Journal of Statistical Physics, Vol. 96, Nos. 5/6, 1999

Note on Two Theorems in Nonequilibrium Statistical Mechanics

E. G. D. Cohen[1] and G. Gallavotti[2]

Received April 6, 1999; final April 30, 1999

An attempt is made to clarify the difference between a theorem derived by Evans and Searles in 1994 on the statistics of trajectories in phase space and a theorem proved by the authors in 1995 on the statistics of fluctuations on phase space trajectory segments in a nonequilibrium stationary state.

KEY WORDS: Chaos; fluctuation theorem; large deviations; chaotic hypothesis; nonequilibrium statistical mechanics; time reversal.

Recently a Fluctuation Theorem (FT) has been proved by the authors (GC),[1] for fluctuations in nonequilibrium stationary states. Considerable confusion has been generated about the connection of this theorem and an earlier one by Evans and Searles (ES)[2] so that it seemed worthwhile to try to clarify the present situation with regards to these two theorems.

In a paper in 1993 by Evans, Cohen, and Morriss,[3] theoretical considerations lead them to a computer experiment about the statistical properties of the fluctuations of a shear stress model (viscous current—or the related entropy production rate—in a thermostatted sheared viscous fluid in a nonequilibrium stationary state.

The Fluctuation Relation found in the simulation,[3] reads in current notation for sufficiently large τ,[1]:

$$\frac{\pi_\tau(p)}{\pi_\tau(-p)} \simeq e^{\tau\sigma_+ p} \tag{1}$$

[1] The Rockefeller University, New York, New York 10021.
[2] Fisica, Universitá di Roma, "La Sapienza," 00185 Rome, Italy.

0022-4715/99/0900-1343$16.00/0 © 1999 Plenum Publishing Corporation

Here $\pi_\tau(p)$ is the probability of observing an average phase space contraction rate (which in the models considered has the interpretation of average entropy production rate) of size $p\sigma_+$ on one of many segments of duration τ on a long phase space trajectory of the dynamical system modeling the shearing fluid in a nonequilibrium stationary state; here $\sigma(x)$ will denote the phase space contraction rate near a phase point x (i.e., the divergence of the equations of motion) and σ_+ is the average phase space contraction rate over positive infinite times so that p is a dimensionless characterization of the phase space contraction (with time average 1). The approximation within which Eq. (1) was observed was very convincing.[3]

Under suitable assumptions, see below, Eq. (1) was *derived* in ref. 1 in the form (clearly already intended by the authors of Eq. (1)):

$$\lim_{\tau \to \infty} \frac{1}{\tau\sigma_+} \ln \frac{\pi_\tau(p)}{\pi_\tau(-p)} = p \tag{2}$$

Later several other computer experiments have confirmed the relation Eq. (2).[4–6]

The original computer experiment,[3] was inspired by a theoretical argument for the relative probabilities to find a phase space trajectory segment of length τ in a state x with phase space contraction rate p and in a state x' with rate $-p$. These theoretical considerations lead to the correct prediction Eq. (1), which was confirmed by the computer experiment whose validity, however, does not rely on the theory. Although the theoretical arguments in favor of (1) in ref. 3 contain hints for a theoretical derivation of (1) based on the SRB distribution and the use of time reversal, these hints cannot be regarded as constituting a proof or justification of (1) and (2).

In 1994, Evans and Searles[2] gave a derivation of a theorem which had a similar form as Eq. (1). More precisely: let E_p be the set of initial conditions of a dynamical system for phase space trajectories along which the phase space contraction is $e^{-p\sigma_+T}$ in a time T. We denote by $\mu_L(E_p)$ its Liouville measure. Similarly, let $\mu_L(E_{-p})$ be the Liouville measure of the corresponding set of phase space trajectories along which the phase space contraction in time T is $e^{p\sigma_+T}$. Quite generally, and in all models considered in the literature relevant here, $E_{-p} = IS_T E_p$, if S_t is the time evolution (Liouville) operator of the system, so that $t \to S_t x$ is the phase space trajectory at time t starting at x at $t=0$, and if I denotes the time reversal operation. Hence E_{-p}, the set of points around which phase space contracts at rate $-p\sigma_+ T$, is obtained by evolving forward over a time T those in E_p (which would contract by $p\sigma_+ T$) and then inverting the velocities by the time reversal operator I. In fact, the sets E_p and E_{-p} are those considered by Evans and Searles in ref. 2.

Then the proof in ref. 2 is the following:

$$\frac{\mu_L(E_p)}{\mu_L(E_{-p})} = \frac{\mu_L(E_p)}{\mu_L(IS_T E_p)} = \frac{\mu_L(E_p)}{\mu_L(S_T E_p)} = \frac{\mu_L(E_p)}{\mu_L(E_p)\, e^{-p\sigma_+ T}} = e^{p\sigma_+ T} \tag{3}$$

where one has used that the Liouville distribution is time reversal invariant, i.e., $\mu_L(E) \equiv \mu_L(IE)$ (although it is not stationary) to get the second equality as well as the definition of phase space contraction in the third equality.

The arbitrary time interval T includes the short times referring to the transient behavior of the system before possibly reaching the nonequilibrium stationary state. In the derivation of Eq. (3) only time reversal symmetry is used. Later,[7] it was argued that under this assumption alone, Eq. (3) also holds in the nonequilibrium stationary state μ_∞, since Eq. (3) is valid for any T and the Liouville distribution μ_L would evolve in a sufficiently long time T into a distribution $S_T\mu_L$ arbitrarily close to a nonequilibrium stationary state μ_∞.

Therefore the Eq. (3) was asserted in ref. 7 to be stronger than Eq. (2) (i.e., to imply it), which, however, refers to the statistics of trajectory segments, along a trajectory in a chaotic nonequilibrium stationary state μ_∞, *not to the statistics of independent trajectory histories emanating from the initial Liouville distribution* μ_L under the time reversibility assumption.

In 1995 the authors proved Eq. (2) based on a dynamical assumption, called *Chaotic Hypothesis* (CH), which assured Strong chaoticity ("Anosov system-like behavior") for the systems for which Eq. (2) held. In that work the name Fluctuation Theorem (FT) was first introduced for Eq. (2), and was proposed as an explanation for the experimental result Eq. (1). We will call this the GCFT.

It is worthwhile to emphasize again that, while the right hand sides of the Eqs. (1) and (3) look very similar, they, as well as the left hand sides of these equations, really refer to entirely different physical situations.

Equation (3)[2] holds for any T on trajectories with initial data sampled from the Liouville distribution at $t=0$ and it can be considered as a simple, but interesting, consequence, for reversible systems, of the very definition of phase space contraction. We will call it here the ESI, where the I refers to "identity." The ESI is much more general than the FT in Eq. (2), which needs, *in addition* to phase space contraction ($\sigma_+ > 0$) and time reversal symmetry, also the Chaotic Hypothesis. The proof of the ESI, fully described in Eq. (3) above, is identical in essence to the proof in ref. 2 which is much more involved.

In order to illustrate the fundamental difference between the two theorems we first give an example of a very simple case where the more general ESI Eq. (3) holds, while the GCFT Eq. (2) does not.

To that end we consider a single charged particle in a periodic box, with charge e moving in an electric field $\mathbf{E}$, i.e., a Lorentz gas without scatterers, and subject to a Gaussian thermostat (to obtain a nonequilibrium stationary state):

$$\dot{\mathbf{q}} = \mathbf{p}, \qquad \dot{\mathbf{p}} = e\mathbf{E} - \alpha\mathbf{p} \tag{4}$$

where the "thermostat" force $-\alpha\mathbf{p}$, with $\alpha = e\mathbf{E}\cdot\mathbf{p}/|\mathbf{p}|^2$, assures the reaching of a nonequilibrium stationary state of this system.

In this case one can solve explicitly the trivial equations of motion Eq. (4) and check that the Liouville distribution μ_L indeed evolves towards a stationary state μ_∞, which is simply a state in which the particle moves with constant speed parallel to $\mathbf{E}$. The ESI Eq. (3) will hold for the phase space trajectories of this system sampled with the initial Liouville distribution μ_L, but it will not be a fluctuation theorem, since there are no fluctuations. Also, GCFT's Eq. (2) will not hold for the phase space trajectory segment fluctuations of this system, which is not a contradiction because the system is not chaotic.

From this simple example and other similar ones, follows that the two theorems cannot be equivalent, and the validity of Eq. (3) cannot imply much, *without extra assumptions*, about the fluctuations (absent in this case) in the stationary state. Note that Eq. (3) is an identity which is always valid in the systems considered. The system of Eq. (4) is therefore a counterexample to the statement that Eq. (3) implies Eq. (2), i.e., to the statement,[(7)] that ESI implies GCFT.

Second, and more interestingly, one can try to derive the GCFT Eq. (2) from the more general ESI Eq. (3). One could then try to proceed as follows.

First one would need to show that on a subsequent trajectory segment of length τ, *after* time T, the ratio of the probabilities of finding a phase space contraction of $+p\sigma_+\tau$ to that of finding $-p\sigma_+\tau$ over this segment of length τ, would be given by $e^{p\sigma_+\tau}$. Here $p\sigma_+\tau$ is any preassigned value of the phase space contraction. However, Eq. (3) gives no information whatsoever about the points in $S_{-T}E_{\pm p}$ which after a time T evolve into points which in the next τ units of time show a phase space contraction $\pm p\sigma_+\tau$. In other words, the ESI does not contain the detailed information needed to derive the GCFT.

If one adds the Chaotic Hypothesis to the time reversal symmetry assumptions made about the dynamical system in the ESI, one could use

Sinai's theorem,[8] to assert that such a system, starting from the initial Liouville distribution μ_L, will indeed approach a chaotic nonequilibrium stationary distribution μ_∞ supported (however) on a fractal attractor A with zero Liouville measure $\mu_L(A)=0$. This is, in this case, the SRB distribution, $\mu_{\mathbf{SRB}}$, of the system, which was used in ref. 1. However, for a proof of the GCFT, details of the SRB distribution are needed, which contain just the details considered in ref. 1. That is, one has to make an appropriate (Markov) partition of the phase space and assign weights to the cells of increasingly finer partitions leading, to the SRB distribution. This then allows one to assign appropriate weights to those regions (of *zero Liouville measure*) in phase space that will give rise, to phase space contractions on trajectory segments τ of $\pm p\sigma_+\tau$ leading to the GCFT.

A more precise comparison between the ESI and GCFT requires a more quantitative statement of the latter result. Namely,[1] Eq. (2) can be derived from the stronger relation:

$$\frac{\pi_\tau(p)}{\pi_\tau(-p)}=e^{(p\sigma_+ + O(T_\infty/\tau))\,\tau} \tag{5}$$

where T_∞ is a time scale of the order of magnitude of the time necessary in order that the distribution $S_T\mu_L$, into which the Liouville distribution μ_L evolves in time T, be "practically" indistinguishable from the stationary state that we denote by μ_∞. Here the validity of the fluctuation relation (1) for asymptotically long times τ is more clearly expressed. The existence of the time T_∞ and its role in bounding the error term in Eq. (5) are among the *main results* of ref. 1. The time T_∞ appears in refs. 1 and 10 as the range of the potential that generates the representation of μ_∞ as a Gibbs state using a symbolic dynamic representation of the SRB distribution on the Markov partition.[8]

Examining the ESI derivation above, one easily sees that the following relation

$$\frac{(S_T\mu_L)(E_p)}{(S_T\mu_L(E_{-p})}=\frac{\mu_L(S_{-T}E_p)}{\mu_L(S_{-T}E_{-p})}=e^{(p\sigma_+ + O(T/\tau))\,\tau} \tag{6}$$

can also be derived.[11]

It is important to note that the argument leading to Eq. (3) *cannot say more than this*: in particular one must justify why T—which in principle should be large, *strictly speaking infinite*, so that $S_T\mu_L$ be identifiable with $\mu_{\mathbf{SRB}}$—*can in fact be taken smaller than* τ.

Justifying this requires assumptions (as the above counterexample indicates) like the mentioned Gibbs property of the SRB distribution which is even stronger than requiring exponential decay of correlations (also implied by the CH). Otherwise, without extra assumptions, one has to say that T has to be taken infinite first with the result that Eq. (6), hence Eq. (3), *becomes empty* in content. Of course one can argue that "on physical grounds" T needs not to be taken infinite but just as large as some characteristic time scale for the approach to the attractor: but the precise meaning of this, and the assumptions under which it can be stated, is precisely what needs to be determined particularly because in non-equilibrium systems the attractor A is *fractal* with $\mu_L(A)=0$ and one can very well doubt that the $S_T\mu_L$ distribution is *ever close enough to the* μ_∞ *distribution to allow comparing* Eq. (5) and (6). This requires a convincing argument, were it only because $(S_T\mu_L)(A)=0$

Very recently a paper appeared,[12] which addresses again the same questions of ref. 7. It gives further evidence of the difference between ESI and GCFT. Their experiment gets ESI when they test it and FT when they test that. From our point of view the results had to be expected: in the first case because ESI is a rigorous identity and in the second case because we believe that out Chaotic Hypothesis applies to the experiment so that CGFT should hold.

ACKNOWLEDGMENTS

E.G.D.C. is very much indebted to L. Rondoni for many very clarifying discussions. The authors are also indebted to C. Liverani for suggestions on the proof of Eq. (3) and to F. Bonetto, J. L. Lebowitz, and L. Rondoni for helpful discussions. E.G.D.C. gratefully acknowledges support from DOE Grant DE-FG02-88-ER13847, while G.G. that from Rutgers University's CNR-GNFM, MPI.

REFERENCES

1. G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in non-equilibrium statistical mechanics, *Phys. Rev. Lett.* **74**:2694–2697 (1995); G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in stationary states, *J. Stat. Phys.* **80**:931–970 (1995).
2. D. Evans and D. Searles, Equilibrium microstates which generate the second law violating steady states, *Phys. Rev. E* **50**:1645–1648 (1994).
3. D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Probability of second law violations in shearing steady flows, *Phys. Rev. Lett.* **71**:2401–2404 (1993).
4. F. Bonetto, G. Gallavotti, and P. Garrido, Chaotic principle: An experimental test, *Physica D* **105**:226–252 (1997).
5. F. Bonetto, N. Chernov, and J. Lebowitz, (Global and local) fluctuations of phase space contraction in deterministic stationary nonequilibrium, *Chaos* **8**:823–833 (1998).
6. S. Lepri, R. Livi, and P. Politi, *Physica D* **119**:140 (1998).

7. D. Evans and D. Searles, The conjugate fluctuation theorem and Green–Kubo relations, preprint (1998), revised version in cond-mat #9902021.
8. Y. G. Sinai, Gibbs measures in ergodic theory, *Russian Math. Surveys* **27**:21–69 (1972); and *Lectures in Ergodic Theory*, Lecture Notes in Mathematics (Princeton University Press, Princeton, 1977).
9. D. Ruelle, Smooth dynamics and new theoretical ideas in non-equilibrium statistical mechanics, Lecture notes, Rutgers University, mp_arc #98-770 (1998), or chao-dyn #9812032, in print on *J. Stat. Phys.*
10. G. Gallavotti, Reversible Anosov maps and large deviations, *Mathematical Physics Electronic Journal*, *MPEJ*, (http://mpej.unige.eh) **1**:1–12 (1995).
11. J. L. Lebowitz, private communication.
12. G. Ayton and D. Evans, On the asymptotic convergence of the transient and steady state fluctuation theorems, cond-mat #9903409.

Communicated by J. L. Lebowitz