

# The Steady State Fluctuation Relation for the Dissipation Function

Debra J. Searles · Lamberto Rondoni · Denis J. Evans

Received: 16 December 2006 / Accepted: 24 June 2007 / Published online: 7 August 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** We give a proof of transient fluctuation relations for the entropy production (dissipation function) in nonequilibrium systems, which is valid for most time reversible dynamics. We then consider the conditions under which a transient fluctuation relation yields a steady state fluctuation relation for driven nonequilibrium systems whose transients relax, producing a unique nonequilibrium steady state. Although the necessary and sufficient conditions for the production of a unique nonequilibrium steady state are unknown, if such a steady state exists, the generation of the steady state fluctuation relation from the transient relation is shown to be very general. It is essentially a consequence of time reversibility and of a form of decay of correlations in the dissipation, which is needed also for, e.g., the existence of transport coefficients. Because of this generality the resulting steady state fluctuation relation has the same degree of robustness as do equilibrium thermodynamic equalities. The steady state fluctuation relation for the dissipation stands in contrast with the one for the phase space compression factor, whose convergence is problematic, for systems close to equilibrium. We examine some model dynamics that have been considered previously, and show how they are described in the context of this work.

**Keywords** Nonequilibrium phenomena · Entropy production · Chaotic hypothesis · Time reversibility · Correlation decay

---

D.J. Searles (✉)

Nanoscale Science and Technology Centre and School of Biomolecular and Physical Sciences,  
Griffith University, Brisbane, Qld 4111, Australia  
e-mail: D.Bernhardt@griffith.edu.au

L. Rondoni

Dipartimento di Matematica and CNISM, Politecnico di Torino, Corso Duca degli Abruzzi 24,  
10129 Torino, Italy  
e-mail: lamberto.rondoni@polito.it

D.J. Evans

Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia  
e-mail: evans@rsc.anu.edu.au

## 1 Introduction

Steady state fluctuation relations provide information about the probability distribution of time averages of certain phase variables for systems in nonequilibrium steady states: they express the ratio of the probabilities of observing positive and negative values of time averages of such variables. These relationships have received much attention over the past decade (see e.g. [1–4]). In fact, they are among the few exact results regarding nonequilibrium systems; close to equilibrium they lead to the Green–Kubo relations for linear transport coefficients [5]; they provide information about fluctuations in nanoscale systems and they show how irreversibility can emerge from time reversible dynamics.

The first such fluctuation relation (FR) for a nonequilibrium particle system was given in Ref. [6], in 1993. It concerned fluctuations of time averages of the energy dissipation or equivalently the phase space contraction rate, for an isoenergetic shearing interacting particle system. In the isoenergetic dissipative system of Ref. [6], the instantaneous energy dissipation equals the instantaneous phase space contraction rate. The form of the relation was the following:

$$\frac{P_\tau(A)}{P_\tau(-A)} = e^{\tau A} \quad (1)$$

where  $A$  and  $-A$  were the time averaged values of the dissipation on steady state trajectory segments of duration  $\tau$ , and  $P_\tau(A)$  is the steady state probability of observing the average value  $A$  (with some tolerance), in a time  $\tau$ . In analogy with the periodic orbit expansion for dynamical systems [7–9], the relation was derived heuristically using the “Lyapunov weights” in the long  $\tau$  limit, and was not expected to work at short  $\tau$ 's.

Paper [6] motivated a number of works, in which various derivations of formulae formally similar to (1) were given for different systems, although for non-isoenergetic systems the phase space contraction rate and the energy dissipation are no longer instantaneously equal. Beginning in 1994, a series of papers by Evans and Searles focused on the fluctuation properties of the “dissipation function”  $\Omega$ , cf. Refs. [1, 10–16], while Gallavotti and Cohen, in 1995 *et seq.*, considered the fluctuations of the time-averaged phase space contraction rate,  $-\Lambda$ , under the assumption that the dynamics is time reversible, transitive and Anosov or that the system satisfies the Chaotic Hypothesis (*i.e.* although the dynamics is not Anosov, for the purposes of this FR, it behaves as if it was Anosov), cf. Refs. [17–20]. Various generalizations and extensions of these relations have been produced by different authors, see e.g. Refs. [12, 21–32].

In the literature, FR's which superficially resemble the form of (1), are often collectively referred to as the Gallavotti–Cohen Fluctuation Relations, or instances of the Gallavotti–Cohen symmetry, although they may have subtly different meanings. Therefore, for the sake of clarity, we refer to the relation for the phase space contraction rate of Refs. [6, 17, 18], as the  $\Lambda$ -FR, while we refer to the relation for the dissipation function of Refs. [1, 6, 10, 11], as the  $\Omega$ -FR. In the cases in which  $\Omega = -\Lambda$ , like those of [6], the two relations have the same content.

$\Omega$ -FR's have been widely used and tested numerically and experimentally, for various transient [10, 11] as well as steady state systems (see e.g. [2, 13, 16, 21, 32–39]). Although the transient  $\Omega$ -FR's have been proven more than a decade ago, beginning with Ref. [10, 11], there has been some question over the derivation of the corresponding steady state relations [40]. Therefore, we present a derivation of the  $\Omega$ -FR's which is more thorough than the ones that have appeared thus far in the literature [1]. The purpose is to shed light on the mechanisms which allow the  $\Omega$ -FR's to hold, and which may be used to better

understand properties of nonequilibrium systems. In particular, we are interested in nonequilibrium steady states where the dissipation function takes the form

$$\Omega = \Sigma = -JV F_e / (k_B T), \quad (2)$$

where  $J$  is the dissipative flux in the system of interest,  $V$  is the volume of the system of interest,  $k_B$  is Boltzmann's constant,  $T$  is the equilibrium thermodynamic temperature of the thermal reservoir with which the system of interest is in contact, and  $\Sigma$  is the generalized entropy production (or energy dissipation) induced by the dissipative field  $F_e$  applied to the system of interest [1, 13]. For the equality  $\Omega = \Sigma$  to hold, we require that the unthermostated (*i.e.* adiabatic) field dependent dynamics preserves phase space volumes instantaneously, and that the initial distribution of phases could in principle be generated by a single (exceedingly long) zero-field phase space trajectory. Systems that satisfy the condition of the adiabatic incompressibility of phase space are called *AIF* in Ref. [41] (cf. Appendix 1).

When referring to steady state systems, we only consider systems which relax to a *unique* steady state. The necessary and sufficient conditions for the system of interest to relax to a unique steady state are unknown. Certainly, it is known that under certain circumstances nonequilibrium steady states are not possible [42]. If the systems of interest are near equilibrium, where the thermodynamic notions of temperature and entropy can be used, then as is shown in the theory of linear irreversible thermodynamics the time or ensemble average of  $\Omega$  is equal to leading order in the dissipative field, to the average so-called spontaneous entropy production rate [43]. Outside this local equilibrium regime, the entropy production rate cannot be defined. While the notions of a thermodynamic temperature, thermodynamic entropy or an entropy production rate cannot be defined outside the local thermodynamic equilibrium regime, in a nonequilibrium steady state the ensemble or time averaged phase space contraction rate is equal to the steady state time derivative of the fine grained Gibbs entropy. The fact that the fine grained Gibbs entropy is not constant in a steady state (unlike the pressure or the energy), is the reason why it is so difficult to attach meaning to standard thermodynamic quantities such as temperature, in steady states outside the local equilibrium regime.

In the derivation of a relation like (1), one may follow different approaches. The one which has led to the  $\Delta$ -FR requires the full knowledge of the physical measure of the dynamics, which is then assumed to be of the Anosov type, hence to have a Sinai–Ruelle–Bowen (SRB) measure,  $\mu_{\text{SRB}}$  [17, 18, 44]. This is the mathematical approach, which aims to identify the class of dynamical systems for which a relation like (1) can be rigorously obtained, regardless of the physical relevance of such dynamical systems or of the quantities appearing in (1). How far one can go along this line and which other condition could replace the Anosov condition are highly nontrivial and interesting mathematical questions, which do not have an answer, at present.

The explicit knowledge of the physical measure gives the maximum possible amount of information about the phase space distribution: all statistical properties of the dynamics can be obtained from it. However, if one is only interested in why the FR holds, less detailed knowledge of the phase space distribution is likely to suffice, and special features of the dynamics, such as the Anosov property, need not play a role. This may explain why many *physical* systems have an observable which obeys a relation like (1). As a matter of fact, hardly any particle system of physical interest is of the Anosov kind, therefore it is interesting to ask whether a physical system obeying a fluctuation relation does so because it shares certain properties with Anosov systems, or whether it obeys such relations for other reasons. To investigate this question, which is part of the physical approach, we follow Evans

and Searles [13], since they do not require the invariant phase space probability measure to be explicitly determined. This will lead to transient fluctuation relations for  $\Sigma$  as well as for many other functions of phase, including  $\Lambda$ . If  $\phi$  is the observable of interest, we speak of  $\phi$ -FR.

The purpose of the present paper is to understand why the  $\Sigma$ -FR holds for many physically interesting systems, and why this can be observed within physically relevant time scales. This is useful to further develop the present theory.

*Remark* This issue is not completely unrelated to, but is quite different from the identification of the class of dynamical systems which allow a rigorous mathematical derivation of the  $\Lambda$ -FR. One may say that the two issues are as much related as Khinchin's approach to the ergodic hypothesis [45] is related to ergodic theory.

In Sect. 2 we derive transient FR's for time reversal invariant dynamics. These relations are called "transient" because they describe the time evolving statistics of an ensemble of finite time averages of given phase functions. They hold under very general conditions, similarly to other transient relations, like the Jarzynski and the Crooks relations [46, 47]. In particular, the transient relations hold even if a steady state with positive and negative fluctuations of the observable under consideration is not reached. If a fluctuating steady state is not reached, the transient relations hold and express properties of evolving ensembles of phase space trajectories, while the Anosov structure cannot characterize the dynamics.

In Sect. 3, we show how the results of Sect. 2 can be used to derive asymptotic FR's for reversible systems, under the technical assumption that  $\Omega$  is bounded. In Sect. 4, the decay of the  $\Omega$ -autocorrelation is shown to be sufficient for the steady state  $\Omega$ -FR to hold. Furthermore, in isoenergetic systems which satisfy  $AIT$  (e.g. systems whose adiabatic evolution is Hamiltonian), the phase space contraction rate and the dissipation function are instantaneously equal. In these cases, the predictions of the  $\Lambda$ -FR and of the  $\Omega$ -FR can be compared directly, and the  $\Omega$ -FR at high dissipation implies a different behavior from the one predicted by the  $\Lambda$ -FR under the axiom-C condition of Ref. [48], but which agrees with the results of [16, 49]. This is possible because the conjectures of [48, 49] are not necessarily verified by particle systems. The approach followed in Sects. 2, 3 and 4 leads to a number of  $\phi$ -FR's and to various other relations. In Sect. 5 we discuss our results. Appendix 1 contains a detailed description of the dissipation function for the common cases. Appendix 2 discusses the problem of axiom-C systems and the decay of correlations, and explains that our approach is based on exact relations, which are then valid even when the conditions of Sect. 4 are not met. Our conclusions are that:

- to satisfy the steady state  $\Omega$ -FR, the time reversal invariance of the dynamics and the decay of the  $\Omega$ -autocorrelations, with respect to the *initial* (non-singular) probability measure, suffice. This decay of correlations is not the one characterizing the steady states of axiom-C systems, which concerns all observables and is referred to the *invariant* (singular) measure.
- For  $\Omega$  to equal  $\Sigma$ , hence to obtain a steady state  $\Sigma$ -FR,  $AIT$  must hold and the initial distribution of phases must in principle be obtainable from a single field free ( $F_e = 0$ ) phase space trajectory. In such a case, we say that the initial distribution is ergodically consistent with the field free dynamics [1].
- The  $\Omega$ -FR's avoid the difficulties encountered by the  $\Lambda$ -FR close to equilibrium. The  $\Lambda$ -FR, requires ever longer convergence times, the closer the steady state is to equilibrium [5, 13, 33–35, 50]. The  $\Omega$ -FR's hold in a range which does not shrink for decreasing

fields, and are verified within the physically relevant ( $\Omega$ -autocorrelation) times, both for systems far and close to equilibrium.

- The direct derivation of the steady state  $\Sigma$ -FR's avoids the difficulties connected with obtaining a relation for the energy dissipation via the  $\Lambda$ -FR, justifies the convergence rates, and leads to new testable relations. However, the class of dynamical systems which enjoy the required decay of correlations remains to be identified.
- It goes without saying that the steady state relations require that a steady state is reached (hence that the correlations decay with respect to the initial state), and that fluctuations of opposite sign in the dissipation function should be observable. If such a steady state is not attained or if the steady state is sufficiently far from equilibrium that negative values of the dissipation function cannot be observed, the steady state FR's cannot be applied, but the transient FR's (including their asymptotic forms) remain valid as properties of the evolving ensembles, because they only require time reversibility.

In this sense, our relations are like the usual thermodynamic relations which are extraordinarily general, being independent of the nature of interparticle interactions and dynamics. All that is required of the dynamics is that it should be time reversible and, for steady state FR's, that the autocorrelation of  $\Omega$  decays and that the steady state is unique.<sup>1</sup> Furthermore our derivation explains why the  $\Omega$ -FR is verified within times which can be expressed in terms of material properties.

## 2 Transient Fluctuation Relation

Because we will be interested in thermodynamic systems which possibly relax to a nonequilibrium steady state, our dynamics will consist of the Hamiltonian dynamics of an  $N$ -particle system of interacting particles to which a dissipative field is applied. The resultant equations of motion may or may not be derivable from a Hamiltonian [41]. In order for such a system to be capable of relaxing to a nonequilibrium steady state, the system must be allowed on average, to loose heat. This can be accomplished by surrounding the system of interest with thermostating walls which serve two purposes: to confine the system of interest and also to remove dissipative heat generated (on average) by the dissipative field applied to it. It has been shown that any loss of heat from a Hamiltonian system requires that the dynamics over the degrees of freedom in the system of interest is no longer volume preserving [51, 52]. These details are explained further in the Appendix. In what follows however, we will consider a more general setting that is time reversible and does not necessarily preserve the phase space volumes.

Given the phase space  $\mathcal{M}$  of a particle system, consider a probability measure  $\mu$  on  $\mathcal{M}$ , with density  $f$ , i.e.  $d\mu(\Gamma) = f(\Gamma)d\Gamma$  for every point  $\Gamma \in \mathcal{M}$ . The measure  $\mu$  does not need to be produced by any dynamics on  $\mathcal{M}$ , although it could represent, for instance, the initial equilibrium distribution of the system. Further, given a (sufficiently well behaved) phase function  $\phi : \mathcal{M} \rightarrow \mathbb{R}$ , the probability that it takes values in the interval  $(a, b) \subset \mathbb{R}$ , according to  $\mu$ , is given by

$$\int_{\phi|_{(a,b)}} d\mu(\Gamma) = \int_{\phi|_{(a,b)}} f(\Gamma)d\Gamma \tag{3}$$

---

<sup>1</sup>This is analogous to the equilibrium thermodynamic requirement of a unique allotrope.

where

$$\phi|_{(a,b)} := \{\Gamma \in \mathcal{M} : \phi(\Gamma) \in (a, b)\}, \tag{4}$$

is the set of points of the phase space for which  $\phi$  takes values in  $(a, b)$ .<sup>2</sup>

Let  $S^\tau : \mathcal{M} \rightarrow \mathcal{M}$  be the time evolution operator, which takes any point  $\Gamma \in \mathcal{M}$  to its evolved image  $S^\tau \Gamma$ , under some dynamics applied for the time  $\tau$ .<sup>3</sup> We refer to time reversal invariant dynamics, *i.e.* dynamics which obey

$$i S^\tau \Gamma = S^{-\tau} i \Gamma \quad \text{for all } \Gamma \in \mathcal{M} \text{ and all } \tau \in \mathbb{R} \tag{5}$$

where  $i : \mathcal{M} \rightarrow \mathcal{M}$ , which obeys  $ii = i^2 = \text{identity}$ , is an involution representing the time inversion operator (e.g.  $i \Gamma \equiv i(\mathbf{q}, \mathbf{p}) \equiv (\mathbf{q}, -\mathbf{p})$ ) for the dynamics. Consider the phase functions defined by:

$$\begin{aligned} \bar{\phi}_{t_0, t_0+\tau}(\Gamma) &:= \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \phi(S^s \Gamma) ds, & \phi_{t_0, t_0+\tau}(\Gamma) &:= \tau \bar{\phi}_{t_0, t_0+\tau}(\Gamma), \\ \text{for } t_0, \tau \in \mathbb{R}; \Gamma \in \mathcal{M}. \end{aligned} \tag{6}$$

Take  $\delta > 0$ ,  $t_0 = 0$ , introduce the sets  $A_\delta^+ = (A - \delta, A + \delta)$ ,  $A_\delta^- = (-A - \delta, -A + \delta)$  and, for an odd phase variable  $\phi$  (*i.e.*  $\phi(i\Gamma) = -\phi(\Gamma)$ ), consider the ratio

$$\frac{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^-})} = \frac{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(\Gamma) d\Gamma}{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^-}} f(\Gamma) d\Gamma}, \tag{7}$$

*i.e.* the probability according to  $\mu$  that  $\bar{\phi}_{0,\tau} \in A_\delta^+$ , divided by the probability according to the same  $\mu$  that  $\bar{\phi}_{0,\tau} \in A_\delta^-$ .

To compute this quantity, observe that the points of  $\mathcal{M}$  which fall in  $\bar{\phi}_{0,\tau}|_{A_\delta^-}$  are those, and only those, obtained by doing the time inversion of the evolution, for a time  $\tau$ , of the points in  $\bar{\phi}_{0,\tau}|_{A_\delta^+}$ , *i.e.*

$$\bar{\phi}_{0,\tau}|_{A_\delta^-} = i S^\tau \bar{\phi}_{0,\tau}|_{A_\delta^+}. \tag{8}$$

Indeed, take any  $\Gamma \in \bar{\phi}_{0,\tau}|_{A_\delta^-}$ , invert it and evolve it backward for a time  $\tau$ ; the resulting point  $Y = S^{-\tau} i \Gamma$  is the initial condition of a trajectory segment of duration  $\tau$  over which one obtains  $\bar{\phi}_{0,\tau} \in A_\delta^+$ , hence  $Y \in \bar{\phi}_{0,\tau}|_{A_\delta^+}$ . Moreover, there are no points of  $i S^\tau \bar{\phi}_{0,\tau}|_{A_\delta^+}$  which do not lie in  $\bar{\phi}_{0,\tau}|_{A_\delta^-}$ . This allows us to compute the denominator of (7) in terms of  $\bar{\phi}_{0,\tau}|_{A_\delta^+}$ , through the coordinate transformation

$$\Gamma = i S^\tau X, \quad \text{whose Jacobian is } J = \left| \frac{d\Gamma}{dX} \right| = \exp\left(\int_0^\tau \Lambda(S^s X) ds\right) = e^{\Lambda_{0,\tau}(X)}. \tag{9}$$

<sup>2</sup> $\phi|_{(a,b)}$  is assumed to be  $\mu$ -measurable, like all sets in this paper.

<sup>3</sup>In particular, the dynamics could be determined by the equations of motion of a particle system subjected to a dissipative field  $F_e$ , and to a deterministic thermostat meant to remove the excess energy pumped into the system by the field, so that a nonequilibrium steady state can be reached [41]. In this case,  $\Gamma = (\mathbf{q}, \mathbf{p})$  denotes the generalized coordinates and momenta of all the particles comprising the system.

Here, the quantity  $\Lambda$  is the phase space expansion rate (the opposite of the contraction rate) which, for dynamics  $\dot{\Gamma} = G(\Gamma)$  on  $\mathcal{M}$ , is defined by

$$\Lambda(\Gamma) = \nabla \cdot \dot{\Gamma} = \nabla \cdot G(\Gamma). \tag{10}$$

This leads to:

$$\int_{\bar{\phi}_{0,\tau}|_{A_\delta^-}} f(\Gamma)d\Gamma = \int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(iS^\tau X) \left| \frac{d\Gamma}{dX} \right| dX = \int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(iS^\tau X)e^{\Lambda_{0,\tau}(X)} dX. \tag{11}$$

A necessary condition for the derivation of the transient fluctuation relations is that

$$\int_{\bar{\phi}_{0,\tau}|_{A_\delta^-}} f(\Gamma)d\Gamma \neq 0 \quad \text{for all } A \text{ for which } \int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(\Gamma)d\Gamma \neq 0.$$

This condition is satisfied by appropriate choices of  $\mu$ , for the given  $S^\tau$ . For example,  $\mu$  could be selected to be the probability distribution generated by the field-free (equilibrium) dynamics of a particle system,  $S_0^\tau$  say, for which  $S^\tau$  is the nonequilibrium dynamics; *i.e.*  $\mu$  could be generated by  $S_0^\tau$ , obtained by setting the dissipative field of the nonequilibrium dynamics  $S^\tau$  to zero. In fact, this selection ensures that  $f(iS^\tau \Gamma) \neq 0$  whenever  $f(\Gamma) \neq 0$ .<sup>4</sup> Then, assuming that  $f$  is also time reversal invariant<sup>5</sup> (*i.e.* that  $f(\Gamma) = f(i\Gamma)$  for every  $\Gamma \in \mathcal{M}$ ) the ratio of probabilities (7) becomes:

$$\begin{aligned} \frac{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^-})} &= \frac{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(\Gamma)d\Gamma}{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(iS^\tau X) \exp(\Lambda_{0,\tau}(X))dX} \\ &= \frac{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(\Gamma)d\Gamma}{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(S^\tau X) \exp(\Lambda_{0,\tau}(X))dX}. \end{aligned} \tag{12}$$

The restriction on the choice of  $\mu$  for the given  $S^\tau$ , *i.e.* that  $f(iS^\tau \Gamma) \neq 0$  whenever  $f(\Gamma) \neq 0$ , is referred to as *ergodic consistency* of  $f$  with  $S^\tau$  [1]. The main quantity used below can now be introduced.

**Definition** The time averaged *dissipation function*  $\bar{\Omega}_{t_0,t_0+\tau} : \mathcal{M} \rightarrow \mathbb{R}$ , for a time reversal invariant phase space probability density  $f$  is defined by [1]:

$$\begin{aligned} \bar{\Omega}_{t_0,t_0+\tau}(\Gamma) &:= \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \Omega(S^s \Gamma)ds := \frac{1}{\tau} \left[ \ln \frac{f(S^{t_0} \Gamma)}{f(iS^{t_0+\tau} \Gamma)} - \int_{t_0}^{t_0+\tau} \Lambda(S^s \Gamma)ds \right] \\ &= \frac{1}{\tau} \left[ \ln \frac{f(S^{t_0} \Gamma)}{f(S^{t_0+\tau} \Gamma)} - \Lambda_{t_0,t_0+\tau}(\Gamma) \right] \end{aligned} \tag{13}$$

provided  $f$  is ergodically consistent with  $S^t$ . The instantaneous dissipation function  $\Omega(\Gamma)$  is obtained from differentiation with respect to  $\tau$  of  $\bar{\Omega}_{0,\tau}$ , for sufficiently regular  $f$ .

<sup>4</sup>Here, it is assumed that equilibrium dynamics corresponds to no net dissipation, and that typical trajectories explore all the phase space.

<sup>5</sup>The assumption that  $f$  be time reversal invariant is not necessary for our main results. A different choice will simply result in an alternative definition of the dissipation function, [1]. However, if  $\mu$  is selected to be an equilibrium measure, then this condition is guaranteed, and relying on it makes the calculations more elegant.

Here the condition of ergodic consistency plays an important role, although it may not be so evident in general. Indeed, given a particle system with certain conserved quantities under its equilibrium dynamics  $S_0^\tau$ , the nonequilibrium dynamics  $S^\tau$  may not conserve the same quantities, and a trajectory  $S^\tau \Gamma$  may wander outside the support of  $\mu$ , making the corresponding  $\Omega$  unusable in the following derivations. Thus, in physical applications  $f$  must be chosen with care, although many phase space densities are acceptable from a mathematical point of view. Using (13), (12) implies what has been referred to as the

**Transient  $\phi$ -FR:**

$$\frac{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^-})} = \frac{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} f(\Gamma) d\Gamma}{\int_{\bar{\phi}_{0,\tau}|_{A_\delta^+}} \exp[-\Omega_{0,\tau}(X)] f(X) dX} := \langle \exp(-\Omega_{0,\tau}) \rangle_{\bar{\phi}_{0,\tau} \in A_\delta^+}^{-1}. \tag{14}$$

Here  $\langle \exp(-\Omega_{0,\tau}) \rangle_{\bar{\phi}_{0,\tau} \in A_\delta^+}$  is the ensemble average of  $\exp[-\Omega_{0,\tau}(X)]$ , over the set of trajectories which satisfy the constraint that  $\bar{\phi}_{0,\tau} \in A_\delta^+$  [1, 12]. In the case that  $\bar{\phi}_{0,\tau} = \bar{\Omega}_{0,\tau}$  the fluctuation relation assumes the particularly elegant form, called

**Transient  $\Omega$ -FR:**

$$\frac{\mu(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} = \langle \exp(-\Omega_{0,\tau}) \rangle_{\bar{\Omega}_{0,\tau} \in A_\delta^+}^{-1} = e^{[A+\varepsilon(\delta,A,\tau)]\tau}. \tag{15}$$

Here,  $\varepsilon$  is an error term of magnitude  $|\varepsilon| \leq \delta$ , which depends on  $\delta, A$  and  $\tau$ , and appears because  $\bar{\Omega}_{0,\tau}$  is not necessarily constant and equal to  $A$ . Therefore,  $\varepsilon(\delta, A, \tau) \rightarrow 0$  as  $\delta \rightarrow 0$ , for all  $A$  for which the ratio of (15) exists [1, 10, 11].

*Remark* Different choices of  $f$  are possible, which lead to different  $\Omega$ -FR’s. In particular, equilibrium probability densities of many-particle systems which obey  $AIF$  lead to an  $\Omega$ -FR which concerns the entropy production, or the energy dissipation rate  $\Sigma$  (cf. Appendix 1). The uniform density in a compact phase space,  $f(\Gamma) = 1/|\mathcal{M}|$  say, yields  $\Omega = -A$ . A density such that  $f d\Gamma$  approximates as accurately as desired the (possibly singular) steady state measure is also allowed.

The above relations are called “transient” because, independently of the length of  $\tau$ , they express properties of the measure  $\mu$  and not of the possible steady state. The time  $\tau$ , indeed, enters the  $\phi$ -FR only in the definition of the observable  $\bar{\phi}_{0,\tau}$ , and not of the phase space probability density. Equations (14), (15) are exact and do not need any limit to be taken either on  $\tau$  or on  $\delta$ . As the calculation shows, the only properties that are required for (14), (15) to hold are that the dynamics  $S^\tau$  be time reversal invariant, to ensure that, for odd  $\phi$ ,  $-A$  can be observed if  $A$  can, and that  $f(iS^\tau \Gamma) \neq 0$  for any  $\Gamma$  for which  $f(\Gamma) \neq 0$ . The range of  $A$ ’s that can be observed depends both on  $S^\tau$  and  $\phi$ .

If the initial distribution can be generated by a single field free phase space trajectory, and  $AIF$  holds, the dissipation function  $\Omega$  takes the form given in (2) and at equilibrium (where  $F_e = 0$ ),  $\Omega = 0$ , hence

$$\frac{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^-})} = \langle e^0 \rangle_{\bar{\phi}_{0,\tau} \in A_\delta^+}^{-1} = 1 \tag{16}$$



as it should be for the equilibrium measure  $\mu$ , and for any odd variable  $\phi$  (for instance,  $\phi$  could be the instantaneous current  $J$ , which does not need to vanish at equilibrium, although its mean certainly does). For symmetric intervals  $(-\delta, \delta)$ , one also has  $A = 0$  and can write

$$\frac{\mu(\bar{\phi}_{0,\tau}|_{(-\delta,\delta)})}{\mu(\bar{\phi}_{0,\tau}|_{(-\delta,\delta)})} = 1, \quad \text{hence} \quad \langle e^{-\Omega_0,\tau} \rangle_{\bar{\phi}_{0,\tau} \in (-\delta,\delta)} = 1. \tag{17}$$

In the limit that  $\delta$  tends to infinity, the second equality in (17) tends to a full phase space average, and produces the NonEquilibrium Partition Identity (also referred to as the Kawasaki normalization)[11, 41]. This identity can be used to test the accuracy of the numerical simulations, and has even been used to calibrate experimental equipment [36–38]. The highly asymmetric convergence of numerical estimates to unity means that estimating the statistical uncertainty of experiments is problematic without this relationship [11]. If the dissipation function is bounded (i.e.  $|\Omega| \leq \Omega^*$ , for some  $\Omega^* > 0$ ), one obtains

$$e^{-\tau\Omega^*} \leq \frac{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{0,\tau}|_{A_\delta^-})} \leq e^{\tau\Omega^*}, \tag{18}$$

for all odd  $\phi$ . Because different dissipation functions are allowed and the above ratio holds for all of them, a sharp bound can be given for all  $\phi$  and independently of the physically relevant dissipation function, by taking the smallest possible  $\Omega^*$ . This constitutes a new prediction, which may be interesting to test in nonequilibrium systems.

### 3 Asymptotic Fluctuation Relations

In order to derive steady state fluctuation relations, one may develop further the transient  $\phi$ -FR’s obtained in Sect. 2, considering the time averaged phase variable  $\bar{\phi}_{t_0,t_0+\tau}$  (cf. (6)), where the time averaging begins at a time  $t_0 > 0$ , rather than at time 0:

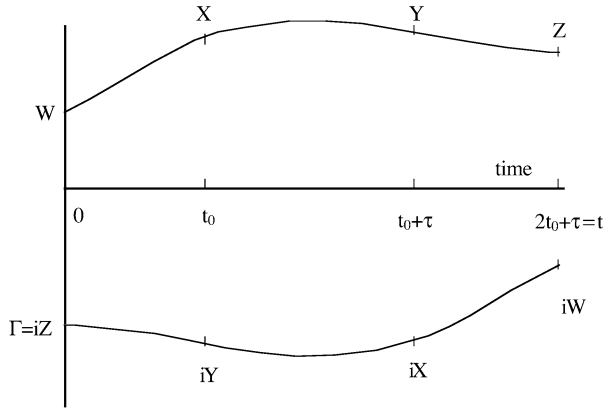
$$\frac{\mu(\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+})}{\mu(\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-})} = \frac{\int_{\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+}} f(\Gamma)d\Gamma}{\int_{\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-}} f(\Gamma)d\Gamma}. \tag{19}$$

Here,  $\bar{\phi}_{t_0,t_0+\tau}|_E = \{\Gamma \in \mathcal{M} : \bar{\phi}_{t_0,t_0+\tau}(\Gamma) \in E\}$  is the set of all phase points which are initial conditions of trajectory segments of duration  $t_0 + \tau$ , over which  $\bar{\phi}_{t_0,t_0+\tau}$  takes values in  $E$ , and  $E = A_\delta^+$  or  $A_\delta^-$ . The calculation of the ratio (19) requires that the points of  $\mathcal{M}$  lying in  $\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-}$  be identified. To do that, consider a trajectory of length  $t = 2t_0 + \tau$ , and take  $W$  in  $\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+}$ . Then, let  $X = S^{t_0}W$ ,  $Y = S^{t_0+\tau}W$ ,  $Z = S^{2t_0+\tau}W = S^tW$ , and  $\Gamma = iZ = iS^tW$  (see Fig. 1). The trajectory segment between  $iY$  and  $iX$  produces the opposite of the average produced between  $X$  and  $Y$ , in the time interval  $(t_0, t_0 + \tau)$ , so that

$$\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-} = iS^t\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+} \tag{20}$$

and, as in the previous section,

**Fig. 1** Schematic diagram illustrating the time-reversible evolution of two points in phase space that are related by a time reversal mapping,  $i$ . If the average of  $\phi$  between  $X$  and  $Y$  is  $A$ , the average between  $iY$  and  $iX$  is  $-A$



$$\int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^-}} f(\Gamma) d\Gamma = \int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+}} f(iS^t W) \left| \frac{d\Gamma}{dW} \right| dW$$

$$= \int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+}} f(S^t W) \exp(\Lambda_{0,t}(W)) dW \tag{21}$$

where  $J = \left| \frac{d\Gamma}{dW} \right| = \exp[\Lambda_{0,t}(W)]$  and we use time reversal invariance of  $f$ .

Note that the integral defining  $J$  is along the full trajectory of duration  $t$ , although  $\bar{\phi}_{t_0, t_0+\tau}$  concerns only one of its parts of duration  $\tau$ . From the definition of  $\bar{\Omega}_{0,t}$  (see (13)), one obtains:

$$\Lambda_{0,t}(W) = \ln \frac{f(W)}{f(S^t W)} - \Omega_{0,t}(W) \tag{22}$$

and

$$\int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^-}} f(\Gamma) d\Gamma = \int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+}} f(W) e^{-\Omega_{0,t}(W)} dW. \tag{23}$$

The probability ratio of (19) can then be written as:

$$\frac{\mu(\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+})}{\mu(\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^-})} = \left[ \frac{\int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+}} f(W) e^{-\Omega_{0,t}(W)} dW}{\int_{\bar{\phi}_{t_0, t_0+\tau} |_{A_\delta^+}} f(W) dW} \right]^{-1} := \langle \exp(-\Omega_{0,t}) \rangle_{\bar{\phi}_{t_0, t_0+\tau} \in A_\delta^+}^{-1} \tag{24}$$

which is the inverse of the average of  $\exp(-\Omega_{0,t})$ , conditioned to the constraint  $\bar{\phi}_{t_0, t_0+\tau} \in A_\delta^+$ ; and the special case  $\phi = \Omega$ , yields

$$\frac{\mu(\bar{\Omega}_{t_0, t_0+\tau} |_{A_\delta^+})}{\mu(\bar{\Omega}_{t_0, t_0+\tau} |_{A_\delta^-})} = \langle \exp(-\Omega_{0,t}) \rangle_{\bar{\Omega}_{t_0, t_0+\tau} \in A_\delta^+}^{-1}. \tag{25}$$

The right hand side of this expression can be written as:

$$\langle \exp(-\Omega_{0,t}) \rangle_{\bar{\Omega}_{t_0, t_0+\tau} \in A_\delta^+}^{-1} = e^{[A+\varepsilon(\delta, t_0, A, \tau)]\tau} \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0, t_0+\tau} \in A_\delta^+}^{-1}. \tag{26}$$

Taking the logarithm, and dividing by  $\tau$  one then obtains:

$$\frac{1}{\tau} \ln \frac{\mu(\bar{\Omega}_{t_0, t_0+\tau}^+ | A_\delta^+)}{\mu(\bar{\Omega}_{t_0, t_0+\tau}^- | A_\delta^-)} = A + \varepsilon(\delta, t_0, A, \tau) - \frac{1}{\tau} \ln \langle e^{-\Omega_{0, t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0, t_0+\tau} \in A_\delta^+} \tag{27}$$

where the correction term  $\varepsilon$  depends on  $\delta, t_0, A$  and  $\tau$  and, like in Sect. 2,  $|\varepsilon| \leq \delta$ . This result is exact, and holds for all  $t_0, \tau$  and  $\delta$ , and for any  $A$  for which both the probabilities in the left hand side of (27) do not vanish. It rests only on the time reversibility of  $S^t$ , and on the ergodic consistency of  $f$  with  $S^t$ .

Equations (24) and (27) are valid for all  $\tau$ , hence also in the limit  $\tau \rightarrow \infty$ , but remain transient relations because no matter how large  $\tau$  is; they refer to the initial measure  $\mu$ , even if the dynamics makes  $\mu$  evolve into  $\mu_\tau$ . However, the above derivation suggests that for systems that reach a steady state, it may be possible to derive steady state FR’s from the transient FR’s. The idea is that for sufficiently large times  $t_0$ , (24), (27) should concern trajectory segments of length  $\tau$  distributed according to the invariant measure, or very close to that. Furthermore, taking  $\tau$  long, and  $\delta$  small, should make the last two terms in (27) negligible.

To investigate this possibility, we firstly transform (24) and (27) by transferring the time evolution up to  $t_0$  from the sets of phase space points  $\bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}$ , to the phase space probability distributions. Given an initial phase space probability density  $f_0$ , and the corresponding probability measure  $\mu_0$ , we denote their evolved counterparts after a time  $t_0$  by  $f_{t_0}$  and  $\mu_{t_0}$  respectively:

$$\mu_{t_0}(S^{t_0} E) = \mu_0(E) \quad \text{or equivalently} \quad \int_{S^{t_0} E} f_{t_0}(X) dX = \int_E f_0(W) dW \tag{28}$$

for every  $t_0 \in \mathbb{R}$  and  $E \subset \mathcal{M}$ , where  $S^{t_0} E$  is the set of points reached by all points of  $E$  after an evolution of duration  $t_0$ . Equations (28) are consequences of the conservation of probability by time evolution.<sup>6</sup> Physically, they are an instance of what is referred to as the *Heisenberg–Schrödinger equivalence*, while mathematically they are the result of the coordinate transformations  $X = S^{t_0} W$  and  $W = S^{-t_0} X$ . The Jacobian of this transformation is given by

$$\left| \frac{\partial W}{\partial X} \right| = e^{A_{0, -t_0}(X)} = e^{\int_0^{-t_0} \Lambda(S^s X) ds} = e^{-\int_{-t_0}^0 \Lambda(S^s X) ds} = e^{-\Lambda_{-t_0, 0}(X)} \tag{29}$$

hence one can write

$$\begin{aligned} \mu_0(\bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}) &= \int_{\bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}} f_0(W) dW \\ &= \int_{S^{t_0} \bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}} f_0(S^{-t_0} X) e^{-\Lambda_{-t_0, 0}(X)} dX \end{aligned} \tag{30}$$

and

$$\mu_0(\bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}) = \mu_{t_0}(S^{t_0} \bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}) = \int_{S^{t_0} \bar{\phi}_{t_0, t_0+\tau} |_{(a,b)}} f_{t_0}(X) dX. \tag{31}$$

<sup>6</sup>They merely state that the initial probability of  $E$  becomes the probability of the evolved set  $S^{t_0} E$  after a time  $t_0$ .

In (30), (31), we could have integrated over any measurable set  $E$ , hence, equating (30) and (31) finally gives:

$$f_{t_0}(X) = f_0(S^{-t_0} X)e^{-\Lambda_{-t_0,0}(X)}. \tag{32}$$

This is known as the Lagrangian form of the Kawasaki distribution function [11].<sup>7</sup> From (22), which yields

$$-\Lambda_{-t_0,0}(X) = \int_0^{-t_0} \Lambda(S^s X) ds = \ln\left(\frac{f(X)}{f(S^{-t_0} X)}\right) - \Omega_{0,-t_0}(X), \tag{33}$$

(32) can be rewritten as

$$f_{t_0}(X) = f_0(X)e^{-\Omega_{0,-t_0}(X)}, \tag{34}$$

and gives the evolution of the phase space probability density  $f_0$ . This well known exact result for the time dependent phase space density  $f_{t_0}(X)$ , was first derived for thermostatted particle dynamics where the dissipation function takes the form given in (2) by Evans and Morriss in 1984 [53]. Observe that

$$S^{t_0} \bar{\phi}_{t_0,t_0+\tau}|_{(a,b)} = \bar{\phi}_{0,\tau}|_{(a,b)} \tag{35}$$

because  $\phi$  is a phase variable, hence does not explicitly depend on time, and because we assume the dynamics to be autonomous. Indeed, a point  $X \in S^{t_0} \bar{\phi}_{t_0,t_0+\tau}|_{(a,b)}$  is the evolution by a time  $t_0$  of a point  $\Gamma$  for which  $\bar{\phi}_{t_0,t_0+\tau}$  takes the same value of  $\bar{\phi}_{0,\tau}(X)$ , i.e.

$$\bar{\phi}_{0,\tau}(S^{t_0} \Gamma) = \frac{1}{\tau} \int_0^\tau \phi(S^s(S^{t_0} \Gamma)) ds = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \phi(S^s \Gamma) ds = \bar{\phi}_{t_0,t_0+\tau}(\Gamma). \tag{36}$$

Thus, (31) expresses the time evolving probability of the fixed set  $\bar{\phi}_{0,\tau}|_{(a,b)}$ ,

$$\mu_{t_0}(\bar{\phi}_{0,\tau}|_{(a,b)}) = \mu_{t_0}(S^{t_0} \bar{\phi}_{t_0,t_0+\tau}|_{(a,b)}). \tag{37}$$

Provided convergence to the steady state occurs, this equation yields a more and more accurate approximation to the steady state measure of  $\bar{\phi}_{0,\tau}|_{(a,b)}$ , as  $t_0$  grows, at fixed  $\tau$ .

Equation (24), which is expressed in terms of the initially chosen phase space probability distribution, and of evolving phase space sets, can now be given in terms of the evolving PDF, since

$$\frac{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^-})} = \frac{\mu_{t_0}(S^{t_0} \bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+})}{\mu_{t_0}(S^{t_0} \bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-})} = \frac{\mu_0(\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^+})}{\mu_0(\bar{\phi}_{t_0,t_0+\tau}|_{A_\delta^-})}. \tag{38}$$

This produces what we call

**$\phi$ -FR:**

$$\frac{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^-})} = \langle \exp(-\Omega_{0,t}) \rangle_{\bar{\phi}_{t_0,t_0+\tau} \in A_\delta^+}^{-1}. \tag{39}$$

Relation (39) is exact and holds for all  $t_0, \tau$  and possible pairs  $A$  and  $-A$  (whose range must be determined case by case). Letting  $\bar{\phi}_{t_0,t_0+\tau} = \bar{\Omega}_{t_0,t_0+\tau}$  and rearranging, (27) yields the

<sup>7</sup>This can also be written,  $f_{t_0}(S^{t_0} W) = f_0(W)e^{-A_{0,t_0}(W)}$ .

**$\Omega$ -FR:**

$$\frac{1}{\tau} \ln \frac{\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} = A + \varepsilon(\delta, t_0, A, \tau) - \frac{1}{\tau} \ln \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+}, \tag{40}$$

where the conditional ensemble average in the last term is computed with respect to the initial distribution  $f$ .

Let us consider some direct consequences of the above relations. For  $A = 0$ ,  $\delta > 0$ ,  $t_0, \tau \in \mathbb{R}$ , and any odd  $\phi$ , the conditional average of (39) yields

$$\langle \exp(-\Omega_{0,t}) \rangle_{\bar{\phi}_{t_0,t_0+\tau} \in (-\delta, \delta)} = \frac{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{(-\delta, \delta)})}{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{(-\delta, \delta)})} = 1 \tag{41}$$

which, in the  $\delta \rightarrow \infty$  limit, leads again to the NonEquilibrium Partition Identity (17), for the full ensemble average.

Taking  $A = 0$  in (40), one obtains

$$\langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in (-\delta, \delta)} = e^{\varepsilon(\delta, t_0, A=0, \tau)\tau} \tag{42}$$

which shows that this particular conditional average is uniformly bounded in  $t_0$ , and tends to 1 when  $\delta \rightarrow 0$ , at fixed  $\tau$ . Equation (41) with  $\phi = \Omega$  differs from (42) because of the integration times in the argument of the exponential function. For fixed  $\delta$  and  $\tau$ , such that  $\delta\tau \ll 1$ , one can write

$$\langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in (-\delta, \delta)} \approx 1, \tag{43}$$

even for  $t_0 \rightarrow \infty$ , which is interesting for the validity of the steady state  $\Omega$ -FR. Indeed, the last term in (40) ought not to dominate over the others, if the steady state  $\Omega$ -FR is to hold. Exchanging the roles of  $A_\delta^+$  and  $A_\delta^-$  in (40), one obtains

$$\langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^-} = \frac{e^{\varepsilon(\delta, t_0, A, \tau) + \varepsilon(\delta, t_0, -A, \tau)}}{\langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+}}, \tag{44}$$

for any  $\delta > 0$ .

One last property that is worthwhile to consider is the following asymptotic result, which is particularly close to a steady state FR. As it happens in many situations of interest, let  $|\Omega|$  be bounded by some  $\Omega^* > 0$ , so that:

$$|\ln \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+}| \leq 2t_0\Omega^* \tag{45}$$

for all  $\tau$ . Then, taking  $\delta < \gamma$ , which implies  $|\varepsilon| < \gamma$ , and

$$\tau(t_0) \geq \frac{2t_0\Omega^*}{\gamma - \delta} = \tau^*(t_0), \tag{46}$$

one obtains the following

**$\Omega_\infty$ -FR:**

$$A - \gamma \leq \lim_{t \rightarrow \infty} \frac{1}{\tau(t)} \ln \frac{\mu_t(\bar{\Omega}_{0,\tau(t)}|_{A_\delta^+})}{\mu_t(\bar{\Omega}_{0,\tau(t)}|_{A_\delta^-})} \leq A + \gamma. \tag{47}$$

Recalling the connection between  $\Omega$  and the physical dissipation, which holds under proper choices of the initial  $f$ , the above condition can be written as

$$\tau \geq \frac{2t_0}{\gamma - \delta} \cdot \frac{J^* F_e V}{k_B T}, \tag{48}$$

where  $J^*$  is the maximum of the flux. This shows that, differently from the case of the steady state  $\Lambda$ -FR, the convergence time of the  $\Omega_\infty$ -FR does not increase when the dissipative field tends to 0; to the contrary, it may decrease.

As in the transient case, an asymptotic  $\phi$ -FR holds for  $A$  in a range which depends on the dynamics and on the observable, but for large  $\tau$  it remains close to  $(-\Omega^*, \Omega^*)$ . Indeed, one has

$$-\frac{2t + \tau}{\tau} \Omega^* \leq \frac{1}{\tau} \ln \frac{\mu_t(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu_t(\bar{\phi}_{0,\tau}|_{A_\delta^-})} \leq \frac{2t + \tau}{\tau} \Omega^* \tag{49}$$

and taking any  $\gamma > 0$  and letting  $t$  grow, with  $\tau(t) \gg t$ , one can write

$\phi_\infty$ -FR:

$$-\Omega^* - \gamma \leq \lim_{t \rightarrow \infty} \frac{1}{\tau(t)} \ln \frac{\mu_t(\bar{\phi}_{0,\tau(t)}|_{A_\delta^+})}{\mu_t(\bar{\phi}_{0,\tau(t)}|_{A_\delta^-})} \leq \Omega^* + \gamma, \tag{50}$$

for any odd  $\phi$ , and with the smallest  $\Omega^*$ , among the upper bounds of the allowed dissipation functions. Because at equilibrium the dissipation function given by (2) is  $\Omega = -F_e J V / k_B T = 0$ , one has

$$-\gamma \leq \frac{1}{\tau} \ln \frac{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^+})}{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^-})} \leq \gamma, \tag{51}$$

whatever odd variable  $\phi$  is considered,  $\Omega$ ,  $\Lambda$  or  $J$  for instance. This yields the usual equilibrium symmetry, in the full range of  $J$ , not just in 0, and without the need for long convergence times.

Given the dynamics  $S^\tau$ , either there is a unique allowed density  $f_0$ , hence a unique dissipation function  $\Omega$  or, given a second initial density  $\tilde{f}_0$ , there is a corresponding  $\tilde{\Omega} \neq \Omega$ . If this is the case, and the attractor for both  $f_0$  and  $\tilde{f}_0$  is the same, one can write

$$\langle \exp(-\tilde{\Omega}_{0,t}) \rangle_{\bar{\phi}_{t_0,t_0+\tau} \in A_\delta^+} \approx \frac{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^-})}{\mu_{t_0}(\bar{\phi}_{0,\tau}|_{A_\delta^+})} \approx \langle \exp(-\Omega_{0,t}) \rangle_{\bar{\phi}_{t_0,t_0+\tau} \in A_\delta^+}, \tag{52}$$

for large  $t_0$ , all  $\tau$  and all allowed  $A$ . Then, as far as the asymptotic fluctuation relations are concerned, the different dissipation functions are equivalent.

Let us note that (47) does not necessarily imply the steady state  $\Omega$ -FR, even when there is a single steady state and all initial measures  $\mu_0$  with density  $f_0$  are attracted by the same  $\mu_\infty$ , because the limit in (50) could be singular. Although this is not impossible, it is not the typical situation of particle systems of physical interest (cf. Sect. 4 and Appendix 2), and (47) is a result rather close to the steady state  $\Omega$ -FR.

In Sect. 4 we discuss the conditions under which (47) describes the statistics of trajectory segments sampled from a single steady state trajectory. Before we do this, we note that whereas in Sects. 2 and 3, the initial distribution  $f_0$  can have arbitrary form, provided it is

ergodically consistent with the dynamics considered, a further condition on  $f_0$  is imposed, in order to obtain the steady state  $\Omega$ -FR: the decay of the  $\Omega$ -autocorrelation, with respect to  $f_0$ . This is necessary for the initial state to relax to a steady state. The further requirement that the ensemble is not only invariant under  $S_0^t$ , but that is also generated by a single trajectory, has the purpose to avoid the presence of distinct steady states, selected by different choices of the initial conditions. These states, indeed, would have different properties, and a unique  $\mu_\infty$  would not be possible. For instance, if  $f_0$  was selected to be a uniform distribution (so that  $\Omega = -A$ ) and isokinetic dynamics was carried out, a distinct  $\mu_\infty$  would be expected for each distinct value of the kinetic energy.

If the system under investigation does not converge to a steady state with positive and negative fluctuations, there is no need for a steady state relation, but the  $\Omega$ -FR, the  $\phi$ -FR, the  $\Omega_\infty$ -FR and the  $\phi_\infty$ -FR remain valid for the given ensemble of trajectories.

#### 4 Steady State Fluctuation Relations

We assume the steady state exists and that it is unique. This will not always be the case, for example if we start from a canonical phase space distribution at time zero and have isokinetic dynamics subsequently, there will be distinct steady state attractors for each trajectory with a different kinetic energy. Each of the steady state attractors will have different properties because they are at different kinetic temperatures. By saying that the steady state is unique we mean that, apart from a set of measure zero, the statistics generated by each trajectory is independent of the initial phase, in the relevant equilibrium phase space.

Equations (27) and (40) are both exact, but contain information on the trajectories before the steady state is reached. To be of practical use in the description of the steady state fluctuations of the dissipation function, the last two terms in the equation have to be negligible. Therefore, for  $A \neq 0$ , some assumption on the nature of the logarithmic term of (40) is necessary since, in principle, this term could be of order  $O(t_0/\tau)$ , and  $t_0$  should tend to infinity, in order to reach the steady state, before  $\tau$  does [40]. We argue that, in most cases of interest, the dynamics of particle systems used to model nonequilibrium fluids satisfy the following:

**Steady state  $\Omega$ -FR** For any tolerance  $\gamma > 0$ , there exists  $\delta_\gamma, \tau_\gamma > 0$  such that

$$A - \gamma \leq \frac{1}{\tau} \ln \frac{\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} \leq A + \gamma \tag{53}$$

holds for  $0 < \delta \leq \delta_\gamma$  and  $\tau \geq \tau_\gamma$ , if both the values  $A$  and  $-A$  are  $\delta$ -possible.

Here, the following definition has been used:

**Definition** The value  $A \in \mathbb{R}$  is called  $\delta$ -possible if, given  $\delta > 0$ , there exists  $\tau_{A,\delta} > 0$  such that

$$\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^+}) > 0$$

for all  $\tau > \tau_{A,\delta}$ .<sup>8</sup>

<sup>8</sup>Because of the finiteness of  $\tau$ , this can be the case even though  $\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^+}) \rightarrow 0$ , for  $\tau \rightarrow \infty$ , as it does if the steady state ensemble average of  $\Omega$ ,  $\langle \Omega \rangle_\infty$  say, does not belong to  $A_\delta^+$ .

It follows that, for  $A$  to be  $\delta$ -possible,  $\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^+})$  must decay with  $\tau$  more slowly than the error term  $[\varepsilon(\delta, t_0, A, \tau) - \frac{1}{\tau} \ln \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{0,t_0+\tau} \in A_\delta^+}]$  of (40). Indeed, given  $\delta > 0$ , a pair  $(-A, A)$  and a fixed  $\tau$ , the logarithmic term in the right hand side of (40) may either take arbitrarily large values or be bounded, in the  $t_0 \rightarrow \infty$  limit. In the first case, because  $A, \delta$  and  $\tau$  are fixed real numbers and  $|\varepsilon| \leq \delta$ , the corresponding divergence is reflected on the measures  $\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^+})$  and  $\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^-})$ , one of which at least either tends to zero or does not converge at all. Indeed, if for every  $R, \tau, \hat{t}_0 > 0$  there is  $t_0 > \hat{t}_0$  such that

$$R < \frac{\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^-})}, \tag{54}$$

one has

$$R < \frac{1}{\mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^-})}, \quad \text{or} \quad \mu_{t_0}(\bar{\Omega}_{0,\tau}|_{A_\delta^-}) < \frac{1}{R} \tag{55}$$

which corresponds either to  $\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^-}) = 0$ , or to the  $\mu_\infty$  non-measurability of  $\bar{\Omega}_{0,\tau}|_{A_\delta^-}$ . In either case,  $-A$  is not  $\delta$ -possible, and similarly one may treat  $A$ .

Clearly,  $A$  is of no interest for a steady state FR, i.e. it does not belong to the domain of the steady state  $\Omega$ -FR with tolerance  $\gamma$ , if  $A$  and  $-A$  are not both  $\delta$ -possible; in a number of cases that domain may even be empty.<sup>9</sup> Therefore, we focus only on the  $A$  values with bounded logarithmic term in (40), and let  $M(A, \delta, \tau)$  be the real number to which the conditional average in (40) tends, in the  $t_0 \rightarrow \infty$  limit. We may then write

$$A - \frac{1}{\tau} \ln M(A, \delta, \tau) - \delta \leq \frac{1}{\tau} \ln \frac{\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\mu_\infty(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} \leq A - \frac{1}{\tau} \ln M(A, \delta, \tau) + \delta. \tag{56}$$

If  $(1/\tau) \ln M(A, \delta, \tau)$  tends to zero, for the given  $A, \delta$  and growing  $\tau$ , the steady state  $\Omega$ -FR holds for  $A$ . This is the case if correlations decay, so that the thermodynamic behavior sets in. In particular, if the  $\Omega$ -autocorrelation function is a  $\delta$ -function (as e.g. in Refs. [49, 50]), one can write:

$$\langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{0,t_0+\tau} \in A_\delta^+} = \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle = \langle e^{-\Omega_{0,t_0}} \rangle \langle e^{-\Omega_{t_0+\tau, 2t_0+\tau}} \rangle, \tag{57}$$

and

$$1 = \langle e^{-\Omega_{0,s} - \Omega_{s,t}} \rangle = \langle e^{-\Omega_{0,s}} \rangle \langle e^{-\Omega_{s,t}} \rangle, \tag{58}$$

which implies

$$\langle e^{-\Omega_{s,t}} \rangle = 1 \quad \text{for all } s, t \tag{59}$$

because of the NonEquilibrium Partition Identity (17). Then,  $(1/\tau) \ln M(A, \delta, \tau)$  identically vanishes for all  $\tau$ . This is an idealized situation and, indeed, the tests performed on molecular dynamics systems,<sup>10</sup> cf. Ref. [54], indicate that typically there exists a constant  $K$  (often not

<sup>9</sup>There are various reasons for which  $A$  or  $-A$  might not be in the domain of the steady state  $\Omega$ -FR. For example,  $A$  could be larger than the largest value of  $\Omega$ , or there could be no fluctuations in the steady state (cf. Appendix 2).

<sup>10</sup>The conditional average of (40) can, of course, be computed independently of the other terms in (40), although this is not strictly necessary, because (40) is an exact identity.



too far from 1), such that

$$0 < \frac{1}{K} \leq \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \leq K, \tag{60}$$

which may be understood as follows. In the first place, the conditional average is limited for growing  $t_0$ , if  $-A, A$  are  $\delta$ -possible, and equals 1 for  $\tau = 0$ . Then, if the  $\Omega$ -autocorrelation decays with time scale  $t_M$ ,<sup>11</sup> one expects  $\ln K$  to be of order  $O(t_M)$ . Taking  $\tau$  sufficiently larger than  $t_M$ , and letting  $c_A$  be a natural number dependent on the given  $A$ , one may write

$$\begin{aligned} & \langle e^{-\Omega_{0,t_0} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \\ &= \langle e^{-\Omega_{0,t_0} - c_A t_M} e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \cdot e^{-\Omega_{t_0-c_A t_M, t_0} - \Omega_{t_0+\tau, t_0+\tau+c_A t_M}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \\ &\approx \langle e^{-\Omega_{0,t_0} - c_A t_M} \cdot e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \\ &\approx \langle e^{-\Omega_{0,t_0} - c_A t_M} \cdot e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \rangle \\ &\approx \langle e^{-\Omega_{0,t_0} - c_A t_M} \rangle \langle e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \rangle = \langle e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \rangle \end{aligned} \tag{61}$$

with an accuracy which improves with increasing  $t_0$  and  $\tau$ , if  $c_A$  is the number of Maxwell times necessary to obtain the desired decorrelation at the chosen  $A$ . Indeed, one would eventually obtain  $t_0, \tau \gg c_A t_M$ .

If (61) holds, the  $t_0 \rightarrow \infty$  limit of  $\langle e^{-\Omega_{t_0+\tau+c_A t_M, 2t_0+\tau}} \rangle$  is finite for all  $\delta$ -possible pairs  $(-A, A)$ , and is insensitive to variations of  $\tau$  above a certain threshold. Indeed, let  $\hat{\tau} = \tau + c_A t_M$  and observe that

$$\begin{aligned} \langle e^{-\Omega_{t_0+\hat{\tau}, 2t_0+\tau}} \rangle &= \int e^{-\Omega_{t_0+\hat{\tau}, 2t_0+\tau}(\Gamma)} f_0(\Gamma) d\Gamma = \int e^{-\Omega_{t_0, t_0-c_A t_M}(\Gamma)} f_{t_0+\hat{\tau}}(\Gamma) d\Gamma \\ &:= \langle e^{-\Omega_{t_0, t_0-c_A t_M}} \rangle_{t_0+\hat{\tau}} \end{aligned} \tag{62}$$

is the average of  $e^{-\Omega_{t_0, t_0-c_A t_M}}$  computed with the probability distribution obtained from the evolution of the initial one, for the time  $t = t_0 + \hat{\tau}$ . The value of this average may be estimated noting that, for  $\delta$ -possible pairs  $(-A, A)$ , and growing  $t_0$ , it converges approximately to the positive number  $M(A, \delta, \hat{\tau})$ , if (61) holds. Furthermore, for fixed  $t_0 > 0$  and for  $t$  of the order of the Maxwell time,  $\langle e^{-\Omega_{t_0, t_0-c_A t_M}} \rangle_t$  approaches its steady state mean value,  $\langle e^{-\Omega_{t_0, t_0-c_A t_M}} \rangle_\infty$  which is equal to 1 at equilibrium for all  $t_0$  and  $c_A$ . Then, for sufficiently large  $t_0$  and  $t = t_0 + \hat{\tau}$ , one may write

$$\begin{aligned} M(A, \delta, \hat{\tau}) &\approx \langle e^{-\Omega_{t_0, t_0-c_A t_M} - \Omega_{t_0+\tau, 2t_0+\tau}} \rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \approx \langle e^{-\Omega_{t_0+\hat{\tau}, 2t_0+\tau}} \rangle = \langle e^{-\Omega_{t_0, t_0-c_A t_M}} \rangle_t \\ &\approx \langle e^{-\Omega_{t_0, t_0-c_A t_M}} \rangle_\infty \end{aligned} \tag{63}$$

thanks also to the fact that  $\hat{\tau}$  is by itself larger than the decorrelation time. Because the left hand side of (63) does not depend on  $t_0$  while the right hand side does not depend on  $\tau$ , one concludes that further growths of  $t_0$  or of  $\tau$  do not cause any substantial changes in the conditional average in the middle. Close to equilibrium, this term should only have a

<sup>11</sup>This time is called Maxwell time; its order of magnitude is that of the mean free time. The Maxwell time expresses a physical property of the system and depends only mildly on the external field; usually,  $t_M(F_e) = t_M(0) + O(F_e^2)$ .

correction of order  $O(F_e^2)$  to its equilibrium average, 1. Numerical studies show that this is precisely what happens to the quantities in (60), (61), (62), [54].

Note that  $c_A$  may vary with  $A$ , as the set of trajectories involved in the conditional average changes with  $A$ ;  $c_A$  could then even diverge. However, close to equilibrium at least, the existence of the transport coefficients implies that the conditional correlation functions must not grow too much (cf. Appendix 2).

Note also that, for the steady state  $\Omega$ -FR to hold, it suffices that  $(1/\tau) \ln M(A, \delta, \tau)$  can be made as small as one wishes, by taking sufficiently large  $\tau$  and small  $\delta$ . Thus, it is not necessary that (60) holds; it suffices that  $M(A, \delta, \tau)$  grows less than exponentially fast with  $\tau$ , or even that it grows exponentially fast, with a rate which tends to zero when  $\delta$  does, as in (42), where  $A = 0$ . However, the numerical tests performed so far indicate that (60) is usually satisfied in the full range of numerically or experimentally accessible  $A$ 's.

If the  $\Omega$ -autocorrelation decays, and (60) holds with the condition on  $\Omega$  replaced by an equivalent one for another observable  $\phi$ , one obtains the

**Steady State  $\phi$ -FR** For sufficiently large  $\tau$ ,

$$\lim_{t_0 \rightarrow \infty} \frac{\ln \langle e^{-\Omega_{t_0, t_0+\tau}} \rangle_{\bar{\phi}_{t_0, t_0+\tau} \in A_\delta^+}^{-1}}{\tau} - \gamma \leq \frac{1}{\tau} \ln \frac{\mu_\infty(\bar{\phi}_{0, \tau} | A_\delta^+)}{\mu_\infty(\bar{\phi}_{0, \tau} | A_\delta^-)}$$

$$\leq \lim_{t_0 \rightarrow \infty} \frac{\ln \langle e^{-\Omega_{t_0, t_0+\tau}} \rangle_{\bar{\phi}_{t_0, t_0+\tau} \in A_\delta^+}^{-1}}{\tau} + \gamma. \tag{64}$$

This expression states that the ratio of the probabilities that  $\bar{\phi}_{0, \tau} \in A_\delta^+$  and  $\bar{\phi}_{0, \tau} \in A_\delta^-$  is determined by the average of the exponential of the values taken by  $\Omega$  in the same time intervals in which  $\phi$  has average in  $A_\delta^+$ . Because  $\phi$  can be any odd observable,  $\Omega$  appears to be a rather special function of phase.

One may wonder how the above could be reconciled with the modifications of the steady state  $\Lambda$ -FR suggested for a special kind of non-transitive dynamical systems [40, 48, 49]. As explained in Appendix 2, (60) is not in contradiction to the scenarios of Refs. [40, 48, 49], because the decay of correlations enjoyed by their dynamical systems concern the invariant measures, while the conditional average in (60) is computed with respect to the initial measure. Therefore, the arguments of this section and those of Refs. [48, 49] are compatible, since (60) would not hold if the cases of Refs. [40, 48, 49] were realized. The scenarios under which the approximation (60) does not hold, appear however quite peculiar, and this section explains why quite commonly the steady state  $\Omega$ -FR does hold: time reversibility and the decay of the  $\Omega$ -autocorrelation characterize the most common deterministic models of nonequilibrium statistical mechanics. Through the connection with the Maxwell time  $t_M$ , this approach also justifies why the convergence times of the steady state  $\Omega$ -FR do not grow with decreasing dissipation: the fact is that the  $\Omega$ -FR depends only on material properties of the system, and these do not change much around equilibrium. This is quite different from the case of the  $\Lambda$ -FR, whose convergence times increase as the driving is decreased, and typically diverge in the equilibrium limit. As stated above, the calculations of this section hold for any regular density  $f$ , i.e. for all possible dissipation functions  $\Omega$ , but the  $\Omega$ -FR turns into the  $\Sigma$ -FR only when the initial distribution can be generated by single field free trajectories, and  $AIF$  holds.

## 5 Conclusions

In this paper we have investigated the mechanisms underlying the validity of the transient and steady state FR's, following the prescriptions given by Evans and Searles over the last decade. As already evidenced in the literature, the transient relations hold under very general conditions; it suffices that the dynamics of the particle systems at hand be deterministic, autonomous, and time reversal invariant. If the initial state is properly selected, the corresponding  $\Omega$  equals the energy dissipation rate  $\Sigma$ . Different initial densities  $f$  are allowed, for any given dynamics  $S^t$ ; they are not completely equivalent, as in general they generate different dissipation functions  $\Omega$ , but to some extent they are interchangeable (cf. (52)). The transient relations describe the properties of evolving ensembles of systems, even in their asymptotic form.

The steady state fluctuation relations need further assumptions. Firstly and quite obviously the system must relax to a nonequilibrium steady state. This requirement is nontrivial. The necessary and sufficient conditions for this to occur are unknown, but necessarily involve some loss of memory of the initial state, and the corresponding decay of correlations. Certainly there are examples of dissipative particle systems in contact with a thermal heat bath that do not relax to a steady state. Also, there may be distinct basins of attraction in  $\mathcal{M}$  with distinct steady state measures. For example if  $\mathcal{M}$  is taken to be all of phase space with an energy less than some bound and the initial ensemble is canonical, but the dynamics is isoenergetic, then obviously there will be no mixing between trajectories that have different initial energies, and the subensemble members will generate distinctly different energy dependent steady states. Different steady states may coexist also when dissipation is too high [55]. Transitivity (i.e. that a typical trajectory explores the whole phase space relevant to the given dynamics) avoids these difficulties.

As a matter of fact, most of the molecular dynamics studies performed so far indicate that convergence to a steady state independent of the initial phase is a very common situation. This convergence is consistent with laboratory experiments where for systems with low Reynolds/Raleigh numbers the nonequilibrium steady state (if such a state exists) is quite independent of the initial starting state or the preparative history of the system. The low shear rate viscosity of fluids is empirically observed to be a function of the temperature and density (or pressure) of the fluid. The independence of the initial phase or preparative history is the nonequilibrium analogue of the situation where at equilibrium thermodynamic free energies are commonly observed to be thermodynamic state functions. Even at equilibrium the conditions required for this to be the case are not known (they are violated in glassy systems for example).

If a single steady state is approached, its statistics may be recovered from an individual trajectory, and this may have no knowledge of the initial ensemble  $f$ . This loss of knowledge of the initial ensemble amounts to a decay of correlations, which may be further weakened. It must be stressed that the decay of correlations required by the validity of the  $\Omega$ -FR reduces to the  $\Omega$ -autocorrelation decay, with respect to the initial ensemble  $f$ . As is well known, this is required, close to equilibrium, for the existence of the transport coefficients.

In the case that transitivity and boundedness of  $\Omega$  suffice to obtain the steady state  $\Omega$ -FR (something that needs further study), we observe that these are much weaker conditions than the Anosov assumption, which also implies the boundedness of the phase space contraction rate.<sup>12</sup> This is interesting, because one can name a number of cases which do

<sup>12</sup>Recall that the Anosov assumption was considered appropriate for the particle system of [6], in which  $-\Lambda = \Omega$ , although  $\Omega$  was not bounded.

have bounded  $\Omega$ , like the isokinetic (or isoenergetic) hard particle systems; the isokinetic or Nosé–Hoover thermostatted systems for heat current and shear flow systems with bounded interaction potentials; the color conductivity isokinetic systems with unbounded potentials etc. Transitivity, in turn, is almost always verified.

We note that in order to obtain a steady state  $\Sigma$ -FR, the initial distribution,  $f_0$ , must be obtainable from a single the field free evolution, and  $AIG$  must hold. These conditions are easily obtainable for systems of physical interest.

The  $\Sigma$ -FR's directly refer to the dissipative fluxes or to quantities simply related to them, rather than to the phase space contraction rate. This avoids the difficulty of justifying how a  $\Lambda$ -FR can speak of the fluctuations of the entropy production rate. As discussed in [5], these difficulties are more evident close to equilibrium, where one observes that the convergence times of the  $\Lambda$ -FR typically diverge as the dissipative applied field decreases and the system approaches equilibrium. This is due to the fact that  $\Lambda$  contains undesirable terms, which have to average to zero. As one approaches equilibrium the dissipation decreases as the square of the dissipative field while the undesirable terms remain constant (to leading order in the field). Therefore, longer and longer averaging times are required before the  $\Lambda$ -FR converges. It has been argued [33, 34] that the  $\Sigma$ -FR can be obtained from the  $\Lambda$ -FR because the time averages of  $\Sigma$  and  $-\Lambda$  become equivalent in the long time limit. However, this argument can not explain the satisfaction of the  $\Sigma$ -FR at physically accessible times in non-isoenergetic systems at low fields. Moreover, if the Van Zon–Cohen picture [25, 26] applies to deterministic particle systems, as argued in Refs. [5, 50, 56], the asymptotic equivalence between  $\Sigma$ -FR and  $\Lambda$ -FR is restricted to a limited fraction of the possible fluctuation values. As a matter of fact,  $\Sigma$  equals  $-\Lambda$  only if  $f$  is uniform in  $\mathcal{M}$ , which requires  $\mathcal{M}$  to be compact. When this is not the case,  $\Sigma$  and  $\Lambda$  are substantially different quantities, as even Ref. [50] shows. The difficulties with the weak field convergence of  $\Lambda$ -FR make the derivation of Green–Kubo relations from  $\Lambda$ -FR, highly problematic [5]. Differently, if a relation for the dissipation is obtained independently, this difficulty is not present. In particular, the convergence times are understood in terms of material properties related to known, physical properties of the system (e.g. the Maxwell time is related to the infinite frequency shear modulus and the zero shear rate shear viscosity), and the derivation of the Green–Kubo relations from the steady state  $\Sigma$ -FRs is reasonably straightforward [5].

It is also important to consider the consistency of the results obtained in this manuscript with those obtained for the axiom C systems postulated in Refs. [48, 49]. As discussed in Appendix 2, these systems may enjoy the decay of correlations of all observables, with respect to an invariant measure supported on a surface  $\mathcal{M}_r$ , of dimension lower than that of the phase space  $\mathcal{M}$ . Then, assuming as in Ref. [49], that  $\Lambda$  is proportional to the contraction rate of volumes in  $\mathcal{M}_r$ ,  $\Lambda_r$  say,  $\Lambda$  should obey a modified FR. So far there is no direct evidence, in particle systems, of the validity of that assumption (supported by numerical studies of certain hydrodynamic models [27–29]). In fact, in the case considered in Ref. [16], the dissipation is sufficiently high that the dynamics is not transitive yet the  $\Sigma$ -FR holds without the modifications conjectured in [49]. In other words, our work and Refs. [48, 49] usually consider different observables, and when they consider the same observable, it is not necessary that they lead to the same conclusions of Refs. [48, 49]. If, on the other hand, the treatment of Refs. [48, 49] applies, the decay of correlations of Sect. 4 is compatible with it, as explained in Appendix 2, because it is referred to the initial measure, and not the invariant

one.<sup>13</sup> The energy dissipation  $\Sigma$  remains related to  $\Lambda$ , not to  $\Lambda_r$ , even if the steady state is confined inside a lower dimensional manifold  $\mathcal{M}_r$ .

As pointed out by one of the anonymous referees, the existence of a unique stationary measure may be considered a strong assumption, and the necessary and sufficient condition required for it to hold are not known. Therefore, the study of Anosov systems constitutes a natural starting point for the purpose of classifying the dynamical systems that verify some kind of FR. The purpose of the present paper is different: it amounts to analyzing systems which cannot be chosen at will, but are selected by physical conditions, in order to understand further their properties as well as the physical mechanism responsible for FR's to hold in natural systems. We have found that time reversibility and ergodic consistency imply, without further assumptions, a number of transient and asymptotic FRs (cf. Sects. 2 and 3), which are amenable to experimental verification. If the system of interest does not reach a unique steady state, assuming that it is Anosov does not help of course, but those relations remain valid. When present, convergence to a unique steady state is the manifestation of a certain decay of correlations of the observables, which typically are required to relax to a steady state. The approach proposed in this paper thus makes it clear that the details of the microscopic dynamics (like the very strong hypothesis of uniform hyperbolicity) do not play any role in the validity of the FR's, as appropriate for thermodynamic relations. The class of dynamical systems which obey (60) cannot be determined, however, by this approach.

These facts may help in developing further the theory of nonequilibrium phenomena, as the derivation of the novel relations given above indicates.

**Acknowledgements** We wish to thank the Australian Research Council, and the Italian Embassy in Canberra for support. We also thank the Erwin Schrödinger Institute for support for the “Workshop on Stochastic and Deterministic Dynamics in Equilibrium and Nonequilibrium Systems” where this work was discussed. We thank Barbara Johnston for her input to this work, and Stephen Williams, Emil Mittag, Angelo Vulpiani, E.G.D. Cohen, Federico Bonetto and Carlos Mejia Monasterio for their comments. Thanks are in order also to the anonymous referees, for useful remarks on the purpose of our work.

### Appendix 1

Using the general definition of the dissipation function (13), it can be shown that the instantaneous value of the *dissipation function* is directly related to the instantaneous *dissipative flux* for a wide class of systems. Here we consider three cases: isoenergetic, isokinetic and Nosé–Hoover thermostatted dynamics.

A general form of the equations of motion for a field-driven  $N$ -particle, thermostatted nonequilibrium system is

$$\begin{aligned} \dot{\mathbf{q}}_i &= \mathbf{p}_i/m + \mathbf{C}_i(\Gamma) \cdot \mathbf{F}_e, \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{D}_i(\Gamma) \cdot \mathbf{F}_e - \alpha \mathbf{p}_i \end{aligned} \tag{65}$$

where  $\mathbf{F}_e$  is the applied field that is coupled to the system via the phase functions  $\mathbf{C}_i(\Gamma)$  and  $\mathbf{D}_i(\Gamma)$ . The term  $-\alpha \mathbf{p}_i$  is a deterministic, time reversible term used to add or remove heat from the system [41]. The adiabatic equations of motion

$$\begin{aligned} \dot{\mathbf{q}}_i &= \mathbf{p}_i/m + \mathbf{C}_i(\Gamma) \cdot \mathbf{F}_e, \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{D}_i(\Gamma) \cdot \mathbf{F}_e \end{aligned} \tag{66}$$

<sup>13</sup>One should also observe that (40) is exact, hence cannot lead to any contradiction about reversible dynamical systems.

simply lack the thermostating term. If (66) are Hamiltonian, the phase space expansion rate  $\Lambda$  obviously vanishes, a condition referred to as *adiabatic incompressibility of phase space*,  $AI\Gamma$  [41], but  $AI\Gamma$  may hold even if the general adiabatic equations are not Hamiltonian. Here, for simplicity, we concentrate on systems which satisfy  $AI\Gamma$ , because their dissipative flux and dissipation function are simply related. In the other cases, the relationship between dissipative flux and dissipation function is only more complicated. We also assume that the initial distribution of phases  $f_0$  may in principle be generated by a single field-free ( $F_e = 0$ ) thermostatted dynamics. By this we mean that  $f_0$  must not only be preserved by the equilibrium dynamics, but also that a single field free trajectory may explore all of its support. In other words, that this support is not the union of disjoint invariant sets, like the constant energy surfaces corresponding to different initial energies of Hamiltonian systems. This is consistent with nonlinear response theory that shows that if one is interested in the statistics of the physically relevant steady state observables,  $f_0$  should be that which corresponds to the field free dynamics [41, 57]. The possible difficulties that might be encountered if  $f_0$  is not selected to be the distribution function generated by the field free dynamics can easily be seen by considering the distribution of  $\phi_{0,\tau}$  along an equilibrium trajectory. In this case the correct physical picture requires the right hand side of (39) to be 1 for all  $A, t_0, \delta$  and  $\tau$ . This can only be the case if  $\Omega(\Gamma) = 0$  for all  $\Gamma$ . Using (34), this implies that  $f_0(S_0^t\Gamma) = f_t(S_0^t\Gamma)$  under the field free dynamics  $S_0^t$ , and that  $f_0 = f_t$  must be preserved by the field free dynamics. Note that a single trajectory can't possibly generate the ensemble in physically relevant times, in any system made of more than a few particles, but it may do it, in principle, allowing it to evolve for exceedingly long. This is needed for the possible initial equilibrium state to be uniquely determined.

In general, the dissipative flux,  $\mathbf{J}$ , is obtained from the adiabatic time-derivative of the internal energy,  $H_0 = \sum_{i=1}^N \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m} + \Phi(\mathbf{q})$ , i.e. from the derivative of the internal energy under the dynamics of (66):

$$\dot{H}_0^{ad} = - \sum_{i=1}^N \left( \frac{\mathbf{p}_i}{m} \cdot \mathbf{D}_i - \mathbf{F}_i \cdot \mathbf{C}_i \right) \cdot \mathbf{F}_e := -\mathbf{J}V \cdot \mathbf{F}_e, \tag{67}$$

while the full dynamics of (65) yields

$$\dot{H}_0 = -\mathbf{J}V \cdot \mathbf{F}_e - \alpha \sum_{i=1}^N \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{m} = -\mathbf{J}V \cdot \mathbf{F}_e - 2K\alpha \tag{68}$$

where  $K$  is the kinetic energy of the system. Consider now three widely used deterministic, reversible thermostating methods.

- Gaussian isoenergetic (ergostatted) system. In this case  $H_0$  is fixed, i.e.  $\dot{H}_0 = 0$ , and (68) gives

$$\alpha(\Gamma) = - \frac{\mathbf{J}(\Gamma)V \cdot \mathbf{F}_e}{2K(\Gamma)}. \tag{69}$$

For the ergostatted system described by equations (65), the phase space expansion rate is,

$$\begin{aligned} \Lambda(\Gamma) &= \nabla \cdot \dot{\Gamma} = \sum_{i=1}^N \left( \frac{\partial \mathbf{C}_i(\Gamma)}{\partial \mathbf{q}_i} + \frac{\partial \mathbf{D}_i(\Gamma)}{\partial \mathbf{p}_i} \right) \cdot \mathbf{F}_e - dN\alpha(\Gamma) + O_N(1) \\ &\stackrel{AI\Gamma}{=} -dN\alpha(\Gamma) + O_N(1) \end{aligned} \tag{70}$$

where  $AI\Gamma$  is used to obtain the final equality, and  $O_N(1)$  is a correction which is order 1 in  $N$ . The  $O_N(1)$  term can be explicitly determined from the partial derivative of the Gaussian ergostating term with respect to  $\mathbf{p}$ . The equilibrium phase space distribution function is microcanonical on a surface  $H_0 = \text{const}$ , hence  $f(iS^\tau \Gamma) = f(S^\tau \Gamma) = f(\Gamma)$ , and (13) yields:

$$\bar{\Omega}_{0,\tau}^{isoE}(\Gamma) = -\bar{\Lambda}_{0,\tau}(\Gamma) \tag{71}$$

for all  $\tau$ , so

$$\Omega(\Gamma) = -\Lambda(\Gamma) = dN\alpha(\Gamma) + O_N(1) = -\mathbf{J}(\Gamma) \cdot \mathbf{F}_e\beta(\Gamma)V + O_N(1) \tag{72}$$

where  $\beta(\Gamma) = (dN - d - 1)/(2K(\Gamma))$ .

- Gaussian isokinetic system. In this case the kinetic energy is fixed  $K(\Gamma) = K_0 = (dN - d - 1)k_B T = (dN - d - 1)/\beta$ . Equation (68) can therefore be used to show,

$$\alpha(\Gamma) = -\frac{(\dot{H}_0(\Gamma) + \mathbf{J}(\Gamma)V \cdot \mathbf{F}_e)\beta}{dN - d - 1} \tag{73}$$

while the phase space expansion rate  $\Lambda$  is still expressed by (70) to order  $N$ , with a different  $O_N(1)$  term. The equilibrium phase space distribution function takes the form  $f(\Gamma) \sim e^{-\beta H_0} \delta(K(\Gamma) - K_0)$ , on a surface  $K(\Gamma) = \text{const}$ , hence  $f(\Gamma)/f(iS^\tau \Gamma) = e^{\beta \int_0^\tau \dot{H}_0(\Gamma(s))ds}$ , and (13) leads to:

$$\bar{\Omega}_{0,\tau}(\Gamma) = \beta \int_0^\tau \dot{H}_0(\Gamma(s))ds - \bar{\Lambda}_{0,\tau}(\Gamma) = -(\mathbf{J} \cdot \bar{\mathbf{F}}_e)_{0,\tau} V\beta + O_N(1) \tag{74}$$

for all  $\tau$ . Combining this with (73) and (70), one obtains

$$\Sigma(\Gamma) = -\mathbf{J}(\Gamma) \cdot \mathbf{F}_e V\beta + O_N(1). \tag{75}$$

- Nosé–Hoover thermostat. In this case, (65) are supplemented with the equation,

$$\dot{\alpha} = \frac{1}{Q}(2K(\Gamma) - dNk_B T) \tag{76}$$

where  $Q$  is a constant related to the relaxation time of the thermostat, and  $T$  is the imposed average temperature. In this case,

$$\alpha(\Gamma) = -(\dot{H}_0(\Gamma) + \mathbf{J}(\Gamma)V \cdot \mathbf{F}_e)/(2K(\Gamma)) \tag{77}$$

and

$$\Lambda(\Gamma) = \nabla \cdot \dot{\Gamma} + \frac{\partial \dot{\alpha}}{\partial \alpha} AI\Gamma \stackrel{\text{def}}{=} -dN\alpha. \tag{78}$$

The equilibrium phase space distribution function is the Nosé–Hoover extended canonical distribution  $f(\Gamma) \sim e^{-\beta(H_0 + \frac{1}{2}Q\alpha^2)}$ , hence  $f(\Gamma)/f(iS^\tau \Gamma) = e^{\beta \int_0^\tau (\dot{H}_0\Gamma(s) + Q\alpha\dot{\alpha})ds}$ . So (13) can be used to show that:

$$\bar{\Sigma}_{0,\tau}^{NH}(\Gamma) = \beta \int_0^\tau (\dot{H}_0(s) + Q\alpha\dot{\alpha})ds - \bar{\Lambda}_{0,\tau}(\Gamma) = -(\mathbf{J} \cdot \bar{\mathbf{F}}_e)_{0,\tau} V\beta \tag{79}$$

for all  $\tau$ , and then

$$\Sigma(\Gamma) = -\mathbf{J}(\Gamma) \cdot \mathbf{F}_e V\beta. \tag{80}$$

Therefore, for field driven nonequilibrium systems whose equations of motion satisfy  $A\Gamma$ , the dissipation function and the dissipative flux are simply related. In fact, for thermostatted systems with a constant field,  $\Sigma(\Gamma) = kJ(\Gamma)$  where  $k$  is a constant.

It is straightforward to show that the same expressions (72), (75), (80) can be obtained even if only a fraction of the system's particles is subjected to a thermostat (e.g. the wall particles), see [13, 14] for example. Furthermore, if the walls are large, the results are not sensitive to the details of the thermostating mechanism [15].

## Appendix 2

The derivation of the steady state relations from transient ones needs a careful analysis of the conditional average

$$\left\langle e^{-\Omega_{0,t_0}} \cdot e^{-\Omega_{t_0+\tau,2t_0+\tau}} \right\rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \quad (81)$$

which appears in (27), (40), and of its limit

$$M(A, \delta, \tau) = \lim_{t_0 \rightarrow \infty} \left\langle e^{-\Omega_{0,t_0}} \cdot e^{-\Omega_{t_0+\tau,2t_0+\tau}} \right\rangle_{\bar{\Omega}_{t_0,t_0+\tau} \in A_\delta^+} \quad (82)$$

which is needed for the transient states  $\mu_{t_0}$  to eventually reach the steady state  $\mu_\infty$ . In particular,  $M(A, \delta, \tau)$  must exist and be positive for  $\tau$  larger than a certain  $\tau_{A,\delta}$ , for the pair  $(-A, A)$  to be  $\delta$ -possible, i.e. for fluctuations of size close to  $A$  to occur, and for  $A$  to be in the domain of the fluctuation relation with some tolerance  $\gamma > \delta$ .

As observed in the text, (27) and (40) are exact, hence cannot contradict any correct result on deterministic, reversible dynamical systems. The question is whether one can extract sufficient information from them, that they can be of practical use for the description of the statistics of steady state trajectories. This requires that either the dynamics of interest be explicitly given, or that the possible situations be explored. In Sect. 4, we have considered the possibility that the  $\Omega$ -autocorrelation decays in time, as the most common for the deterministic reversible particle models of nonequilibrium fluids. In fact, less is needed for the steady state  $\Omega$ -FR to hold for  $A$ ; it suffices that  $M(A, \delta, \tau)$  does not grow too fast with  $\tau$ : subexponential growths, and exponential growths with rate not larger than  $\delta$  (as for  $A = 0$ ) are all acceptable.

In the linear regime, the decay of correlations in the steady state can be well approximated by their decay at equilibrium, and for this reason, in the linear regime, the transport coefficients are given by the integral of the equilibrium time correlation function, multiplied by the field [41]. Lack of decay of correlations of  $\Omega$  corresponds to the non-existence of the transport coefficient associated with  $\Omega$ . Further away from equilibrium, nonlinear response theory applies. Now the time correlation function is still determined with respect to the initial measure  $\mu$ , but the dynamics used to compute the nonlinear transport coefficients is the nonequilibrium one (cf. the Transient Time Correlation Function [41]). This is analogous to what is done in calculations of  $M(A, \delta, \tau)$ , and if the nonlinear transport coefficients exist in the nonlinear regime, then there should be a decay of the  $\Omega$ -autocorrelation, with respect to the equilibrium measure. Note also that the growth of  $\tau$  increases the separation of the end of the integral in the first exponential of (81) from the beginning of the integral in the second exponential, hence that the growth of  $\tau$  should contribute to the decorrelation.

Therefore, it looks quite plausible that, in most cases, one of the above considerations on the decay of correlations be satisfied. There is, however, one competing effect, which may become important in particular circumstances: the set of trajectories over which the



conditional average (81) is computed changes with  $\tau$ , and this may balance the decorrelation phenomena in some cases.

To illustrate these facts, consider a simple model that has been discussed in connection to the fluctuation theorems in the past [40], and apply the procedure described in this manuscript to that example. The model consists of a particle moving in empty space, under the action of a constant external force  $\mathbf{F}_e$  and a Gaussian thermostat:

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{F}_e - \frac{\mathbf{F}_e \cdot \mathbf{p}}{\mathbf{p} \cdot \mathbf{p}} \mathbf{p} \end{aligned}$$

where the kinetic energy  $K = \mathbf{p} \cdot \mathbf{p}/2$  is a constant of motion. The dynamics of this particle is rather simple: for initial conditions with  $\mathbf{p}$  pointing in the direction opposite to  $\mathbf{F}_e$ ,  $\mathbf{p}$  is a constant of motion; for all other initial conditions the direction of  $\mathbf{p}$  tends to the direction of  $\mathbf{F}_e$ , while the magnitude of  $\mathbf{p}$ ,  $\sqrt{2K}$ , is constant. The treatment of Sects. 2 and 3 applies to this system, since it is reversible and an initial distribution that is ergodically consistent with the final distribution can be selected. However there are various reasons why the treatment in Sect. 4 cannot be applied to produce a steady state  $\Sigma$ -FR in this case. For instance, the distribution generated by the field free equations of motion is not ergodically consistent with the steady state dynamics (see Appendix 1). In addition, the attractor is a fixed point in  $\mathbf{p}$ -space, and the time-dependence of  $\mathbf{q}$  is irrelevant; so, for any  $\Omega$  that is a function of  $\mathbf{p}$  only, there are no steady state fluctuations - that is for small  $\delta$ , no pair  $(-A, A)$  is  $\delta$ -possible, hence that the domain of the steady state  $\Omega$ -FR is empty. The theory outlined in this paper also anticipates these facts, since correlations for any phase function do not decay, and therefore the term  $M$  diverges in the  $t_0 \rightarrow \infty$  limit. In particular, selecting the dissipation function that would correspond to a uniform initial distribution, we obtain  $\Omega = -\Lambda = (d - 1)\mathbf{F}_e \cdot \mathbf{p}/2K$  where  $d$  is the number of dimensions in configuration space. The range of values of this  $\Omega$  is  $[-(d - 1)\sqrt{2K}|\mathbf{F}_e|, (d - 1)\sqrt{2K}|\mathbf{F}_e|]$  and any smooth probability distribution in it (e.g. the uniform one) evolves in time in such a way that

$$\mu_{t_0}(\Omega_{0,\tau} | A_\delta^+) \xrightarrow{t_0 \rightarrow \infty} \begin{cases} 0 & \text{if } (d - 1)\sqrt{2K}|\mathbf{F}_e| \notin (A - \delta, A + \delta), \\ 1 & \text{if } (d - 1)\sqrt{2K}|\mathbf{F}_e| \in (A - \delta, A + \delta). \end{cases} \tag{83}$$

As all trajectories, except those with initial condition  $\mathbf{p}(0) = -(d - 1)\sqrt{2K}\mathbf{F}_e/|\mathbf{F}_e|$ , tend to have  $\Omega = (d - 1)\sqrt{2K}\mathbf{F}_e/|\mathbf{F}_e|$ , this means that the integrals  $\Omega_{0,t_0}$  and  $\Omega_{t_0+\tau,2t_0+\tau}$  diverge linearly with  $t_0$ .

Therefore, (60) does not apply in this case. This will be the case for any steady state that has no fluctuations. Nevertheless, the transient and asymptotic relations remain valid for the evolving ensembles. The transient  $\Omega$ -FR expresses the rate at which the initial probability of observing  $-A$  vanishes, compared to the initial probability of  $A$ , and the  $\Omega_\infty$ -FR expresses the corresponding asymptotic rate; but the domain of the steady state  $\Omega$ -FR is empty. A similar situation, with no fluctuations of opposite sign in the steady state, may be produced in more realistic models by applying extremely high fields, making the motion ordered. In these cases, the domain of the steady state  $\Omega$ -FR would also be empty. However, intermediate situations are possible, like those of Ref.[16], in which the attractors have small dimension, but the initial correlations decay, and the steady state  $\Omega$ -FR holds.

References [48, 49] consider systems with steady states confined within manifolds  $\mathcal{M}_r$  of dimension smaller than  $\mathcal{M}$ , whose volumes fluctuate in time, so that the corresponding phase space contraction rate,  $\Lambda_r$ , obeys the standard fluctuation relation. In these systems,

correlations of all observables decay with respect to the invariant measure  $\mu_\infty$ , supported on  $\mathcal{M}_r$ . Furthermore, if the phase space contraction rate  $\Lambda$ , in the space of the field-free dynamics  $\mathcal{M}$ , is proportional to  $\Lambda_r$ , as assumed in [49], a modified FR holds for it, which can be expressed as

$$cA - \gamma \leq \frac{1}{\tau} \ln \frac{\mu_\infty(\bar{\Lambda}_{0,\tau}|_{A_\delta^+})}{\mu_\infty(\bar{\Lambda}_{0,\tau}|_{A_\delta^-})} \leq cA + \gamma \quad (84)$$

with  $c < 1$ . The relevance of such a possibility for particle systems is not obvious, but no difficulty emerges with the theory developed here, since the decay of correlations with respect to  $\mu_\infty$  does not imply the decay of correlations with respect to a non-singular measure supported on  $\mathcal{M}$ . Many different scenarios are possible, and should be analyzed case by case, but this goes beyond the purpose of the present paper. Here it suffices to observe that no scenario that can be realized can contradict the framework of the present paper; to the contrary, it could be explained in it, through the analysis of the conditional average (81).

## References

1. Evans, D.J., Searles, D.J.: The fluctuation theorem. *Adv. Phys.* **52**, 1529–1585 (2002)
2. Searles, D.J., Evans, D.J.: Fluctuation relations for nonequilibrium systems. *Aust. J. Chem.* **57**, 1119–1123 (2004)
3. Gallavotti, G.: *Statistical Mechanics, a Short Treatise*. Springer, Berlin (2000)
4. Bustamante, C., Liphardt, J., Ritort, F.: The nonequilibrium thermodynamics of small systems. *Phys. Today* **58**, 43–48 (2005)
5. Evans, D.J., Searles, D.J., Rondoni, L.: Application of the Gallavotti–Cohen fluctuation relation to thermostatted steady states near equilibrium. *Phys. Rev. E* **71**, 056120 (2005)
6. Evans, D.J., Cohen, E.G.D., Morriss, G.P.: Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **71**, 2401–2404 (1993)
7. Parry, W.: Synchronisation of canonical measures for hyperbolic attractors. *Commun. Math. Phys.* **106**, 267–275 (1986)
8. Vance, W.N.: Unstable periodic orbits and transport properties of nonequilibrium steady states. *Phys. Rev. Lett.* **69**, 1356–1359 (1992)
9. Cvitanović, P., Artuso, R., Mainieri, R., Tanner, G., Vattay, G.: *Chaos: Classical and Quantum*, Chaos-Book.org. Niels Bohr Institute, Copenhagen (2005)
10. Evans, D.J., Searles, D.J.: Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E* **50**, 1645–1648 (1994)
11. Evans, D.J., Searles, D.J.: Steady states, invariant measures and response theory. *Phys. Rev. E* **52**, 5839–5848 (1995)
12. Searles, D.J., Ayton, G., Evans, D.J.: Generalised fluctuation formula. *AIP Conf. Ser.* **519**, 271–280 (2000)
13. Searles, D.J., Evans, D.J.: Ensemble dependence of the transient fluctuation theorem. *J. Chem. Phys.* **113**, 3503–3509 (2000)
14. Searles, D.J., Evans, D.J.: Fluctuation theorem for heat flow. *Int. J. Thermophys.* **22**, 123–124 (2001)
15. Williams, S.R., Searles, D.J., Evans, D.J.: Independence of the transient fluctuation theorem to thermostating. *Phys. Rev. E* **70**, 066113 (2004)
16. Williams, S.R., Searles, D.J., Evans, D.J.: Numerical study of the steady state fluctuation relations far from equilibrium. *J. Chem. Phys.* **124**, 194102/1-9 (2006)
17. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in stationary states. *J. Stat. Phys.* **80**, 931–970 (1995)
18. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.* **74**, 2694–2697 (1995)
19. Gallavotti, G.: Reversible Anosov diffeomorphisms and large deviations. *Math. Phys. Electron. J.* **1**, 1–12 (1995)
20. Gallavotti, G.: Fluctuation theorem revisited. [arXiv:cond-mat/0404699](https://arxiv.org/abs/cond-mat/0404699) (2004).
21. Searles, D.J., Evans, D.J.: The fluctuation theorem for stochastic systems. *Phys. Rev. E* **60**, 159–164 (1999)

22. Kurchan, J.: Fluctuation theorem for stochastic dynamics. *J. Phys. A* **31**, 3719–3729 (1998)
23. Lebowitz, J.L., Spohn, H.: A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **95**, 333–365 (1999)
24. Maes, C.: The fluctuation theorem as a Gibbs property. *J. Stat. Phys.* **95**, 367–392 (1999)
25. van Zon, R., Cohen, E.G.D.: Extension of the fluctuation theorem. *Phys. Rev. Lett.* **91**, 110601 (2003)
26. van Zon, R., Cohen, E.G.D.: Extended heat-fluctuation theorems for a system with deterministic and stochastic forces. *Phys. Rev. E* **69**, 056121 (2004)
27. Gallavotti, G.: Dynamical ensembles equivalence in fluid mechanics. *Physica D* **105**, 163–184 (1997)
28. Rondoni, L., Segre, E.: Fluctuations in two-dimensional reversibly damped turbulence. *Nonlinearity* **12**, 1471–1487 (1999)
29. Gallavotti, G., Rondoni, L., Segre, E.: Lyapunov spectra and nonequilibrium ensembles equivalence in 2D fluid mechanics. *Physica D* **187**, 338–357 (2004)
30. Gallavotti, G.: A local fluctuation theorem. *Physica A* **263**, 39–50 (1999)
31. Rondoni, L., Tèl, T., Vollmer, J.: Fluctuation theorems for entropy production in open systems. *Phys. Rev. E* **61**, R4679–R4682 (2000)
32. Ayton, G., Evans, D.J., Searles, D.J.: A local fluctuation theorem. *J. Chem. Phys.* **115**, 2033–2037 (2001)
33. Zamponi, F., Ruocco, G., Angelani, L.: Fluctuations of entropy production in the isokinetic ensemble. *J. Stat. Phys.* **115**, 1655–1668 (2004)
34. Giuliani, A., Zamponi, F., Gallavotti, G.: Fluctuation relation beyond linear response theory. *J. Stat. Phys.* **119**, 909–944 (2005). Note that the fluctuation relation that is validated in these papers is the  $\Omega$ -FR
35. Dolowschiak, M., Kovacs, Z.: Fluctuation formula in the Nosé–Hoover thermostated Lorentz gas. *Phys. Rev. E* **71**, 025202 (2005)
36. Wang, G.M., Sevick, E.M., Mittag, E., Searles, D.J., Evans, D.J.: Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.* **89**, 050601 (2002)
37. Garnier, N., Ciliberto, S.: Nonequilibrium fluctuations in a resistor. *Phys. Rev. E* **71**, 060101 (2005)
38. Wang, G.M., Reid, J.C., Carberry, D.M., Williams, D.R.M., Sevick, E.M., Evans, D.J.: Experimental study of the fluctuation theorem in a nonequilibrium steady state. *Phys. Rev. E* **71**, 046142 (2005)
39. Puglisi, A., Rondoni, L., Vulpiani, A.: Relevance of initial and final conditions for the fluctuation relation in Markov processes. *J. Stat. Mech.* **8**, P08010 (2006)
40. Cohen, E.G.D., Gallavotti, G.: Note on two theorems in nonequilibrium statistical mechanics. *J. Stat. Phys.* **96**, 1343–1349 (1999)
41. Evans, D.J., Morriss, G.P.: *Statistical Mechanics of Nonequilibrium Liquids*. Academic, London (1990)
42. Liem, S.Y., Brown, D., Clarke, J.H.R.: Investigation of the homogeneous-shear nonequilibrium-molecular-dynamics method. *Phys. Rev. A* **45**, 3706–3713 (1992)
43. de Groot, S.R., Mazur, P.: *Non-Equilibrium Thermodynamics*. Dover, New York (1984)
44. Ruelle, D.: Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *J. Stat. Phys.* **95**, 393–468 (1999)
45. Khinchin, A.I.: *Mathematical Foundations of Statistical Mechanics*. Dover, New York (1949)
46. Jarzynski, C.: Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690–2693 (1997)
47. Crooks, G.E.: Path ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **61**, 2361–2366 (2000)
48. Bonetto, F., Gallavotti, G.: Reversibility, coarse graining and the chaoticity principle. *Commun. Math. Phys.* **189**, 263 (1997)
49. Bonetto, F., Gallavotti, G., Garrido, P.L.: Chaotic principle: an experimental test. *Physica D* **105**, 226–252 (1997)
50. Bonetto, F., Gallavotti, G., Giuliani, A., Zamponi, F.: Chaotic hypothesis, fluctuation theorem and singularities. *J. Stat. Phys.* **123**, 39–54 (2006)
51. Ruelle, D.: Conversations on nonequilibrium physics with an extraterrestrial. *Phys. Today* **57**(5), 48–53 (2004)
52. Rondoni, L.: Deterministic thermostats and fluctuation relations. In: Garbaczewski, P., Olkiewicz, R. (eds.) *Dynamics of Dissipation*. Springer, Berlin (2002)
53. Evans, D.J., Morriss, G.: Nonlinear response theory for steady planar Couette flow. *Phys. Rev. A* **30**, 1528–1530 (1984)
54. Evans, D.J., Searles, D.J., Rondoni, L.: in preparation
55. Giberti, C., Rondoni, L., Vernia, C.: Coexistence of chaotic and non-chaotic states in the two-dimensional Gauss–Navier–Stokes dynamics. *Physica D* **187**, 358–369 (2004)
56. Gilbert, T.: Fluctuation theorem applied to the Nosé–Hoover thermostatted Lorentz gas. *Phys. Rev. E* **73**, 035102 (2006)
57. Evans, D.J., Searles, D.J.: Nonlinear response theory and the fluctuation theorem. *J. Chem. Phys.* (2007, submitted), <http://arxiv.org/abs/cond-mat/0612105>