Cotunneling thermopower of single electron transistors

M. Turek* and K. A. Matveev

Department of Physics, Duke University, Durham, North Carolina 27708-0305

We study the thermopower of a quantum dot weakly coupled to two reservoirs by tunnel junctions. At low temperatures the transport through the dot is suppressed by charging effects (Coulomb blockade). As a result the thermopower shows an oscillatory dependence on the gate voltage. We study this dependence in the limit of low temperatures where transport through the dot is dominated by the processes of inelastic cotunneling. We also obtain a crossover formula for intermediate temperatures which connects our cotunneling results to the known sawtooth behavior in the sequential tunneling regime. As the temperature is lowered, the amplitude of thermopower oscillations increases, and their shape changes qualitatively.

I. INTRODUCTION

In the last few years a number of experiments have been performed in order to investigate the transport of electrons through small conductors, such as metallic particles and semiconductor quantum dots.[1,2] One of the most commonly studied types of devices is the *single electron transistor* which consists of a quantum dot (or a small metal particle) connected to two leads by tunnel junctions. The particle is usually also capacitively coupled to an additional gate electrode, Fig. 1. The transport of electrons through the quantum dot is strongly affected by charging effects. Indeed, when an electron tunnels into the dot from a lead, the electrostatic energy of the system increases by $\sim E_C \equiv e^2/2C$, where e is the elementary charge and C is the capacitance of the dot. In a typical experiment the temperature is low, $T \ll E_C$, and the tunneling is strongly suppressed as only a very small fraction of electrons have energy of the order of E_C necessary for the tunneling to occur. This phenomenon is commonly referred to as the Coulomb blockade. In a single electron transistor, the charging energy can be controlled by the gate voltage V_g. For instance, by applying a positive voltage to the gate, one can lower the increase in electrostatic energy caused by adding an extra electron to the dot. As a result at certain values of V_g the electrostatic gap vanishes, and the transport of electrons through the system is strongly enhanced. This leads to a sequence of periodic peaks in the conductance of the single electron transistor as a function of V_g, which is often observed in experiments.[2] The positions and shape of the peaks are in good agreement with the so-called sequential tunneling theory[3,4] of conductance which accounts for real processes of electron tunneling between the dot and leads.

In a number of more recent experiments the thermopower S of a single electron transistor has been studied.[5–8] Similarly to the case of conductance, the thermopower shows periodic oscillations as a function of the gate voltage. The theory of the thermopower in a single electron transistor in the framework of the sequential tunneling approach was developed by Beenakker and Staring.[9] It was confirmed by the experimental data of Ref. 5, where the sawtooth shape of the thermopower oscillations predicted in Ref. 9 was observed. On the other hand, in a recent experiment[7] the observed thermopower oscillations had a shape different from the saw-

tooth, and also the amplitude of the oscillations was far smaller than the one predicted in Ref. 9.

The most significant difference between the experiments of Ref. 5 and Ref. 7 was that the temperature in the latter work was very small, $T \sim 0.006 e^2/C$, in comparison with the estimated temperature $T \sim 0.065 e^2/C$ in Ref. 5. The authors of Ref. 7 attributed the deviation of their data from the theory[9] to the fact that the theory neglected the effects of virtual tunneling (*cotunneling*) of electrons through the dot. Indeed, it is known that in the case of very low temperatures cotunneling processes[11] give a dominant contribution to the conductance of single electron transistors. One can therefore expect that this mechanism will result in a different behavior of the thermopower at low temperatures.

In this paper we develop the theory of the thermopower of single electron transistors in the regime of inelastic cotunneling. This mechanism dominates the low-temperature electron transport in the case of relatively large dots, where the effects of finite quantum level spacing can be neglected. We find the thermopower oscillations of the shape qualitatively similar to that observed in the experiment,[7] and the amplitude of the oscillations is significantly lower than the result of the sequential tunneling theory.[9] We also study the crossover from
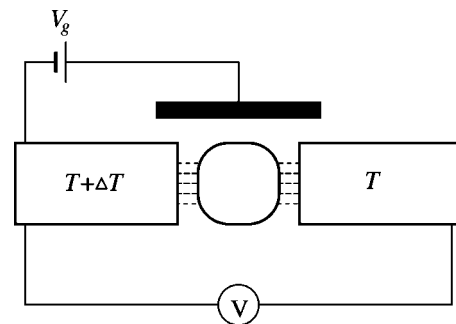

FIG. 1. Thermopower measurement in a single electron transistor. The quantum dot is coupled to the two leads by tunnel barriers. The electrostatics of the system is controlled by the voltage V_g applied to a gate coupled capacitively to the quantum dot. The left and right leads are held at temperatures $T + \Delta T$ and T, respectively. As a result a small voltage V is generated between the leads. The resulting thermopower $S = -V/\Delta T$ is measured as a function of V_g.

the cotunneling behavior of the thermopower to the sawtooth regime of sequential tunneling which occurs as one raises the temperature T of the system above a certain crossover temperature T_c. The latter is found to be $T_c \simeq E_C / \ln[e^2/\hbar(G_l + G_r)]$, where G_l and G_r are the conductances of the tunneling junctions connecting the quantum dot to the left and right leads. Throughout the paper G_l and G_r are assumed to be small, $G_l, G_r \ll e^2/\hbar$. The opposite regime of strong tunneling, more relevant for the conditions of the experiment,[8] was recently addressed in Ref. 10.

In the next section we introduce the theoretical model of a single electron transistor used in this paper and discuss the relevant mechanisms of electronic transport in the device. In Sec. III we review the known results for the thermopower S in the regime of sequential tunneling and obtain S in the regime of inelastic cotunneling. In Sec. IV these results are unified in a single formula that correctly describes the crossover between the two regimes.

II. MECHANISMS OF TRANSPORT

To describe the single electron transistor, Fig. 1, we introduce the Hamiltonian in the form $\hat{H} = \hat{H}_0 + \hat{V}$, where

$$\hat{H}_0 = \sum_k \xi_k a_k^\dagger a_k + \sum_p \xi_p a_p^\dagger a_p + \sum_q \xi_q a_q^\dagger a_q + \frac{\hat{Q}^2}{2C} + \phi \hat{Q},$$
(1)

$$\hat{V} = \sum_{k,p_1} (t_{k,p_1} a_k^\dagger a_{p_1} + t_{k,p_1}^* a_{p_1}^\dagger a_k)$$

$$+ \sum_{p_2,q} (t_{p_2,q} a_{p_2}^\dagger a_q + t_{p_2,q}^* a_q^\dagger a_{p_2}),$$
(2)

$$\hat{Q} = -e \sum_p [a_p^\dagger a_p - \theta(-\xi_p)].$$
(3)

Here a_k, a_p, and a_q are the annihilation operators for the electrons in the left lead, quantum dot, and the right lead, respectively; the electron energies ξ_k, ξ_p, and ξ_q are measured from the Fermi level; the tunneling in and out of the dot is described by matrix elements t_{k,p_1} and $t_{p_2,q}$. Operator $\hat{Q}$ represents the charge of the quantum dot, e is the elementary charge, and the potential ϕ is proportional to the gate voltage V_g.

Our goal is to find the thermopower S of the system as a function of the potential ϕ for a given temperature. The thermopower is defined in terms of the voltage V generated across the single electron transistor when the temperatures of the left and right leads T_l and T_r differ by $\Delta T \ll T_l, T_r$ and no current I through the system is allowed:

$$S \equiv - \lim_{\Delta T \to 0} \frac{V}{\Delta T} \bigg|_{I=0}.$$
(4)

In the linear response regime the current can be presented as

$$I = G_T \Delta T + G V,$$
(5)

where G is the conductance of the system and the kinetic coefficient G_T describes the current response to the temperature difference. The condition $I = 0$ then results in the following expression for the thermopower:

$$S = \frac{G_T}{G}.$$
(6)

Thus, one can find the thermopower S by calculating the kinetic coefficients G_T and G.

In this paper we concentrate on the case of relatively large quantum dots, where the effects of finite quantum level spacing δ in the dot can be ignored. In the limit $\delta \to 0$ the transport of electrons in single electron transistors can be accomplished via either sequential tunneling or inelastic cotunneling mechanisms. The respective contributions to the linear conductance G are well known. The conductance as a function of the gate voltage ϕ shows periodic peaks centered at points

$$\phi_N = \frac{e}{C}\left(N + \frac{1}{2}\right),$$
(7)

where the charging energies of the dot with N and $N+1$ additional electrons are equal. Sequential tunneling theory[3,4] accounts for the real events of electron tunneling between the leads and the dot. When the $(N+1)$st electron tunnels into the dot, the charging energy changes by $e(\phi_N - \phi)$. At low temperatures $T \ll e|\phi_N - \phi|$, the density of electrons with energy sufficient to charge the dot is exponentially small, resulting in conductance peaks with exponential tails:

$$G^{sq} = \frac{G_l G_r}{2(G_l + G_r)} \frac{e(\phi - \phi_N)/T}{\sinh[e(\phi - \phi_N)/T]}.$$
(8)

On the other hand, the inelastic cotunneling mechanism[11] accounts for the second-order tunneling processes when, e.g., an electron tunnels from the left lead into the dot and then another electron tunnels from the dot to the right lead. The initial and final states of such a process have the same charge in the dot. The Coulomb blockade only affects the energy of the virtual state, resulting in only a power-law suppression of the conductance at low temperatures:

$$G^{co} = \frac{\pi}{3} \frac{\hbar}{e^2} G_l G_r \frac{T^2}{[e(\phi - \phi_N)]^2}.$$
(9)

The above expression formally diverges at $\phi = \phi_N$, because the calculation in Ref. 11 neglected the contribution of the quasiparticle energies to the energy of the virtual state. To estimate the cotunneling contribution at the center of a peak, i.e., at $\phi = \phi_N$, one can replace the charging energy difference $e(\phi - \phi_N)$ in the denominator of Eq. (9) by the temperature T. This results in $G^{co} \sim (\hbar/e^2) G_l G_r$. The contribution (8) of the sequential tunneling mechanism at the peak is $G^{sq} \sim G_l G_r / (G_l + G_r)$. Thus, near the centers of the peaks the conductance is dominated by sequential tunneling.

On the other hand, between the peaks the sequential tunneling contribution (8) decays exponentially at $T \to 0$, as opposed to the relatively slow dependence $G^{co} \propto T^2$ of the co-

tunneling conductance (9). Therefore, at low enough temperatures the cotunneling mechanism will dominate the conduction in the valleys between the peaks. By comparing the contributions (8) and (9) in the middle of a valley, i.e., for $\phi - \phi_N = e/2C$, we find that the sequential tunneling dominates the conductance at any gate voltage only at temperatures $T > T_c$, where

$$T_c \simeq \frac{E_C}{\ln[e^2/\hbar(G_l + G_r)]}. \tag{10}$$

At lower temperatures $T < T_c$, the conductance in the regions of width

$$\Delta \phi = \frac{e}{2C} \frac{T}{T_c} \simeq \frac{T}{e} \ln\left[\frac{e^2}{\hbar(G_l + G_r)}\right] \tag{11}$$

around each peak is determined by the sequential tunneling processes, Eq. (8), whereas outside those regions cotunneling processes dominate, Eq. (9). Note that in our case of weak tunneling, $G_l, G_r \ll e^2/\hbar$, the region (11) is wider than the thermal width of the peak, T/e. Thus, the cotunneling mechanism becomes important only away from the peaks, where the conductance is much smaller than its peak value.

The change in the transport mechanism from sequential tunneling to cotunneling is more dramatic in the case of the thermopower. Unlike the conductance, the thermopower in the sequential tunneling regime does not have the form of sharp peaks near $\phi = \phi_N$; in fact it reaches its maximum values near the centers of the valleys between the peaks of $G(\phi)$ (Ref. 9). At $T < T_c$ the transport in those regions is strongly affected by the cotunneling mechanism, which leads to qualitative changes in both the amplitude and shape of the Coulomb blockade oscillations of the thermopower.

III. THERMOPOWER IN THE REGIMES OF SEQUENTIAL TUNNELING AND COTUNNELING

The thermopower S of a system is a direct measure of the average energy that the electrons carry during the tunneling processes. This follows from an Onsager relation[12] between the Peltier coefficient $\Pi = -\langle \xi \rangle / e$ and the thermopower:

$$S = \frac{\Pi}{T} = -\frac{\langle \xi \rangle}{eT}. \tag{12}$$

The two transport mechanisms in single electron transistors—sequential tunneling and cotunneling—involve electrons with typical energies $\xi \sim E_C$ and $\xi \sim T$, respectively. Thus the thermopower (12) is strongly affected by the crossover between the two regimes at $T \sim T_c$.

A. Sequential tunneling regime

In this section we review the results for the thermopower in the sequential tunneling regime found by Beenakker and Staring.[9] The sequential tunneling current is determined by the following two elementary real transition processes: (a) an electron tunnels between the left lead and the dot and (b) an electron tunnels between the dot and the right lead. Both

processes are of the first order in the perturbation parameter $G_{l,r}/(e^2/\hbar)$. The current in the stationary state can then be derived by means of a kinetic equation which involves the probabilities for the system being in a certain charge state and the tunneling rates obtained by Fermi's golden rule.

In the linear-response regime the current can be due to either a bias voltage or a temperature difference between the two leads. The conductance G and the kinetic coefficient G_T are found in Refs. 4 and 9, respectively:

$$G^{sq} = \frac{G_l G_r}{G_l + G_r} \sum_N W_N^{(0)} f\left(\frac{E_N - E_{N-1}}{T}\right), \tag{13}$$

$$G_T^{sq} = -\frac{1}{2e} \frac{G_l G_r}{G_l + G_r} \sum_N W_N^{(0)} f\left(\frac{E_N - E_{N-1}}{T}\right) \frac{E_N - E_{N-1}}{T}. \tag{14}$$

Here $E_N \equiv E_C N^2 - Ne\phi$ is the electrostatic energy of the dot containing N electrons, $f(x) \equiv x/(1 - e^{-x})$, and $W_N^{(0)}$ stands for the equilibrium probability distribution of the dot charge, $W_N^{(0)} \equiv e^{-E_N/T} / \sum_N e^{-E_N/T}$.

At low temperatures $T \ll E_C$ at most two charge states contribute significantly to the sums in Eqs. (13) and (14). One can then neglect the exponentially small contributions of the other charge states and obtain Eq. (8) for the conductance G as well as an analogous expression for G_T:

$$G_T^{sq} = \frac{G_l G_r}{4(G_l + G_r)} \frac{e(\phi - \phi_N)^2/T^2}{\sinh[e(\phi - \phi_N)/T]}. \tag{15}$$

Using Eq. (6) one then obtains the low-temperature limit of the thermopower:

$$S^{sq} = \frac{\phi - \phi_N}{2T}. \tag{16}$$

Here, ϕ_N defines the position (7) of the conductance peak closest to ϕ. This is the sawtoothlike behavior described in Ref. 9. The amplitude of the Coulomb blockade oscillations of the thermopower is given by $S_{max}^{sq} = e/(4CT)$.

This result can be understood in terms of the average energy of tunneling electrons. Exactly in the middle between two conductance peaks (e.g., $\phi = [\phi_{N-1} + \phi_N]/2$) the same amount E_C of energy is required to either add or remove an electron from the dot. Therefore, the two processes involving electrons with energies $\xi \approx E_C$ and $\xi \approx -E_C$ contribute equally to the electronic transport, and the average energy of tunneling electrons $\langle \xi \rangle$ vanishes. However, if the gate potential ϕ is slightly changed by $\sim T/e$ towards one of the conductance peaks, one of the two charge states is exponentially suppressed. In this case the average energy is found to be $\langle \xi \rangle \approx \pm E_C$. This means that the thermopower exhibits a sharp jump between two conductance peaks. If the gate voltage is tuned further towards ϕ_N, the average tunneling electron energy is given by $\langle \xi \rangle \sim u_1 \equiv E_{N+1} - E_N = e(\phi_N - \phi)$ and thus changes linearly with ϕ. Using this qualitative considerations the result (16) is reproduced except for an additional numerical factor of 1/2. This factor 1/2 is due to the fact that not only electrons with energy u_1 above the Fermi

level are involved in the transport. The probability for an electron with energy ξ to tunnel into the dot is proportional to the average occupation number $\sim e^{-\xi/T}$ in the lead and to the density of holes $\sim e^{-(u_1-\xi)/T}$ in the dot. Therefore, all electrons with energies less than u_1 above the Fermi level have the same probability to tunnel into the dot. For that reason, one obtains for the average tunneling electron energy $\langle\xi\rangle = u_1/2$ which then reproduces the result (16).

B. Cotunneling regime

As we discussed in Sec. II, at temperatures below the crossover temperature (10) cotunneling is the main mechanism of transport in a single electron transistor. A cotunneling process consists of two steps. First, an electron close to the Fermi level in one lead tunnels into the dot, thus changing the energy by $\sim u_1$. Since the energy is not conserved during this process, the system is in a virtual intermediate state. A second tunneling process has to follow which involves an electron tunneling from the dot to the other lead and thus restoring the conservation of energy. Since the charge of the dot before and after a cotunneling process is the same, the electronic transport is primarily due to electrons with energies $\xi \sim T$. Therefore, unlike sequential tunneling, the cotunneling contribution to the current is not exponentially suppressed at $T \ll E_C$, and it dominates the transport in the valleys of the Coulomb blockade. In the following only inelastic cotunneling processes which involve two different electrons are considered, since the elastic ones are suppressed at small level spacing in the dot, $\delta \ll T^2/E_C$ (Ref. 11).

The cotunneling contribution to the current is of second order in the small parameter $G_{l,r}/(e^2/\hbar)$ because two coherent tunneling processes have to take place. For a given number N of electrons in the dot there are two types of second-order processes which can transfer an electron from one lead to the other (see Fig. 2). The current then reads

$$I = -\frac{2\pi}{\hbar}e\sum_{k,p_1,p_2,q}[n_k(1-n_{p_1})n_{p_2}(1-n_q)$$
$$-n_q(1-n_{p_2})n_{p_1}(1-n_k)]\delta(\xi_k-\xi_{p_1}+\xi_{p_2}-\xi_q)$$
$$\times\left|\frac{t_{k,p_1}t_{p_2,q}}{\xi_k-\xi_{p_1}-u_1}-\frac{t_{k,p_1}t_{p_2,q}}{\xi_k-\xi_{p_1}+u_{-1}}\right|^2, \tag{17}$$

where n_k, n_{p_1}, n_{p_2}, and n_q are the occupation numbers of the respective states; $u_{\pm1} \equiv E_{N\pm1} - E_N$, and N is determined by minimizing the electrostatic energy of the dot for a given potential ϕ. Note that the first product of the occupation numbers corresponds to the current from the left to the right lead, whereas the second one accounts for the current in the opposite direction.

If the bias voltage V is zero and all three electrodes are at the same temperature T, the current from the left lead to the right lead cancels with the one flowing into the opposite direction. This can most easily be seen by exchanging $k \leftrightarrow q$
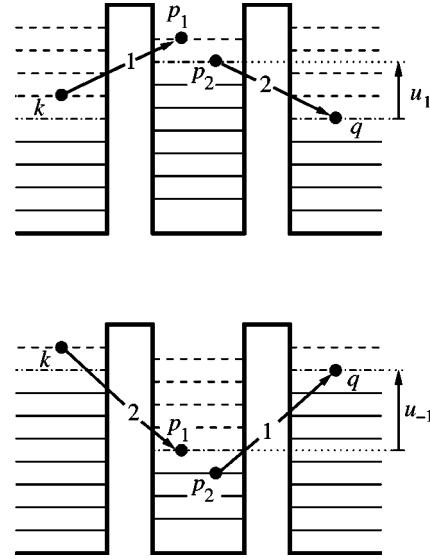

FIG. 2. The two types of inelastic cotunneling processes transferring electrons from the left to the right lead. The first process consists of following two steps: an electron tunnels from a state k in the left lead to a state p_1 in the dot, and then an electron from a state p_2 in the dot tunnels to state q in the right lead. The second process transfers electrons in the opposite order: it starts with an electron tunneling from the dot to the right lead and finishes with another electron tunneling from the left lead into the dot. The energies of the virtual states in the two types of processes are affected by the electrostatic energies u_1 and u_{-1} required to either increase or decrease the charge of the dot by that of one electron. The dash-dotted lines show the positions of the Fermi level in the leads and the dot.

and $p_1 \leftrightarrow p_2$ in one of the products of occupation numbers which is possible because of the apparent left-right symmetry of the system.

The cotunneling contribution G^{co} to the conductance can be derived by linearizing Eq. (17) with respect to the bias voltage V. After replacing the sums by integrals and the occupation numbers by Fermi functions, the conductance can be presented as a single integral:

$$G^{co} = \frac{\hbar G_l G_r}{8\pi e^2 T}\int\frac{\xi^2}{\sinh^2\frac{\xi}{2T}}\left|\frac{1}{\xi-u_1}-\frac{1}{\xi+u_{-1}}\right|^2 d\xi. \tag{18}$$

This expression formally diverges at $\xi = \pm u_{\pm1}$; we will discuss the proper regularization procedure in Sec. IV A. At low temperatures $T \ll u_1, u_{-1}$ the contributions of the regions near the singularities are exponentially small, and one can evaluate G^{co} by neglecting the quasiparticle energies $\xi \sim T$ in the denominators. Then the conductance G^{co} is found[11] to be

$$G^{co} = \frac{\pi\hbar}{3e^2}G_l G_r T^2\left(\frac{1}{u_1}+\frac{1}{u_{-1}}\right)^2. \tag{19}$$

If the potential ϕ is significantly closer to the peak at ϕ_N than to ϕ_{N-1}, the term $1/u_{-1}$ can be neglected compared to $1/u_1$, and Eq. (9) is reproduced.

In order to calculate the kinetic coefficient G_T we assume a slightly higher temperature $T+\Delta T$ in the left lead compared to the dot and the right lead. Then the current (17) is linearized with respect to the small temperature difference $\Delta T \ll T$. After replacing the sums by integrals and occupation numbers by Fermi functions, the following expression for G_T is obtained:

$$G_T^{co} = \frac{\hbar G_l G_r}{2\pi e^3} \int d\xi_k d\xi_{p_1} d\xi_{p_2} d\xi_q \delta(\xi_k - \xi_{p_1} + \xi_{p_2} - \xi_q)$$

$$\times \left| \frac{1}{\xi_k - \xi_{p_1} - u_1} - \frac{1}{\xi_k - \xi_{p_1} + u_{-1}} \right|^2 \frac{\xi_k}{T} \frac{dn(\xi_k)}{d\xi_k}$$

$$\times \{ [1 - n(\xi_{p_1})] n(\xi_{p_2}) [1 - n(\xi_q)]$$

$$+ n(\xi_{p_1}) [1 - n(\xi_{p_2})] n(\xi_q) \}. \tag{20}$$

The same expression is obtained in a more careful treatment where the temperature of the dot is not necessarily equal to the temperature of the right lead. Three of the four integrals can be calculated exactly, resulting in the following expression for G_T:

$$G_T^{co} = -\frac{\hbar G_l G_r}{16\pi e^3} \frac{1}{T^2} \int \frac{\xi^3}{\sinh^2(\xi/2T)} \left| \frac{1}{\xi - u_1} - \frac{1}{\xi + u_{-1}} \right|^2 d\xi. \tag{21}$$

As was the case for the conductance G, the main contribution to G_T comes from energies ξ of the order of $T \ll u_1, u_{-1}$. However, setting those terms in the denominators to zero yields a vanishing G_T as the integrand in Eq. (21) becomes an odd function in ξ. The first nonvanishing contribution to G_T is obtained by expanding the fractions in the integral up to first order in ξ/u_1. The final result for G_T^{co} can be written as

$$G_T^{co} = -\frac{4\pi^3}{15} \frac{\hbar}{e^3} G_l G_r T^3 \left(\frac{1}{u_1} + \frac{1}{u_{-1}} \right) \left(\frac{1}{u_1^2} - \frac{1}{u_{-1}^2} \right). \tag{22}$$

Using the expression (6) for the thermopower we find from Eqs. (19) and (22)

$$S^{co} = \frac{4\pi^2}{5} \frac{T}{e^2} \left(\frac{1}{\phi - \phi_N} + \frac{1}{\phi - \phi_{N-1}} \right). \tag{23}$$

This expression corresponds to the potential range between two conductance peaks, $\phi_{N-1} < \phi < \phi_N$, with the exception of the values of ϕ very close to the peak: namely, $(\phi_N - \phi), (\phi - \phi_{N-1}) \gg T/e$.

The result (23) can be understood in terms of the average energy of tunneling electrons, Eq. (12). The transport in the cotunneling regime is mainly due to the electrons with energies $\xi \sim \pm T$. In the low-temperature limit $T \ll u_1$ the tunneling probability from a state with energy ξ into an intermediate state with energy ξ_p is proportional to

$$w(\xi) \propto \frac{1}{(\xi_p - \xi + u_1)^2} \simeq \frac{1}{u_1^2} \left(1 + 2\frac{\xi - \xi_p}{u_1} \right). \tag{24}$$

This expression clearly shows that the tunneling probability $w(\xi)$ is enhanced for electrons above the Fermi level $\xi > 0$, and therefore the average energy $\langle \xi \rangle$ will not vanish. In fact, with a typical energy $\xi \sim T$ for electrons involved in the tunneling and the probability (24) we find for the average energy $\langle \xi \rangle \sim T^2/u_1$. Using the relation (12), the cotunneling contribution to the thermopower (23) is reproduced in the correct order of magnitude.

IV. THERMOPOWER AT ARBITRARY GATE VOLTAGE

In the previous section expressions for the thermopower in both the sequential tunneling and cotunneling regime were presented. The sequential tunneling result is given by Eq. (16). It dominates the thermopower for gate potentials ϕ close to the positions of the conductance peaks ϕ_N. On the other hand, the cotunneling result (23) gives the correct description of the thermopower between the conductance peaks $\phi_{N-1} < \phi < \phi_N$ if the temperature T is below the crossover temperature T_c. In this section we find the combined contributions of both transport mechanisms and obtain an expression for the thermopower valid at any gate voltage.

The current through the quantum dot is the sum of the two contributions: sequential tunneling and cotunneling. Therefore using Eqs. (5) and (6) we find for the thermopower

$$S = \frac{G_T^{sq} + G_T^{co}}{G^{sq} + G^{co}}. \tag{25}$$

The well-known[4,9] sequential tunneling contributions G^{sq} and G_T^{sq} are valid for the entire potential range. However, the cotunneling results (19) and (22) are only valid away from the centers of the conductance peaks, i.e., $(\phi_N - \phi), (\phi - \phi_{N-1}) \gg T/e$, and diverge close to the peaks. In the following we will show how to correctly regularize these divergences. First we discuss the thermopower (25) in the limit of very low temperatures $T \ll T_c$. Then we present the more general result valid for higher temperatures.

A. Low-temperature thermopower in the vicinity of a conductance peak

In the regime of very low temperatures $T \ll T_c$ there are at most two charge states—e.g., N and $N+1$ electrons in the dot—that contribute significantly to the total current while the contributions of all other states are exponentially suppressed. The crossover between the sequential tunneling and the cotunneling regimes takes place at a gate potential close to the conductance peak $|\phi - \phi_N| \sim \Delta\phi$, where $\Delta\phi$ is given by Eq. (11). In order to find the behavior of the thermopower (25) in this potential range we can neglect terms $\sim 1/u_{-1}$ occurring in the current (17).

To regularize the singularities at $\xi_k - \xi_{p_1} = u_1$ in the expression (17) for the cotunneling current we follow the approach presented in Ref. 13 and add a small imaginary part Γ to the energy of the intermediate virtual states in Eq. (17):

$$I = -\frac{2\pi}{\hbar}e \sum_{k,p_1,p_2,q} [n_k(1-n_{p_1})n_{p_2}(1-n_q)$$

$$-n_q(1-n_{p_2})n_{p_1}(1-n_k)]\delta(\xi_k-\xi_{p_1}+\xi_{p_2}-\xi_q)$$

$$\times \left| \frac{t_{k,p_1}t_{p_2,q}}{\xi_k-\xi_{p_1}-u_1-i\Gamma} \right|^2. \tag{26}$$

In the limit $\Gamma \to 0$ this expression can be divided into a large part $\sim 1/\Gamma$ and a smaller one that is independent of Γ using the following procedure:

$$\int dE \frac{f(E)}{E^2+\Gamma^2} = \int dE \frac{f(0)}{E^2+\Gamma^2} + \int dE \frac{f(E)-f(0)}{E^2+\Gamma^2}$$

$$\to \frac{\pi}{\Gamma}f(0) + \int dE \frac{f(E)-f(0)}{E^2}, \tag{27}$$

where the last integral is to be understood as the principal value. As we show in the Appendix A, the first term corresponds to sequential tunneling. In particular, the results (8) and (15) are restored in the linear-response regime using the scheme (27) and Eq. (A6) for Γ. The second term represents the correct nondivergent expression for the cotunneling. We apply the scheme (27) by rewriting the cotunneling conductance (18) and the kinetic coefficient (21) according to the following rule:

$$\int dE \frac{f(E)}{E^2} \to \int dE \frac{f(E)-f(0)}{E^2}$$

$$= \int dE \frac{f(E)+f(-E)-2f(0)}{2E^2}. \tag{28}$$

This regularization procedure coincides with the one discussed in Ref. 14 for the case of multichannel tunneling junctions. The result for G^{co} and G_T^{co} then reads

$$G^{co} = \frac{\hbar}{4\pi e^2} G_l G_r \mathcal{F}\left[\frac{e(\phi-\phi_N)}{2T} \right], \tag{29}$$

$$G_T^{co} = \frac{\hbar}{4\pi e^3} G_l G_r \mathcal{F}_T\left[\frac{e(\phi-\phi_N)}{2T} \right], \tag{30}$$

where the functions $\mathcal{F}$ and $\mathcal{F}_T$ are defined as

$$\mathcal{F}(x) \equiv |x| \int_0^\infty \frac{dz}{z^2} \left(\frac{(1+z)^2}{\sinh^2[x(1+z)]} \right.$$

$$\left. + \frac{(1-z)^2}{\sinh^2[x(1-z)]} - \frac{2}{\sinh^2[x]} \right), \tag{31}$$
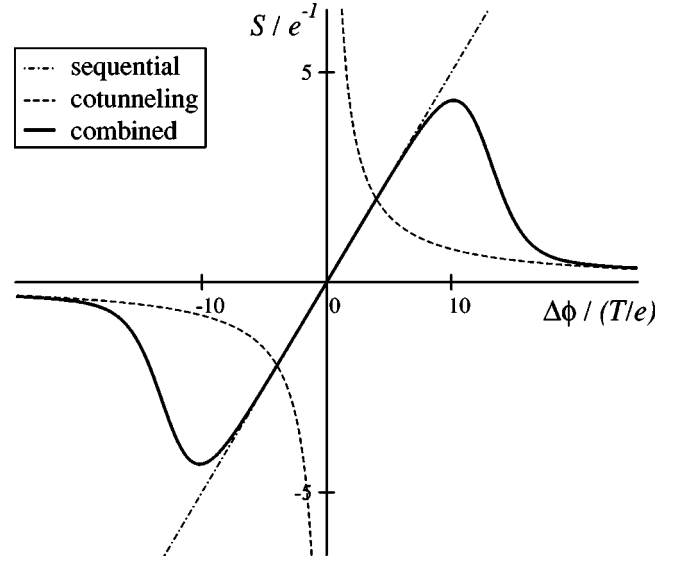

FIG. 3. The thermopower S of a single electron transistor in the vicinity of a conductance peak. The distance from the peak $\Delta\phi$ is assumed to be small compared to the distance to the next peak, $\Delta\phi \ll e/C$. The solid curve is calculated for $G_l+G_r=10^{-3} \times 2\pi e^2/\hbar$ by substituting the expressions (8), (15), (29), and (30) into Eq. (25). The asymptotics of cotunneling and sequential tunneling are shown by dashed and dash-dotted lines, respectively.

$$\mathcal{F}_T(x) \equiv x|x| \int_0^\infty \frac{dz}{z^2} \left(\frac{(1+z)^3}{\sinh^2[x(1+z)]} \right.$$

$$\left. + \frac{(1-z)^3}{\sinh^2[x(1-z)]} - \frac{2}{\sinh^2[x]} \right). \tag{32}$$

The final result for the thermopower is then obtained by substituting the sequential tunneling results (8) and (15) and the regularized cotunneling results (29) and (30) into Eq. (25). The dependence of the thermopower on the gate voltage obtained from those expressions is illustrated in Fig. 3.

As can be seen in Fig. 3, the thermopower is dominated by the sequential tunneling contribution for small $\Delta\phi = \phi - \phi_N$. After it reaches its maximum value S_{max}, the thermopower falls off sharply to merge with the cotunneling result (23) for large $\Delta\phi$. To understand this behavior one should notice that the crossover between the sequential tunneling and the cotunneling occurs at different values $\Delta\phi_1$ and $\Delta\phi_2$ for the conductance G and for G_T. The two crossover values of the gate voltage can be estimated by setting $G^{sq}=G^{co}$ and $G_T^{sq}=G_T^{co}$. At small conductance $g \equiv \hbar(G_l+G_r)/(2\pi e^2) \ll 1$, Eqs. (8), (19), (15), and (22) result in

$$\frac{e\Delta\phi_1}{T} \simeq \ln\left[\frac{1}{g}\left(\ln\frac{1}{g} \right)^3 \right], \tag{33}$$

$$\frac{e\Delta\phi_2}{T} \simeq \ln\left[\frac{1}{g}\left(\ln\frac{1}{g} \right)^5 \right]. \tag{34}$$

The thermopower at $\Delta\phi < \Delta\phi_1$ is given by the sequential tunneling result $S \simeq \Delta\phi/2T$. In the narrow range $\Delta\phi_1 < \Delta\phi < \Delta\phi_2$ the conductance is already dominated by the cotun-

neling, while the main contribution to G_T in the numerator of Eq. (25) is still due to the sequential tunneling. Therefore we find a steeply descending thermopower $S \propto \exp[-e\Delta\phi/T]$ in that region. Eventually, at $\phi > \phi_2$ both the numerator and denominator in Eq. (25) are governed by the cotunneling contribution, and $S \sim T/(e\Delta\phi)$. Therefore the maximum of the thermopower is reached at $\Delta\phi \simeq \Delta\phi_1$ and can be obtained using the sequential tunneling expression (16):

$$S_{max} \simeq \frac{1}{2e} \ln \frac{1}{g}. \tag{35}$$

To find the value of the thermopower after the exponential falloff, one can substitute $\Delta\phi_2$ into the cotunneling expression (23), which gives $S \simeq (4\pi^2/5)e^{-1} \ln^{-1}(1/g) \ll S_{max}$.

B. Thermopower at arbitrary temperatures

In Sec. IV A we studied the thermopower at very low temperatures $T \ll T_c$. In this regime the thermopower in the valleys between the conductance peaks centered at $\phi = \phi_N$ is described by the simple cotunneling result (23); sequential tunneling contributes only in narrow regions around $\phi = \phi_N$. Another interesting regime is that of the temperatures of the order of T_c, where one can explore the crossover from the sawtooth behavior (16) to the correct low-temperature limit.

The crossover takes place in the valleys between the peaks, where more than two charge states give comparable contributions to the current through the device. Thus we present the total cotunneling current as a weighted sum of the contributions of states with a different number N of electrons in the dot, $I = \Sigma W_N^{(0)} I_N^{co}$, where I_N^{co} is given by

$$I_N^{co} = -\frac{2\pi}{\hbar} e \sum_{k,p_1,p_2,q} [n_k(1-n_{p_1})n_{p_2}(1-n_q)$$
$$-n_q(1-n_{p_2})n_{p_1}(1-n_k)]\delta(\xi_k-\xi_{p_1}+\xi_{p_2}-\xi_q)$$
$$\times \left| \frac{t_{k,p_1}t_{p_2,q}}{\xi_k-\xi_{p_1}-E_{N+1}+E_N} - \frac{t_{k,p_1}t_{p_2,q}}{\xi_k-\xi_{p_1}+E_{N-1}-E_N} \right|^2. \tag{36}$$

Note that one has to keep terms $\sim 1/u_{-1}=1/(E_{N-1}-E_N)$ if the expression is to be valid not only close to the peaks at ϕ_N but also in the valleys between the peaks. Similarly to the discussion in the previous section, we linearize the current with respect to a small bias voltage V and to a small temperature difference ΔT, yielding G and G_T. The cotunneling contributions have then to be regularized applying the rule (28) to each of the divergent terms. The result is

$$G^{co} = \frac{\hbar}{e^2} \frac{G_l G_r}{4\pi} \sum_{N=-\infty}^{\infty} \left[(W_{N-1}^{(0)}+W_N^{(0)})\mathcal{F}\left(\frac{E_N-E_{N-1}}{2T}\right) \right.$$
$$\left. -\frac{4TC}{e^2}(W_{N-1}^{(0)}-W_N^{(0)})\mathcal{F}^*\left(\frac{E_N-E_{N-1}}{2T}\right) \right], \tag{37}$$
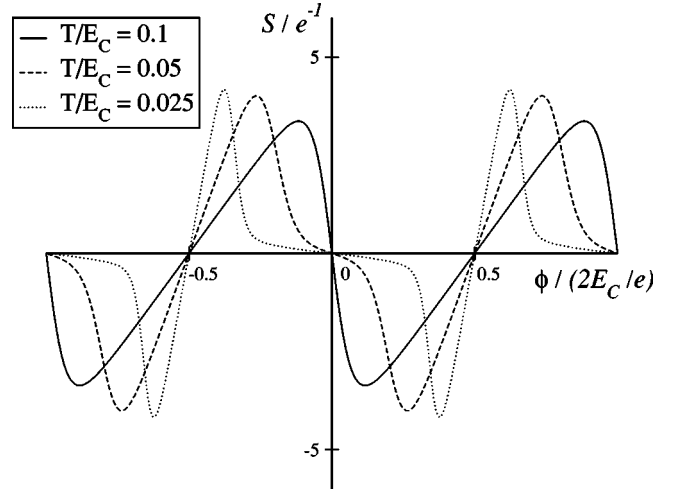

FIG. 4. The thermopower of a single electron transistor at different temperatures. Below the crossover temperature T_c the thermopower in the valleys of Coulomb blockade is suppressed, and the shape of the oscillations changes from the sawtooth (solid line) typical of sequential tunneling to the low-temperature behavior shown by dotted line. The curves were calculated for $G_l + G_r = 10^{-3} \times 2\pi e^2/\hbar$ by substituting Eqs. (13), (14), (37), and (38) into Eq. (25).

$$G_T^{co} = -\frac{\hbar}{e^3} \frac{G_l G_r}{4\pi} \sum_{N=-\infty}^{\infty} \left[(W_{N-1}^{(0)}+W_N^{(0)})\mathcal{F}_T\left(\frac{E_N-E_{N-1}}{2T}\right) \right.$$
$$\left. -\frac{4TC}{e^2}(W_{N-1}^{(0)}-W_N^{(0)})\mathcal{F}_T^*\left(\frac{E_N-E_{N-1}}{2T}\right) \right], \tag{38}$$

where the functions $\mathcal{F}^*$ and $\mathcal{F}_T^*$ are defined by

$$\mathcal{F}^*(x) \equiv x|x| \int_0^\infty \frac{dz}{z}\left(\frac{(1+z)^2}{\sinh^2[x(1+z)]} - \frac{(1-z)^2}{\sinh^2[x(1-z)]}\right), \tag{39}$$

$$\mathcal{F}_T^*(x) \equiv x^2|x| \int_0^\infty \frac{dz}{z}\left(\frac{(1+z)^3}{\sinh^2[x(1+z)]} - \frac{(1-z)^3}{\sinh^2[x(1-z)]}\right). \tag{40}$$

The nondivergent cotunneling contributions G^{co} and G_T^{co} together with the well-known sequential tunneling results (13) and (14) for G^{sq} and G_T^{sq} then give the thermopower according to Eq. (25). This expression for the thermopower S is the final result of the calculation. It is valid at any gate voltage ϕ. As the temperature changes, this result shows a crossover from the sawtooth shape of the Coulomb blockade oscillations of the thermopower at $T \gg T_c$ to the low-temperature behavior discussed in Sec. IV A at $T \ll T_c$ (see Fig. 4). The low-temperature curve of $S(\phi)$ shown by dotted line in Fig. 4 is qualitatively similar to the experimentally measured thermopower of Ref. 7.

V. SUMMARY

We have studied the thermopower of a single electron transistor based on a quantum dot weakly coupled to two

leads, Fig. 1. The transport through the dot is governed by two different electronic transport mechanisms: sequential tunneling and cotunneling. At temperatures above the crossover temperature T_c given by Eq. (10) sequential tunneling dominates. The thermopower in this regime was studied in Ref. 9. At temperatures below T_c the cotunneling mechanism gives the main contribution to the transport in the valleys between the Coulomb blockade peaks of conductance. This changes the shape of the Coulomb blockade oscillations of the conductance dramatically, Fig. 4. We derived a single expression for the thermopower at arbitrary temperature including both contributions. It is given by Eq. (25) in combination with Eqs. (13), (14), (37), and (38). As can be seen in Fig. 4, the resulting thermopower significantly changes its shape from sawtooth behavior for temperatures above T_c to low-temperature behavior at $T \ll T_c$. In the regime of low temperatures the amplitude of the thermopower becomes temperature independent, Eq. (35). Near the maximum of the thermopower the transport is determined by the balance of the two mechanisms. The shape of the thermopower peaks is studied in detail in Sec. IV A and illustrated in Fig. 3.

APPENDIX: SEQUENTIAL TUNNELING LIMIT

In this appendix we show how the sequential tunneling contribution I^{sq} can be extracted from the current (26) by taking the $\Gamma \to 0$ limit and keeping only large terms $\sim 1/\Gamma$. Using the scheme (27) the fraction $1/[(\xi_k - \xi_{p_1} - u_1)^2 + \Gamma^2]$ is replaced by $\pi \delta(\xi_k - \xi_{p_1} - u_1)/\Gamma$. The result can be presented as

$$I^{sq} = -e \frac{R^l_{N \to N+1} R^r_{N+1 \to N} - R^l_{N+1 \to N} R^r_{N \to N+1}}{2\Gamma/\hbar}, \quad (A1)$$

where the tunneling rates R are given by Fermi's golden rule. For tunneling between the left lead and the dot they are

$$R^l_{N \to N+1} = \frac{2\pi}{\hbar} \sum_{k,p} |t_{k,p}|^2 n_k (1 - n_p) \delta(\xi_k - \xi_p - u_1),$$

$$R^l_{N+1 \to N} = \frac{2\pi}{\hbar} \sum_{k,p} |t_{k,p}|^2 n_p (1 - n_k) \delta(\xi_k - \xi_p - u_1),$$

and R^r can be obtained from the expressions for R^l by replacing the indices $k \to q$.

On the other hand, in sequential tunneling theory[4] the current is presented as

$$I^{sq} = -e(W_N R^l_{N \to N+1} - W_{N+1} R^l_{N+1 \to N})$$

$$= -e(W_{N+1} R^r_{N+1 \to N} - W_N R^r_{N \to N+1}), \quad (A2)$$

where W_N represents the probability of the system being in a state with N electrons in the dot. These two expressions for the current I^{sq} arise from the condition that the current through the left barrier equal the current through the right barrier. Employing this equality the current (A2) can be rewritten as

$$I^{sq} = -eW_N \frac{R^l_{N \to N+1} R^r_{N+1 \to N} - R^l_{N+1 \to N} R^r_{N \to N+1}}{R^l_{N+1 \to N} + R^r_{N+1 \to N}}. \quad (A3)$$

This is the same relation as Eq. (A1) if we make the identification

$$\Gamma = \frac{\hbar}{2} \frac{R^l_{N+1 \to N} + R^r_{N+1 \to N}}{W_N}. \quad (A4)$$

This expression coincides with the one proposed in Ref. 13:

$$\Gamma = \frac{\hbar}{2}(R^l_{N+1 \to N} + R^r_{N+1 \to N} + R^l_{N \to N+1} + R^r_{N \to N+1}), \quad (A5)$$

if one takes into account the balance equation (A2) and the condition $W_N + W_{N+1} = 1$. Our derivation shows that replacing the denominator in Eq. (26) with a δ function reproduces the result of the sequential tunneling even if eV and/or ΔT are not small compared to the temperature T. In Sec. IV A we use this result in the linear response regime $eV, \Delta T \ll T$ to conclude that this approximation reproduces the results (8) for G^{sq} and (15) for G_T^{sq}. In this case the Γ is determined by the equilibrium rates and can be calculated by replacing the occupation numbers with Fermi functions,

$$\Gamma = \frac{\hbar}{2e^2}(G_l + G_r) u_1 \coth\left[\frac{u_1}{2T}\right], \quad (A6)$$

in agreement with Ref. 13.

*Present address: University of Regensburg, Institute for Theoretical physics, D-93040 Regensburg, Germany.

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

[2] See, e.g., L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M. Westervelt, and N. S. Wingreen, in *Mesoscopic Electron Transport*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schön, Vol. 345 of *NATO Advanced Study Institute, Ser.*

E (Kluwer, Dordrecht, 1997), pp. 105–214.

[3] L. I. Glazman and R. I. Shekhter, J. Phys.: Condens. Matter **1**, 5811 (1989).

[4] C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).

[5] A. A. M. Staring, L. W. Molenkamp, B. W. Alphenhaar, H. van Houten, O. J. A. Buyk, M. A. A. Mabesoone, C. W. J. Beenakker, and C. T. Foxon, Europhys. Lett. **22**, 57 (1993).

[6] A. S. Dzurak, C. G. Smith, M. Pepper, D. A. Ritchie, J. E. F.

Frost, G. A. C. Jones, and D. G. Hasko, Solid State Commun. **87**, 1145 (1993).

[7] A. S. Dzurak, C. G. Smith, C. H. W. Barnes, M. Pepper, L. Martin-Moreno, C. T. Liang, D. A. Ritchie, and G. A. C. Jones, Phys. Rev. B **55**, 10 197 (1997).

[8] S. Möller, H. Buhmann, S. F. Godijn, and L. W. Molenkamp, Phys. Rev. Lett. **81**, 5197 (1998).

[9] C. W. J. Beenakker and A. A. M. Staring, Phys. Rev. B **46**, 9667 (1992).

[10] A. V. Andreev and K. A. Matveev, Phys. Rev. Lett. **86**, 280 (2001); K. A. Matveev and A. V. Andreev, cond-mat/0201161 (unpublished).

[11] D. V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **65**, 2446 (1990).

[12] See, e.g., A. A. Abrikosov, *Fundamentals of the Theory of Metals* (North-Holland, Amsterdam, 1988).

[13] D. V. Averin, Physica B **194-196**, 979 (1994).

[14] J. Koenig, H. Schoeller, and G. Schön, Phys. Rev. Lett. **78**, 4482 (1997).