Coulomb Drag for Strongly Localized Electrons: A Pumping Mechanism

M. E. Raikh[1] and Felix von Oppen[2]

[1]*Department of Physics, University of Utah, Salt Lake City, Utah 84112*
[2]*Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

The mutual influence of two layers with strongly localized electrons is exercised though the random Coulomb shifts of site energies in one layer caused by electron hops in the other layer. We trace how these shifts give rise to a voltage drop in the passive layer, when a current is passed through the active layer. We find that the microscopic origin of drag lies in the *time correlations* of the occupation numbers of the sites involved in a hop. These correlations are neglected within the conventional Miller-Abrahams scheme for calculating the hopping resistance.

Since the subject was first introduced in pioneering papers [1,2], Coulomb drag between two parallel electronic layers has commanded a lot of attention. Stimulated by the early experiments [3,4], the development of the theory proceeded in two directions: (i) progress on the formalism for calculating drag between conventional two-dimensional electron gases [5–11]; (ii) accommodation of various realizations of interacting layers. These include electron-hole layers [12–14], layers in the superconducting state [9], electronic layers with tunneling links [15], strongly disordered layers at the onset of Anderson localization [16], diffusive layers with correlated disorder [17], and double-layer systems in a perpendicular magnetic field. In the latter case, the picture of drag depends on the magnetic-field regime, namely, classically strong fields [18], quantizing fields [19,20], the vicinity of the integer quantum Hall transition [21], and the fractional quantum Hall regime [22–25].

For all these interacting two-dimensional systems, the theories of Coulomb drag shared the common scenario of quasiparticles (electrons, holes, or composite fermions [22–25]) in two contacting layers scattering off one another in the course of ballistic motion, diffusion [7] or anomalous diffusion [21]. This scattering results in a nonzero average momentum transfer between the active (current-carrying) and passive (open-circuit) layers. A voltage drop is then induced across the passive layer to ensure the absense of a net momentum.

Consider now the deeply insulating regime, where the localization radius of electronic states is smaller than the interelectronic distance. Obviously, the momentum is not a good quantum number in this case, so that the conventional scenario of Coulomb drag does *not* apply. In addition, the picture of long-range time fluctuations of the electron density, and thus the language of *spatially averaged* response functions, is inadequate in the strongly localized regime. This is because the electron motion is due to hopping, which is characterized by an exponentially wide spread in the *local* transition frequencies.

Instead of momentum exchange due to collisions, the coupling between the contacting strongly localized systems is based on random shifts of energy levels in the passive layer caused by electron hops in the active layer, and vice versa. In this situation, it is not immediately obvious how these shifts "communicate" the overall direction of current from the active to the passive layer.

This question is addressed in the present paper. We find that Coulomb drag in the strongly localized regime is due to the *correlated character* of hopping transport. To clarify this point, consider two neighboring sites, 1 and 2, which belong to the current-carrying cluster [26]. At nonzero temperature the time evolution, $n_1(t)$ and $n_2(t)$, of the occupation numbers of sites has the form of telegraph noise. In equilibrium, the average $K_{12}(\tau) = \langle n_1(t)n_2(t + \tau)\rangle$ is an *even* function of τ. Suppose now that current flows in the direction $1 \rightarrow 2$. Then we have $K_{12}(\tau) > K_{12}(-\tau)$ for $\tau > 0$. This reflects the fact that, as the hops within a pair of sites occur preferentially from 1 to 2, the occupation numbers change *in a certain sequence*. This asymmetry is the analog of the current-induced asymmetry between wave vectors $\mathbf{q}$ and $-\mathbf{q}$ of the thermal density fluctuations in the metallic regime, and thus, it is responsible for Coulomb drag between layers with strongly localized electrons.

Note in passing that conventional theories of hopping transport neglect the asymmetry in $K_{12}(\tau)$. In both the noninteracting and the interacting [26,27] cases, the resistance of an elementary hop is usually computed under the assumption of uncorrelated occupation numbers, i.e., $\langle n_1(t)n_2(t')\rangle = \langle n_1\rangle\langle n_2\rangle$ although *equal-time* correlations are sometimes taken into account [26,28].

To show how the preferential sequence in the change of the occupation numbers transforms into drag we consider the simplest possible model as illustrated in Fig. 1. Within this model, the active and the passive layer are each represented by a pair of sites (1 and 2) coupled to metallic contacts (l and r). Both "active" and "passive" pairs of sites can be either empty or singly occupied. The corresponding conditions are $\varepsilon_{1,2} + U > E_F + eV$ for the active layer, where U is the Coulomb interaction between two electrons on sites 1 and 2 and V is the applied voltage, and $\varepsilon_{1',2'} + U > E_F$ for the passive layer. Because of this
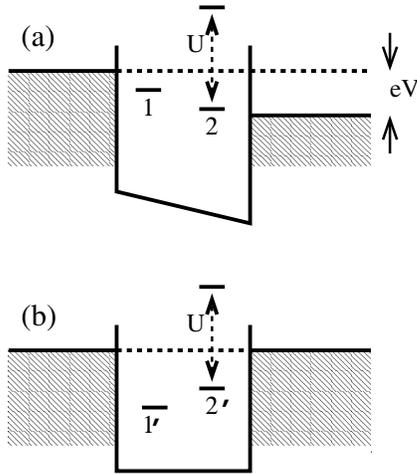
FIG. 1. Schematic drawing of (a) active layer with applied voltage V and two localized states 1 and 2. U indicates the Coulomb energy by which level 1 (2) is elevated if level 2 (1) is occupied. (b) The passive layer also involves two localized states $1'$ and $2'$, but has no voltage applied.
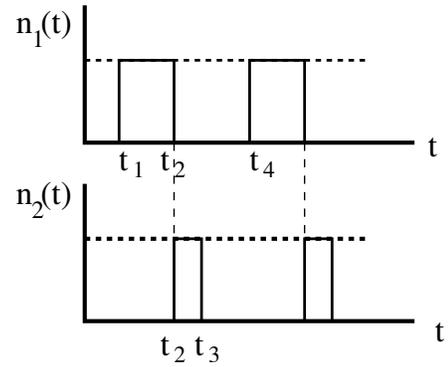

FIG. 2. Schematic time evolution of occupations $n_1(t)$ and $n_2(t)$ of levels 1 and 2, respectively, in the active layer under conditions of a finite current.

condition an electron occupying, say, site 2 elevates the energy level of site 1 above the Fermi level, thus forbidding the tunneling process $l \to 1$ (see Fig. 1). We will treat this model first in the strongly nonlinear regime, in which the temperature T is much lower than V. Subsequently, we will consider the Ohmic regime, where $V \ll T$.

In the strongly nonlinear regime, we can neglect all activation processes and thus, only a *single* sequence of hops is possible in the active layer. Within this sequence, the occupation numbers n_1 and n_2 of sites 1 and 2 undergo the transformations $(1, 0) \to (0, 1) \to (0, 0) \to (1, 0)$. With each repeated cycle of this sequence, an electron is transferred from the left to the right contact. The time evolution of n_1 and n_2 during one cycle is illustrated in Fig. 2. The strong asymmetry in $K_{12}(\tau)$ is evident. The average current through the active layer is equal to [29,30] $\langle I_a \rangle = e/(\tau_1 + \tau_2 + \tau_3)$, where τ_1, τ_2, and τ_3 are the average waiting times for the transitions $l \to 1$, $1 \to 2$, and $2 \to r$, respectively. In the strongly nonlinear regime, the current $\langle I_a \rangle$ is independent of the voltage drop, V.

Consider now the response of the passive layer to a single cycle of $[n_1(t), n_2(t)]$. In the configuration $n_1 = 1$, $n_2 = 0$, the occupied site 1 elevates the level $\varepsilon_{1'}$ by some W_1, and the level $\varepsilon_{2'}$ by some $W_2 = W_1 - \Delta < W_1$. Conversely, when $n_1 = 0$, $n_2 = 1$, the level $\varepsilon_{1'}$ is elevated by W_2, while the level $\varepsilon_{2'}$ is elevated by W_1. (We define the level positions $\varepsilon_{1'}$, $\varepsilon_{2'}$ in the passive layer with respect to an "empty" active layer, $n_1 = n_2 = 0$.) As we will see, the pair $1'$, $2'$ is most "sensitive" when the conditions

$$E_F > \varepsilon_{1'} + W_1 > \varepsilon_{2'} + W_2, \qquad (1)$$

$$\varepsilon_{2'} + W_1 > E_F \qquad (2)$$

are met. Indeed, by virtue of Eq. (1), the passive layer can only make the transition $1' \to 2'$ during the interval $t_1 < t < t_2$ when site 1 in the active layer is occupied. This is illustrated in Fig. 3(a). If the hop $1' \to 2'$ does take place, then two transitions within the passive layer become possible during the next time interval $t_2 < t < t_3$: (i) the electron on site $2'$ goes back to $1'$; cf. Fig. 3(b). This transition is energetically favorable due to Eq. (2); (ii) the electron tunnels from $2'$ into the right contact, r', emptying the passive layer. Since $\varepsilon_{1'} + W_2 < E_F$, this opens the possibility for the hop $l' \to 1'$ (see Fig. 3(c)).

In the first case, there is no transfer of charge from the left to the right contact in the passive layer during the cycle in the active layer. In the second case, however, provided that both transitions $2' \to r'$ and $l' \to 1'$ took place during the interval $t_2 < t < t_3$, the cycle in the active layer results in the transfer of an electron from l' to r'. This transfer is nothing but the drag current.

This drag mechanism closely resembles the operation of a classical *electron pump* [31–33] where two phase-shifted rf signals are applied either to the barriers [31] or to the gate electrodes [32,33] of a single or multiple quantum dot structure. In contrast to adiabatic quantum pumping [34,35], classical pumping is due to the Coulomb blockade, which forces an electron to enter the dot from one contact and to leave into the other contact during each rf cycle. In hopping drag, the pulses $n_1(t)$ and $n_2(t)$ replace the rf signals, with the phase shift determined by the direction of current in the active layer.

We now turn to the calculation of the average drag current, $\langle I_d \rangle$. Denote with $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ the waiting times corresponding to the transitions $l' \to 1'$, $1' \to 2'$, and $2' \to r'$ in the passive layer. The simple system considered here mimics drag between two-dimensional hopping layers when the transitions $1 \to 2$ and $1' \to 2'$ constitute "bottlenecks" for the transport in active and passive layer, respectively, i.e., τ_2 and $\mathcal{T}_2$ are the "long" waiting times. In this limit we have $\langle I_a \rangle \approx e/\tau_2$. The calculation of $\langle I_d \rangle$ is greatly simplified if the "short" waiting times are related as follows $\mathcal{T}_1 \ll \mathcal{T}_3 \ll \tau_3$. The latter conditions ensure that the passive layer returns to the ground state after each
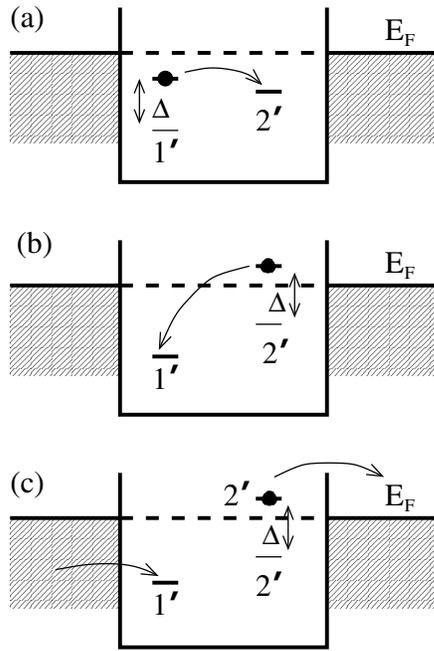
FIG. 3. Variants of evolution of the level occupations in the passive layer during a single cycle in the active layer.

cycle in the active layer. Indeed, if the transition $1' \to 2'$ took place during the time interval $t_1 < t < t_2$ (the corresponding probability is equal to $p_{1'2'} = 1 - \exp[-(t_2 - t_1)/\mathcal{T}_2]$), then the probability that, during the subsequent interval $t_2 < t < t_3$, *both* transitions $2' \to r'$ and $l' \to 1'$ take place is close to 1. This is because the characteristic duration of this interval $(t_3 - t_2) \sim \tau_3$ is much longer than $\mathcal{T}_3$ and $\mathcal{T}_1$. In principle, the transition $2' \to r'$ opens the possibility for two followup hops, namely, $l' \to 1'$ and $r' \to 2'$. However, since $\mathcal{T}_1 \ll \mathcal{T}_3$, tunneling of an electron from the left contact onto site $1'$, thereby returning the passive layer into its ground state, occurs *before* the back hop $r' \to 2'$. Once the transition $l' \to 1'$ occurred, the passive layer remains in the ground state until the end of the cycle in the active layer $t = t_4$ (see Fig. 2). Thus, calculating the drag current reduces to averaging $p_{1'2'}$ which yields

$$\langle I_d \rangle = \frac{\tau_2}{\tau_2 + \mathcal{T}_2} \langle I_a \rangle = \frac{e}{\tau_2 + \mathcal{T}_2}. \tag{3}$$

Remarkably, the drag current is *not* exponentially small compared to the current in the active layer. Instead, both currents are of *comparable magnitude*. This result differs clearly from naive predictions based on spatially-averaged response functions such as a widely used Fermi-golden-rule-type expression for the drag resisitivity [7].

In principle, hops in the passive layer might, in turn, affect transport in the active layer. In particular, occupation of site $2'$ shifts level 2 upward. As a result, the hop $2 \to 1$, in the direction opposite to the net current, might become energetically favorable. The condition that such "feed-back" does not occur is $\varepsilon_1 > \varepsilon_2 + \Delta$. As a concluding remark on our simplified model, we note that for a given polarity of voltage across the active layer (see Fig. 1) the activationless drag exists only if $\varepsilon'_1 < \varepsilon'_2$.

We now turn to the linear regime $V \to 0$ where transport in the active layer becomes more complicated in two respects. First, charge-transfer processes become possible in *both* directions, $l \to r$ and $r \to l$, with only a small difference in their frequencies due to the applied voltage. Second, the dynamics involves "round-trip" processes such as $l \to 1 \to l$ or $l \to 1 \to 2 \to 1 \to l$ that do not result in charge-transfer between the contacts.

It is important to note that the traditional description of hopping transport [26], based on the Miller-Abrahams network, does not capture the realistic occupation dynamics of the sites 1 and 2. Indeed, in this description, a pair of sites is considered as an *isolated system* so that the characteristic times for the changes in the occupations n_1 and n_2 are long ($\sim \tau_2$ in our notations). This is indeed the case for the non-Ohmic regime considered above where an electron *has* to spend a time $\sim \tau_2$ at site 1 before the transition $1 \to 2$ takes place, since it cannot return to the contact l without activation. In the Ohmic regime, however, while the waiting time, τ_2, is long, the occupation of sites 1 and 2, constituting a "bottleneck," changes many times between successive hops $1 \to 2$ (or $2 \to 1$). In other words, a typical charge-transfer process is preceded by many round-trip processes.

As in the non-Ohmic regime, we restrict our considerations to a particular domain of parameters for which the analysis of drag is greatly simplified. We assume that (i) the energies of both sites in the active layer are above the Fermi level ($\varepsilon_{1,2} > E_F$). Thus their average occupations are small, $\langle n_1(t) \rangle, \langle n_2(t) \rangle \ll 1$, i.e., the states 1 and 2 are empty for long periods before an electron enters for a time of the order of τ_1 or τ_2. In addition, we assume for the passive layer that (ii) the sites $1'$ and $2'$ are symmetric in space, so that $\mathcal{T}_1 \approx \mathcal{T}_3$, and close in energy (within temperature). As a result, unlike the situation in Fig. 3(a), the site $1'$ is also elevated *above* the Fermi level, when the site 1 in the active layer is occupied. Finally, we demand that (iii) the "long" time $\mathcal{T}_2$ in the passive layer is *shorter* than the typical time interval during which the active layer is empty, but longer than the time when one of sites 1 or 2 are occupied. The latter assumption implies that an electron in the passive layer hops many times from $1'$ to $2'$ and back between successive round-trip processes in the active layer. We also demand $\mathcal{T}_1, \mathcal{T}_2 \ll \tau_1, \tau_2$.

In this limit, we find the remarkable result that the drag current is simply half the current in the active layer. To understand how this comes about, consider a single charge-transfer process in the active layer, $l \to 1 \to 2 \to r$. Imagine first that the passive layer is in the state $n_{1'} = 0$, $n_{2'} = 1$ when the transition $l \to 1$ takes place. In this case, nothing happens in the passive layer until the transition $1 \to 2$ takes place, after which the passive layer will

undergo the transition $2' \rightarrow r'$ in essentially all cases. This will be immediately followed by $l' \rightarrow 1'$. The transition $2 \rightarrow r$ effectively returns the system to its initial state and one electron has been transferred in *both* layers. As the passive layer has a 50% chance of being in the state $n_{1'} = 0$, $n_{2'} = 1$ at the outset, this implies

$$\langle I_d \rangle = \frac{1}{2} \langle I_a \rangle. \tag{4}$$

Equation (4) assumes that there is *no* comparable drag current when $1'$ is initially occupied in the passive layer. In this case, $l \rightarrow 1$ causes the passive-layer transitions $1' \rightarrow l'$ and then $r' \rightarrow 2'$, which are reversed after $1 \rightarrow 2$ in the active layer. For similar reasons, round-trip processes in the active layer do not lead to a net drag current.

An actual sample in the deeply localized regime will consist of a network of transresistors of the type described above. Hence, we should compute the transresistance of this network. This is by no means a trivial task, even if the transresistances between elements of the conducting networks in active and passive layer are known. This can be illustrated by considering the way in which transresistances combine when connected in sequence or in parallel. These situations are readily analyzed in terms of Kirchhoff's laws. Two transresistors in sequence have a transresistance equal to the sum of the individual transresistances, $R_t^{(1+2)} = R_t^{(1)} + R_t^{(2)}$, similar to ordinary resistors. Two transresistances in parallel are less simple. Here, one can show that

$$R_t^{(1\|2)} = \frac{R_1 R_1' R_t^{(2)} + R_2 R_2' R_t^{(1)}}{(R_1 + R_2)(R_1' + R_2')}, \tag{5}$$

where R_i (R_i') denote the resistances of the active (passive) layer of the ith transresistor. These expressions already allow one to draw an important conclusion about the net transresistance of the entire network. Even though the resistances R_i and R_i' are exponentially large, adding transresistances in sequence or in parallel does not lead to exponential changes in the transresistance. In this sense, the drag current is of comparable magnitude as the current in the active layer, even for the entire network.

In conclusion, the analysis of particular realizations of the active and passive layers, carried out in the present paper illuminates the physics underlying the strong drag in the localized regime. The *discreteness* of the hopping electrons gives rise to Coulomb blockade, which, in turn, opens the possibility of classical pumping.

---

[1] M. B. Pogrebinskii, Fiz. Tekh. Poluprovodn. **11**, 637 (1977) [Sov. Phys. Semicond. **11**, 372 (1977)].
[2] P. M. Price, Physica (Amsterdam) **117B**, 750 (1983).
[3] T. G. Gramila *et al.*, Phys. Rev. Lett. **66**, 1216 (1991).
[4] U. Sivan, P. M. Solomon, and H. Shtrikman, Phys. Rev. Lett. **68**, 1196 (1992).
[5] H. C. Tso, P. Vasilopoulos, and F. M. Peeters, Phys. Rev. Lett. **68**, 2516 (1992).
[6] A.-P. Jauho and H. Smith, Phys. Rev. B **47**, 4420 (1993).
[7] L. Zheng and A. H. MacDonald, Phys. Rev. B **48**, 8203 (1993).
[8] K. Flensberg and B. Y.-K. Hu, Phys. Rev. Lett. **73**, 3572 (1994).
[9] A. Kamenev and Y. Oreg, Phys. Rev. B **52**, 7516 (1995).
[10] K. Flensberg *et al.*, Phys. Rev. B **52**, 14 761 (1995).
[11] K. Flensberg and B. Y.-K. Hu, Phys. Rev. B **52**, 14 796 (1995).
[12] H. C. Tso, P. Vasilopoulos, and F. M. Peeters, Phys. Rev. Lett. **70**, 2146 (1993).
[13] L. Świerkowski, J. Szymański, and Z. W. Gortel, Phys. Rev. Lett. **74**, 3245 (1995).
[14] G. Vignale and A. H. MacDonald, Phys. Rev. Lett. **76**, 2786 (1996).
[15] Y. Oreg and A. Kamenev, Phys. Rev. Lett. **80**, 2421 (1998).
[16] E. Shimshoni, Phys. Rev. B **56**, 13 301 (1997).
[17] I. V. Gornyi, A. G. Yashenkin, and D. V. Khveshchenko, Phys. Rev. Lett. **83**, 152 (1999).
[18] A. V. Khaetskii and Yu. V. Nazarov, Phys. Rev. B **59**, 7551 (1999).
[19] M. C. Bønsager *et al.*, Phys. Rev. Lett. **77**, 1366 (1996); Phys. Rev. B **56**, 10 314 (1997).
[20] F. von Oppen, S. H. Simon, and A. Stern, Phys. Rev. Lett. **87**, 106803 (2001).
[21] E. Shimshoni and S. L. Sondhi, Phys. Rev. B **49**, 11 484 (1994).
[22] I. Ussishkin and A. Stern, Phys. Rev. B **56**, 4013 (1997).
[23] A. Sakhi, Phys. Rev. B **56**, 4098 (1997).
[24] Y.-B Kim and A. J. Millis, Physica (Amsterdam) **4E**, 171 (1999).
[25] B. N. Narozhny, I. L. Aleiner, and A. Stern, Phys. Rev. Lett. **86**, 3610 (2001).
[26] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer, Berlin, 1984).
[27] A. L. Efros and B. I. Shklovskii, in *Electron-Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).
[28] E. I. Levin, N. V. Lien, and B. I. Shklovskii, Fiz. Tekh. Poluprovodn. **16**, 815 (1982) [Sov. Phys. Semicond. **16**, 523 (1982)]; K. S. Chase and D. J. Thouless, Phys. Rev. B **39**, 9809 (1989).
[29] M. E. Raikh and A. Asenov, Superlattices Microstruct. **11**, 325 (1992).
[30] Y. A. Kinkhabwala and A. N. Korotkov, Phys. Rev. B **62**, R7727 (2000).
[31] L. P. Kouwenhoven *et al.*, Phys. Rev. Lett. **67**, 1626 (1991).
[32] H. Pothier *et al.*, Physica (Amsterdam) **169B**, 573 (1991).
[33] H. Pothier *et al.*, Europhys. Lett. **17**, 249 (1992).
[34] B. L. Altshuler and L. I. Glazman, Science **283**, 1864 (1999).
[35] M. Switkes *et al.*, Science **283**, 1905 (1999), and references therein.