

Article

Time Reversibility, Correlation Decay and the Steady State Fluctuation Relation for Dissipation

Debra J. Searles ^{1,2,*}, Barbara M. Johnston ², Denis J. Evans ³ and Lamberto Rondoni ^{4,5}

¹ Australian Institute of Bioengineering and Nanotechnology and School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Qld 4072, Australia

² Queensland Micro-and Nanotechnology Centre and School of Biomolecular and Physical Sciences, Griffith University, Brisbane, Qld 4111, Australia; E-Mail: barbara.johnston@griffith.edu.au

³ Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia; E-Mail: evans@rsc.anu.edu.au

⁴ Dipartimento di Scienze Matematiche, Politecnico di Torino, Torino 10129, Italy; E-Mail: lamberto.rondoni@polito.it

⁵ INFN, Sezione di Torino, Via P. Giuria 1, Torino 10125, Italy

* Author to whom correspondence should be addressed; E-Mail: d.bernhardt@uq.edu.au; Tel.: +61-7-3346-3939; Fax: +61-7-3346-3992.

Received: 4 March 2013; in revised form: 17 April 2013 / Accepted: 19 April 2013 /

Published: 25 April 2013

Abstract: Steady state fluctuation relations for nonequilibrium systems are under intense investigation because of their important practical implications in nanotechnology and biology. However the precise conditions under which they hold need clarification. Using the dissipation function, which is related to the entropy production of linear irreversible thermodynamics, we show time reversibility, ergodic consistency and a recently introduced form of correlation decay, called T-mixing, are sufficient conditions for steady state fluctuation relations to hold. Our results are not restricted to a particular model and show that the steady state fluctuation relation for the dissipation function holds near or far from equilibrium subject to these conditions. The dissipation function thus plays a comparable role in nonequilibrium systems to thermodynamic potentials in equilibrium systems.

Keywords: steady state fluctuations; dissipation; nonequilibrium molecular dynamics simulation; entropy production

Classification: PACS 05.40.-a, 05.70.Ln, 05.20.-y

1. Introduction

Fluctuation relations (FRs) have been studied intensively over the past twenty years. Transient Fluctuation relations are the best understood. They are known to be exact for systems of arbitrary size, arbitrarily near or far from equilibrium. Previously, very few exact results were known in nonequilibrium statistical mechanics. Their application to small systems coincided with an upsurge of interest in nanotechnology and nano-biology and the study of small bio-engines. The applicability of the transient FRs far from equilibrium meant that they were an important touchstone for large deviation theory.

Asymptotic steady state fluctuation relations are far less well understood. Some such relations were known to have convergence problems close to equilibrium. In this paper we prove that for the so-called T-mixing systems, the Evans–Searles steady state fluctuation relation is free of convergence difficulties, *i.e.*, it applies to systems of arbitrary size and applies arbitrarily near, or far from, equilibrium. This paper identifies the conditions required for the steady state FR involving the dissipation function, to be derived from the exact Evans–Searles transient fluctuation theorem, and provides an illustration of the results using numerical simulations.

In 1993 [1], Evans, Cohen and Morriss studied the steady state fluctuations of the entropy production rate in isoenergetic SLLOD (SLLOD is the term used to refer to the standard algorithm used to model homogeneous flow [2]) shearing systems, defined by:

$$\dot{\mathbf{q}}_i = \mathbf{p}_i/m + \mathbf{n}_x \dot{\gamma} y_i, \quad \dot{\mathbf{p}}_i = \mathbf{F}_i - \mathbf{n}_x \dot{\gamma} p_{yi} - \alpha_{IE} \mathbf{p}_i \quad (1)$$

where $i = 1, \dots, N$; N is the number of particles in a periodic system with a unit cell of volume V ; m is the mass of each particle; $\dot{\gamma}$ is the shear rate; $\mathbf{n}_x$ is the unit vector in the x -direction; $\mathbf{F}_i$ is the internal force on particle i ; $\mathbf{q}_i = (x_i, y_i)$ is the position of each particle; $\mathbf{p}_i = (p_{xi}, p_{yi})$ is the momentum of each particle; and α_{IE} is a Lagrange multiplier that fixes the internal energy of the system, $H_0 = \sum_{i=1}^N \mathbf{p}_i \cdot \mathbf{p}_i / (2m) + \Phi(\mathbf{q})$, where $\Phi(\mathbf{q})$ is the interparticle potential. The value of α_{IE} is related to: the rate of energy dissipation under unthermostatted (adiabatic) conditions, $\dot{H}_0^{ad}$; the entropy production rate close to equilibrium, Σ ; and the phase space expansion rate, $\Lambda = \partial/\partial\Gamma \cdot \dot{\Gamma}$ where $\Gamma = (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) = (\mathbf{q}, \mathbf{p})$. This relationship is given by,

$$\alpha_{IE} = \frac{\dot{H}_0^{ad}}{\sum_{i=1}^N p_i^2/m} = \frac{\Sigma}{N_{df}} = -\frac{\Lambda}{N_{df}} = \frac{-\dot{\gamma} P_{xy} V}{\sum_{i=1}^N p_i^2/m} \quad (2)$$

Here P_{xy} is the off-diagonal part of the pressure tensor and $N_{df} = dN - d - 1$ is the number of degrees of freedom. In [1] the spatial dimension was $d = 2$. In this and similar molecular dynamics models, unphysical thermostating terms like α_{IE} may be confined to arbitrarily remote boundaries. Then, the bulk properties of the systems of interest should not be affected by the thermostating details. To set our notation, let $S^s\Gamma$ be a point in phase space evolved via the time evolution operator, S , (given by the equations of motion) for a period s from initial condition Γ . Given an observable $\varphi(\Gamma)$, let us denote its time average and integral, between times t_1 and t_2 , along the phase space trajectory segment with initial condition Γ as

$$\bar{\varphi}_{t_1, t_2}(\Gamma) \equiv \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} \varphi(S^s\Gamma) ds \equiv \frac{\varphi_{t_1, t_2}(\Gamma)}{(t_2 - t_1)} \quad (3)$$

As the time average $\bar{\Lambda}_{0,\tau}$ over the period τ , equals the sum of the Lyapunov exponents $\sum_n \lambda_n$ for the system in the long τ limit, Evans, Cohen and Morriss [1] proposed and tested the asymptotic steady state relation ($\Lambda - FR$)

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \frac{\mu_i}{\mu_{i^*}} &= \frac{\exp \left[-\sum_n^+ \lambda_{i,n} \tau \right]}{\exp \left[-\sum_n^+ \lambda_{i^*,n} \tau \right]} \\ &= \exp \left[-\sum_n \lambda_{i,n} \tau \right] = \exp \left[Nd\bar{\alpha}_{IE;0,\tau;i} \tau \right] \end{aligned} \quad (4)$$

Here i, i^* label trajectory segments generated over the period τ and related by time reversal symmetry so that the averages of Λ over the segments i, i^* obey $\bar{\Lambda}_{0,\tau;i^*} = -\bar{\Lambda}_{0,\tau;i}$. The n -th phase space expansion rate along trajectory segment i is given by $\lambda_{i,n}$, and the sums run over all the positive λ 's. The values of μ_i and μ_{i^*} are the probabilities of observing trajectories initially within a small volume $d\Gamma$ about i or i^* [3]. As in periodic orbit expansions [4,5], Equation (4) can then be predicted to hold within a bounded range $|\bar{\Lambda}_{0,\tau;i}| \equiv |Nd\bar{\alpha}_{IE;0,\tau;i}| < B$, $B > 0$. Later it was discovered that for non-isoenergetic systems, the physically relevant asymptotic limit in (4) is really $\lim(\tau\dot{\gamma}^2 \rightarrow \infty)$ while B was estimated to be $B = O(\dot{\gamma}^2)$ [6]. These two facts can make the observability of the steady state $\Lambda - FR$ difficult close to equilibrium.

Evans *et al.* [1] show that if one considers Equation (4) and sums over all i, i^* for which $\bar{\alpha}_{IE;0,\tau;i}$ (or equivalently the entropy production) takes on a particular value within some tolerance, the fluctuation relation (FR) for the entropy production rate can be obtained. However, the proportionality between Λ , the energy dissipation and the entropy production given in Equation (2) holds only for Gaussian isoenergetic systems. For instance, systems with constant kinetic energy differ from isoenergetic systems, because,

$$\alpha_{IK} = \frac{\sum_{i=1}^N \mathbf{F}_i \cdot \dot{\mathbf{q}}_i - \dot{\gamma} P_{xy} V}{\sum_{i=1}^N \mathbf{p}_i^2 / m} \quad (5)$$

replaces α_{IE} in Equation (1), and Λ becomes the sum of two terms, only one of which is related to energy dissipated by the strain rate:

$$\Lambda = \frac{c_1 \sum_{i=1}^N \mathbf{F}_i \cdot \dot{\mathbf{q}}_i}{\sum_{i=1}^N \mathbf{p}_i^2 / m} - \frac{c_2 \dot{\gamma} P_{xy} V}{\sum_{i=1}^N \mathbf{p}_i^2 / m} \quad (6)$$

Here, the constants c_1, c_2 depend on N and d .

In 1994, Evans and Searles [7,8] developed an alternative approach for deriving fluctuation relations for the energy dissipation that could be generalised to non-isoenergetic systems, through the introduction of the dissipation function Ω [9]. For field-driven systems, Ω equals $-JV F_e / k_B T$ [9], which, close to equilibrium, is the rate of entropy production [2]. Here, J is the dissipative flux due to the force F_e , V is the volume, k_B is Boltzmann's constant and T is the temperature of the equilibrium state that the system would relax to if the driving force (strain rate) was set to zero. The numerical value of this temperature is, for this system, the kinetic temperature [7,8] of the nonequilibrium driven system. The first set of such relations, called transient Ω -FRs, were obtained under virtually no hypothesis, except for time reversibility and ergodic consistency; they concern phase space distributions that are not preserved by the dynamics. No bounds limit the argument values for these transient Ω -FRs. They are valid arbitrarily near or far from equilibrium. In the far from equilibrium case unlike dissipation, the entropy production

cannot be defined. These transient Ω -FRs are (like related identities such as the Jarzynski equality, see e.g., [10]) exact for all times.

Transient Ω -FRs have been verified experimentally [11]. Many tests and laboratory experiments have also validated the extension of the transient Ω -FRs to nonequilibrium steady states [12,13]. This supports the use of thermostatted dynamics in this context. However, the understanding of the physical mechanisms underlying the validity of the steady state FRs is less complete [6,14–16]. The present paper clarifies some of these outstanding issues.

Following the approach developed by Evans and Searles, we demonstrate how time reversibility, ergodic consistency and a form of the decay of correlations, known as T-mixing [14,15,17], lead to the steady state Ω -FR for systems of arbitrary size, near or far from equilibrium.

2. Calculations

For the sake of clarity, and without loss of generality, we focus our discussion and simulations on the colour conductivity model [2] defined by:

$$\dot{\mathbf{q}}_i = \mathbf{p}_i/m, \quad \dot{\mathbf{p}}_i = \mathbf{F}_i + c_i \mathbf{F}_e - \alpha \mathbf{p}_i; \quad i = 1, \dots, N \quad (7)$$

where F_e is the applied colour field that is coupled to the system via the colour charges, $c_i = (-1)^i$, and $\alpha = \sum_{i=1}^N (\mathbf{F}_i \cdot \mathbf{p}_i + c_i \mathbf{F}_e \cdot \mathbf{p}_i) / \sum_{i=1}^N \mathbf{p}_i^2$. This model is one of the first used in nonequilibrium molecular dynamics, for calculation of the self-diffusion coefficient. It is a simplified model of an ionic liquid where the particles do not have Coulomb interactions with each other but experience a force proportional to their colour charge when subject to a field. The term $-\alpha \mathbf{p}_i$ is a deterministic, time reversible term used to add or remove energy—in the form of heat—from the system [2] in order to keep the kinetic energy fixed. If α is set to zero, one obtains the adiabatic equations of motion, which are Hamiltonian and hence preserve phase space volumes, a condition referred to as adiabatic incompressibility of phase space, AIF [2]. The thermostat might only apply to the wall particles, as discussed above. As this makes no difference for the derivation of the FRs, we employ a homogenous thermostat.

In the simulations, we choose a system in two Cartesian dimensions with $F_e = (F_{ex}, 0)$, and all results are given in Lennard-Jones reduced units. The initial distribution is taken to be the equilibrium isokinetic phase space distribution function that takes the form

$$f(\mathbf{\Gamma}; 0) = \frac{\exp[-\beta H_0(\mathbf{\Gamma})] \delta(\mathbf{P}) \delta(K(\mathbf{\Gamma}) - K_0)}{\int \exp[-\beta H_0(\mathbf{\Gamma})] \delta(\mathbf{P}) \delta(K(\mathbf{\Gamma}) - K_0) d\mathbf{\Gamma}} \quad (8)$$

where $\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i$ is the total momentum of the system, K is its kinetic energy, which is fixed at K_0 and $\beta = 1/k_B T$. We consider a two-dimensional Gaussian isokinetic system, with kinetic energy fixed at the equipartition value $K_0 = (2N - 3)\beta^{-1}/2$. Note that the first argument of $f(\mathbf{\Gamma}; t)$ refers to the phase point where the distribution function is evaluated and the second argument refers to the time to which the phase space distribution function has evolved, starting from an initial distribution $f(\mathbf{\Gamma}; 0)$.

The time integral from time 0 to time $2t + \tau$ of the dissipation function, along a trajectory originating at $\mathbf{\Gamma}$, is formally defined as

$$\Omega_{0,2t+\tau}(\mathbf{\Gamma}) \equiv \ln \left[\frac{f(\mathbf{\Gamma}; 0)}{f(S^{2t+\tau}\mathbf{\Gamma}; 0)} \right] - \int_0^{2t+\tau} \Lambda(S^s \mathbf{\Gamma}) ds \quad (9)$$

A key point to the definition of dissipation is that Γ and $M^T S^{2t+\tau} \Gamma$ are the origin phases for a trajectory and its conjugate anti-trajectory respectively where M^T is the time reversal operator. This places constraints on the propagator $S^{2t+\tau}$. We break the interval $(0, 2t + \tau)$ into a set of symmetric subintervals: $(0, t), (t, t + \tau), (t + \tau, 2t + \tau)$ to ensure that $M^T S^{2t+\tau} \Gamma$ is the origin of the conjugate anti-trajectory of a trajectory that starts from Γ at $t = 0$. Later we will let $t \rightarrow \infty$ so that the middle trajectory segment $(t, t + \tau)$ approaches a steady state segment. To ensure that the dissipation function is well-defined over the ostensible phase space domain, we assume the system is ergodically consistent *i.e.*, $f(\Gamma; 0) \neq 0 \Rightarrow f(M^T S^{2t+\tau} \Gamma; 0) \neq 0$ [18].

For our system, the instantaneous dissipation function is [14]

$$\Omega(\Gamma) = -\beta \mathbf{J}(\Gamma) V \cdot \mathbf{F}_e \quad (10)$$

where the dissipative flux, $\mathbf{J}$, is obtained from the adiabatic time-derivative of the internal energy H_0 :

$$\dot{H}_0^{ad}(\Gamma) = \sum_{i=1}^N c_i \frac{\mathbf{p}_i}{m} \cdot \mathbf{F}_e \equiv -\mathbf{J}(\Gamma) V \cdot \mathbf{F}_e \quad (11)$$

while the fully thermostatted dynamics of Equation (7) yields

$$\begin{aligned} \dot{H}_0(\Gamma) &= -\mathbf{J}(\Gamma) V \cdot \mathbf{F}_e - 2K_0 \alpha \\ &= -\mathbf{J}(\Gamma) V \cdot \mathbf{F}_e + \Lambda(\Gamma) k_B T \end{aligned} \quad (12)$$

by equipartition, the phase space expansion rate is $\Lambda = -(2N - 3) \alpha$.

Time reversibility implies that for every initial condition Γ that yields a value $\Omega_{0,2t+\tau}(\Gamma)$, there is an initial condition Γ^* yielding the value $\Omega_{0,2t+\tau}(\Gamma^*) = -\Omega_{0,2t+\tau}(\Gamma)$. To obtain Γ^* the time reversal map, M^T , is applied to the time-evolved phase, $S^{2t+\tau} \Gamma$, *i.e.*, $\Gamma^* = M^T S^{2t+\tau} \Gamma$. We also assume that $f(M^T \Gamma; 0) = f(\Gamma; 0)$, which guarantees that we measure properties in a co-moving frame since the average velocity of particles in the system is zero. From (8):

$$\frac{f(\Gamma; 0)}{f(S^{2t+\tau} \Gamma; 0)} = \exp \left(\beta \int_0^{2t+\tau} \dot{H}_0(S^s \Gamma) ds \right) \quad (13)$$

Substituting (13) into (9) using (11) gives,

$$\begin{aligned} \bar{\Omega}_{0,2t+\tau}(\Gamma) &= \frac{\beta}{(2t + \tau)} \int_0^{2t+\tau} \dot{H}_0(S^s \Gamma) ds - \bar{\Lambda}_{0,2t+\tau}(\Gamma) \\ &= -\bar{\mathbf{J}}_{0,2t+\tau}(\Gamma) \cdot \mathbf{F}_e V \beta \end{aligned} \quad (14)$$

It is very important to note that in Equations (9) and (13), the phase point evolves but we only refer to the initial phase space distribution function. This is the Schrödinger picture.

Consider firstly the transient Ω -FR, which is just a consequence of time reversibility and ergodic consistency. In the case of the transient Ω -FR, we wish to consider evolution of the trajectory from Γ , and therefore t in the equations above is 0. Time reversibility implies that the coordinate transformation $\Gamma^* = M^T S^\tau \Gamma$ has Jacobian,

$$\begin{aligned} \left| \frac{d\Gamma^*}{d\Gamma} \right| &= \exp \left[\int_0^\tau \Lambda(S^s \Gamma) ds \right] \\ &= \exp \left[\beta \int_0^\tau \dot{H}_0(S^s \Gamma) ds + \tau \bar{\mathbf{J}}_{0,\tau}(\Gamma) \cdot \mathbf{F}_e V \beta \right] \end{aligned} \quad (15)$$

i.e., the exponent involves the change in the heat. Let $A_\delta^+ = (A - \delta, A + \delta)$ and $A_\delta^- = (-A - \delta, -A + \delta)$ be small intervals of width 2δ centred in A and $-A$, and denote by $\phi|_E$ the set of Γ such that $\phi(\Gamma) \in E$ where $E = A_\delta^+$ or A_δ^- . The above leads to

$$\begin{aligned} & \frac{\text{Prob}(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V|_{A_\delta^+})}{\text{Prob}(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V|_{A_\delta^-})} = \frac{\int_{\bar{\Omega}_{0,\tau}|_{A_\delta^+}} f(\Gamma; 0) d\Gamma}{\int_{\bar{\Omega}_{0,\tau}|_{A_\delta^-}} f(\Gamma^*; 0) d\Gamma^*} \\ &= \frac{\int_{\bar{\Omega}_{0,\tau}|_{A_\delta^+}} f(\Gamma; 0) d\Gamma}{\int_{\bar{\Omega}_{0,\tau}|_{A_\delta^+}} \exp[-\Omega_{0,\tau}(\Gamma)] f(\Gamma; 0) d\Gamma} \\ &= e^{[A + \varepsilon(\delta, A, \tau)]\tau} \end{aligned} \quad (16)$$

where $|\varepsilon| \leq \delta$. Equation (16) is called transient Ω -FR because the system is not time invariant as it evolves from $s = 0$ to $s = \tau$. Time reversibility is essentially the only requirement for its derivation, which is why the Ω -FR is so general and is exact.

Now, let us summarise the derivation of the steady state FR given in [14], in the framework of our model. First of all, consider the time average of the dissipation function over the period t to $t + \tau$: from Equation (14), $\bar{\Omega}_{t,t+\tau}(\Gamma) = -\bar{\mathbf{J}}_{t,t+\tau}(\Gamma) \cdot \mathbf{F}_e V \beta$. We can write,

$$\begin{aligned} & \frac{\text{Prob}(\beta \bar{\mathbf{J}}_{t,t+\tau} \cdot \mathbf{F}_e V|_{A_\delta^+})}{\text{Prob}(\beta \bar{\mathbf{J}}_{t,t+\tau} \cdot \mathbf{F}_e V|_{A_\delta^-})} = \frac{\text{Prob}(\bar{\Omega}_{t,t+\tau}|_{A_\delta^+})}{\text{Prob}(\bar{\Omega}_{t,t+\tau}|_{A_\delta^-})} \\ &= \frac{\int_{\bar{\Omega}_{t,t+\tau}|_{A_\delta^+}} f(\Gamma; 0) d\Gamma}{\int_{\bar{\Omega}_{t,t+\tau}|_{A_\delta^-}} f(\Gamma^*; 0) d\Gamma^*} \end{aligned} \quad (17)$$

Performing the coordinate transformation $\Gamma^* = M^T S^{2t+\tau} \Gamma$ in the denominator, noting that due to the time reversibility of the dynamics, $\bar{\Omega}_{t,t+\tau}(\Gamma^*) = -\bar{\Omega}_{t,t+\tau}(\Gamma)$, and using conservation of probability in phase space, one eventually obtains [14]:

$$\begin{aligned} & \frac{\text{Prob}(\bar{\Omega}_{t,t+\tau}|_{A_\delta^+})}{\text{Prob}(\bar{\Omega}_{t,t+\tau}|_{A_\delta^-})} = \frac{\text{Prob}_t(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\text{Prob}_t(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} \\ &= \left\langle e^{(-\Omega_{0,t} - \Omega_{t,t+\tau} - \Omega_{t+\tau,2t+\tau})} \right\rangle_{\bar{\Omega}_{t,t+\tau} \in A_\delta^+}^{-1} \end{aligned} \quad (18)$$

where Prob_t is the probability distribution at time t , *i.e.*, the distribution whose density $f(\Gamma; t)$ is the solution of the phase continuity equation with initial condition $f(\Gamma; 0)$. $\langle \dots \rangle_{\bar{\Omega}_{t,t+\tau} \in A_\delta^+}^{-1}$ denotes a conditional average, according to the equilibrium density $f(\Gamma; 0)$, under the condition $\bar{\Omega}_{t,t+\tau} \in A_\delta^+$. Taking the logarithm and dividing by τ produces:

$$\begin{aligned} & \frac{1}{\tau} \ln \frac{\text{Prob}_t(\bar{\Omega}_{0,\tau}|_{A_\delta^+})}{\text{Prob}_t(\bar{\Omega}_{0,\tau}|_{A_\delta^-})} \\ &= A + \varepsilon(A, t, \tau, \delta) \end{aligned}$$

$$-\frac{1}{\tau} \ln \left\langle e^{(-t\bar{\Omega}_{0,t} - t\bar{\Omega}_{t+\tau, 2t+\tau})} \right\rangle_{\bar{\Omega}_{t,t+\tau} \in A_{\delta}^+} \\ \equiv A + \varepsilon(A, t, \tau, \delta) - \frac{1}{\tau} \ln \left(\Xi(t, \tau; A_{\delta}^+) \right) \quad (19)$$

where $\Xi(t, \tau; A_{\delta}^+) = \left\langle \exp(-t\bar{\Omega}_{0,t} - t\bar{\Omega}_{t+\tau, 2t+\tau}) \right\rangle_{\bar{\Omega}_{t,t+\tau} \in A_{\delta}^+}$. The ensemble average is with respect to the initial distribution and $|\varepsilon(A, t, \tau, \delta)| \leq \delta$. This result is exact for all t, τ, δ , and observable pairs $(A, -A)$ [14]. By “observable pairs $(A, -A)$ ” we mean that, for large t , $\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_{\delta}^-} \right) \neq 0$ when $\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_{\delta}^+} \right) \neq 0$.

Under the assumption of ergodic consistency, observability is clearly guaranteed at any finite time t , for all values A within the range allowed by the definition of Ω and by the possible constraints on the dynamics. In other words, the choice of an initial ergodically consistent state guarantees the existence of conjugate trajectory/anti-trajectory pairs in the initial distribution of phases, hence guarantees that Equation (19) holds at all finite times t .

One may distinguish between two types of observability. The first type is easy observability, in which one considers the time averaged dissipative flux within a few standard deviations of the mean. This is an important criterion for experiments or computer simulations. The second type is difficult observability, concerning large deviations from the mean (measured in terms of standard deviations). Of course, the probability of observing sets of trajectories with an average flux many standard deviations from the mean may be exceedingly small! However, unless precluded because of some constant of the motion or of other reasons concerning the definition of Ω , ergodic consistency implies that at any finite time t there is a nonzero probability of observing the conjugate set of anti-trajectories.

Let us begin by assuming that correlations between $e^{-t\bar{\Omega}_{0,t}}$ and $e^{-t\bar{\Omega}_{t+\tau, 2t+\tau}}$ decay instantaneously, with respect to $f(\Gamma; 0)$. The Nonequilibrium Partition Identity [14] (which corresponds to the integral fluctuation theorem in stochastic dynamics) yields:

$$\left\langle e^{-t\bar{\Omega}_{0,t} - t\bar{\Omega}_{t+\tau, 2t+\tau}} \right\rangle_{\bar{\Omega}_{t,t+\tau} \in A_{\delta}^+} = \left\langle e^{-t\bar{\Omega}_{0,t} - t\bar{\Omega}_{t+\tau, 2t+\tau}} \right\rangle \\ = \left\langle e^{-t\bar{\Omega}_{0,t}} \right\rangle \left\langle e^{-t\bar{\Omega}_{t+\tau, 2t+\tau}} \right\rangle = 1, \quad \forall t, \tau \quad (20)$$

Therefore, $\ln \left[\Xi(t, \tau, A_{\delta}^+) \right] / \tau = 0, \forall t, \tau$, and we derive an exact steady state FR that holds for all averaging times τ .

If correlations do not decay instantaneously with respect to $f(\Gamma; 0)$, but vanish to some arbitrarily fine tolerance, for times greater than some cut-off time, $c\tau_M$ (some multiple of the Maxwell time, τ_M , which depends on the tolerance required), one has $\lim_{t \rightarrow \infty} \left(\Xi(t, \tau; A_{\delta}^+) \right) = \text{const} = O(e^{\tau_M})$ [14]. Then the steady state FR is equally obtained in the $\tau \rightarrow \infty$ limit, but with corrections of order $O(\tau_M/\tau)$, which also requires the averaging time τ to be large.

By expanding the exponential terms and noting $\int_s^t \Omega(S^u \Gamma) du = \int_0^{t-s} \Omega(S^{s+z} \Gamma) dz$, it can be seen that the decay in correlation of the two exponential terms is a special instance of the condition,

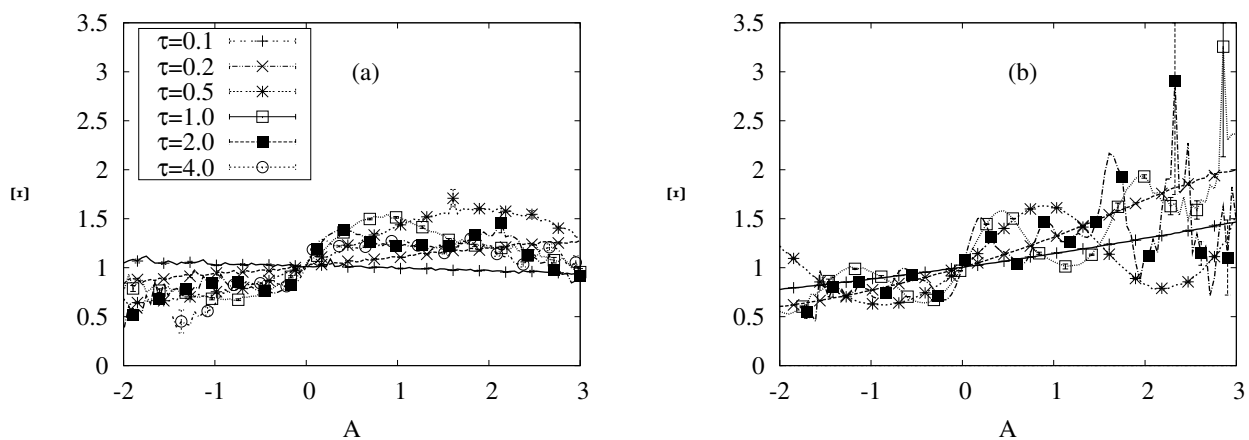
$$\lim_{t \rightarrow \infty} [\langle \varphi(t) \Omega(0) \rangle - \langle \varphi(t) \rangle \langle \Omega(0) \rangle] < O(1/t) \rightarrow 0 \quad (21)$$

now known as T-mixing (first defined in [17]). This guarantees the time integrability of the transient time correlations function $\langle \varphi(t) \Omega(0) \rangle$. Together with the dissipation theorem [14,19], which states that beyond some correlation time t_c phase variables become constant and are given by

$$\begin{aligned}
 \langle \varphi(t) \rangle &\equiv \int \varphi(\mathbf{\Gamma}) f(\mathbf{\Gamma}; t) d\mathbf{\Gamma} \\
 &= \langle \varphi(0) \rangle + \int_0^{t_C} ds \langle \varphi(s) \Omega(0) \rangle \quad \forall t > t_C
 \end{aligned}
 \tag{22}$$

one has that Prob_t tends to a steady state Prob_∞ when $t \rightarrow \infty$ and the averages of smooth phase functions according to Prob_∞ are given by the asymptotic statistics of single dynamical trajectories. The physical conditions for the existence and uniqueness of Prob_∞ are common. For example, the transport coefficients of fluids at low Reynolds or Rayleigh numbers are observed to be single valued functions of temperature, density and possibly the driving force, such as the shear rate or temperature gradient respectively. If they were not single valued functions of these state variables, then the system could not be T-mixing because correlation functions of phase space measures or order parameters for these different states would never decay in time. In T-mixing systems, the $t \rightarrow \infty$ limit of Equation (19) turns from a statement about the time evolving ensemble $f(\mathbf{\Gamma}; t)$ to a statement concerning also the evolution of typical single dynamical trajectories. Then, for time reversal invariant, ergodically consistent, T-mixing systems, the steady state Ω -FR holds with a correction that vanishes as $O(\tau_M/\tau)$.

Figure 1. Plots of the conditional ensemble average, $\Xi(t, \tau; A)$, which is defined in Equation (19), as a function of τ and A . The systems consist of 8 particles in a periodic box undergoing colour diffusion. Here $T = 1$, $\delta = 0.048$, $F_{ex} = 2.0$, $t = 1.0$ for (a) density = 0.8 and (b) density = 1.0 and τ varies between 0.1 and 4.0 for (a) and between 0.1 and 2.0 for (b). The error bars are one standard error in the mean of these (a) 10 and (b) 2 sets of data. The graphs in (a) and in a similar system with 6 particles and density = 0.8 (not presented) are consistent with our proposal that the conditional ensemble average is bounded above and below. This may also be the case in (b), allowing for errors in the mean for large A (not all of which are shown), when $\tau = 1.0$ and 2.0.

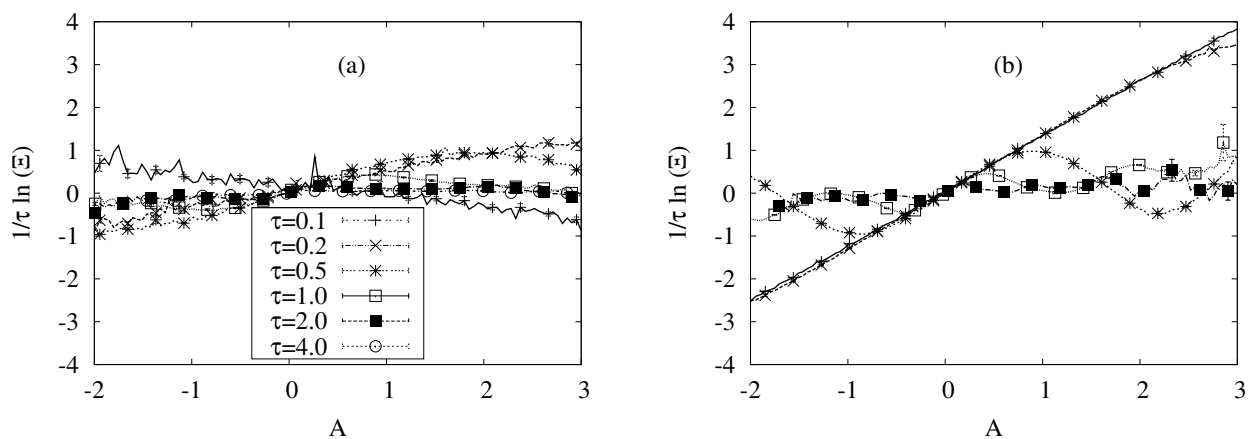


It is now necessary to quantify this process. In Figure 1, we show the results of simulations of 8 WCA [20] particle systems with densities of 0.8 and 1.0 at a temperature of 1. The ensemble averages were over a set of at least 120 million trajectories with $t = 1.0$ and τ varying between 0.1 and 4.0. The results in Figure 1 indicate that, indeed, the conditional ensemble average appearing in Equations (18)

and (19) is bounded. We note that at large A the statistics are poor for $A > 1.5$ when $\tau = 1.0$ and 2.0 , and in all cases for sufficiently large A this will be the case, because we are sampling the tails of the probability distribution of the dissipation function. Furthermore, the value of A at which the statistics become poor will depend on the distribution for the particular system we are considering. However, in both cases (and additional investigations with $N = 6$ gave similar results) the simulations suggest that there exists a constant $K = O(e^{\tau_M}) \approx 2$ for our system, such that $0 < 1/K \leq \Xi(t, \tau, A_\delta^+) \leq K$.

If these scenarios are realised, the last term of Equation (19) vanishes as $1/\tau$, as confirmed by the plots of this quantity in Figure 2, for the same simulations as in Figure 1. An interesting observation from Figures 1 and 2 is the periodic nature of the quantities plotted, with the period appearing to depend both on τ and density. For example, in Figure 1 the period decreases as τ increases and also as the density increases. This is likely to be related to the fact that the simulations were carried out at high densities and therefore one component of the particles' motion involved oscillations within regions formed by the nearest neighbours. As the density decreases, the regions become larger, the periods increase and this component of the motion would diminish.

Figure 2. Plots of $1/\tau \ln(\Xi(t, \tau; A))$, which appears as the final term in Equation (19), as a function of τ and A . The systems are the same as those described in Figure 1, that is, in (a) density = 0.8 and in (b) density = 1.0. The graphs show that the final term in Equation (19) tends to 0 as τ increases. This was also the case for a similar system of 6 particles with density = 0.8 (not presented).



3. Discussion and Conclusions

In accord with our observations, the validity of the steady state Ω -FR,

$$\begin{aligned} & \frac{1}{\tau} \ln \frac{\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_\delta^+} \right)}{\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_\delta^-} \right)} \\ &= A + O\left(\frac{\tau_M}{\tau}\right) + \varepsilon(A, t, \tau, \delta), \quad |\varepsilon| \leq \delta \end{aligned} \quad (23)$$

can be attributed to the following conditions. Firstly we assume the system is time reversal invariant, ergodically consistent, and the initial distribution is even under time reversal. Then, the main technical

assumptions are that the system is reversible and T-mixing. This implies that at a fixed number of standard deviations from the mean, the distribution of the dissipative flux $\bar{\mathbf{J}}_{0,\tau}$ must be Gaussian and furthermore typical average values for the time averaged flux scale as $A \approx O(F_e) \pm O(\tau^{-1/2})$ near the mean for weak fields. For a fixed number of samples, the easily observable range of fields goes to zero as $\tau^{-1/2}$. However, although the easily observable range of values shrinks to zero as $\tau \rightarrow \infty$, the error induced by truncating the correlation function $O(\tau_M/\tau)$ vanishes even faster, ensuring that the steady state Ω -FR

$$\begin{aligned} & \lim_{\tau/\tau_M \rightarrow \infty} \frac{1}{\tau} \ln \frac{\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_\delta^+} \right)}{\text{Prob}_t \left(\beta \bar{\mathbf{J}}_{0,\tau} \cdot \mathbf{F}_e V \Big|_{A_\delta^-} \right)} \\ &= A + \varepsilon(A, t, \tau, \delta), \quad |\varepsilon| \leq \delta \end{aligned} \quad (24)$$

can be easily verified to arbitrary accuracy. By going to larger τ , we can observe the steady state Ω -FR to arbitrarily high accuracy by taking ever more samples. Of course if the correlation time τ_M is zero, the asymptotic limit is unnecessary since $\tau/\tau_M = \infty$, $\forall \tau$ and (24) is exact for all values of τ .

The lack of bounds on the range of values for the argument is highly significant. It means that provided the system is T-mixing, the steady state Ω -FR is valid and verifiable outside the region of Gaussian statistics near the mean, and actually describes a symmetry of rare events in nonequilibrium steady states.

This goes to the root of the fundamental difference between the second law of thermodynamics and the fluctuation theorems. The second law says that certain probabilities are actually zero. In contradistinction, the fluctuation theorems say that certain probabilities may be exceedingly small—unobservably small for macroscopic bodies over accessible timescales—but they precisely quantify these probabilities. Thus the transient Ω -FR and the steady state Ω -FR for T-mixing systems are valid for all values of their argument that can be generated by trajectories originating in the ostensible phase space. Because of ergodic consistency, regardless of how strong or weak the field is, we can prove that the steady state Ω -FR for T-mixing systems will eventually be satisfied. The asymptotic time required for the convergence of the steady state Ω -FR does not scale with the applied field. T-mixing is to be contrasted with the mixing property of stationary measures for dynamical systems. T-mixing does not consider stationary averages but rather concerns itself with transient averages. Our approach is novel and, among other issues, identifies the physically relevant time scales (the decorrelation times with respect to $f(\mathbf{\Gamma}; 0)$). If, on the other hand, the system is not T-mixing and a steady state is not reached, the transient relations, as well as Equation (19), still hold for ensembles of systems at all finite times.

Previous work has also resulted in transient and steady state fluctuation relations for arbitrary odd phase variables, referred to as generalised fluctuation relations [14,18,21]. These can be applied to obtain a fluctuation relation for the phase space expansion. Combining Equations (12) and (10) gives a relationship between the instantaneous dissipation function and the phase space expansion rate for a thermostatted system. Then, using the generalised transient fluctuation relation, the transient fluctuation relation for the phase space expansion for the thermostatted system can be written as:

$$\frac{\text{Prob}(\bar{\Lambda}_{0,\tau} \Big|_{A_\delta^+})}{\text{Prob}(\bar{\Lambda}_{0,\tau} \Big|_{A_\delta^-})} = \frac{\int_{\bar{\Lambda}_{0,\tau} \Big|_{A_\delta^+}} f(\mathbf{\Gamma}; 0) d\mathbf{\Gamma}}{\int_{\bar{\Lambda}_{0,\tau} \Big|_{A_\delta^-}} \exp[-\Omega_{0,\tau}(\mathbf{\Gamma})] f(\mathbf{\Gamma}; 0) d\mathbf{\Gamma}}$$

$$\begin{aligned}
&= \frac{\int_{\bar{\Lambda}_{0,\tau}|_{A_\delta^+}} f(\Gamma; 0) d\Gamma}{\int_{\bar{\Lambda}_{0,\tau}|_{A_\delta^+}} \exp[\Lambda_{0,\tau}(\Gamma) - \beta(H_0(S^\tau \Gamma) - H_0(\Gamma))] f(\Gamma; 0) d\Gamma} \\
&= e^{[-A + \varepsilon(\delta, A, \tau)]\tau} \left\langle e^{-\beta(H_0(S^\tau \Gamma) - H_0(\Gamma))} \right\rangle_{\bar{\Lambda}_{0,\tau} \in A_\delta^+}^{-1} \quad (25)
\end{aligned}$$

In the special case where there are no correlations between $\bar{\Lambda}_{0,\tau}$ and $H_0(S^\tau \Gamma) - H_0(\Gamma)$, this will simplify to:

$$\begin{aligned}
\frac{\text{Prob}(\bar{\Lambda}_{0,\tau}|_{A_\delta^+})}{\text{Prob}(\bar{\Lambda}_{0,\tau}|_{A_\delta^-})} &= e^{[-A + \varepsilon(\delta, A, \tau)]\tau} \left\langle e^{-\beta(H_0(S^\tau \Gamma) - H_0(\Gamma))} \right\rangle^{-1} \\
&= e^{[-A + \varepsilon(\delta, A, \tau)]\tau} \quad (26)
\end{aligned}$$

where the final equality is obtained by recognising that the probability ratio must be equal to unity when $A = 0$. We note that this can then be used to write a fluctuation relation for joint probability distributions of $\bar{\Lambda}_{0,\tau}$ and $H_0(S^\tau \Gamma) - H_0(\Gamma)$ considered for stochastic systems in [22–24]. In the deterministic systems studied here, these correlations will be significant, however limits under which these correlations become negligible are important to establish and will be considered in future work.

Similarly, a steady state fluctuation relation for the phase space contraction over the period of length τ can be obtained:

$$\begin{aligned}
\frac{\text{Prob}(\bar{\Lambda}_{t,t+\tau}|_{A_\delta^+})}{\text{Prob}(\bar{\Lambda}_{t,t+\tau}|_{A_\delta^-})} &= \frac{\text{Prob}_t(\bar{\Lambda}_{0,\tau}|_{A_\delta^+})}{\text{Prob}_t(\bar{\Lambda}_{0,\tau}|_{A_\delta^-})} \\
&= \left\langle e^{(-\Omega_{0,t} + \Lambda_{t,t+\tau} - \beta(H_0(S^{t+\tau} \Gamma) - H_0(S^t \Gamma)) - \Omega_{t+\tau, 2t+\tau})} \right\rangle_{\bar{\Lambda}_{t,t+\tau} \in A_\delta^+}^{-1} \quad (27)
\end{aligned}$$

$$= e^{-A + \varepsilon(A, t, \tau, \delta)} \left\langle e^{(-\Omega_{0,t} - \beta(H_0(S^{t+\tau} \Gamma) - H_0(S^t \Gamma)) - \Omega_{t+\tau, 2t+\tau})} \right\rangle_{\bar{\Lambda}_{t,t+\tau} \in A_\delta^+}^{-1} \quad (28)$$

In the simplistic case where there are no correlations between $\bar{\Lambda}_{t,t+\tau}$ and $H_0(S^{t+\tau} \Gamma) - H_0(S^t \Gamma)$ and assuming that T-mixing holds would enable a simple form of the limiting steady state fluctuation relation for the phase space expansion to be obtained.

Acknowledgements

The authors are grateful to Carlos Mejia-Monasterio for insightful remarks and the High Performance Computing unit at Griffith University for their support. LR thanks the European Research Council, for funding under the Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement n. 202680. The EC is not liable for any use that can be made on the information contained herein. DJS and DJE thank the Australian Research Council for its support.

References

1. Evans D.J.; Cohen, E.G.D.; Morriss, G.P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **1993**, *71*, 2401–2404.

2. Evans, D.J.; Morriss, G.P. *Statistical Mechanics of Nonequilibrium Liquids*; Academic: London, UK, 1990.
3. Jepps, O.; Searles, D.J.; Evans D.J. The fluctuation theorem and Lyapunov weights. *Physica D* **2004**, *187*, 326–337.
4. Parry, W. Synchronisation of canonical measures for hyperbolic attractors. *Comm. Math. Phys.* **1986**, *106*, 267–275.
5. Vance, W.N. Unstable periodic orbits and transport properties of nonequilibrium steady states. *Phys. Rev. Lett.* **1992**, *69*, 1356–1359.
6. Evans, D.J.; Searles, D.J.; Rondoni, L. Application of the Gallavotti-Cohen fluctuation relation to thermostated steady states near equilibrium. *Phys. Rev. E* **2005**, *71*, 056120.
7. Evans, D.J.; Searles D.J. Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E* **1994**, *50*, 1645–1648.
8. Evans, D.J.; Searles D.J. Steady states, invariant measures, and response theory. *Phys. Rev. E* **1995**, *52*, 5839–5848.
9. Searles, D.J.; Evans, D.J. Ensemble dependence of the transient fluctuation theorem. *J. Chem. Phys.* **2000**, *113*, 3503–3509.
10. Jarzynski, C. Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Cond. Matt. Phys.* **2011**, *2*, 329–351.
11. 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.* **2002**, *89*, 050601.
12. Wang, G.M.; Carberry, D.M.; Reid, J.C.; Sevick, E.M.; Evans, D.J. Demonstration of the steady-state fluctuation theorem from a single trajectory. *J. Phys. Cond. Matt.* **2005**, *17*, S3239–S3244.
13. 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* **2005**, *71*, 046142.
14. Searles, D.J.; Evans, D.J.; Rondoni, L. The steady state fluctuation relation for the dissipation function. *J. Stat. Phys.* **2007**, *128*, 1337–1363.
15. Williams, S.R.; Searles, D.J.; Evans, D.J. Numerical study of the steady state fluctuation relations far from equilibrium. *J. Chem. Phys.* **2006**, *124*, 194102.
16. Marconi, U.M.B.; Puglisi, A.; Rondoni, L.; Vulpiani, A. Fluctuation-dissipation: Response theory in statistical physics. *Phys. Rep.* **2008**, *461*, 111–195.
17. Evans, D.J.; Searles, D.J.; Williams, S.R. On the probability of violations of Fourier’s law for heat flow in small systems observed for short times. *J. Chem. Phys.* **2010**, *132*, 024501.
18. Evans, D.J.; Searles, D.J. The fluctuation theorem. *Ad. Phys.* **2002**, *52*, 1529–1585.
19. Evans, D.J.; Searles, D.J.; Williams, S.R. On the fluctuation theorem for the dissipation function and its connection with response theory. *J. Chem. Phys.* **2008**, *128*, 014504.
20. Weeks, J.D.; Chandler, D.; Andersen, H.C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.* **1971**, *54*, 5237–5247.

21. Searles, D.J.; Ayton, G.; Evans, D.J. Generalised fluctuation formula. *AIP Conference Series*, **519**, 2000, 271–280.
22. García-García, R.; Domínguez, D.; Lecomte, V.; Kolton, A.B. Unifying approach for fluctuation theorems from joint probability distributions. *Phys. Rev. E* **2010**, *82*, 030104.
23. García-García, R.; Lecomte, V.; Kolton A.B.; Domínguez, D. Joint probability distributions and fluctuation theorems. *J. Stat. Mech.* **2012**, doi:10.1088/1742-5468/2012/02/P02009.
24. Noh, J.D.; Park, J.-M. Fluctuation relation for heat. *Phys. Rev. Lett.* **2012**, *108*, 240603.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).