RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 44, NUMBER 11 15 SEPTEMBER 1991-I

Variation of the Coulomb staircase in a two-junction system by fractional electron charge

A. E. Hanna and M. Tinkham
Department of Physics and Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138
(Received 17 June 1991)

We report a measurement of the Coulomb staircase in a two-junction system where the fractional residual charge Q_0 on the center electrode is varied without an external electrode. We present a simple analytic equation for $I(V)$ that depends on the parameters of the junctions, $C_{1,2}$, $R_{1,2}$, and Q_0, and allows us to obtain them directly from the data. Full "orthodox" theory simulations incorporating these parameters are in remarkable agreement with our data. Asymmetric gaplike features in the I-V curve are seen to vary with Q_0, and can be well understood by use of a C_2/C_1-Q_0 phase diagram.

We present high-quality Coulomb-staircase data, taken with a low-temperature scanning tunneling microscope and a granular gold film, which confirm a simple analytic expression for the current in a voltage-biased two-junction system in which the ratio of tunneling resistances $R_2/R_1 \gg 1$. Within this framework we show how I-V curves can be separated into four distinct cases, depending on the capacitance ratio C_2/C_1 and the fractional residual charge Q_0. We explain how the individual parameters of the system can be obtained from the I-V curve and how asymmetries are achieved when $Q_0 \neq 0$. The width of the Coulomb blockade region is also shown to depend on Q_0 and vary from the often quoted value e/C_Σ.

Our analysis is based on the "orthodox" theory of correlated electron tunneling, and follows the notation and derivation presented by Averin and Likharev.[1] A recent paper by Amman *et al.*[2] gives a concise description of the two-junction system, a schematic diagram of which is shown in Fig. 1(a), and provides a general analytic solution; however, that solution is not intuitive by inspection. To make our discussion self-contained, we will first rederive the more intuitive equation of Averin and Likharev by looking at a limiting case.

The particle tunneling rate for the jth junction is represented by $\Gamma_j^{\pm}(n)$, where the $+/-$ refers to electrons tunneling on/off the center electrode ($n \rightarrow n \pm 1$). $\Gamma_j^{\pm}$ can be easily obtained from a basic golden-rule calculation:[1]

$$\Gamma_j^{\pm}(n) = \frac{1}{R_j e^2}\left[\frac{-\Delta E_j^{\pm}}{1-\exp(\Delta E_j^{\pm}/k_B T)}\right], \tag{1}$$

where ΔE is the energy change of the system when the electron tunnels across the barrier and R_j is the tunneling resistance of the jth junction. The equations for ΔE are obtained from electrostatic energy considerations ($e > 0$):

$$\Delta E_1^{\pm} = \Delta U^{\pm} \pm \frac{eC_2}{C_\Sigma} V ,$$

$$\Delta E_2^{\pm} = \Delta U^{\pm} \mp \frac{eC_1}{C_\Sigma} V ,$$

$$\Delta U^{\pm} = \frac{(Q \pm e)^2}{2C_\Sigma} - \frac{Q^2}{2C_\Sigma} ,$$

where Q is the excess charge on the center electrode before the electron tunnels, C_j is the capacitance of the jth junction, and $C_\Sigma = (C_1 + C_2)$. ΔU is just the change in charging energy of the center electrode as it gains or loses one electron. The second term in $\Delta E_{1,2}^{\pm}$ is the potential difference across the barrier times the electron charge. By letting $Q = (ne - Q_0)$, where n is the integer nearest Q/e, i.e., $|Q_0| \leq e/2$, we can rewrite the above equations as

$$\Delta E_1^{\pm} = \frac{e}{C_\Sigma}\left[\frac{e}{2} \pm (ne - Q_0) \pm C_2 V\right] ,$$

$$\Delta E_2^{\pm} = \frac{e}{C_\Sigma}\left[\frac{e}{2} \pm (ne - Q_0) \mp C_1 V\right] . \tag{2}$$

Q_0 represents the fractional electron charge present on the particle when the voltage electrode in Fig. 1(a) is floating. The origin of the fractional charge Q_0 will be discussed

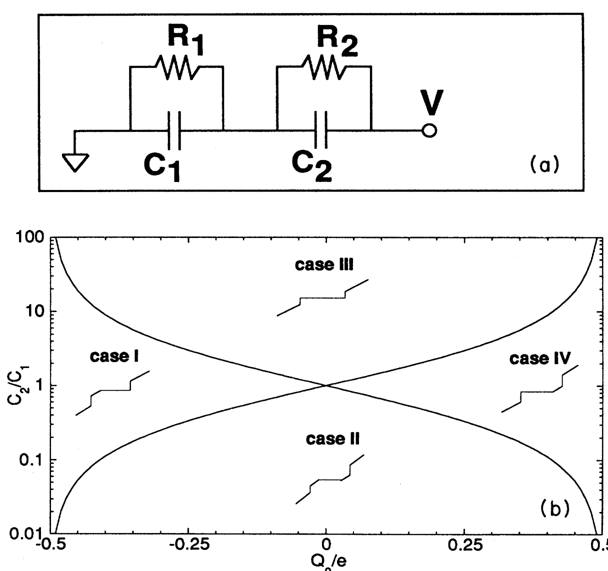

FIG. 1. (a) Schematic of the experiment. (b) The four qualitatively different cases of I-V curves. Within each case is shown a small trace which is representative of data in that region, although the shape of the trace can be varied greatly within each case by changing Q_0 or C_2/C_1. It is the order of the conduction onsets which is fixed in each case.

later. The current is then given by

$$I(V)=e\sum_{n=-\infty}^{\infty}\sigma(n)[\Gamma_2^+(n)-\Gamma_2^-(n)]$$
$$=e\sum_{n=-\infty}^{\infty}\sigma(n)[\Gamma_1^-(n)-\Gamma_1^+(n)]\,, \qquad (3)$$

where $\sigma(n)$ is the ensemble distribution of the number of electrons on the center electrode. The distribution $\sigma(n)$ is obtained by noting that the net probability for making a transition between any two adjacent states in steady state is zero; thus:

$$\sigma(n)[\Gamma_1^+(n)+\Gamma_2^+(n)]$$
$$=\sigma(n+1)[\Gamma_1^-(n+1)+\Gamma_2^-(n+1)]\,. \qquad (4)$$

Since the $\Gamma_j^\pm$ are known from (1) and (2), this allows one to solve for the distribution $\sigma(n)$, subject to the normalization condition, $\sum_{n=-\infty}^{\infty}\sigma(n)=1$. We can thus numerically solve for $I(V)$ from (3).

It is possible, however, to obtain a simple analytic expression for $I(V)$ by considering the limit where $R_2/R_1 \gg 1$. The most probable number of electrons on the center electrode, n_0, i.e., the value for which $\sigma(n_0) \geq \sigma(n_0 \pm 1)$, is then primarily determined by junction 1 because $\Gamma_1 \gg \Gamma_2$ in (4). Combined with (4) and (1), this maximum probability condition requires that

$$e^{-1}(-C_2V+Q_0-e/2)\leq n_0 \leq e^{-1}(-C_2V+Q_0+e/2)\,. \qquad (5)$$

$\sigma(n)$ is expected to be sharply peaked if $|\Delta E_1| \gg k_BT$. That is, $\sigma(n)\approx \delta_{n,n_0}$ at low temperatures. The net current (3) is then given by

$$I(V)=e[\Gamma_2^+(n_0)-\Gamma_2^-(n_0)]\,.$$

Note that junction 1, having a much higher tunneling rate, determines n_0, while junction 2, with the smaller tunneling rate, responds to this constant n_0 by adjusting the current correspondingly. For low temperatures such that $|\Delta E_2^\pm(n_0)| \gg k_BT$, we can simplify $\Gamma_2(n_0)$ to obtain

$$\Gamma_2^\pm(n_0)=\begin{cases}\dfrac{1}{R_2C_\Sigma}[-\tfrac{1}{2}\mp(n_0-Q_0/e)\pm C_1V/e], & \text{for } \Delta E_2^\pm<0\\ 0, \ \text{for } \Delta E_2^\pm>0\,. & \end{cases}$$

This means that $I(V)=0$ when

$$(-e/2+n_0e-Q_0)/C_1\leq V\leq (e/2+n_0e-Q_0)/C_1\,. \qquad (6a)$$

This is the so-called Coulomb blockade. Outside of this voltage range $I(V)$ is given by:[1]

$$I(V)=\frac{1}{R_2C_\Sigma}\left[-(n_0e-Q_0)+C_1V-\frac{e}{2}\,\mathrm{sgn}(V)\right], \qquad (6b)$$

where n_0 is obtained from (5). Equations (5) and (6) form the basis for interpreting our data. They are used qualitatively to identify the "case" to which a given I-V curve belongs and to determine parameter values from specific features in the data. These parameter values are then used in (3), the exact $I(V)$ solution, for quantitative comparison with the data.

Our experimental setup consisted of a low-temperature scanning tunneling microscope (STM), tunneling into a gold surface. The STM used a chemically etched platinum-iridium tip. Images of the gold surface show a very grainy surface. It is believed that the two-junction system was created by the presence of a grain or some equivalent contaminant[3] between the tip and the bulk of the sample. The sample was voltage biased and the data were taken at 4.2 K.

In a two-junction system, the most commonly observed feature of a Coulomb blockade is an offset in the asymptote of the I-V curve.[4] Our analysis gives more detail. Equation (5) shows that $Q=(n_0e-Q_0)$ jumps from $(-C_2V-e/2)$ to $(-C_2V+e/2)$ each time n_0 jumps by 1. From (6b) the corresponding currents are V/R_2 and $(V-e/C_\Sigma)/R_2$, for $V>0$. Thus (6b) describes a series of sloping steps within an envelope with slope $dI/dV=1/R_2\approx 1/R_\Sigma$, and *average* offset $e/2C_\Sigma$. When the capacitances are as small as 10^{-18}–10^{-19} F, however, the asymptotic limit is not well defined in our range of measurement, which is only a few volts. The data in this relatively small voltage region show detailed features about the conduction onset in a two-junction system. In particular, our work shows that there are distinct cases which differ according to whether the onset of conduction is caused by (i) n_0 changing, because V reaches the limits in (5) before those in (6a), or, (ii) overcoming the Coulomb blockade of C_1, because V reaches the limits in (6a) before those in (5). The former causes a discrete jump in current, while the latter causes the onset of a linear increase in current. Taking account of positive and negative voltages separately, we distinguish four cases which are illustrated schematically in Fig. 1(b).

The four cases identified above are exactly what is observed in our experiments. By noting the sequence of conduction onsets, one can identify the case to which a measured I-V belongs, and then proceed to experimentally determine the model parameters, including Q_0. In our approximation, where $R_2 \gg R_1$ and $T\approx 0$, Eqs. (5) and (6) give simple expressions for distinct features in the data. For example, from (5) we obtain that the difference between the threshold voltages for the first positive step and first negative step is e/C_2, while their sum is $2Q_0/C_2$. Similarly, (6) shows that the plateau slope on a step (n_0 constant) is C_1/R_2C_Σ. Some of these quantities are noted in Fig. 2. These preliminary parameter values are refined for finite temperatures by using (3), (4), and the normalization of $\sigma(n)$ to compute the dotted curve,[5] which has $r\equiv R_2/R_1=\infty$ and $T=4.2$ K. Since, the position of the noted features are *not* affected be changing r, we then simply decrease r until the best fit is obtained. This is the solid theoretical trace (usually hidden by the experimental data points). The dashed curve is (6) calculated with the same parameters except with $r=\infty$ and $T=0$, to show explicitly the nature of the finite-temperature effects in our data.

Rounding in the steps arises from two sources, a nonzero temperature and a finite R_2/R_1. Finite temperature causes a symmetric broadening of the current step jumps. In our experiment, at 4.2 K, $k_BT<0.02(e^2/C_\Sigma)$ and thermal effects are small. The finite ratio R_2/R_1 leaves the linear conduction onsets very sharp, but the step onsets ($n_0 \rightarrow n_0 \pm 1$) are no longer vertical jumps connecting plateaus with universal slope C_1/R_2C_Σ, as in the simplified result (6). Rather, the onsets have a finite slope and the step top rounds toward a plateau slope of C_1/R_2C_Σ [Fig. 3(a)]. The step onset can be distinguished from a linear increase by noting that the slope of the linear increase is the same as the slope of all the step plateaus, while the step onset will have a much larger slope. This distinction can be clearly seen in Figs. 2(a) and 3(a).

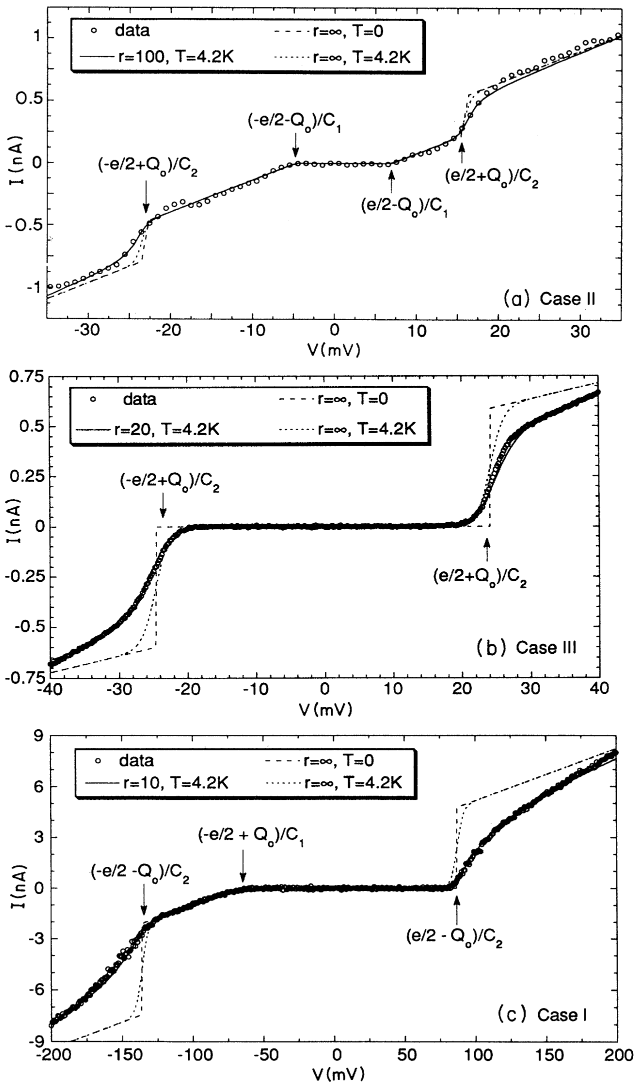

FIG. 2. Data for various cases compared to exact theoretical curves and two levels of approximation. (a) Case II, (b) case III, and (c) case I. The measured parameter values are (a) $C_1=1.36\times10^{-17}$ F, $C_2=4.05\times10^{-18}$ F, $R_1=0.3$ MΩ, $R_2=29.3$ MΩ, and $Q_0=-0.096e$, (b) $C_1=1.64\times10^{-18}$ F, $C_2=3.28\times10^{-18}$ F, $R_1=2.0$ MΩ, $R_2=39.2$ MΩ, and $Q_0=-0.005e$, and (c) $C_1=7.2\times10^{-19}$ F, $C_2=4.05\times10^{-19}$ F, $R_1=1.7$ MΩ, $R_2=16.6$ MΩ, and $Q_0=-0.11e$. All data were taken at $T=4.2$ K and fitted using that value.

Figure 3(a) shows data in a wider voltage range where successive steps become apparent, forming the Coulomb staircase. From (5) and (6) it is clear that the width of the individual steps is e/C_2, and that the slope on the step plateau is C_1/R_2C_Σ, in all cases; this agrees extremely well with our data, where the values of C_1 and C_2 are already uniquely obtained from the central region. Also, as Q_0 approaches $\pm e/2$ in any case, the width of the zero conductance region goes to zero as shown in Fig. 3(b). This is the single-electron transistor effect in which by changing Q_0 from 0 to $\pm e/2$ we change the conductance of the center region from zero to a finite value. A close look at measurements reported by other researchers shows that their results are well accounted for in our model. In particular some unexplained asymmetric traces are perfect examples of cases I and IV [Fig. 1(b)].[6-8] It is also worth noting how closely class-III traces resemble superconductor-normal-metal STM measurements reported as showing large superconducting gaps.[9,10]

In our experiment Q_0 is thought to originate from the

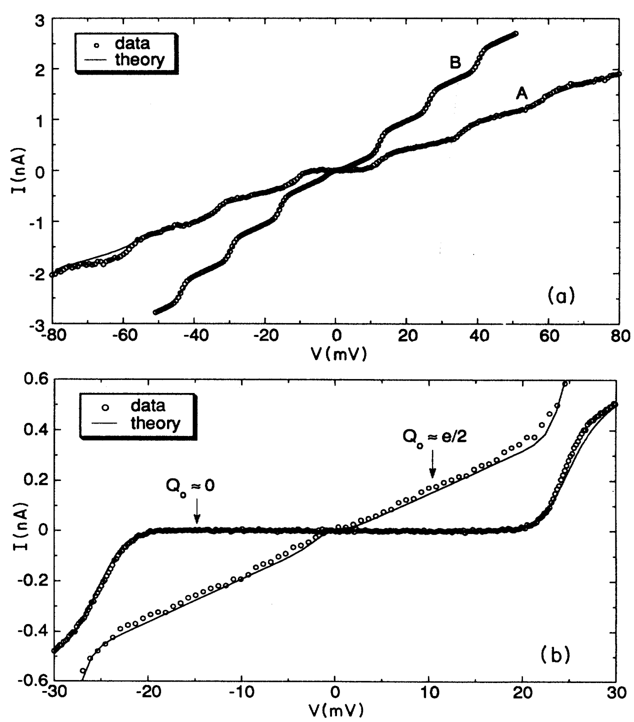

FIG. 3. (a) Coulomb-staircase data continue to fit theory outside of the central region. By moving the STM tip closer to the sample, we go from A to B, decreasing the total resistance of the system (asymptotic slope). This movement also changes the geometric capacitance, C_1 or C_2, and hence changes Q_0. This capacitively induced Q_0 change can be seen more dramatically in the central region of data in another junction, Fig. 3(b).

TABLE I. Width of the zero conduction region, ΔV_0, for different cases. The parenthetical restrictions show that the necessary symmetry of the $Q_0 - C_2/C_1$ phase space is preserved. Note that in cases I and IV, $\Delta V_0 \rightarrow 0$ as $|Q_0| \rightarrow e/2$, while in cases II and III, $\Delta V_0 = e/\max(C_2, C_1)$.

Case	ΔV_0
I ($Q_0 < 0$)	$(e/2 + Q_0)(C_\Sigma/C_1C_2)$
II ($C_1 > C_2$)	e/C_1
III ($C_2 > C_1$)	e/C_2
IV ($Q_0 > 0$)	$(e/2 - Q_0)(C_\Sigma/C_1C_2)$

difference in work functions of the different metals used in the junctions. If so, the more fundamental variable is the contact potential across the junction. $Q_0 (\mathrm{mod}\ e)$ is obtained from

$$Q_0 = \frac{1}{e}[C_1(\Delta\phi_1) - C_2(\Delta\phi_2)]$$

where the $\Delta\phi_1$ and $\Delta\phi_2$ are the contact potentials across junctions 1 and 2, respectively. By changing the tip-grain distance we can change the capacitance C_1 or C_2 and hence change Q_0 as well. This allows us to move systematically around the phase space of Fig. 1(b), taking data in the various sections. We have observed the oscillation of the width of the zero conduction region caused by varying the tip-to-grain capacitance sufficiently to change Q_0 through multiples of e. A detailed report of these data is in preparation.

We would like to thank M. Tuominen, M. Rzchowski, and R. Fitzgerald for their valuable input on presenting and explaining the data, and also D. W. Abraham, C. J. Lobb, and E. Ganz on whose STM design our instrument is based. This research was supported in part by NSF Grants No. DMR89-12927 and No. DMR89-20490, and by ONR Grants No. N00014-89-J-1565 and No. N00014-89-J-1023.

[1]D. V. Averin and K. K. Likharev, *Mesoscopic Phenomena in Solids,* edited by B. L. Altshuler, P. A. Lee, and R. A. Webb (Elsevier, Amsterdam, 1991), p. 169.

[2]M. Amman, R. Wilkins, E. Ben-Jacob, P. D. Maker, and R. C. Jaklevic, Phys. Rev. B **43**, 1146 (1991).

[3]R. Berthe and J. Halbritter, Phys. Rev. B **43**, 6880 (1991).

[4]P. J. M. van Bentum, H. van Kempen, L. E. C. van de Leemput, and P. A. A. Teunissen, Phys. Rev. Lett. **60**, 369 (1988).

[5]The program was created by A. Karatkov and K. K. Likharev.

[6]K. A. McGreer, J-C. Wan, N. Anand, and A. M. Goldman, Phys. Rev. B **39**, 12260 (1989). The asymmetric trace *d* in Fig. 3 shows linear conductance on the negative bias side of the central region. This is characteristic of case I.

[7]R. Wilkins, M. Amman, E. Ben-Jacob, and R. C. Jaklevic, Phys. Rev. B **42**, 8698 (1990). Trace *D* in Fig. 2, is asymmetric due to a linear conductance on the positive bias side of the central region. This is characteristic of case IV.

[8]P. J. M. van Bentum, R. T. M. Smokers, and H. van Kempen, Phys. Rev. Lett. **60**, 2543 (1988). Trace *a* in Fig. 4, is asymmetric due to a linear conductance on the positive bias side of the central region. This is characteristic of case IV.

[9]Chen Wang, B. Giambattista, C. G. Slough, R. V. Coleman, and M. A. Subramanian, Phys. Rev. B **42**, 8890 (1990).

[10]S. Vieira, J. G. Rodrigo, M. A. Ramos, K. V. Rao, and Y. Makino, Phys. Rev. B **40**, 11403 (1989).