nearly compensate each other leading only to a small quantum correction (46) to the classical result[1] in the limit of small $g \ll 1$.

Combining the results presented here with the analysis developed in Ref. 6 one can also obtain information about the linear transport properties of the junction at finite frequencies and/or temperature. For example, one can consider the configuration with the external voltage source V_x switched to the junction via an Ohmic resistor R_s and define the linear conductance $G_{\rm tot}(\omega)$ of the system "tunnel junction + Ohmic resistor" from the linear current response in the circuit $I(\omega) = G_{\rm tot}(\omega)V_x(\omega)$. Then with the aid of Ref. 6 after the analytic continuation to real frequencies one finds

$$G_{\rm tot}(\omega) = \frac{i\omega C}{i\omega R_s C - 1} + \frac{\langle I_t(\omega)I_t(-\omega)\rangle}{i\omega(i\omega R_s C - 1)^2}, \qquad (47)$$

where the correlation function for the "tunneling current" $\langle I_t(\omega)I_t(-\omega)\rangle$ (Ref. 10) is calculated in imaginary time with the effective action (5) and then is continued to real times. Expanding to the second order in α_t after a simple algebra we get for $\omega \ll E_c$ (Ref. 11)

$$\langle I_t(\omega)I_t(-\omega)\rangle = \frac{4\alpha_t\omega^2 C}{\pi^2} + \frac{2e^2\alpha_t^2}{15\pi^3}\frac{(-i\omega)^7}{E_c^6}. \qquad (48)$$

Then substituting this result into (47) we obtain for $Q_x < e/2$, $T \to 0$ and small ω

$$G_{\rm tot}(Q_x,\omega) = \frac{1}{R_{\rm eff}(\omega) + i/\omega C_{\rm eff}},$$
$$R_{\rm eff}(\omega) = R_s + \frac{16\alpha_t^2 R_q}{15\pi^4}\left(\frac{\omega}{E_c}\right)^4, \qquad (49)$$

where $C_{\rm eff}(Q_x) = C(d\langle Q\rangle/dQ_x)^{-1}$ is the effective (renormalized) junction capacitance calculated before. The result (49) is valid for $R_S \gg \alpha_t R_q$.[6] We see that at low frequencies and $Q_x < e/2$ single electron tunneling across the junction gives rise to the renormalization of both the junction capacitance and the Ohmic resistance (49). The latter effect vanishes at $\omega \to 0$ in a complete agreement with the result (43) $\Gamma(Q_x < e/2) = 0$.

IV. DISCUSSION

It is interesting to compare the results of our analysis with those obtained by means of an instanton technique in the limit of large junction conductance.[5,6] Both approaches show that for small values of the external charge $Q_x \ll e$ Coulomb blockade is *not* destroyed by quantum fluctuations of the charge. In both cases of small and large α_t these fluctuations yield a capacitance renormalization defined by Eq. (8) for relatively small α_t and by the expression[5,6]

$$C_{\rm eff} = \frac{C}{16\alpha_t^2}\exp(2\alpha_t - \gamma), \qquad (50)$$

for large α_t, $\gamma = 0.577$ is the Euler constant. In our present analysis we essentially use the small parameter $g = \alpha_t/\pi^2 \ll 1$. The corresponding small parameter for the instanton technique of Refs. 5 and 6 is $\exp(-2\alpha_t) \ll 1$. In a certain window for the parameter α_t of order one (more precisely $\alpha_t \sim 1$–3) these conditions overlap and one can expect both theories to provide a qualitatively correct description of the problem. Indeed, the expressions for the effective capacitance obtained within these two approaches match approximately at $\alpha_t \simeq 1.5$–2, i.e., at the borderline of applicability of these approaches. Therefore, a combination of (8) and (50) with a sufficient accuracy describes the effective junction capacitance $C_{\rm eff}$ for all values of α_t.

In the vicinity of the Coulomb-blockade threshold $Q_x = e/2$ the situation becomes somewhat more complicated. For small values of g the analysis developed here yields a finite slope for the ground-state energy $E_0(Q_x)$ (and, therefore, nonzero expectation value $\langle Q\rangle$) for all $Q_x \neq e/2$. Although the band $E_0(Q_x)$ has an obvious tendency to flattening with increasing g one can hardly extrapolate our results to relatively large values of $g \gtrsim 1$ to draw any definite conclusion about the form of $E_0(Q_x)$ for such values of g. To estimate the validity range of our analysis let us combine (37) and (38) with an obvious requirement $\langle Q\rangle \geq 0$. Then within the accuracy of our calculation we immediately get

$$\alpha_t \lesssim \pi^2/b \simeq 2.6. \qquad (51)$$

For larger values of α_t the approximation of noncrossing diagrams is not sufficient and a more sophisticated technique has to be developed. Then in principle one can expect qualitative changes in the behavior of $E_0(Q_x)$ close to the point $Q_x = e/2$. For example, the instanton analysis of Refs. 5 and 6 yields the flat band and $\langle Q\rangle = 0$ in the vicinity of this point. Accordingly, the junction acquires a nonzero conductance already for $V_x < e/2C$. The transition between this flat band picture for large α_t and one described here in the limit of small α_t can take place as a sharp crossover or a phase transition for values $\alpha_t \sim 2$–3. We, however, would like to emphasize that the presence of such a crossover or a phase transition by no means implies destruction of the Coulomb blockade for small values of Q_x in which case the only effect of quantum fluctuations of the charge is the capacitance renormalization (8) and (50).

As was already pointed out, the ground-state energy of a small tunnel junction in the limit $g \ll 1$ has been previously calculated by Matveev[7] who made use of the renormalization group approach developed for the anisotropic Kondo problem. This calculation has been carried out under specific assumptions about the details of the model for electron tunneling across the insulating layer. Our analysis demonstrates that the junction behavior is practically insensitive to such details. For example, assuming that the tunneling matrix element is independent on the momentum values of tunneling electrons and following Ref. 8 one can immediately recover the effective action (5) and (6). The same effective action can be recovered for a "wide junction" Hamiltonian of Ref. 7 as well as for other tunneling models provided the junction cross section area is much larger than the square of the Fermi wavelength. This is always the case for metallic tunnel junctions. The only physical quantity that is sensitive to particular features of the tunneling model is the transparency of a tunnel barrier whereas the junction properties expressed in terms of its tunneling conductance $1/R_t$ remain the same.

Recently, the problem investigated here was also considered by Falci, Schön, and Zimanyi.[12] Starting from the effective action (5) and (6) and proceeding within the framework of the noncrossing diagram approximation these authors reduced the problem to the effective two-state one which then was treated by means of a renormalization group technique. The result for $\langle Q\rangle$ obtained in Ref. 12 coincides with our result (37) and (38) within the main logarithmic approximation. To establish a connection between our analysis and that developed in Ref. 12 let us follow[12] and restrict our attention only to the noncrossing two-state diagrams, neglecting diagrams that describe virtual tunneling to higher energy states. For Q_x close to $e/2$ contributions from these neglected diagrams are proportional to the lower than the main powers of $\ln a(Q_x)$. Therefore, the accuracy of the method[12] coincides with that of the approximation of main logarithms. Furthermore, proceeding within the noncrossing two-state diagram analysis it is straightforward to check that— in contrast to the analysis presented here (see the Appendix)— it turns out to be impossible to regularize the problem by means of a unique shift of the system energy. In other words, the effective two-state system free energy essentially depends on the high frequency cutoff parameter ω_c. Calculating this free energy perturbatively in α_t and making use of the standard scaling arguments we arrive at the equations

$$d\Delta_{10}/d\ln\omega_c = 2g\Delta_{10}, \quad dE_0/d\ln\omega_c = -g\Delta_{10}, \quad dg/d\ln\omega_c = 2g^2, \tag{52}$$

which leave the free energy invariant under the procedure of successive decreasing of ω_c. These equations essentially coincide with those derived in Ref. 12 [the first Eq. (52) has been derived in an earlier paper by Guinea and Schön[4]]. It is interesting to point out that this scaling procedure cannot be generalized to describe our initial problem which includes all charge states because the (regularized) free energy for this problem does not depend on ω_c. Thus strictly speaking the renormalization group Eqs. (50) are valid only for the reduced two-state problem[12] but not for the initial one. As we already discussed, the results of both for $\langle Q\rangle$ coincide in the main logarithmic approximation. Some other results differ. For example, it is easy to see from (52) that the midgap line $E_0(e/2) = E_0(Q_x) + \Delta_{10}/2$ remains unchanged whereas our analysis yields $E_0(e/2) = E_c(1 - 8g\ln 2)/4$.

Finally let us note that the problem dicussed here also shows many similarities to the two-level system with linear Ohmic dissipation.[13] For example, the first Eq. (50) for the level splitting Δ_{10} coincides with the analogous equation[13] if we identify g with the square of the tunneling matrix element between two levels (or the fugacity). The third Eq. (52) is also somewhat similar (although not identical) to the equation for the fugacity[13] for a particular case of a dimensionless dissipative parameter equal to one. Both problems can be mapped onto that of logarithmically interacting gas of blips. An important formal difference between them is that in our problem only pairs of blips interact between each other whereas for a dissipative two-level system[13] such interaction occurs between all blips. This results in different physical behavior obtained for these two problems.

ACKNOWLEDGMENTS

We would like to thank C. Bruder, G. Falci, F. Guinea, S.V. Panyukov, G. Schön, and G.T. Zimanyi for useful discussions. We acknowledge the support by the NATO Linkage Grant No. 920525, by INTAS under Grant No. 93-790 and by the Russian Foundation for Fundamental Research under Grant No. 93-02-14052.

APPENDIX

Let us consider the grand partition function (4) with the action

$$S[\varphi] = \int d\tau \frac{C}{2}\left(\frac{\dot\varphi}{2e}\right)^2 + \int d\tau \int d\tau' \alpha(\tau - \tau') \times \left[\exp(-E_c|\tau - \tau'|) - \cos\left(\frac{\varphi(\tau) - \varphi(\tau')}{2}\right)\right]. \tag{A1}$$

To illustrate the main idea of our regularization procedure we first expand the partition function to the first order in α_t. Making use of (A1) we get

$$\mathcal{Z}^{(1)}(\beta,Q_x)=\exp(-Q_x^2\beta/2C)\int_0^\beta d\tau\int_0^\beta ds\alpha(s)B_1(s), \tag{A2}$$

where

$$B_1(s)=\exp\left[-\frac{e}{C}\left(\frac{e}{2}+Q_x\right)s\right]+\exp\left[-\frac{e}{C}\left(\frac{e}{2}-Q_x\right)s\right]-2\exp\left(-\frac{e^2}{2C}s\right). \tag{A3}$$

In the limit of small s we have $B_1(s)\propto s^2$. This compensates the singular behavior of $\alpha(s)\propto 1/s^2$ for small s and the integral (A2) remains finite.

In order to demonstrate that such regularization remains in higher orders of the perturbation theory let us expand the partition function with the action (A1) in powers of α_t. Then we get

$$\mathcal{Z}(\beta,Q_x)=\exp(-Q_x^2\beta/2C)\left(1+\sum_{n=1}^{\infty}\sum_{\nu_i}\frac{1}{n!}\int_0^\beta\ldots\int_0^\beta ds_1ds_1'\ldots ds_nds_n'\alpha(s_1-s_1')\ldots\alpha(s_n-s_n')\right.$$
$$\left.\times\exp\left[-E_c\sum_{j=1}^{n}|s_j-s_j'|\right]B(s_1,\ldots,s_n,s_1',\ldots,s_n')\right), \tag{A4}$$

where

$$B(s_1,\ldots,s_n,s_1',\ldots,s_n')=(-1)^{n-\sigma_1}2^{-\sigma_2}\exp\left[\frac{eQ_x}{C}\sum_{j=1}^{n}\nu_j(s_j'-s_j)\right.$$
$$\left.-E_c\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\nu_i\nu_j(|s_i'-s_j|-|s_i'-s_j'|-|s_i-s_j|+|s_i-s_j'|)\right]. \tag{A5}$$

Here we imply $\nu_j=-1;0;+1$ for a summation over "charge" configurations ν and denote

$$\sigma_1=\sum_j\nu_j,\quad \sigma_2=\sum_j\nu_j^2. \tag{A6}$$

We are interested in the behavior of the expression (A3) for small $|s_k-s_k'|$ where the kernel $\alpha(s_k-s_k')$ (6) diverges as $1/(s_k-s_k')^2$. To find the function B in this limit we perform a summation over $\nu_k=-1;0;+1$ $(k\neq j)$ in (A3). Then we obtain

$$-B(s_1,\ldots,s_n,s_1',\ldots,s_n')=(-1)^{n-\sigma_1'}2^{-\sigma_2'}\exp\left[\frac{eQ_x}{C}\sum_{j\neq k}\nu_j(s_j'-s_j)\right.$$
$$\left.-\frac{E_c}{2}\sum_{i,j\neq k,i\neq j}(1-\delta_{ij})\nu_i\nu_j(|s_i'-s_j|-|s_i'-s_j'|-|s_i-s_j|+|s_i-s_j'|)\right]$$
$$\times\left(\cosh\left[\frac{eQ_x}{C}(s_k'-s_k)-E_c\sum_{j\neq k}\nu_j(|s_k'-s_j|-|s_k'-s_j'|-|s_k-s_j|+|s_k-s_j'|)\right]-1\right), \tag{A7}$$

where

$$\sigma_1'=\sum_{j\neq k}\nu_j,\quad \sigma_2'=\sum_{j\neq k}\nu_j^2. \tag{A8}$$

For small $|s_k-s_k'|$ the argument of the hyperbolic cosine in (A5) is close to zero and we have $B\propto(s_k-s_k')^2$. Therefore, the product $B\alpha(s_k-s_k')$ remains finite for $|s_k-s_k'|\to 0$ and no divergencies appear in the problem defined by the effective action (A1), (6).

Finally, let us note that the frequently used form of the effective action in which $\cos\left(\frac{\varphi(\tau)-\varphi(\tau')}{2}\right)$ is replaced by $\cos\left(\frac{\varphi(\tau)-\varphi(\tau')}{2}\right)-1$ (see, e.g., Refs. 2 and 4) does not allow one to avoid high frequency divergencies in the perturbation theory. Indeed, the first order correction to the grand partition function in this case again has the form (A2) with

$$B_1(s)=\exp\left[-\frac{e}{C}\left(\frac{e}{2}+Q_x\right)s\right]+\exp\left[-\frac{e}{C}\left(\frac{e}{2}-Q_x\right)s\right]-2. \tag{A9}$$

For $s\to 0$ we have $B_1(s)\propto s$ and the integral over s in (A2) diverges logarithmically at small s.

[1] D.V. Averin and K.K. Likharev, in *Mesoscopic Phenomena in Solids*, edited by B.L. Altshuler, P. Lee, and R.A. Webb (Elsevier, Amsterdam, 1991), p. 167.
[2] G. Schön and A.D. Zaikin, Phys. Rep. **198**, 237 (1990).
[3] P. Lafarge *et al.*, Z. Phys. B **85**, 327 (1991).
[4] F. Guinea and G. Schön, Europhys. Lett. **1**, 585 (1986); J. Low Temp. Phys. **69**, 219 (1987).
[5] S.V. Panyukov and A.D. Zaikin, Phys. Rev. Lett. **67**, 3168 (1991).
[6] A.D. Zaikin and S.V. Panyukov, Phys. Lett. A **183**, 115 (1993).
[7] K.A. Matveev, Zh. Eksp. Teor. Fiz. **99**, 1598 (1991) [Sov. Phys. JETP **72**, 892 (1991].
[8] V. Ambegaokar, U. Eckern, and G. Schön, Phys. Rev. Lett. **48**, 1745 (1982).
[9] See, e.g., K. Huang, *Quarks, Leptons and Gauge Fields* (World Scientific, Singapore, 1982).
[10] E. Ben-Jacob, E. Mottola, and G. Schön, Phys. Rev. Lett. **51**, 2064 (1983).
[11] The ω^7 dependence of the correlation function $\langle I_t(\omega)I_t(-\omega)\rangle$ (with a slightly different prefactor) has been also obtained by Yu.V. Nazarov, Fiz. Nizk. Temp. (Kiev) **16**, 718 (1990) [Sov. J. Low Temp. Phys. **16**, 422 (1990)].
[12] G. Falci, G. Schön, and G.T. Zimanyi (unpublished).
[13] For a review see, e.g., A.J. Leggett *et al.*, Rev. Mod. Phys. **59**, 1 (1987).