Extension of Onsager's reciprocity to large fields and the chaotic hypothesis

Giovanni Gallavotti
Fisica, Universitá di Roma, "La Sapienza", 00185 Roma, Italia
(April 3, 1996)

The fluctuation theorem (FT), the first derived consequence of the *Chaotic Hypothesis* (CH) of [1], can be considered as an extension to arbitrary forcing fields of the fluctuation dissipation theorem (FD) and the corresponding Onsager reciprocity (OR), in a class of reversible nonequilibrium statistical mechanical systems.

47.52.+j, 05.45.+b, 05.70.Ln, 47.70.-n

A typical system studied here will be N point particles subject to *(a)* mutual and external conservative forces with potential $V(\vec{q}_1,\ldots,\vec{q}_N)$, *(b)* external (non conservative) forces, *forcing agents*, $\{\vec{F}_j\}$, $j=1,\ldots,N$, whose strength is measured by parameters $\{G_j\}$, $j=1,\ldots,s$, and *(c)* also to forces $\{\vec{\varphi}_j\}$, $j=1,\ldots N$, generating constraints that provides a model for the thermostatting mechanism that keeps the energy of the system from growing indefinitely (because of the continuing action of the forcing agents).

An observable $O(\{\vec{q},\dot{\vec{q}}\})$ evolves in time under the time evolution S_t solving the equations of motion:

$$m\ddot{\vec{q}}_j = -\partial_{\vec{q}_j} V(\vec{q}_j) + \vec{F}_j(\{G\}) + \vec{\varphi}_j \tag{1}$$

with $m=$ particles mass, and $\vec{\varphi}_j$ "thermostatting" forces assuring that the system reaches a (non equilibrium) stationary state. The time evolution of the observable O on the motion with initial data $x=(\vec{q},\dot{\vec{q}})$ is the function $t\longrightarrow O(S_t x)$ so that the motion *statistics* is the probability distribution μ_+ on the *phase space* $\mathcal{C}$ such that:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T O(S_t x)\,dt = \int_{\mathcal{C}} O(y)\mu_+(dy) \stackrel{def}{=} \langle O\rangle_+ \tag{2}$$

for all data $x\in\mathcal{C}$ except a set of zero measure with respect to the volume $\overline{\mu}_0$ on $\mathcal{C}$. The distribution μ_+ is assumed to exist: a property called *zero-th law*, [2,3].

The thermostatting mechanism will be described by force laws $\vec{\varphi}_j$ which enforce the constraint that the kinetic energy (or the total energy) of the particles, or of subgroups of the particles, remains constant, [4]. It is convenient also to imagine that the constraints keep the total kinetic energy bounded (hence phase space is bounded). The constraint forces $\{\vec{\varphi}_j\}$ will be supposed such that the system is *reversible*: this means that there will be a map i, defined on phase space, anticommuting with time evolution: *i.e.* $S_t\, i \equiv i\, S_{-t}$.

Reversibility is a key assumption, [1,5].

In [2] the divergence of the r.h.s. of Eq. (1) is a quantity $-\sigma(x)$ defined on phase space that has been identified with the *entropy production rate*.

The *chaotic hypothesis*, (CH), of [1] implies a *fluctuation theorem* or (FT) which, [2], is a property of the fluctuations of the entropy production rate. Namely if we denote $\langle\sigma\rangle_+ = \int_{\mathcal{C}}\sigma(y)\mu_+(dy)$ the time average, over an infinite time interval by Eq. (2), then the *dimensionless* finite time average $p=p(x)$:

$$\frac{1}{\tau}\int_{-\tau/2}^{\tau/2}\sigma(S_t x)\,dt \stackrel{def}{=} \langle\sigma\rangle_+\, p \tag{3}$$

has a statistical distribution $\pi_\tau(p)$ with respect to the stationary state distribution μ_+ such that:

$$\frac{1}{\tau\langle\sigma\rangle_+\, p}\log\frac{\pi_\tau(p)}{\pi_\tau(-p)} \longrightarrow_{\tau\to\infty} 1 \tag{4}$$

provided (of course) $\langle\sigma\rangle_+>0$. Following [1] a reversible system for which $\langle\sigma\rangle_+>0$ will be called *dissipative.*

Ruelle's H-theorem, [5], states that $\langle\sigma\rangle_+>0$ unless the stationary distribution μ_+ has the form $\rho(x)\overline{\mu}_0(dx)$.

Hence we shall suppose that the system is dissipative when the forcing $\vec{G}$ does not vanish and, without real loss of generality, that $\sigma(ix)=-\sigma(x)$ and that $\sigma(x)\equiv 0$ when the external forcing vanishes, writing:

$$\sigma(x) = \sum_{i=1}^{s} G_i J_i^0(x) + O(G^2) \tag{5}$$

Assuming that the σ–σ correlations have a fast decay (a fact to be expected if the (CH) is accepted, [6]) then by a result of Sinai [6] the entropy verifies a *limit theorem*; *i.e.*:

$$\lim_{\tau\to\infty}\frac{1}{\tau}\log\pi_\tau(p) = -\zeta(p) \tag{6}$$

where $\zeta(p)$ is analytic for p in the interval $[-p^*,p^*]$ within which it can vary (model dependent) [7,8]. The function $\zeta(p)$ can be conveniently computed because its transform $\lambda(\beta) = \lim_{\tau\to\infty}\frac{1}{\tau}\log\int e^{\beta\tau\,(p-1)\langle\sigma\rangle_+}\pi_\tau(p)dp$ can be expressed by a *cumulant expansion.* Once $\lambda(\beta)$ is "known" then $\zeta(p)$ is recovered via a Legendre transform; $\zeta(p) = \max_\beta\big(\beta\langle\sigma\rangle_+(p-1)-\lambda(\beta)\big)$, [7,8].

By using the cumulant expansion for $\lambda(\beta)$ we find that $\lambda(\beta)=\frac{1}{2!}\beta^2C_2+\frac{1}{3!}\beta^3C_3+\ldots$ where the coefficients C_j are $\int_{-\infty}^{\infty}\langle\sigma(S_{t_1}\cdot)\sigma(S_{t_2}\cdot)\ldots\sigma(S_{t_{j-1}}\cdot)\sigma(\cdot)\rangle_+^T\,dt_1\ldots$ if $\langle\ldots\rangle_+^T$ denote the cumulants of the variables $\sigma(x)$.

In our case the cumulants of order j have size $O(G^j)$, by Eq. (5), so that:

$$\zeta(p) = \frac{\langle\sigma\rangle_+^2}{2C_2}(p-1)^2 + O((p-1)^3(G^3) \qquad (7)$$

(note the first term in r.h.s. giving the central limit theorem). Eq. (7), together with the (FT) (4), yields at fixed p our main relations:

$$\langle\sigma\rangle_+ = \frac{1}{2}C_2 + O(G^3) \qquad (8)$$

We define, [7], the *current* $J_i(x) = \partial_{G_i}\sigma(x)$ and the *transport coefficients* $L_{ij} = \partial_j\langle J_i(x)\rangle_+|_{G=0}$ and we study L_{ij}.

To derive (FD), we first look at the r.h.s. of the first relation in Eq. (8) discarding $O(G^3)$: the r.h.s becomes a quadratic form in G with coefficients:

$$\frac{1}{2}\int_{-\infty}^{\infty} dt\, \left(\langle J_i^0(S_t\cdot)J_j^0(\cdot)\rangle_+ - \langle J_i^0\rangle_+\langle J_j^0\rangle_+\right) \qquad (9)$$

On the other hand the expansion of $\langle\sigma\rangle_+$ in the l.h.s. of Eq. (8) to second order in G gives:

$$\langle\sigma\rangle_+ = \frac{1}{2}\sum_{ij} G_iG_j\left(\partial_i\partial_j\langle\sigma\rangle_+\big|_{G=0}\right) \qquad (10)$$

because the first order term vanishes (by Eq. (5), or (8)). The r.h.s. of (10) is the sum of $\frac{1}{2}G_iG_j$ times $\partial_j\partial_j\int\sigma(x)\mu_+(dx)$ which equals the sum of the following three terms: the first is $\int\partial_i\partial_j\sigma(x)\mu_+(dx)$, the second is $\int\partial_i\sigma(x)\partial_j\mu_+(dx) + (i \leftrightarrow j)$ and the third is $\int\sigma(x)\partial_i\partial_j\mu_+(dx))$, all evaluated at $G=0$. The first addend is 0 (by time reversal), the third addend is also 0 (as $\sigma = 0$ at $G = 0$). Hence:

$$\partial_i\partial_j\langle\sigma\rangle_+|_{G=0} = \Big(\partial_j\langle J_i^0\rangle_+ + \partial_i\langle J_j^0\rangle_+\Big)|_{G=0} \qquad (11)$$

and it is easy to check, again by using time reversal, that:

$$\partial_j\langle J_i^0\rangle_+|_{G=0} \;=\; \partial_j\langle J_i\rangle_+|_{G=0} \;=\; L_{ij} \qquad (12)$$

Therefore equating the r.h.s and the l.h.s. Eq. (8), as expressed respectively by Eq. (9) and (11) we get a formula for the matrix $\frac{L_{ij}+L_{ji}}{2}$ giving (FD) (*i.e.* Green–Kubo's formula) at least if $i=j$.

We want to show that the above ideas *also* suffice to prove (OR), *i.e.* $L_{ij} = L_{ji}$. The main remark is that we can *extend* the (FT) theorem to give properties of the *joint* distribution of the average of σ, (3), and of the corresponding average of $G_j\partial_j\sigma$. In fact define the *dimensionless* j-current $q = q(x)$ as:

$$\frac{1}{\tau}\int_{-\tau/2}^{\tau/2} G_j\partial_j\sigma(S_tx)dt \stackrel{def}{=} G_j\langle\partial_j\sigma\rangle_+ q \qquad (13)$$

where the factor G_j is there only to keep σ and $G_j\partial_j\sigma$ with the same dimensions.

Then if $\pi_\tau(p,q)$ is the joint probability of p,q the *same* proof of the (FT) in [1] yields also:

$$\lim_{\tau\to\infty}\frac{1}{\tau\langle\sigma\rangle_+p}\log\frac{\pi_\tau(p,q)}{\pi_\tau(-p,-q)} = 1 \qquad (14)$$

and the limit theorem in (7) is extended, [6], to:

$$\zeta(p,q) = \lim_{\tau\to\infty}\frac{1}{\tau}\log\pi_\tau(p,q) \qquad (15)$$

We can compute $\zeta(p,q)$ in the same way as $\zeta(p)$ by considering first the transform $\lambda(\beta_1,\beta_2)$:

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int e^{\tau(\beta_1\,(p-1)\langle\sigma\rangle_+ + \beta_2\,(q-1)\langle G_j\partial_j\sigma\rangle_+)}\pi_\tau(p,q)dpdq \qquad (16)$$

and then the Legendre trasform:

$$\zeta(p,q) = \max_{\beta_1,\beta_2}\big(\beta_1\,(p-1)\langle\sigma\rangle_+ + \beta_2\,(q-1)\langle G_j\partial_j\sigma\rangle_+ -$$

$$-\lambda(\beta_1,\beta_2)\big) \qquad (17)$$

The function $\lambda(\vec\beta)$, $\vec\beta = (\beta_1,\beta_2)$, is evaluated by the cumulant expansion, as above, and one finds:

$$\lambda(\vec\beta) = \frac{1}{2}\,(\vec\beta, C\,\vec\beta) + O(G^3) \qquad (18)$$

where C is the 2×2 matrix of the second order cumulants. The coefficient C_{11} is given by C_2 appearing in (7); C_{22} is given by the same expression with σ replaced by $G_j\partial_j\sigma$ while C_{12} is the mixed cumulant:

$$\int_{-\infty}^{\infty}\left(\langle\sigma(S_t\cdot)\,G_j\partial_j\sigma(\cdot)\rangle_+ - \langle\sigma(S_t\cdot)\rangle_+\,\langle G_j\partial_j\sigma(\cdot)\rangle_+\right)dt \qquad (19)$$

Hence if $\vec w = \begin{pmatrix}(p-1)\langle\sigma\rangle_+\\ (q-1)\langle G_j\partial_j\sigma\rangle_+\end{pmatrix}$ we get:

$$\zeta(p,q) = \frac{1}{2}\left(C^{-1}\vec w,\vec w\right) + O(G^3) \qquad (20)$$

completely analogous to (7). But the (FT) in (14), implies that $\zeta(p,q)-\zeta(-p,-q)$ is q *independent*: this means, as it is immediate to check:

$$(C^{-1})_{22}\langle G_j\partial_j\sigma\rangle_+ - (C^{-1})_{21}\langle\sigma\rangle_+ = 0 + O(G^3) \qquad (21)$$

which because of (8), and of $(C^{-1})_{22} = C_{11}/\det C$, becomes the analogue of (8):

$$\langle G_j\partial_j\sigma\rangle_+ = \frac{1}{2}C_{12} + O(G^3) \qquad (22)$$

and, proceeding as in the derivation of (9) through (12) (*i.e.* expanding both sides of (22) *to first order* in the G_i's and using (19)) we get that $\partial_i \langle \partial_j \sigma \rangle_+$ is given by the integral in (9). This means that $L_{ij} = L_{ji}$ and the (FD) theorem follows toghetr with the (OR).

Thus Eq. (8),(22) and the ensuing (FD), and (OR), are a consequence of (FT), (4), and of its (obvious) extension, (14), in the limit $G \to 0$, when combined with the expansion (7) for entropy fluctuations. Those theorems and the fast decay of the $\sigma\sigma$ correlations are all natural consequences of (CH) for reversible statistical mechanical systems, which is the starting point of our considerations. Note that reversibility is here assumed *both in equilibrium and in non equilibrium*: this is a feature of gaussian thermostat models, [2], but by no means of all models, [4,9].

Of course while the (OR) and (FD) only hold around equilibrium, *i.e.* they are properties of G–derivatives evaluated at $G = 0$, and the expansion for $\lambda(\beta)$ is a general consequence of the correlation decay, the (FT) also holds far from equilibrium, *i.e.* for large G and can be considered a generalization of the (OR) and (FD).

Acknowledgements: I have profited from many discussions and hints from F. Bonetto and P. Garrido, [11]. I am particularly indebted and grateful to E.G.D. Cohen for important comments, criticism, suggestions and much needed encouragement, and to G. Gentile for pointing out an error in the first version of the letter and for many comments. Partial financial support from Rockefeller U., CNR-GNFM, ESI and the EU program: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460.

[1] Gallavotti, G., Cohen, E.G.D.: Physical Review Letters, **74**, 2694–2697, 1995.

[2] Gallavotti, G., Cohen, E.G.D.: J. Stat. Phys., **80**, 931–970, 1995.

[3] Gallavotti, G.: in *mp_arc@ math. utexas. edu*, #95-288, 1995. In print in J. Stat. Phys.

[4] Examples of such thermostats are the gaussian thermostats and a few concrete cases can be found in [2]. The importance of the gaussian thermostats, and of reversibility, and the possibility of linking them with the dynamical systems theory of the SRB distributions, [6], has been stressed in the fundamental paper [9] whose experimental results and theoretical hints led to the formulation of the (CH) as a reinterpretation of Ruelle's principle, [5].

[5] Ruelle, D.: *Positivity of entropy production in non equilibrium statistical mechanics*, IHES preprint, 1995.

[6] Sinai, Y.G.: Russian Mathematical Surveys, **27**, 21–69, 1972. Also: *Lectures in ergodic theory*, Lecture notes in Mathematics, Princeton, 1977.

[7] Gallavotti, G.: Math. Phys. Electronic J., MPEJ, **1**, 1–12, 1995, (http:// mpej. unige. ch). Strictly speaking this paper deals with discrete time systems; the corresponding continuous time case is discussed in [8].

[8] Gentile, G.: *Large deviation rule for Anosov flows*, IHES preprint, 1996.

[9] Evans, D.J.,Cohen, E.G.D., Morriss, G.P.: Phys. Rev. Lett., **71**, 2401–2404, 1993.

[10] Ruelle, D.: *Chaotic motions and strange attractors*, Lezioni Lincee, notes by S. Isola, Cambridge U. Press, 1989; and: Ruelle, D.: *Measures describing a turbulent flow*, Ann. of the N. Y. Acad. of Sci., **357**, 1–9, 1980.

[11] Bonetto, F., Gallavotti, G., Garrido, P.: *Chaotic principle: an experimental test*, in preparation.