

## Sources of negative differential resistance in electric nanotransport

Ioan Bâldea\* and Horst Köppel

*Theoretische Chemie, Universität Heidelberg, Im Neuenheimer Feld 229, D-69120 Heidelberg, Germany*

(Received 26 February 2010; published 3 May 2010)

A negative differential resistance (NDR) in nanotransport is often ascribed to electron correlations. We present a simple example revealing that finite electrode bandwidths and energy-dependent electrode density of states can cause a significant NDR, which may occur even in uncorrelated systems. So, special care is needed in assessing the role of electron correlations in the NDR.

DOI: [10.1103/PhysRevB.81.193401](https://doi.org/10.1103/PhysRevB.81.193401)

PACS number(s): 73.63.-b, 85.35.Be, 85.35.Gv, 85.65.+h

The fact that the current-voltage ( $I$ - $V$ ) characteristics of the dc transport can exhibit a negative differential resistance (NDR) in systems described within a single-particle picture and is not necessarily related to electron correlations is well known in semiconductor physics.<sup>1</sup> However, in the nanophysics community the NDR in the  $I$ - $V$  curve is often ascribed to (presumably strong) electron correlations. In fact, some calculations performed on simple but nontrivial models of correlated electrons, like the interacting resonant level model, found no NDR effect far away from resonance,<sup>2,3</sup> while other calculations revealed a more<sup>4</sup> or less<sup>5,6</sup> pronounced NDR effect at resonance. At the end of this note, we shall return to the NDR effect within the interacting resonant level model. Beforehand—and this is the main aim of the present work—we want to emphasize that other, more common sources of the NDR are relevant for nanotransport as well. Therefore, special care is needed if one attempts to ascribe the NDR to electron correlations.

The naive “argument” behind the confusion that the NDR is an electron correlation effect seems to be the following. Within the Landauer approach of the transport in uncorrelated systems, the current resulting from the imbalance between the source and drain chemical potentials  $\mu_S = \varepsilon_F + eV_{sd}/2$  and  $\mu_D = \varepsilon_F - eV_{sd}/2$  is expressed as an integral of the transmission coefficient  $T(\varepsilon)$  over energies from  $\varepsilon = \mu_D$  to  $\varepsilon = \mu_S$ . A NDR cannot occur because the current monotonically increases since the integrand is positive [ $T(\varepsilon) \geq 0$ ] and the integration range increases as the voltage  $V_{sd}$  becomes higher.

To illustrate that this is not the case, let us consider a two-terminal setup (Fig. 1), consisting of a nanosystem [quantum dot(s) or molecule(s)] linked to semi-infinite leads (source and drain) at zero temperature. For simplicity, their bandwidth  $4t$  as well as their coupling to (say, the dot)  $\tau$  will be supposed to be identical. By gradually rising the source-drain voltage  $V_{sd}$  starting from  $V_{sd}=0$ , the drain current  $I_{sd}$  will first progressively increase because the energy window  $\Delta E$  of the (elastic) electron tunneling processes allowed by Pauli’s principle becomes broader [Fig. 1(a)]. However, further increasing  $V_{sd}$  beyond half of the electrode bandwidth ( $eV_{sd}^* \equiv 2t$ ) will diminish this energy window [Fig. 1(b)], and this will be accompanied by a current reduction, which becomes more and more pronounced as the electrode band edge is approach. For  $eV_{sd} \geq 4t$ , elastic tunneling is no longer possible, and the current is completely blocked ( $I_{sd}=0$ ). This fact that the current  $I_{sd}$  should diminish as  $V_{sd}$  exceeds  $V_{sd}^*$

and is completely suppressed above the band edge ( $4t$ ) applies for a general two-terminal setup for a sufficiently weak hybridization  $\Gamma_0 \equiv 2\tau^2/t$ .

To make the analysis more specific, let us consider a point contact (noninteracting resonant level) model, wherein the nanosystem consists of a single nondegenerate energy level  $\varepsilon_g$  linked to one-dimensional semi-infinite electrodes. The second-quantized Hamiltonian reads as

$$\begin{aligned}
 H = & -t \sum_{l \leq -1} (c_l^\dagger c_{l-1} + \text{H.c.}) \\
 & + \mu_S \sum_{l \leq -1} c_l^\dagger c_l - t \sum_{l \geq 1} (c_l^\dagger c_{l+1} + \text{H.c.}) \\
 & + \mu_D \sum_{l \geq 1} c_l^\dagger c_l + \varepsilon_g c_0^\dagger c_0 - \tau (c_{-1}^\dagger c_0 + c_1^\dagger c_0 + \text{H.c.}). \quad (1)
 \end{aligned}$$

As usual, we set  $t=1$  and  $\varepsilon_F=0$ . We assume  $\varepsilon_g \geq 0$  ( $n$ -type conduction) for simplicity, but because model (1) possesses particle-hole symmetry, one can replace  $\varepsilon_g$  by  $|\varepsilon_g|$  below. The electrode-dot coupling  $\tau$  yields well-known expressions of the embedding self-energies  $\Sigma_x(\varepsilon) = \Delta_x(\varepsilon) - i\Gamma_x(\varepsilon)/2$  ( $x = S, D$ ), where<sup>7,8</sup>

$$\begin{aligned}
 \Delta_x(\varepsilon) &= \Delta(\varepsilon - \mu_x), \quad \Gamma_x(\varepsilon) = \Gamma(\varepsilon - \mu_x), \\
 \Delta(\varepsilon) &= \frac{\tau^2 \varepsilon}{2t^2}, \quad \Gamma(\varepsilon) = \frac{\tau^2}{t^2} \sqrt{4t^2 - \varepsilon^2} \theta(2t - |\varepsilon|). \quad (2)
 \end{aligned}$$

They can be inserted into the Dyson equation

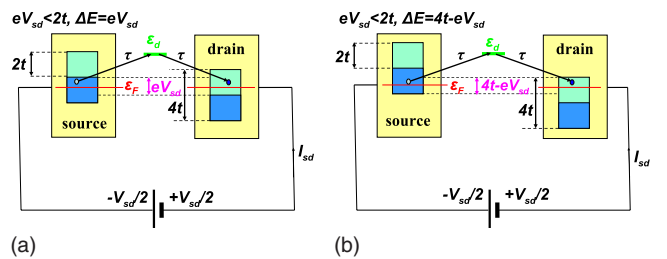


FIG. 1. (Color online) Schematic representation of a typical two-terminal setup. By gradually increasing the source-drain voltage  $V_{sd}$  the energy window  $\Delta E$  of the allowed elastic tunneling processes (a) increases for  $eV_{sd} < 2t$ , but (b) beyond the point  $eV_{sd} = 2t$  (electrode half-bandwidth) it decreases. Elastic tunneling cannot occur for  $eV_{sd} \geq 4t$ .

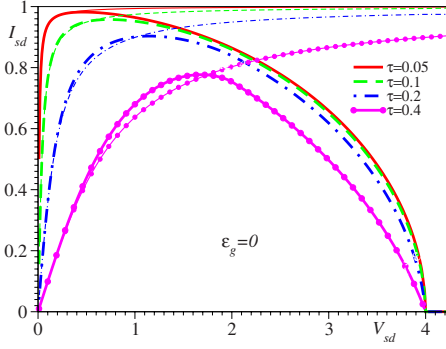


FIG. 2. (Color online)  $I$ - $V$  curves at resonance ( $\varepsilon_g = \varepsilon_F = 0$ ) for  $\tau = 0.05, 0.1, 0.2, 0.4$  computed exactly (thick lines) and within approximation (i) described in the text (thin lines). Current  $I_{sd}$  in units  $I_{sd}^s = \pi e \Gamma_0 / h$ .

$$G^{-1}(\varepsilon) = \varepsilon - \varepsilon_g - \Sigma_S(\varepsilon) - \Sigma_D(\varepsilon) \quad (3)$$

to obtain the retarded Green's function  $G(\varepsilon)$  of the embedded dot. With the aid of the latter, the electric current can be expressed as (electron spin is disregarded)

$$\begin{aligned} I_{sd} &= \frac{e}{h} \int_{\mu_D}^{\mu_S} d\varepsilon T(\varepsilon) = \frac{e}{h} \int_{\mu_D}^{\mu_S} d\varepsilon \Gamma_S(\varepsilon) \Gamma_D(\varepsilon) |G(\varepsilon)|^2 \\ &= \frac{e}{h} \int_{\mu_D}^{\mu_S} d\varepsilon \frac{\Gamma_D(\varepsilon) \Gamma_S(\varepsilon)}{[\varepsilon - \varepsilon_g - \bar{\Delta}(\varepsilon)]^2 + \bar{\Gamma}(\varepsilon)^2 / 4}, \end{aligned} \quad (4)$$

where  $\bar{\Gamma}(\varepsilon) \equiv \Gamma_D(\varepsilon) + \Gamma_S(\varepsilon)$  and  $\bar{\Delta}(\varepsilon) \equiv \Delta_D(\varepsilon) + \Delta_S(\varepsilon)$ .

$I$ - $V$  characteristics computed exactly by means of Eq. (4) at resonance ( $\varepsilon_g = 0$ ) are depicted by the thick lines in Fig. 2. These curves show that, indeed, the current is suppressed as the bias approaches the bandwidth and disappears beyond  $eV_{sd} > 4t$ . Away from resonance ( $\varepsilon_g \neq 0$ ), another aspect is visible in Fig. 3(a). The current vanishes even below the bandwidth  $4t$ . Practically, the suppression is complete at  $V_{sd} = 4t - \varepsilon_g$ ; beyond this value, the  $I$ - $V$  curves only exhibit negligible tails of widths  $\sim \Gamma_0 = 2\tau^2/t$ . On the other side, the exact  $I$ - $V$  characteristics of Figs. 2 and 3(a) reveal that the current decreases well before reaching the value  $V_{sd}^* = 2t/e$ , which one could expect from Fig. 1. This demonstrates that the finite bandwidth effect discussed above is only *one* reason why the NDR should occur.

Significant physical insight can be gained by examining three limits of Eq. (4):

(i) One can approximate the embedding energies by their values at  $\varepsilon = \mu_x$  ( $\Sigma_{S,D} \approx -i\Gamma_0/2$ ) in the *whole* integration range, which means to simply ignore the  $\theta$  step functions in Eq. (2). One then gets the current

$$I_{sd}^{low} = \frac{e\Gamma_0}{h} \left( \arctan \frac{eV - 2\varepsilon_g}{2\Gamma_0} + \arctan \frac{eV + 2\varepsilon_g}{2\Gamma_0} \right). \quad (5)$$

As this amounts to assume that the electrode bandwidth is the largest energy scale (more precisely, for  $V_{sd}, \varepsilon_g, \tau \ll t$ ), Eq. (5) is usually referred to as the wide band limit.

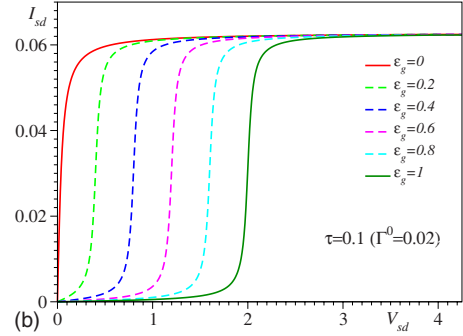
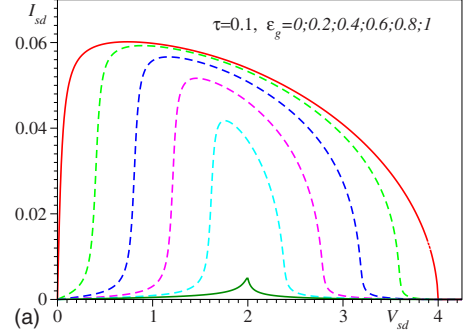


FIG. 3. (Color online)  $I$ - $V$  curves out of resonance for  $\tau = 0.1$  ( $\Gamma_0 = 0.02$ ) computed (a) exactly and (b) within approximation (i) described in the text for  $\varepsilon_g = 0, 0.2, 0.4, 0.6, 0.8, 1$  (values increasing downwards). Current in units  $et/h$ .

(ii) Next, one can compute the current using the electrode density of states (DOS)  $\Gamma_x$  for  $\varepsilon = \mu_x$ , but unlike above, considering the Heaviside  $\theta$  functions in Eq. (2)

$$I_{sd}^{fb} = \frac{e\Gamma_0}{h(1 - \tau^2/t^2)} \left( \arctan \frac{\Lambda_+}{2\Gamma_0} + \arctan \frac{\Lambda_-}{2\Gamma_0} \right), \quad (6)$$

where  $\Lambda_{\pm} \equiv [\min(eV_{sd}, 4t - eV_{sd}) \pm 2\varepsilon_g] \times (1 - \tau^2/t^2)$ . Similar to approximation (i), the electrode DOS is assumed constant, but the fact that the electrode bandwidths are *finite* (the main physical aspect underlying Fig. 1) is taken into account by this approximation.

(iii) Because the main contribution to the integral in Eq. (4) comes from the pole of the Green's function of the isolated dot, one can use the embedding energies calculated at  $\varepsilon = \varepsilon_g$ . In fact, this approximation yields very accurate  $I$ - $V$  curves, which are not shown because they could be hardly distinguished from the exact curves within the drawing accuracy of Figs. 2, 3(a), 4, and 5. More instructive is however to furthermore assume that the voltage  $V_{sd}$  is sufficiently high and extend the integration in Eq. (4) from  $-\infty$  to  $+\infty$ . The result is

$$I_{sd}^{high} = \frac{e}{\hbar} \frac{\Gamma(\varepsilon_g - eV/2)\Gamma(\varepsilon_g + eV/2)}{\Gamma(\varepsilon_g - eV/2) + \Gamma(\varepsilon_g + eV/2)}. \quad (7)$$

$I$ - $V$  curves in the limit (i) are depicted in Fig. 2 (thin lines), Figs. 3(b) and 4. They show a monotonically increasing current, which exhibits a step at  $eV_{sd} \approx 2\varepsilon_g$  of width  $\delta V_{sd}$  increasing with  $\tau$  and rapidly saturates at an  $\varepsilon_g$ -independent

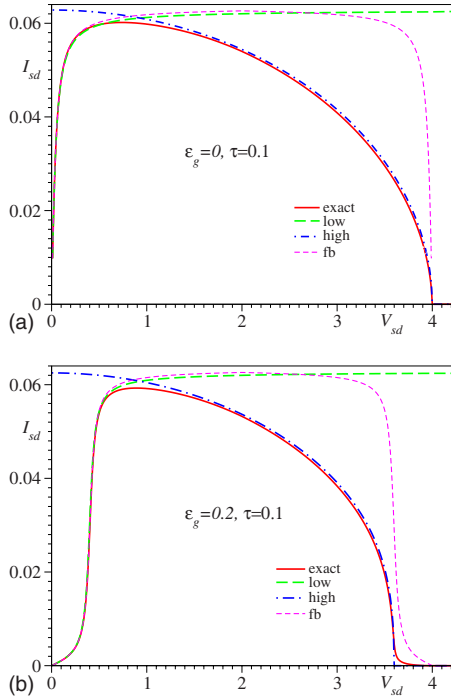


FIG. 4. (Color online)  $I$ - $V$  curves for  $\tau=0.1$  computed exactly and within the approximations described in the text: (a) at resonance  $\varepsilon_d=0$  and (b) out of resonance,  $\varepsilon_d=0.2$ . Current in units  $et/h$ . Labels as in Eqs. (5)–(7).

value  $I_{sd}^s = \pi e \Gamma_0 / h$ . Such curves are usually shown in textbooks, and this feeds the lore of the absent NDR in uncorrelated systems.

What is wrong with the naive argument against the NDR in uncorrelated systems is that the transmission is *not* independent of  $V_{sd}$ . The  $V_{sd}$  dependence enters via the electrode densities of states  $\Gamma_{S,D}$  [cf. Eq. (2)].

On one side, this dependence is considered by the  $\theta$  functions of Eq. (2), which diminish the window of allowed tunneling processes. Approximation (ii) that accounts for this yields two qualitatively correct results: an NDR beyond  $V_{sd}^*$  where the predicted  $I$ - $V$  curve exhibits a cusp (Fig. 5) and a

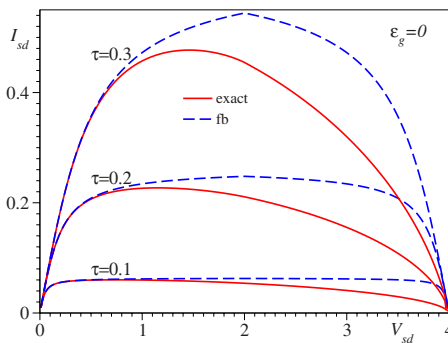


FIG. 5. (Color online)  $I$ - $V$  curves on resonance ( $\varepsilon_g=0$ ) for the three electrode-dot couplings  $\tau$  specified in the inset computed exactly and within approximation (ii) described in the text (label *fb*). Notice that the latter exhibit a cusp at  $eV_{sd}=2t$  that marks the NDR onset in this approximation, which can be substantially higher than the exact NDR onset.

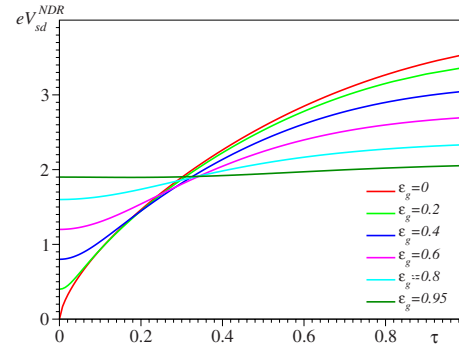


FIG. 6. (Color online) Curves for the NDR onset voltage  $V_{sd}^{NDR}$  computed from Eq. (8) for several level energies  $\varepsilon_g$ . Notice that for smaller electrode-dot couplings  $\tau$  and not too far away from resonance,  $V_{sd}^{NDR}$  is significantly smaller than 2 (half of electrode's bandwidth).

vanishing current for  $eV_{sd} \geq 4t$ . Quantitatively, the NDR onset (at  $V_{sd}=V_{sd}^*$ ) is unsatisfactory; compare these approximate curves (label *fb*) with the exact ones in Figs. 4 and 5. The NDR occurs well below the point predicted by this approximation.

On the other side, not only the  $\theta$  functions but also the  $\varepsilon$  dependence of the electrode DOS [the square roots in Eq. (2)] is important. It is this fact that makes the finite bandwidth argument incomplete. The  $\varepsilon$  dependence of  $\Gamma_{S,D}$  is accounted for within approximation (iii). The comparison with the exact curves (Fig. 4) reveals an excellent agreement at sufficiently higher voltages (as assumed within this approximation) and demonstrates that, to describe quantitatively the NDR, one has to consider both the allowed energy window, which is finite, and the energy dependence of the electrode DOS.

In Fig. 4, we present exact  $I$ - $V$  characteristics from Eq. (4) along with those computed within the three aforementioned approximations [Eqs. (5)–(7)]. As visible there, approximation (i) is accurate for lower voltages, while approximation (iii) is accurate for higher voltages. The crossover occurs at a voltage  $V_{sd}^{NDR}$ , which can be identified with the NDR onset. This value can be obtained by equating

$$I_{sd}^{low}(V_{sd}^{NDR}) = I_{sd}^{high}(V_{sd}^{NDR}). \quad (8)$$

Curves for  $V_{sd}^{NDR}$  are presented in Fig. 6. They show that for situations not very far away from resonance and sufficiently weak electrode-dot couplings  $\tau$ ,  $V_{sd}^{NDR}$  is considerably smaller than the value  $eV_{sd}^* = 2t$  expected from the finite bandwidth argument. The significant departure of the NDR onset predicted exactly and within approximation (ii) is also clearly depicted in Fig. 5. For smaller  $\tau$ 's one can deduce an analytical estimate ( $c \simeq 4$ )

$$V_{sd}^{NDR} \simeq 2\varepsilon_g + c(t\tau^2)^{1/3}. \quad (9)$$

Interesting for nanotransport are the electron level(s) not too misaligned with electrode's Fermi level; otherwise, as illustrated by the curve for  $\varepsilon_g=t$  in Fig. 3(a), the current is very small. Therefore, the results on  $V_{sd}^{NDR}$  expressed by Eq. (9) and Fig. 6 are perhaps the most relevant ones from an

experimental perspective. At resonance and realistic parameters [ $t \approx 1$  eV and  $\tau \approx 1$  meV (Ref. 9)], Eq. (9) yields  $V_{sd}^{NDR} \approx 40$  meV. Based on this estimate, we argue that the NDR discussed here can be observed. On one side, correlations are important only at much lower voltages; in single-electron transistors (SETs) (Ref. 9), the relevant scale is the Kondo temperature  $T_K$  ( $eV_{sd} \lesssim k_B T_K \lesssim 0.1$  meV). For voltages of tens of mV, correlation effects (e.g., Kondo's) are suppressed; the present uncorrelated limit is justifiable. On the other side, the estimated NDR onset voltages ( $\sim 10$  mV) are much lower than the electrode bandwidth ( $\sim 1$  eV), and a material damage prior to the NDR onset can be ruled out. For Si-based SETs, the material can support even much higher values,  $V_{sd} \sim 1$  V.<sup>10</sup> So, we hope that the present estimate will stimulate experimentalists to search NDR effects at moderate  $V_{sd}$ . Again quite relevant for experiments, the NDR onset can be controlled by tuning the level's energy  $\varepsilon_g$  with the aid of a gate potential. Gating methods were routinely employed for nanosystems in the past<sup>9</sup> and recently also in molecular transport.<sup>11</sup> In (weakly correlated) molecules, the level  $\varepsilon_g$  would be either the highest occupied molecular orbital (HOMO) (Ref. 11) or the lowest unoccupied molecular orbital (LUMO) (as in Fig. 1), depending on which is closer to  $\varepsilon_F$ . There,  $\tau \sim 1$  eV and  $|\varepsilon_g| \sim 1$  eV.<sup>11</sup> So, the NDR-onset [cf. Eq. (8) and Fig. 6] is expected at  $V_{sd}$  values of a few eV, slightly higher than used in experiment.<sup>11</sup>

The present analysis can be extended without difficulty to nanosystems or molecules with several "active" electron levels. As long as these levels  $\varepsilon_{g1}, \varepsilon_{g2}, \dots$  are well separated energetically and the hybridization is weak enough (a different situation can also be encountered, see Ref. 12), they manifest themselves as current steps at the voltages  $eV_{sd} \approx 2\varepsilon_{g1}, 2\varepsilon_{g2}, \dots$ . However, even in this case the finite elec-

trode bandwidth and the energy dependence of the electrode DOS remain possible important sources of an NDR.

Similar to other situations encountered in nanotransport,<sup>13,14</sup> we believe that the results for uncorrelated systems are instructive and could also be useful to correctly interpret the nanotransport in correlated systems. In the present concrete case, they could help to unravel the physical origin of the NDR. In the light of the present analysis, it is plausible to ascribe an NDR as an electron correlation effect in cases where the NDR was found within calculations to a correlated nanosystem carried out within the wideband limit. This is, e.g., the case of Refs. 5 and 6, where a weaker NDR effect was obtained at resonance at stronger Coulomb contact interactions. As suggested by Fig. 3, the farther away from resonance, the more is the NDR onset pushed toward higher voltages ( $eV_{sd}^{NDR} > 2|\varepsilon_g|$ ). The values of  $V_{sd}$  chosen in the figures shown in Ref. 3 do not belong to this range and the absence of an NDR could be related to this fact. Unlike the wide (infinite) band limit assumed in the aforementioned references, a discrete model of the electrodes, with a finite bandwidth  $4t$ , *exactly* as in Eq. (1), has been utilized for the numerical calculations of Ref. 4 at resonance. The  $I$ - $V$  curves reported there exhibit a pronounced NDR effect. However, in view of the finite bandwidth assumed in that work, attributing this effect to electron correlations at rather high voltages should be made with special care. We believe that in order to interpret this effect reliably one should first carefully subtract the contribution to the NDR due to the finite bandwidth and the energy-dependent electrode DOS discussed above.

The financial support for this work from the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

\*Also at National Institute for Lasers, Plasma, and Radiation Physics, ISS, RO, 077125 Bucharest, Romania.

<sup>1</sup>W. R. Frensley, *Rev. Mod. Phys.* **63**, 215 (1991).

<sup>2</sup>P. Mehta and N. Andrei, *Phys. Rev. Lett.* **96**, 216802 (2006).

<sup>3</sup>P. Mehta, S. Chao, and N. Andrei, [arXiv:cond-mat/0703426](https://arxiv.org/abs/cond-mat/0703426) (unpublished).

<sup>4</sup>E. Boulat, H. Saleur, and P. Schmitteckert, *Phys. Rev. Lett.* **101**, 140601 (2008).

<sup>5</sup>A. Nishino, T. Imamura, and N. Hatano, *Phys. Rev. Lett.* **102**, 146803 (2009).

<sup>6</sup>B. Doyon, *Phys. Rev. Lett.* **99**, 076806 (2007).

<sup>7</sup>C. Caroli, R. Combescot, P. Nozières, and D. Saint-James, *J. Phys. C* **4**, 916 (1971).

<sup>8</sup>A. Nitzan, *Annu. Rev. Phys. Chem.* **52**, 681 (2001).

<sup>9</sup>D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, *Nature (London)* **391**, 156 (1998).

<sup>10</sup>H. Liu, T. Fujisawa, H. Inokawa, Y. Ono, A. Fujiwara, and Y. Hirayama, *Appl. Phys. Lett.* **92**, 222104 (2008); A. Fujiwara (private communication).

<sup>11</sup>H. Song, Y. Kim, Y. H. Youngsang, H. Jeong, M. A. Reed, and T. Lee, *Nature (London)* **462**, 1039 (2009).

<sup>12</sup>M. C. Toroker and U. Peskin, *J. Phys. B* **42**, 044013 (2009).

<sup>13</sup>I. Bâldea and H. Köppel, *Phys. Rev. B* **78**, 115315 (2008).

<sup>14</sup>I. Bâldea and H. Köppel, *Phys. Rev. B* **80**, 165301 (2009).