(2)

Annual Progress Report
ONR/ARPA Grant # N00014-93-1-0880

Semiconductor Single-electron Digital Devices and Circuits

Principal Investigator: Prof. D.V. Averin
Department of Physics
State University of New York
Stony Brook, NY 11794-3800

Project Period: June 20, 1993 — June 20, 1996
Report Period: June 20, 1993 — June 20, 1994

Research Objectives:

Theoretical investigation of single-electron tunneling phenomena in semiconductor heterostructures, and their application possibilities in the development of ultradense logic and memory circuits.

Report Period Objectives:

1. To design and model simple digital circuits based on capacitively coupled single-electron transistors.

2. To extend the "orthodox" theory of correlated single-electron tunneling to ultrasmall quantum dots.

Accomplishments:

1. Analysis of digital circuits based on capacitively-coupled single-electron transistors — Prof. K. Likharev (Co-P.I.), R. Chen, A. Korotkov.

We have analyzed the possibility of using capacitively-coupled single-electron (SET) transistors to build logic and memory circuits. The analysis

shows that while the potential density of such circuits is very high, up to 10^{10} cells/cm^2, their operating temperature is limited to about 1/100th of the transistor charging energy. For 10-nm technology this confines possible circuits to liquid helium temperatures. However, the physics of SET transistors perm the cells to be scaled down further, with a corresponding increase in operation temperature and circuit density. The approach developed in our work enables us to calculate all relevant circuit parameters, including switching delays, power dissipation, and noise margins.

2. Numerical analysis of quantum tunneling of charge in complex single-electron circuits. — Prof. K. Likharev (Co-P.I.), L. Fonseca, A. Korotkov.

One of the problems of single-electron circuits based on the controlled transfer of electrons one-by-one is macroscopic quantum tunneling (mqt) of charge. It leads to unwanted electron transition through the imposed Coulomb energy barriers, and thus causes errors in device operation. Calculation of the mqt rates in realistic circuits presents a complicated problem that requires extensive numerical simulations. We have developed the software package that performs such a simulation for arbitrary circuits with the maximum mqt order 8–9. In future we plan to use this software extensively for the design of single-electron circuits.

3. Extension of the "orthodox" theory of single-electron tunneling to ultrasmall quantum dots.

3.1 Tunneling in an arbitrary electromagnetic environment. — Prof. D. Averin, H. Imam, A. Korotkov, V. Ponomarenko (visiting scientist).

We extended the standard theory of single-electron tunneling in two directions making it applicable to ultrasmall quantum dots. Firstly, the main feature of such dots that determines their electron transport properties is the formation of discrete zero-dimensional (0D) electron states.

We developed a theory of electron tunneling via one non-degenerate 0D state localized in such a dot in the presence of an arbitrary electromagnetic environment. One of the conclusions of this theory is that in sufficiently resistive environment (the characteristic resistance larger than the so-called resistance quantum equal to 13 KOhm) electron transfer between the 0D states has an *irreversible* character, and in principle can be used to design single-electron circuits in complete analogy to circuits with larger dimensions. This result removes the main potential limitation on caling down of single-electron devices.

Another limitation of the orthodox theory is related to an assumption of instantaneous electron tunneling. We have developed a theory of electron tunneling with finite transversal time and arbitrary electromagnetic environment. It shows that although there are quantitative corrections to the orthodox theory the qualitative picture of single-electron tunneling remains valid.

3.2 Numerical modeling of electron-electron interaction in a few-electron quantum dots. — Prof. J.K. Jain (Co-P.I.), L. Belkhir.

The quantitative design of single-electron devices is based on the description of the charging energy in terms of the geometrical capacitances C. Such a description is very accurate for relatively large metallic structures, but can not be, strictly speaking, justified for small semiconductor heterostructures. We performed exact numerical simulations of electron-electron interaction in few-electron quantum dots and showed that the classical expression for the charging energy should be semi-quantitatively (with about 30% accuracy) valid even for quantum dots which contain as few as five electrons.

POSSIBLE EXTREMELY-HIGH-DENSITY MEMORY BASED ON SINGLE-ELECTRON TRANSISTORS

R. Chen, A.N. Korotkov, and K.K. Likharev

Department of Physics, State University of New York
Stony Brook, NY 11794-3800, 516-632-8159

Recently discovered effects of correlated single-electron tunneling (for reviews see, e.g., Refs. 1, 2) form the physical basis for a new generation of electronic devices, in particular, very dense memories. Digital bits in these memories may be represented by either single electrons or bundles of a few electrons. In this work we are suggesting a static random-access memory (SRAM) of the latter type, where the number N of electrons in the bundle (N~10) oscillates in time by $\delta N = \pm 1\text{-}2$. (Larger fluctuations of N, implying digital errors, occur with vanishing probability.) In contrast to memory cells based on the trapping of single electrons [2], the new approach implies somewhat higher power consumption, but promises much higher speed and wider parameter margins.

The memory cell consists of two single-electron transistors (SETs) [1, 2] connected in a positive-feedback loop. In each of two stable states of this symmetrical flip-flop, one SET is open, while another is in the closed (Coulomb-blockade) state. Two additional SETs enable write-0, write-1 and read-out operations, in a mode very similar to that of the usual MOSFET SRAMs. The only substantial difference is the high output impedance of the SETs, which makes it necessary to use at least a two-level hierarchy of sense amplifiers (the first level should use SETs, while the next levels may be based on usual FETs).

We have used the "orthodox" theory of single-electron tunneling, and its extension to co-tunneling processes [2], to optimize the SET SRAM cell and calculate its parameter tolerances, switching speed, and power consumption. These results have enabled us to estimate the possible performance of the SET SRAMs for several levels of their fabrication technology. For example, a 10-nm silicon-based technology should allow implementation of this type of memory with density close to 10 Gb/cm^2, cycle time of ~300 ps, and power consumption ~3 W/cm^2, operating at liquid helium temperatures (~4 K). It is important that the physics of SET transistors permits the memory cells to be scaled down further, with a corresponding increase in cell density and operation temperature.

The work was supported in part by AFOSR and ONR/ARPA.

1. D.V. Averin and K.K. Likharev, in: "Mesoscopic Phenomena in Solids", ed. by B. Altshuler et al., Elsevier, 1991, pp. 173-271.

2. "Single Charge Tunneling", ed. by H. Grabert and M.H. Devoret, Plenum, 1992.

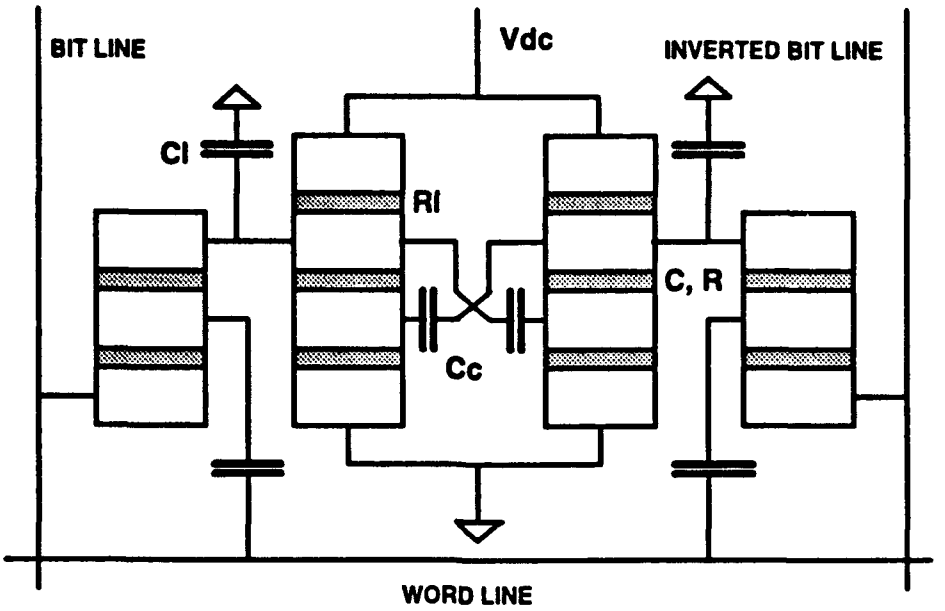

Figure 1. Possible circuit diagram of the SET SRAM cell. Open rectangles denote conducting electrodes, while gray rectangles show tunnel junction barriers.

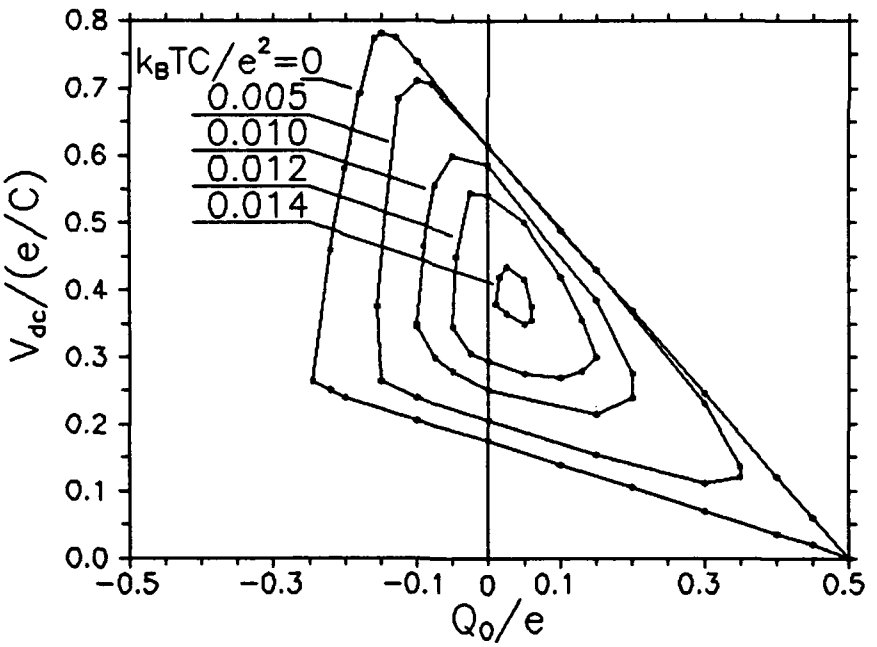

Figure 2. Parameter window of the correct operation of the cell with Cc=2C, Cl>>C, and Rl=10R, for several operation temperatures T (in units e^2/k_BC). Q_0 is the background charge. One can see that the window closes rapidly at temperatures above ~$0.01e^2/k_BC$.

Coulomb blockade of resonant tunneling

H.T. Imam,[1] V.V. Ponomarenko,[1,2] and D.V. Averin[1,3]

[1] *Department of Physics, State University of New York, Stony Brook, NY 11794;*

[2] *A.F. Ioffe Physical Technical Institute, 194021 St. Petersburg, Russia;*

[3] *Department of Physics, Moscow State University, Moscow 119899 GSP, Russia*

We have considered the influence of electromagnetic fluctuations on electron tunneling via one non-degenerate resonant level, the problem that is relevant for electron transport through quantum dots in the Coulomb blockade regime. We show that the overall effect of such an influence depends on whether the electron bands in external electrodes are empty or filled. In the empty band case, depending on the relation between the tunneling rate Γ and characteristic frequency Ω of the fluctuations, the field either simply shifts the conductance peak (for rapid tunneling, $\Gamma \gg \Omega$) or broadens it (for $\Gamma \ll \Omega$). In the latter case, the system can be in three different regimes for different values of the coupling g between electrons and the field. Increasing interaction strength in the region $g < 1$ leads to gradual suppression of the conductance peak at the bare energy of the resonant level ε_0, while at $g \gg 1$ it leads to the formation of a new peak of width $E_c/g^{1/2}$ at the energy $\varepsilon_0 + E_c$, where E_c is a charging energy. For intermediate values of g the conductance is non-vanishing in the entire energy range from ε_0 to $\varepsilon_0 + E_c$. These results provide a possible explanation for experimentally observed extra width of the conductance resonances at low temperatures. For filled bands the problem is essentially multi-electron in character. One consequence of this is that, in contrast to the situation with the empty band, the fluctuations of

the resonant level do not suppress conductance at resonance for $g < 1$. At $g = 1$ a phase transition occurs leading to the appearance of a Coulomb gap in the position of the resonant level as a function of its bare energy, and results in the suppression of conductance.

PACS numbers: 73.20.Dx, 73.40.Gk, 73.40.Kp

I. INTRODUCTION

Several recent experiments [1]–[5] demonstrated resonant tunneling under Coulomb blockade conditions, when the resonant level is localized in a mesoscopic quantum dot and is affected by electron-electron interaction in the dot. The interaction shifts the energy of th resonant level by an amount which is roughly proportional to n, the number of electrons in the dot. This makes the effective energy spectrum of the dot a nearly periodic function of n, with each period corresponding to the addition of one electron to the dot. This periodicity leads to several new phenomena which are attracting considerable interest.

The effect of the interaction has such a simple form of an energy shift only if the response of electrons in the dot and/or external electrodes to the transfer of an electron to or from the resonant level has a time scale which is incompatible with the time scale of tunneling. Our aim in this work is to consider the situation when such a condition is not satisfied and the characteristic response time can be comparable to the tunneling time. Following the theory of electron tunneling in small metallic tunnel junctions (see, e.g., [6]), one can model the electron-electron interaction in this regime as the interaction of electrons with fluctuations of the electromagnetic field in a given electromagnetic environment. Thus, we can reduce the problem under consideration to one of inelastic resonant tunneling of electrons coupled to bosonic degrees of freedom. The relation between the tunneling time and characteristic frequency of the boson modes is important in this model because, depending on whether the boson modes have enough time to adjust to the changing charge of the resonant level or not, the effective energy of this level either coincides with the bare energy ε_0 or is shifted by the charging energy E_c.

This fact, together with the renormalization of the effective tunneling rate by the

interaction, leads to several different regimes in the dynamics of tunneling through the resonant level. Identification of various regimes of electron tunneling and calculation of the current-voltage characteristics of the resonant level in these regimes are the aims of this work.

The paper is organized as follows. In Sec. 2 we derive the general expression for the current through the resonant level, and prove that in the large-bias limit the current reaches an asymptotic value which is independent of interaction. In Sec. 3 we simplify the general expression by adopting a one-electron approximation which was originally developed in [7]–[9], and establish limits on the validity of this approximation. It is shown that the one-electron approximation is valid when the electron band in at least one of external electrodes is empty and the rate of tunneling into this electrode is much larger than the rate of tunneling into the other electrode. Both of these conditions are typically satisfied in vertical resonant tunneling structures. In Sec. 4 we calculate the current-voltage characteristic of such a structure in the one-electron approximation. In Sec. 5 we consider the situation (typical for lateral quantum dots) when the bands in both external electrodes are filled and the one-electron approximation is incorrect, and calculate the linear conductance and the current within two approximation schemes, adiabatic approximation and perturbation theory in tunneling. It is shown that if interaction with the bosonic mode is strong enough and the bare tunneling rate is small, then, due to a strong renormalization of the tunneling amplitude a Coulomb gap appears in the effective position of the resonant level as a function of its bare energy . In Conclusion we discuss the relation of our results to existing experiments on electron tunneling in quantum dots, and also discuss them from the point of view of general approaches to transport of interacting electrons.

II. MODEL AND BASIC EXPRESSION FOR CURRENT

We consider tunneling of electrons between two bulk electrodes via a quantum dot with a resonant non-degenerate state using a standard tunneling Hamiltonian:

$$H_{el} = \sum_{j,kj} \varepsilon_{kj} c_{kj}^{\dagger} c_{kj} + \varepsilon_0 c^{\dagger} c + \sum_{j,kj} (t_j c_{kj}^{\dagger} c + h.c.) . \qquad (1)$$

Here index $j = 1, 2$ denotes, respectively, left and right electrode, and ε_0 is the energy of the resonant level. The assumption of a non-degenerate level can be justified in several situations. In particular, strong magnetic field applied to the dot may result in the Zeeman splitting of the spin-degenerate levels. Another possibility is a strong intrasite electron repulsion which may lead to significant splitting of the initially degenerate levels when the Kondo resonance is destroyed by non-vanishing temperature or bias voltage.

To describe the relevant low-energy properties of tunneling through the resonant level, we need to take into account the interaction of the tunneling electron with low-energy excitations in the quantum dot structure. In the mesoscopic structures, the relevant excitations are low-frequency modes of the longitudinal electromagnetic field associated with fluctuations of the total electric charge of electrodes (see, e.g., Ref. [6]). Using standard gauge transformation we can express the interaction with these modes in the form of field-induced phases of the electron tunneling amplitudes t_j:

$$t_j \to t_j e^{i\phi_j(t)} , \qquad \dot{\phi}_j(t) = e\tilde{V}_j(t) , \qquad (2)$$

where $\tilde{V}_j$ is the fluctuating voltage between the resonant level and jth external electrode.

Below we consider a standard structure with two bias external electrodes and one gate electrode (Fig. 1). In such a structure there are two types of low-frequency

modes, one associated with fluctuations of the bias voltage, and another one associated with the fluctuations of the gate voltage. Adding the energy of these two types of modes to the electron Hamiltonian (1) with tunneling amplitudes (2) we get the total Hamiltonian of our model:

$$H = H_{el} + \sum_{\omega} \omega a_{\omega}^{\dagger} a_{\omega} + \sum_{\omega} \omega b_{\omega}^{\dagger} b_{\omega} \,. \tag{3}$$

An important characteristic of the low-frequency photon modes is that the spatial distribution of the electric field in them is determined solely by the geometry of the system and is frequency-independent. Hence, we can introduce frequency-independent factors λ, $\lambda_{1,2}$ which determine the distribution of the voltages in the electrodes of the system and write the phases ϕ_j as sums of the two terms corresponding to fluctuations of the bias voltage and of the gate voltage:

$$\phi_1 = \lambda \phi_V + \lambda_1 \phi_g \,, \quad \phi_2 = -(1-\lambda)\phi_V + \lambda_2 \phi_g \,. \tag{4}$$

Here λ and λ_1 give, respectively, the fraction of the bias voltage and of the gate voltage that drop between the dot and left external electrode; $(1-\lambda)$ and λ_2 have the same meanings for the dot and right external electrode.

The phases ϕ_V and ϕ_g can be directly expressed in terms of the photon modes a and b:

$$\phi_V = \sum_{\omega} \eta_{\omega} (a_{\omega}^{\dagger} + a_{\omega}) \,, \quad \phi_g = \sum_{\omega} \nu_{\omega} (b_{\omega}^{\dagger} + b_{\omega}) \,. \tag{5}$$

where the coupling constants η_ω and ν_ω together with the density of modes are related to the effective impedances of the structure [6]:

$$\sum_{\omega} \eta_{\omega}^2 \ldots = \int \frac{d\omega}{\omega} F_V(\omega) \ldots \,, \quad \sum_{\omega} \nu_{\omega}^2 \ldots = \int \frac{d\omega}{\omega} F_g(\omega) \ldots \,.$$

Here spectral densities $F_{V,g}(\omega) = (e^2/2\pi\hbar)\mathrm{Re}Z_{V,g}(\omega)$ are given by the real part of impedances of the bias circuit and of the gate electrode — see Fig. 1.

Below we will study in detail two specific examples of the spectral densities $F(\omega)$. In the Ohmic case, the impedance is reduced to the frequency-independent resistor R. Then

$$F(\omega) = g/(1 + (\omega/\Omega)^2)\,, \tag{6}$$

where $g \equiv e^2R/\pi$, and the cut-off frequency Ω is determined by the finite capacitance C between the electrodes, $\Omega = 1/RC$. Another example corresponds to an inductor L in the external circuit which leads to a one-mode spectral density:

$$F(\omega) = g\Omega \sum_{\pm} \delta(\omega \pm \Omega)\,, \tag{7}$$

where $\Omega = (LC)^{-1/2}$ is the mode frequency, and $g = (e^2/\pi)(L/C)^{1/2}$ is the coupling constant. Spectral densities (6), (7) satisfy the sum rule:

$$\int d\omega F(\omega) = \frac{e^2}{C} \equiv 2E_c\,. \tag{8}$$

which should be satisfied by arbitrary $F(\omega)$ [6].

Although we have used, so far, the language appropriate specifically for the photon modes, the model is obviously valid for other bosonic modes. Moreover, to the extent that the "bosonization" procedure is applicable to excitations of arbitrary system, the model describes resonant tunneling of electrons interacting with arbitrary dissipative environment. As an example, one can note that the microscopic model of Coulomb screening of the resonant level by a Fermi sea of electrons [10], [11] can be reduced to the bosonic form considered in this work. In the microscopic model an electron on the resonant level creates a self-consistent potential for electrons in the Fermi sea, which gives rise to shift of scattering phases φ_j of these electrons. Bosonization of electron-hole excitations of the Fermi sea reduces this model to the one discussed above with

$$F_g(\omega) = \sum_j (\varphi_j/\pi)^2 e^{-\omega/\Omega}\,, \tag{9}$$

where the exponential cut-off is commonly used for simplicity.

From a general point of view, the spectral density (6) corresponds to an environment with a relaxation-type response to charging of the resonant level characterized by a single relaxation-time constant Ω^{-1}, while the spectral density (7) corresponds to the environment with an oscillatory response and a single frequency Ω.

Our aim is to calculate the dc current I through the resonant level which can be expressed in terms of either left or right currents,

$$I = I_1 = I_2\,,$$

where

$$I_1 = -2e\mathrm{Im}\sum_{k1} t_1 \langle c_{k1}^\dagger e^{i\phi_1} c\rangle\,, \quad I_2 = 2e\mathrm{Im}\sum_{k2} t_2 \langle c_{k2}^\dagger e^{i\phi_2} c\rangle\,. \tag{10}$$

Writing down the expansion of the averages in eq. (10) using perturbation theory with respect to the tunneling terms in the Hamiltonian, and constructing the corresponding Dyson equation, we express the currents in terms of non-equilibrium Green's functions of the resonant level. (A similar transformation was used in [12].) For instance, the left current is:

$$I_1 = -2e \mid t_1 \mid^2 \mathrm{Im}\sum_{k1}\int \frac{d\varepsilon}{2\pi i}[G_{k1}^<(\varepsilon)G_1^t(\varepsilon) - G_{k1}^{\tilde{t}}(\varepsilon)G_1^<(\varepsilon)]\,, \tag{11}$$

where $G_1(\varepsilon)$ is the exact (i.e., calculated in presence of tunneling) Green's function of the resonant level dressed by the phase $e^{i\phi_1}$, e.g.,

$$G_1^<(\varepsilon) = \int dt e^{i\varepsilon t} G_1^<(t)\,, \quad G_1^<(t) = i\langle c^\dagger e^{-i\phi_1(0)} e^{i\phi_1(t)} c(t)\rangle\,,$$

while $G_{k1}(\varepsilon)$ is a free Green's function of the left electrode. (Notations for G's in eq. (11) are the same as in the book by Mahan [13], Sec. 2.9.) Using the fact that $G^>(\varepsilon)$

and $G^<(\varepsilon)$ are purely imaginary and $\mathrm{Im}G^t(\varepsilon) = [G^>(\varepsilon) + G^<(\varepsilon)]/2i$, and plugging expressions for G_{k1}, $G^<_{k1}(\varepsilon) = 2\pi i f_{k1}\delta(\varepsilon - \varepsilon_{k1})$, $\mathrm{Im}G^t_{k1}(\varepsilon) = \pi(2f_{k1} - 1)\delta(\varepsilon - \varepsilon_{k1})$ in eq. (11) we get finally:

$$I_1 = -\frac{e\Gamma_1}{2\pi i}\int d\varepsilon[f_1(\varepsilon)G_1^>(\varepsilon) + (1 - f_1(\varepsilon))G_1^<(\varepsilon)]\,. \qquad (12)$$

Here $\Gamma_1 = 2\pi \mid t_1 \mid^2 \sum_{k1}\delta(\varepsilon - \varepsilon_{k1})$ is the rate of tunneling into the left electrode, and $f_1(\varepsilon)$ is Fermi distribution function in this electrode.

The current between the resonant level and the right electrode can be expressed similarly:

$$I_2 = \frac{e\Gamma_2}{2\pi i}\int d\varepsilon[f_2(\varepsilon)G_2^>(\varepsilon) + (1 - f_2(\varepsilon))G_2^<(\varepsilon)]\,, \qquad (13)$$

where $G_2(\varepsilon)$ is the Green's function of the resonant level dressed by the phase $e^{i\phi_2}$.

Equations (12) and (13) will be the starting point of our calculations in the next sections. They show, in particular, that in the large-bias limit the current reaches a constant value that is independent of interaction. Indeed, in this limit $f_1(\varepsilon) = 1$ and $f_2(\varepsilon) = 0$ for all relevant energies, and we have from (12) and (13):

$$I_1 = -\frac{e\Gamma_1}{2\pi i}\int d\varepsilon G_1^>(\varepsilon) = e\Gamma_1(1 - n)\,, \quad I_2 = \frac{e\Gamma_2}{2\pi i}\int d\varepsilon G_2^<(\varepsilon) = e\Gamma_2 n\,, \qquad (14)$$

where $n = \langle c^\dagger c\rangle$ is the occupation probability of the resonant level. Since the left and right currents should be equal, we get from eq. (14) that in the large-bias limit both n and I are the same as in the non-interacting case:

$$n = \frac{\Gamma_1}{\Gamma_1 + \Gamma_2}\,, \qquad I = e\frac{\Gamma_1\Gamma_2}{\Gamma_1 + \Gamma_2}\,. \qquad (15)$$

Using eqs. (12) and (13) we can in principle determine the current for arbitrary bias voltage adopting various approximations for Green's functions of the resonant level.

III. ONE-ELECTRON APPROXIMATION

We begin by considering the one-electron approximation, which was originally developed in [7] - [9]. The starting point of this approach is an assumption of free evolution of the bosonic modes. With this assumption the dynamics of electron operators can be determined straightforwardly. In order to obtain the limits on the validity of this approximation we start with a complete set of equations of motion, both for electron operators c and the bose fields ϕ_j that follow from Hamiltonian (1) with tunneling amplitudes (2):

$$i\dot{c}_{kj} = \varepsilon_{kj}c_{kj} + t_j e^{i\phi_j}c\,, \quad i\dot{c} = \varepsilon_0 c + \sum_{j,kj} t_j^* e^{-i\phi_j}c_{kj}\,,$$

$$\phi_j(t) = \phi_j^{(0)} + 2\lambda_j \sum_i \lambda_i \int_0^\infty \frac{d\omega}{\omega} F(\omega) \int_{-\infty}^t d\tau sin(\omega(t-\tau))I_i(\tau)/e\,. \qquad (16)$$

Here $\phi_j^{(0)}$ denotes free bose field, and $I_{1,2}(\tau)$ are operators of the left and right currents (10). Appropriate sum over the two spectral densities $F_g(\omega)$ and $F_V(\omega)$ is assumed in the last equation (16). Equations of motion for electron operators can be solved explicitly provided that the density of states in external electrodes is constant on the energy scale associated with tunneling:

$$c(t) = -i\int_{-\infty}^t d\tau e^{-(i\varepsilon_0+\Gamma)(t-\tau)} \sum_{j,kj} t_j^* e^{-i\phi_j(\tau)} c_{kj}^{(0)}(\tau)\,. \qquad (17)$$

Here $\Gamma = (\Gamma_1 + \Gamma_2)/2$ is the tunneling width of the resonant level, and $c_{kj}^{(0)}$ are free electron operators which determine the occupation of electron states in external electrodes. Equation (17) gives for the Green's functions of the resonant level:

$$G_1^<(\varepsilon) = i\sum_{i,j} t_j^* t_i \int dt e^{i\varepsilon t} \int_{-\infty}^0 d\tau d\tau' e^{\Gamma(\tau+\tau')} e^{i\varepsilon_0(\tau-\tau')}$$
$$\langle c_{kj}^{(0)+}(\tau') e^{i\phi_j(\tau')} e^{-i\phi_1(0)} e^{i\phi_1(t)} e^{-i\phi_i(t+\tau)} c_{ki}^{(0)}(t+\tau)\rangle\,, \qquad (18)$$

$$G_1^>(\varepsilon) = -i\sum_{i,j} t_j^* t_i \int dt e^{i\varepsilon t} \int_{-\infty}^0 d\tau d\tau' e^{\Gamma(\tau+\tau')} e^{i\varepsilon_0(\tau-\tau')}$$
$$\langle e^{i\phi_1(t)} e^{-i\phi_j(t+\tau)} c_{kj}^{(0)}(t+\tau) c_{ki}^{(0)+}(\tau') e^{i\phi_i(\tau')} e^{-i\phi_1(0)}\rangle\,.$$

In the one-electron approximation, the evolution of the bose field in eq. (18) is taken to be free. Under this assumption eq. (18) is reduced to:

$$G_1^<(\varepsilon) = i\int dt e^{iet}\int_{-\infty}^{0} d\tau d\tau' e^{\Gamma(\tau+\tau')}e^{ie_0(\tau-\tau')}\sum_{j,kj}\mid t_j\mid^2 f_j(\varepsilon_{kj})$$
$$e^{-ie_{kj}(t+\tau-\tau')}\langle e^{i\phi_j(\tau')}e^{-i\phi_1(0)}e^{i\phi_1(t)}e^{-i\phi_j(t+\tau)}\rangle\,, \qquad (19)$$
$$G_1^>(\varepsilon) = -i\int dt e^{iet}\int_{-\infty}^{0} d\tau d\tau' e^{\Gamma(\tau+\tau')}e^{ie_0(\tau-\tau')}\sum_{j,kj}\mid t_j\mid^2 (1-f_j(\varepsilon_{kj}))$$
$$e^{-ie_{kj}(t+\tau-\tau')}\langle e^{i\phi_1(t)}e^{-i\phi_j(t+\tau)}e^{i\phi_j(\tau')}e^{-i\phi_1(0)}\rangle\,.$$

Combining eq. (19) with eq. (12) we find the current:

$$I = \frac{e\Gamma_1\Gamma_2}{2\pi}\int d\varepsilon d\varepsilon'[f_1(\varepsilon)(1-f_2(\varepsilon'))P_{12}(\varepsilon,\varepsilon') - f_2(\varepsilon')(1-f_1(\varepsilon))P_{21}(\varepsilon',\varepsilon)]\,, \qquad (20)$$

where $P_{12}(\varepsilon,\varepsilon')$ and $P_{21}(\varepsilon,\varepsilon')$ are transition probabilities from the state with the energy ε in one electrode to the state ε' in another electrode. They can be written as follows:

$$P_{12}(\varepsilon,\varepsilon') = \frac{1}{2\pi}\int dt e^{i(\varepsilon-\varepsilon')t}\langle A(t)A^\dagger\rangle\,,\quad P_{21}(\varepsilon',\varepsilon) = \frac{1}{2\pi}\int dt e^{i(\varepsilon-\varepsilon')t}\langle A^\dagger A(t)\rangle\,, \qquad (21)$$

where A has the meaning of transition amplitude:

$$A = \int_{-\infty}^{0} d\tau e^{[i(e_0-\varepsilon')+\Gamma]\tau}e^{i\phi_1(0)}e^{-i\phi_2(\tau)}\,. \qquad (22)$$

Representation (21) shows straitforwardly that the left→right and the right→left transitions are related by the detailed balance condition:

$$P_{12}(\varepsilon,\varepsilon') = e^{\beta(\varepsilon-\varepsilon')}P_{21}(\varepsilon',\varepsilon)\,. \qquad (23)$$

Due to this condition the terms that describe transition within one electrode drop out from expression for the current (20). It should be noted that eq. (23) does not

imply left↔right symmetry of transition probabilities, and the current can be an asymmetric function of the voltage if the resonant level is placed asymmetrically with respect to the electrodes (i.e. $\lambda \neq 1/2$).

To understand why the assumption of free bosonic evolution constitutes one-electron approximation one can note that deviations from free evolution are mediating electron-electron interaction, since disturbance of the bosonic mode by one electron affects other electrons. This interaction is disregarded, when the bosonic evolution is assumed to be free.

In order to obtain conditions for the validity of the one-electron approximation, the full evolution of the bose field φ should be taken into account. Deviations of φ from $\varphi^{(0)}$ give rise to corrections in the expressions for G_1 (19) and the current (20). In the first nonvanishing order in electron-boson coupling, correction to the current contains the terms of the following structure:

$$\Delta I \propto \Gamma_1\Gamma_2 \sum_i \Gamma_i \int\int d\epsilon d\epsilon' f_i(\epsilon')[(1-f_1(\epsilon))f_2(\epsilon) - (1-f_2(\epsilon))f_1(\epsilon)][\dots],$$

where $[\dots]$ denotes a function of ϵ, ϵ' precise form of which is not relevant for our argument. This expression shows that all corrections resulting from deviations of the boson evolution from free evolution can be omitted when we can neglect products of the type $\Gamma_i f_i(\epsilon)$, e.g. if the conduction band of the second electrode is empty and transparency of the first barrier is much smaller than that of the second:

$$f_2(\epsilon) = 0, \quad \Gamma_1 \ll \Gamma_2. \tag{24}$$

In simple terms conditions (24) mean that the resonant level is practically empty, so that electrons indeed tunnel independently. Another way to understand eq. (24) is to note that in the one-electron approximation the tunneling width of the resonant level is unaffected by interaction. However, it is obvious on simple physical grounds that since the interaction changes the tunneling rates, the width Γ should also be

modified. Thus, the conditions (24) mean that Γ is determined solely by the rate of tunneling from the resonant level into the empty band which is really unaffected by interaction.

If the conditions (24) are not satisfied, the one-electron approximation is valid only for energies away from resonance, where we can neglect the width of the resonant level altogether, and eq. (20) reduces basically to an expression for the current that follows from the second order perturbation theory in tunneling (see Sec. 5).

Now we proceed with the calculation of transition rates (21). Since the phases ϕ_1, ϕ_2 are combinations of free boson operators, the averages of the four exponents in eq. (19) can be evaluated directly (see, e.g., [13], Sec. 4.3). For instance:

$$\langle e^{i\phi_1(t)}e^{-i\phi_2(t+\tau)}e^{i\phi_2(\tau')}e^{-i\phi_1(0)}\rangle = \exp\{\langle(\phi_1(t)-\phi_1(0))\phi_1(0)+(\phi_2(t+\tau-\tau')-$$

$$\phi_2(0))\phi_2(0)+\phi_1(t)(\phi_2(t+\tau)-\phi_2(\tau'))+(\phi_2(\tau')-\phi_2(t+\tau))\phi_1(0)\rangle\}\,. \quad (25)$$

Together with expression for equilibrium phase correlators,

$$\langle\phi_{g,V}(t)\phi_{g,V}(0)\rangle = \int d\omega \frac{F_{g,V}(\omega)}{\omega}\frac{e^{-i\omega t}}{1-e^{-\beta\omega}}\,,$$

eq. (19) gives the finally:

$$P_{12}(\varepsilon,\varepsilon') = \frac{1}{2\pi}\int dt e^{i(\varepsilon-\varepsilon')t}\int_{-\infty}^{0} d\tau d\tau' e^{\Gamma(\tau+\tau')}e^{i(\varepsilon_0-\varepsilon')(\tau-\tau')}$$

$$\exp\{\int\frac{d\omega}{\omega}\frac{1}{1-e^{-\beta\omega}}[F_{11}(\omega)(e^{-i\omega t}-1)+F_{22}(\omega)(e^{-i\omega(t+\tau-\tau')}-1)+$$

$$F_{12}(\omega)(e^{i\omega\tau}+e^{-i\omega\tau'}-e^{-i\omega(t-\tau')}-e^{-i\omega(t+\tau)}]\}\,. \quad (26)$$

Here

$$F_{11}(\omega) = \lambda_1^2 F_g(\omega)+\lambda^2 F_V(\omega)\,,\quad F_{22}(\omega)=\lambda_2^2 F_g(\omega)+(1-\lambda)^2 F_V(\omega)\,,$$

$$F_{12}(\omega) = \lambda_1\lambda_2 F_g(\omega)-\lambda(1-\lambda)F_V(\omega)\,.$$

For $F_V(\omega)=0$ equation (26) coincides with the results of [7]—[9].

IV. TUNNELING INTO AN EMPTY BAND

In this section we will deal with the case when the conditions (24) are satisfied and the one-electron approximation (expressed by eqs. (20) and (26)) is valid for all energies including those close to ε_0. Under these conditions the expression for the current can be simplified further. Making a change of variables in integral over t in eq. (26), $t \to t - \tau + \tau'$ and integrating $P_{12}(\varepsilon, \varepsilon')$ over the final energy ε' we get (in this Section we will suppress indices 1,2 of P_{12} and F_{11}):

$$I = \frac{e\Gamma_1}{\pi} \int d\varepsilon f_1(\varepsilon) P(\varepsilon), \quad P(\varepsilon) = \mathrm{Re} \int_{-\infty}^{0} d\tau e^{[i(\varepsilon_0 - \varepsilon) + \Gamma_2/2]\tau} e^{J(\tau)},$$

$$J(\tau) = \int d\omega \frac{F(\omega)}{\omega} \frac{e^{i\omega\tau} - 1}{1 - e^{-\beta\omega}}. \tag{27}$$

Equation (27) describes the current step as the current increases from 0, for voltages below the resonance, to $e\Gamma_1$ for voltages above the resonance. We limit ourselves mostly to the calculation of the zero-temperature differential conductance of the structure $G = dI/dV$ which characterizes the shape of the current step and coincides with the transition probability $P(\varepsilon)$:

$$G(\varepsilon) = \frac{e^2 \lambda \Gamma_1}{\pi\hbar} P(\varepsilon). \tag{28}$$

Substitution $\varepsilon = \lambda eV$ in eq. (28) gives the conductance as a function of the bias voltage V.

The shape of the conductance-versus-energy peak depends on the spectral density $F(\omega)$ which characterizes the dissipative environment of the resonant level. In the one-mode case (10) we get from eqs. (27), (28):

$$G = \frac{e^2 \lambda \Gamma_1 \Gamma_2}{2\pi\hbar} e^{-g} \sum_{n=0}^{\infty} \frac{g^n}{n!} \frac{1}{(\delta - n\Omega)^2 + \eta^2}. \tag{29}$$

where $\delta \equiv \varepsilon - \varepsilon_0$, and $\eta \equiv \Gamma_2/2$. For large coupling constant $g \gg \max\{1, (\Gamma_2/\Omega)^2\}$ eq. (29) shows that the envelope of the system of elementary peaks is

$$G = \frac{e^2\lambda\Gamma_1}{\hbar(2\pi E_c\Omega)^{1/2}}\exp\{-\frac{(\delta - E_c)^2}{2E_c\Omega}\}\,, \tag{30}$$

so that the total conductance peak is Gaussian and centered around $\varepsilon_0 + E_c$.

However, when $\Gamma_2 \gg \max\{1, g^{1/2}\}\Omega$ (i.e., response of the photon mode is much slower than the tunneling rate), the peak becomes Lorentzian and has no fine structure. The conductance exhibits then the same peak as in the non-interacting case, but with the peak position shifted to the energy $\varepsilon_0 + E_c$:

$$G = \frac{e^2\lambda\Gamma_1\Gamma_2}{2\pi\hbar}\frac{1}{(\delta - E_c)^2 + \eta^2}\,. \tag{31}$$

Equation (31) can be obtained at large Γ_2 directly from the sum rule (8) and eq. (27), and, hence, is valid for arbitrary frequency dependence of $F(\omega)$. Thus, we see that for the one-mode environment, the conductance peak evolves, as a function of the coupling constant g (for fixed energy $E_c = g\Omega$), from the Lorentzian at $\varepsilon = \varepsilon_0$ for $g \to 0$ to a similar peak centered at $\varepsilon = \varepsilon_0 + E_c$ for $g \gg 1$.

A similar transition takes place for the Ohmic environment (3). To follow this transition we first consider the shape of $G(\varepsilon)$ at small energies, $\mid \varepsilon - \varepsilon_0 \mid \ll \Omega$, $\Gamma_2 \ll \Omega$, and $g < 1$. In this case the conductance is determined by the long-time asymptote of $J(\tau)$ in eq. (27). Using the well-known expression for this asymptote (see, e.g., [6]), we get:

$$G = \frac{e^2\lambda\Gamma_1}{\pi\hbar}\mathrm{Re}\int_{-\infty}^{0} d\tau e^{[-i\delta+\eta]\tau}\frac{e^{-(i\pi/2+\gamma)g}}{(\Omega\tau)^g} = \frac{e^2\lambda\Gamma_1}{\pi\hbar\Omega}e^{-\gamma g}\Gamma(1-g)\times$$
$$\sin[(1-g)\Theta](\frac{\Omega^2}{\delta^2+\eta^2})^{(1-g)/2}\,, \quad \Theta \equiv \frac{\pi}{2} + \tan^{-1}(\delta/\eta)\,, \tag{32}$$

where γ is Euler's constant, and $\Gamma(1-g)$ is the gamma function. Equation (32) shows that for $\delta \gg \eta$ the conductance peak at $\varepsilon = \varepsilon_0$ is suppressed and turns into a plato at $g = 1$:

$$G = \frac{e^2\lambda\Gamma_1 e^{-\gamma}}{\pi\hbar\Omega}[\frac{\pi}{2} + \tan^{-1}(\frac{\delta}{\eta})]\,. \tag{33}$$

For arbitrary g, the conductance can be calculated either numerically or analytically if the Lorentzian cut-off $1/(1+(\omega/\Omega)^2)$ in the Ohmic spectrum is replaced with the exponential cut-off $e^{-\omega/\Omega}$. The integrals in (27) can then be calculated explicitly and give:

$$G = \frac{e^2\lambda\Gamma_1}{\pi\hbar\Omega}\mathrm{Im}\{z^{g-1}\Gamma(1-g,z)\}\,, \quad z \equiv \exp\{-i\Theta\}\frac{(\delta^2+\eta^2)^{1/2}}{\Omega}\,, \tag{34}$$

where $\Gamma(1-g,z)$ is the incomplete gamma function. At low energies, $\mid \varepsilon - \varepsilon_0 \mid \ll \Omega$, $\Gamma_2 \ll \Omega$ eq. (34) gives:

$$G = \frac{e^2\lambda_1\Gamma_1}{\pi\hbar\Omega}\{\Gamma(1-g)\sin[(1-g)\theta](\frac{\Omega^2}{\delta^2+\eta^2})^{(1-g)/2} + \frac{\eta}{(g-2)\Omega}\}\,. \tag{35}$$

Equation (35) generalizes eq. (32) to $g > 1$. It shows that as g is increased further in the region $g > 1$ the conductance around ε_0 is suppressed stronger and, moreover, starts to grow as one goes away from the original resonance at $\varepsilon = \varepsilon_0$. This indicates that another conductance peak is formed at larger energies. The next (after $g = 1$) change in the asymptotic behavior of conductance at low energies occur at $g = 2$, when the first term in brackets in eq. (35) should be neglected and G is dominated by the second term. This means that the point $g = 2$ can be viewed as the starting point for the formation of the conductance peak at larger energies. To describe this peak we can expand $J(\tau)$ at small τ in eq. (27) and get:

$$G = \frac{e^2\lambda_1\Gamma_1 e^{-\gamma}}{(2\pi\hbar X)^{1/2}}e^{-\frac{(\delta-\Omega g)^2}{2X}}, \quad X \equiv \int_0^\infty F(\omega)\omega d\omega\,. \tag{36}$$

We see that the new conductance peak is formed at the energy of the resonant level shifted by the charging energy, $\varepsilon = \varepsilon_0 + E_c$ and has a width $g^{1/2}\Omega$. Hence, it becomes well-defined at $g \gg \max\{1,(\Gamma_2/\Omega)^2\}$.

Figure 2 shows numerically calculated conductance of the resonant level coupled to the Ohmic environment (without the approximation of exponential cutoff), which exhibits transition from the peak at $\varepsilon = \varepsilon_0$ to $\varepsilon = \varepsilon_0 + E_c$ with increasing g. The

formation of the new peak is described quantitatively by eq. (36), but is slightly slower because of the divergence of X (36) for the Ohmic environment.

Another interesting aspect of the above results is that, for intermediate values of g, the conductance is non-vanishing in the entire energy interval between $\varepsilon = \varepsilon_0$ and $\varepsilon = \varepsilon_0 + E_c$. This means that the width of the current step in this regime is determined by the charging energy E_c and can be much larger than the tunneling width of the resonant level Γ_2 or temperature T. Carrying out the integration in the first of eqs. (27) we get for the Ohmic environment:

$$I(\varepsilon) = e\Gamma_1[\frac{1}{2} - T\int_{-\infty}^{0} d\tau \frac{\sin[(\varepsilon_0 - \varepsilon)\tau + (\pi g/2)(e^{\Omega\tau} - 1)]}{\sinh[\pi T\tau]} \exp\{\Gamma_2\tau/2 + \mathrm{Re}[J(\tau)]\}] \,. \tag{37}$$

Figure 3 show the current (37) for several values of the temperature. At $T \ll E_c$ the width of the current step saturates, and the current step becomes asymmetric. At first the current rises sharply on the energy scale Γ_2 at $\varepsilon = \varepsilon_0$ and then increases monotonously and levels-off on the scale $E_c \gg \Gamma$ (in accordance with eq. (33)). As will be discussed in the Conclusion, these features might be relevant for interpretation of some aspects of experiments on resonant tunneling in quantum dots.

V. FILLED BANDS

As we saw in the previous sections, the main characteristic feature of the one-electron approximation is that the tunneling width of the resonant level is unaffected by interaction. As a result, this approximation is only valid when the energy band in at least one of the external electrodes is empty. In this section we turn to the case when the energy bands in both electrodes are filled, so that the one-electron approximation would give qualitatively incorrect results.

It will be shown below that renormalization of the tunneling width of the resonant level by interaction plays a crucial role in this case, and has two major consequences. Firstly, at small coupling constants g, the value of conductance at resonance is not suppressed by the interaction but the width of the resonance decreases with increasing g. The second consequence is that at sufficiently large g the effective tunneling width Γ vanishes and a Coulomb gap is formed in the position of resonant level as a function of its bare energy ε_0. The Coulomb gap suppresses the conductance.

We consider at first the situation when bosonic modes are coupled only to the resonant level itself (i.e., $Z_V(\omega) = 0$ and $\lambda_1 = \lambda_2$). A qualitative understanding of this situation can be obtained from the relation between the zero-temperature linear conductance G of the resonant level and its occupation probability n [14]:

$$G = \frac{e^2}{2\pi\hbar}\frac{\Gamma_1\Gamma_2}{\Gamma^2}\sin^2\pi n\,. \tag{38}$$

Equation (38) shows that if n is a *continuous* function of energy ε_0 of the resonant level, then the conductance reaches at $n = 1/2$ the same maximum value as without interaction. The fact that the resonant occupation probability is fixed at $1/2$ is a manifestation of the electron-hole symmetry of the problem. Conversely, eq. (38) implies that the maximum value of the conductance can be suppressed only if n becomes a discontinuous function of ε_0, i.e. if a Coulomb gap appears in the effective renormalized energy of the resonant level.

Quantitatively, we can first analyze the problem at small g when a direct perturbation calculation is possible. This calculation shows that suppression of the linear conductance due to the suppression of the elastic transmission probability $P(\epsilon, \epsilon)$, which is a hallmark of the one-electron approximation (elastic transmission probability is suppressed due to appearance of the inelastic channels), is precisely

compensated for by corrections arising from deviations of the boson evolution from free evolution discussed in Sec. 3.

To go beyond the perturbation calculations we discuss at first the one-mode environment with frequency Ω that is larger than all frequencies associated with electron tunneling: ε_0, eV, T, $\Gamma \ll \Omega$. In this case, the fast boson mode adjusts itself instantly to the slow electron tunneling and we can employ the adiabatic approximation and average the total Hamiltonian of the system over the fast motion of the boson mode to obtain an effective electron Hamiltonian. As a result of such averaging we get that the tunneling amplitudes (2) are renormalized as follows:

$$t_j^* = \langle t_j e^{i\phi_j} \rangle = t_j e^{-g/2} . \tag{39}$$

This means that at low temperatures $T \ll \Gamma^*$, conductance should exhibit the same Lorentzian peak as in the noninteracting case but with renormalized width Γ^*:

$$G = \frac{e^2 \Gamma_1^* \Gamma_2^*}{2\pi\hbar} \frac{1}{(\varepsilon - \varepsilon_0)^2 + (\Gamma^*)^2} , \tag{40}$$

where $\Gamma_j^* = \Gamma_j e^{-g}$. We see that in accordance with our qualitative discussion, the conductance reaches a maximum value at $\varepsilon = \varepsilon_0$ which is the same maximum as in the non-interacting case.

Away from the resonance, when $| \varepsilon - \varepsilon_0 | \gg \Gamma^*$, the conductance can be calculated by perturbation theory in tunneling. If the Fermi levels in external electrodes are below ε_0, then the perturbation expansion should be performed starting with the empty resonant level as the zero-order approximation. Following the standard calculation of the second-order tunneling (see, e.g., [15]), we get, naturally, the same eqs. (20)–(22) for the current, without the tunneling width Γ in the definition of the transition amplitude A (22). However, when the Fermi levels are above ε_0, the perturbation expansion should be performed around occupied resonant level. The

perturbation theory gives then the same expression for the current (without Γ), but with the reversed sign of the energy difference $\varepsilon - \varepsilon_0$ in the expression for A (22). This change of sign reflects the fact that for $\varepsilon > \varepsilon_0$ the tunneling is more naturally interpreted as a tunneling of holes and not electrons (see the discussion below).

Expanding the exponent in eq. (26) and integrating each term of the series over t we get that only the first term contributes to the zero-temperature linear conductance G. Taking into account that for $\varepsilon > \varepsilon_0$ we should change the sign of $\varepsilon - \varepsilon_0$ we get for G:

$$G(\varepsilon) = \frac{e^2\Gamma_1\Gamma_2}{2\pi\hbar} \mid \int_{-\infty}^{0} d\tau e^{i|\varepsilon-\varepsilon_0|} e^{J(\tau)} \mid^2 , \tag{41}$$

where $J(\tau)$ is given by eq. (27) (note that for $Z_V(\omega) = 0$ and $\lambda_1 = \lambda_2$ all factors F coincide, $F_{11} = F_{22} = F_{12} \equiv F$). Taking the small-$\tau$ limit of $J(\tau)$ in eq. (41) we get an asymptotic expression for the conductance for $\mid \varepsilon - \varepsilon_0 \mid + E_c \gg \Omega$:

$$G(\varepsilon) = \frac{e^2}{2\pi\hbar} \frac{\Gamma_1\Gamma_2}{(\mid \varepsilon - \varepsilon_0 \mid + E_c)^2} . \tag{42}$$

It is interesting to note that if we had directly applied one-electron approximation (expressed by eqs. (20) - (22)) both for $\varepsilon < \varepsilon_0$ and $\varepsilon > \varepsilon_0$, we would get the same eq. (42) but with $\varepsilon_0 - \varepsilon$ instead of $\mid \varepsilon - \varepsilon_0 \mid$. Equation (42) would imply then that there is a resonant peak of the linear conductance at $\varepsilon = \varepsilon_0 + E_c$. Simple physical reasoning demonstrate why this conclusion is incorrect. To see this, notice that one of the assumption behind the one-electron approximation is that the photon modes are in equilibrium when the resonant level is empty, so that electron tunneling to the level displaces them out of equilibrium. Such a displacement results in the shift of the effective resonant level energy from ε_0 to $\varepsilon_0 + E_c$. The correct picture for $\varepsilon > \varepsilon_0$ is, however, that the photon modes are in equilibrium when the resonant level is occupied, and their displacement is caused by electron tunneling *from* the level or, in other words, tunneling of holes to the level. Thus, at $\varepsilon > \varepsilon_0$

the tunneling holes experience the same influence of the interaction with photons as electrons at $\varepsilon < \varepsilon_0$.

To complete our analysis of the *one-mode environment* we note that eq. (41) is reduced in this case to:

$$G = G_0 e^{-2g} \mid \sum_{n=0}^{\infty} \frac{g^n}{n!} \frac{1}{z+n} \mid^2 = G_0 e^{-2g} \Gamma^2(z) \gamma^*(z,-g)^2 , \quad G_0 = \frac{e^2 \Gamma_1 \Gamma_2}{2\pi\hbar\Omega^2} , \quad (43)$$

where $z \equiv \mid \varepsilon - \varepsilon_0 \mid /\Omega$, and γ^* is another incomplete gamma function [16]. At $g \gg 1$ and $\mid \varepsilon - \varepsilon_0 \mid \gg E_c e^{-2g}$ the sum in eq. (43) is dominated by the terms around $n \simeq g$ and eq. (43) coincides with eq. (42). At small energies $\mid \varepsilon - \varepsilon_0 \mid \ll E_c e^{-2g}$ the sum is dominated by the first term ($n = 0$), so that $G(\varepsilon) = (e^2/2\pi\hbar)\Gamma_1\Gamma_2 e^{-2g}/(\varepsilon - \varepsilon_0)^2$. Since this expression coincides with the result (40) of adiabatic approximation in the wide energy range, $\Gamma e^{-2g} \ll \mid \varepsilon - \varepsilon_0 \mid \ll E_c e^{-2g}$, we can conclude that the combination of eqs. (40) and (43) gives the $G(\varepsilon)$ dependence that is valid for all energies.

These results show that in the limit $g \to \infty$ eq. (42) becomes exact at all energies and describes the Coulomb gap appearing in the renormalized energy of the resonant level. Namely, eq. (42) implies that when the bare energy is above the Fermi levels in the electrodes, $\varepsilon_0 \to +0$ the renormalized energy is $\varepsilon_0 + E_c$, but as soon as it is below the Fermi level, $\varepsilon_0 \to -0$, the renormalized energy is $-(\varepsilon_0 + E_c)$. The discontinuity of the renormalized energy at $\varepsilon_0 = 0$ is a result of the discontinuity of the occupation probability n of the resonant level at this energy.

For *Ohmic environment*, the renormalization of the tunneling width can be written following (39) as $\Gamma^* = \Gamma(\Gamma/\Omega)^{g/(1-g)}$ if $\Gamma \ll \Omega$. It means that Γ is renormalized to zero at $g > 1$. In this case the perturbation expression (41) will be valid at all energies. With the exponential cutoff this expression gives for conductance:

$$G = G_0 e^{2z} z^{2(g-1)} \Gamma^2(1-g,z) = G_0 e^{2z} [\Gamma(1-g) z^{g-1} - \sum_{n=0} \frac{(-1)^n}{n!} \frac{z^n}{1-g+n}]^2 . \quad (44)$$

For Ohmic spectrum with Lorentzian cutoff we can use the asymptotic form of $J(\tau)$ for $\tau \gg \Omega^{-1}$ to find the conductance at small energy ($z \ll 1$), and $g < 1$:

$$G = G_0 e^{-2\gamma g} \Gamma^2 (1-g) z^{2(g-1)} . \tag{45}$$

Equation (45) agrees with eq. (44) at $z \ll 1$, $g < 1$. At arbitrary energy and g, the conductance can be calculated numerically, and is shown in Fig. 4. This figure confirms that the evolution of the conductance with g has a transition point at $g = 1$ (as indicated also by eqs. (44), (45)). For $g < 1$ the conductance has a peak at $\varepsilon \simeq \varepsilon_0$, while for $g > 1$ the peak is completely suppressed. At the transition point $g = 1$ we get from eq. (44):

$$G = G_0 (\ln z + \gamma)^2 . \tag{46}$$

For $g > 1$ the form of the conductance curves agrees qualitatively with (42) which gives $G = G_0/(g-1)^2$ at small z. At large g this expression also agrees with eq. (42).

These results imply that in the case of Ohmic environment the Coulomb gap is formed at finite g, namely, $g = 1$. The reason for this is the logarithmic divergence of the bosonic correlators $J(\tau)$ in this case. Thus, formation of the Coulomb gap for Ohmic environment is a zero-temperature first-order phase transition. The physical picture of the Coulomb gap is the same as discussed above for the one-mode case. The Coulomb gap provides a physical explanation of the low-energy singularity obtained recently [17] in an impurity model similar to our model.

We now turn to the case of the bias voltage fluctuations ($Z_V(\omega) \neq 0$, $Z_g(\omega) = 0$). In this case the photon modes are coupled to the charge of the external electrodes, and the picture of tunneling is much closer to the usual Coulomb blockade in small tunnel junctions than in the case of the resonant level fluctuations. In the latter case electron interact with the photon modes only when it occupies the resonant level

during the tunneling, while the states of the modes before and after the tunneling are the same. In contrast to this, in the case of voltage fluctuations initial and final states of the photon modes are different because the tunneling changes the charge of external electrodes. This leads to stronger suppression of conductance, in particular, the linear conductance is completely suppressed at $T = 0$, and the current-voltage characteristic has a power-law singularity at small voltages.

To calculate the current in the small-voltage limit we notice that in this limit the relevant time scale for integration over t in eq. (26) is much larger than the time scales for τ and τ'. (The small-voltage limit implies that $\epsilon' - \epsilon \ll \mid \epsilon_0 - \epsilon \mid$.) Neglecting τ and τ' in comparison with t we obtain:

$$P(\epsilon, \epsilon') = \frac{1}{2\pi} \int dt e^{i(\epsilon' - \epsilon)t} e^{J(t)} \mid \int_{-\infty}^{0} d\tau e^{i|\epsilon_0 - \epsilon|\tau} e^{-\lambda(1-\lambda)J(\tau)} \mid^2 , \qquad (47)$$

where $J(\tau)$ is given by eq. (27) with $F(\omega) = F_V(\omega)$.

We see that expression (47) for the transition rate factorizes at small voltages in the two terms. One describes its voltage dependence (dependence on $\epsilon - \epsilon'$ in eq. (47)) and coincides with the similar term for the first-order tunneling in small tunnel junctions [6]. The second term has the meaning of the transition matrix element and coincides with the transition element (41) in the case of the resonant level fluctuations. Combining these two terms we get, for example, for Ohmic environment:

$$I(V) = \frac{e\hbar\Gamma_1\Gamma_2}{2\pi} \frac{e^{-\gamma(g-2\bar{g})}\Gamma^2(1+\bar{g})}{\Omega^{(g-2\bar{g})}\Gamma(1+g)} \frac{\mid eV \mid^{(1+g)}}{\mid \epsilon - \epsilon_0 \mid^{2(1+\bar{g})}} , \quad \bar{g} = \lambda(1-\lambda)g . \qquad (48)$$

We see that the term responsible for the voltage dependence of the current suppresses the current uniformly for all ϵ, while the second term sharpens the resonance suppressing the current for ϵ away from ϵ_0. Thus, in contrast to the resonant level fluctuations which suppress the resonance, fluctuations of the voltage make the resonance sharper with increasing interaction strength.

VI. CONCLUSION

In conclusion, we have considered resonant tunneling of electrons interacting with fluctuations of electromagnetic field or some other dissipative environment. Broadly speaking, interaction changes position and shape of the resonance. Our calculations might be relevant for experiments on electron transport in quantum dots. In particular, the one-electron approximation used in Sec. 3, is adequate for vertical double-barrier heterostructures similar, e.g., to those studied in [1], [2], [4]. In such structures the current appears only at large bias voltages, at which both conditions of applicability of the one-electron approximation are satisfied: first, the relevant electron states in the collector electrode are empty, and second, transparency of the collector barrier is much larger than that of the emitter barrier.

One interesting, and so far unexplained result of experiment [1] is the large low-temperature width of the current step that corresponds to the tunneling through the lowest 0D state in the quantum dot. In this experiment conditions (24) of applicability of the one-electron approximation were not, strictly speaking, satisfied: the rate of tunneling through the emitter barrier Γ_1 was not always smaller than the collector tunneling rate Γ_2. However, the ratio Γ_1/Γ_2 was at most on the order of one, so that one-electron approximation should be at least qualitatively valid. Then the results obtained above in Sec. 3 are applicable and suggest that a possible explanation of the extra width of the current step could lie in a moderately strong ($g \simeq 1$) interaction of electrons with some dissipative environment, since, as was shown in Sec. 3, at $g \simeq 1$ the width of the current step is expected to be roughly equal to E_c (i.e., much larger than the tunneling width Γ). In the vertical double-barrier heterostructures, the most probable candidate for dissipative environment interacting strongly with tunneling electrons are the charged dopants

in the depletion layer of the collector electrode. For interaction with dopants in the depletion layer the energy E_c is the difference between resonant level energy under the two conditions, when the charges in the depletion layer are frozen, and when they are allowed to rearrange to screen the electron on the resonant level. Qualitatively, the reason for the widening of the current step in this situation can be understood as follows. The current appears at voltages at which electron can tunnel through the resonant level when the level energy is lowered to ϵ_0 by the rearrangement of the environment in response to the transfer of electron to the level. However, the current saturates only at voltages when electron can tunnel even if the environment is not responding to electron transfer so that the effective energy of the resonant level is $\epsilon_0 + E_c$. An additional argument in favor of such an explanation is the asymmetric shape of the current step, with steeper rise of the current at small voltage and more gradual leveling-off at larger voltages [1]. As was discussed in Sec. 3, a similar shape of the current step results from the interaction with environment with Ohmic spectral density of excitations.

Another aspect of our results is that they provide a specific example of the situation when the calculation of conductance is a multi-electron problem despite the fact that there is no direct electron-electron interaction in the Hamiltonian (3). This implies that the frequently used ideology of the "generalized Landauer formula" based on the solution of the one-electron problem for a given bosonic field [18], [19], is not valid as a general approach, and establishing conditions of its validity is an important problem.

ACKNOWLEDGMENTS

Useful discussions with V.J. Goldman and K.K. Likharev are gratefully acknowledged. One of us (V.P.) appreciate useful discussions with R. Suris at the beginning

of this work and with B. Spivak at the completion of it. This work was supported by ONR Grant # N00014-93-1-0880.

REFERENCES

[1] B. Su, V.J. Goldman, and J.E. Cunningham, Science **255**, 313 (1992); Phys. Rev. B **46**, 7644 (1992); and private communication.

[2] P. Gueret, N. Blank, R. Germann, and H. Rothuizen, Phys. Rev. Lett. **68**, 1896 (1992).

[3] A.T. Johnson, L.P. Kouwenhoven, W. de Jong, N.C. van der Vaart, C.J.P.M. Harmans, and C.T. Foxon, Phys. Rev. Lett. **69**, 1592 (1992).

[4] M. Tewordt, L. Martin-Moreno, J.T. Nicholls, M. Pepper, M.J. Kelly, V.J. Law, D.A. Ritchie, J.E.F. Frost, and G.A.C. Jones, Phys. Rev. B **45**, 14407 (1992).

[5] E.B. Foxman, P.L. McEuen, U. Meirav, N.S. Wingreen, Y. Meir, P.A. Belk, N.R. Belk, M.A. Kastner, and S.J. Wind, Phys. Rev. B **47**, 10020 (1993).

[6] G.-L. Ingold and Yu.V. Nazarov, in: *Single Charge Tunneling*, ed. by H. Grabert and M. Devoret (Plenum, New York, 1992).

[7] L.I. Glazman and R.I. Shekhter, Sov. Phys. JETP **67**, 163 (1988).

[8] N.S. Wingreen, K.W. Jacobsen, and J.W. Wilkins, Phys. Rev. Lett. **61**, 1396 (1988).

[9] M. Jonson, Phys. Rev. B **39**, 5924 (1989).

[10] V.V. Ponomarenko, Europhys. Lett. **22**, 293 (1993).

[11] K.A. Matveev and A.I. Larkin, Phys. Rev. B **46**, 15337 (1992).

[12] Y. Meir and N.S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).

[13] G.D. Mahan, *Many-Particle Physics* (Plenum, New York, 1991).

[14] V.V. Ponomarenko, Phys. Rev. B **48**, 5265 (1993).

[15] D.V. Averin and Yu.V. Nazarov, in: *Single Charge tunneling*, ed. by H. Grabert and M. Devoret (Plenum, New York, 1992).

[16] ***Handbook of Mathematical Functions***, ed. by M. Abramowitz and I.A. Stegun (Dover, New York, 1972), Ch. 6.

[17] T. Giamarchi, C.M. Varma, A.E. Ruckenstein, P. Nozieres, Phys. Rev. Lett. **70**, 3967 (1993).

[18] S. Feng, Phys. Lett. A **143**, 400 (1990).

[19] F. Hekking, Yu.V. Nazarov, and G. Schön, Europhys. Lett. **20**, 255 (1992).

FIGURES

FIG. 1. Schematic diagram of the resonant tunneling structure considered in this work. $Z_g(\omega)$ and $Z(\omega)$ are effective impedances of the structure representing the density of the photon modes responsible for fluctuations of the resonant level energy, and fluctuations of the bias voltage V, respectively.

FIG. 2. Differential conductance of the resonant tunneling structure with the Ohmic spectrum of photon modes in the one-electron approximation. G_{max} denotes the maximum conductance in the non-interacting case, $G_{max} \equiv (2e^2\lambda\Gamma_1/\pi\hbar\Gamma_2)$. The curves illustrate he transition between low-energy and high-energy conductance peaks with increasing strength of the coupling to photon modes.

FIG. 3. DC current through the resonant tunneling structure with the Ohmic spectrum of photon modes for several values of the temperature T and coupling constant $g = 1$ – see eq. (37). In the low-temperature limit the width of the current step is determined by the charging energy E_c and is much larger than the tunneling width of the resonant level.

FIG. 4. Linear conductance of the resonant tunneling structure with Ohmic spectrum of the photon modes coupled to the resonant level. The plot shows suppression of the resonance at $\varepsilon = \varepsilon_0$ when the coupling strength is increased beyond $g = 1$. The conductance scale G_0 is defined as in eq. (43).

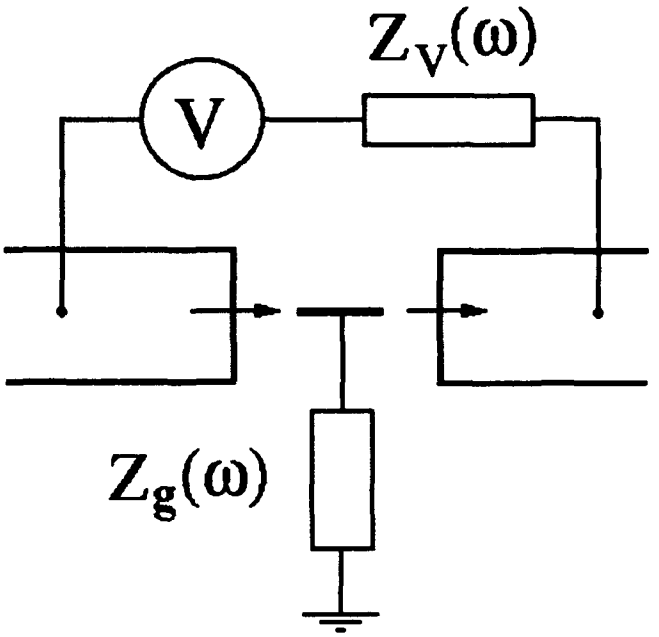

Fig. 1

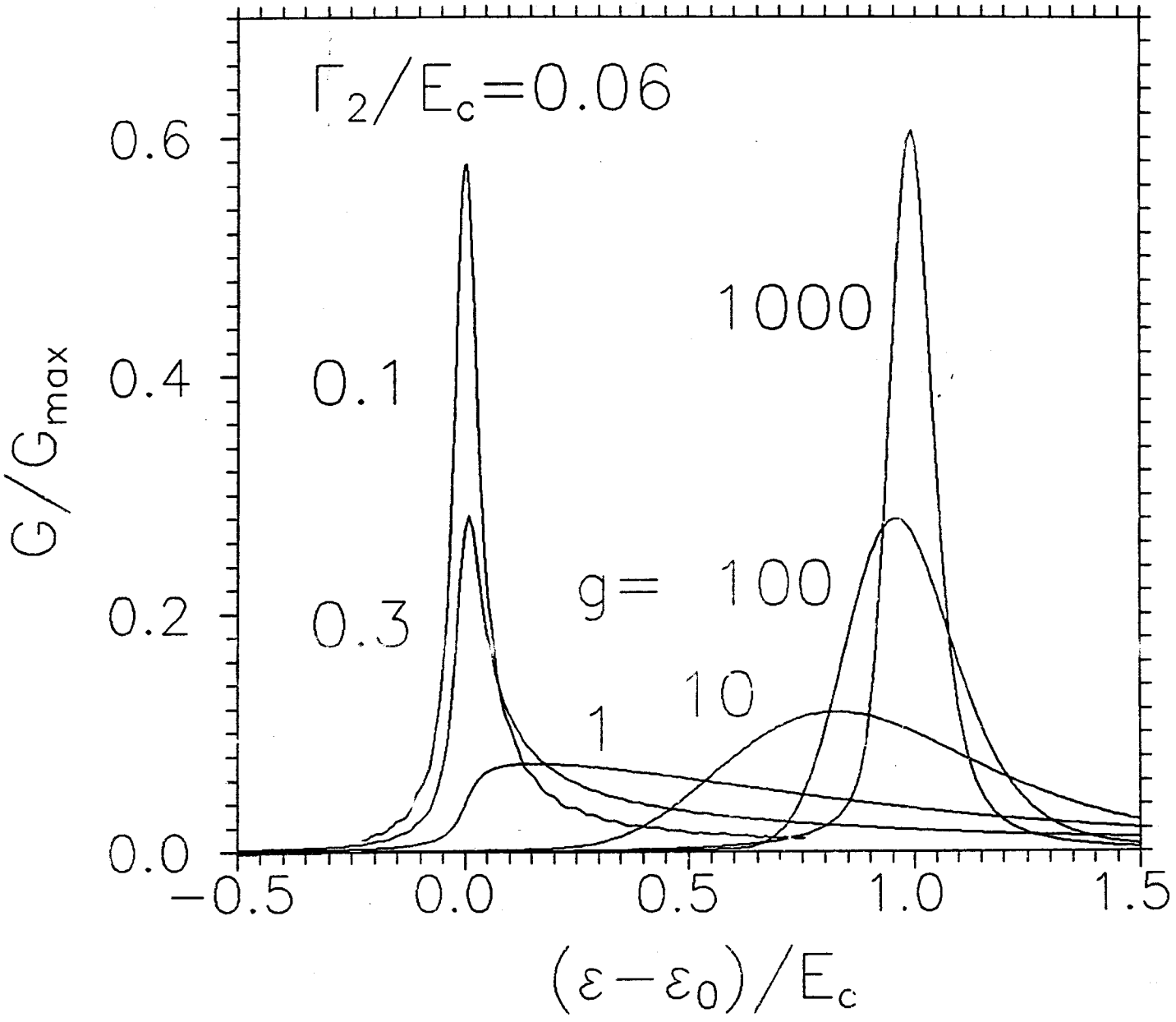
$\Gamma_2/E_c=0.06$
0.6
0.4
0.2
0.0
G/G_{max}
0.1
0.3
1000
g= 100
10
1
−0.5
0.0
0.5
1.0
1.5
$(\varepsilon-\varepsilon_0)/E_c$

Fig. 2

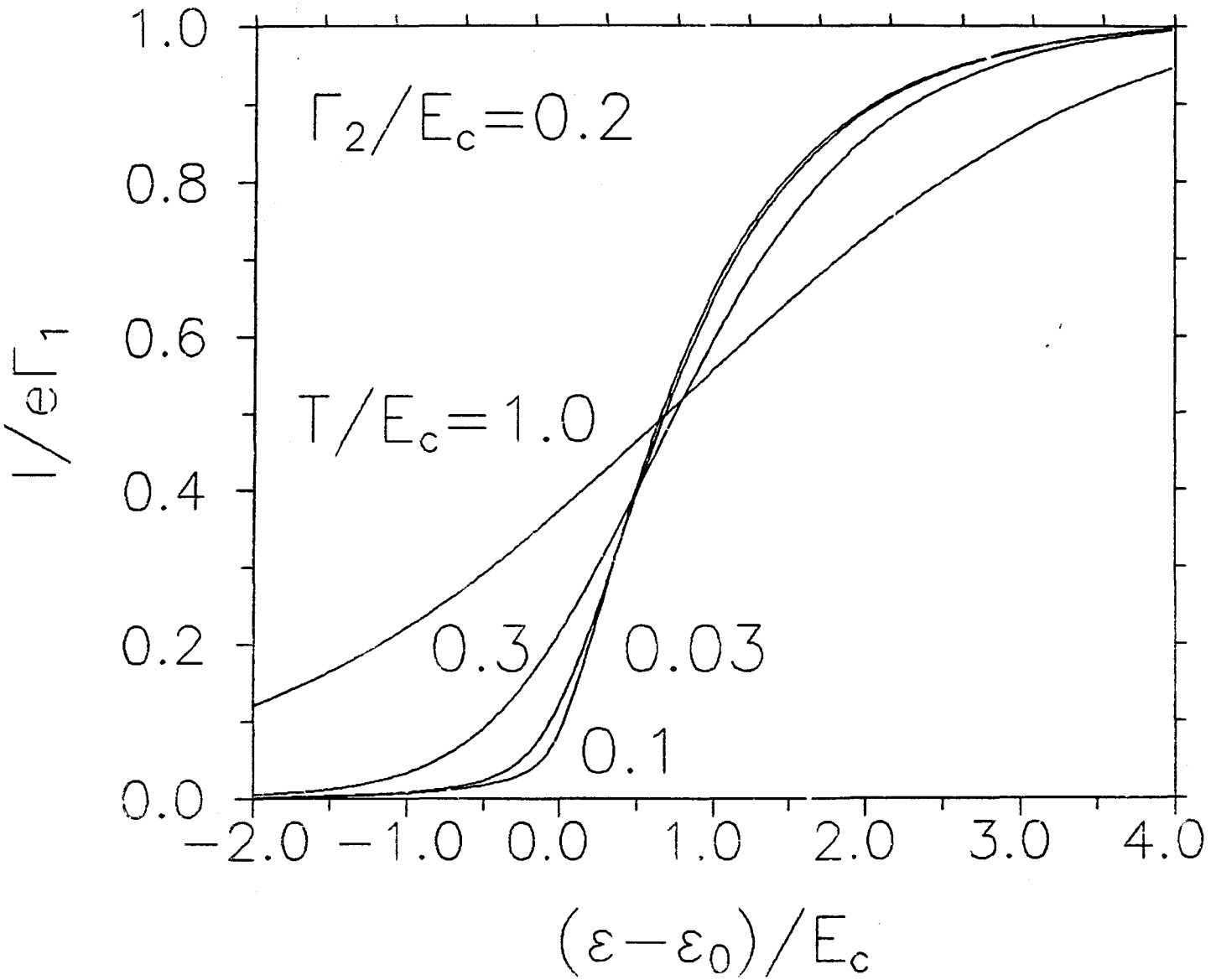
1.0
0.8
0.6
0.4
0.2
0.0
$\Gamma_2/E_c=0.2$
$T/E_c=1.0$
0.3
0.03
0.1
$I/e\Gamma_1$
−2.0 −1.0 0.0 1.0 2.0 3.0 4.0
$(\varepsilon-\varepsilon_0)/E_c$

Fig. 3

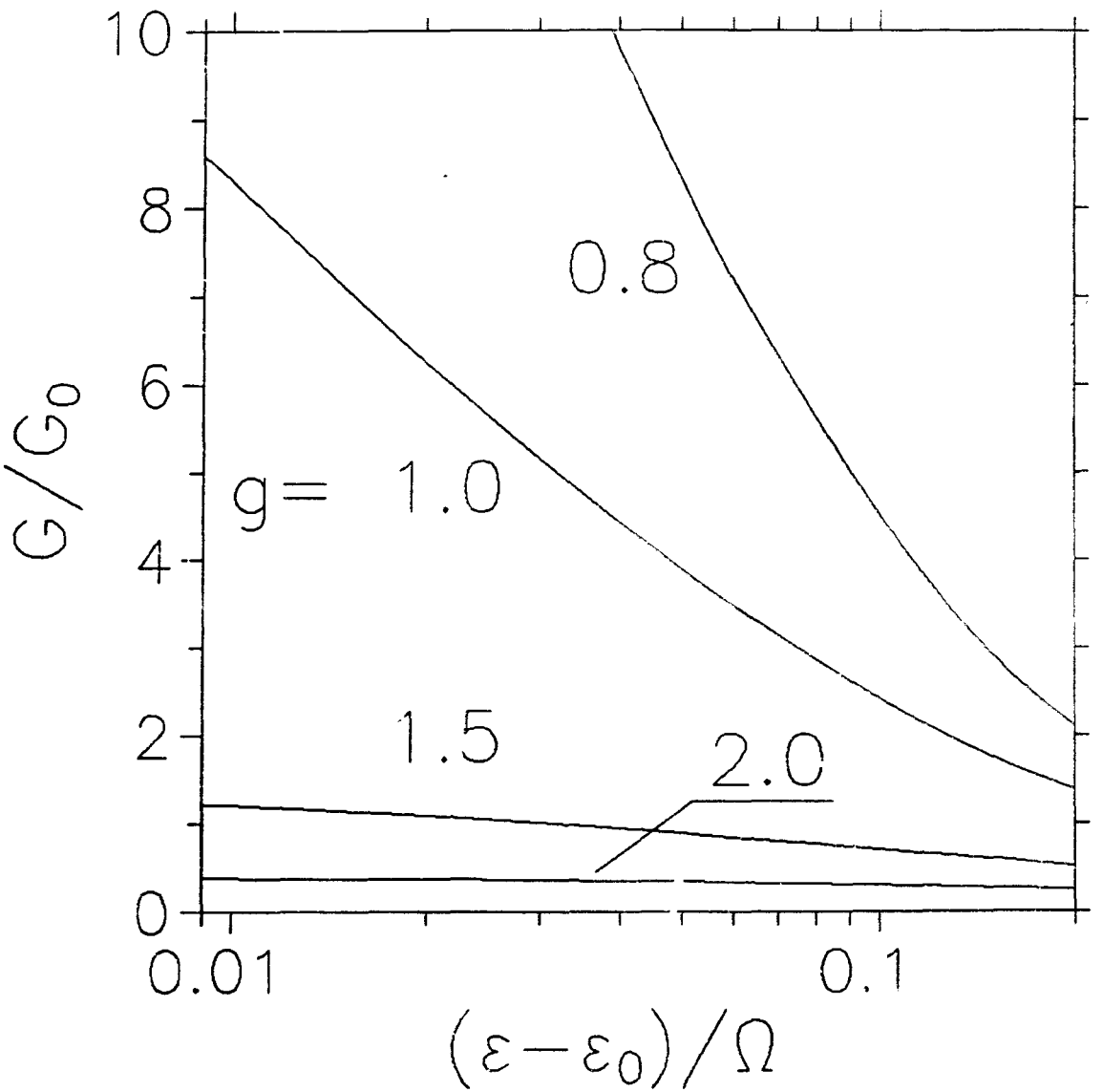

Fig. 4

Submitted to Phys. Rev. B

Effect of image charge on single-electron tunneling

A. N. Korotkov

Department of Physics, State University of New York,
Stony Brook, NY 11794-3800, USA,
and Institute of Nuclear Physics, Moscow State University,
Moscow 119899 GSP, Russia.

The existence of the image charge causes a modification of the tunnel barrier shape which depends on the effective junction capacitance C. The tunneling rate can be calculated using the expression of the "orthodox" theory of single-electron tunneling, with an additional prefactor of the order of $\exp(\tau e^2/\hbar C)$ where τ is the traversal time of tunneling.

Correlated tunneling in systems of ultrasmall tunnel junctions is a rapidly developing field of mesoscopic physics.[1-3] The simple "orthodox" theory of single-electron tunneling[1] provides the basis for theoretical analysis of these processes. There are, however, several effects not taken into account by "orthodox" theory.[3] In particular, in Refs. 4-6 the influence of the finite traversal time of tunneling τ was considered (in the "orthodox" theory, τ is assumed to be infinitesimal). It was shown that the effect becomes important when τ is of the order of $\hbar C/e^2$, where C is the junction capacitance.

The main focus of Refs. 4-6 was on the shape of the dc I-V curve. In contrast, in the present paper we calculate the multiplicative correction of the order of $\exp(\tau e^2/\hbar C)$ to the tunneling rate, which is independent (in our approximation) of the dc voltage applied to the system. The origin of this correction is the variation of the image charge at the edges of the tunnel junction.

First, consider a tunnel junction biased by a fixed dc voltage $V>0$. Assume that the temperature T is zero, then tunneling is possible only in one direction (say, from left to right). The effect of the image charge on the tunneling was considered in a number of papers (see, e.g., Ref. 7 and references therein). In the simplest model we assume that the image charge follows the position of the electron inside the barrier as in the static case. This so-called "static" image[7] model is valid when the frequencies of the surface plasmons are much larger than $1/\tau$. If we also assume that the Thomas-Fermi screening length in electrodes is much less than the thickness of the barrier, the image charge can be

calculated by a simple multiple reflection procedure (Fig. 1a), and the effective barrier shape is the sum of the initial shape $U_0(x)$ and the correction $U_{IM}(x)$ due to image charges.

Note that in this case the total image charges Q_l at the left side of the junction and Q_r at the right side depend linearly on the position x of the electron inside the barrier (measured, say, from the surface of the left electrode)

$$Q_l=(x/L-1)e, \qquad Q_r=(-x/L)e, \tag{1}$$

where L is the barrier thickness. In reality these charges are located at the electrode surfaces and supplied by the voltage source.

Now consider the same tunnel junction separated from the voltage source. Let the initial voltage V be greater than e/2C (after the tunneling of one electron this voltage becomes V-e/C). In contrast to the fixed voltage case, now the total charge of each electrode is fixed. Hence, in comparison with the previous case, there are additional charges $-Q_l-e$, $-Q_r$ uniformly distributed along the electrode surfaces (Fig. 1b). This leads to the additional electric field $E=(Q_r-Q_l-e)/2CL=-xe/(CL^2)$ which depends on the position of the electron. Hence, the effective barrier becomes $U_0(x)+U_{IM}(x)+U_{SET}(x)$ where

$$U_{SET}(x) = -e\int_0^x E(x')\,dx' = (x^2/L^2)(e^2/2C). \tag{2}$$

Let us emphasize that at the point x=L this additional energy coincides with the change $e^2/2C$ of the electrostatic energy of the

system after the tunneling. Because of this fact (which is valid only in the "static" image model), taking into account only the linear part of U_{SET},

$$U_{LIN}(x) = (x/L)(e^2/2C), \tag{3}$$

we would exactly reduce the case of the separated junction to the case of the tunnel junction biased by the fixed voltage V-e/2C. Then the tunneling rate Γ could be calculated using the expression

$$\Gamma = \frac{I(V-e/2C)}{e} \qquad (V>e/2C,\ T=0) \tag{4}$$

where I(V) is the I-V curve of the junction biased by a fixed voltage. Equation 4 exactly coincides with the equation used in the "orthodox" theory of single-electron tunneling.[1-3]

The correction to Eq. 4 is caused by the remaining part of the barrier modification

$$U_{CORR}(x) = U_{SET} - U_{LIN} = -(x/L)(1-x/L)(e^2/2C). \tag{5}$$

Assuming $U_{CORR} \ll U_0+U_{IM}$ and using the WKB approximation it is easy to calculate the tunneling rate:

$$\Gamma = K\,\frac{I(V-e/2C)}{e}\,, \tag{6}$$

$$K = \exp\left(-\frac{1}{\hbar}\int_0^L \left(\frac{2m}{U_0(x)+U_{IM}(x)}\right)^{1/2} U_{CORR}(x)\,dx\right). \tag{7}$$

The expression for the correction factor K depends on the barrier shape and in the general case cannot be exactly expressed

in terms of capacitance C and traversal time τ. However, for an estimate let us assume $U_0(x)+U_{IM}(x)=\text{const}$. This gives a simple expression

$$K = \exp\left(\frac{e^2}{6C}\frac{\tau}{\hbar}\right), \qquad \tau = L((U_0+U_{IM})/2m)^{-1/2}. \tag{8}$$

Thus, similar to the nonlinear effects considered in Refs. 4-6, the correction is essential when the traversal time τ is not too small in comparison with $e^2/\hbar C$. In our approximation this correction does not depend on the voltage.

Equations (5-8) can be easily extended to the case of a tunnel junction inside an arbitrary single-electron circuit containing other tunnel junctions, capacitances, voltage sources and resistances, with the only restriction (usual for the "orthodox" theory) that any resistance should be either much smaller or much greater than $\hbar/e^2$ and τ/C (in the most interesting case the last two values are of the same order). Then it is straightforward[3] to introduce the effective capacitance C_{eff} of the tunnel junction and the only change in Eqs. (5)-(8) is the substitution $C \to C_{eff}$. The effective capacitance is defined via the difference between the voltage $V_i=V$ before tunneling and the voltage V_f after tunneling,

$$C_{eff} = e/(V_i-V_f). \tag{9}$$

For example, in the system of two junctions connected in series (the "single-electron transistor") the effective capacitance is the sum of the junction capacitances, $C_{eff}=C_1+C_2$.

The simple substitution $C \to C_{eff}$ in Eqs. (5)-(8) is possible

only if the circuit size is much less than τc (c is the speed of light), so it will not be valid for too large circuits. In this case as well as for arbitrary resistances in the circuit, a more complicated theory based on the equations of Refs. 4-5 is necessary.

Generalization to the case of finite temperature T is also quite simple because the barrier change $U_{CORR}(x)$ does not depend on the temperature. The general expression

$$\Gamma = K\, I(V^*)/(1-\exp(-eV^*/T)), \qquad (10)$$

$$V^* = V - e/2C_{eff} = \frac{1}{2}(V_i+V_f), \quad K = K(C_{eff})$$

is similar to that of the "orthodox" theory; the only difference is the prefactor K. The existence of this prefactor depending on the effective capacitance of the junction is the main point of the present paper.

Now let us discuss the possibility of observing the considered effect in experiment. The simplest way is to compare the dc I-V curve of the single tunnel junction biased by a fixed voltage and the dc I-V curve of the double-junction system. At $\tau \sim \hbar(C_1+C_2)/e^2$ the current in the double-junction system should be larger than that predicted by "orthodox" theory. The simplest check is to compare the low-voltage resistance of one junction with the differential resistance of the double-junction system at the voltage just above the Coulomb blockade threshold. In "orthodox" theory these two values coincide, if the background charge is not close to zero and the temperature is low.

Let us estimate possible experimental parameters. For the tunnel junctions metal-insulator-metal the typical traversal time τ is about $3 \cdot 10^{-15}$ s. The correction factor K in this case is essential for $e/C > 0.3$ V. Hence, in principle, the effect can be observed using the scanning tunneling microscope.[1-3] However, in this case it is practically impossible to prepare identical tunnel junctions for single-junction and double-junction experiments.

The traversal time in semiconductor tunnel junctions can be made much longer by the use of low tunnel barriers. For τ long enough the model of "static" image charge can be applicable[8] in spite of the fact that the plasmon frequencies in semiconductors are much lower than those in metals. In Ref. 8 tunnel junctions having traversal time up to $3 \cdot 10^{-13}$ s were used. For our estimate, let us take the more moderate value $\tau = 10^{-13}$ s. Then for observation of the effect considered in the present paper, the typical voltage e/C should be of the order of $\hbar/\tau e \sim 7$ mV. Hence, the typical capacitance may be about $2 \cdot 10^{-17}$ F. Note that the condition $\tau \sim \hbar C_{eff}/e^2$ means that the typical voltage of the exponential nonlinearity of the I-V curve should be of the order of Coulomb blockade threshold.

In conclusion, we have found a correction to the tunneling rate which is used in the "orthodox" theory of single-electron tunneling. The correction is essential when $\tau e^2/\hbar C > 1$. It leads, in particular, to an increase of the current through the double-junction system in comparison with that calculated using the "orthodox" theory.

ACKNOWLEDGEMENTS

Fruitful discussions with D. V. Averin, K. K. Likharev and Yu. V. Nazarov are gratefully acknowledged. The work was supported in part by Russian Fund for Fundamental Research, Grant #93-02-14136 and ONR Grant #N00014-93-1-0880.

REFERENCES

[1] K.K. Likharev, IBM J. Res. Dev. **32**, 144 (1988).

[2] D.V. Averin and K.K. Likharev, in: *Mesoscopic Phenomena in Solids*, ed. by B. Altshuler *et al.* (Elsevier, Amsterdam, 1991), p. 173.

[3] *Single Charge Tunneling*, ed. by H. Grabert and M.H. Devoret (Plenum, New York, 1992).

[4] Yu.V. Nazarov, Solid State Commun. **75**, 669 (1990).

[5] Yu.V. Nazarov, Phys. Rev. B **43**, 6220 (1991).

[6] A.N. Korotkov and Yu.V. Nazarov, Physica B **173**, 217 (1991).

[7] B.N. Persson and A. Baratoff, Phys. Rev. B **38**, 9616 (1998).

[8] P. Gueret, E. Marclay, and H. Meier, Appl. Phys. Lett. **53**, 1617 (1988).

Figure caption

Fig. 1. (a) Image charges for voltage biased tunnel junction and (b) additional charges in the case of the separated junction.

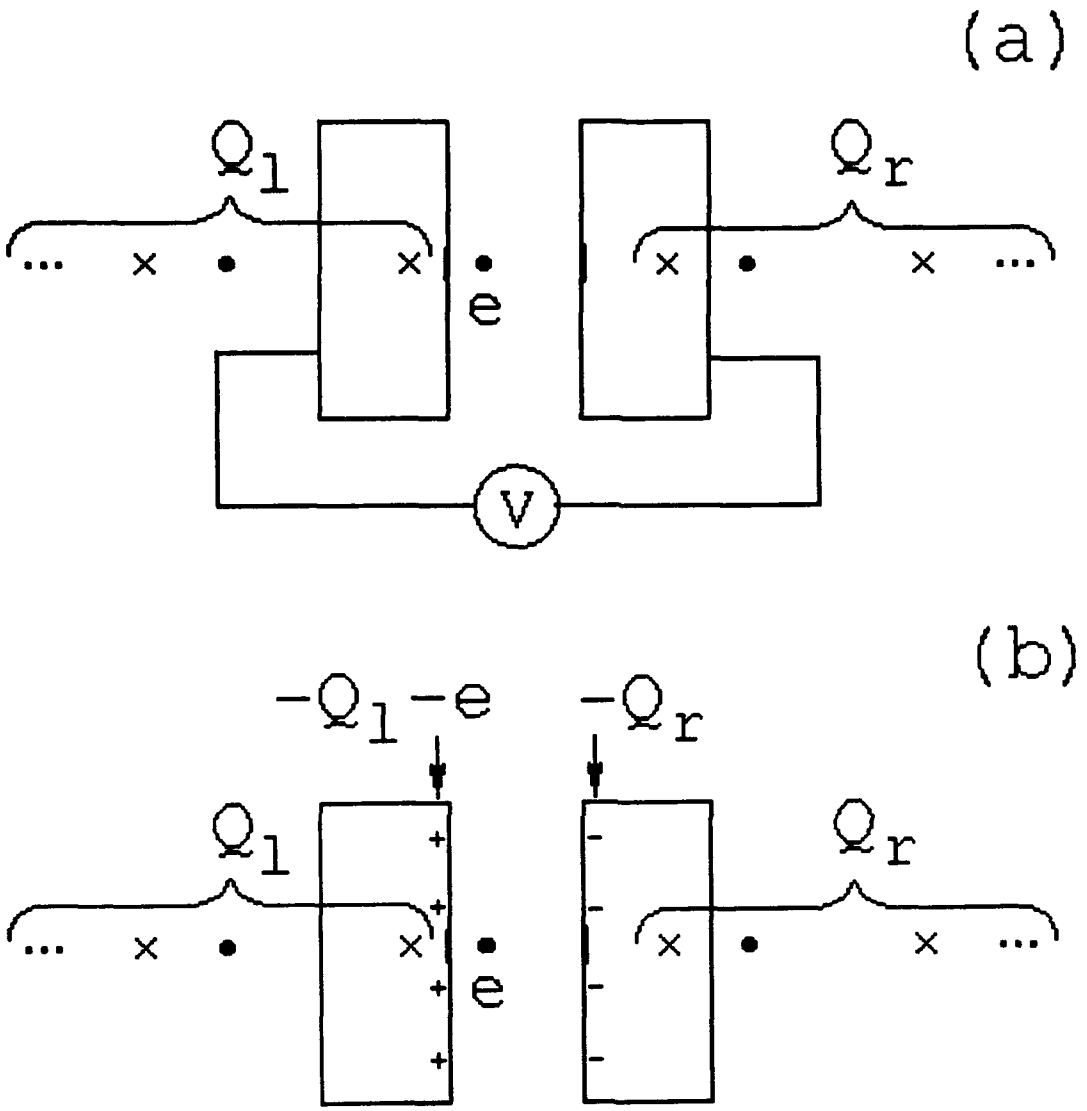

Fig. 1

Quantum and Many-Body Effects on the Capacitance of a Quantum Dot

Lotfi Belkhir

Department of Physics,
State University of New York,
Stony Brook, New York 11794-3800

Abstract

We calculate exactly, using finite size techniques, the quantum mechanical and many-body effects to the self-capacitance of a spherical quantum dot in the regime of extreme confinement, where the radius of the sphere is much smaller than the effective Bohr radius. We find that the self capacitance oscillates as a function of the number of electrons close to its classical value. We also find that the electrostatic energy as a function of the number of electrons extrapolates to zero when $N = 1$, suggesting that the energy scales like $e^2N(N-1)$ instead of $(N\,e)^2$. We also provide evidence that the main deviations from the semiclassical description are due to the exchange interaction between electrons. This establishes, at least for this configuration, that the semiclassical description of Coulomb charging effects in terms of capacitances holds to a good approximation even at very small scales.

PACS numbers: 73.20.Dx, 73.40.Gk, 72.15.Qm

With the rapid advances in the fabrication of increasingly smaller quantum dots, approaching the atomic scale, the question of quantum and many-body effects on the essential characteristics of these objects has become a central issue. The current stage of theoretical understanding of quantum dots relies essentially on a semi-classical picture[1,2], which makes the assumption that the Coulomb charging effects can be described in terms of classical capacitances, where the resonance energies can be separated in a single-particle confinement energy, and a constant Coulomb charging energy terms. The subject has been further studied in many recent theoretical[3–8] and experimental investigations[9–13]. The classical picture of Coulomb blockade, however, has recently been questioned by Johnson and Payne[14] who, using a harmonic model interaction that is exactly solvable[15], showed that, in presence of magnetic field, the resonance energies could not be written as the sum of single-particle and charging energies terms. They argued that the model interaction shows a behavior similar to a Coulomb interaction with a cutoff[16] for a certain range of electron-electron separation. On the other hand, however, the semiclassical description seems so far to provide a qualitatively correct picture, given that some experiments[17,13], performed in the regime where confinement and charging energies are of equal importance, can be well explained by this model.

We investigate this issue further, using exact finite size calculations techniques. We consider an isolated spherical quantum dot in zero magnetic field, and solve numerically the full Coulomb interaction problem, for up to 30 electrons, in the regime where the quantum effects are expected to be maximal, i.e when the radius of the dot is much smaller than the effective Bohr radius (i.e $R \ll a_0$). It is assumed that the added electrons move on the surface of the sphere, which could be a reasonable model for a metallic sphere. It has recently been shown[18], in the framework of density-functional theory, that in most experimental situations the main contribution to the capacitance of a quantum dot, in the presence of leads and backgates, comes from the self-capacitance. The contributions of the leads and backgates was found to be 30% at most. The results for an isolated dot are therefore not irrelevant to actual experiments.

We do the calculation for both spin unpolarized and spin polarized cases. We find that: (i) the interaction energy spectrum scales like $N(N-1)e^2/2C$, where N is the number of electrons at the surface of the dot, and R its radius. This corroborates the semi-classical expression for the charging energy suggested recently by some authors[19,20], instead of the more widely used expression $(Ne)^2/2C$. (ii) the resonance energies *do* separate into confinement and charging energies to a good approximation; (iii) the self-capacitance of the isolated dot, defined from the charging energy, oscillates as a function of N, around its classical value, i.e R, and gets closer to R as N increases. (iv) The main deviation of the self-capacitance from the classical value are due to the exchange interaction between electrons, which, in agreement with Hund's first rule, tend to lower the ground state energy of the system, and thus increase the value of the capacitance. This increase is however never greater than 25% for $N > 2$.

Our Hamiltonian for an isolated spherical dot, is given by

$$H = \frac{1}{2m^* R^2} \sum_i^N |\mathbf{L}_i|^2 + \sum_{i<j} \frac{e^2}{\epsilon |\mathbf{r}_i - \mathbf{r}_j|}, \tag{1}$$

where $\mathbf{L}_i = -i\hbar \mathbf{R}_i \times \nabla_i$ and $\mathbf{r}_i$ are the angular momentum and the position of the i-th

particle, and ϵ the dielectric constant. The eigenvalues of $|\mathbf{L}|^2$ are equal to $l(l+1)$ with l an integer. The kinetic energy of an electron in the shell l is $\epsilon_l = \hbar^2/(2m^*R^2)\ l(l+1)$. The maximum number of electrons in a shell of angular momentum l is $2l+1$ and $2(2l+1)$ for a spin polarized and spin unpolarized cases respectively.
The dot contains N electrons, with an effective mass m^*, and charge $-|e|$. Notice the absence of a confinement term, due to the use of a spherical geometry, where the electrons are constrained to move on the surface of the sphere of constant radius R. Previous quantum mechanical calculations of the capacitance of quantum dots[8,18] involved disc geometries with parabolic confinement potential. It was found in those calculations that the effective size of the dot increases with the number of electrons. in our case however, the size of the dot is absolutely rigid, which will facilitate considerably the comparison of our calculations with the classical results.
The Coulomb interaction between two electrons moving on the surface of a sphere with radius R can be rewritten as

$$V(\hat{\Omega}_1, \hat{\Omega}_2) = \frac{e^2}{R}\frac{1}{|\hat{\Omega}_1 - \hat{\Omega}_2|} = \frac{e^2}{R}\sum_{l,m}\left(\frac{4\pi}{2l+1}\right)\ Y^*_{lm}(\Omega_1)Y_{lm}(\Omega_2), \tag{2}$$

In second quantization form, the interaction operator is given by

$$\hat{V} = \frac{1}{2}\sum_{\sigma_1,\sigma_2}\int\ d\Omega_1\ d\Omega_2 \Psi^\dagger_{\sigma_1}(\Omega_1)\Psi^\dagger_{\sigma_2}(\Omega_2)V(\Omega_1,\Omega_2)\Psi_{\sigma_2}(\Omega_2)\Psi_{\sigma_1}(\Omega_1), \tag{3}$$

where the $Y^*_{lm}(\hat{\Omega})$ are the usual spherical harmonics, and $\Psi^\dagger_\sigma(\hat{\Omega}) = \sum_{l,m} Y_{lm}(\hat{\Omega})\ a^\dagger_{lm\sigma}$.
To carry out our numerical calculations, we choose the convenient single-particle basis states defined by

$$< \Omega | lm\sigma >= Y_{lm}(\Omega)\ \chi_\sigma. \tag{4}$$

In this basis the two body interaction operator is given by

$$\hat{V} = \sum_{allindices}\frac{1}{2} v_{l_1 m_1 l_2 m_2 l_3 m_3 l_4 m_4} c^\dagger_{l_2 m_2 \sigma} c^\dagger_{l_3 m_3 \sigma'} c_{l_4 m_4 \sigma'} c_{l_1 m_1 \sigma}, \tag{5}$$

where the matrix elements are given by

$$v_{l_1 m_1 l_2 m_2 l_3 m_3 l_4 m_4} = \sum_{L,M}(-1)^{l_1+l_4-l_2-l_3}\left[\frac{(2l_1+1)(2l_4+1)}{(2l_2+1)(2l_3+1)}\right] < l1, m1; L, M | l_3, m_3 > \tag{6}$$

$$< l_1, 0; L, 0 | l_3, 0 >< l_4, m_4; L, M | l_2, m_2 > < l_4, 0; L, 0 | l_2, 0 > \tag{7}$$

where the terms $< l, m; l_1, m_1 | l_2, m_2 >$ are the usual Clebsh-Gordon coefficients, which are non zero only when $m + m_1 - m_2 = 0$.
We also assume the limit of $R \to 0$, when the energy separation between successive angular momentum shells is large compared to the Coulomb interaction energy, so that mixing between shells can be neglected. This is equivalent to assuming that $R \ll a_0$, where

$a_0 = \frac{\epsilon\hbar^2}{m^* e^2}$ is the effective Bohr radius, determined solely by the material's properties. In GaAs quantum dots, $a_0 = 10nm$, and the confinement energy $\hbar^2/(2m^*R^2)$ is typically about 15/,meV, which for a spherical dot yields a radius $R = 8nm$. The strong confinement regime could be attained by either reducing the size of the dots, or using materials with a higher dielectric constant ϵ, which would increase the effective Bohr radius a_0. The assumption of $R \to 0$ drastically reduces the Hilbert size of the quantum system, since intershell transitions can be completely ignored, which allows us to do the calculations for up to 30 electrons. The calculation done in this limit is similar to the finite size calculations done in the context of the fractional quantum Hall effect[21], where the limit of infinite magnetic field is assumed in order to ignore transitions to higher Landau levels.

We calculate the energy spectrum by exact diagonalization of the Hamiltonian. The calculation is done by filling the angular momentum shells by adding electrons one by one, and calculating the ground state of the whole many-body system. In less than half-filled shell, all electrons tend to have the same spin in order to gain the exchange energy in accordance with Hunds' rule. As half-filling is reached, there is a sudden increase in the interaction energy due to the fact that the additional electrons must have opposite spins, in order to satisfy the Pauli principle, and thus lose the exchange energy. Our results for the ground state interaction energy $E_c(N)$ as a function of N, spin unpolarized case, are shown in figure 1. It shows an almost linear behavior, and extrapolates to 0 when $N = 1$, which is consistent with the semiclassical expression $N(N-1)e^2/R$.

We first define the chemical potential as

$$\mu(N) = E_c(N) - E_c(N-1) \tag{8}$$

We define the self-capacitance of our dot as $C = \Delta Q/\Delta V$ which can be readily obtained from the chemical potential. For a single electron $\Delta Q = e$ and $\Delta V = [\mu(N+1) - \mu(N)]/e$. Thus

$$C_{dot}(N) = e^2/[\mu(N+1) - \mu(N)] \tag{9}$$

In the case where tunneling occurs through the channel where the dot is at its ground state before and *after* the event, the resonance energies are identical to the chemical potential. Notice that the definition of the chemical potential is usually taken as the difference between the total energies, rather than the interaction energies. In our case, both definitions are identical for electrons in the same shell, since the kinetic energy within one shell is constant. The only difference occurs when N electrons correspond to a state of completely filled shells, and the $(N+1)$th electron must occupy the next empty shell. The jump in the chemical potential will then be incremented by the confinement energy, in addition to the charging energy. Since we are choosing the confinement energy to be extremely large, we always subtract it from the total energy to keep only the interaction part.

Figure (2a) and (2b) show the numerical results for the chemical potential, and the self-capacitance as a function of N in units of R for a spin unpolarized electron system. We see clearly that the capacitance is quite close to its classical value, modulo some quantum fluctuations. The peaks observed are due to the sudden increase in energy that occurs at half filling of each angular momentum shell. Notice also that, except for half filled shells, the quantum and many-body effects tend to increase the capacitance only slightly . The average

over all the 30 electrons, including the values at the peaks is $1.117R$. Figure (3a) and (3b) show the results for the chemical potential and the self-capacitance of the spin polarized case. Observe now how the capacitance has became consistently greater than its classical value. However the difference is not greater than 25% for all $N \leq 30$. The average capacitance for the spin polarized case over $N_{max} = 29$ is $1.123R$. Also notice that the deviation tends to decrease asymptotically as N increases. This deviation is clearly due to the exchange energy between electrons of same spin. However this exchange energy becomes smaller in higher angular momentum shells, which explains the asymptotic behavior of the self-capacitance. We also did the calculation for a system of spinless electrons (i.e no exchange interaction), and found that the capacitance become then precisely centered around its classical value R. In conclusion we have exactly calculated the energy spectrum and self-capacitance of a spherical quantum dot in the strong confinement limit where quantum effects are expected to be predominant. Remarkably, we have found that the semi-classical theory remains valid on average in this regime. We have also found that the main deviations from the semi-classical result are due to the exchange interaction between electrons, but that theses deviations do not exceed 25%.

We gratefully acknowledge insighteful conversations with J. K. Jain, and K. K. Likharev, and T. Kawamura. This work was supported in part by the office of Naval Research under grant No. N00014-93-1-0880

REFERENCES

[1] D. V. Averin and K. K. Likharev in '*Mesoscopic Phenomena in Solids*', edited by B. L. Altshuler, P. A. Lee and R. A. Webb, (Elsevier Science Publishers B.V., 1991).

[2] H. van Houten, C. W. Beenaker, and A. A. M. Staring, in '*Single Charge Tunneling*', edited by H. Grabert and M. H. Devoret (Plenum Press, New York, 1992).

[3] G. W. Bryant, *Phys. Rev. Lett.*, **59**, 1140 (1987)

[4] L. I. Glazman and R. I. Shekhter, *J. Phys. Cond. Matt.* **1**, 5811 (1989).

[5] A. Kumar, S. E. Laux, and F. Stern, *Phys. Rev. B.*, **42**, 1566 (1990).

[6] Y. Meir, N. S. Wingreen, and P. A. Lee, *Phys. Rev. Lett.*, **66**, 3048 (1991).

[7] A. Grochev, T. Ivanov, and V. Valtchinov, *Phys. Rev. Lett.*, **66**, 1082 (1991)

[8] M. Stopa, *Phys. Rev. B.*, **48**, 18340 (1993).

[9] U. Meirav, M. A. Kastner, and S. J. Wind, *Phys. Rev. Lett.*, **65**, 771 (1990).

[10] L. P. Kouwenhoven *et al.*, *Z. Phys. B.*, **85**, 367 (1991).

[11] U. Meirav *et al.*, *Z. Phys. B.*, **85**, 357 (1991).

[12] R. C. Ashoori *et al.*, *Phys. Rev. Lett.*, **68**, 3088 (1992).

[13] V. J. Goldman, B. Su and J. E. Cunningham, *Phys. Rev. B.*, **46**, 7644 (1993).

[14] N. F. Johnson and M. C. Payne, *Phys. Rev. B.*, **45**, 3819, 1992.

[15] N. F. Johnson and M. C. Payne, *Phys. Rev. Lett.*, **67**, 1157, 1991.

[16] L. Banyai, *et al.*, *Phys. Rev. B.*, **36**, 6099, 1987.

[17] P. L. McEuen, E. B. Foxman, U. Meirav, M. A. Kastner, Y. Meir, N.S. Wingreen and S. J. Wind, *Phys. Rev. Lett.*, **66**, 1926, 1991.

[18] M. Macucci, K. Hess, G. J. Iafrate, *Phys. Rev. B.*, **48**, 17354, 1993.

[19] C. S. Lent, in '*Proceedings of the International Symposium on Nanostructures and Mesoscopic Systems, Santa Fe* ', edited by W. P. Kirk and M. A. Reed (Academic, Boston, 1991).

[20] D. V. Averin, A. N. Korotov, and K. K. Likharev, *Phys. Rev. B.*, **44**, 6199, 1991.

[21] F.D.M. Haldane in *The Quantum Hall Effect*, Eds. R.E. Prange and S.M. Girvin (Springer Verlag, New York, second edition).

FIGURES

FIG. 1. Ground state interaction energy $E(N)/N$ as a function of N for the spin unpolarized case. Notice that it extrapolates to 0 when $N = 1$.

FIG. 2. (a)The chemical potential and (b)the Self Capacitance of a spherical quantum dot as a function of N for the spin unpolarized electron system.

FIG. 3. (a)The chemical potential and (b)the Self Capacitance of a spherical quantum dot as a function of N for the spin polarized electron system.

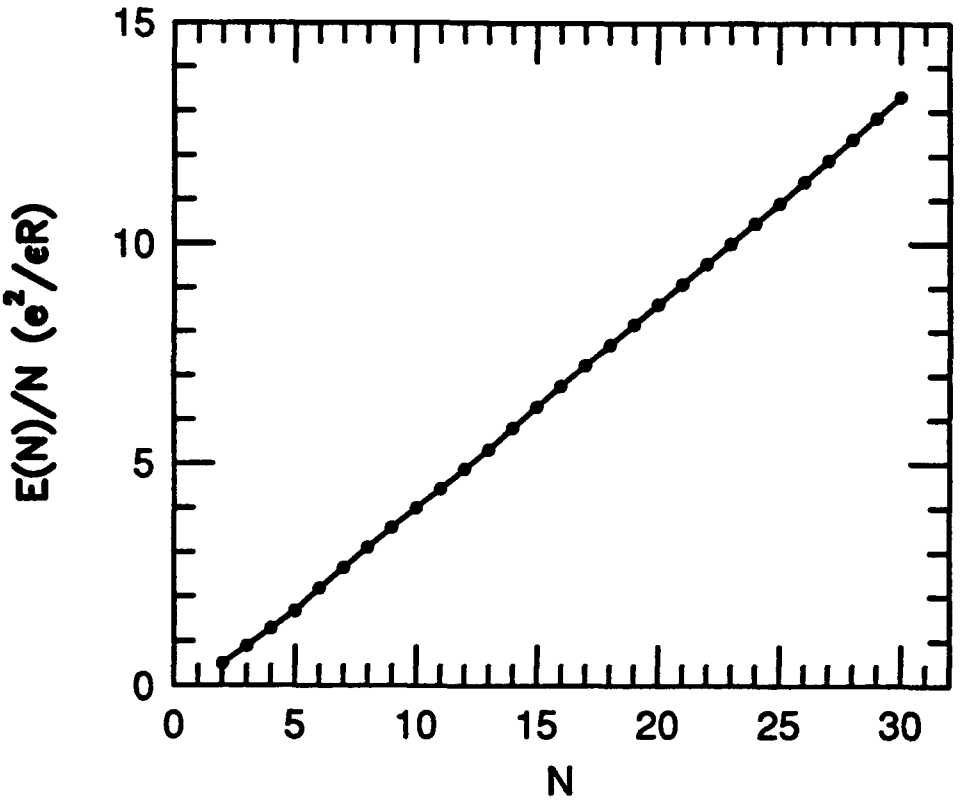

Fig 1.

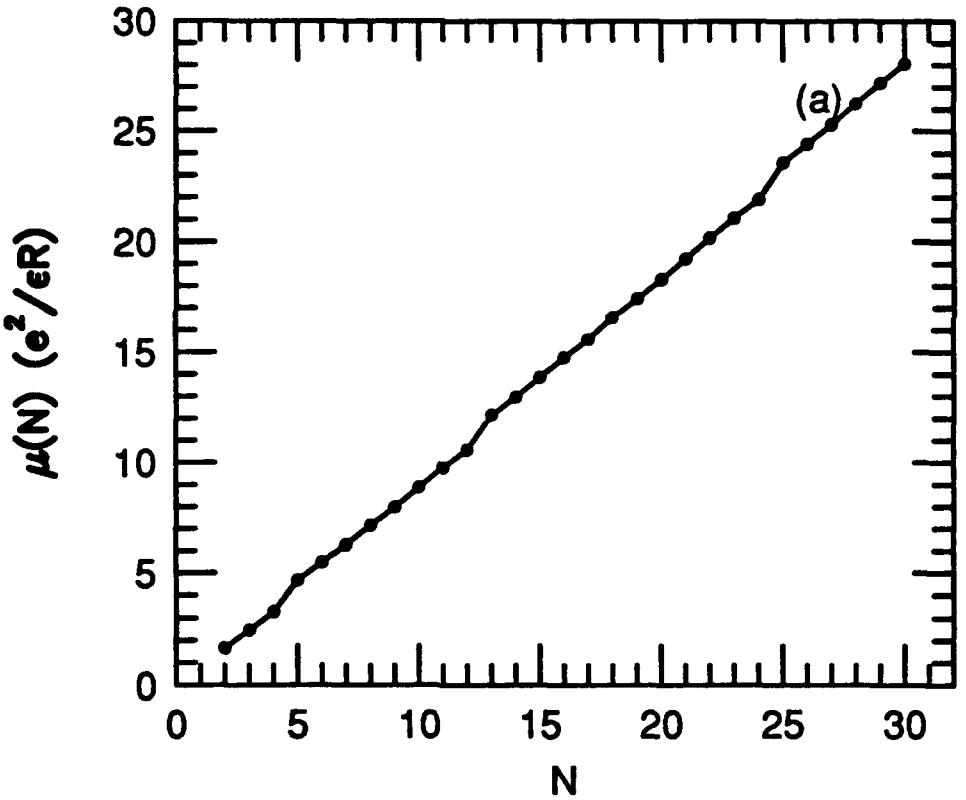

fig 2.a

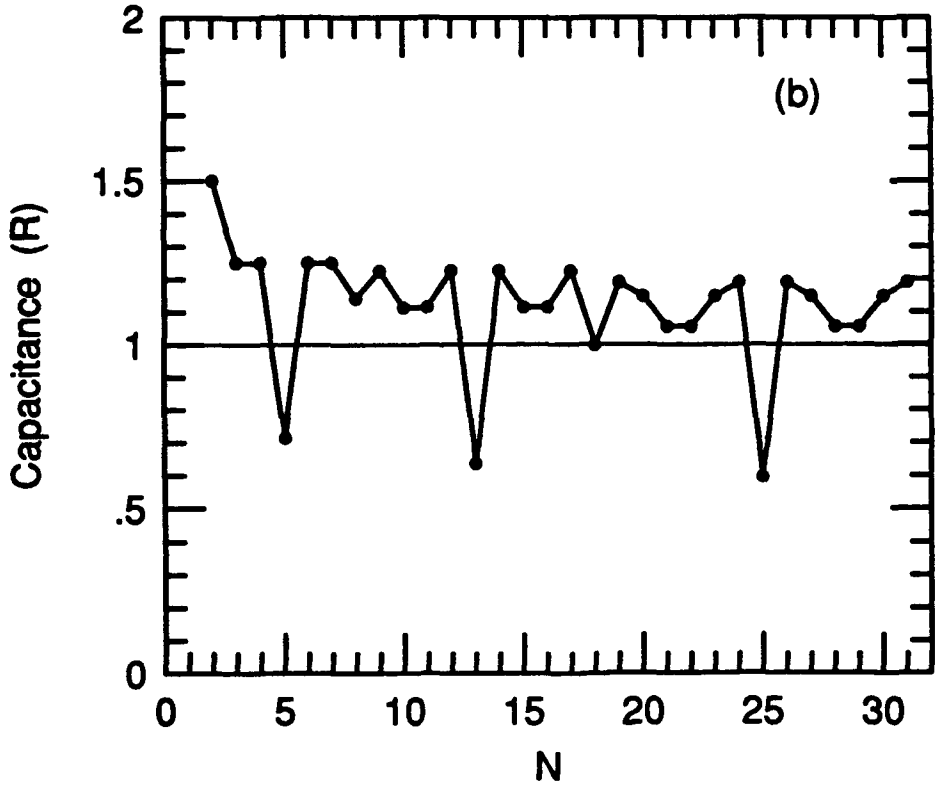

Fig 2.b

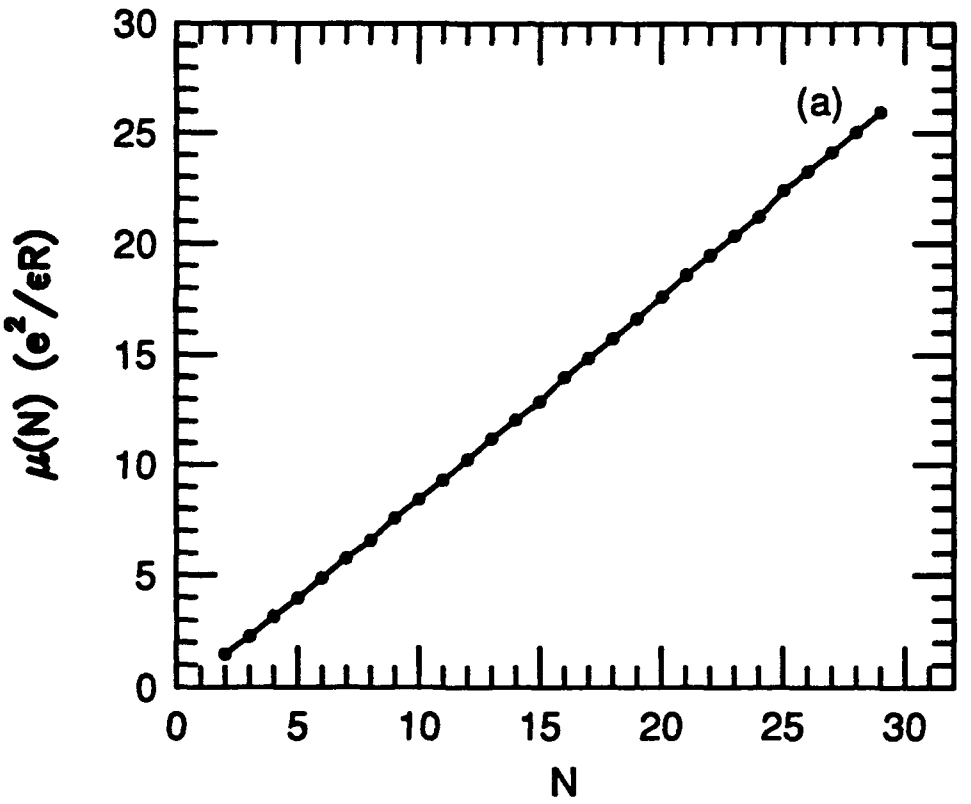

Fig3.a

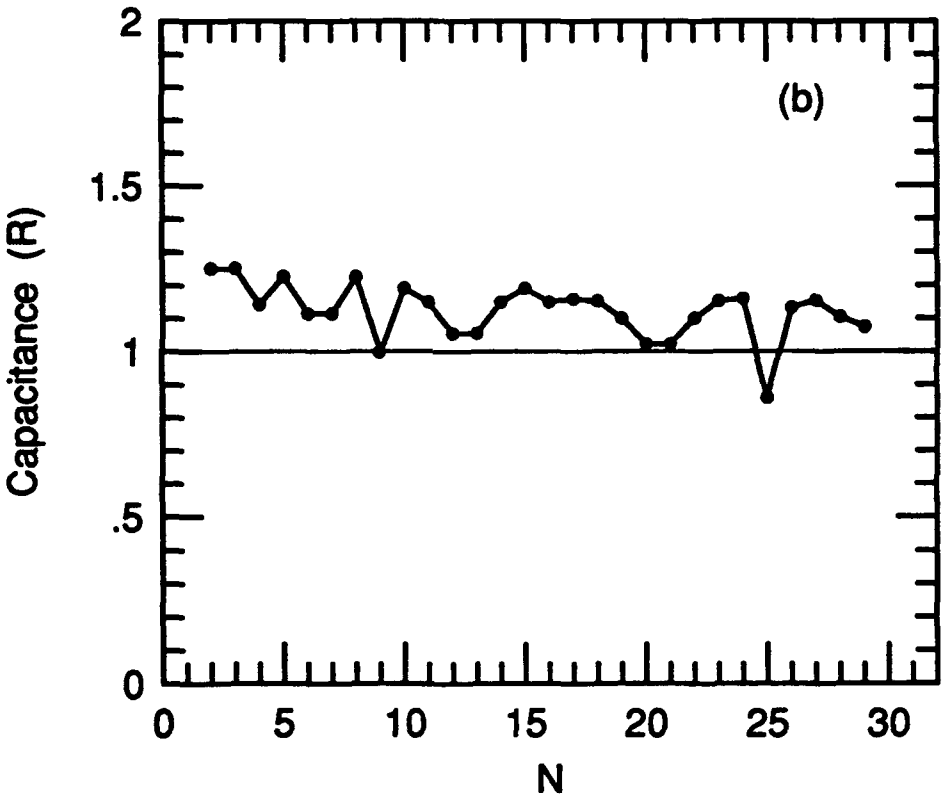

Fig 3b