Current Fluctuations in Nonequilibrium Diffusive Systems: An Additivity Principle

T. Bodineau and B. Derrida

1Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, Universités Paris VI et VII, 4 place Jussieu, Case 188, F-75252 Paris, Cedex 05, France
2Laboratoire de Physique Statistique, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France

(Rceived 22 January 2004; published 3 May 2004)

We formulate a simple additivity principle allowing one to calculate the whole distribution of current fluctuations through a large one dimensional system in contact with two reservoirs at unequal densities from the knowledge of its first two cumulants. This distribution (which in general is non-Gaussian) satisfies the Gallavotti-Cohen symmetry and generalizes the one predicted recently for the symmetric simple exclusion process. The additivity principle can be used to study more complex diffusive networks including loops.

DOI: 10.1103/PhysRevLett.92.180601

Understanding the fluctuations of the steady state current through a system in contact with two (or more) heat or particle reservoirs is one of the simplest and most fundamental problems of nonequilibrium physics [1–3]. For quantum particles such as (weakly interacting) electrons which satisfy the Pauli principle, the whole distribution (the full counting statistics) of the number of particles transferred between the two reservoirs during a long time interval is known [4] and it can be calculated by a number of theoretical approaches [3,5–7], ranging from the theory of random matrices [4,8] to the Boltzmann-Langevin semiclassical description [9].

For systems of purely classical interacting particles [1,2] in contact with two reservoirs the theory is, to our knowledge, less developed. However, for a number of stochastic models of classical interacting particles [10–13], the cumulants of the current fluctuations were found to coincide with those previously known of noninteracting quantum particles. It is, of course, an important issue to know under what condition a classical particle system could present the same distribution of current as in the quantum case.

For most theoretical approaches developed in the quantum or in the classical description, the calculation of the cumulants becomes harder and harder as the degree of the cumulants increases. The goal of the present Letter is to show that for classical stochastic models, if one postulates a simple additivity principle for the current fluctuations, the whole distribution of current fluctuations can be calculated from the knowledge of the first two cumulants of the current.

We consider here a one dimensional diffusive open system of length \( N \) (with \( N \) large) in contact, at its two ends, with two reservoirs of particles at densities \( \rho_a \) and \( \rho_b \). In the bulk, the system evolves under some conservative stochastic dynamics and, at the boundaries, particles are created or annihilated to match the densities of the reservoirs.

Let \( Q_t \) be the integrated current up to time \( t \), i.e., the number of particles that went through the system during time \( t \). For large \( N \), we shall see that the whole distribution of the fluctuations of \( Q_t \) depends on only two macroscopic parameters \( D(\rho) \) and \( \sigma(\rho) \) defined as follows: Suppose that for \( \rho_a = \rho + \Delta \rho \) and \( \rho_b = \rho \) with \( \Delta \rho \) small, we know that in the steady state Fick’s law holds,

\[
\frac{\langle Q_t \rangle}{t} = \frac{1}{N} D(\rho)\Delta \rho. \quad (1)
\]

Suppose that for \( \rho_a = \rho_b = \rho \) (in which case \( \langle Q_t \rangle = 0 \)), we also know that, for large \( t \), the variance is

\[
\frac{\langle Q_t^2 \rangle}{t} = \frac{1}{N} \sigma(\rho). \quad (2)
\]

The main result of the present Letter is that, using a simple additivity principle (10) and (11), one can predict all the cumulants of \( Q_t \), for arbitrary \( \rho_a \) and \( \rho_b \). If we define the integrals \( I_n \) by

\[
I_n = \int_{\rho_b}^{\rho_a} D(\rho)\sigma(\rho)^n d\rho,
\]

the first cumulants of \( Q_t \), for large \( t \), are given by

\[
\frac{\langle Q_t \rangle}{t} = \frac{1}{N} I_1, \quad \frac{\langle Q_t^2 \rangle - \langle Q_t \rangle^2}{t} = \frac{1}{N} I_2, \quad (3)
\]

\[
\frac{\langle Q_t^3 \rangle_c}{t} = \frac{1}{N} \frac{3(I_4I_1 - I_2^2)}{I_1^3}, \quad (4)
\]

\[
\frac{\langle Q_t^4 \rangle_c}{t} = \frac{1}{N} \frac{3(5I_4I_1^2 - 14I_4I_2I_3 + 9I_3^2)}{I_1^6}. \quad (5)
\]

The case \( \rho_a = \rho_b \) can be obtained by letting \( \rho_a \) tend to \( \rho_b \) in the previous expressions.

More generally, all the higher cumulants can be obtained from the knowledge of \( \mu_N \) which characterizes the large \( t \) growth of the generating function of \( Q_t \),

\[
\mu_N(\lambda, \rho_a, \rho_b) = \lim_{t \to \infty} t^{-1} \ln \langle e^{\lambda Q_t} \rangle. \quad (6)
\]

We are going to show that, for large \( N \), \( \mu_N \) takes the...
following parametric form:

$$\mu_N(\lambda, \rho_a, \rho_b) = -\frac{K}{N} \left[ \int_{\rho_a}^{\rho_b} \frac{D(\rho) d\rho}{\sqrt{1 + 2K\sigma(\rho)}} \right]^2 + O \left( \frac{1}{N} \right), \tag{7}$$

where $K = K(\lambda, \rho_a, \rho_b)$ is the solution of

$$\lambda = \int_{\rho_a}^{\rho_b} \frac{D(\rho)}{\sigma(\rho)} \frac{1}{\sqrt{1 + 2K\sigma(\rho)}} - 1. \tag{8}$$

As $\mu_N = (\lambda(Q_1) + \lambda^2(Q_1^2)/2 + \lambda^3(Q_1^3)/6 + \ldots)/t$, one simply needs to expand (7) and (8) in powers of $K$ and eliminate $K$ to obtain $\mu_N$ as a power series of $\lambda$ and the cumulants such as (3)–(5).

Note that (7) and (8) are valid only for $\rho_a \neq \rho_b$ and in the range of values of $\lambda$ where $K$ is such that the argument of the square root in the integrants does not vanish. We checked that they can also be analytically continued to cover the other ranges of $\lambda$ and the case $\rho_a = \rho_b$.

Our derivation of (7) and (8) is based on an additivity principle that we formulate now. The probability $P_N(q, \rho_a, \rho_b, t)$ of observing an integrated current $Q_t = qt$ exponential in $t$ for large $t$,

$$P_N(q, \rho_a, \rho_b, t) \sim \exp \left[ t F_N(q, \rho_a, \rho_b, \text{contacts}) \right], \tag{9}$$

where $F_N(q, \rho_a, \rho_b, \text{contacts})$ depends on the length $N$ of the system, on $q$, on the densities $\rho_a$ and $\rho_b$ in the two reservoirs, and on the nature of the contacts of the system with the two reservoirs. ($F_N$ is negative and vanishes only when $q$ takes its most likely value $\langle Q_t \rangle/t$). When $N$ is large and $q$ is of order $1/N$, the effect of the contacts becomes negligible and asymptotically $F_N(q, \rho_a, \rho_b)$ depends only on $q, \rho_a, \rho_b$, on the length $N$, and on the bulk properties of the system. We then assume that, for large $N$ and $q$ of order $1/N$, the large deviation function $F_N(q, \rho_a, \rho_b)$ satisfies the following additivity principle:

$$F_{N+N}(q, \rho_a, \rho_b) \simeq \max_{\rho} \{ F_N(q, \rho_a, \rho) + F_N(q, \rho, \rho_b) \}. \tag{10}$$

This property simply means that the two subsystems are independent, except that they try to adjust the density $\rho$ at their contact to maximize the following product:

$$P_{N+N}(q, \rho_a, \rho_b, t) \sim \max_{\rho} \{ P_N(q, \rho_a, \rho, t)P_N(q, \rho, \rho_b, t) \}. \tag{11}$$

We also make the following scaling hypothesis:

$$F_N(q, \rho_a, \rho_b) \simeq N^{-1} G(Nq, \rho_a, \rho_b). \tag{12}$$

This hypothesis, which is valid, in particular, for the symmetric simple exclusion process, means that $\mu_N$ defined by (6) is of order $1/N$ for large $N$ (see [13]).

If we write $N = (N + N')x$, i.e., we split a system of macroscopic unit length into two parts of lengths $x$ and $1 - x$, then (10) and (11), lead to

$$G(q, \rho_a, \rho_b) = \max_{\rho} \left\{ \frac{G(qx, \rho_a, \rho)}{x} + \frac{G(q(1 - x), \rho, \rho_b)}{1 - x} \right\}, \tag{13}$$

If we keep dividing the system into smaller and smaller pieces and use that for a piece of small (macroscopic) size $\Delta x$ (i.e., of $N\Delta x$ sites), one has (1), (2), (10), and (11),

$$\frac{1}{\Delta x} G(q\Delta x, \rho, \rho + \Delta\rho) \approx -\frac{[q\Delta x + D(\rho)\Delta\rho]^2}{2\sigma(\rho)\Delta x}; \tag{14}$$

one finds a variational form for $G$,

$$G(q, \rho_a, \rho_b) = -\min_{\rho(0)} \left[ \int_0^{\rho(1)} \frac{[q + D(\rho(x))\rho'(x)]^2}{2\sigma(\rho(x))} dx \right], \tag{15}$$

where the minimum is over all the functions $\rho(x)$ with boundary conditions $\rho(0) = \rho_a$ and $\rho(1) = \rho_b$.

The optimal $\rho(x)$ in (14) satisfies

$$q^2 a'(\rho) - c'(\rho) \left( \frac{d\rho}{dx} \right)^2 - 2c(\rho) \frac{d^2\rho}{dx^2} = 0,$$

where $a(\rho) = (2\sigma(\rho))^{-1}$ and $c(\rho) = D^2(\rho)a(\rho)$. Multiplying the above equation by $d\rho(x)/dx$, one obtains after one integration

$$D^2(\rho) \left( \frac{d\rho}{dx} \right)^2 = q^2(1 + 2K\sigma(\rho)), \tag{16}$$

where $K$ is a constant of integration.

To proceed further one needs to determine the sign of $d\rho/dx$. The simplest case is when $\rho(x)$ is monotone, and this happens when $q$ is close enough to its average value for $\rho_a \neq \rho_b$ (this corresponds to values of $K$ small enough for the right-hand side (rhs) of (15) to vanish). The investigation of this regime is enough to determine all the cumulants. If, for example, $\rho_a > \rho_b$, the optimal $\rho(x)$ is decreasing for small $K$,

$$\frac{d\rho}{dx} = -\frac{q}{D(\rho)} \sqrt{1 + 2K\sigma(\rho)}, \tag{17}$$

and this leads to the following expression for $G$:

$$G = q \int_{\rho_a}^{\rho_b} \frac{D(\rho)}{\sigma(\rho)} \left[ 1 - \frac{1 + K\sigma(\rho)}{\sqrt{1 + 2K\sigma(\rho)}} \right] d\rho. \tag{18}$$

where the constant $K$ is determined by

$$q = \int_{\rho_a}^{\rho_b} d\rho \frac{D(\rho)}{\sqrt{1 + 2K\sigma(\rho)}}. \tag{19}$$

One can then show that

$$\frac{\partial G}{\partial q} = \frac{G}{q} + Kq = \int_{\rho_a}^{\rho_b} d\rho \frac{D(\rho)}{\sigma(\rho)} \left[ 1 - \frac{1}{\sqrt{1 + 2K\sigma(\rho)}} \right], \tag{20}$$

where the derivative is taken keeping $\rho_a$ and $\rho_b$ fixed, and using the fact that $\mu_N = N^{-1} \max_{\lambda} \{ \lambda q + G(q, \rho_a, \rho_b) \}$, one obtains (7) and (8).

When the optimal $\rho(x)$ is no longer monotone, i.e., $K$ is negative enough for the right-hand side of (15) to vanish, the expressions (7), (8), (17), and (18) of $\mu_N, \lambda, G, q$ are modified. We checked that their new
expressions are simply the analytic continuations of (7), (8), (17), and (18).

In general, when the system is in equilibrium ($\rho_a = \rho_b = \rho$) the fluctuations given by (14) are non-Gaussian. However when $\rho_a = \rho_b = \rho^*$, where $\rho^*$ is the density for which $\sigma(\rho)$ is maximum, the optimal $\rho(x)$ in (14) satisfies $\rho'(x) = 0$ and the fluctuations become Gaussian $\{G(q, \rho^*, \rho^*) = -q^2/[2\sigma(\rho^*)]\}$ in agreement with the conjecture made in [13] for a specific model, the symmetric simple exclusion process.

By expanding the square in (14), we see that the optimal profile $\rho(x)$ is the same for $q$ and $-q$ so that

$$G(-q, \rho_a, \rho_b) = G(q, \rho_a, \rho_b) - 2q \int_{\rho_a}^{\rho_b} \frac{D(\rho)}{\sigma(\rho)} d\rho,$$

which is the Gallavotti-Cohen relation [13–15].

The optimum is achieved when $q'(y_1 + y_2) = qy_2$; thus,

$$G_{\text{loop}}(q, \rho_a, \rho_b) = G(qu, \rho_a, \rho_b)/u,$$

with $u = x_1 + (y_1^{-1} + y_2^{-1})^{-1} + x_2$. So the current fluctuations for the system with a loop are the same as for a linear system with a length given by Kirchoff’s law for the addition of resistors.

We consider now two specific examples of stochastic dynamics on a 1d lattice, the symmetric simple exclusion process (SSEP) and the zero range process (ZRP). The number of particles at site $i \in \{0, N\}$ is denoted by $\eta_i$.

$$\sum_{i,j} \partial_t \langle Q_i^j \rangle = 2 \sum_{i,j} \langle Q_i^j(\eta_i(1 - \eta_{i+1}) - \eta_{i+1}(1 - \eta_i)) \rangle = 2 \sum_{i,j} \langle Q_i^j(\eta_i - \eta_{i+1}) \rangle + 2 \sum_i \langle \eta_i(1 - \eta_{i+1}) \rangle.$$

The first term simplifies as

$$\sum_{i,j} \langle Q_i^j(\eta_i - \eta_{i+1}) \rangle = \left(\sum_{i} \langle Q_i^i \rangle\right) \langle \eta_0 \rangle - \left(\sum_{i} \langle Q_i^i \rangle\right) \langle \eta_N \rangle$$

if sites 0 and $N$ are kept in equilibrium with the left-hand and right-hand reservoirs and it vanishes for $\rho = \rho_a = \rho_b$. For $\rho_a = \rho_b$, the stationary measure is product (i.e., the occupation numbers of the sites are independent) so that $\sigma(\rho) = 2\rho(1 - \rho)$ according to (2). The cumulants derived in [13], as well as the expression conjectured for $\mu_N$,

$$\mu_N(\lambda) = -N^{-1}[\sin^{-1}(\sqrt{\omega})]^2, \quad \omega \leq 0,$$

where $\omega = (1 - e^{-\lambda})[e^{\lambda} \rho_a - \rho_b - (e^\lambda - 1)\rho_a \rho_b]$ can be recovered from (7) and (8). This can be seen by noticing that the optimal profile solution of (16) is

$$\rho(x) = \frac{1}{2} \left(1 + \frac{\sin[2(\theta_a + (\theta_b - \theta_a)x)]}{\sin(2f)}\right),$$

Note that $G$ given by (17) vanishes for $K = 0$, in which case the most likely profile [16] satisfies (16). Consider now a system composed of four parts, as in Fig. 1. The left-hand reservoir is connected to $C$ by a chain of length $Nx_1$. Between $C$ and $D$ there is a loop made of two chains in parallel of lengths $Ny_1$ and $Ny_2$, and $D$ is connected to the right-hand reservoir by a chain of length $Nx_2$. From the additivity principle, one should have

$$N \partial_t \langle Q_i^j \rangle = \sum_{0 \leq i < N-1} \langle \eta_i - \eta_{i+1} \rangle,$$

thus,

$$\sum_i \langle Q_i^j \rangle^2 = \sum_i \langle \Phi(\eta_i) - \Phi(\eta_{i+1}) \rangle,$$

where the parameters $f, \theta_a, \theta_b$ are fixed by $K = \tan^2(2f), \rho(0) = \rho_a, \rho(1) = \rho_b$. In terms of these parameters, $\alpha$ and $\mu_N$ take the form

$$\alpha = \log \left[\frac{\cos(f + \theta_a)\cos(f - \theta_b)}{\cos(f - \theta_a)\cos(f + \theta_b)}\right],$$

$$\mu_N = -(\theta_a - \theta_b)^2.$$

For the ZRP the number of particles on each site can be arbitrary and the jump rate $\Phi(\eta_i)$ from site $i$ to each of its neighbors is an increasing function of the number of particles $\eta_i$ at this site. We choose $\Phi(0) = 0$. We have

$$\sum_i \partial_t \langle Q_i^j \rangle = \sum_i \langle \Phi(\eta_i) - \Phi(\eta_{i+1}) \rangle = \Psi(\rho_a) - \Psi(\rho_b),$$

with the expectation of $\Phi$ under the stationary measure at density $\rho$ is denoted by $\Psi(\rho)$. We also have

$$N^2 \partial_t \langle Q_i^j \rangle^2 = 2 \sum_i \langle Q_i^j(\Phi(\eta_i) - \Phi(\eta_{i+1})) \rangle + 2 \sum_i \langle \Phi(\eta_i) \rangle.$$

Note that $G$ given by (17) vanishes for $K = 0$, in which case the most likely profile [16] satisfies (16). Consider now a system composed of four parts, as in Fig. 1. The left-hand reservoir is connected to $C$ by a chain of length $Nx_1$. Between $C$ and $D$ there is a loop made of two chains in parallel of lengths $Ny_1$ and $Ny_2$, and $D$ is connected to the right-hand reservoir by a chain of length $Nx_2$. From the additivity principle, one should have

$$N \partial_t \langle Q_i^j \rangle = \sum_{0 \leq i < N-1} \langle \eta_i - \eta_{i+1} \rangle,$$

thus,

$$\sum_i \langle Q_i^j \rangle^2 = \sum_i \langle \Phi(\eta_i) - \Phi(\eta_{i+1}) \rangle,$$
As for the SSEP, the first term in the rhs of the above equation vanishes when \( \rho_a = \rho_b \), and we get \( D(\rho) = \sigma(\rho) / 2 \) and \( \sigma(\rho) = 2 \Psi(\rho) \) according to (1) and (2). Thus from (7) and (8),

\[
\mu_N(\lambda) = (1 - e^{-\lambda})[e^\lambda \sigma(\rho_a) - \sigma(\rho_b)] / 2N.
\]

This generalizes the case of noninteracting particles for which \( \sigma(\rho) = 2\rho \). The optimal profile is obtained by

\[
\sigma(\rho(x)) = (\theta_a + (\theta_b - \theta_a)x^2 - 1) / (2K),
\]

where \( \theta_a, \theta_b \) are fixed by \( \rho(0) = \rho_a \) and \( \rho(1) = \rho_b \). In particular, the expression of \( \mu_N \) follows from

\[
\lambda = \log \left( 1 + \frac{\theta_b}{\theta_a} \right), \quad \mu_N(\lambda) = -\frac{(\theta_a - \theta_b)^2}{4K}.
\]

The additivity principle (9) and (10) formulated here and its variational expression (14) can be derived (work in progress) from the hydrodynamic large deviation theory [17–19]. This theory was extended recently by Bertini et al. [20] to calculate the density large deviation functional of the steady state as the optimal cost for a space/time density fluctuation. For diffusive systems, the probability of observing an atypical space/time density profile over a time \( t \) can be estimated by the exponential of a functional depending only on \( D(\rho), \sigma(\rho) \), and on the density \( \rho(x, t) \) [see, e.g., (17–19)]. The optimal strategy to create a fluctuation of the current \( Q_t = q t \) over a very long time \( t \) is to create a fixed density profile \( \rho(x) \) in order to facilitate the deviation of the current, and (14) can be understood as the cost for maintaining this atypical density profile. The optimal profile controlling here the current fluctuations is time independent, in contrast to the one which controls the steady static density fluctuations that Bertini et al. [20] had to calculate. This is why our task here was easier and the additivity principle (10) is simpler than the one obtained in [21] for the steady state fluctuations of the density.

The large deviation function \( F \) defined in (9) seems to give a microscopic definition of the nonequilibrium free energy introduced in [22] where an additivity relation similar to (10) was conjectured.

It would be interesting to see whether the Bertini et al. macroscopic fluctuation theory satisfies a generalized additivity principle for time-dependent densities in the reservoirs. Other interesting extensions of the present work include the study of the effect of asymmetry in the bulk dynamics (i.e., of a field which favors jumps of particles from left to right) [23–27] or the analysis of more complex networks, in particular, of systems in contact with three or more reservoirs [28–32].

Of course, a challenging issue would be to see whether the additivity principle could be valid for mechanical systems satisfying (1) and (2) without an intrinsic source of noise as in the stochastic systems considered here.

We thank B. Douçot, G. Giacomin, J. L. Lebowitz, P.-E. Roche, and F. van Wijland for helpful discussions. T.B. acknowledges the hospitality of Rutgers University and support by NSF Grant No. DMR 01-279-26.