Fluctuation Theorem, non linear response and the regularity of time reversal symmetry

Marcello Porta
Dipartimento di Fisica, Università di Roma “La Sapienza”,
Piazzale Aldo Moro 2, 00185 Roma Italy

The Gallavotti – Cohen Fluctuation Theorem (FT) implies an infinite set of identities between correlation functions that can be seen as a generalization of Green Kubo formula to the nonlinear regime. As an application, we discuss a perturbative check of the FT relation through these identities for a simple Anosov reversible system; we find that the lack of differentiability of the time reversal symmetry implies a violation of the Gallavotti – Cohen fluctuation relation. Finally, a brief comparison with Lebowitz – Spohn FT is reported.

I. INTRODUCTION

Despite many proposals have been advanced, a general theory of the steady state of dissipative systems is still lacking. Nevertheless, it is a remarkable fact that under suitable hypothesis something can be said, and physical predictions can be made; for example, chaotic hypothesis (CH), [1], stating that for the purpose of studying macroscopic properties, a system exhibiting chaotic motions may be regarded as a transitive hyperbolic (that is, Anosov) one, see [2], implies two remarkable results: the Gallavotti – Cohen fluctuation theorem, [3], [4], holding for reversible systems, and a formula describing the linear response of nonequilibrium systems, due to Ruelle, [5], [6]. These results have been and are still widely studied in the physical literature; see [7] (where a relation which inspired the FT was empirically discovered), [8], [9], [10], [11], [12] for instance. In this note we focus on the first on these two results; in particular, in section II we show that FT has some nice implication on nonlinear response (in [13] it has been already pointed out that FT implies the usual linear response theory, that is Green Kubo formula (GK) and Onsager reciprocity relations), while in section III we discuss a check of FT in a simple Anosov reversible system. Interestingly, a simple perturbative calculation shows that the lack of differentiability of the time reversal symmetry operator implies a violation of the Gallavotti – Cohen fluctuation relation. Finally, in section IV we show that in presence of a non differentiable time reversal symmetry an identity equivalent to the result proved by Lebowitz and Spohn in [14] is true.

The connection between FT and nonlinear response consists in the fact that FT implies identities between correlation functions of physical observables in a nonequilibrium steady state; this has been pointed out first in [15], where the authors considered systems whose evolution was stochastic and ruled by a master equation, and independently in [16], in the context of deterministic systems satisfying CH. The results of [16] are presented in sections II and III of this paper.

Before turning to our results, we spend a few words on some of the main features of Anosov systems; we refer the interested reader to [2] for a modern introduction to the subject. Consider a generic discrete dynamical system \( x_k = S^k x_0 \) (discreteness only causes technical problems, see [17], since \( S \) can be thought as the map arising from the Poincaré section of a system evolving in continuous time), and assume that \( S \) is Anosov; informally, this means that given a point \( x \) the nearby points separate exponentially fast from \( x \) in the future and in the past, except when located on a surface \( W_s(x) \) (stable manifold) or \( W_u(x) \) (unstable manifold), respectively for the future and for the past.

It is a well known result that Anosov systems admit an invariant measure \( \mu_+ \), the so – called SRB measure; in fact, given a sufficiently regular observable \( F(x) \) the following equality holds:

\[
\lim_{T \to +\infty} \frac{1}{T} \sum_{j=0}^{T-1} F(S^j x) = \int \mu_+(dx) F(x), \tag{1.1}
\]

apart from a set of points of zero volume measure. It is a remarkable fact that the SRB measure
admit in principle an explicit representation, similar to the equilibrium Gibbs distribution.

Notice that at equilibrium, that is when the system is stationary and non dissipative, the chaotic hypothesis implies the ergodic hypothesis, in the sense that assuming CH the SRB measure is the Liouville one; but in general, when dissipation in present the SRB measure is singular with respect to the volume, that is it is concentrated on a zero volume set.

To conclude, the SRB distribution verifies a large deviation theorem (see for example [17] for a proof of this statement for a special choice of $F$ and in the more complex case of Anosov flows). In fact, consider the finite time average $f = \frac{1}{\tau} \sum_{j=1}^{\tau-1} F(S^j x)$; then, it is possible to prove that there are values $f_1, f_2$ such that if $[a, b] \in (f_1, f_2)$ then $\text{Prob}_{\mu^+} (f \in [a, b]) \sim e^{*\xi_\tau(f)}$, 

\[
\lim_{\tau \to +\infty} \frac{1}{\tau} \log \text{Prob}_{\mu^+} (f \in [a, b]) = \max_{f \in [a, b]} \xi_\tau(f),
\]

and $\xi_\tau(f)$ is analytic and convex in $(f_1, f_2)$.

II. FLUCTUATION THEOREM AND NONLINEAR RESPONSE

At present time, no universally accepted definition of entropy for a dissipative system has been given. Nevertheless, the rate of entropy production is a well defined quantity, and is proportional to the work per unit time made by the thermostats on the system; the proportionality factor is the inverse of the temperature of the thermostats (setting to 1 the Boltzmann constant). In particular, for a special class of thermostats, the gaussian ones, the entropy production rate corresponds to the phase space contraction, that is to minus the divergence of the equation of motion, see [18]; this fact can be taken as a general definiton of entropy production, if one assumes that the steady state of a large system is not affected by the details of the thermostating mechanism which ensures the existence of the steady state.

In the case of a system evolving in discrete time, which is the case that we want to consider, the entropy production rate $\sigma$ is given by $\sigma(x) = -\log |\det \partial S(x)|$, where $\partial S$ is the jacobian of the time evolution $S$.

Now, let $p$ be the adimensional average over a time $\tau$ of $\sigma(x)$, that is $p = \frac{1}{\tau \sigma^+} \sum_{j=1}^{\tau-1} \sigma(S^j x)$, where $\sigma^+$ is the SRB expectation of $\sigma(x)$, i.e. $\sigma^+ = \int (dx) \sigma(x)$, and call $\zeta(p)$ the large deviation functional of $\sigma$, as defined in (1.2); assume that CH holds, and that the system is reversible, which means that there exists a differentiable isometry $I$ such that $I \circ S = S^{-1} \circ I$, $I \circ I = 1$. Then, as proven by Gallavotti and Cohen, see [3] or [19] for a proof detailed from a formal viewpoint, the following result holds.

**Fluctuation theorem (FT):** There is $p^* \geq 1$ such that for $|p| < p^*$

\[
\zeta(-p) = \zeta(p) - p \sigma^+ ,
\]

(2.3)

This result has an interesting corollary. Setting $\pi_\tau(q) dq = \text{Prob}_{\mu^+} (p \in [q, q+ dq])$, define $\lambda(\beta)$ as

\[
\lambda(\beta) = \lim_{\tau \to +\infty} \frac{1}{\tau} \log \int e^{\tau (q-1) \sigma^+ + \beta \pi_\tau(q)} dq ;
\]

(2.4)

clearly, $\lambda(\beta)$ is related to $\zeta(p)$ through a Legendre transform, that is

\[
\zeta(p) = \max_{\beta} (\beta \sigma^+(p-1) - \lambda(\beta)) ,
\]

(2.5)

and moreover $\lambda(\beta)$ admits the following expansion:

\[
\lambda(\beta) = \sum_{n \geq 2} \sum_{t_1, \ldots, t_{n-1} = -\infty}^{+\infty} < \sigma(\cdot) \sigma(S^{t_1} \cdot) \ldots \sigma(S^{t_{n-1}} \cdot) \geq_T \frac{\beta^n}{n!} \equiv \sum_{n \geq 2} C_n \frac{\beta^n}{n!} ,
\]

(2.6)

where by $< ... >_T$ we denote the cumulant with respect to the SRB measure $\mu^+$. It is straightforward to see that the fluctuation theorem implies an identity for the generating functional.
\( \lambda(\beta) \), see [14], [15], [12] and [16] for instance: it follows that, as a consequence of the relation 
\[ \pi_\tau(p) \sim e^{\tau p^* + \pi_\tau(-p)} , \]
valid under the hypothesis of FT,
\[ \lambda(\beta) = \lambda(-1 - \beta) - \sigma_+(2\beta + 1) . \] (2.7)

Notice that the generating functional of the cumulants is usually defined as (see [14], [15], [12])
\[ \lambda(\beta) = \lim_{\tau \to +\infty} -\frac{1}{\tau} \log \langle e^{-\tau \sigma_+ \beta} \rangle , \]
and with this definition the relation (2.7) is replaced by the more familiar \( \lambda(\beta) = \lambda(1 - \beta) \); but the two definitions are equivalent, since \( \lambda(\beta) = -\beta \sigma_+ - \lambda(\beta) \). Formula (2.7) translates immediately in a relation for \( \sigma_+ \); in fact, (2.7) evaluated at \( \beta = 0 \) becomes:
\[ 0 = \lambda(-1) - \sigma_+ \Rightarrow \sigma_+ = \sum_{n \geq 2} C_n \frac{(-1)^n}{n!} . \] (2.8)

Assuming that the entropy production \( \sigma(x) \) has the form \( \sigma(x) = \sum_i G_i J_i^{(0)}(x) + O(G^2) \), where \( \{G_i\} \), and \( \{J_i^{(0)}(x)\} \) are respectively the forcing parameters and the corresponding currents, it has been shown in [13] that the identity
\[ \sum_{i,j} G_i G_j \partial_{G_i G_j} \sigma_+ \big|_{\mathcal{G} = 0} = \sum_{i,j} G_i G_j \partial_{G_i G_j} \sigma_+ \big|_{\mathcal{G} = \mathcal{G}^2} , \] (2.9)
which is nothing else that (2.8) at second order, is equivalent to the Green Kubo formula, stating that
\[ L_{ij} \equiv \partial_{G_i} J_i^{(0)} > + \big|_{\mathcal{G} = 0} = \frac{1}{2} \sum_{t=\pm\infty} < J_i^{(0)}(S^t) J_i^{(0)}(\cdot) > , \] (2.10)
where by \( < >_0 \) we denote the expectation with respect to the invariant measure at zero forcing
(by CH is the Liouville measure). Hence, formula (2.8) can be seen as a generalization of GK formula to the nonlinear regime, being an identity for \( \sigma_+ \) valid for \( \mathcal{G} \neq 0 \). Moreover, by taking derivatives with respect to \( \beta \) in the r.h.s. and in the l.h.s. of (2.7) we find that
\[ C_n = \sum_{k \geq 0} \frac{(-1)^{k+n} C_{k+n}}{k!} \quad n \geq 2 , \] (2.11)
which is a nontrivial identity for the cumulants valid at \( \mathcal{G} \neq 0 \). Finally, these identities can be considerably extended by using a generalized version of FT. Consider a generic observable \( O \) odd under time reversal symmetry, i.e. such that \( O(-x) = -O(x) \), let \( w \) be its adimensional average over a finite time \( \tau \), that is \( w = \frac{1}{\tau} \sum_{j=1}^{\tau-1} O(S^j x) \) where \( O_+ \) is the SRB average of \( O \), and call \( \zeta(p, w) \) the large deviation functional of the joint probability distribution \( \pi_\tau(p, w) \) of \( p \) and \( w \); then, under the same hypothesis of FT the following result holds, as a special case of a much more general result in [20].

**Generalized Fluctuation Theorem (GFT):** There are \( w^* \geq 1, p^* \geq 1 \) such that for \( |p| < p^* \) and \( |w| < w^* \)
\[ \zeta(-p, -w) = \zeta(p, w) - p \sigma_+ . \] (2.12)

One can define the generating functional \( \lambda(\beta_1, \beta_2) \) of the mixed cumulants of \( O, \sigma \) in a way analogous to (2.4),
\[ \lambda(\beta_1, \beta_2) = \lim_{\tau \to +\infty} -\frac{1}{\tau} \log \int e^{\tau(q-1)\sigma_+ + \tau(q-1)O_+ + \beta_2 \sigma_+(q, t)} dq dt ; \] (2.13)
again, \( \lambda(\beta_1, \beta_2) \) is related to \( \zeta(p, w) \) through a Legendre transform, and moreover it can be expressed as
\[ \lambda(\beta_1, \beta_2) = \sum_{k \geq 2} \sum_{m,n \geq 0 : m+n=k} \frac{\beta_1^m \beta_2^n}{m!n!} C_{n,m} , \] (2.14)
where $C_{n,m} = \partial_{\beta_1} \partial_{\beta_2} \lambda(\beta_1, \beta_2) |_{\beta = 0}$, that is $C_{n,m}$ is given by a sum over times of mixed cumulants of $\sigma(S^i x), O(S^j x)$, $0 \leq i \leq n$, $0 \leq j \leq m$. Then, in full analogy with what has already been discussed, it is easy to show that GFT translates into an identity for the generating functional, namely

$$\lambda(\beta_1, \beta_2) = \lambda(-1 - \beta_1, -\beta_2) - (2\beta_1 + 1)\sigma_+ - 2\beta_2 O_+ ,$$  

(2.15)

which implies the following relations:

$$O_+ = \sum_{k \geq 2, \text{even}} \frac{1}{(k-1)!} \frac{(-1)^k}{2^{k-1}} C_{1,k-1} + \sum_{k \geq 3, \text{odd}} \frac{1}{k!} \frac{(-1)^k}{2^{k-1}} C_{0,k} + O(G^2)$$  

(2.16)

$$C_{l,n-l} = \sum_{k \geq 0} \frac{(-1)^{n+k}}{k!} C_{k+l,n-l}, \quad n \geq 2, \ l \leq n ;$$  

(2.17)

equation (2.16) is obtained setting $\beta_1 = \beta_2 = -\frac{1}{2}$ in (2.15), while (2.17) can be proved differentiating with respect to $\beta_1, \beta_2$ the r.h.s. and the l.h.s. of (2.15). It is interesting to see what happens to formula (2.16) in the linear regime. Assuming that $\sigma(x) = \sum_l G_l j_l^{(0)}(x) + O(G^2)$ we can rewrite (2.16) as

$$O_+ = \sum_{k \geq 2, \text{even}} \frac{1}{(k-1)!} \frac{(-1)^k}{2^{k-1}} C_{1,k-1} + \sum_{k \geq 3, \text{odd}} \frac{1}{k!} \frac{(-1)^k}{2^{k-1}} C_{0,k} + O(G^2)$$  

(2.16)

$$= \frac{1}{2} C_{1,1} + \sum_{k \geq 3, \text{odd}} \frac{1}{k!} \frac{(-1)^k}{2^{k-1}} C_{0,k} - 1 \frac{1}{2} C_{1,1} + O(G^2) ,$$  

(2.18)

and from (2.17) we find that for $k$ odd $C_{0,k} - \frac{1}{2} C_{1,1} = O(G^2)$. Hence,

$$O_+ = \frac{1}{2} C_{1,1} + O(G^2) = \frac{1}{2} \sum_{l=-\infty}^{\infty} \sigma(S^l x)O(x) >_+ + O(G^2) ,$$  

(2.19)

which gives:

$$\partial_G O_+ |_{G=0} = \frac{1}{2} \sum_{l=-\infty}^{\infty} \sigma(S^l x)O(x) >_0 .$$  

(2.20)

Formula (2.20) describes the linear response of a generic observable $\text{odd}$ under time reversal symmetry; this result can be seen as a special case of the remarkable linear response formula obtained by Ruelle, [5], valid in a much more general context.

**III. CHECK OF FLUCTUATION THEOREM FOR A SIMPLE ANOSOV SYSTEM**

In this section we will perform a check of the fluctuation relation (2.3) for a simple Anosov model, the perturbed Arnold cat map, starting from the identities (2.8), (2.11); as we are going to see, the lack of differentiability of the time reversal symmetry operator implies a violation of (2.3). Consider the discrete evolution on the bidimensional torus $\mathbb{T}^2$

$$x_k = S^k x_0 \mod 2\pi ,$$  

(3.21)

where

$$S^k x = S^k x + \varepsilon f(x), \quad S \equiv \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} , \quad f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} ,$$  

(3.22)
and $f_1(x)$, $f_2(x)$ are trigonometric polynomials. The map $S$ is the so-called Arnol’d cat map, which is the simplest example of Anosov map: in fact, the eigenvalues $\lambda_+, \lambda_-$ of $S$ are such that $\lambda_+ > 1$, $\lambda_- < 1$. Moreover the map $S$ is reversible, that is there exist $I$ such that $I \circ S = S^{-1} \circ I$, $I \circ I = 1$, where

$$I = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}.$$  

(3.23)

Notice that since $S$ is conservative (det $S = 1$) $\sigma_+ = 0$, which makes meaningless the fluctuation relation (2.3) (the adimensional quantity $p$ is not defined); but one can derive the analogous of (2.3) for the dimensional quantity $p' = \sigma_+ p$, and this relation becomes trivial if the evolution is conservative because in this case $p' = 0$.

Consider now $\varepsilon \neq 0$. By the structural stability of Anosov systems, the evolution generated by $S_\varepsilon$ is still Anosov provided $\varepsilon$ is chosen small enough. In fact, for $\varepsilon < \varepsilon_0$ a conjugation $H_\varepsilon$, defined by the identity $S_\varepsilon \circ H_\varepsilon = H_\varepsilon \circ S$, can be explicitly constructed through a convergent power series in $\varepsilon$, and it turns out that $H_\varepsilon(x)$ is Hölder continuous in $x$, [2]. In an analogous way, the SRB measure can be explicitly constructed, and it follows that the expectations of Hölder continuous functions exists and are analytic in $\varepsilon$, [2]. In particular it follows that, generically,

$$\sigma_+ \equiv \langle - \log |\partial S_\varepsilon| \rangle_+ > 0,$$

(3.24)

which means that the system is dissipative and so that the invariant measure is singular with respect to the volume; hence, the check of the fluctuation relation (2.3) is nontrivial in this case. Notice that the proof of FT, [19], requires that the evolution is reversible, and in particular that the time reversal symmetry operator is differentiable; in our specific case the existence of $H_\varepsilon$ implies that $I_\varepsilon = H_\varepsilon \circ I \circ H_\varepsilon^{-1}$ verifies $I_\varepsilon \circ S_\varepsilon = S_\varepsilon^{-1} \circ I_\varepsilon$, but due to the mild regularity properties of $H_\varepsilon$, $I_\varepsilon(x)$ is likely to be not differentiable. Then one can ask what happens to the fluctuation relation; to understand this, we make the choice $f_1(x) = \sin(x_1) + \sin(2x_1)$, $f_2(x) = 0$. An explicit computation shows that the linear response is still valid, as expected, and that

$$C_4 = -12\varepsilon^3 + O(\varepsilon^4),$$

(3.25)

i.e. formula (2.11) with $n = 3$ is false at the lowest nontrivial order in perturbation theory, since it tells that $C_4 = C_4/2 + O(\varepsilon^3)$, and $C_4 = O(\varepsilon^4)$. This is enough to say that (2.3) is violated; in fact, from (2.5) it follows that, [16],

$$\zeta(p) = -\frac{(p - 1)^2}{2} \left[ \sigma_+ - \frac{C_2}{4} + O(\varepsilon^4) \right] + \frac{(p - 1)^3}{48} \left[ C_3 + O(\varepsilon^4) \right] + O((p - 1)^4\varepsilon^4),$$

(3.26)

that is

$$\zeta(p) - \zeta(-p) = ps\sigma_+ + p \left[ \sigma_+ - \frac{C_2}{2} + \frac{C_3}{8} + O(\varepsilon^4) \right] + p^3 \left[ \frac{C_4}{24} + O(\varepsilon^4) \right] + O(p^5\varepsilon^5),$$

(3.27)

and (3.25) implies that the difference $\zeta(p) - \zeta(-p)$ is not linear in $p$ (the cubic term is nonzero). This result also shows that, as expected, $I_\varepsilon$ cannot be differentiable: a check that would be not so easy without using the FT.

IV. CONCLUSIONS AND COMPARISON WITH [14]

The Gallavotti – Cohen Fluctuation Theorem implies nontrivial identities between correlation functions valid at non zero forcing $G$, which reduce to the usual linear response in the limit $G \to 0$; we have shown that it is essential that the time reversal transformation be (continuously) differentiable. By checking the FT relation through these identities, in a simple Anosov system where the time reversal symmetry operator exists but is not differentiable, we have found that the relation expected from the possible validity of FT cannot hold (while the linear response is still valid, as expected). Notice, however, that in physical applications the time
reversal symmetry operation is usually regular; for instance, it can correspond to the inversion of velocities (or to more subtle permutations of coordinates, see [21]).

To conclude, it is interesting to note that in presence of a non differentiable time reversal symmetry operator $I$ a different FT holds, the Lebowitz – Spohn one, see [14]; strictly speaking this theorem has been proved in the context of general Markov processes, but it can be understood also in the case of deterministic chaotic dynamics. This result applies in particular to systems that are (suitably small) perturbations of reversible Anosov ones, since, as it has been pointed out in the previous section, by the structural stability of Anosov dynamics the time reversal symmetry is not destroyed by the perturbation (although it will be in general only Hölder continuous in $x$).

The fact that a fluctuation relation still holds is a natural consequence of the Gibbian nature of the invariant measure describing the steady state; this has been pointed out in [22], where a FT for the one dimensional Ising model in an external field was derived, and then systematized in [23], where large deviation rules for Gibbs states involving transformations different from time reversal were discussed.

Following step by step the proof of Gallavotti – Cohen FT, [19], it is easy to see that the large deviation functional $\tilde{\zeta}(\tilde{p})$ of the dimensionless quantity $\tilde{p} = \frac{1}{\sigma_0} \sum_{j=-\infty}^{-1} \tilde{\sigma}(S^j x)$ where $\tilde{\sigma}(x) = -\lambda_\sigma(x) + \lambda_u(Ix)$ and $\lambda_\sigma(x)$ is the sum of the positive Lyapunov exponents corresponding to the local unstable manifold $W_u(x)$ of $S$ verifies

$$\zeta(-\tilde{p}) = \zeta(\tilde{p}) - \tilde{p} \tilde{\sigma}_+,$$  \hspace{1cm} (4.28)

(of course under the physical restriction on $\tilde{p}$ to vary within the analyticity interval of $\tilde{\zeta}$, see [19]) which reduces to (2.5) if $I$ is differentiable, since in this case $\sigma_\sigma(Ix) = -\lambda_\sigma(x)$. Formula (4.28) is equivalent to the Lebowitz – Spohn fluctuation theorem; in fact, $\tilde{\sigma}(x)$ is proportional to the logarithm of the ratio of the SRB probabilities of the trajectories $x, x+1, ..., x_{-}\frac{1}{2}$ and of its time reversed, which is precisely the “action functional” introduced in [14].

---