Nonlinear Conductance in Resonant Tunneling

C. A. Stafford
Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland
(Received 5 April 1996)

A multiterminal conductance formula describing resonant tunneling through an interacting mesoscopic system is derived and used to investigate the nonlinear conductance of a quantum dot. An explicit gauge-invariant expression for the *I-V* characteristic which depends sensitively on the full capacitance matrix is obtained. A voltage probe is found to have a dramatic effect on the nonlinear conductance. [S0031-9007(96)01247-1]

PACS numbers: 72.10.Bg, 73.20.Dx, 73.40.Gk

The problem of coherent transport through an interacting mesoscopic system coupled to macroscopic electron reservoirs is of considerable current interest [1–10]. Recently, substantial progress has been made in treating transport through correlated systems using nonequilibrium Green's functions [5–7] and bosonization [8,9]. However, Büttiker has emphasized [10] that a gauge-invariant description of nonlinear transport requires a proper treatment of the long-range Coulomb interaction which explicitly includes the external gates and reservoirs. In this Letter, we derive a Breit-Wigner type formula for the resonant conductance through an arbitrary interacting system with a nondegenerate ground state, coupled weakly via tunnel barriers to several electron reservoirs, and use this formula to investigate nonlinear transport through a quantum dot, including inelastic processes explicitly. It is shown that the inclusion of long-range Coulomb interactions via a capacitive charging model [11] leads to a gauge-invariant *I-V* characteristic which depends explicitly on the full capacitance matrix.

An arbitrary interacting mesoscopic conductor coupled to M macroscopic electron reservoirs is described by the Hamiltonian

$$H = H_{\mathrm{int}}(\{d_n^\dagger, d_n\}) + \sum_{\alpha=1}^{M} \sum_{k\in\alpha} \epsilon_k c_k^\dagger c_k + \sum_{\alpha=1}^{M} \sum_{k\in\alpha} \sum_n (V_{kn} c_k^\dagger d_n + \mathrm{H.c.}) , \qquad (1)$$

where $\{d_n^\dagger\}$ creates a complete set of single-particle states in the mesoscopic system, $c_{k\in\alpha}^\dagger$ creates an electron in state k of reservoir α, and H_{int} is a polynomial in $\{d_n^\dagger, d_n\}$ which commutes with the electron number $N = \sum_n d_n^\dagger d_n$. We denote the ground state of H_{int} for each N by $|0_N\rangle$ and the ground state energy by E_N^0. We assume E_N^0 to be nondegenerate, as is generically the case in a nonzero magnetic field.

The expectation value of the current flowing into reservoir α can be expressed using the formalism of Meir and Wingreen as [5]

$$I_\alpha = -\frac{e}{h} \int d\epsilon \, \mathrm{Im}\,\mathrm{Tr}\{\Gamma^\alpha(\epsilon)[G^<(\epsilon) + 2f_\alpha(\epsilon)G^r(\epsilon)]\} , \qquad (2)$$

where $\Gamma_{nm}^\alpha(\epsilon) = 2\pi \sum_{k\in\alpha} V_{kn} V_{km}^* \delta(\epsilon - \epsilon_k)$ is a matrix characterizing the tunnel barrier connecting reservoir α to the system, $G_{nm}^{<,r}(\epsilon)$ are Fourier transforms of the Green's functions $G_{nm}^<(t) = i\langle d_m^\dagger(0) d_n(t)\rangle$ and $G_{nm}^r(t) = -i\theta(t)\langle\{d_n(t), d_m^\dagger(0)\}\rangle$, which describe propagation within the mesoscopic system in the presence of coupling to the leads, and $f_\alpha(\epsilon) = \{\exp[(\epsilon - \mu_\alpha)/k_B T] + 1\}^{-1}$ is the Fermi function for reservoir α. If the tunneling barriers to the reservoirs are sufficiently large, and if the temperature and bias are small compared to the energy of an excitation, then the current through the system will be determined by transitions $|0_{N-1}\rangle \to |0_N\rangle$ between nondegenerate ground states. In the vicinity of such a resonance, the Green's functions can be shown to have the form

$$G_{nm}^r(\epsilon) = \frac{\langle 0_{N-1}|d_n|0_N\rangle\langle 0_N|d_m^\dagger|0_{N-1}\rangle}{\epsilon - E_N^0 + E_{N-1}^0 + i\Gamma_N/2} + \text{additional poles}, \qquad (3)$$

$$G_{nm}^<(\epsilon) = \frac{i\langle 0_{N-1}|d_n|0_N\rangle\langle 0_N|d_m^\dagger|0_{N-1}\rangle \sum_\alpha \Gamma_N^\alpha f_\alpha(\epsilon)}{(\epsilon - E_N^0 + E_{N-1}^0)^2 + (\Gamma_N/2)^2} + \text{additional poles}, \qquad (4)$$

where $\Gamma_N = \sum_{\alpha=1}^M \Gamma_N^\alpha$, and

$$\Gamma_N^\alpha = 2\pi \sum_{k\in\alpha} \sum_{n,m} \langle 0_{N-1}|V_{kn} d_n|0_N\rangle \times \langle 0_N|V_{km}^* d_m^\dagger|0_{N-1}\rangle \delta(\epsilon_k - E_N^0 + E_{N-1}^0) . \qquad (5)$$

Inserting $G_{nm}^{<,r}$ into Eq. (2), one finds the multiprobe current formula

$$I_\alpha = \frac{e}{h} \sum_{\beta=1}^{M} \int d\epsilon \sum_N \frac{\Gamma_N^\alpha \Gamma_N^\beta [f_\alpha(\epsilon) - f_\beta(\epsilon)]}{(\epsilon - E_N^0 + E_{N-1}^0)^2 + (\Gamma_N/2)^2} . \qquad (6)$$

The low-temperature transport through such a correlated many-body system weakly coupled to multiple leads thus exhibits resonances of the Breit-Wigner type [12], where the positions and intrinsic widths of the resonances are determined by the *many-body* states of the system. In particular, Eq. (5) implies that the partial widths of the conductance resonances are strongly suppressed for large N due to the orthogonality catastrophe [13]. Equation (6), which expresses the current in terms of transmission probabilities, is a generalization of the multiterminal conductance formula for a noninteracting system derived by Büttiker [14] to the case of resonant tunneling through an interacting system.

In deriving Eq. (6), we have neglected the additional poles in $G_{nm}^{<,r}(\epsilon)$, which is justified provided k_BT, $\Gamma_N \ll \Delta E_N$ and $\Delta\mu < \Delta E_N$, where $\Delta E_N = \min(E_N^1 - E_N^0, E_{N-1}^1 - E_{N-1}^0, E_{N+1}^0 - E_N^0 - \mu_\alpha, \mu_\alpha - E_{N-1}^0 + E_{N-2}^0)$, E_N^1 being the energy of the lowest lying excited state of the N-electron system [15]. Equation (6) is thus appropriate to describe resonant tunneling through semiconductor nanostructures [1,2] or ultrasmall metallic/superconducting systems [3,16] under conditions of low temperature and bias, where transport is dominated by a *single* ground state to ground state transition $|0_{N-1}\rangle \rightarrow |0_N\rangle$.

We next specialize to the three-terminal configuration shown in Fig. 1(a). Transport occurs between the left (L) and right (R) reservoirs. The auxiliary reservoir (A) serves as a voltage probe [17]. An ideal voltage probe should have an infinite impedance, so we demand that the expectation value of the current flowing into reservoir A be zero, which fixes μ_A via Eq. (6). Eliminating $f_A(\epsilon)$ from Eq. (6), one finds the current flowing from left to right

$$I = -\frac{e}{h}\sum_N \frac{\Gamma_N^L\Gamma_N^R}{\Gamma_N^L + \Gamma_N^R} \times \int \frac{\Gamma_N[f_L(\epsilon) - f_R(\epsilon)]\,d\epsilon}{(\epsilon - E_N^0 + E_{N-1}^0)^2 + (\Gamma_N/2)^2}\,. \quad (7)$$

The total width of the Nth resonance is $\Gamma_N = \Gamma_N^L + \Gamma_N^R + \Gamma_N^A$, where the quantity $\Gamma_N^A/\hbar$ may be interpreted as the *inelastic scattering rate* due to phase-breaking processes in the auxiliary reservoir. Such processes arise when an electron in the dot escapes into reservoir A and is replaced by an electron from the reservoir, whose phase is uncorrelated with that of the previous electron [12]. It should be emphasized that both the position of the resonance $E_N^0 - E_{N-1}^0$ and, in principle, the resonance width Γ_N depend on the external electrochemical potentials and gate voltages due to Coulomb interactions in the system. This dependence must be determined from the explicit form of H_{int} in order to calculate the I-V characteristic.

Let us consider the simple example of a quantum dot defined electrostatically in a 2D electron gas by several metallic gates, as indicated in Fig. 1(a), treating electron-electron interactions using a capacitive charging model [11], as indicated in Fig. 1(b). The Hamiltonian for the interacting region, including the work done by the external voltage sources, is

$$H_{\text{int}} = \sum_n \varepsilon_n d_n^\dagger d_n + Q^2/2C_\Sigma + Q\sum_i C_iV_i/C_\Sigma\,, \quad (8)$$

where ε_n are the single-particle energies of the (quasi)-bound states in the confining potential of the quantum dot, $Q = -e\sum_n d_n^\dagger d_n$ is the charge operator for the quantum dot, $C_\Sigma = \sum_i C_i$, and the voltages in the reservoirs are defined by $\mu_\alpha = \epsilon_F - eV_\alpha$, where ϵ_F is the equilibrium electrochemical potential. The spin degeneracy of the system is assumed to be broken, e.g., by an external magnetic field [18]. From Eq. (8), one finds the resonance positions

$$E_N^0 - E_{N-1}^0 = \varepsilon_N + e^2(N - 1/2)/C_\Sigma - e\sum_i C_iV_i/C_\Sigma \quad (9)$$

and widths

$$\Gamma_N^\alpha = 2\pi\sum_{k\in\alpha}|V_{kN}|^2\delta(\epsilon_k - E_N^0 + E_{N-1}^0)\,. \quad (10)$$

For this simple charging model, which neglects intradot correlations, the many-body corrections to the resonance widths implicit in Eq. (5) are absent. The matrix elements V_{kN} depend on the external voltages which define the point contacts 1 and 3. However, their variation in the vicinity of a single resonance may be neglected, so Γ_N^α will be taken to be independent of V_i. Inserting these expressions into Eq. (7), and performing the energy integral at

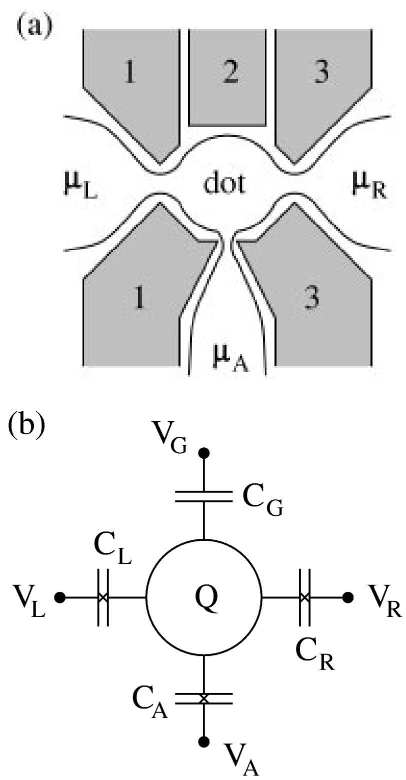

FIG. 1. (a) Schematic diagram of a quantum dot in a three-terminal configuration. (b) Equivalent circuit with $C_G = C_1 + C_2 + C_3$ and $C_GV_G = C_1V_1 + C_2V_2 + C_3V_3$.

$T = 0$, one finds the I-V characteristic for the Nth resonance

$$I/I_{\max} = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{\Delta\epsilon + e\sum_i C_i V_{iR}/C_\Sigma}{\Gamma_N/2}\right) - \tan^{-1}\left(\frac{\Delta\epsilon + e\sum_i C_i V_{iL}/C_\Sigma}{\Gamma_N/2}\right)\right], \tag{11}$$

where $I_{\max} = (e/\hbar)\Gamma_N^L\Gamma_N^R/(\Gamma_N^L + \Gamma_N^R)$, $\Delta\epsilon = \epsilon_F - [\varepsilon_N + e^2(N - 1/2)/C_\Sigma]$, and $V_{ij} \equiv V_i - V_j$. Similarly, the mean charge on the quantum dot at $T = 0$ in the vicinity of the Nth resonance is found to be [19]

$$\langle Q\rangle = -e\left[N - \frac{1}{2} + \frac{\Gamma_N^R/\pi}{\Gamma_N^L + \Gamma_N^R}\tan^{-1}\left(\frac{\Delta\epsilon + e\sum_i C_i V_{iR}/C_\Sigma}{\Gamma_N/2}\right) + \frac{\Gamma_N^L/\pi}{\Gamma_N^L + \Gamma_N^R}\tan^{-1}\left(\frac{\Delta\epsilon + e\sum_i C_i V_{iL}/C_\Sigma}{\Gamma_N/2}\right)\right]. \tag{12}$$

The voltage in probe A is determined by the constraint $I_A = 0$, and satisfies

$$\sum_i C_i(V_A - V_i)/C_\Sigma = (\Gamma_N/2e)\tan[\pi(\langle Q\rangle/e + N - 1/2)] + \Delta\epsilon/e\,. \tag{13}$$

For $C_A \neq 0$, both I and $\langle Q\rangle$ depend on V_A, and the nonlinear conductance must be determined from a solution of the coupled Eqs. (11)–(13). Equations (11) and (12) are functions of voltage differences only, and are thus gauge invariant, i.e., invariant under a global shift of the external electric potentials. Evidently, the multiprobe current given by Eq. (6) is quite generally gauge invariant provided the long-range Coulomb interaction is included in H_{int}, so that E_N^0 transforms as $E_N^0 \to E_N^0 - eN\delta V$ under a global shift $V_i \to V_i + \delta V$.

From Eq. (11), one finds the nonlinear conductance for bias $V = V_L - V_R$,

$$\frac{\partial I}{\partial V} = \frac{e^2}{h}\frac{\Gamma_N^L\Gamma_N^R}{\Gamma_N^L + \Gamma_N^R}\sum_{\sigma=\pm1}\frac{\Gamma_N[(1 + \sigma\eta)/2 + \sigma(C_A/C_\Sigma)\partial V_{A0}/\partial V]}{[\Delta\epsilon + (\sigma + \eta)eV/2 + e(C_A V_{A0} + Q_G)/C_\Sigma]^2 + (\Gamma_N/2)^2}\,, \tag{14}$$

where $\eta = (C_L - C_R)/C_\Sigma$, $Q_G = C_G(V_G - V_0)$, and $V_{A0} = V_A - V_0$, with $V_0 \equiv (V_L + V_R)/2$ (C_G and V_G are defined in the caption of Fig. 1). As a function of the polarization charge Q_G induced by the external gates, the linear response conductance exhibits Lorentzian peaks with spacing $\Delta Q_G = e + C_\Sigma\Delta\varepsilon_N/e$, where $\Delta\varepsilon_N$ is the single-particle level spacing in the dot, leading to periodic Coulomb blockade oscillations in the small capacitance limit, as discussed in Ref. [11]. While the linear response conductance depends only on Q_G, Γ^α, and the total capacitance C_Σ, the nonlinear conductance (14) depends explicitly on the asymmetry η in the capacitances of the tunnel barriers. Such an asymmetry leads to a shift in the resonance for large bias, as shown in Fig. 2. More significantly, the I-V characteristic is in general *asymmetric* for nonzero η, leading to rectification. The differential conductance typically exhibits a doublet structure (see Fig. 3), with peaks when the Fermi levels in the right and left reservoirs are aligned with the resonance; for $C_A = 0$, the peaks occur at $V = \pm 2(\Delta\epsilon/e + Q_G/C_\Sigma)/(1 \pm \eta)$ and the ratio of the conductance at the two peaks is $(1 - \eta)/(1 + \eta)$. Thus, to treat the nonlinear conductance, it is not sufficient to characterize the electron-electron interactions simply by the charging energy $U = e^2/C_\Sigma$ and the polarization charge Q_G; the full capacitance matrix must be used.

Equation (12) predicts that the jump [20] in the mean charge on the quantum dot as a function of Q_G is split into two jumps at large bias, of width $C_\Sigma\Gamma/e$ and heights $e\Gamma^L/2(\Gamma^L + \Gamma^R)$ and $e\Gamma^R/2(\Gamma^L + \Gamma^R)$, separated by $\Delta Q_G = (C_\Sigma - C_A)V$, which arise when the resonant energy crosses the Fermi levels in the left and right reservoirs. This behavior is illustrated by the dashed curve in Fig. 2.

Let us next consider the effect of capacitive coupling to the voltage probe. Because the voltage in the probe adjusts so as to prevent charge accumulation in reservoir A whatever the dc bias, a capacitive coupling between the probe and the dot tends to suppress charge accumulation

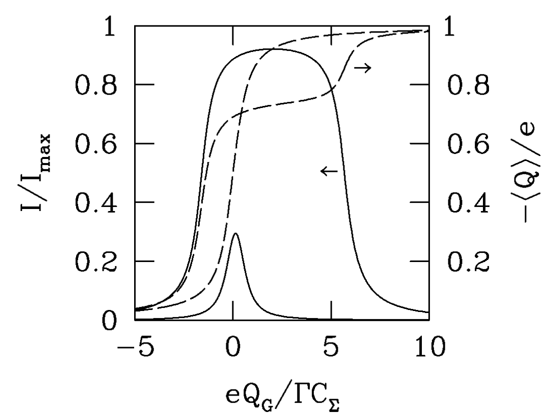

FIG. 2. Resonant current in units of $I_{\max} = (e/\hbar)\Gamma^L\Gamma^R/(\Gamma^L + \Gamma^R)$ (solid curves) and mean charge $\langle Q\rangle$ (dashed curves) for a quantum dot with $\eta = -1/2$, $\Gamma^R/\Gamma^L = 2$, $\Delta\epsilon = 0$, and $C_A/C_\Sigma = 0.1$ as a function of the polarization charge $Q_G = C_G[V_G - (V_L + V_R)/2]$; the bias $V = V_L - V_R = \Gamma/2e$ and $8\Gamma/e$ (curves marked with arrows). For large bias, the jump in the mean charge on the quantum dot as a function of Q_G is split into two jumps, which arise when the resonant energy crosses the Fermi levels in the left and right reservoirs.

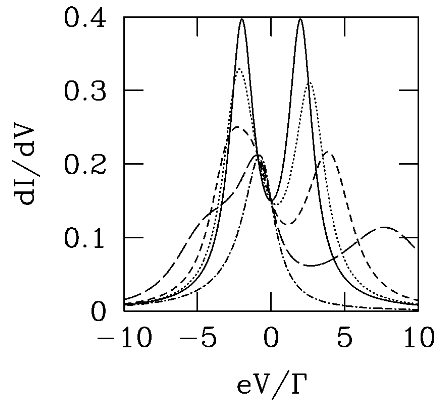

FIG. 3. Differential conductance $\partial I/\partial V$ in units of $(1 - \Gamma^A/\Gamma)e^2/h$ for a quantum dot with $\eta = 0$, $\Gamma^L/\Gamma^R = 3$, and $\Delta\epsilon/\Gamma = 1$ ($Q_G = 0$) for $C_A/C_\Sigma = 0$ (solid curve); 1/4 (dotted curve); 1/2 (short dashed curve); 3/4 (long dashed curve); 1 (dash-dotted curve). Note that the linear response conductance is independent of C_A.

in the dot, leading to a mean dot charge $\langle Q\rangle = -e[N - 1/2 + (1/\pi)\tan^{-1}(2\Delta\epsilon/\Gamma_N)]$ which is *independent* of the external voltages in the limit $C_A \to C_\Sigma$. Such a suppression of charge accumulation—over and above the Coulomb blockade—leads to a dramatic change in the nonlinear conductance, as shown in Fig. 3. In addition, for $C_A \neq 0$, the I-V characteristic is generally asymmetric even if $\eta = 0$, provided $\Gamma^L \neq \Gamma^R$ (see Fig. 3). Thus a seemingly noninvasive voltage measurement involving negligible dephasing $\Gamma^A/\hbar$ can significantly modify the nonlinear I-V characteristic due to electrostatic effects.

The theoretical results presented above are consistent with several recent experiments on coherent resonant tunneling through quantum dots [1,2]. The thermally broadened Lorentzian conductance peaks predicted by Eq. (7) and the shift of the peak positions at finite bias for $\eta \neq 0$ are consistent with the experimental results of Foxman *et al.* [1]. Furthermore, the fact that the tunneling current can be expressed in terms of the transmission probabilities through the many-body system [Eq. (6)], just as in the noninteracting case [14], is consistent with the intriguing observation of coherent modulation of the Coulomb blockade peak heights as a function of gate voltage and magnetic field [2], which suggested the applicability of a Landauer-type formula to resonant tunneling through an interacting system.

In conclusion, a multiterminal conductance formula describing resonant tunneling through an interacting mesoscopic system was derived and used to investigate the nonlinear conductance of a quantum dot. It was shown that the inclusion of the Coulomb interaction between the quantum dot and the external reservoirs and gates leads to a gauge-invariant I-V characteristic which depends sensitively on the full capacitance matrix.

The author thanks T. Christen and M. Büttiker for useful discussions, and for making available their results [21] prior to publication. This work was supported by the Swiss National Science Foundation.

[1] E. B. Foxman *et al.*, Phys. Rev. B **47**, 10 020 (1993).
[2] A. T. Johnson *et al.*, Phys. Rev. Lett. **69**, 1592 (1992); T. Heinzel *et al.*, Europhys. Lett. **26**, 689 (1994).
[3] D. C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **74**, 3241 (1995); C. T. Black, D. C. Ralph, and M. Tinkham, *ibid.* **76**, 688 (1996).
[4] A. Yacoby, M. Heiblum, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **74**, 4047 (1995).
[5] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).
[6] S. Hershfield, J. H. Davies, and J. W. Wilkins, Phys. Rev. B **46**, 7046 (1992); A. Levy Yeyati, A. Martín-Rodero, and F. Flores, Phys. Rev. Lett. **71**, 2991 (1993); N. S. Wingreen and Y. Meir, Phys. Rev. B **49**, 11 040 (1994), and references therein.
[7] C. Bruder, R. Fazio, and H. Schoeller, Phys. Rev. Lett. **76**, 114 (1996).
[8] A. Schiller and S. Hershfield, Phys. Rev. B **51**, 12 896 (1995).
[9] D. L. Maslov, Phys. Rev. B **52**, 14 368 (1995); I. Safi and H. J. Schulz, *ibid.* **52**, 17 040 (1995); R. Egger and H. Grabert, Phys. Rev. Lett. **77**, 538 (1996).
[10] M. Büttiker, J. Phys. Condens. Matter **5**, 9361 (1993).
[11] D. Averin, A. N. Korotkov, and K. K. Likharev, Phys. Rev. B **44**, 6199 (1991); H. van Houten, C. W. J. Beenakker, and A. A. M. Staring, in *Single Charge Tunneling,* edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992); Y. Meir, N. S. Wingreen, and P. A. Lee, Phys. Rev. Lett. **66**, 3048 (1991).
[12] M. Büttiker, IBM J. Res. Dev. **32**, 63 (1988); A. D. Stone and P. A. Lee, Phys. Rev. Lett. **54**, 1196 (1985).
[13] J. M. Kinaret *et al.*, Phys. Rev. B **46**, 4681 (1992); K. A. Matveev, L. I. Glazman, and H. U. Baranger, *ibid.* **53**, 1034 (1996).
[14] M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
[15] Equation (6) is not applicable to systems with a (spin) degenerate ground state [$\Delta E(N) = 0$], where the low-temperature physics is that of the Kondo effect, as discussed in Refs. [5–8].
[16] Equation (6) is applicable to tunneling through superconducting systems in which the charging energy exceeds the superconducting gap, so that Andreev reflection is energetically prohibited.
[17] H.-L. Engquist and P. W. Anderson, Phys. Rev. B **24**, 1151 (1981).
[18] The general considerations presented here on the effect of the entire electrical environment on the nonlinear conductance should also be applicable to models with spin degeneracy, such as those considered in Refs. [5–8].
[19] The mean charge on the quantum dot is determined by $\langle Q\rangle = -e\int_{-\infty}^{\infty} d\epsilon\, \mathrm{Im}\,\mathrm{Tr}\{G^<(\epsilon)\}/2\pi$.
[20] H. Grabert, Phys. Rev. B **50**, 17 364 (1994).
[21] T. Christen and M. Büttiker (unpublished) develop a gauge-invariant theory of nonlinear transport based on the scattering approach, where electron-electron interactions are treated within the Hartree approximation; however, the single-electron charging effects discussed here are outside the scope of that approach.