Fluctuation theorem for stochastic dynamics

1998 J. Phys. A: Math. Gen. 31 3719

(http://iopscience.iop.org/0305-4470/31/16/003)

J. Phys. A: Math. Gen. **31** (1998) 3719–3729. Printed in the UK PII: S0305-4470(98)88861-6

Fluctuation theorem for stochastic dynamics

Jorge Kurchan†

Laboratoire de Physique Théorique de l' Ecole Normale Supérieure de Lyon, Allée d'Italie, Lyon, France

Abstract. The fluctuation theorem of Gallavotti and Cohen holds for finite systems undergoing Langevin dynamics. In such a context all non-trivial ergodic theory issues are bypassed, and the theorem takes a particularly simple form. As a particular case, we obtain a nonlinear fluctuation–dissipation theorem valid for equilibrium systems perturbed by arbitrarily strong fields.

1. Introduction

The fluctuation theorem (FT) concerns the distribution of entropy production over long time intervals. It states that the ratio of the probabilites of having a given entropy production σ_t averaged over a (large) time interval to that of having $(-\sigma_t)$ is $e^{t\sigma_t}$. It was stated and proved in [1] (in what follows GC) for thermostated Hamiltonian systems driven by external forces, under certain 'chaoticity' assumptions for the dynamics [2].

The relevance of this apparently bizarre result became clear when it was shown [3] that it reduces to the fluctuation–dissipation theorem and the Onsager relations in the limit of zero power input (i.e. in equilibrium).

In this paper we show how to derive the GC fluctuation theorem for systems undergoing Langevin dynamics. The purpose of the exercise is threefold.

• The Langevin dynamics is trivially 'ergodic', in the sense that for purely conservative forces, bounded systems with finitely many degrees of freedom reach the Gibbs–Boltzmann distribution irrespective of the form of the interaction.

For this reason, one can make a proof of the FT that is as simple as it can possibly be, having bypassed every non-trivial question of ergodic theory. For example, the stationary states in this context are the zero eigenvalues of a certain (non-Hermitian) Schrödinger-like operator, and are of a rather familiar nature.

Because of this extreme simplicity, one can use the Langevin systems as an heuristic tool to find new results, and try to see whether they hold for more general thermostated Hamiltonian systems.

• In order to prove the FT in GC, in addition to making some assumptions regarding the 'chaoticity' of the models, some conditions of boundedness and finiteness of the number of degrees of freedom were also required. Here, because all 'ergodicity' aspects have been done by hand, one can study how the FT can be violated *in problems with (and as a consequence of) having infinitely many degrees of freedom*. Hence, we have a formalism that allows us to isolate the violations due to 'complexity' (i.e. non-trivial features specific to the large-size limit) from violations due to the possible non-applicability of the chaotic hypothesis.

† Also at: Institute for Theoretical Physics, University of California at Santa Barbara, CA 93106, USA. E-mail address: Jorge.Kurchan@enslapp.ens-lyon.fr

There are situations for which in the limit of zero forcing the fluctuation–dissipation theorem is violated in a stationary state *of an infinite system* [4]. Because the FT reduces to the fluctuation–dissipation theorem in the limit of zero non-conservative forces [3], this is a particular instance in which the FT is violated.

• There are many interesting driven systems that can be well represented by Langevin problems, e.g. Burgers–KPZ, phase separation under shear, turbulence, etc.

This paper is organized as follows. In section 2 we review the equations describing the evolution of the probability distribution for a Langevin process. We do this for the case with inertial (second time-derivative) term, corresponding to the Kramers equations. We discuss the (crucial) property of detailed balance, and how it is modified in the presence of non-conservative forces.

In section 3 we show that the modified detailed balance propery leads to the FT. We also present the *limit theorem* for the entropy production [5], as applied to the stochastic case. We then follow the steps of Gallavotti [3] in showing how the FT reduces to the Green–Kubo formula in the purely conservative limit.

We construct a particular form of the FT corresponding to purely conservative driving forces which yields a nonlinear generalization of the usual fluctuation–dissipation theorem.

In section 4 we show that the arguments of the preceding sections can be applied to a Langevin process without inertial term, corresponding to the Fokker–Planck equation. We also present a direct proof of the nonlinear fluctuation–dissipation theorem for this case.

In the conclusions we discuss the possible violations of the FT equality in systems with infinitely many degrees of freedom.

2. Langevin and Kramers equations

We will consider the Langevin dynamics

$$m\ddot{x}_i + \gamma\dot{x}_i + \partial_{x_i}U(\boldsymbol{x}) + f_i = \Gamma_i \tag{2.1}$$

where $i = 1, \ldots, N$. Γ_i is a delta-correlated white noise with variance $2\gamma T$.

The f_i are velocity-independent forces that do not necessarily derive from a potential.

We will not deal here with the limits $\gamma = 0$ (Hamiltonian dynamics), $T = 0$ (noiseless dynamics) and $N \to \infty$, for reasons that will become clear.

We shall first treat, in detail, the case with inertia $m \neq 0$, leading to Kramer's equation, and later indicate how to treat the case in which $m = 0$ which leads to Fokker–Planck equation.

2.1. Kramers equation

If $m \neq 0$, the probability distribution at time t for the process (2.1) is expressed in terms of the phase-space variables x_i, v_i and is given by

$$P(\boldsymbol{x}, \boldsymbol{v}, t) = \mathrm{e}^{-tH} P(\boldsymbol{x}, \boldsymbol{v}, 0) \tag{2.2}$$

where H is the Kramers operator [6]:

$$H = \partial_{x_i} v_i - \frac{1}{m}\partial_{v_i}\left(\gamma v_i + (\partial_{x_i}U(\boldsymbol{x})) + f_i + \gamma\frac{T}{m}\partial_{v_i}\right). \tag{2.3}$$

We find it convenient to use bracket and operator notation:

$$\begin{aligned} &P(\boldsymbol{x}, \boldsymbol{v}, t) = \langle x, v|\phi(t)\rangle \\ &|\phi(t)\rangle = \mathrm{e}^{-tH}|\phi(0)\rangle. \end{aligned} \tag{2.4}$$

Expectation values of a variable $O(x, v)$ are obtained as:

$$\langle O(t)\rangle = \langle -|O|\phi(t)\rangle \tag{2.5}$$

where we have defined the flat distribution:

$$\langle -|x, v\rangle = 1 \qquad \forall x, v. \tag{2.6}$$

Introducing the Hermitian operators $\hat{p}_{x_i}$, $\hat{p}_{v_i}$ as:

$$\begin{aligned} \langle x, v|\hat{p}_{x_i}|\phi(t)\rangle &= -\mathrm{i}\frac{\partial}{\partial x_i}\langle x, v|\phi(t)\rangle \\ \langle x, v|\hat{p}_{v_i}|\phi(t)\rangle &= -\mathrm{i}\frac{\partial}{\partial v_i}\langle x, v|\phi(t)\rangle \end{aligned} \tag{2.7}$$

the Kramers Hamiltonian reads:

$$H = \mathrm{i}v_i\hat{p}_{x_i} - \frac{\mathrm{i}}{m}\hat{p}_{v_i}\left(\gamma v_i + \partial_{x_i}U(\boldsymbol{x}) + f_i + \gamma\frac{\mathrm{i}T}{m}\hat{p}_{v_i}\right). \tag{2.8}$$

One can explicitate 'conservative' and 'forced' parts of H (the latter being non-conservative if f_i do not derive from a potential):

$$H = H^c - \frac{\mathrm{i}}{m}\hat{p}_{v_i}f_i. \tag{2.9}$$

Probability conservation is guaranteed by

$$\langle -|H = 0. \tag{2.10}$$

A stationary state satisfies:

$$H|\,\mathrm{stat}\rangle = 0 \qquad \langle -|\,\mathrm{stat}\rangle = 1. \tag{2.11}$$

2.2. Detailed balance and time reversibility

The evolution of the system satisfies in the absence of non-conservative forces a form of detailed balance:

$$\langle x', v'|\mathrm{e}^{-tH^c}|x, v\rangle\mathrm{e}^{-\beta E_K(\boldsymbol{x},\boldsymbol{v})} = \langle x, -v|\mathrm{e}^{-tH^c}|x', -v'\rangle\mathrm{e}^{-\beta E_K(\boldsymbol{x}',-\boldsymbol{v}')} \tag{2.12}$$

where the total energy is $E_K(\boldsymbol{x}, \boldsymbol{v}) = \frac{1}{2}\sum_i mv_i^2 + U(\boldsymbol{x})$. This leads to a symmetry property, which in operator notation reads:

$$Q_K^{-1}H^cQ_K = H^{c\dagger} \tag{2.13}$$

where the operator Q_K is defined by:

$$Q_K|x, v\rangle \equiv \mathrm{e}^{-\beta E_K(\boldsymbol{x},\boldsymbol{v})}|x, -v\rangle. \tag{2.14}$$

In the presence of arbitrary forces f_i, equation (2.13) is modified to:

$$Q_K^{-1}HQ_K = H^\dagger + \beta f_i v_i \equiv H^\dagger - S^\dagger. \tag{2.15}$$

The operator S is the power exerted on the system divided by the temperature

$$S^\dagger = -\beta\boldsymbol{f}\cdot\boldsymbol{v} \tag{2.16}$$

and this is the entropy production in the case of a stationary system. We also have:

$$Q_K^{-1}SQ_K = -S^\dagger. \tag{2.17}$$

Clearly, in the particular case in which the forces f_i derive from a potential $f_i = \frac{\partial A}{\partial x_i}$:

$$S = -\beta f_i v_i = -\beta\frac{\mathrm{d}A}{\mathrm{d}t} \tag{2.18}$$

whose average value at stationarity is zero.

A non-increasing $\mathcal{H}$-function may be defined as [7]

$$\mathcal{H}(t) = \int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\,P(\boldsymbol{x},\boldsymbol{v})(T\ln P(\boldsymbol{x},\boldsymbol{v}) + E(\boldsymbol{x},\boldsymbol{v})) \tag{2.19}$$

and may be interpreted as a 'generalized free energy'. Writing $P_{\mathrm{stat}}(\boldsymbol{x},\boldsymbol{v}) \equiv \langle x, v|\,\mathrm{stat}\rangle$, we have that $\dot{\mathcal{H}} = 0$ implies that

$$-\langle -|\boldsymbol{f}\cdot\boldsymbol{v}|\,\mathrm{stat}\rangle = \gamma\sum_i\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}\,\frac{(mv_iP_{\mathrm{stat}} + T\partial_{v_i}P_{\mathrm{stat}})^2}{m^2P_{\mathrm{stat}}} \equiv T\langle\sigma\rangle \geqslant 0. \tag{2.20}$$

The averaged entropy production at stationarity $\langle\sigma\rangle$ is non-negative†.

Let us now write a (Green–Kubo) fluctuation–dissipation theorem for a *purely conservative* system perturbed by an arbitrary small force field $-h(t)\boldsymbol{f}(\boldsymbol{x})$. The current operator is

$$J \equiv \boldsymbol{f}\cdot\boldsymbol{v} \tag{2.21}$$

and $S = -h\beta J$. Linear response theory implies, for any observable $O(x,v)$ (cf (2.4) and (2.8)):

$$\frac{\delta\langle O(t)\rangle}{\delta h(t')}|_{h=0} = -\frac{\mathrm{i}}{m}\langle -|O\mathrm{e}^{-(t-t')H^c}\hat{p}_{v_i}f_i\mathrm{e}^{-t'H^c}|\,\mathrm{init}\rangle. \tag{2.22}$$

In equilibrium $|\,\mathrm{init}\rangle = |\,\mathrm{GB}\rangle$, the Gibbs–Boltzmann distribution and:

$$-\frac{\mathrm{i}}{m}\hat{p}_{v_i}|\,\mathrm{GB}\rangle = \beta v_i|\,\mathrm{GB}\rangle \qquad \mathrm{e}^{-t'H}|\,\mathrm{GB}\rangle = |\,\mathrm{GB}\rangle \tag{2.23}$$

which introduced in (2.22) implies the fluctuation–dissipation theorem:

$$\left.\frac{\delta\langle O(t)\rangle}{\delta h(t')}\right|_{h=0} = \beta\langle -|O\mathrm{e}^{-(t-t')H^c}J|\,\mathrm{GB}\rangle = \beta\langle O(t)J(t')\rangle\theta(t-t'). \tag{2.24}$$

2.3. Power and entropy production distribution

Consider the power $T\sigma_t$ done by the forces f_i in a time t for a given path in phase space $(\boldsymbol{x}(t'),\boldsymbol{v}(t'))$:

$$T\sigma_t \equiv \int_0^t \boldsymbol{f}(\boldsymbol{x})(t')\cdot\boldsymbol{v}(t')\,\mathrm{d}t'. \tag{2.25}$$

We wish to study the distribution $\Pi_t(\sigma_t)$ of σ_t for different noise realizations. Let us show that:

$$\Pi_t(\sigma_t) = t\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty}\mathrm{d}\lambda\langle -|\mathrm{e}^{-t(H+\lambda S)}|\,\mathrm{init}\rangle\mathrm{e}^{t\lambda\sigma_t}. \tag{2.26}$$

This is most easily seen in the path-integral representation. Denoting $\mathcal{S}$ the action associated with H along a path we have

$$\begin{aligned}\Pi_t(\sigma_t) &= t\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty}\mathrm{d}\lambda\langle -|\mathrm{e}^{-t(H+\lambda\beta\boldsymbol{f}\cdot\boldsymbol{v}-\lambda\sigma_t)}|\,\mathrm{init}\rangle\\ &= t\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty}\mathrm{d}\lambda\int\mathcal{D}(\mathrm{paths})\mathrm{e}^{-\mathcal{S}(\mathrm{path})-\lambda(\beta\int_0^t\boldsymbol{f}\cdot\boldsymbol{v}(t')\mathrm{d}t'-t\langle\sigma\rangle p)}\\ &= t\int\mathcal{D}(\mathrm{paths})\mathrm{e}^{-\mathcal{S}(\mathrm{path})}\delta\left(\beta\int_0^t\boldsymbol{f}\cdot\boldsymbol{v}(t')\,\mathrm{d}t' - t\sigma_t\right).\end{aligned} \tag{2.27}$$

† In the pure Hamiltonian $\gamma = 0$ case $\langle\sigma\rangle = 0$ at all times, a consequence of Liouville's theorem.

Bearing in mind that the factor $e^{-\mathcal{S}(\text{path})}$ is precisely the probability of each path, the last equality implies (2.26).

In section 3 we shall assume that there is a non-zero average $\langle\sigma\rangle$, and following GC we shall work with the adimensional variable

$$p = \frac{\sigma_t}{\langle\sigma\rangle} \tag{2.28}$$

and study the distribution of p given by $\pi_t(p) \equiv \langle\sigma\rangle \Pi_t(\langle\sigma\rangle p)$.

3. Modified detailed balance and the fluctuation theorem

3.1. A first version of the fluctuation theorem

The FT follows from the modified form of detailed balance equations (2.15) and (2.17). These two imply, for any λ:

$$Q^{-1}(H + \lambda S)Q = H^\dagger - (1 + \lambda)S^\dagger = [H - (1 + \lambda^*)S]^\dagger \tag{3.1}$$

so that $(H+\lambda S)$ and $(H-(1+\lambda^*)S)$ have conjugate spectra. This relation has consequences for the distribution of power. Let us see what the implications are for $\pi_t(p)$, starting from an initial distribution $|\,\text{init}\rangle$.

$$\begin{aligned}
\pi_t(p) &= t\langle\sigma\rangle \int_{-i\infty}^{+i\infty} d\lambda \langle -|e^{-t(H+\lambda S)}|\,\text{init}\rangle e^{t\lambda\langle\sigma\rangle p} \\
&= t\langle\sigma\rangle \int_{-i\infty}^{+i\infty} d\lambda \langle -|\,Q e^{-t[H-(1+\lambda^*)S]^\dagger} Q^{-1}|\,\text{init}\rangle e^{t\lambda\langle\sigma\rangle p} \\
&= t\langle\sigma\rangle \int_{-i\infty}^{+i\infty} d\lambda \langle \text{init}\,|Q^{-1\dagger} e^{-t[H-(1+\lambda^*)S]} Q^\dagger|-\rangle^* e^{t\lambda\langle\sigma\rangle p}.
\end{aligned} \tag{3.2}$$

Using the fact that H, S and Q are real:

$$\pi_t(p) = t\langle\sigma\rangle \int_{-i\infty}^{+i\infty} d\lambda \langle \text{init}\,|Q^{-1\dagger} e^{-t[H-(1+\lambda)S]} Q^\dagger|-\rangle e^{t\lambda\langle\sigma\rangle p}. \tag{3.3}$$

Making $\lambda \to -1-\lambda$:

$$\pi_t(p) = t\langle\sigma\rangle e^{-t\langle\sigma\rangle p} \int_{-1-i\infty}^{-1+i\infty} d\lambda\, \langle \text{init}\,|Q^{-1\dagger} e^{-t[H+\lambda S]} Q^\dagger|-\rangle e^{t\lambda\langle\sigma\rangle(-p)}. \tag{3.4}$$

Now, $\langle A|e^{-t(H+\lambda S)}|B\rangle$ is for given $|A\rangle$, $|B\rangle$ an analytical function of λ, and we can deform the contour in the integral of the last line to $\int_{-i\infty}^{+i\infty}$.

Consider first the case in which we start from a Gibbs–Boltzmann distribution $|\,\text{GB}\rangle$ [8], which *need not be* a stationary distribution in the presence of the non-conservative forces†. We then have:

$$\langle \text{GB}\,|Q^{-1\dagger} \propto \langle -| \qquad Q^\dagger|-\rangle \propto |\,\text{GB}\rangle \tag{3.5}$$

and equation (3.4) implies, *for all times*:

$$\pi_t^{\text{GB}}(p) = e^{-t\langle\sigma\rangle p}\pi_t^{\text{GB}}(-p) \tag{3.6}$$

$$-\langle\sigma\rangle p = \frac{\ln(\pi_t^{\text{GB}}(p)) - \ln(\pi_t^{\text{GB}}(-p))}{t}. \tag{3.7}$$

† This situation can be experimentally created by switching on the non-conservative forces at $t = 0$ on an equilibrated conservative system.

Here we have added the superscript GB to indicate the initial condition.

What about other initial conditions? Already at this stage it becomes intuitive that if the system is such that any initial condition (in particular the Gibbs–Boltzmann distribution) evolves *in finite time* $\sim\tau_{\rm erg}$ to the same stationary distribution $|\,{\rm stat}\rangle$, then we should have a result like (3.7) (but only for $t \gg \tau_{\rm erg}$) irrespective of the initial condition. Note that, surprisingly, the stationary distribution does not appear to play a special role here, but the Gibbs–Boltzmann distribution does! However, this statement has to be qualified if we wish to identify the power made by the non-conservative forces as an 'entropy production', something we can do only in the stationary regime.

3.2. Long-time distributions of $\pi(p)$.

Let us make the above remarks more precise. If, under certain assumptions we have that for long times there is a single limiting function $\zeta(p)$ [5] such that *irrespective of the initial conditions*

$$\lim_{t\to\infty}\frac{1}{\langle\sigma\rangle t}\ln(\pi_t) \to -\zeta(p) \tag{3.8}$$

then the FT (3.7) will hold for long times for any initial condition, and will read:

$$\zeta(p) - \zeta(-p) = \langle\sigma\rangle p. \tag{3.9}$$

In order to derive (3.8) we shall make the two following assumptions.

(1) The lowest (zero) eigenvalue of H is non-degenerate.

(2) The initial state has a non-zero overlap with the left eigenvector of H; $\langle{\rm stat}\,|\,{\rm init}\rangle \neq 0$.

Any of these assumptions may fail to hold in unbounded systems or in systems with infinitely many degrees of freedom. Indeed, conservative systems with slow dynamics such as glasses and coarsening are known to have a gap-less spectrum of the Fokker–Planck operators (the gap goes to zero with the system's number of degrees of freedom).

Furthermore, the gap vanishes in the purely Hamiltonian $\gamma = 0$ limit, as there are many long-lived phase-space distributions in that case (e.g. invariant tori, etc), as well as in the $T = 0$ case. Note that H loses the second derivative in these cases. This is the main reason why we only consider $\gamma > 0, T > 0$ here.

We proceed as follows. Introducing the right and left eigenvectors:

$$(H + \lambda S)|\psi_i^R(\lambda)\rangle = \mu_i(\lambda)|\psi_i^R(\lambda)\rangle \qquad \langle\psi_i^L(\lambda)|(H + \lambda S) = \mu_i(\lambda)\langle\psi_i^L(\lambda)| \tag{3.10}$$

we have:

$$\pi_t^{|\,{\rm init}\rangle}(\sigma_t) = t\langle\sigma\rangle\int_{-{\rm i}\infty}^{+{\rm i}\infty}{\rm d}\lambda\sum_i\langle-|\psi_i^R(\lambda)\rangle\langle\psi_i^L(\lambda)|\,{\rm init}\rangle{\rm e}^{-t(\mu_i(\lambda)-\lambda\langle\sigma\rangle p)}. \tag{3.11}$$

Let us denote $\mu_0(\lambda)$ the eigenvalue with lowest real part and the corresponding left and right eigenvectors $|\psi_0^R(\lambda)\rangle$ and $|\psi_0^L(\lambda)\rangle$. Under the assumptions above, there will be at least a range of values of λ around zero such that the eigenvalue $\mu_0(\lambda)$ will be non-degenerate. Then, the integral over λ will be dominated for large t by the saddle-point value:

$$\zeta(p) = \mu_0(\lambda_{\rm sp}) - \lambda_{\rm sp}\langle\sigma\rangle p \tag{3.12}$$

where the saddle point $\lambda_{\rm sp}$ is a function of p determined by:

$$\frac{{\rm d}\mu_0}{{\rm d}\lambda}(\lambda_{\rm sp}) = \langle\sigma\rangle p. \tag{3.13}$$

The dependence upon the initial distribution is, within these assumptions, subdominant for large t.

In order to check that the distribution of p obtained from (3.12) and (3.13) is indeed peaked at $p = 1$, we calculate

$$\frac{\mathrm{d}\zeta(p)}{\mathrm{d}p} = \frac{\mathrm{d}\mu_0}{\mathrm{d}\lambda}\bigg|_{\lambda_{\rm sp}} \frac{\mathrm{d}\lambda}{\mathrm{d}p}\bigg|_{\lambda_{\rm sp}} - \frac{\mathrm{d}\lambda}{\mathrm{d}p}\bigg|_{\lambda_{\rm sp}} \langle\sigma\rangle p - \langle\sigma\rangle\lambda_{\rm sp} = -\langle\sigma\rangle\lambda_{\rm sp}. \tag{3.14}$$

Hence, the minimum of ζ occurs at $\lambda_{\rm sp} = 0$. We can further calculate the derivative of an eigenvalue using first-order perturbation theory. Equation (3.13) for $\lambda_{\rm sp} = 0$ then reads:

$$\langle\sigma\rangle p|_{\lambda_{\rm sp}=0} = \frac{\mathrm{d}\mu_0}{\mathrm{d}\lambda}\bigg|_{\lambda_{\rm sp}=0} = \frac{\langle\psi_0^L(0)|S|\psi_0^R(0)\rangle}{\langle\psi_0^L(0)|\psi_0^R(0)\rangle} = \langle -|S|\,\mathrm{stat}\rangle = \langle\sigma\rangle \tag{3.15}$$

where we have used the fact that in the absence of perturbation the lowest eigenvalue is zero and corresponds to the (unique) stationary state. Since then $\mu_0(0) = 0$, we obtain that $\zeta(p)$ takes its minimum (zero) value at precisely $p = 1$ (i.e. $\sigma_t = \langle\sigma\rangle$), and for large t the distribution is sharply peaked at $p = 1$, as it should.

3.3. The fluctuation–dissipation theorem as the 'conservative limit' limit of the FT

Gallavotti has shown [3] that the FT gives the fluctuation–dissipation theorem in the conservative limit of zero entropy production. Here, we shall paraphrase that derivation, as applied to the Langevin case.

Before doing so, let us first note that the detailed balance symmetry (2.15) is responsible in the purely conservative $S = 0$ case for the existence of fluctuation–dissipation and reciprocity relations. As we have seen, it is also responsible in the driven $S \neq 0$ case for the FT relation. The result in [3] is that the FT formula is, *on its own*, enough to give us back the fluctuation–dissipation and the reciprocity relations in the purely conservative limit.

Let us rewrite a form of the fluctuation–dissipation theorem for a conservative systems. In equation (2.24) we set $O = J$, and compute the response to a force $-h\boldsymbol{f}$ constant in time. Integrating (2.24) over t, t':

$$\int_0^t \mathrm{d}t' \frac{\partial\langle J(t')\rangle}{\partial h}|_{h=0} = \beta \int_0^t \mathrm{d}t'\,\mathrm{d}t''\langle -|J\mathrm{e}^{-(t'-t'')H^c} J|\,\mathrm{GB}\rangle\theta(t-t'). \tag{3.16}$$

The right-hand side of (3.16) can be re-expressed as follows

$$\frac{\beta}{2}\int_0^t \mathrm{d}t'\,\mathrm{d}t''\langle -|\mathrm{e}^{-(t-t')H^c} J\mathrm{e}^{-(t'-t'')H^c} J\mathrm{e}^{-t''H^c}|\,\mathrm{GB}\rangle = \frac{1}{2\beta}\frac{\partial^2}{\partial\lambda^2}\langle -|\mathrm{e}^{-t(H^c+\lambda\beta J)}|\,\mathrm{GB}\rangle|_{\lambda=0} \tag{3.17}$$

where we used the fact that H^c annihilates both $\langle -|$ and $|\,\mathrm{GB}\rangle$. Similarly, the left-hand side of (3.16) can be re-expressed as

$$\frac{\partial}{\partial h}\int_0^t \mathrm{d}t'\langle -|J\mathrm{e}^{-t'H(h)}|\,\mathrm{GB}\rangle|_{h=0} = \frac{\partial}{\partial h}\int_0^t \mathrm{d}t'\langle -|\mathrm{e}^{-(t-t')H(h)} J\mathrm{e}^{-t'H(h)}|\,\mathrm{GB}\rangle|_{h=0}$$
$$= \frac{1}{\beta}\frac{\partial^2}{\partial h\partial\lambda}\langle -|\mathrm{e}^{-t(H(h)+\lambda\beta J)}|\,\mathrm{GB}\rangle|_{\substack{h=0\\ \lambda=0}} \tag{3.18}$$

where $H(h)$ is the perturbed Hamiltonian. Then, we can rewrite the fluctuation–dissipation theorem as:

$$\left[\frac{1}{2}\frac{\partial^2}{\partial\lambda^2} - \frac{\partial^2}{\partial h\partial\lambda}\right]\langle -|\mathrm{e}^{-t(H(h)+\lambda\beta J)}|\,\mathrm{GB}\rangle|_{\substack{h=0\\ \lambda=0}} = 0. \tag{3.19}$$

Using the fact that $\langle J \rangle = 0$ in equilibrium for a conservative system, the first derivatives with respect to λ and with respect to h vanish and equation (3.19) can be rewritten as:

$$\left[\frac{1}{2}\frac{\partial^2}{\partial\lambda^2} - \frac{\partial^2}{\partial h \partial\lambda}\right] \ln\{\langle - | \mathrm{e}^{-t(H(h)+\lambda\beta J)} | \mathrm{GB}\rangle\}|_{\substack{h=0\\ \lambda=0}} = 0. \tag{3.20}$$

To lowest (quadratic) order in h and λ, the general solution of (3.20) is:

$$\ln\{\langle - | \mathrm{e}^{-t(H(h)+\lambda\beta J)} | \mathrm{GB}\rangle\} = A(t)(\lambda^2 + (\lambda + h)^2) + B(t)\lambda(h+\lambda) \tag{3.21}$$

where A, B are model dependent. Equation (3.21) is a form of the fluctuation–dissipation theorem.

Let us now show that the FT directly implies (3.21) in the purely conservative ($h = 0$) limit. We have that:

$$\ln\{\langle - | \mathrm{e}^{-t(H(h)+\lambda\beta J)} | \mathrm{GB}\rangle\} = \ln\{\langle - | \mathrm{e}^{-t(H(h)-\frac{\lambda}{h}S)} | \mathrm{GB}\rangle\} = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \mathrm{d}p\, \pi(p) \mathrm{e}^{\frac{-t\langle\sigma\rangle p\lambda}{h}}. \tag{3.22}$$

Now, *with the only assumption of the* FT applied to the term on the right we easily obtain:

$$\ln\{\langle - | \mathrm{e}^{-t(H(h)+\frac{\lambda}{h}S)} | \mathrm{GB}\rangle\} = \ln\{\langle - | \mathrm{e}^{-t(H(h)-(1+\frac{\lambda}{h})S} | \mathrm{GB}\rangle\} \tag{3.23}$$

order in λ, h implies (3.21).

3.4. A nonlinear fluctuation–dissipation theorem in the 'conservative limit'

So far we have only concentrated on the case in which the forces f_i do not derive from a potential and thus generate entropy at stationarity. However, the calculations can be performed in the case in which f_i derive from a potential:

$$f_i = -h \frac{\partial A(\boldsymbol{x})}{\partial x_i}. \tag{3.24}$$

In that case, S is no longer an entropy production, it represents the rate of variation of A. For example:

$$S = \beta h \frac{\partial A(\boldsymbol{x})}{\partial x_i} v_i = \beta h \frac{\mathrm{d}A}{\mathrm{d}t}. \tag{3.25}$$

If we start at $t = 0$ with the equilibrium distribution *in the absence of forces* f_i and only then switch on the forces, we easily obtain a version of the FT:

$$\frac{\pi_h(A(t) - A(0) = a)}{\pi_h(A(t) - A(0) = -a)} = \mathrm{e}^{\beta h a} \tag{3.26}$$

valid for *arbitrary* h. Here the subindex h means that the distribution depends on the field conjugate to A that has been on from $t = 0$ to t.

Using (3.26) in the limit of small h, we recover the usual fluctuation–dissipation theorem as follows

$$\begin{aligned}\langle A(t) - A(0)\rangle|_h &= \int a\mathrm{d}a\, \pi_h(A(t) - A(0) = a) = -\int a\, \mathrm{d}a\, \pi_h(A(t) - A(0) = -a) \\ &= -\int \mathrm{d}a\, \mathrm{e}^{-\beta h a} a \pi_h(A(t) - A(0) = a).\end{aligned} \tag{3.27}$$

This yields, to lower order in h:

$$\begin{aligned}\langle A(t) - A(0)\rangle|_h &= -\int a\mathrm{d}a\, \pi_h(A(t) - A(0) = a)(1 - \beta h a) \\ &= -\int a\, \mathrm{d}a\, \pi_h(A(t) - A(0) = a) - h\beta \int a^2\, \mathrm{d}a\, \pi_{h=0}(A(t) - A(0) = a)\end{aligned} \tag{3.28}$$

which implies:

$$\frac{\mathrm{d}}{\mathrm{d}h}\langle A(t) - A(0)\rangle|_{h=0} = \frac{\beta}{2}\langle (A(t) - A(0))^2\rangle|_{h=0} \tag{3.29}$$

which is a usual form of the fluctuation–dissipation theorem.

4. Fokker–Planck equation

If the inertial term in the Langevin equation vanishes the probability distribution at time t for the process (2.1) can be expressed only in terms of the positions (in fact, the velocities are in this case undefined). The evolution is now given by

$$P(\boldsymbol{x}, t) = \mathrm{e}^{-tH_{\mathrm{FP}}} P(\boldsymbol{x}, 0) \tag{4.1}$$

where H_{FP} is the Fokker–Planck operator:

$$H_{\mathrm{FP}} = -\partial_{x_i}(T\partial_{x_i} + \partial_{x_i} U(\boldsymbol{x}) + f_i). \tag{4.2}$$

We have set $\gamma = 1$ for this case. In bracket notation, we have:

$$\begin{aligned} P(\boldsymbol{x}, t) &= \langle x|\phi(t)\rangle \\ |\phi(t)\rangle &= \mathrm{e}^{-tH_{\mathrm{FP}}}|\phi(0)\rangle \end{aligned} \tag{4.3}$$

$$H_{\mathrm{FP}} = \hat{p}_{x_i}(T\hat{p}_{x_i} - \mathrm{i}\partial_{x_i} U - \mathrm{i} f_i) \tag{4.4}$$

which can again be separated in conservative and 'forced' parts:

$$H_{\mathrm{FP}} = H^c_{\mathrm{FP}} - \mathrm{i}\hat{p}_{x_i} f_i. \tag{4.5}$$

The total energy of the conservative part for this case is simply U.

The evolution of the system satisfies the usual detailed balance *in the absence of non-conservative forces*:

$$\langle x'|\mathrm{e}^{-tH^c_{\mathrm{FP}}}|x\rangle \mathrm{e}^{-\beta U(\boldsymbol{x})} = \langle x|\mathrm{e}^{-tH^c_{\mathrm{FP}}}|x'\rangle \mathrm{e}^{-\beta U(\boldsymbol{x}')} \tag{4.6}$$

leading to:

$$Q^{-1}_{\mathrm{FP}} H^c_{\mathrm{FP}} Q_{\mathrm{FP}} = H^{c\dagger}_{\mathrm{FP}} \tag{4.7}$$

where the operator Q_{FP} is defined by:

$$Q_{\mathrm{FP}}|x\rangle \equiv \mathrm{e}^{-\beta U(\boldsymbol{x})}|x\rangle. \tag{4.8}$$

4.1. Power and entropy production for Fokker–Planck processes

In a Langevin process without inertia it is not *a priori* obvious how to define the power done by the bath. In order to do this, we shall first make a heuristic argument and only then formalise it. Let us compute the power as

$$T\sigma_t = \int_0^t \mathrm{d}t'\, \dot{x}_i f_i. \tag{4.9}$$

Because the velocity is not well defined, we shall have to be careful about the meaning of this expression. Let us for the moment continue naively, writing a functional expression for the distribution $\Pi(\sigma_t)$:

$$\begin{aligned} \Pi(\sigma_t)t &= \left[\int \mathrm{D}(\boldsymbol{x})\boldsymbol{J}(\boldsymbol{x})\delta(\dot{x}_i + \partial_{x_i} U(\boldsymbol{x}) + f_i - \Gamma_i)\delta\left(t\sigma_t - \beta\int_0^t \mathrm{d}t\, \dot{x}_i f_i\right)\right]_\Gamma \\ &= t\left[\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \mathrm{d}\lambda \int \mathrm{D}(\boldsymbol{x})\,\mathrm{D}(\boldsymbol{p})\,\boldsymbol{J}(\boldsymbol{x})\mathrm{e}^{-\int_0^t[\mathrm{i}p_i(\dot{x}_i + \partial_{x_i} U(\boldsymbol{x}) + f_i - \Gamma_i) + \lambda\beta\dot{x}_i f_i]}\right]_\Gamma \mathrm{e}^{\lambda t\sigma_t} \end{aligned}$$

where $\boldsymbol{J}$ is the Jacobian associated with the equation of motion delta. The square brackets denote averaging over the noise. Performing this average, we obtain:

$$\Pi(\sigma_t) = t\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \mathrm{d}\lambda \int \mathrm{D}(\boldsymbol{x})\,\mathrm{D}(\boldsymbol{p})\,\boldsymbol{J}(\boldsymbol{x})\mathrm{e}^{-\int_0^t \mathrm{d}t'\,\mathrm{i}p_i(\dot{x}_i+\partial_{x_i}U(\boldsymbol{x})+f_i-\mathrm{i}Tp_i)} \times \mathrm{e}^{-\lambda\int_0^t \mathrm{d}t'\,(2\mathrm{i}f_ip_i-\beta(f_i+\partial_{x_i}U(\boldsymbol{x}))-\lambda^2\beta\int_0^t \mathrm{d}t'\,f_i^2)}\mathrm{e}^{\lambda t\sigma_t}.$$

Equation (4.10) is a functional expression of the equality:

$$\Pi(\sigma_t) = t\int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \mathrm{d}\lambda \langle -|\mathrm{e}^{-t[H_{\mathrm{FP}}-\lambda(\mathrm{i}f_i(\hat{p}_{x_i}-\mathrm{i}\beta\partial_{x_i}U(\boldsymbol{x})-\mathrm{i}\beta f_i)+\mathrm{i}\hat{p}_{x_i}f_i)-\beta\lambda^2 f_i^2]}|\,\mathrm{init}\rangle \mathrm{e}^{t\lambda\sigma_t}. \tag{4.10}$$

This expression, to be compared with (2.26), can be taken as the definition of 'power done by the bath', yielding the entropy production at stationarity. Note that in this definition we have made a precise choice of factor orderings (we have associated the c-number $2p_i f_i$ with the operator $\hat{p}_{x_i} f_i + f_i \hat{p}_{x_i}$).

We are now in a position to explore the consequences of a modified detailed balance. Proposing a tranformation such as (4.7) we find that:

$$\begin{aligned}\mathrm{e}^{\beta U}[H_{\mathrm{FP}} &- \lambda(\mathrm{i}f_i(p_i - \mathrm{i}\beta\partial_{x_i}U(\boldsymbol{x}) - \mathrm{i}\beta f_i) + \mathrm{i}p_i f_i) - \beta\lambda^2 f_i^2]\mathrm{e}^{-\beta U} \\ &= [H_{\mathrm{FP}} - \tilde{\lambda}(\mathrm{i}f_i(p_i - \mathrm{i}\beta\partial_{x_i}U(\boldsymbol{x}) - \mathrm{i}\beta f_i) + \mathrm{i}p_i f_i) - \beta\tilde{\lambda}^2 f_i^2]^\dagger\end{aligned} \tag{4.11}$$

where $\tilde{\lambda} \equiv -(1+\lambda)$. Performing the change of variables $\lambda \to \tilde{\lambda}$ in the integral and following the same steps as in the Kramers case we obtain the fluctuation theorem for $\Pi(\sigma_t)$ defined as in (4.10).

4.2. Nonlinear fluctuation–dissipation theorem

A direct derivation of the nonlinear fluctuation–dissipation theorem (3.26) can be made for dynamics which (like Fokker–Planck's, heat bath, etc) satisfies detailed balance. Denoting $H(h)$, the evolution Hamiltonian associated with a potential $U - hA(x)$, we then have:

$$\mathrm{e}^{\beta(U-hA)}H(h)\mathrm{e}^{-\beta(U-hA)} = H^\dagger(h). \tag{4.12}$$

We write

$$\pi_h(A(t) - A(0) = a) = \int \mathrm{d}\lambda \langle -|\mathrm{e}^{\lambda A}\mathrm{e}^{-tH(h)}\mathrm{e}^{-\lambda A}|\,\mathrm{GB}\rangle \mathrm{e}^{-\lambda a} \tag{4.13}$$

where $|\,\mathrm{GB}\rangle$ is the canonical distribution *at* $h = 0$. Now,

$$\begin{aligned}\langle -|\mathrm{e}^{\lambda A}\mathrm{e}^{-tH(h)}\mathrm{e}^{-\lambda A}|\,\mathrm{GB}\rangle &= \langle -|\mathrm{e}^{\lambda A}\mathrm{e}^{-\beta(U-hA)}\mathrm{e}^{-tH^\dagger(h)}\mathrm{e}^{\beta(U-hA)}\mathrm{e}^{-\lambda A}|\,\mathrm{GB}\rangle \\ &= \langle -|\mathrm{e}^{(\beta h+\lambda)A}\mathrm{e}^{-tH(h)}\mathrm{e}^{-(\beta h+\lambda)A}|\,\mathrm{GB}\rangle.\end{aligned} \tag{4.14}$$

Introducing this in (4.13) and performing the integral over λ one readily obtains (3.26).

5. Conclusions

In this paper we have derived the FT for Langevin processes of systems with finitely many degrees of freedom. We do not require any properties of the potential apart from boundedness, since the Langevin equation is 'as ergodic as possible'.

However, we have noted in several places that the derivations do not carry through automatically if the zero eigenvalue of the operator H is degenerate. This happens surely in the case that the system is disconnected from the bath ($\gamma = 0$) and in the $T = 0$ case. More interestingly, it may also happen in an infinite $N = \infty$ system.

Indeed, because the 'gap' in the spectrum is the inverse of a timescale, a gapless spectrum is an indication of 'slow' dynamics. This suggests that the FT might be violated for those (infinite) driven systems which in the absence of drive have a slow relaxational dynamics that does not lead them to equilibrium in finite times (as is the case of glassy systems, coarsening, etc).

In such systems it is known [4] that the fluctuation–dissipation theorem can be violated even in the limit of vanishing power input, although the frequency range of the violation respects some bounds [10]. The violation of the fluctuation–dissipation theorem (or alternatively, the appearance of 'effective temperatures' different from the bath temperature) seems to be a signature of the dynamics of conservative or near-conservative complex systems [9, 11]. This raises the intriguing possibility that the violation of FT might exist and play a similar role for strongly driven infinite systems.

Acknowledgments

I wish to thank L F Cugliandolo and P LeDoussal for useful suggestions. This work was supported in part by Natural Science Foundation under grant no PHY9407194.

References

[1] Gallavotti G and Cohen E D G 1995 *Phys. Rev. Lett.* **74** 2694
Gallavotti G and Cohen E D G 1995 *J. Stat. Phys.* **80** 931
[2] The interested reader will find a complete list of relevant publications that have followed GC in the reference (ZZ) of Gallavotti G 1998 *Preprint* cond-mat 9802012
[3] Gallavotti G 1996 *Phys. Rev. Lett.* **77** 4334
[4] Crisanti A and Sompolinsky H 1987 *Phys. Rev.* A **36** 4922
Horner H 1996 *Z. Phys.* B **100** 243
Here we follow: Cugliandolo L F, Kurchan J, Le Doussal P and Peliti L 1997 *Phys. Rev. Lett.* **78** 350 and [9].
[5] This is, within the Langevin context, the result of Sinai Y G 1972 *Russ. Math. Surv.* **27** 21
Sinai Y G 1977 *Lectures in Ergodic Theory (Lecture Notes in Mathematics)* (Princeton, NJ: Princeton University Press)
[6] Risken H 1985 *The Fokker–Planck Equation* (Berlin: Springer)
[7] Kubo R, Toda M and Hashitume N 1992 *Statistical Physics II. Nonequilibrium Statistical Mechanics* (Berlin: Springer)
De Groot S R and Mazur P 1984 *Non-equilibrium Thermodynamics* (New York: Dover)
[8] My attention to the role of the Gibbs Boltzmann distribution was drawn by LeDoussal P Private communication.
[9] Cugliandolo L F, Kurchan J and Peliti L 1997 *Phys. Rev.* E **55** 3898
[10] Cugliandolo L F, Dean D S and Kurchan J 1997 *Phys. Rev. Lett.* **79** 2168
[11] For a review on recent progress in theoretical aspects of glass models see Bouchaud J-P, Cugliandolo L F, Kurchan J and Mézard M 1997 Out of equilibrium dynamics in spin-glasses and other glassy systems *Spin-glasses and Random Fields* ed A P Young (Singapore: World Scientific)