Fluctuation theorem for Hamiltonian systems: Le Chatelier's principle

Denis J. Evans,[1] Debra J. Searles,[2] and Emil Mittag[1]
[1]*Research School of Chemistry, Australian National University, Canberra ACT 2601, Australia*
[2]*School of Science, Griffith University, Brisbane Qld 4111, Australia*
(Received 26 December 2000; published 16 April 2001)

For thermostated dissipative systems, the fluctuation theorem gives an analytical expression for the ratio of probabilities that the time-averaged entropy production in a finite system observed for a finite time takes on a specified value compared to the negative of that value. In the past, it has been generally thought that the presence of some thermostating mechanism was an essential component of any system that satisfies a fluctuation theorem. In the present paper, we point out that a fluctuation theorem can be derived for purely Hamiltonian systems, with or without applied dissipative fields.

DOI: 10.1103/PhysRevE.63.051105 PACS number(s): 05.20.−y, 47.10.+g

The fluctuation theorem [1–3] (FT) gives a general formula for the logarithm of the probability ratio that in a thermostated dissipative system, the time-averaged entropy production $\overline{\Sigma}_t$ takes a value A to minus the value, $-A$,

$$\Pr(\overline{\Sigma}_t/k_B=A)/\Pr(\overline{\Sigma}_t/k_B=-A)=\exp[At]. \tag{1}$$

From this equation it is obvious that as the averaging time or system size increases, it becomes exponentially likely that the entropy production will be positive. The fluctuation theorem is important for at least three reasons: first, it gives an expression for the probability that in a finite system observed for a finite time, the Second Law will be violated; secondly, it gives one of the very few exact fluctuation relations that are known for nonequilibrium steady states, even far from equilibrium; thirdly, it can be derived using some of the standard results of the mathematical theory of dynamical systems theory [3].

The theorem was initially proposed [1] for nonequilibrium steady states that are thermostated in such a way that the total energy of the system is constant. Subsequently, it was shown by Gallavotti and Cohen [3] that the theorem could be proved for sufficiently chaotic, isoenergetic nonequilibrium systems using the Sinai-Ruelle-Bowen measure [3]. For transient trajectory segments that start at $t=0$ from an ensemble of initial phase-space vectors $\mathbf{\Gamma}(0)$, and evolve in time under the influence of a reversible deterministic thermostat and an applied dissipative field towards a unique nonequilibrium steady state, a transient fluctuation theorem can be derived using the Liouville measure [2,4]. It has also been shown that the theorem is valid for a wide class of stochastic nonequilibrium systems [5].

It has recently be shown [4] that if initial phases are sampled from a known N-particle phase-space distribution function, $f(\mathbf{\Gamma},0)$, and if we define a dissipation function, $\Omega(\mathbf{\Gamma})$, by

$$t\overline{\Omega}_t \equiv \int_0^t ds\, \Omega(\mathbf{\Gamma}(s)) = \ln\left\langle \frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}\right\rangle - \int_0^t \Lambda(\mathbf{\Gamma}(s))ds, \tag{2}$$

where $\Lambda(\mathbf{\Gamma})\equiv \partial/\partial\mathbf{\Gamma}\cdot\dot{\mathbf{\Gamma}}$ is the phase-space compression factor, then one can derive a transient fluctuation theorem,

$$\frac{\Pr(\overline{\Omega}_t=A)}{\Pr(\overline{\Omega}_t=-A)}=\exp[At]. \tag{3}$$

Thermostats lead to nonzero expressions for the phase-space compression factor. Equation (2) is consistent with all known deterministic transient fluctuation theorems covering a wide variety of initial ensemble types and thermostating mechanisms [4].

This relationship (3) has been tested using computer simulations for a range of nonequilibrium steady-state systems in which the phase-space contraction is nonzero [2,4,5(b),6]. It predicts that the dissipation function has a definite sign that is consistent with the Second Law of thermodynamics. In the present paper, we point out that Eq. (3) can be applied to purely conservative systems where there is no phase-space contraction, $\Lambda(t)\equiv 0$.[1] We consider two cases in detail: the adiabatic (unthermostated) response of a system to a dissipative applied field, and the free relaxation of density inhomogeneities in a system to which no dissipative fields or thermostats are applied. These examples are of general interest since the fine-grained Gibbs entropy $S_G(t) \equiv \int d\mathbf{\Gamma}\, f(\mathbf{\Gamma}(t))\ln[f(\mathbf{\Gamma}(t))]$ is a constant of motion in both systems.

Consider a system of N interacting particles subject to a color field F_c. The total Hamiltonian is $H(\mathbf{\Gamma})=H_0(\mathbf{\Gamma}) + F_c\Sigma_{i=1}^N c_i y_i$, where $c_i=(-1)^i$ and $H_0(\mathbf{\Gamma})=K(\mathbf{p})+\Phi(\mathbf{q})$ is the Hamiltonian for N particles interacting via the WCA potential $\Phi(\mathbf{q})=\Sigma_{i=1}^{N-1}\Sigma_{j>i}^N \phi(|\mathbf{q}_i-\mathbf{q}_j|)$ with $\phi(q)=4[q^{-12} - q^{-6}]$, $q<2^{1/6}$; $=0$ otherwise. We assume the N particles

[1]We note that over many years there have been treatments of quasi-steady-state systems by modeling a small Hamiltonian dissipative subsystem in contact with a *large* Hamiltonian thermal reservoir. In the limit where the thermal reservoir has infinitely many degrees of freedom, the subsystem can be regarded as being in a steady state. For example, see Ref. [7].

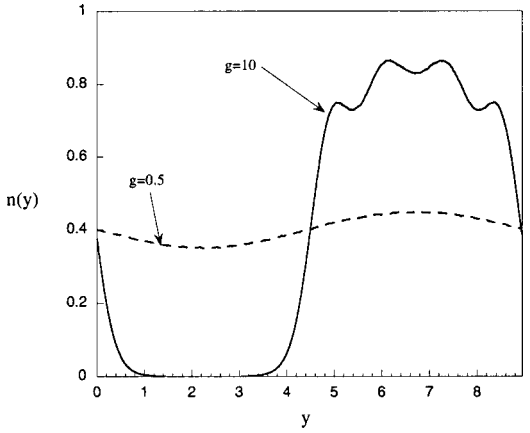

FIG. 1. The density profile for systems of $N=32$ particles in two Cartesian dimensions at $n=0.4$ and $T=1.0$ with $\Phi_g = \Sigma_{i=1}^{N} g \sin ky_i$, where $g=0.5$ (dashed line) or 10.0 (solid line) and $k=2\pi/\sqrt{V}$.

populate a cubic cell of volume V in d Cartesian dimensions and that the system is periodic in the spatial coordinates $\mathbf{q}$.

We assume that the initial ensemble of phases characterized by a normalized N-particle phase-space distribution $f(\mathbf{\Gamma},0) \sim \exp[-\beta H_0(\mathbf{\Gamma})]$, where β is the usual Boltzmann factor $\beta = 1/k_B T$ and T is the absolute temperature. The dissipative flux $J(\mathbf{\Gamma})$ is easily seen to be $\dot{H}_0^{\rm ad} \equiv -J(\mathbf{\Gamma})VF_c = -F_c \Sigma c_i \dot{y}_i$. The superscript ''ad'' denotes that the time derivative is taken in the absence of any thermostats.

Applying the master fluctuation equation (3) to this problem shows that

$$\ln\left[\frac{\Pr(-\beta \overline{J_t} VF_c = A)}{\Pr(-\beta \overline{J_t} VF_c = -A)}\right] = At. \tag{4}$$

It is important to note that the Boltzmann factor appearing in this equation refers to the temperature of the *initial* canonical ensemble. It does *not* refer to the time-dependent temperature of the dissipative system. Because the color field does work on the system, on average the system will heat up as time increases. The entropy production inferred from linear irreversible thermodynamics [8] would be $\Sigma(t) = -J(t)VF_c/T(t)$, whereas Eq. (4) refers to the dissipation function defined in Eq. (2), $\Omega \equiv -\beta J(t) VF_c = -J(t)VF_c/k_B T(0)$. To stress the difference between the dissipation function and the entropy production, we will sometimes refer to Ω as the remnant entropy production.

The second system we consider is the same periodic set of interacting WCA particles, this time with no applied color field. However, at $t=0$, the system is at equilibrium under the influence of a nondissipative sinusoidal gravity field, $\Phi_g = \Sigma_{i=1}^{N} g \sin ky_i$ $(k = 2\pi/V^{1/d})$. This field establishes a sinusoidal density variation across the unit cell in the limit of small g. At $t=0$, this field is removed and the system is monitored as it relaxes to equilibrium. As g becomes large, the particles become confined to the upper half of the box if g is positive (see Fig. 1). The initial distribution function for this system is chosen to be canonical:

$$f(\mathbf{\Gamma},0) \sim \exp\{-\beta[K(\mathbf{p}) + \Phi(\mathbf{q}) + \Phi_g(\mathbf{q}_y)]\}. \tag{5}$$

For $t>0$, we observe the free relaxation of the density modulations towards equilibrium. It is straightforward, using Eq. (2), to find that the dissipation function for this system is

$$t\overline{\Omega}_t \equiv \beta g \sum_{i=1}^{N} [\sin ky_i(t) - \sin ky_i(0)], \tag{6}$$

and from Eq. (3) the fluctuation theorem is

$$\ln \frac{\Pr\left(\beta g \sum_{i=1}^{N} [\sin ky_i(t) - \sin ky_i(0)] = A\right)}{\Pr\left(\beta g \sum_{i=1}^{N} [\sin ky_i(t) - \sin ky_i(0)] = -A\right)} = A. \tag{7}$$

In this equation, β again refers to the temperature of the initial ensemble rather than the time-dependent temperature of the relaxing system. Equation (7) shows that as time increases, it becomes overwhelmingly likely that the initial density inhomogeneities will disappear. This particular proof of Le Chatelier's principle can be generalized to arbitrary nondissipative perturbing fields, Φ_g.

In order to test the fluctuation theorem for the two cases, we carried out molecular-dynamics simulations of $N=32$ Weeks-Chandler-Andersen particles in two Cartesian dimensions. In both cases, the initial system had a particle density of $n=0.4$ and a temperature of $T=1.0$. Molecular-dynamics simulations using a Nosé-Hoover thermostat [9,10] with a heat bath of mass $Q=10$ were used to generate the initial canonical ensembles for the two systems. Since the initial systems are ergodic, Monte Carlo techniques could also have been used to generate the initial canonical distributions.

In the first case, we measure the adiabatic response of the system to a color field applied in the x direction. From a single trajectory, initial canonically distributed phases $\mathbf{\Gamma}(0)$ were sampled at regular intervals. For these phases, a color field was applied and the thermostat disabled and transient trajectories were generated according to the equations of motion:

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathbf{i} c_i F_c, \end{aligned} \tag{8}$$

where $\mathbf{F}_i = -\partial\Phi/\partial\mathbf{q}_i$, F_c is the imposed color field, and $c_i = (-1)^i$ is the color of the ith particle, which determines its response to the color field. Simulations were carried out for systems employing a color field of $F_c = 0.4$.

Figure 2 shows the probability histogram for the time-averaged value of the remnant entropy production, $\overline{\Omega}_t = -\beta \overline{J_t} VF_c$, for different trajectory segment lengths $t = 0.2$, 0.8, 1.6, and 4.0. As the trajectory segment becomes longer, the average value of the remnant entropy production increases due to the work done by the field, and the standard deviation of the distribution decreases due to the longer

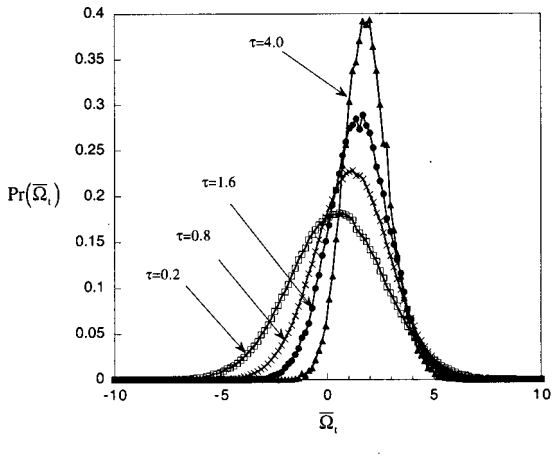

FIG. 2. Probability histograms for the time-averaged remnant entropy production obtained for a system of $N=32$ particles in two Cartesian dimensions at $n=0.4$ and $T=1.0$ subject to a color field of $F_c=0.4$ and averaged over trajectory segments of length $t=0.2$, 0.8, 1.6, and 4.0.

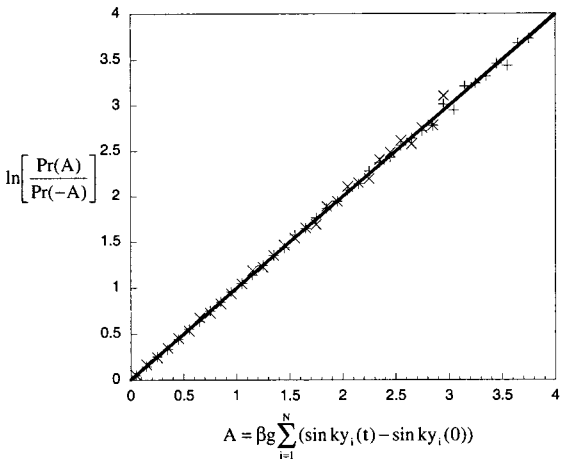

FIG. 4. A test of the FT given by Eq. (7) for a system of $N=32$ particles in two Cartesian dimensions at $n=0.4$ and $T=1.0$, initially subject to a gravity field of $g=0.5$. If Eq. (7) is valid, a slope of unity is obtained, which is shown by the solid line. The trajectories were of length $t=2.0$ $(\times)$ and $t=4.0$ $(+)$.

averaging time. In Fig. 3, the value of $\ln[\Pr(A)/\Pr(-A)]$ is plotted as a function of A, where $A=\overline{\Omega}_t$ or $A=\overline{\Sigma}/k_B=-\overline{\beta J_t}VF_c$. If the FT given by Eq. (4) is valid, a line of slope t will be obtained. Clearly, as expected, Eq. (4) is verified. However, if we also test the relationship using the irreversible thermodynamic entropy production rather than the remnant entropy production, a divergence from linear behavior is observed. We note that for the heating rates that occur in this experiment, local thermodynamic equilibrium is an excellent approximation. We therefore use the instantaneous kinetic temperature to estimate the thermodynamic temperature required in the thermodynamic definition of the entropy production [8].

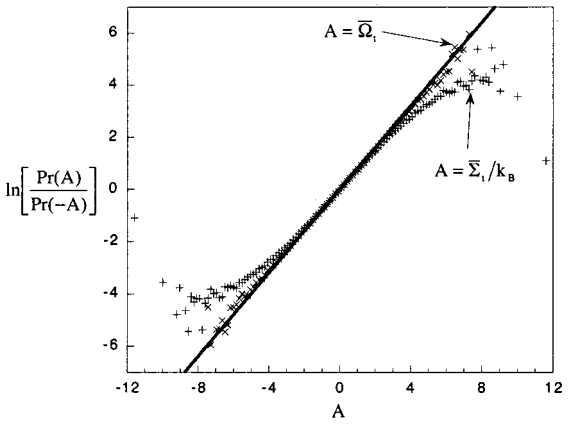

FIG. 3. The logarithm probability ratio of the remnant entropy production $(\times)$ and the irreversible thermodynamic entropy production $(+)$ as a function of the time-averaged remnant entropy production and the irreversible thermodynamic entropy production, respectively, for a system of $N=32$ particles in two Cartesian dimensions at $n=0.4$ and $T=1.0$, subject to a color field of $F_c=0.4$ and averaged over trajectory segments of length $t=0.8$. If the FT given by Eq. (4) is valid, then the data will fall on a straight line of slope $t(0.8)$, as shown by the solid line.

In the second case, we monitor a system with an initially inhomogeneous density profile relaxing towards equilibrium. During this relaxation, the equations of motion for the particles are simply Newtonian (no thermostats or external fields). The mean of the probability histogram for $\overline{\Omega t}=\beta g\Sigma_{i=1}^{N}[\sin ky_i(t)-\sin ky_i(0)]$ shifts from zero when t approaches zero to $-\beta g\langle \sin ky(0)\rangle$ as $t\to\infty$ since we expect that $\beta g\langle \sin ky(t)\rangle$ will approach zero as the density becomes homogeneous and the system approaches equilibrium. In this case, the variance of the distribution will approach a constant value as t increases. We test the FT for this system, given by Eq. (7), in Fig. 4. The initial sinusoidal gravity field was $g=0.5$ with trajectory segments of length $t=2.0$ and 4.0. In both cases, a straight line of slope unity is obtained, as expected from Eq. (7).

The FT given by Eq. (7) suggests that if $\overline{\Omega t}=\beta g\Sigma_{i=1}^{N}[\sin ky_i(t)-\sin ky_i(0)]$ is positive, the logarithm of the probability ratio is positive and hence $\overline{\Omega t}$ is more likely to be positive than negative. This means that the initial density fluctuation must decay from an inhomogeneous system where if $g>0$, $\beta g\langle \sin ky\rangle<0$ (see the density profiles in Fig. 1) to a homogenous system where $\beta g\langle \sin ky\rangle=0$.

We have shown that phase-space compression that results from the application of deterministic thermostats is *not* an essential element of systems that satisfy the fluctuation theorem. We have developed a generalization of the fluctuation theorem that applies to an ensemble of adiabatic dissipative systems. The FT so developed shows that with increasing system size and observation time, it becomes exponentially likely that the dissipative flux will flow in the direction predicted by the Second Law of thermodynamics. However, unlike the situation for *thermostated* nonequilibrium systems, the new adiabatic FT does not involve time averages of the entropy production that one would infer from standard irreversible thermodynamics. Further, we have verified that the

corresponding FT, which employs the standard expressions for the entropy production, is not valid.

We have also developed a version of the fluctuation theorem that applies to the free relaxation of isolated Hamiltonian systems towards equilibrium. This constitutes a statistical mechanical proof of Le Chatelier's principle, in the sense that as the observation time and system size increase, it becomes overwhelmingly likely that any initial deviations from equilibrium will decay rather than grow spontaneously. This proof can trivially be generalized to arbitrary nondissipative fields, $\Phi_g(\mathbf{q})$.

The fluctuation theorem would appear to run counter to Loschmidt's Umkehreinwand [11]. There are two points we make in this regard. First, in our proof, *perfect* time-reversal symmetry is broken by our assumption of causality [2(c)]. We compute the required probabilities of phase-space trajectory time averages from *initial* rather than from *final* states. Few would object to this assumption. However, had we computed the required probabilities from the final rather than the initial states, we would have derived an antifluctuation theorem [2(c)], which would predict the overwhelming violation of the Second Law of thermodynamics.

Second, since the equations of motion are time-reversible, as Loschmidt observed, for every phase-space trajectory, its conjugate time-reversed *antitrajectory* is also observable dynamically. The fluctuation theorem that we have derived gives an expression for the ratio of probabilities of observing time-averaged properties that arise from these time-reversed trajectory pairs. Our computer simulation results confirm the validity of the theorem. The particular properties and ratio given by the fluctuation theorem confirm that in adiabatic dissipative systems, the sign of the entropy production is overwhelmingly likely to be in accord with the Second Law of thermodynamics. In isolated Hamiltonian systems, time evolution of initial inhomogeneities will, with overwhelming likelihood, be in accord with Le Chatelier's principle and the system will, with overwhelming likelihood, relax toward equilibrium.

We would like to thank the Australian Research Council for the support of this project. We would also like to thank Professor E. G. D. Cohen for helpful discussions. D.J.E. would like to thank NIST, Boulder, CO for partial support of this research.

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morris, Phys. Rev. Lett. **71**, 2401 (1993).

[2] (a) D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994); (b) **52**, 5839 (1995); (c) **53**, 5808 (1996); (d) D. J. Searles, G. Ayton, and D. J. Evans, in *Statistical Physics*, edited by Michio Tokuyama and H. Eugene Stanley, AIP Conf. Proc. No. 519 (AIP, Melville, NY, 2000), p. 271.

[3] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995); J. Stat. Phys. **80**, 931 (1995).

[4] D. J. Searles and D. J. Evans, J. Chem. Phys. **113**, 3503 (2000).

[5] (a) J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999); (b) D. J. Searles and D. J. Evans, Phys. Rev. E **60**, 159 (1999).

[6] D. J. Searles and D. J. Evans, Int. J. Thermophys. (to be published).

[7] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).

[8] S. R. De Groot and P. Mazur, *Non-equilibrium Thermodynamics* (Dover, New York, 1984).

[9] D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Academic, London, 1990).

[10] S. Nosé, J. Chem. Phys. **81**, 511 (1984); Mol. Phys. **52**, 255 (1984); W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[11] An objection or counterargument based on reasoning involving the reversal of time. J. Loschmidt, Wien Ber. **73**, 139 (1876); **75**, 67 (1877).