Theory of single-electron charging of quantum wells and dots

D. V. Averin*
Department of Physics, Moscow State University, Moscow 119899 GSP, U.S.S.R.

A. N. Korotkov
Institute of Nuclear Physics, Moscow State University, Moscow 119899 GSP, U.S.S.R.

K. K. Likharev*
Department of Physics, Moscow State University, Moscow 119899 GSP, U.S.S.R.
and Department of Applied Physics, Delft University of Technology, 2600 GA Delft, The Netherlands
(Received 16 April 1991)

Single-electron charging effects similar to those in small-area metallic tunnel junctions should take place in semiconductor heterostructures, in particular, small-area quantum wells. Our analysis shows that dc current-voltage characteristics of such a well should exhibit an interplay between single-electron charging and energy-quantization effects. Relative magnitude of the single-electron charging effects is determined by the same parameter which scales multielectron charging in the conventional (large-area) quantum wells.

I. INTRODUCTION

Single-electron charging effects in small-area metal tunnel junctions are well understood now (see, e.g. the review[1]). The origin of these effects is the fact that due to small electric capacitance C of a small-area junction, its electrostatic energy is changed considerably in the result of tunneling of even a single electron. This change leads to a certain correlation between electron tunneling events, and gives rise to a variety of new phenomena that can be observed in single junctions as well as multijunction systems.

In particular, electron-electron correlations in the system of two tunnel junctions connected in a series result in the "Coulomb staircase," a periodic modulation of the dc current-voltage characteristics of the system. Each period of the Coulomb staircase corresponds to the addition of one more electron to the central electrode of the system. This effect was clearly observed in a number of experiments.[1]

In semiconductors, a similar structure is the famous double-barrier quantum well—see, e.g., the review.[2] In the usual (large-area) wells the multielectron charging effects are known to be important. In particular, such a charging can lead to an intrinsic bistability of the current flow through this structure.[3]

An objective of the present work was to show that in the small-area quantum wells (and other small-area heterostructures) the discreteness of the charge accumulation should become important, since the charging of the well by even a single electron changes the potential profile of the structure considerably. A brief report on our results was presented earlier.[4] Similar ideas were put forward in Refs. 5–7.

We will show that the single-electron charging effects arising in semiconductor heterostructures may be similar to those in metal junction systems. For instance, the dc current-voltage characteristics of the quantum well may exhibit a Coulomb staircase type structure similar to that observed in its metallic counterpart.

However, the single-electron charging phenomena in semiconductors and metals should differ in some important respects, mainly due to two circumstances. First, the Bohr radius a_B in semiconductors is much larger ($a_B \cong 100$ Å for GaAs), and heterostructures with dimensions smaller than a_B can be fabricated. As a result, the discreteness Δ of the electron kinetic energy spectra in a conducting region of heterostructures ("electrode") can be comparable to the characteristic charging energy δ, whereas for typical metal junctions, $\Delta \ll \delta$. Second, the Fermi energy in semiconductor structures can be comparable to the charging energy δ, and the absolute number of free electrons in the electrodes is not necessarily much larger unity.

In this work we extend the conventional description[1] of the single-electron charging effects to semiconductor heterostructures (specifically, the quantum wells) with an account of these peculiarities. In Sec. II we write down the master equation that describes the tunneling through the quantum well. This equation takes into account both the electron-electron correlations that arise due to single-electron charging, and discreteness of the energy spectrum of the well. Using this equation we check the adopted model by calculating the dc current-voltage characteristics of a conventional (large-area) quantum well (Sec. III). Next, we consider the most important case of a small-area quantum well and calculate its dc I-V curves (Sec. IV). An emphasis is placed on a discussion of the interplay between two structures of these two curves, one related to the discreteness of the charge accumulation in the well, and another one related to the energy discreteness. In Sec. V we consider electron transport through a laterally confined quantum dot that also can be described with our master equation. In this case the gate

voltage dependence of the quantum dot characteristics is of the special interest. We conclude (in Sec. VI) with a discussion of published experimental results on the electron transport through small-area quantum wells and dots and possible further experiments.

II. THE MODEL AND BASIC EQUATIONS

We begin by considering the conventional double-barrier structure with the conduction-band-edge profile (along the direction of the current flow) shown schematically in Fig. 1. The Hamiltonian of this structure includes energy H_w of the well itself, Hamiltonians $H^{e,c}$ of the external electrodes (emitter and collector), and terms $H_T^{e,c}$ describing tunneling through the energy barriers,

$$H = H_W + H^c + H^e + H_T^e + + H_T^c \ . \tag{1}$$

We assume that electrons in the emitter and collector form a free electron gas,

$$H^{e,c} = \sum_i \varepsilon_i^{e,c} c_i^\dagger c_i \ , \tag{2}$$

with a certain Fermi energy ε_F determined by the doping density of these electrodes.

Besides that, we assume that addition of a single electron to the well does not change substantially the shape of the confining potential of the well (that is the sum of the potential imposed by the structure, and the electrostatic potential due to the charge accumulated in the well) and changes only the average well potential relative to the external electrodes.[8] In this case, by treating the electron-electron interaction in the Hartree approximation one can write down the well Hamiltonian in the following form:

$$H_W = \sum_k \varepsilon_k c_k^\dagger c_k + U \ , \tag{3}$$

where $\{\varepsilon_k\}$ is the electron-energy spectrum in the confining potential, and U is the electrostatic energy.

We will consider the "ideal" model of the quantum well in which electron motion along the layers of the structure and motion in the transverse direction (in the direction of the current flow) can be separated.[2] We assume that electrons in the well occupy only the lowest-energy level E_0 (see Fig. 1) with respect to motion in the direction of the current flow and introduce the energy E of this level relative to the conductance-band edge in the emitter at vanishing bias voltage. Hence, the average interval ρ^{-1} between eigenenergies ε_k is determined by the two-dimensional (2D) density of states ρ, $\rho = \pi\hbar^2/mS$.

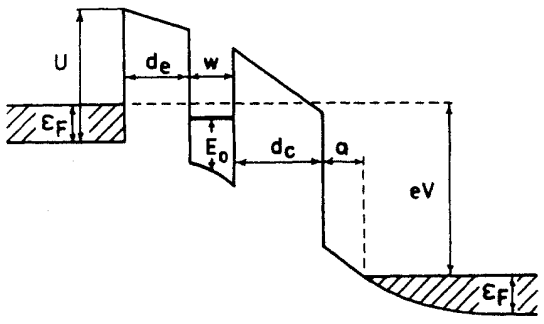

FIG. 1. Band-edge-profile model of the quantum well.

For a description of electrostatic energy U of the well, we borrow the simple expressions

$$U = \frac{Q^2}{2C} - \eta VQ - eVN, \quad Q = en, \quad n = \sum_k c_k^\dagger c_k \ , \tag{4}$$

which work so well for the tunneling between metallic electrodes.[1] Here electric capacitance C of the well is the sum of emitter-well and collector-well capacitances,

$$C = C_e + C_c \ , \tag{5a}$$

which are given approximately by the following equations:

$$C_e = \epsilon\epsilon_0 S \Big/ \left[d_c + \frac{w}{2} \right], \quad C_e = \epsilon\epsilon_0 S \Big/ \left[\frac{w}{2} + d_c + a \right] , \tag{5b}$$

where S is the area of the well, and other notations are explained by Fig. 1. Parameter η in Eq. (4) denotes the part of the total voltage V across the structure that drops between the emitter and the well, $\eta \cong C_c/C$; N is the number of electrons that have tunneled through the well.

One should note that an approximate description of the electron-electron interaction in the well, based on Eqs. (3) and (4), is not strictly valid when the number of free electrons or electron density in the well is small. In this case electron-electron interaction inside the well is not screened out, so that both exchange interaction and electron correlations inside the well [which are not taken into account in the Hartree approximation used in Eqs. (3) and (4)] may become essential. In this case one can improve Eq. (4) somewhat by taking into account part of the exchange interaction that is responsible for the fact that an electron does not contribute to the average electrostatic potential for the same electron. This can be achieved by replacing $Q^2/2C$ in Eq. (4) with $(e^2/2C)n(n-1)$.

Following Refs. 9 and 10 we write down the tunneling terms $H_T^{e,c}$ as the standard tunnel Hamiltonians

$$H_T^{e,c} = \sum_{i,k} T_{ik}^{e,c} c_i^\dagger c_k + \text{H.c.} \ , \tag{6}$$

and assume that the tunneling matrix elements $T^{e,c}$ and the corresponding tunneling rates

$$\Gamma_k^{e,c}(\varepsilon) = \frac{2\pi}{\hbar} \sum_i |T_{ik}^{e,c}|^2 \delta(\varepsilon - \varepsilon_i) \ , \tag{7}$$

are small

$$\hbar\Gamma^{e,c} \ll \min\{\rho^{-1}, \delta\} \ . \tag{8}$$

In relation (8), $\delta \equiv e^2/C$ is the characteristic charging energy. This relation implies that the average conductance of the whole structure is also small, $G \ll R_Q^{-1}$ ($R_Q \equiv \pi\hbar/2e^2 \cong 6.5$ KΩ).[11] When this condition is

satisfied and temperatures are not too low ($T \gg \hbar\Gamma^{e,c}$), the dynamics of tunneling can be described by a simple master equation. Validity of this equation is closely related to the fact that for $T \gg \hbar\Gamma^{e,c}$ the resonant current peak associated with each energy level (described by the conventional Breit-Wigner formula) is strongly broadened by temperature, and the resonant tunneling process cannot be distinguished from the sequential tunneling.[9]

In addition to assumption (8) we assume that the tunneling rates $\Gamma^{e,c}$ are smaller than the electron-energy relaxation rate τ_ϵ^{-1} in the well. In this case the current flow does not disturb the electron equilibrium, and the master equation can be constructed from the following simple considerations. (The effects of electron nonequilibrium have been considered in Ref. 11.)

Electron tunneling changes the number n of electrons in the well, and hence, the electrostatic energy U, $\Delta U = U(n+1) - U(n)$. When an electron tunnels between the emitter and well or between the well and collector, the energy change is

$$\Delta U_n^e = \delta(n+\tfrac{1}{2}) - \eta eV, \quad \Delta U_n^c = \delta(n+\tfrac{1}{2}) + (1-\eta)eV , \tag{9}$$

respectively. Thus, the tunneling connects the state ε_k in the well with an emitter or collector state with the energy $\varepsilon_k + \Delta U_n^{e,c}$.

The total rates $w_n^{\pm}$ of electron tunneling to or from the quantum well are the sum of all partial tunneling rates to or from specific energy levels ε_k,

$$\begin{aligned} w_+^{e,c}(n) &= \sum_k \Gamma_k^{e,c}(\varepsilon_+) f(\varepsilon_+)[1 - g_n(\varepsilon_k)] , \\ w_-^{e,c}(n) &= \sum_k \Gamma_k^{e,c}(\varepsilon_-)[1 - f(\epsilon_-)] g_n(\varepsilon_k) , \\ w_n^{\pm} &= w_\pm^e(n) + w_\pm^c(n) , \end{aligned} \tag{10a}$$

and

$$\varepsilon_+ = \varepsilon_k + E + \Delta U_n^{e,c}, \quad \varepsilon_- = \varepsilon_k + E + \Delta U_{n-1}^{e,c} . \tag{10b}$$

Here $g_n(\varepsilon_k)$ is the single-particle distribution function associated with the equilibrium Gibbs distribution $F(\varepsilon_{k_1}, \ldots, \varepsilon_{k_n})$ of n electrons in the well,

$$g_n(\varepsilon_k) = \sum_{k_1, \ldots, k_{n-1}} F(\varepsilon_k, \varepsilon_{k_1}, \ldots, \varepsilon_{k_{n-1}}) , \tag{11a}$$

$$F(\varepsilon_{k_1}, \ldots, \varepsilon_{k_n}) = Z^{-1} \exp\left[-\frac{1}{T} \sum_{i=1}^{n} \varepsilon_k \right] , \tag{11b}$$

$$Z = \sum_{k_1, \ldots, k_n} \exp\left[-\frac{1}{T} \sum_{i=1}^{n} \varepsilon_{k_i} \right] , \tag{11c}$$

and the sums in Eqs. (11a) or (11c) are taken over the sets of different energy levels, $k_i \neq k, k_j$. In the two limits $T \ll \rho^{-1}$ and $T \gg \rho^{-1}$, all $g_n(\epsilon)$ coincide with the Fermi distribution function $f(\epsilon)$. For intermediate temperatures, $T \simeq \rho^{-1}$, we calculate them numerically from Eqs. (11).

Transitions (10) lead to an evolution of the probabilities σ_n to find n electrons in the well in accordance with the master equation

$$\dot{\sigma}_n = w_{n+1}^- \sigma_{n+1} + w_{n-1}^+ \sigma_{n-1} - (w_n^+ + w_n^-)\sigma_n . \tag{12}$$

Using the evident relation

$$\langle \dot{n} \rangle = \sum_n n \dot{\sigma}_n = \sum_n (w_n^+ \sigma_n - w_{n+1}^- \sigma_{n+1})$$

one can easily be convinced that the stationary solution of Eq. (12) is given by the following recurrent formula:

$$\sigma_{n+1} = \sigma_n (w_n^+ / w_{n+1}^-) . \tag{13}$$

The dc current flowing through the quantum well can be calculated from this solution as follows:

$$\begin{aligned} I &= I^e = I^c , \\ I^{e,c} &= \pm e \sum_n \sigma_n [w_+^{e,c}(n) - w_-^{e,c}(n)] . \end{aligned} \tag{14}$$

Equations (10)–(14) are used below to calculate the dc current-voltage characteristics of large-area and small-area quantum wells.

III. LARGE-AREA QUANTUM WELL: MULTIELECTRON CHARGING

Both the electric capacitance C and density of states ρ of the quantum well with a large area S are large too: $C, \rho \propto S$. In such a well the effects associated with the charge or energy spectrum discreteness are suppressed, for instance, by thermal fluctuations, $T \gg \delta, \Delta$. As a result, the dc I-V curves have no fine structure on the voltage scales of δ/e or Δ/e.

The large-scale shape of the dc I-V curves (on the voltage scale of ε_F/e) is sensitive to the dependence of the tunneling rates $\Gamma_k^{e,c}(\varepsilon_\pm)$ on the energies $\varepsilon_\pm$. We will adopt the usual approximation[2] that not only the total-electron energy, but also the electron mode of propagation along the layers of the structure is not changed during the tunneling. In this case the tunneling rates depend only on the energy ε_{tr} of electron motion in the transport direction. Under the assumption that heights of the energy barriers are much larger than ε_F, one can get the following expression for the emitter tunneling rate (see Appendix A):

$$\begin{aligned} \Gamma_k^e(\varepsilon_\pm) &= \Gamma^e \begin{cases} (\varepsilon_{tr}/\varepsilon_F)^{1/2}, & \varepsilon_{tr} > 0 \\ 0, & \varepsilon_{tr} < 0 \end{cases} \\ \varepsilon_{tr} &= e(\varphi - \eta V) + E . \end{aligned} \tag{15}$$

Here we have introduced the electric potential φ associated with the accumulated charge Q, $\varphi \equiv Q/C$. By Γ^e we denote the emitter tunneling rate at the threshold voltage V_t, at which the tunnel current through the quantum well appears, $e\eta V_t \equiv E - \varepsilon_F$. We will use the same convention for the collector tunneling rate Γ^c, but here we neglect its dependence [Eq. (A4)] on transport energy because the relative variation of Γ^c in the relevant voltage region, $V_t < V \lesssim V_t + \varepsilon_F/e\eta$, is much smaller than the relative variation of Γ^e.

For the large-area quantum well its charge can be con-

sidered as continuous ($Q \gg e$). Fluctuations δQ arising due to the shot noise of the tunnel current are small as compared to the average charge Q_0, $\delta Q \simeq (eQ_0)^{1/2} \ll Q_0$. Thus, in order to calculate a current through a well it is sufficient to find the average steady-state potential φ from the equation

$$I^e(\varphi) = I^c(\varphi) \ . \tag{16}$$

Inserting Eqs. (15) into Eqs. (10) and (14) we get the following expression for the emitter-well and well-collector currents $I^{e,c}(\varphi)$:

$$\begin{aligned} I^e(\varphi) &= e\rho\Gamma^e(\varepsilon_{tr})[e\eta(V - V_t) - e\varphi - \mu] \ , \\ I^c(\varphi) &= e\rho\Gamma^c\mu \ , \end{aligned} \tag{17}$$

where we have introduced the Fermi energy μ in the well, $\mu \equiv n/\rho$. Since both the electrostatic potential φ and the Fermi energy μ are proportional to the number n of electrons in the well, μ is directly related to φ,

$$\mu = \frac{e\varphi}{\beta}, \quad \beta \equiv \rho\delta = \frac{e^2\rho}{C} = \frac{4d_{\text{eff}}}{a_B} \ . \tag{18}$$

Here $a_B = 4\pi\epsilon\epsilon_0/\hbar^2/me^2$ is the Bohr radius, and $d_{\text{eff}}^{-1} = (d_e + w/2)^{-1} + (d_c + a + w/2)^{-1}$ characterizes the total capacitance of the well—see Eqs. (5).

Inserting Eqs. (17) into Eq. (16) we find for the dc current I flowing through the well,

$$I = [1 + \alpha(\varepsilon_{tr})]^{-1} e^2\rho\eta(V - V_t)\frac{\Gamma^e(\varepsilon_{tr})\Gamma^c}{\Gamma^e(\varepsilon_{tr}) + \Gamma^c} \ , \tag{19}$$

where

$$\alpha(\varepsilon_{tr}) \equiv \beta\frac{\Gamma^e(\varepsilon_{tr})\Gamma^c}{\Gamma^e(\varepsilon_{tr}) + \Gamma^c}, \quad \alpha(\varepsilon_{tr})|_{V=V_t} \equiv \alpha \ . \tag{20}$$

Another solution is

$$I = 0 \quad \text{for } V - V_t > \varepsilon_F/e\eta \ . \tag{19'}$$

Equations (19) and (20) in the approximation $\Gamma^{e,c}(\varepsilon_{tr}) = \text{const}$ were first obtained by Sheard and Toombs.[9] They pointed out that parameter α determines the importance of charge-accumulation effects, in particular, the magnitude of the bistability region. This conclusion remains qualitatively unchanged in our somewhat more realistic model, in which we take into account the dependence of the emitter tunneling rate on the transport energy ε_{tr}. To find the dc I-V curve of the well in this model, one needs to solve Eq. (16) for the well potential φ and substitute it in Eq. (19). The results of these calculations are shown in Fig. 2. One can see that the additional element introduced by ε_{tr} dependence of Γ^e (and related to the fact that $\Gamma^e(\varepsilon_{tr})$ vanishes at $\varepsilon_{tr} \to 0$) is a rounding of the resonant current peak. It leads to a decrease of the peak height and shrinking of the voltage region where there are two stationary states with different current values (19). The degrees of this rounding depends on the tunneling rates ratio Γ^e/Γ^c [Fig. 2(b)].

The shape of the resonant current peak shows that in the voltage region above $\varepsilon_F/e\eta$ there is the third current

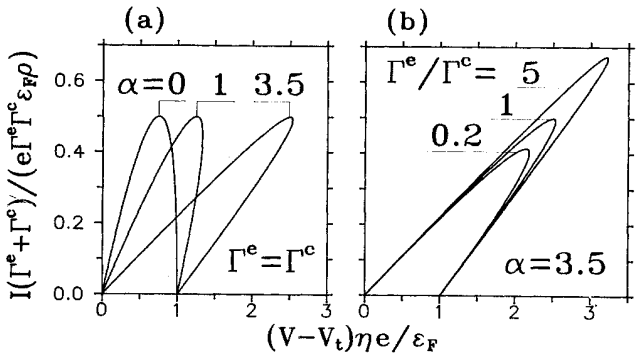

FIG. 2. dc I-V curves of a large-area quantum well for (a) various values of the parameter α and (b) various ratios of the tunneling rates Γ^e/Γ^c.

value which satisfies Eqs. (16) and (17). However, one can get convinced (in contrast to the suggestion made in Ref. 12) that this solution is unstable—see Appendix B.

Thus, we can conclude that the charge accumulation in the large-area quantum well can be essential if the parameter α (20) is sufficiently large, $\alpha \gtrsim 1$. This inequality implies that the two conditions are satisfied. First, the barriers of the structure are rather thick: $4d_{\text{eff}} \gtrsim a_B$ (so that $\beta \gtrsim 1$) and second, the collector tunneling rate is not much larger than the emitter rate, $\Gamma^e \gtrsim \Gamma^c$. As it will be shown in the next section, the same parameters α, β govern the single-electron charging of the small-area quantum wells, so that the same conditions are necessary for such a charging to take place.

IV. SMALL-AREA QUANTUM WELLS

In small-area quantum wells both the average energy interval $\Delta \equiv 2\rho^{-1}$ between 2D electron states in the well (that are at least double degenerate) and the single-electron charging energy δ become large. In particular, they can be larger than the energy scale of the thermal fluctuations: $\Delta, \delta \gg T$. Hence, one should take into account the discreteness of the 2D electron spectrum of the well, as well as the discreteness of the electric charge in it.

In this case, the dc current-voltage characteristics of the well can be calculated from Eqs. (9)–(14). The Fermi energy in the external electrodes is typically[3,13–15] of the order of 0.1 eV, and it is larger than both internal energies Δ, δ that are of the order of 10 meV or smaller even for submicron-size structures,

$$\varepsilon_F \gg \Delta, \delta \ . \tag{21}$$

It means that the "global" ($V \simeq \varepsilon_F/e$) shape of the I-V curve is the same as that of the large-area quantum well, if the current is scaled down proportionally to the area.

The only difference in the global shapes can arise from the fact that for small-area quantum wells deviations from the "ideal" resonant tunneling model considered in Sec. III can become essential. Since the lateral dimensions of such a well are comparable to the depletion depth, the laterally confining potential can be different in

different layers of the structure, and electron motion along the layers can be changed at tunneling. As a result, the emitter tunneling rate can depend on energy in another way than in the ideal model described by Eq. (15). In particular, Γ^e does not necessarily vanish when the conduction-band edge of the emitter is higher in energy than the subband edge of the well. This means that the current through the well does not vanish at large voltages, so that the peak-to-valley current ratio of the small-area quantum well should be typically smaller than that of the similar large-area structure.

However, these variations of the large-scale ($\simeq\varepsilon_F$) energy dependence of the tunneling rates do not influence a fine structure of the I-V curves, arising either due to the discreteness of the 2D electron-energy spectrum of the well, or due to the discreteness of the electric charge in it. This fine structure (which is of main interest in this work) has a voltage scale of the order of Δ/e or δ/e, i.e., much smaller than ε_F/e (21), and can be described adequately within the same resonant tunneling model (15) employed above.

A. General scheme of the fine structure

Equation (9) (corrected for the case of the empty well, as it was discussed in Sec. II) and Eqs. (10)–(14) imply that in the low-temperature limit ($T \ll \Delta, \delta$) the dc I-V curves of the well with $E > \varepsilon_F$ should exhibit current steps at dc voltage levels $V_{n,k}$ which obey a very simple rule

$$e\eta(V_{n,k} - V_t) = \delta n + \varepsilon_k, \quad n = 0,1,2,\ldots, \quad k = 1,2,\ldots . \tag{22}$$

The physical meaning of Eq. (22) is quite simple: the step appears when the voltage aligns the Fermi level of the emitter with the kth energy level of the well, with an account of the well charging by n electrons already residing there. Such an alignment opens new channels of the dc current flow through the well. (When $E < \varepsilon_F$ similar current steps can also arise due to the alignment of the well energy levels with the collector Fermi level.)

The system of voltages (22) is shown schematically in Fig. 3(a). One can see that despite the simplicity of Eq. (22) the fine structure of the I-V curve can be quite complex, and the mere identification of the experimentally observed singularities can present a hard problem.

Nevertheless, in some particular cases the fine structure may be more simple, because not all of the steps (22) have equal heights, and some of them can be not observed at all. Let us analyze these ultimate cases.

B. $\delta \ll \Delta$: Energy quantization

We begin by considering the case $\delta \ll \Delta$, when the energy spectrum discreteness is dominant. Let the temperature be not too low, so that the charging energy δ is negligible,

$$\Delta, T \gg \delta .$$

In this case one can neglect the terms containing δ in Eqs. (9), i.e., neglect the variations of the well potential arising due to the charge variations. Hence, the summation over k in Eqs. (10a) for the transition probabilities partially can be carried out explicitly, and one gets

$$w^e_+(n) - w^e_-(n) = \Gamma^e(\varepsilon_{tr}) \left[\sum_k f(\varepsilon_k + \varepsilon_{tr}) - n \right] , \qquad \varepsilon_{tr} = E - e\eta V \tag{23a}$$

$$w^c_-(n) - w^c_+(n) = \Gamma^c \left[n - \sum_k f(\varepsilon_k + \varepsilon_{tr} + eV) \right] . \tag{23b}$$

Equations (22) allow one to find the current I (14) without solving the master equation (12). Inserting them into Eqs. (14) one can find the average steady-state number $\langle n \rangle$ of electrons in the well,

$$\langle n \rangle = [\Gamma^e(\varepsilon_{tr}) + \Gamma^c]^{-1} \left[\Gamma^e(\varepsilon_{tr}) \sum_k f(\varepsilon_k + \varepsilon_{tr}) + \Gamma^c \sum_k f(\varepsilon_k + \varepsilon_{tr} + eV) \right] ,$$

and hence, the current I

$$I = \frac{e\Gamma^e(\varepsilon_{tr})\Gamma^c}{\Gamma^e(\varepsilon_{tr}) + \Gamma^c} \sum_k [f(\varepsilon_k + \varepsilon_{tr}) - f(\varepsilon_k + \varepsilon_{tr} + eV)] . \tag{24}$$

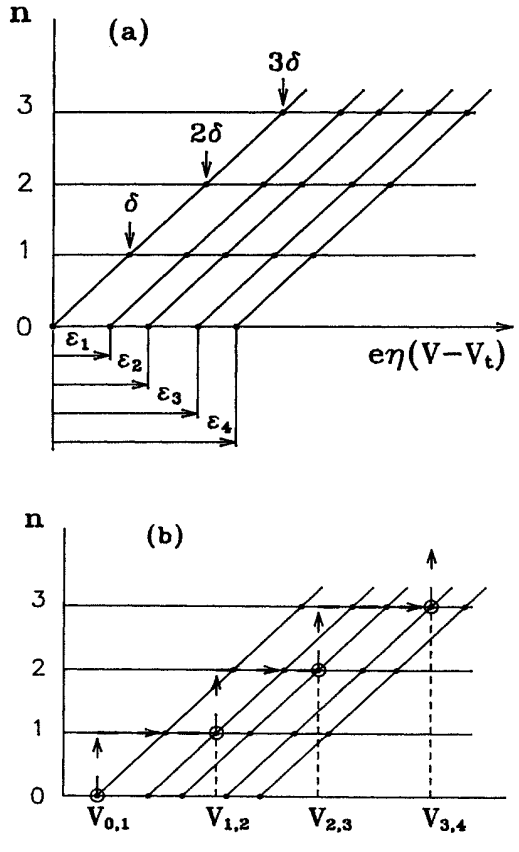

FIG. 3. (a) The general scheme of the dc-current step positions on the dc I-V curve of a small-area quantum well and (b) steps visible at $T \to 0$ and $\Gamma^e \gg \Gamma^c$ (indicated by circles).

One consequence of Eq. (24) is that at $T \ll \Delta$ and $E > \varepsilon_F$ the threshold voltage is increased by $\varepsilon_1/e\eta$, where ε_1 is the lowest eigenenergy of the well energy spectrum $\{\varepsilon_k\}$, in comparison to the threshold voltage of a large-area quantum well with the same parameters of the band structure.

Besides this, Eq. (24) implies that only the current steps corresponding to $n=0$ in Eq. (22) are visible. The shape of the current steps at $V \cong V_k$ for a finite temperature T is

$$I(V)=\frac{2e\Gamma^e(E-e\eta V)\Gamma^c}{\Gamma^e(E-e\eta V)+\Gamma^c} \times(k-1+\{1+\exp[-e\eta(V-V_k)/T]\}^{-1}) . \tag{25}$$

In the initial part of the global I-V curve where the current increases, these singularities are real current steps somewhat rounded by temperature. The decrease of the tunneling rate $\Gamma^e(E-e\eta V)$ with increasing voltage gives rise to the small ($\propto 1/\rho\varepsilon_F$) negative slope of these steps. For voltages well beyond the threshold voltage V_t the fine structure adjusts itself to the global shape of the I-V curve, due to the voltage dependence of Γ^e (Fig. 4).

C. $\delta \gg \Delta$: Single-electron charging

Now consider the opposite limit, when the discreteness of the well energy spectrum is negligible,

$$\delta, T \gg \Delta .$$

Such a well is very similar to the system of two metallic tunnel junctions, so that the two effects, characteristic for the latter system, also should take place in the well. The first effect is the Coulomb blockade of tunneling (a suppression of the tunnel current at small-bias voltages), and the second one is the Coulomb staircase (a periodic modulation of the I-V curve at voltages above the Coulomb blockade interval)—see, e.g., Ref. 1.

Calculations in this limit are essentially simplified by

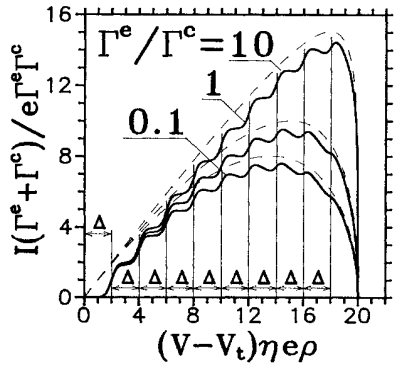

FIG. 4. The current steps associated with the single-particle energy levels in the dc I-V curves of a small-area quantum well with negligible charging and an equidistant energy spectrum for various Γ^e/Γ^c ratios and $T=0.1\Delta$. Dashed lines show the dc I-V curves of a similar but large-area quantum wells, formally scaled to the same small area.

two circumstances. The sum over k in Eqs. (10a) for the tunneling rates can be changed into integral over ε_k, and the electron distribution function $g_n(\varepsilon_k)$ at $T \gg \rho^{-1}$ is the Fermi distribution, with the Fermi energy μ defined by the following equation:

$$\rho\int_0^\infty d\varepsilon\, g_n(\varepsilon)=\rho\int_0^\infty d\varepsilon\{\exp[(\varepsilon-\mu)/T]+1\}^{-1}=n,$$

which yields

$$\exp\left[\frac{\mu}{T}\right]=\exp\left[\frac{n}{\rho T}\right]+1 . \tag{26}$$

Carrying out the integration in Eq. (10a) with Γ_k^e given by Eq. (15) and $\Gamma_k^c=\Gamma^c=\text{const}$, one gets

$$w_\pm^{e,c}(n)=\pm\rho T\Gamma^{e,c}(\varepsilon_\pm)\{1-\exp[\pm(\mu+\varepsilon_\pm-\varepsilon_F)/T]\}^{-1} \times\ln\left[\frac{1+\exp[(\varepsilon_F+\varepsilon_\pm)/T]}{1+\exp(\mu/T)}\right], \tag{27}$$

where $\varepsilon_+=E+\Delta U_n^{e,c}$ and $\varepsilon_-=E+\Delta U_{n-1}^{e,c}$.

If the conduction subband edge E of the well lies sufficiently below the Fermi level of the external electrodes at zero-bias voltage V ($E<\varepsilon_F$), the Fermi energy μ is large: $\mu \gg T, \delta$. In addition, if the voltage is small, so that $\varepsilon_F-\varepsilon_\pm \gg T$, Eq. (27) takes the following form:

$$w_\pm^{e,c}(n)=\pm\rho\Gamma^{e,c}(\varepsilon_\pm)(\varepsilon_F-\varepsilon_\pm-\mu) \times\{1-\exp[\pm(\mu+\varepsilon_\pm-\varepsilon_F)/T]\}^{-1} . \tag{28}$$

For small voltages, $V \ll \varepsilon_F/e$, one can neglect the energy dependence of the tunneling rates $\Gamma^{e,c}$, and the transition rates (28) coincide with those for the system of two metallic tunnel junctions with conductances $G_{1,2}$ and capacitances $C_{1,2}$,

$$G_{1,2}=e^2\rho\Gamma^{e,c},\quad C_1=\eta C,\quad C_2=(1-\eta)C . \tag{29a}$$

Parameter Q_0 which describes the zero-bias potential difference between external electrodes and central electrode, is

$$Q_0=\frac{C}{e}(\varepsilon_F-E_0-\mu)+\frac{e}{2} . \tag{29b}$$

Thus, transport behavior of the quantum well with $E<\varepsilon_F$ and $\delta \gg \Delta$ is completely similar to that of the system of two metal tunnel junctions with parameters defined by Eqs. (29), and the dc I-V curves of such a well coincide at small voltages with the dc I-V curves of the latter system.[1]

Now let us consider another situation, when the conduction subband edge E of the well lies above the Fermi levels of external electrodes at $V=0$ ($E>\varepsilon_F$, as is shown in Fig. 1), and Eqs. (28) for the transition rate are not applicable. In this case the chemical potential μ is small, and for not too small temperatures ($T \gg n\rho^{-1}$) Eqs. (27) can be simplified as follows:

$$w_+^{e,c}(n)=\rho T\Gamma^{e,c}(\varepsilon_+)\ln\{1+\exp[(\varepsilon_F-\varepsilon_+)/T]\} , \tag{30a}$$

$$w_{-}^{e,c}(n)=n\Gamma^{e,c}(\varepsilon_{-})\exp[(\varepsilon_{-}-\varepsilon_F)/T] \times\ln\{1+\exp[(\varepsilon_F-\varepsilon_{-})/T]\}\ . \quad (30b)$$

Although at arbitrary temperatures the tunneling rates (30) are different from those described by Eq. (28), they can be cast in the form similar to the tunneling rates in the system of two metal tunnel junctions at $T\ll\delta$ and voltages near the threshold voltage. The parameters of equivalent junctions are, however, different from those given by Eq. (29):

$$G_1=e^2\rho\Gamma^e,\quad G_2=C\Gamma^c,\quad C_1=0,\quad C_2=C,\quad Q_0=\frac{e}{2}\ , \quad (31)$$

and the effective voltage for the equivalent metallic system is $\eta(V-V_t)+e/C$.

These parameters of the equivalent metallic system imply that the tunnel current appears at $V=V_t$, so that there is no Coulomb blockade of electron tunneling (that could manifest itself as an increase of the threshold voltage) in the well with $E>\varepsilon_F$. The reason is that when the voltage across the structure reaches V_t the first electron tunnels in the empty well, and electrostatic potential created by this electron does not affect the electron itself,[16] as it was discussed in Sec. II. As a result, there is no shift of the threshold voltage.

Nevertheless, from the analogy (31) with the metal tunnel junction system it follows that the dc I-V curve of the quantum well should exhibit the Coulomb staircase, if the parameters (31) satisfy the condition $G_1/G_2=\beta\Gamma^c\gg1$. For $\beta\gg1$ this inequality is close to the inequality $\alpha\gg1$ with α given by Eq. (20). The origin of the Coulomb staircase can be understood as follows. If $\beta\Gamma^e\gg\Gamma^c$, the well-collector conductance is small and the current through the well is limited by the well-collector tunneling rate $w_{-}^{c}(n)$. In this case the number n of electrons in the well [as well as $w_{-}^{c}(n)$ (30b) that is proportional to n] increases discretely with increasing voltage. Thus, the I-V curve looks like a staircase, each step corresponding to the addition of one more electron to the well. The period of the structure is $\delta/e\eta$.

Decrease of the ratio $\beta\Gamma^e/\Gamma^c$ and/or raising the temperature suppress the Coulomb staircase as illustrated by Fig. 5. The dc I-V curves of the well shown in this figure were calculated numerically according to Eq. (14) by solving the master equation (12) with the transition rates (30).

D. Coexistence of the two structures

If the temperature is sufficiently low,

$$T\ll\delta,\Delta\ ,$$

the structures arising due to discreteness of the charge (the Coulomb staircase), and due to discreteness of the well energy spectrum, should coexist. The resulting I-V curve of the well depends strongly on the ratio of the tunneling rates Γ^e/Γ^c.

For $\Gamma^e\ll\Gamma^c$ the charge is not accumulated in the well and the only visible structure is the one related to energy

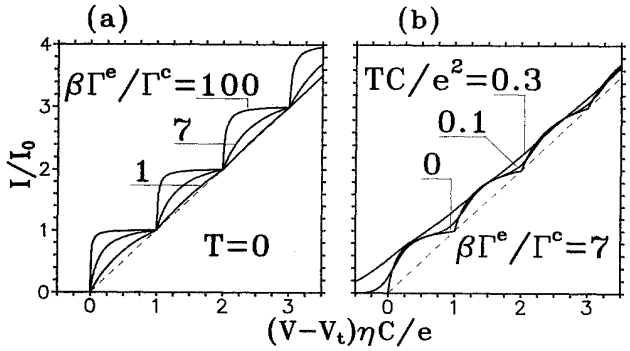

FIG. 5. The Coulomb staircase structure on the dc I-V curves of a small-area quantum well with negligible energy spectrum discreteness ($\beta\gg1$) at (a) various values of the parameter $\beta\Gamma^e/\Gamma^c$ and (b) various temperatures T. The dashed line denotes the same as in Fig. 4.

quantization which is described by Eqs. (24) and (25) [i.e., $n=0$ in Fig. 3(a)]. In the opposite limit, $\Gamma^e\gg\Gamma^c$, the choice of the steps is entirely different [Fig. 3(b)]. Let us increase dc voltage starting from $V=V_t$ when the well is empty. First we will reach voltage $V_{0,1}$ and populate the level ε_1. Due to small Γ^c, occupancy of this level will be virtually complete, so that $n=1$ now. Hence, we should pass to the line $n=1$ in Fig. 3 if we consider the further voltage increase. On this way we reach the value $V_{1,1}$, but since the level ε_1 is already occupied, no current step of a considerable amplitude is formed. Reaching the value $V_{1,2}$ we will populate the level ε_2, form the new current step on the I-V curve, and pass to the line $n=2$, etc. [see the dashed line in Fig. 3(b)].

As a result, of all the systems shown in Fig. 3(a), only the current steps at values $n=k-1$, i.e.,

$$V=V_{k-1,k}=V_t+[\varepsilon_k+\delta(k-1)]/e\eta\ , \quad (32)$$

will be pronounced. Thus, the main effect introduced by the charging of the well in the limit $\Gamma^e\gg\Gamma^c$ is a shift of the eigenenergies ε_k by $(e^2/C)(k-1)$. In particular, this means that degeneracy of the energy levels is lifted by the charging.

In order to find the dc current through the structure in this limit quantitatively it is sufficient to calculate the stationary probability density for $\Gamma^c=0$. Let's assume that for n electrons in the well the highest populated level is a p-fold-degenerate level ε_r, and q electrons of n occupy this level. (Here and below in Sec. V we use index r instead of k when we want to distinguish only nondegenerate levels in the well. If the well energy levels are nondegenerate $r=k$.) For voltage $V\cong V_{n,r}$ [see Eq. (22)] and low temperatures ($T\ll\delta,\Delta$), only two probabilities σ_n and σ_{n+1} are nonvanishing. The transition rates (10a) between these two states are

$$w_{+}^{e}(n)=(p-q)\Gamma^e g(\varepsilon)\ ,$$
$$w_{-}^{e}(n+1)=(q+1)\Gamma^e[1-g(\varepsilon)]\ , \quad (33)$$

where $\varepsilon=\varepsilon_r+E+\Delta U_n^e$. From Eqs. (33) and (13) we get

the probabilities

$$\sigma_n = \frac{(q+1)[1-g(\varepsilon)]}{(p-2q-1)g(\varepsilon)+q+1},$$
$$\sigma_{n+1} = -\frac{(p-q)g(\varepsilon)}{(p-2q-1)g(\varepsilon)+q+1}. \quad (34)$$

Inserting Eqs. (34) into Eq. (14) one can find the current I through the well. When the subband edge E lies above ε_F ($E > \varepsilon_F$), the backward collector-well tunneling vanishes ($w_+^c = 0$) and we get

$$I = e\Gamma^c[n\sigma_n + (n+1)\sigma_{n+1}]$$
$$= e\Gamma^c\left[n + \frac{p-q}{p-q+(q+1)\exp[-e\eta(V-V_n)/T]}\right]. \quad (35)$$

It means that in the case of a degenerate level, $p \neq 1$, the shape of the current steps (35) is different from steps (25). This is the consequence of the Coulomb correlations between electrons that occupy what was the degenerate level.

For intermediate values of the Γ^e/Γ^c ratio, the I-V curves exhibit not only two limiting series of the steps discussed above ($V_{0,k}$ and $V_{k-1,k}$), but also all steps between them [below the dashed-arrow line in Fig. 3(b)]. Let us have a look on examples shown in Figs. 6–8. They were calculated numerically from Eqs. (10)–(14) under the assumption that the energy spectrum of the well is equidistant and all energy levels are doubly degenerate.

When $\Delta \leq \delta$, (e.g., $\delta = 3.5\Delta$—see Fig. 6) the main structure is one associated with charging and described by Eqs. (32) and (35). However, each current step of this charging-related structure has a "superfine" structure related to the energy spectrum discreteness Δ [Fig. 6(b)]. With a decreasing Γ^e/Γ^c ratio and/or increasing voltage, the amplitude of the main structure decreases and only the superfine structure remains visible. On the contrary,

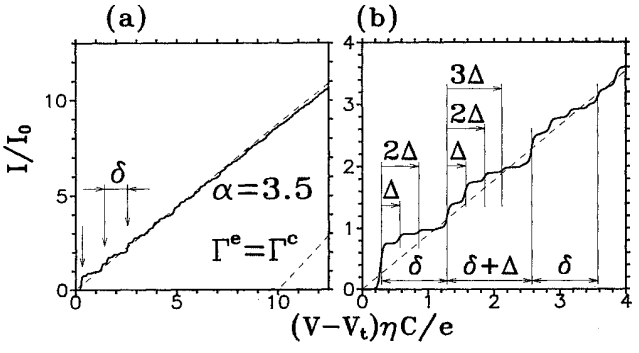

FIG. 6. dc I-V curve of a quantum well with a relatively large charging energy $\delta = 3.5\Delta$, for $e^2/C\varepsilon_F = 0.1$, $\Gamma^e = \Gamma^c$, and $T = 0.2e^2/C$: (a) the global shape of the curve; (b) an enhancement of its initial part. The curve exhibits the coexistence of the fine structure with the period $\delta/e\eta$ due to the single-electron charging and superfine structure with the period Δ/e due to the energy quantization. The dashed line meaning is the same as in Fig. 4.

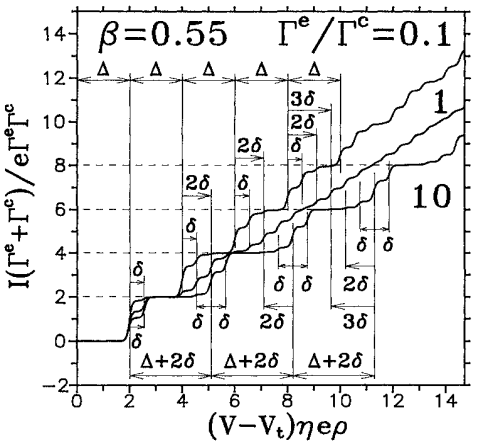

FIG. 7. dc I-V curves of a small-area quantum well with a relatively small charging energy ($\delta \cong 0.28\Delta$) and various ratios of the tunneling rates, at $T = 0.1e^2/C$. Here, the structure with a larger period in the upper curve is due to the energy quantization, whereas its smaller period is associated with the single-electron charging. The bottom curve shows the splitting of the current steps associated with the doubly degenerate energy levels, due to the single-electron charging of the well.

at large Γ^e/Γ^c ratios the superfine structure is suppressed.

If the energy spectrum discreteness Δ is dominant, $\Delta \gtrsim \delta$, and $\Gamma^e \lesssim \Gamma^c$, the two fine structures also can be easily separated: the energy quantization structure with the larger period Δ/e modulates the charging related superfine structure with the smaller period δ/e (Fig. 7). One can note that at large Δ the current steps associated with the single-electron charging persist to larger voltages than the usual Coulomb staircase at small Δ. For $\Gamma^e \gtrsim \Gamma^c$ the modulation period is increased due to the charge accumulation, and equals $(\Delta + 2\delta)/e$. In this case the superfine structure can be viewed as the splitting of

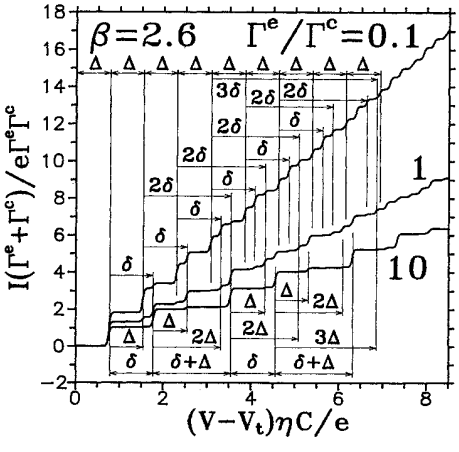

FIG. 8. dc I-V curves of a small-area quantum well with nearly equal characteristic energies ($\delta = 1.3\Delta$) and various ratios of the tunneling rates at $T = 0.02e^2/C$.

the large current steps corresponding to the double degenerate levels into two smaller steps separated by the voltage interval δ/e—see the bottom curve in Fig. 7.

When the periods of the two fine structures are nearly equal, $\Delta \cong \delta$, they form a more complicated pattern. In the limits $\Gamma^e \ll \Gamma^c$ and $\Gamma^e \gg \Gamma^c$ the largest current steps in this pattern are those described by Eqs. (24) and (25) and Eqs. (32) and (35), respectively (Fig. 8). At intermediate Γ^e/Γ^c ratios and/or large voltages, additional current steps at voltages $(k\Delta + j\delta)/e$ (where k,j are arbitrary integers) become pronounced. As a result, the visible quasiperiod of the I-V curve fine structure equals neither Δ/e nor δ/e—see the two upper curves in Fig. 8.

Thus, we can conclude that for the most realistic case when the characteristic charging energy δ and the average interval Δ between energy levels in the well are comparable, the dc I-V curve of the well should exhibit a complicated steplike fine structure. In order to identify current steps with the single-particle energy levels one should take into account the charging contribution to the well energy.

V. LATERAL TRANSPORT THROUGH THE QUANTUM DOT

The master equation (12) allows one to describe not only the vertical transport through the quantum-well structure shown in Fig. 1, but also the lateral transport through a "quantum dot" formed by a laterally confined region of the 2D electron gas. These cases differ only in the energy dependence of the tunneling rates $\Gamma^{e,c}$. For driving voltages much smaller than ε_F/e, where ε_F is the Fermi energy of the 2D electron gas outside the quantum dot, one can neglect this dependence and adopt the approximation $\Gamma^{e,c}(\varepsilon) = \text{const}$. Within this approximation the above discussion of the two types of the I-V curve fine structures can be applied to the quantum dot as well.

If the effective distance d_{eff} between the dot and the nearest conducting layer of the structure is smaller than the lateral dot dimension a, the dot electric capacitance C can be calculated as that of the plane capacitor. In this case the type of the I-V curve fine structure is determined by the same parameters α (20) and β (18). In the opposite limit $d_{\text{eff}} \gg a$, capacitance C can be estimated as the self-capacitance of the disk with radius a, $C = 8\epsilon\epsilon_0 a$, and the parameter β should be expressed as

$$\beta = \frac{e^2\rho}{C} = \frac{\pi a}{2a_B} \quad \text{for } d_{\text{eff}} \gg a \ . \tag{36}$$

Typically, $\min\{4d_{\text{eff}}, \pi a/2\} \gg a_B$, so that the Coulomb modulation of the I-V curve of the quantum dot should dominate.

The lateral transport through the quantum dot has one specific feature in comparison to the vertical transport through the quantum well: the electric potential of the dot can be readily varied (at fixed potentials of the external electrodes) by that of some additional gate electrode. Hence, one can study the dependence of the dot characteristics on the gate voltage V_g.

Influence of this voltage can be described by adding the term $-\lambda V_g Q$ to the electrostatic energy (4), where $\lambda \cong C_g/C$, C_g is the dot-gate capacitance, and the total capacitance $C = C_e + C_c + C_g$ now. This term leaves the above reasoning intact, if it is added to the right-hand part of Eq. (9). One can, however, simplify the resonance condition [see Eq. (22)] by restricting the observations to those of the zero-bias differential conductance G_d and differential capacitance C_d of the structure as functions of V_g. In this case the resonance condition similar to condition (22) takes the form

$$(V_g)_{n,k} = \frac{1}{e\lambda}[\delta(n+\tfrac{1}{2}) + \varepsilon_k + E - \varepsilon_F] \ . \tag{37a}$$

Moreover, in contrast to the finite-voltage case, the quantum dot remains in the thermodynamic equilibrium, so that at low temperatures ($T \ll \Delta, \delta$) the electrons fill completely all states below the kth level, irrespective of the Γ^e/Γ^c ratio. Hence one can take $n = k-1$, and the resonant series becomes quite simple,

$$(V_g)_{n,k} = (V_g)_{k-1,k} = \frac{1}{e\lambda}[\delta(k-\tfrac{1}{2}) + \varepsilon_k + E - \varepsilon_F] \ , \tag{37b}$$

One can say that the physical origin of these resonances is a V_g dependence of the energy gaps between the state with the n electrons in the dot, which provides the minimum of the dot energy (4), and excited states with $n \pm 1, \ldots$ electrons. These gaps (for $\beta \gg 1$) are of the order of characteristic charging energy δ, except for the "resonant" values of $V_g = (V_g)_{n,k}$, when one of the gaps, e.g., one between states with n and $n+1$ electrons, vanishes. It means that at low temperatures, $T \ll \delta$, the number of electrons in the dot is fixed, and both C_d and G_d vanish at any gate voltage except for these resonant voltage values.

Thus, the differential capacitance and conductance of the quantum dot as functions of the gate voltage V_g should exhibit a sequence of the resonance peaks. Each peak corresponds to the addition of one more electron to the quantum dot. In the most realistic case $\beta \gg 1$ the voltage interval between the peaks is determined by the charging energy, since this energy gives the main contribution to the dot energy when an electron is transferred to the dot. Hence, in this case the sequence of the resonance peaks is almost periodic with the period $\delta/e\lambda \cong e/C_g$.

The shape of the $C_d(V_g)$ and $G_d(V_g)$ peaks at low temperatures can be easily calculated from the stationary solution of the general master equation (12) for $\beta \gg 1$. If the states with n and $n+1$ electrons are in resonance, only the probabilities σ_n and σ_{n+1} are nonvanishing,

$$\sigma_n = \frac{w_{n+1}^-}{w_n^+ + w_{n+1}^-}, \quad \sigma_{n+1} = \frac{w_n^+}{w_n^+ + w_{n+1}^-} \ .$$

From these equations we get the following expressions for the differential capacitance of the dot:

$$C_d(V_g) = \frac{d\langle Q\rangle}{dV_g} = e\frac{d}{dV_g}(1 + w_{n+1}^-/w_n^+)^{-1} \ , \tag{38}$$

and [after substituting them into Eq. (11)] for the

differential conductance:

$$G_d(V_g)=\left.\frac{dI}{dV}\right|_{V=0} =e\frac{d}{dV}\frac{w_+^e(n)w_-^c(n+1)-w_+^c(n)w_-^e(n+1)}{w_n^+ +w_{n+1}^-} . \tag{39}$$

If the temperature is not too low, $T\gg\Delta$ (but still $T\ll\delta$), the transition rates $w_n^{\pm}$ are given by Eqs. (27) with $\Gamma^{e,c}=$const, and we obtain

$$C_d(V_g)=\lambda\frac{e^2}{4T}\left[\mathrm{ch}\frac{\varepsilon_F-\varepsilon-\mu}{2T}\right]^{-2} , \tag{40}$$

$$G(V_g)=\frac{\rho e^2}{2}\frac{\Gamma^e+\Gamma^c}{\Gamma^e+\Gamma^c}\ln\left[\frac{1+\exp[(\varepsilon_F-\varepsilon)/T]}{1+\exp(\mu/T)}\right] \times\left[\mathrm{sh}\frac{\varepsilon_F-\varepsilon-\mu}{T}\right]^{-1} . \tag{41}$$

Here $\varepsilon=E+\delta(n+\frac{1}{2})-e\lambda V_g$, and chemical potential μ is defined by Eq. (26). When the electron gas in the dot is degenerate, $\mu\gg T$, the logarithm in Eq. (41) can be replaced with the linear function $(\varepsilon_F-\varepsilon)/T$, and Eq. (41) coincides with that obtained earlier in Refs. 17 and 6.

At low temperatures, $T\ll\Delta$, the shape of the resonance peaks depends on the degree of degeneracy of the one-electron energy levels of the dot. As in Sec. IV C, let us assume that the highest populated level in the dot is a p-fold degenerate level ε_r, and q electrons occupy this level while there are n electrons in the dot as a whole. In this case the transition rates w_n^+, w_{n+1}^- [cf. Eqs. (33)] are

$$w_n^+=(p-q)\sum_{e,c}\Gamma^{e,c}g(\varepsilon) , \quad w_{n+1}^-=(q+1)\sum_{e,c}\Gamma^{e,c}[1-g(\varepsilon)] , \tag{42}$$

where $\varepsilon=\varepsilon_r+E+\Delta U_n^{e,c}$. Equations (42) inserted into Eq. (38) show that the capacitance peak $C_d(V_g)$ is described by the same expression (40) as for $T\gg\Delta$ if one replaces the chemical potential μ with $\varepsilon_r+T\ln[(p-q)/(q+1)]$. From Eqs. (39) and (42) one gets

$$G_d(V_g)=\frac{e^2}{T}\frac{\Gamma^e\Gamma^c}{\Gamma^e+\Gamma^c}\frac{[(p-q)(q+1)]^{1/2}}{p+1+2[(p-q)(q+1)]^{1/2}\mathrm{ch}[(\varepsilon_F-\varepsilon-\mu')/T]} , \tag{43a}$$

$$\mu'\equiv\varepsilon_r+(T/2)\ln[(p-q)/(q+1)] . \tag{43b}$$

Equations (40), (41), and (43) imply that in accordance with Eqs. (37), $C_d(V_g)$ and $G(V_g)$ have sharp maxima, when $\varepsilon\simeq\varepsilon_F-\mu$, i.e., at $V_g\cong[\mu-\varepsilon_F+E+\delta(n+1/2)]/e\lambda$. The interval between the resonance peaks of $C_d(V_g)$ and $G(V_g)$ varies slightly from peak to peak due to the n dependence of the chemical potential μ. These small deviations from the charging-determined period are of the order of T or Δ for, respectively, nondegenerate and degenerate electron gas in the dot.

An interesting property of the conductance peaks (43) is that for multidegenerate levels, $p\geq 4$, the amplitude of the peaks $G_{\max}(n)$ exhibits large-scale oscillations even for $\Gamma^{e,c}=$const. One can see from Eq. (43a) that this amplitude as a function of the "filing factor" q reaches its maximum at $q\simeq p/2$ (the energy level is half-filled with electrons) and decreases both for $q\to 0$ and for $q\to p$. These oscillations arise from the Coulomb correlations of electrons occupying the degenerate energy level.

VI. EXPERIMENTAL SITUATION

There already exists a number of experiments on the transport properties of the small-area quantum wells and quantum dots. A steplike fine structure of the current-voltage characteristics of the submicron quantum wells was observed in a series of experiments[12–14] pioneered by Read *et al.*[13] In the first two of these experiments[13,14] the double-barrier structure was symmetrical, so that at relevant voltages $V>V_t$ the condition $\Gamma^e\ll\Gamma^c$ was satisfied. Hence, charge accumulation in the well was negligible and the observed structure was apparently due to the 2D electron-energy quantization rather than single-electron charging. The latter effect can be presumably observed in almost similar heterostructures, but with a thicker and/or higher collector barrier, providing $\Gamma^e\gtrsim\Gamma^c$.

Such an asymmetrical heterostructure was used in experiment;[15] however, only the results obtained for positive-bias voltage (when the emitter barrier was the higher one) are reported. According to our results, at negative-bias voltages this heterostructures should exhibit the single-electron charging effects, since a similar large-area quantum well exhibited the hysteretic I-V curve[13] characteristic of the intrinsic bistability[3] due to the charge accumulation in the well.

Barrier thickness of the typical double-barrier heterostructures including those studied in Refs. 13–15 varied between 30 and 100 A. Hence, parameter β (18) for these structures was approximately within the interval 1–4, so that the characteristic charging energy δ was of the order of the energy interval Δ between energy levels. Thus, in these structures the single-electron charging should modify the dc I-V curve similarly to the curves shown in Fig. 8.

However, as it was discussed in Sec. IV C, a detailed interpretation of the fine structure of the I-V curves seems to be a serious problem. In this respect the case of the lateral transport through the quantum dot seems to be more advantageous, since in this case the effective distance d_{eff} between the quantum dot and external elec-

trodes is not limited by the width of the tunnel barriers, and the charging energy can be made much larger than Δ.

Several experiments[18,19] with submicron quantum dots show a periodic structure in the gate voltage dependence of the differential capacitance $C_d(V_g)$. The structure was observed as periodic oscillations in the (dC_d/dV_g)-vs-V_g curves. The oscillations could be associated in principle with the transfer of single electrons to the quantum dot. However, two properties of these oscillations, their large period, $\Delta V \simeq 50$ mV, and their periodicity, cannot be united in a simple way within this approach. In fact, the large period implies that the gate-dot capacitance (and thus the effective size d of the dot) is small, $d \simeq 100$ A. (The effective size of the dot can be much smaller than the lithographically defined size due to a surface depletion layer.) Hence, the parameter β (36) is close to unity and the energy spectrum discreteness should be comparable to δ. In this case, under the realistic assumption that electron-energy levels in the dot are at least double degenerate, oscillations in the capacitance spectrum should not be periodic, since the energy of the order of Δ is added to part of the periods of the oscillations.

Thus, a more plausible explanation of these oscillations relates them[20] to the large-scale fluctuations of the density of states in the dot. However, to prove this explanation one needs to gain more understanding of the origin of these fluctuations, in particular, a reason for there periodicity.

In another series of experiments[21–23] the linear conductance $G_d(V_g)$ through the submicron quantum dot was shown to exhibit periodic oscillation as a function of the gate voltage. These oscillations were attributed to the single-electron charging of the dot.[5] Although this interpretation cannot fully account for the existing experimental results, it seems to be consistent with most of them. The main evidence for the single-electron charging is the almost strict periodicity of the conductance oscillations and a very nice correspondence between the voltage period of these oscillations and estimations of the gate-dot capacitance C_g ($\Delta V_g = e/C_g \simeq 1$ mV).[22,23]

Another strong support for this interpretation of conductance oscillations was provided by experiment,[23] in which these oscillations are accompanied by the Coulomb staircase at larger driving voltages. The absence of such a structure on the dot current-voltage characteristic in the experiment[22] can be explained by the assumption that conductances of the two tunnel barriers confining electrons in the dot were approximately equal in this experiment. In such a case the Coulomb staircase is very weakly pronounced.[1]

An obscure aspect of these experimental results is a periodic modulation of amplitude of the conductance oscillations. This large-scale modulation could be related to the Coulomb correlations of electrons on the multidegenerate levels in the well, as was discussed in Sec. V. However, for this interpretation to be valid, the temperature should be smaller than the energy interval between single-particle levels in the dot. It seems not to be the case at least in the experiment,[22] since the width of the thermally broadened conductance peaks was of the order of the peak separation. Another possible origin of these large-scale oscillations is the large-scale oscillations[20] of the single-particle density of states in the dot, mentioned above. Finally, for the lateral transport through quantum dots the arguments presented in Appendix A are not strictly valid, so that the energy dependence of $\Gamma^{e,c}$ can be more complex than that suggested by Eq. (15). Thus we can conclude that although some experimental evidence of the single-electron charging in semiconductor heterostructures is on hand, further experiments and calculations are needed to obtain more definite results.

Note added in proof. After submission of this paper we learned that our suggestion to use asymmetrical quantum wells for observation of their single-electron charging (first published in Ref. 4) was successfully implemented by Su, Goldman, and Cunningham.[25] Their experimental data are in at least a qualitative agreement with our results for the case $\delta \geq \Delta$.

APPENDIX A

Tunneling rates $\Gamma_k^{e,c}(\varepsilon)$ are defined by Eq. (7) where the sum is taken over emitter or collector energy states. The eigenenergy ε_i can be separated into energy ε_{tr} of electron motion in the transport direction and energy ε_q of electron motion along the layers of the structures, $\varepsilon_i = \varepsilon_{tr} + \varepsilon_q$, so that

$$\sum_i = \sum_q \int_0^\infty d\varepsilon_{tr} \rho(\varepsilon_{tr}) \ . \tag{A1}$$

Here $\rho(\varepsilon_{tr})$ is the one-dimensional density of states, $\rho(\varepsilon_{tr}) \propto (\varepsilon_{tr})^{-1/2}$.

Since the mode of electron propagation along the layers is not changed during the tunneling, $T_{ik} = \delta_{qk} t(\varepsilon_{tr})$, and the tunneling probabilities $\Gamma^{e,c}$ in Eq. (10) are

$$\begin{aligned}\Gamma_k^{e,c}(\varepsilon_\pm) &= \frac{2\pi}{\hbar} \sum_q \delta_{qk} \int_0^\infty d\varepsilon_{tr} \rho(\varepsilon_{tr}) |t(\varepsilon_{tr})|^2 \\ &\quad \times \delta(\varepsilon_\pm - \varepsilon_{tr} - \varepsilon_q) \\ &= \rho(\varepsilon_{tr}) |t(\varepsilon_{tr})|^2 |_{\varepsilon_{tr} = \varepsilon_\pm - \varepsilon_k} \ .\end{aligned} \tag{A2}$$

The tunneling amplitude $t(\varepsilon_{tr})$ can be calculated according to the standard prescription (see e.g., Ref. 24)

$$t(\varepsilon_{tr}) \propto \int dx \, \psi_p(x) \psi_0(x), \quad \varepsilon_{tr} = \frac{p^2}{2m} \ . \tag{A3}$$

Here $\psi_p(x)$ is the wave function of an electron that in the external electrode has momentum p in the transport direction, $\psi_0(x)$ is the electron wave function in the well, and the integral is taken over the width of the well.

One can get convinced that when the height U of the tunnel barriers (Fig. 1) is large, $U \gg \varepsilon_{tr}$, the amplitude of the wave function $\psi_p(x)$ transmitted through the barrier into the well is proportional to p. Thus, in this case $t(\varepsilon_{tr}) \propto (\varepsilon_{tr})^{1/2}$, and we get from Eq. (A2)

$$\Gamma_k^{e,c}(\varepsilon_\pm) \propto (\varepsilon_{tr})^{1/2}, \quad \varepsilon_{tr} = \varepsilon_\pm - \varepsilon_k \ . \tag{A4}$$

For the large-area well, Eq. (A4) gives us Eqs. (15) of the main text.

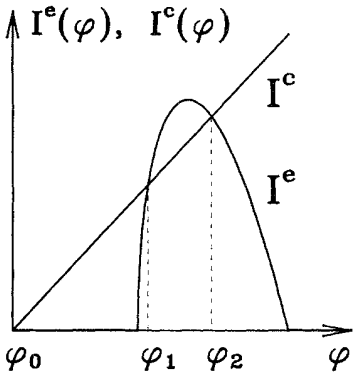

FIG. 9. Emitter→well and well→ collector currents $I^{e,c}(\varphi)$ [Eq. (17)] in the bistability regime.

APPENDIX B

In this appendix we will show that the charge accumulation in the large-area double-barrier structure gives rise only to two stable states with different values of the current flowing through the structure. The third state, which can be obtained by solving Eqs. (16) and (17), is unstable.

In order to prove this statement one needs to consider the time evolution of the well potential φ associated with accumulated charge Q, $\varphi=Q/C$. This evolution is governed by a simple equation,

$$\dot{\varphi}=[I^e(\varphi)-I^c(\varphi)]/C \ , \tag{B1}$$

where emitter→well and well→collector currents $I^{e,c}$ are given by Eqs. (17). These currents are plotted in Fig. 9 as functions of φ for the voltage V in the bistability region ($V > V_t+\varepsilon_F/e\eta$).

The stationary solutions φ_i of Eq. (B1) are defined by the intersections of the two curves in Fig. 9. Equation (B1) determines the evolution of the small deviation $\delta\varphi=\varphi-\varphi_i$ from these stationary solutions,

$$\delta\dot{\varphi}=-\lambda_i\,\delta\varphi, \quad \lambda_i\equiv\left[\frac{dI^c}{d\varphi}-\frac{dI^e}{d\varphi}\right]\Bigg|_{\varphi=\varphi_i} . \tag{B2}$$

One can see from Fig. 9 that for $\varphi=\varphi_0=0$ and $\varphi=\varphi_2$ $(dI^c/d\varphi)>(dI^e/d\varphi)$, so that $\lambda_{0,2}>0$ and these two solutions are stable. In contrast to this, $\lambda_1<0$, the solution $\varphi=\varphi_1$ is unstable [small deviations from φ_1 are spontaneously amplified, $\delta\varphi(t)\propto\exp(|\lambda_1|t)$].

Thus, there are only two stable solutions of Eqs. (16) and (17). It is clear from the above reasoning that this result does not depend on the precise analytic expressions for the currents $I^{e,c}(\varphi)$, but only on the topology of Fig. 9. Hence, our main conclusion that the charge accumulation gives rise to the bistability, but not tristability, is more general than the adopted model with the tunneling rates (A4), and will apparently hold on for any realistic model of the charge accumulation in the quantum wells.

*Present address: Department of Physics, State University of New York, Stony Brook, New York 11794-3800.

[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, Amsterdam, 1991), p. 169.

[2]S. Luryi, in *Heterojunction Band Discontinuities: Physics and Device Applications,* edited by F. Capasso and G. Margaritondo (Elsevier, Amsterdam, 1987), p. 489.

[3]V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. Lett. **58**, 1256 (1987); A. Zaslavsky, V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Appl. Phys. Lett. **53**, 1408 (1988).

[4]A. N. Korotkov, D. V. Averin, and K. K. Likharev, Physica B&C **165&166B**, 927 (1990).

[5]H. van Houten and C. W. J. Beenakker, Phys. Rev. Lett. **63**, 1893 (1989); C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).

[6]L. I. Glazman and R. I. Shechter, J. Phys. Condens. Matter **1**, 5811 (1989).

[7]A. Groshev, Phys. Rev. B **42**, 5895 (1990).

[8]Numerical simulation of the quantum dot energy spectrum carried out by A. Kumar, S. E. Laux, and F. Stern [Phys. Rev. B **42**, 5166 (1990)] indicates that this is a reasonable approximation even for a small number of electrons in the dot.

[9]M. C. Payne, J. Phys. C **19**, 1145 (1986).

[10]F. W. Sheard and G. A. Toombs, Appl. Phys. Lett. **52**, 1228 (1988).

[11]D. V. Averin and A. N. Korotkov, J. Low. Temp. Phys. **80**, 173 (1990).

[12]D. D. Coon, K. M. S. V. Bandara, and H. Zoo, Appl. Phys. Lett. **54**, 2115 (1989).

[13]M. A. Reed, J. N. Randall, R. J. Aggarval, R. J. Matyi, T. M. Moore, and A. E. Wetsel, Phys. Rev. Lett. **60**, 535 (1988).

[14]G. Faini, A. Ramdane, F. Mollot, and H. Launois (unpublished).

[15]Bo Su, V. J. Goldman, M. Santos, and M. Shayegan, Appl. Phys. Lett. **58**, 747 (1991).

[16]One could argue that electrostatic interaction of electron with external electrodes (emitter and collector) gives rise to an image potential which is in a sense the back action of the tunneling electron on itself. However, this potential (as was first pointed out to authors by Yu. V. Nazarov) is independent of the area S of the well, and hence, cannot be interpreted as the Coulomb blockade of tunneling.

[17]I. O. Kulik and R. I. Shechter, Zh. Eksp. Teor. Fiz. **68**, 623 (1975) [Sov. Phys.—JETP **41**, 308 (1975)].

[18]T. P. Smith III, K. Y. Lee, C. M. Knoedler, J. M. Hong, and D. P. Kern, Phys. Rev. B **38**, 2172 (1988); W. Hansen, T. P. Smith III, K. Y. Lee, J. A. Brum, C. M. Knoedler, J. M. Hong, and D. P. Kern, Phys. Rev. Lett. **62**, 2168 (1989).

[19]H. Fang, R. Zeller, and P. J. Stiles, Appl. Phys. Lett. **55**, 1433 (1989).

[20]R. H. Silsbee and R. C. Ashoori, Phys. Rev. Lett. **64**, 1991 (1990).

[21]J. H. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62**, 583 (1989); A. A. M. Staring, H. van Houten, C. W. J. Beenakker, and C. T. Foxon, in *High Magnetic Fields in Semiconductor Physics*

III, edited by G. Landwehr (Springer, Berlin, in press).
[22]U. Meirav, M. A. Kastner, and S. J. Wind, Phys. Rev. Lett. **65**, 771 (1990).
[23]L. P. Kouwenhouven, B. J. van Wees, V. ban der Enden, and K. P. J. M. Harmans, in *Proceedings of the 20th International Conference on the Physics* of Semiconductors, edited by J. Joannopoulos (World Scientific, London, 1990).
[24]E. O. Kane, in *Tunneling Phenomena in Solids,* edited by E. Burstein and S. Lundqvist (Plenum, New York, 1969), p. 1.
[25]Bo Su, V. J. Goldman, and J. E. Cunningham (unpublished).