Nonadiabatic Charge Pumping in a Hybrid Single-Electron Transistor

Dmitri V. Averin[1] and Jukka P. Pekola[2]

[1]*Department of Physics and Astronomy, Stony Brook University, SUNY, Stony Brook, New York 11794-3800, USA*
[2]*Low Temperature Laboratory, Helsinki University of Technology, P.O. Box 3500, 02015 TKK, Finland*

(Received 10 February 2008; published 4 August 2008)

We study theoretically current quantization in the charge turnstile based on the superconductor–normal-metal single-electron transistor. The quantization accuracy is limited by either Andreev reflection or by Cooper-pair–electron cotunneling. The rates of these processes are calculated in the "above-the-threshold" regime when they compete directly with the lowest-order tunneling. By shaping the ac gate voltage drive it should be possible to achieve the metrological accuracy of 10^{-8}, while maintaining the quantized current on the level of 30 pA, just by one turnstile with realistic parameters using aluminum as a superconductor.

DOI: 10.1103/PhysRevLett.101.066801 PACS numbers: 73.23.Hk, 74.45.+c, 84.37.+q

Nanoscale tunneling structures provide the general basis for development of metrological sources of electrical current utilizing controlled transfer of individual charges [1]. Despite the beautiful achievements based on experiments with gated arrays of metallic tunnel junctions [2–5], and with semiconductor surface-acoustic-wave and charge-coupled devices [6–9], no fully satisfactory system in terms of both the accuracy and current magnitude has been realized yet. It was suggested recently [10] that an unexpectedly simple structure, a single-electron (SET) transistor with two hybrid normal-metal (N)–superconductor (S) tunnel junctions (SINIS or NISIN structure, I for insulator) holds promise as a quantized source of current. The first experiments with such a turnstile [10] demonstrated correct operation at the level of classical charge dynamics, but they were not yet conclusive as to its ultimate accuracy. In this Letter we analyze theoretically all the relevant higher-order quantum tunneling processes which limit this accuracy. The main conclusion is that these errors can be suppressed in a single ordinary aluminum-based device to the level mandated by the metrological requirements ($\leq 10^{-8}$), while keeping the absolute current relatively large (see Fig. 4 below).

The basic "classical" dynamics of the hybrid SET transistor (Fig. 1) as a charge turnstile can be described conveniently on the stability diagram shown in Fig. 2. Periodic variation of the gate-induced charge $n_g(t) \equiv C_g V_g(t)/e$ with frequency f (notations are defined by Fig. 1) indicated by the line with arrows in Fig. 2 drives the transistor periodically between the two nearest stability areas, e.g., $n = 0$ and $n = 1$, where n is the equilibrium number of extra electrons on the island. The turnstile operation requires that the lowest-order tunneling transitions are organized so that at finite bias voltage V and low temperature T they transfer precisely one electron per period $1/f$ through the transistor [10]. The properties of the tunneling thresholds (solid lines in Fig. 2) that make this possible in the hybrid transistor but not in the normal-metal one can be seen from Fig. 2. The thresholds in the hybrid are shifted with respect to the normal-metal system (dashed lines in Fig. 2) by the superconducting energy gap Δ, i.e., the shift along the n_g axis is $\delta = \Delta/2E_C$, where $E_C \equiv e^2/2C_\Sigma$ and $C_\Sigma = C_1 + C_2 + C_g$, expanding the stability areas. As a result, the neighboring stability areas overlap, and the gate voltage can drive the system between them keeping it all the time in the region of suppressed tunneling. Also, in this process, when the outgoing gate-voltage trajectory leaves the initial stability area, it crosses only one of the tunneling thresholds that define this area, allowing electron tunneling in only one direction. For instance, if the state $n = 0$ is brought by increase of n_g out of its expanded stability area into the $n = 1$ area (Fig. 2), an electron can tunnel into the transistor island only through the left junction. When the gate voltage decreases back to the $n = 0$ state, electron can tunnel out only through the right junction [10].

This turnstile operation is possible for any, e.g., harmonic, periodic time dependence $n_g(t)$ with the amplitude sufficiently large to move the system between the two stability areas (Fig. 2). The time that the system spends, however, in the overlap region of the two areas only increases the effect of the unwanted transitions. In order to maximize the turnstile output current, one needs then to minimize this time by shaping the waveform $n_g(t)$ as in Fig. 1(b). In this case, the system is switched abruptly between the regions where an electron tunnels in or out of the transistor, and the operation frequency f is limited only by the need to make the probability of missing these transitions $e^{-\gamma/2f}$ sufficiently small. At zero temperature,

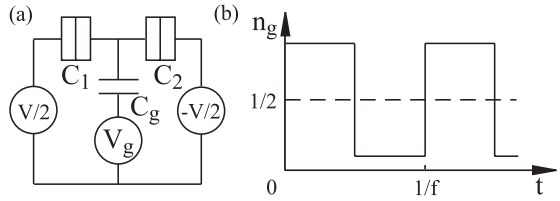

FIG. 1. (a) Hybrid (SINIS or NISIN) SET transistor and (b) time dependence of the ac gate-induced charge $n_g = C_g V_g/e$ oscillating with frequency f around $n_g = 1/2$.

the corresponding tunneling rate is $\gamma(U) = \gamma_0(U^2/\Delta^2 - 1)^{1/2}$, where U is electrostatic energy change due to tunneling, $\gamma_0 \equiv G\Delta/e^2$, and G is the junction tunnel conductance. The use of the rectangular [Fig. 1(b)] instead of the harmonic waveform suppresses the probability to miss a transition. Relative strength of this suppression depends on the amplitude of the n_g modulation, decreasing with increasing amplitude. Even for large amplitude, the exponent in the probability is reduced by the factor $2/\pi$ in the case of the harmonic waveform, leading roughly to a factor of 2 increase of the output current [see Eq. (5)] for the same error probability p, in favor of the rectangular waveform. This optimized waveform should be abrupt on the time scale of the period $1/f$ while being smooth on the scale h/Δ to avoid excitations of the higher-energy states that can lead to errors in the turnstile dynamics. This condition can be satisfied easily, since for a typical current of 100 pA, $f = I/e < 1$ GHz is well below $\Delta/h \simeq 50$ GHz.

In addition to missed cycles of tunneling due to finite frequency f, the basic correct tunneling sequence can be interrupted by thermal excitations due to finite temperature T, or quantum higher-order tunneling processes [11] which set the theoretical limit on the accuracy of the quantized current $I = ef$ produced by the turnstile. The rate of thermal errors depends on how far the gate-voltage trajectory is from the crossing points of the four relevant tunneling thresholds shown as solid lines in Fig. 2. The thresholds are given by the conditions $U_j^{\pm} = \Delta$ on electrostatic energy change $U_j^{\pm}$ due to forward (wanted) or backward (unwanted) electron tunneling in the jth junction: $U_j^{\pm} = \pm 2E_C[\upsilon_j - (-1)^j(n_g - 1/2)]$, $j = 1, 2$, where υ_j is a part of V that drops across the jth junction, $\upsilon_1 = (C_2 + C_g/2)V/e$ and $\upsilon_2 = (C_1 + C_g/2)V/e$. These equations show that at the thresholds of correct tunneling, the energy barriers for unwanted transitions through the opposite junction of the transistor are $\Delta - U_1^- = \Delta - U_2^- = eV$. Thus, with exponential accuracy, the thermal probability of electron tunneling in or out through the wrong junction leading to no net charge transfer in the cycle is e^{-eV/k_BT}. Another type of unwanted thermal transitions is the excitation of an extra electron through the transistor during the part of the period spent in the overlap region of the two stability areas. The electron is transferred by two successive excitations over the energy barriers $\Delta - U_j^+$, so that the thermal excitation exponent for the overall process is $e^{-(2\Delta - eV)/k_BT}$. Comparing the probabilities of the two types of thermal errors, we see that the thermal error rate is minimum for $eV \simeq \Delta$; in practice, the resulting classical error $e^{-\Delta/k_BT}$ is less than 10^{-8} at realistic temperatures $T \simeq 100$ mK.

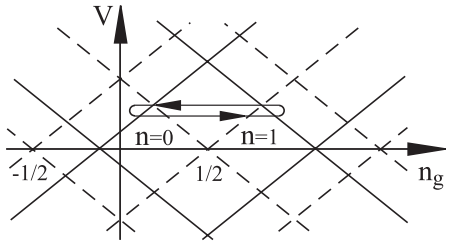

FIG. 2. Charge stability diagram of the hybrid SET transistor. Dashed lines are the tunneling thresholds of the rhombic stability regions $n = 0, 1$ in the normal-metal case. Solid lines show the thresholds in the hybrid transistor shifted by the superconducting gap Δ. Variation of the gate voltage (line with arrows) transfers one electron per period through the transistor.

We consider now quantum errors assuming ideal s-wave BCS superconductors. The rates of "elastic" higher-order processes which transfer electrons coherently, without creating excitations in the electrodes, are different in the NISIN and SINIS structures. In the NISIN transistor, the dominant elastic process is electron cotunneling, the rate of which is smaller than the rate γ of the lowest-order tunneling γ roughly by a factor $(\hbar G/e^2)(\delta E/\Delta)$ [12], where $\delta E/\Delta$ is the level spacing of the transistor island. For typical parameters this suppression factor is small, about 10^{-6}–10^{-7}, but does not quite reach the metrologically required level. In the SINIS transistor, in the relevant regime $eV \simeq \Delta$, the main contribution to elastic leakage is due to rectification of the ac Josephson current through the transistor. The resulting dc current is proportional to the square of the SINIS critical current and is much smaller than the inelastic leakage assisted by Andreev reflection (AR) considered below.

The rates of incoherent "inelastic" processes depend only on the local properties of the tunnel junctions and are the same in the NISIN and SINIS transistors. The intensity of these processes decreases rapidly with the number of involved electron transfers. The simplest process of electron inelastic cotunneling through the transistor is energetically forbidden in the relevant voltage range $eV < 2\Delta$. Two more complex transitions are relevant for the error analysis. One is Andreev reflection, i.e., tunneling of two electrons in a Cooper pair for which the superconducting gap does not provide an energy barrier. It transfers two electrons instead of one to or from the transistor island causing an error in turnstile operation. Another is the third-order process of Cooper-pair–electron (CPE) cotunneling. In it, correct one-electron tunneling in one of the junctions is combined coherently with the transfer of another electron through the whole transistor. To avoid creating superconducting excitations, the tunneling of two electrons in one of the transistor junctions in this process happens as AR. Electrostatic energy gains in the two processes are

$$\text{AR: } U_1^{++} = 4E_C(\upsilon_1 + n_g - 1), \quad U_2^{++} = 4E_C(\upsilon_2 - n_g), \qquad \text{CPE: } W_j^+ = U_j^+ + eV, \tag{1}$$

and the diagram of the corresponding thresholds, $U = 0$ for AR, and $W = \Delta$ for CPE, is shown in Fig. 3. If single-electron charging energy is small, $E_C < \Delta$ (i.e., $\delta > 1/2$ in Fig. 3), AR is allowed in the regions of the lowest-order tunneling needed for turnstile operation. For larger charg-

ing energy, $E_C > \Delta$, turnstile can be operated in the regime with suppressed AR ($\delta < 1/2$ in Fig. 3), with only the higher-order CPE processes causing errors. The CPE processes are allowed for any turnstile parameters and set the limit on the accuracy of current quantization.

We calculate the rates of the two higher-order tunneling processes assuming the simple quasi-1D ballistic geometry of the turnstile junctions, in which different transport modes in the electrodes are not mixed by tunneling. This assumption is reasonable in view of large conductivity of electrodes of practical SET transistors. Because of the nonadiabatic variation of the gate voltage [Fig. 1(b)], both higher-order tunneling processes take place in the "-above-the-threshold" regime, when they coexist with the lowest-order single-particle tunneling. We start with the rate γ_{AR} of the Andreev reflection. Above the single-particle threshold, the standard description of AR as the two-step transition perturbative in the electron tunneling amplitudes t (see, e.g., [13]) should be modified to account for the competing single-particle tunneling with rate $\gamma(U^+)$. Similarly to the theory of Coulomb-blockade threshold [14], this can be done by taking into account the lifetime broadening $i\gamma(U^+)/2$ of the initial state.

Because of the mutual coherency of Cooper pairs in different orbital states in the superconducting electrode, the amplitudes of the Cooper-pair tunneling from different states p within each transport mode into the two single-particle states with energies ϵ_k, ϵ_l in the normal electrode should be summed coherently. The total AR amplitude A is

$$A(\epsilon_k, \epsilon_l) = \sum_p u_p v_p t_{pk} t_{pl} \left(\frac{1}{\Omega_p + \epsilon_k - u} + \frac{1}{\Omega_p + \epsilon_l - u} \right),$$

where $u_p, v_p = [(1 \pm \epsilon_p/\Omega_p)/2]^{1/2}$ are the usual BCS quasiparticle factors, $\Omega_p = (\Delta^2 + \epsilon_p^2)^{1/2}$ is the quasiparticle energy, and $u = U^+ + i\gamma(U^+)/2$. Taking the sum over p under the standard approximation of constant density of states ρ and tunnel amplitudes t in the relevant energy range on the order of energy gap Δ, we get

$$A(\epsilon_k, \epsilon_l) = \rho t^2 \Delta [a(u - \epsilon_k) + a(u - \epsilon_l)],$$
$$a(\epsilon) = (\epsilon^2 - \Delta^2)^{-1/2} \ln \left[\frac{\Delta - \epsilon + (\epsilon^2 - \Delta^2)^{1/2}}{\Delta - \epsilon - (\epsilon^2 - \Delta^2)^{1/2}} \right]. \quad (2)$$

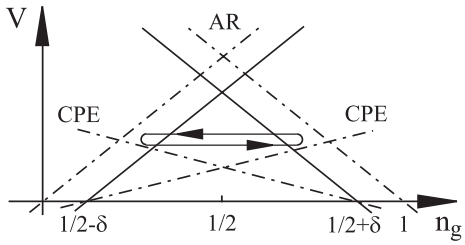

FIG. 3. Tunneling thresholds for Andreev reflection (AR) and Cooper-pair–electron (CPE) cotunneling in a hybrid SET transistor. Solid lines are the thresholds of the lowest-order tunneling. All tunneling processes are allowed above the corresponding threshold.

The main qualitative feature of the amplitude A is the resonance at the gap edge, $\epsilon \simeq \Delta$, where the rate $|A|^2$ diverges as $1/|\epsilon - \Delta|$. Level broadening, in our case due to the single-particle tunneling with rate γ, broadens the resonance and suppresses the divergence.

The amplitude A gives the AR rate at $k_B T \ll \Delta$:

$$\gamma_{\mathrm{AR}} = \frac{2\pi}{\hbar} \sum_{k,l} |A|^2 [1 - f(\epsilon_k)][1 - f(\epsilon_l)] \delta(\epsilon_k + \epsilon_l - U^{++}),$$

where in the adopted quasi-1D model the states k, l in the sum should belong to the same transport mode. The result of summation over these modes can be expressed in terms of the normal-state conductance G within the natural junction model in which transparency t^2 varies exponentially with energy on the scale $\epsilon_0 \gg \Delta$. The effective number $\mathcal{N}$ of the transport modes in the junction is determined then by the decrease of transparency with increasing transverse energy of the mode: $\mathcal{N} = Sm\epsilon_0/\pi\hbar^2$, where S is the junction area and m is electron mass. The sum over modes and integration over the total energy can then be done separately giving

$$\gamma_{\mathrm{AR}} = \frac{\gamma_0 g \Delta}{16\pi\mathcal{N}} \int d\epsilon f(\epsilon - U^{++}/2) f(-\epsilon - U^{++}/2) \times \left| \sum_{\pm} a(\pm\epsilon + E_C - i\gamma/2) \right|^2, \quad g \equiv \hbar G/e^2. \quad (3)$$

If AR transitions are not energetically allowed, the leakage current is determined by the third-order CPE cotunneling which combines AR with one more electron transfer in the opposite junction. The part of the CPE amplitude $\mathcal{A}$ that corresponds to the two-electron AR transfer process is calculated as above for direct AR. Combining terms with different ordering of the three involved electron transfers we get the total CPE amplitude

$$\mathcal{A} = \left(\frac{1}{2E_C + 2U^+ - \epsilon_k - \epsilon_l} + \frac{1}{\epsilon_k + \epsilon_l - U^+ - u} \right) \times [a(\epsilon_k - U^+) + a(\epsilon_l - U^+)] + [a(u - \epsilon_k) + a(u - \epsilon_l)] \times \left(\frac{1}{U^+ + u - \epsilon_k - \epsilon_l} + \frac{1}{2E_C - U^+ - u + \epsilon_k + \epsilon_l} \right). \quad (4)$$

Summing all transitions with this amplitude as above, we obtain the total rate of the CPE cotunneling:

$$\gamma_{\mathrm{CPE}} = \frac{\gamma_0 g^2 \Delta}{32\pi^2 \mathcal{N}} \int_\Delta^\infty d\Omega \frac{\Omega}{\sqrt{\Omega^2 - \Delta^2}} \int d\epsilon_k \int d\epsilon_l |\mathcal{A}|^2 [1 - f(\epsilon_k)][1 - f(\epsilon_l)] f(\Omega + \epsilon_k + \epsilon_l - U^+ - eV).$$

Figure 4(a) shows the gate dependence of the zero-temperature normalized rates of the (wanted) single-particle tunneling, $\tilde{\gamma} \equiv \gamma/\gamma_0$, of the AR transitions, $\tilde{\gamma}_{\mathrm{AR}} \equiv \gamma_{\mathrm{AR}}/(\gamma_0 g/16\pi\mathcal{N})$, and CPE cotunneling, $\tilde{\gamma}_{\mathrm{CPE}} \equiv \gamma_{\mathrm{CPE}}/(\gamma_0 g^2/32\pi^2\mathcal{N})$ at the optimum bias point $eV = \Delta$ for a few values of E_C/Δ. As in Fig. 3, the thresholds of

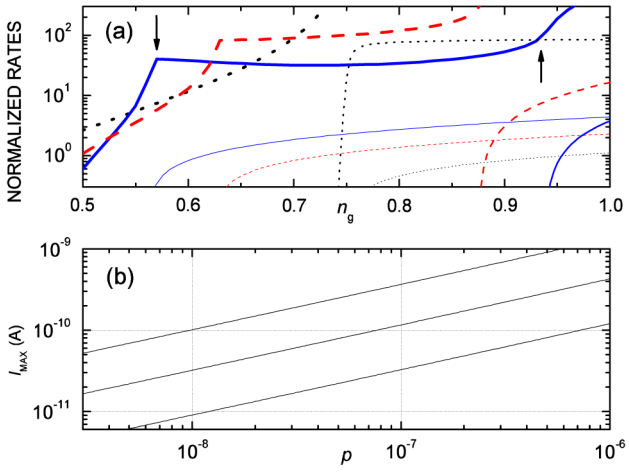

FIG. 4 (color online). Performance of the SINIS hybrid turnstile. (a) Normalized rates (for normalization, see text) of the various processes as functions of gate position at the optimum bias point $eV = \Delta$: CPE cotunneling (curves in the top part), single-particle (lower right corner), and Andreev tunneling (taller curves on the right). The different sets of curves refer to $E_C/\Delta = 1$ (black dash-dotted lines), 2 (red dashed lines), and 4 (blue solid lines). The thresholds at $E_C/\Delta = 4$ for single-particle and AR are indicated by vertical arrows. (b) The maximum pumped current (5) as a function of the allowed error rate p for $E_C/\Delta = 10$, 4, and 2 from top to bottom.

single-particle and Andreev processes coincide for $E_C/\Delta = 1$, but for larger values of this ratio there is a window between the two onsets. Kinks in CPE rate occur at these thresholds; in between, $\gamma_{\rm CPE}$ changes only little. One can see from this plot that an optimum gate value—fast single-particle transfer and errors only by CPE—exists for the case $E_C > \Delta$, and it lies within $1/2 + (2\Delta - eV)/4E_C < n_g < 1 - eV/4E_C$. The turnstile should thus be operated by a gate voltage [Fig. 1(b)] switching between such an n_g and $1 - n_g$.

We note that since the CPE contributes one extra transferred electron, the relative transfer error is $p = 2\gamma_{\rm CPE}/\gamma$ in the operation window discussed above. This gives the (maximum) junction conductance which can still suppress the CPE error to below p as $g = 4\pi[\mathcal{N} p\tilde{\gamma}/\tilde{\gamma}_{\rm CPE}]^{1/2}$. On the other hand, one can drive the turnstile at a frequency $f = \gamma/(2\ln(1/p))$ to suppress the missing cycle errors to the same level. The maximum current of the turnstile at the error rate p is then

$$I_{\rm max} = ef = \frac{e\Delta}{\hbar}\frac{2\pi}{\ln(1/p)}[\mathcal{N} p\tilde{\gamma}^3/\tilde{\gamma}_{\rm CPE}]^{1/2}. \qquad (5)$$

Figure 4(b) shows $I_{\rm max}$ versus p with aluminum as the superconductor, for which $\Delta/k_B \simeq 2.5$ K. In this plot, we also take into account that $\mathcal{N} \propto E_C^{-1}$ because of the junction area dependence of both of these quantities, and use an estimate of the tunnel barrier characteristics $\mathcal{N} = 10^4$ for $E_C = 1$ K. We can see from Fig. 4 that increasing E_C/Δ indeed improves the turnstile performance, and a single turnstile with $E_C/\Delta = 4$ reaches an accuracy of 10^{-8} at about 30 pA current with $\simeq 400$ kΩ junction resistance. With $E_C/\Delta = 10$ (such high E_C's were obtained, e.g., in [15]), 100 pA current can be reached with the same accuracy.

We have shown that a simple hybrid SINIS turnstile should qualify as a metrological source of current. To reach sufficient level of current, either a very large charging energy or a few parallel turnstiles are needed. The latter option is affordable because of the simplicity of the basic device [10]. In practical pumps [16], other sources of fluctuations (e.g., variations of the background charge) that cannot be precisely predicted by theory influence the performance as well.

This work was supported in part by NSF Grant No. DMR-0325551, by Technology Industries of Finland Centennial Foundation, and by the Academy of Finland. We thank M. Möttönen for discussions.

[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 (North-Holland, Amsterdam, 1991), p. 173.
[2] L. J. Geerligs, V. F. Anderegg, P. A. M. Holweg, J. E. Mooij, H. Pothier, D. Esteve, C. Urbina, and M. H. Devoret, Phys. Rev. Lett. **64**, 2691 (1990).
[3] H. Pothier, P. Lafarge, C. Urbina, D. Esteve, and M. H. Devoret, Europhys. Lett. **17**, 249 (1992).
[4] M. W. Keller, A. L. Eichenberger, J. M. Martinis, and N. M. Zimmerman, Science **285**, 1706 (1999).
[5] S. V. Lotkhov, S. A. Bogoslovsky, A. B. Zorin, and J. Niemeyer, Appl. Phys. Lett. **78**, 946 (2001).
[6] J. M. Shilton *et al.*, J. Phys. Condens. Matter **8**, L531 (1996).
[7] A. Fujiwara, N. M. Zimmerman, Y. Ono, and Y. Takahashi, Appl. Phys. Lett. **84**, 1323 (2004).
[8] M. D. Blumenthal *et al.*, Nature Phys. **3**, 343 (2007).
[9] B. Kaestner *et al.*, Phys. Rev. B **77**, 153301 (2008).
[10] J. P. Pekola, J. J. Vartiainen, M. Möttönen, O.-P. Saira, M. Meschke, and D. V. Averin, Nature Phys. **4**, 120 (2008).
[11] D. V. Averin, A. A. Odintsov, and S. V. Vyshenskii, J. Appl. Phys. **73**, 1297 (1993).
[12] D. V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **69**, 1993 (1992).
[13] J. W. Wilkins, in *Tunneling Phenomena in Solids*, edited by E. Burnstein and S. Lundquist (Plenum, New York, 1969), p. 333.
[14] Yu. V. Nazarov, J. Low Temp. Phys. **90**, 77 (1993); D. V. Averin, Physica (Amsterdam) **194/196B**, 979 (1994); H. Schoeller and G. Schön, Phys. Rev. B **50**, 18 436 (1994).
[15] Yu. A. Pashkin, Y. Nakamura, and J. S. Tsai, Appl. Phys. Lett. **76**, 2256 (2000).
[16] R. L. Kautz, M. W. Keller, and J. M. Martinis, Phys. Rev. B **62**, 15 888 (2000); X. Jehl, M. W. Keller, R. L. Kautz, J. Aumentado, and J. M. Martinis, *ibid.* **67**, 165331 (2003).