J Stat Phys (2007) 128: 1307–1320
DOI 10.1007/s10955-007-9368-z

Gallavotti–Cohen Theorem, Chaotic Hypothesis and the Zero-Noise Limit

Jorge Kurchan

Received: 28 February 2007 / Accepted: 14 June 2007 / Published online: 4 August 2007

Abstract The Fluctuation Relation for a stationary state, kept at constant energy by a *deterministic* thermostat—the Gallavotti–Cohen Theorem—relies on the ergodic properties of the system considered. We show that when perturbed by an energy-conserving random noise, the relation follows trivially for any system at finite noise amplitude. The time needed to achieve stationarity may stay finite as the noise tends to zero, or it may diverge. In the former case the Gallavotti–Cohen result is recovered, while in the latter case, the crossover time may be computed from the action of 'instanton' orbits that bridge attractors and repellors. We suggest that the 'Chaotic Hypothesis' of Gallavotti, Cohen and Ruelle can thus be reformulated as a matter of stochastic stability of the measure in trajectory space. In this form this hypothesis may be directly tested.

1 Introduction

The Fluctuation Theorems are relations satisfied by the distribution of entropy production $\bar{\sigma}_\tau = \frac{1}{\tau}\int_o^\tau dt\sigma(t)$ [1–9]

$$\frac{P(\bar{\sigma}_\tau)}{P(-\bar{\sigma}_\tau)} \sim e^{\tau\bar{\sigma}_\tau} \quad \text{or} \quad \langle e^{-\lambda\tau\bar{\sigma}_\tau}\rangle = \langle e^{-(1-\lambda)\tau\bar{\sigma}_\tau}\rangle \quad \forall\lambda. \tag{1.1}$$

They hold in a series of different settings. The proofs are very simple and general when they apply to systems starting from an equilibrium configuration [2], or in contact with a thermal bath involving stochastic noise that guarantees ergodicity [10, 11]. On the contrary, when the system is treated in the stationary state achieved thanks to a *deterministic* thermostat, the derivation is not simple and requires a knowledge of the ergodic properties that is very rarely accessible analytically [3, 4, 13].

In those cases in which the nature of the thermostat is thought to be physically irrelevant, it is quite natural to prefer the stochastic baths because they invoke the least hypotheses,

J. Kurchan (✉)
PMMH-ESPCI, CNRS UMR 7636, 10 rue Vauquelin, 75005 Paris, France
e-mail: jorge@pmmh.espci.fr

making the analysis conceptually simpler. On the other hand, the very fact that for deterministic systems the Fluctuation Relation is not automatically satisfied makes it an interesting tool to study questions of ergodic theory that are still largely open.

For a purely deterministic case, a forced system converges to an 'attractor' set, while the time-reversed dynamics converges to a 'repellor' set. Loosely speaking, the Gallavotti–Cohen theorem is then based on two types of hypotheses: (i) that the attractor is sufficiently chaotic, and (ii) that the attractor and repellor are interwoven fractals with sufficiently overlapping closures. As the forcing is increased, attractor and repellor tend to separate, and even if the attractor remains chaotic the Fluctuation Relation no longer holds in general.[1] For macroscopic systems at reasonable levels of forcing, one can assume that the hypotheses above are effectively (if not always strictly) satisfied, leading to the Chaotic Hypothesis [15].

In this paper we recast the content of the previous paragraph in the language of stochastic stability [16–18]: we first show that adding a small, energy-conserving noise the Fluctuation Relation is trivially satisfied by any model, and we then analyze the limit of noise going to zero. Indeed, all the subtleties of the Gallavotti–Cohen theorem and of the Chaotic Hypothesis are a result of this limit. The small noise limit has moreover the advantage that an approximation scheme (analogous to the semiclassical treatment) becomes exact, so that one can analyze it in detail.

The analysis of this paper suggests the following alternative Chaotic Hypothesis: *The large deviation function of macroscopic observables can be regarded as stochastically stable.* Apart from its conceptual appeal, this formulation has the advantage of being easily testable numerically.

2 Thermostatted Dynamics

We shall consider Hamilton's equations with forcing, a thermostatting mechanism, and energy conserving noise [19–22]:

$$\begin{aligned} \dot{q}_i &= p_i, \\ \dot{p}_i &= -\frac{\partial V(\mathbf{q})}{\partial q_i} + g_{ij}\eta_j - f_i(\mathbf{q}) + \gamma(t) p_i = -\frac{\partial V(\mathbf{q})}{\partial q_i} + g_{ij}(\eta_j - f_j) \end{aligned} \tag{2.2}$$

here and in what follows we assume summation convention.

- η_j are white, independent noises of variance ϵ.
- $g_{ij} = \delta_{ij} - \frac{p_i p_j}{\mathbf{p}^2}$ is the projector onto the space tangential to the energy surface.
- $f_i(\mathbf{q})$ is a forcing term that depends on the coordinates only.
- $\gamma(t) = \frac{\mathbf{f} \bullet \mathbf{p}}{\mathbf{p}^2}$. Multiplying the first of (2.2) by $\frac{\partial V(\mathbf{q})}{\partial q_i}$, the second by p_i and adding, one concludes that energy is conserved.
- The product $g_{ij}\eta_j$ is rather ill defined because both terms depend on time and η_j is discontinuous. The ambiguity is raised by stipulating that it has to be interpreted in the Stratonovitch convention [23], the meaning of which will be made clear below.

The probability evolves through [23]:

$$\dot{P} = -HP \tag{2.3}$$

[1] And then the only hope is that the attractor itself is composed of two intertwined sets, related by some discrete symmetry playing for them the role played by time-reversal for the attractor-repellor pair. See [14].

where H is the operator:

$$H = p_i \frac{\partial}{\partial q_i} - \frac{\partial V(\mathbf{q})}{\partial q_i}\frac{\partial}{\partial p_i} + \frac{\partial}{\partial p_i}[\gamma p_i] - \frac{\partial}{\partial p_i} f_i - \epsilon \frac{\partial}{\partial p_j} g_{ij} g_{il} \frac{\partial}{\partial p_l}. \tag{2.4}$$

The precise factor ordering in the last term is exactly what we meant by (2.2) being in the 'Stratonovitch convention'. In the absence of driving $\mathbf{f} = 0$ it is easy to check that H annihilates any function that depends on the phase-space coordinates only through the energy $E = \frac{\mathbf{p}^2}{2} + V$. Hence, the noise respects the microcanonical measure, in that case.

We shall consider as usual the 'time-reversal transformation' $\tilde{H} = [QHQ^{-1}]^\dagger$, where Q is the operator that reverses momenta: $Qp_iQ^{-1} = -p_i$ and $Q\frac{\partial}{\partial p_i}Q^{-1} = -\frac{\partial}{\partial p_i}$. Under momentum reversal all terms change signs, except the last. Under Hermitean conjugation the factor order is reversed, and derivatives change sign. We get:

$$\begin{aligned} \tilde{H} &= [QHQ^{-1}]^\dagger = H + \sigma, \\ &\qquad \sigma \equiv [\gamma p_i]\frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_i}[\gamma p_i] = -(d-1)\gamma \end{aligned} \tag{2.5}$$

with $2d$ is the dimension of phase-space, which can be checked by straightforward calculation. We hence have:

$$[QHQ^{-1}]^\dagger = H + \sigma, \tag{2.6}$$

$$[Q(H + \lambda\sigma)Q^{-1}]^\dagger = H + (1-\lambda)\sigma \tag{2.7}$$

because the entropy production rate σ changes sign under time-reversal and conjugation.

From here it is standard to show that for every finite value of ϵ the fluctuation theorem holds [10, 11]. For completeness, let us outline the proof: given any initial and final configurations $x_i = (q_i, p_i)$ and $x_f = (q_f, p_f)$ the expectation of the total work is

$$\begin{aligned} \langle e^{-\lambda \int_o^\tau \sigma dt} \rangle &= \langle x_f | e^{-\tau(H+\lambda\sigma)} | x_i \rangle = \langle x_i | e^{-\tau(H+\lambda\sigma)^\dagger} | x_f \rangle \\ &= \langle x_i | Q^{-1} e^{-\tau(H+(1-\lambda)\sigma)} Q | x_f \rangle. \end{aligned} \tag{2.8}$$

Because of the noise, the evolution is ergodic on the energy surface,[2] and thus the real part of the spectrum of H restricted to functions on the energy surface has a gap. Evaluating (2.8) on a basis of eigenfunctions on the energy surface, the 'lowest' right and left eigenfunctions (the ones with the eigenvector $\mu_o(\lambda)$ having the lowest real part) $\langle L_o|$ and $|R_o\rangle$ dominate:

$$\begin{aligned} \langle x_f | e^{-T(H+\lambda\tau\sigma)} | x_i \rangle &\sim \langle x_f | R_o \rangle \langle L_o | x_i \rangle e^{-\tau\mu_o(\lambda)} \\ &= \langle x_f | L_o \rangle \langle R_o | x_i \rangle e^{-\tau\mu_o(1-\lambda)} \end{aligned} \tag{2.9}$$

and using the fact that $H + \lambda\sigma$ and $H + (1-\lambda)\sigma$ have the same spectrum, we conclude that:

$$\lim_{\tau\to\infty} \left\{ \frac{1}{\tau} \ln\langle e^{-\lambda\int_o^T \sigma dt} \rangle - \frac{1}{\tau} \ln\langle e^{-(1-\lambda)\int_o^T \sigma dt} \rangle \right\} = 0 \tag{2.10}$$

[2] In fact, when only half of the equations are noisy, there are some (rather atypical) counterexamples to ergodicity. One can always resort to stochastic equations with noise both on coordinate and momentum equations, see [12].

which is the Laplace transformed version of the Fluctuation Theorem (1.1). Clearly, the whole argument breaks down at strictly zero noise, when the eigenfunctions on the energy surface are no longer smooth, and the overlaps $\langle L_o|x_f\rangle$ and $\langle R_o|x_i\rangle$ may vanish. The problems thus may (and will) arise because we have proven the theorem for

$$\lim_{\epsilon\to 0}\lim_{\tau\to\infty}\frac{1}{\tau}\ln P\left(\int_o^\tau dt\sigma\right) \tag{2.11}$$

and the Gallavotti–Cohen theorem is for:

$$\lim_{\tau\to\infty}\lim_{\epsilon\to 0}\frac{1}{\tau}\ln P\left(\int_o^\tau dt\sigma\right). \tag{2.12}$$

In what follows, we shall analyze the small-noise limit, starting from a simple example.

3 A Simple Example

Consider a particle on a two-dimensional space with toroidal boundary conditions, a constant field E along the direction 1, and an isokinetic thermostat. The example was discussed in Ref. [13]. We write the velocity vector (p_x, p_y) as:

$$p_x = p\cos\theta; \qquad p_y = p\sin\theta. \tag{3.13}$$

There are two stationary situations: parallel velocity $(p_x, p_y) = (p, 0)$, $\theta = 0$ (stable: the attractor) and antiparallel velocity $(p_x, p_y) = (-p, 0)$, $\theta = \pi$ (unstable: the repellor). Equation (2.2) reads, in this case:

$$\begin{aligned}\dot p_x &= E - \frac{Ep_x^2}{p^2} + \eta_x - p_x\frac{p_x\eta_x + p_y\eta_y}{p^2},\\ \dot p_y &= -\frac{Ep_xp_y}{p^2} + \eta_y - p_y\frac{p_x\eta_x + p_y\eta_y}{p^2}.\end{aligned} \tag{3.14}$$

In terms of θ both equations collapse into a single one:

$$\dot\theta = -E\sin\theta + \eta_\theta = -\frac{d}{d\theta}[-E\cos\theta] + \eta_\theta \tag{3.15}$$

where we have defined the angular, isotropic noise: $\eta_\theta \equiv \sin\theta\eta_x - \cos\theta\eta_y$, a white noise whose variance is ϵ. The entropy production rate is $\sigma = E\cos\theta$.

3.1 Weak Noise Limit

Equation (3.15) can be interpreted as a Langevin process in an effective potential $V = -E\cos\theta$, see Fig. 1. The probability evolves according to:

$$\dot P = -\left[\epsilon\frac{\partial^2}{\partial\theta^2} + E\frac{\partial}{\partial\theta}\sin\theta\right]P = -H_{FP}P. \tag{3.16}$$

If we are interested in calculating the average $e^{-\lambda\int\sigma dt}$ in the large-time limit we need to compute the lowest eigenvalue of

$$H_\lambda = H_{FP} + \lambda E\cos\theta \tag{3.17}$$

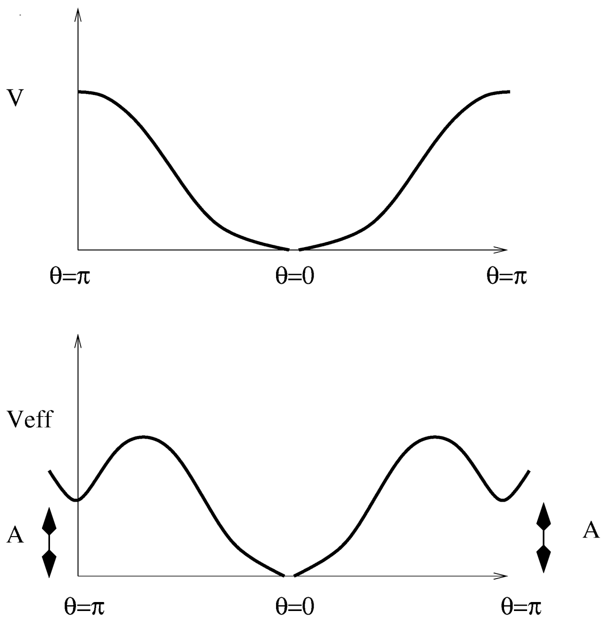

Fig. 1 *Top*: potential in terms of angle. *Bottom*: effective potential associated with the Shroedinger problem (3.18). A is the subdominant bias proportional to $(\lambda - 1/2)$ that favors attractor $(A > 0)$, or repellor $(A < 0)$

which can be taken as usual to the Hermitian form

$$H_\lambda^h = e^{\epsilon E \sin\theta} H_\lambda e^{-\epsilon E \sin\theta} = \frac{1}{\epsilon}\left[-\epsilon^2 \frac{\partial^2}{\partial\theta^2} + \frac{E}{4}\sin^2\theta + \epsilon E\left(\lambda - \frac{1}{2}\right)\cos\theta\right] \tag{3.18}$$

and takes the form of a Shroedinger operator in an effective potential $V_{\text{eff}} = \frac{E}{4}\sin^2\theta + \epsilon(\lambda - \frac{1}{2})E\cos\theta$ depicted in Fig. 1. The average work per unit time at long times is given by the lowest eigenvalue of (3.18). The Gallavotti–Cohen symmetry is manifest making $(\lambda - 1/2) \to -(\lambda - 1/2)$ and $\theta \to \pi - \theta$.

Now, in the limit of small ϵ the situation is as follows: to leading order in ϵ the ground state is doubly degenerate, the eigenfunctions being concentrated in $\theta = 0$ and $\theta = \pi$, respectively – both attractor and repellor are selected by the potential V_{eff} at this order (ϵ^{-1}). It is the next-to-leading order $A = (2\lambda - 1)E\cos\theta$ in V_{eff} (see Fig. 1) that lifts the degeneracy by a quantity of $O(1)$ in ϵ, thus selecting attractor or repellor, depending on the sign of $(\lambda - 1/2)$. Clearly, if $(\lambda - 1/2)$ is of order one, for small ϵ the ground state eigenfunction is peaked in only one well, which one depends on the sign of $(\lambda - 1/2)$, and:

$$\langle e^{-\lambda\tau\bar{\sigma}_\tau}\rangle \sim e^{\tau(|\lambda-1/2|-1/2)E} \tag{3.19}$$

where the last term $1/2$ comes from the zero-point energy of the minimum.

3.2 Attractor, Repellor and Time-Reversal

Let us try now a more direct approach (the techniques can be found in [24]). We can write an expression for the probability of a trajectory in configuration space:

$$P(\text{traj}) = e^{-\frac{1}{\epsilon}\int dt[\frac{1}{4}(\dot\theta + E\sin\theta)^2 - \frac{1}{2}E\epsilon\cos\theta]} \tag{3.20}$$

and we wish to calculate the average

$$\langle e^{-\lambda\sigma}\rangle = \int_{\text{traj}} P(\text{traj}) e^{-\lambda\int dt\cos\theta} = \int_{\text{traj}} e^{-\frac{1}{\epsilon}\int dt[\frac{1}{4}(\dot\theta+E\sin\theta)^2+(\lambda-\frac{1}{2})E\epsilon\cos\theta]}. \tag{3.21}$$

The trajectory weight in (3.21) is composed of two terms, the leading term corresponding to the Lagrangian:

$$L = \frac{1}{4}(\dot\theta + E\sin\theta)^2 = \frac{1}{4}\dot\theta^2 + \frac{1}{4}E^2\sin^2\theta - \frac{1}{2}E\frac{d\cos\theta}{dt} \tag{3.22}$$

and a subleading term proportional to $(\lambda - 1/2)$. The dominant terms must satisfy Lagrange's equations derived from (3.22): these correspond to a problem with inertia in the *inverted potential* $-V_{\text{eff}}$. The solutions are:

$$\dot\theta = \pm\sqrt{4\mathcal{E} + E^2\sin^2\theta} \tag{3.23}$$

where $\mathcal{E}$ is a constant playing the role of an energy associated with the Lagrangian (3.22). It is easy to see that the integral in the exponent (3.20) (the action) can only be finite for large times if the trajectory spends most of the time in attractor or repellor, and this is only possible if $\mathcal{E} = 0$. We thus have that solutions are either solutions of the original noiseless problem:

$$\dot\theta = -E\sin\theta = -\frac{d}{d\theta}[-E\cos\theta] \tag{3.24}$$

which converge to the attractor and have zero action, or 'rare':

$$\dot\theta = E\sin\theta = \frac{d}{d\theta}[-E\cos\theta] \tag{3.25}$$

taking from attractor to repellor, and these have a finite action equal to $E\int d\theta\dot\theta\sin\theta = 2E$. The general trajectory selected by the first term of (3.21) is composed of sejours in the attractor and in the repellor, with rapid descents from repellor to attractor (all three with zero action), plus the 'instanton' trajectories with action $2E$, see Fig. 2.

An important point to note is that the trajectories associated with the Lagrangian (3.22) in the noiseless limit are, as we have mentioned, many more than the trajectories of the original dynamic system in the absence of noise—they include all the non-zero action solutions. Also very important is the fact that the dynamics associated with (3.22) differs from the original

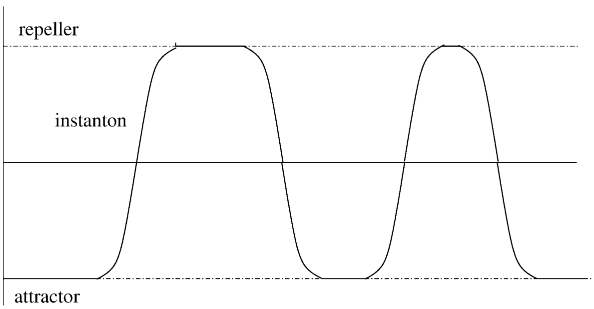

Fig. 2 An 'instanton'–'antiinstanton' gas configuration. The average time spent in the attractor and repellor determines the entropy production

dynamics in that the original stable point $\theta = 0$ is now unstable, due the reversal of the potential. These questions will be relevant in the following sections.

The exact proportion of time spent in attractor $\tau_{\text{attractor}}$ and repellor τ_{repellor} will be determined by the second, subdominant term. This term then gives an action $\sim \epsilon E(\lambda - 1/2)(\tau_{\text{repellor}} - \tau_{\text{attractor}})$. The dominant contribution for the λ-dependent average is given for large times τ by the trajectories that maximize

$$-\ln P(\text{traj}) \sim -\epsilon(\lambda - 1/2)(\tau_{\text{repellor}} - \tau_{\text{attractor}}) + N_{\text{instanton}}\Delta/\epsilon \tag{3.26}$$

where $N_{\text{instanton}}$ is the number of climbings from attractor to repellor. If $\lambda = 0$ the system is unbiased and it will prefer to stay in the attractor. If, on the contrary, we wish to compute the large-deviation function with $\lambda \neq 0$, then the trajectories that dominate will share time in attractor and repellor, so that:

$$(\tau_{\text{repellor}} + \tau_{\text{attractor}})\bar{\sigma}_\tau = \tau_{\text{repellor}}\bar{\sigma}_{\text{repellor}} + \tau_{\text{attractor}}\bar{\sigma}_{\text{attractor}} = (\tau_{\text{attractor}} - \tau_{\text{repellor}})E. \tag{3.27}$$

For the probability of work per unit time $\bar{\sigma}$, at small noise level, we get a linear profile as in Fig. 3 which is in agreement with (3.19).

For infinite times the last term in (3.26) is negligible, but for times such that instanton configurations weigh comparably to the terms proportional to time: $\Delta/\epsilon \sim \ln \tau$ and we shall not see reversals. There are then two possibilities: we may consider a set of initial conditions corresponding to an equilibrium configuration, and get the correct proportion of sejour in attractor and repellor *without needing to perform any activated jumps from attractor to repellor*. This is the transient fluctuation theorem, which holds for zero noise. On the other hand, if we consider any other initial configuration, we need jumps from attractors to repellor and back in order to obtain the right proportion of sejours in the large time limit. Clearly, this second 'stationary' fluctuation theorem needs $\ln \tau \gg \Delta/\epsilon$.

Let us note the close analogy to a one-dimensional Ising model [14] $E = -J\sum_i s_i s_{i+1} - h\sum s_i$ with $\beta J \to \exp(\Delta/\epsilon)$ the wall energy (where Δ is the instanton action), and $\beta h \to \epsilon(\lambda - 1/2)$. Although in our case 'climbing' walls cost, and 'descending' ones do not, the difference is unimportant because the number of instantons and of antiinstantons differ by at most one. Here ϵ enters in the length needed to accommodate many walls, which is necessary in the thermodynamic (large-length) limit. The alternative is to consider a properly weighted set of boundary conditions, so that no walls are necessary: this is analogous to the transient fluctuation theorem.

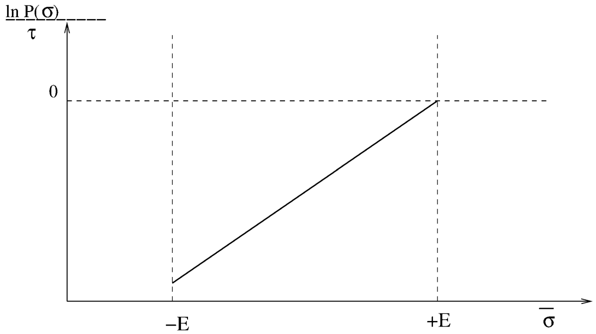

Fig. 3 $\ln P(\sigma)$ vs. σ for the problem of Sect. 2. The profile is linear, and can be seen as the coexistence curve of two phases with free-energy $\pm E$

4 Small Noise Limit, General

We shall now show that the structure of the previous section is quite general, although there are many instanton-antiinstanton trajectories. Consider the evolution equations (2.3) and (2.4). We write this as a path integral

$$P(\mathbf{q},\mathbf{p}) = \int_{[\text{paths}]} e^{-\frac{\Delta_{\text{path}}}{\epsilon}} = \int_{[\text{paths}]} e^{-\frac{1}{\epsilon}\int dt (L(\mathbf{q},\mathbf{p}) - \frac{\epsilon}{2}\sigma)} \tag{4.28}$$

where the paths have the appropriate initial condition in phase-space, and all possible final configurations.

$$\begin{aligned} L &= \hat{q}_i(\dot{q}_i - p_i) + \hat{p}_i(\dot{p}_i + V_i + g_{ij} f_j) + g_{ij}\hat{p}_j g_{il}\hat{p}_l \\ &= \hat{q}_i\dot{q}_i + \hat{p}_i\dot{p}_i - H \end{aligned} \tag{4.29}$$

with

$$H = \hat{q}_i p_i - \hat{p}_i g_{ij} f_j - \hat{p}_i V_i - g_{ij}\hat{p}_j g_{il}\hat{p}_l. \tag{4.30}$$

Here and in what follows we use the notation $V_i = \frac{\partial V}{\partial q_i}$. The second, subdominant term in the exponent comes from the symmetrization (cf. the second of (2.5)):

$$\frac{\partial}{\partial p_i}(\gamma p_i) = \frac{1}{2}\left[\frac{\partial}{\partial p_i}(\gamma p_i) + (\gamma p_i)\frac{\partial}{\partial p_i}\right] - \frac{\sigma}{2} \tag{4.31}$$

where again $\sigma = -(d-1)\gamma$.

We wish to calculate the large-deviation function

$$\langle e^{\lambda \int_o^\tau dt\sigma}\rangle = \int_{[\text{paths}]} e^{-\frac{1}{\epsilon}\int dt[L(\mathbf{q},\mathbf{p}) + \epsilon(\lambda - \frac{1}{2})\sigma]}. \tag{4.32}$$

Just as in the previous section, in the limit of small noise we have to look for saddle-point solutions only taking into account the first term in the exponent. The equations of motion then read:

$$\begin{aligned} \dot{\hat{q}}_i &= -\frac{\partial H}{\partial q_i} = \hat{p}_l g_{lr}\frac{\partial f_r}{\partial q_i} + \hat{p}_l\frac{\partial V_l}{\partial q_i}, \\ \dot{\hat{p}}_i &= -\frac{\partial H}{\partial p_i} = -\hat{q}_i - [\hat{p}_l f_j + \hat{p}_l\hat{p}_j]\frac{\partial g_{lj}}{\partial p_i}, \\ \dot{q}_i &= \frac{\partial H}{\partial \hat{q}_i} = p_i, \\ \dot{p}_i &= \frac{\partial H}{\partial \hat{p}_i} = -g_{ij}(f_j + 2\hat{p}_j) - V_i. \end{aligned} \tag{4.33}$$

Equations (4.33) are Hamilton's equations, and they involve twice as many degrees of freedom as the original system. In fact, the original phase space variables (q_i, p_i) have now become the coordinates, and their conjugate momenta are the new variables $(\hat{q}_i, \hat{p}_i)$. Multiplying the third equation by V_i, the fourth by p_i and adding, we conclude that *all* the solutions in the extended phase-space of (4.33) are still thermostatted to $\sum p_i^2 + V =$ constant.

Let us now rewrite the action in Lagrangian form. We first put $\hat{p}_i g_{ij} f_j = \hat{p}_l g_{li} g_{ij} f_j$ where we have used $g_{li} g_{ij} = g_{lj}$. Next, we separate tangential and normal parts of $\hat{p}_i \dot{p}_i = \hat{p}_l g_{li} \dot{p}_i + \hat{p}_l p_l \dot{p}_i p_i / p^2$. We can complete squares and collect the terms, to get the Lagrangian:

$$L = \hat{q}_i(\dot{q}_i - p_i) - \frac{1}{4}[\dot{p}_i + g_{ij} f_j + V_i]^2 + \left[\frac{1}{2}(\dot{p}_i + g_{ij} f_j + V_i) + g_{ij}\hat{p}_j\right]^2 + \frac{\hat{p}_l p_l}{p^2}(\dot{p}_i p_i + p_i V_i). \tag{4.34}$$

Using the last two equations of motion (4.33), we see that at the saddle point level the last two terms cancel. We hence conclude that the action to leading order equals

$$\Delta = \int dt \tilde{L}(\ddot{\mathbf{q}}, \dot{\mathbf{q}}, \mathbf{q}) = \int dt [\ddot{q}_i + g_{ij} f_j + V_j]^2 \tag{4.35}$$

where we have substituted $p_i = \dot{q}_i$ everywhere in $\tilde{L}$. Note that $\tilde{L}$ has an unusual second time derivative $\ddot{q}_i$. All relevant trajectories are the saddle points of the action (4.35):

$$\frac{d^2}{dt^2}\frac{\partial \tilde{L}}{\partial \ddot{q}_i} - \frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}_i} + \frac{\partial \tilde{L}}{\partial q_i} = 0. \tag{4.36}$$

The solutions are determined by initial and final values of *both* q_i and $\dot{q}_i$. In particular, trajectories of the original system without noise are the ones having zero action:

$$\ddot{q}_i + g_{ij} f_j + V_i = 0. \tag{4.37}$$

Furthermore, in order that the action be finite, we must have that $\ddot{q}_i + g_{ij} f_j + V_i \to 0$ for large t. As an example, in the next section we shall construct the solutions of (4.36) for the Lorentz gas.

In a completely chaotic system, we can construct an expansion in terms of (isolated) trajectories [25–29], and just as before these will commute between attractor(s) and repellor(s)—defined as dynamically stable and unstable solutions of the equations of motion of the original system—where they will spend essentially all the time, since otherwise their action would be infinite. Because the only difference between attractors and repellors is, by definition, their stability, the next-to-leading contribution in the action

$$\Delta_\epsilon = \left(\lambda - \frac{1}{2}\right)\int dt\sigma + \frac{1}{2}\int dt dt' \frac{\delta^2 \tilde{L}}{\delta q_i(t)\delta q_j(t')}\Big|_{\text{traj}} \delta q_i(t)\delta q_j(t') \tag{4.38}$$

will decide the length of the sejours in each for each λ.

In this setting, a necessary condition for the Gallavotti–Cohen theorem (i.e. the Fluctuation relation at zero noise) to hold, is that there be orbits with zero action bridging attractor and repellor.

4.1 Variational Solutions: the Zero-Action Situation

Given a point on the attractor $(q_a, p_a = \dot{q}_a)$ and a point on the repellor $(q_r, p_r = \dot{q}_r)$ and two times t_i and t_f we can obtain the solution going from $(q_a, \dot{q}_a)$ at t_i to $(q_r, \dot{q}_r)$ at t_f as follows. Starting from any curve $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$, contained within the energy surface and having the correct endpoints, its action will be an upper bound for the true, minimal action.

Deforming the curve so as to minimize the action (4.35) we reach a solution. The arrival time may become a variational parameter itself.

If the lowest action $\Delta_{\min}$ associated with any trajectory joining any two points in the attractor and repellor is finite, we know that the system will necessarily need a time at least $\tau \sim e^{\Delta_{\min}/\epsilon}$ to explore them. This can be simply seen by looking at the sum over paths (4.32) as a partition function in a space of trajectories at inverse temperature ϵ, with $\Delta_{\min}$ its 'ground state energy'.

Clearly, for the Gallavotti–Cohen theorem to apply, a necessary condition is that $\Delta_{\min} = 0$. For this to happen, attractor and repellor have to be interwoven so that just advancing from $(q_a, \dot{q}_a)$ along a zero-action trajectory satisfying (4.37) we reach a point that is as close as we wish to some $(q_r, \dot{q}_r)$, providing a zero-action 'bridge'. However, the condition $\Delta_{\min} = 0$ only implies that the time for bridging attractor and repellor grows in a subexponential — for example power law $\tau \sim \epsilon^{-\mu}$ — way, and is not necessarily finite as $\epsilon \to 0$. One can imagine such a situation arising in a case in which attractor and repellor have closures of dimension d_1, but these intersect in a manifold of dimension $d_2 < d_1$. Without noise a trajectory might have vanishing probability of leaving the attractor and entering the repellor, since the time spent in the overlapping region would be negligible; the addition of noise will coarsen both attractor, repellor and intersection by introducing a width, hence making the passage more likely.

5 Instantons for the Lorentz Gas

We shall consider here a Lorentz gas [30], a system with a single thermostatted particle under the action of an electric field as in Sect. 2, but with in addition obstacles where the particle bounces, introducing chaos The Lagrangian (3.22) reads

$$
\begin{aligned}
\tilde{L} &= \frac{1}{4}(\dot{\theta} + E\sin\theta)^2 = p_\theta \dot{\theta} - \mathcal{H}, \\
\mathcal{H} &= \mathcal{E} = 2p_\theta(p_\theta - E\sin\theta)
\end{aligned}
\tag{5.39}
$$

with $\mathcal{E}$ defined in (3.23) and

$$
p_\theta = \frac{1}{2}(\dot{\theta} + E\sin\theta). \tag{5.40}
$$

The motion is punctuated by bounces on the wall. Between bounces, the trajectories satisfy (3.23):

$$
\begin{aligned}
d\theta &= \pm\sqrt{4\mathcal{E} + E^2\sin^2\theta}\,dt, \\
dx &= \cos\theta dt = \frac{\cos\theta}{\dot{\theta}} d\theta = \pm\frac{\cos\theta}{\sqrt{E^2\sin^2\theta + 4\mathcal{E}}} d\theta, \\
dy &= \sin\theta dt = \frac{\sin\theta}{\dot{\theta}} d\theta = \pm\frac{\sin\theta}{\sqrt{E^2\sin^2\theta + 4\mathcal{E}}} d\theta
\end{aligned}
\tag{5.41}
$$

which yields the parametric equations for the trajectory segment. One also has that

$$
\begin{aligned}
\dot{p}_\theta &= -2E\cos\theta p_\theta = 2E p_\theta \dot{x}, \\
p_\theta(t) &= p_\theta(t_o) e^{2Ex|_o^t}
\end{aligned}
\tag{5.42}
$$

where t_o is the initial time of the segment. The action for the trajectory segment reads:

$$\Delta = \int dt p_\theta^2(t) = p_\theta^2(t_o) \int dt e^{4Ex|_o^t} \tag{5.43}$$

where $x|_o^t = \int dt \dot{x}$ is the total distance along x, without subtracting the windings if the space is periodic. It is positive if the particle moved downhill, and negative if it moves contrary to the force.

Next, we need to consider bounces. Because we are solving the instanton dynamics, rather than the original one, we could wonder what happens during a bounce. To make this point clear, in the Appendix we compute a bounce by considering the wall as a limit of continuous potentials. The result is that: (i) the reflection law holds, (ii) the contribution of the bounce to the action vanishes in the hard wall limit, and (iii) p_θ does not change from just before to just after the bounce. Instead, $\mathcal{E}$ is not conserved in bounces, since:

$$\mathcal{E}|_{\text{before}}^{\text{after}} = 2p_\theta E \sin\theta|_{\text{before}}^{\text{after}} \tag{5.44}$$

which is nonzero in general, except when $p_\theta = 0$.

Segments of trajectories belonging to the attractor and repellor of the original noiseless dynamics have zero action and $p_\theta(t) = \mathcal{E} = 0$: this is preserved by both trajectories and bounces at all times (cf. (5.43) and (5.44)).

Trajectories have a starting point $(x, y, \theta, \dot{\theta})$. Thanks to bounces, they move uphill and downhill with respect to the field. Those that diffuse uphill have a smaller and smaller value of p_θ (see (5.43)). Furthermore, as $p_\theta \to 0$ also $\mathcal{E} \to 0$ (although $\mathcal{E}$ can still increase somewhat in a bounce). Hence, a trajectory that had bounces allowing it to diffuse a long way uphill becomes more and more closely a zero-action trajectory. Moreover, it is (or rather, it shadows) a trajectory belonging to the *repellor*, since trajectories in the attractor move on average downhill.

Suppose next that the starting point is a very small perturbation with $p_\theta \neq 0$ of a point on the attractor. At the beginning the point will continue to move on average downhill. However, this implies that the value of p_θ grows exponentially (see (5.43)). In other words, in the full dynamics with instantons, true attractor trajectories are *unstable*, a general fact already mentioned in the previous sections. After some time, these trajectories will bounce uphill and downhill. A few of these perturbations will diffuse a large amount uphill, and as described above those will then be very close to the repellor. Hence, we have identified the instantons interpolating between attractor and repellor as the (rare) trajectories that have initial conditions close to the attractor and such that they keep diffusing on average upwards. Their action is just the integral of the exponential of the uphill distance—a finite quantity.

As is typical with instantons, we need some sort of 'shooting' method to find those solutions that actually end up in (or shadowing an orbit in) the repellor, otherwise a generic solution will just evolve away. Secondly, we see that there are many of these solutions, and not essentially one as in the case without obstacles.

6 Conclusions

We have recasted the problem of the deterministic Fluctuation Theorem as the vanishing noise limit of stochastic one. Large-deviation functions in this limit need not coincide with the zero-noise result, but one can always postulate that this is so in a practical case, thus making an alternative Chaotic Hypothesis [15]: Considering the energy-conserving noise described above (and more generally noise respecting all the constants of motion of the dynamics), *a reversible many-particle system in a stationary state can be regarded as stochastically stable for the purposes of computing probabilities over trajectories of macroscopic observables.* The advantage of this formulation is that it is immediately testable in a numerical (and perhaps also in a real) experiment.

We have argued here that if there are orbits with zero action that bridge attractor and repellor, then the time needed in order that the Fluctuation Relation is satisfied is subexponential in the noise $\epsilon \ln \tau \to 0$. We have, however, fallen short of proving the Gallavotti–Cohen theorem for that case, since we argue that subexponential yet slowly divergent growth of τ as $\epsilon \to 0$ is likely to arise, perhaps in systems in which the closures of attractor and repellor are simultaneously dense in a manifold of low dimension. It seems an interesting question to obtain estimates of passage times in terms of the noise (or temperature) in such cases.

Acknowledgements I wish to thank G. Gallavotti, D. Ruelle and F. Zamponi for discussions and suggestions.

Appendix Instantons and Hard Walls

To model the bounce, we consider walls as regions with a constant repulsive force (perpendicular to the wall), and then take the limit of large force. We shall rotate the axis to a new axes (x', y') so that outside the wall region the field is (E_1, E_2) and inside the wall region the field is $(E^w, 0)$, $E^w \to \infty$ as shown Fig. 4.

In general, in a region of fields (E_1, E_2) the Lagrangian reads

$$\begin{aligned} L &= \frac{1}{4}(\dot{\theta} + E_1 \sin\theta - E_2 \cos\theta)^2, \\ \mathcal{H} &= 2p_\theta(p_\theta - E_1 \sin\theta + E_2 \cos\theta) \end{aligned} \tag{7.45}$$

with

$$p_\theta = \frac{1}{2}(\dot{\theta} + E_1 \sin\theta - E_2 \cos\theta). \tag{7.46}$$

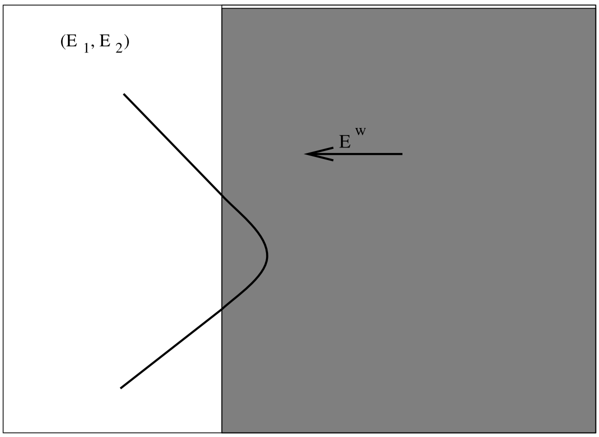

Fig. 4 Bounce trajectory in a wall region

The forces E_i jump when the particle crosses the wall. Across this jump, θ is continuous, but $\dot\theta$ jumps. To calculate this, we write the equations of motion:

$$2\frac{d}{dt}(\dot\theta + E_1\sin\theta - E_2\cos\theta) - 2(\dot\theta + E_1\sin\theta - E_2\cos\theta)(E_1\cos\theta + E_2\sin\theta) = 0 \quad (7.47)$$

which implies that in regions of constant fields:

$$\begin{aligned}\frac{d}{dt}\ln(\dot\theta + E_1\sin\theta - E_2\cos\theta) &= 2(E_1\dot x + E_2\dot y),\\ \ln|\dot\theta + E_1\sin\theta - E_2\cos\theta|_o^t &= 2(E_1x + E_2y)|_o^t = 2\mathbf{x}\bullet\mathbf{E}|_o^t.\end{aligned} \quad (7.48)$$

On entering the wall region, (7.47) implies that

$$p_\theta(t) = \frac{1}{2}(\dot\theta + E_1\sin\theta - E_2\cos\theta) \quad (7.49)$$

is continuous across discontinuities of the potential. Instead, the numerical value of $H = \mathcal{E}$ jumps when the trajectory bounces, except if $p_\theta = 0$.

Within the wall region:

$$L = \frac{1}{4}(\dot\theta - E_w\sin\theta)^2. \quad (7.50)$$

The conserved quantity is:

$$\mathcal{E} = \dot\theta^2 - E_w^2\sin^2\theta \quad (7.51)$$

and we have:

$$\dot\theta = \pm\sqrt{E_w^2\sin^2\theta + 4\mathcal{E}}. \quad (7.52)$$

This is shown in Fig. 5, and it is clear that the reflection law will hold, since the trajectory inside the wall region is symmetric. In terms of the coordinates normal and parallel to the

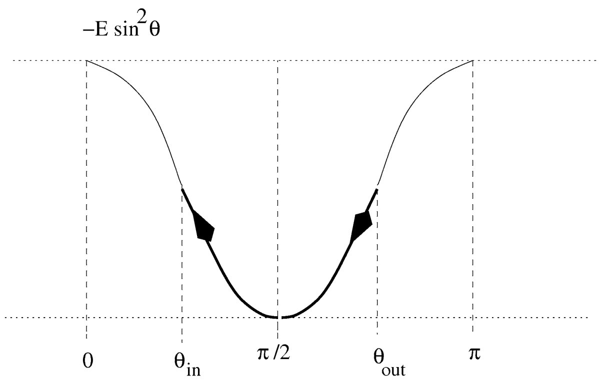

Fig. 5 The angle evolves in a potential $\sim -E\sin^2\theta$, entrance and exit angle a symmetrically disposed, since the distance along x moved is the integral of $\cos\theta(t)$, which is exactly zero if $\pi/2 - \theta_{\rm in} = \theta_{\rm out} - \pi/2$

wall:

$$
\begin{aligned}
dx' &= \cos\theta dt = \frac{\cos\theta}{\dot\theta}d\theta = \frac{\cos\theta}{\sqrt{E_w^2\sin^2\theta + 4\mathcal{E}}}d\theta,\\
dy' &= \sin\theta dt = \frac{\sin\theta}{\dot\theta}d\theta = \frac{\sin\theta}{\sqrt{E_w^2\sin^2\theta + 4\mathcal{E}}}d\theta.
\end{aligned}
\tag{7.53}
$$

We have $\dot{x}' = \cos\theta$ and (7.48) reads:

$$
(\dot\theta - E_w\sin\theta)_t = (\dot\theta - E_w\sin\theta)_{t=0}e^{-2E_w x'}. \tag{7.54}
$$

It is then clear that, in the limit of strong E_w, the time spent in the wall region tends to zero, while the integrand in the action $e^{4E_w x'}$ is itself of order one (since the penetration $x' \sim O(1/E_w)$). We conclude that the action during the bounce tends to zero as $|E_w| \to \infty$. Note that the argument is valid for quite general bounces.

References

1. Evans, D.J., Cohen, E.G.D., Morriss, G.P.: Phys. Rev. Lett. **71**, 2401 (1993)
2. Evans, D.J., Searles, D.J.: Phys. Rev. E **50**, 1645 (1994)
3. Gallavotti, G., Cohen, E.G.D.: Phys. Rev. Lett. **74**, 2694–2697 (1995)
4. Gallavotti, G., Cohen, E.G.D.: J. Stat. Phys. **80**, 931–970 (1995)
5. Ritort, F.: Séminaire Poincare, 6 December 2003, p. 63, cond-mat/0401311
6. Evans, D., Searles, D.: Adv. Phys. **51**, 1529 (2002)
7. Gallavotti, G.: Statistical Mechanics A Short Treatise. Texts and Monographs in Physics. Springer, Berlin (1999), available in http://ipparco.roma1.infn.it/pagine/libri.html
8. Kurchan, J.: cond-mat/0511073
9. Zamponi, F.: cond-mat/0612019
10. Kurchan, J.: J. Phys. A: Math. Gen. **31**, 3719 (1998)
11. Lebowitz, J.L., Spohn, H.: J. Stat. Phys. **95**, 333 (1999)
12. Cepas, O., Kurchan, J.: Eur. Phys. J. B **2**(2), 221 (1998)
13. Cohen, E.G.D., Gallavotti, G.: J. Stat. Phys. **96**, 1343 (1999)
14. Bonetto, F., Gallavotti, G.: Commun. Math. Phys. **189**, 263 (1997)
15. Gallavotti, G.: Phys. Rev. Lett. **77**, 4334 (1996)
16. Sinai, Ya.G.: Ann. Probab. **17**(3), 833 (1989)
17. Young, L.S.: J. Stat. Phys. **108**, 733 (2002)
18. Cowieson, W., Young, L.S.: **25** 1115 (2005)
19. Espanol, P.: Europhys. Lett. **40**(6), 631 (1997)
20. Giardina, C., Kurchan, J.: J. Stat. Mech. **5**, 05009 (2005)
21. Fritz, J., Nagy, K., Olla, S.: J. Stat. Phys. **122**(3), 399 (2006)
22. Tailleur, J., Kurchan, J.: Nature Phys. **3**, 203 (2007)
23. Risken, H.: The Fokker–Planck Equation. Springer, Berlin (1984)
24. Zinn-Justin, J.: Quantum field theory and critical phenomena. Oxford University Press, Oxford (1996)
25. Cvitanovic, P. et al.: Chaos, classical and quantum. http://chaosbook.org/
26. Creagh, S.C., Whelan, N.D.: Phys. Rev. Lett. **82**, 5237 (1999)
27. Creagh, S.C., Whelan, N.D.: Phys. Rev. Lett. **84**, 4084 (2000)
28. Creagh, S.C.: Nonlinearity **17**, 1261 (2004)
29. Smith, G.C., Creagh, S.C.: J. Phys. A. **39**, 8283 (2006)
30. Bonetto, F., Gallavotti, G., Garrido, P.L.: Physica D **105**, 226 (1997)