Journal of Statistical Physics, Vol. 85, Nos. 1/2, 1996

Positivity of Entropy Production in Nonequilibrium Statistical Mechanics

David Ruelle[1]

Received December 22, 1995

We analyze different mechanisms of entropy production in statistical mechanics, and propose formulas for the entropy production rate $e(\mu)$ in a state μ. When μ is a steady state describing the long term behavior of a system we show that $e(\mu) \geqslant 0$, and sometimes we can prove $e(\mu) > 0$.

KEY WORDS: Ensemble; entropy production; folding entropy; nonequilibrium stationary state; nonequilibrium statistical mechanics; SRB state; thermostat.

INTRODUCTION

The study of nonequilibrium statistical mechanics leads naturally to the introduction of *nonequilibrium states.* These are probability measures μ on the phase space of the system, suitably chosen and stationary (in principle) under the nonequilibrium time evolution. In the present paper we analyze the entropy production $e(\mu)$ for such nonequilibrium states, and show that it is positive (more precisely $\geqslant 0$, sometimes one can prove >0). That the positivity of $e(\mu)$ needs a proof was repeatedly pointed out by G. Gallavotti and E. G. D. Cohen.[2] Here we shall emphasize the physics of the problem and be particularly concerned with a proper choice of mathematical framework and definitions; the proof that $e(\mu) \geqslant 0$ will then be relatively easy.

---

[1] IHES, 91440 Bures sur Yvette, France, and Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903.

[2] In their seminal paper,[15] for instance, they state, "positivity rests on numerical evidence," and refer to ref. 12.

822/85/1-2-1

Thermostatting

We shall think of a physical system having a finite (possibly large) number of degrees of freedom. The phase space $\mathscr{S}$ is thus a finite-dimensional manifold, with a symplectic structure and therefore a natural volume element. In the situation of *equilibrium statistical mechanics* there are conservative forces acting on the system. A Hamiltonian H is thus defined on $\mathscr{S}$, and energy is conserved, i.e., time evolution is restricted to an energy shell $\mathscr{S}_E = \{x: H(x) = E\}$ of $\mathscr{S}$, where E typically ranges from some lower bound E_0 to $+\infty$ (this is because the potential energy is $\geqslant E_0$ and the kinetic energy takes all values $\geqslant 0$). While $\mathscr{S}$ is noncompact and has infinite volume, $\mathscr{S}_E$ is compact and has finite volume.

In the case of *nonequilibrium statistical mechanics* we have nonconservative forces and, although we may be able to define a natural energy function, with values in $[E_0, +\infty)$, the energy is in general not conserved. Typically the point representing the system wanders away to infinity in the noncompact phase space $\mathscr{S}$ while the system heats up, i.e., its energy tends to $+\infty$. In this situation it is not possible to define time averages corresponding to a probability measure on $\mathscr{S}$, i.e., it is not possible to introduce nonequilibrium states. This difficulty follows from the noncompactness of phase space and is not tied to the special physical meaning of the energy. (The same difficulty arises in diffusion problems where the energy is constant but the configuration space is infinite).

Physically, the way to avoid heating up the system is to put it in contact with a thermostat. One can idealize the thermostat as a random interaction (with a heat bath). The study of entropy production remains to be done in this framework, and should separate the randomness (or entropy) introduced by the thermostat and that created by the system itself.

It is also possible to constrain the time evolution by brute force to some compact manifold $M \subset S$. Consider, for instance, a system satisfying Hamiltonian equations of motion:

$$\dot{X} = J\,\partial_X H$$

where $X = (p, q)$ and $J\,\partial_X = (-\partial_q, \partial_p)$. The energy is conserved because $\dot{H} = \partial_X H \cdot \dot{X} = \partial_X H \cdot J\,\partial_X H = 0$. Let us add an external driving term F so that the time evolution is now

$$\dot{X} = J\,\partial_X H + F(X)$$

This will in general heat up the system because

$$\dot{H} = \partial_X H \cdot \dot{X} = \partial_X H \cdot F(X) \neq 0$$

but if we replace F by

$$F - \frac{(\partial H \cdot F)}{(\partial H \cdot \partial H)} \cdot \partial H$$

then H is preserved: this is the so-called *Gaussian thermostat.*[18]

To summarize, we want to act on our system to keep it outside of equilibrium, but also impose a thermostat to prevent heating. Physically this means that we pump entropy out of the system, while keeping the energy fixed.

From now on we shall consider a time evolution on a compact manifold M. We shall forget the symplectic structure (this is no longer relevant because we no longer have a Hamiltonian). We shall, however, need the volume element dx to define the statistical mechanical entropy $S(\rho) = -\int dx\, \rho(x) \log \rho(x)$ of a probability density ρ on M. Equivalent volume elements will be equivalent for our purposes because changing dx to $\phi(x)\, dx$ replaces $S(\rho)$ by $S(\rho) + \int dx\, \rho(x) \log \phi(x)$; the additive term is bounded independently of ρ, and will play no significant role in our considerations. We may thus take for dx the volume element associated with any Riemann metric. Note that $S(\rho)$ is the physical entropy when ρ is a thermodynamic equilibrium state, but we can extend the definition to arbitrary ρ such that $S(\rho)$ is finite.

The fact that we take seriously the expression $S(\rho) = -\int dx\, \rho(x) \log \rho(x)$ for the entropy seems to be at variance with the point of view defended by Lebowitz,[20] who prefers to give physical meaning to a *Boltzmann entropy* different from $S(\rho)$. There is, however, no necessary contradiction between the two points of view, which correspond to idealizations of different physical situations. Specifically, Lebowitz discusses the entropy of states which are locally close to equilibrium, while here we analyze entropy production for certain particular steady states (which may be far from equilibrium).

Pumping Entropy Out of the System

We have now reduced our mathematical framework to a smooth time evolution on a compact manifold M. We may also discretize the time (using a time one map f or a Poincaré first return map f) and consider that the time evolution is given by iterates of $f: M \to M$. Even though the mathematical setup is now just that of a smooth dynamical system (M, f), there remains the problem to study how entropy is pumped out of the system, and how nonequilibrium states are defined. We shall consider three cases.

(i) f is a diffeomorphism (hence f^{-1} is defined). Nonequilibrium states μ may be defined by time averages corresponding to orbits $(f^k x)$, where $x \in \mathscr{V}(\mu)$ and $\mathscr{V}$ has positive Riemann volume: vol $\mathscr{V}(\mu) > 0$. More precisely, let $\delta(x)$ denote the unit mass at x; we may say that μ is a nonequilibrium state if $\mu = \lim_{n \to \infty} (1/m) \sum_{k=0}^{m-1} \delta(f^k x)$ for all $x \in \mathscr{V}(\mu)$, and $\mathscr{V}(\mu) > 0$; special examples are the so-called SRB states. We shall see that entropy is pumped out of μ because f contracts volume elements (in the average).[3]

(ii) f is an noninvertible map. Here the folding of the phase space caused by f acts to pump entropy out of the system.[4] Nonequilibrium states may be defined as limits of states $(1/m) \sum_{k=0}^{m-1} f^k \rho$ with ρ absolutely continuous with respect to the volume.

(iii) f has a nonattracting set A which carries a nonequilibrium state μ associated with a diffusion process.[17, 10] Specifically, let A be an f-invariant subset of M which is not attracting. If U is a small neighborhood of A, $fU$is not contained in U. Let ρ be the Riemann volume normalized to U. Then $f\rho$ is not supported in U. We multiply by the characteristic function of U and normalize to obtain a new probability measure $\rho_1 = \|\chi_U \cdot f\rho\|^{-1} \chi_U \cdot f\rho$. Iterating this process m times, we obtain ρ_m, and define

$$\rho^{(m)} = \frac{1}{m} \sum_{k=1}^{m} f^{-k} \rho_m$$

In the Axiom A case we shall see (Section 3 below) that $\rho^{(m)}$ tends to an f-invariant probability measure μ giving to the quantity $[h(\mu - \sum$ positive Lyapunov exponents of $\mu]$ its maximum value P (h is the *Kolmogorov–Sinai entropy* and the *pressure* P is $\geqslant 0$). One can argue (see, refs. 19 and 11, and below) that the volume of the points $x \in U$ such that $fx,..., f^m x \in U$ behaves like e^{mP}. Here again entropy is pumped out of the system by getting rid of the part of $f\rho$ outside of U, and μ may be interpreted as nonequilibrium state.

For a recent physically oriented review of nonequilibrium statistical mechanics we refer the reader to Dorfman.[9] He discusses in particular calculations using periodic orbits, as advocated by Cvitanović *et al.*.[8, 1]

Toward Physical Applications

The *ergodic hypothesis* states that the Liouville measure restricted to an energy shell $\mathscr{S}_E$ (for a Hamiltonian system) is ergodic under time evolution. This serves to justify the ensembles of statistical mechanics and, while the ergodic hypothesis is likely to be false in general, it is apparently almost

[3] See refs. 5 and 6 for models with phase space contraction.

[4] A model with folding of phase space has been considered by Chernov and Lebowitz.[7]

true in the sense that the application of equilibrium statistical mechanics to real systems has been extremely successful.

One may try to base nonequilibrium statistical mechanics on a principle similar to the ergodic hypothesis. Here one assumes that the dynamical system (M, f) describing time evolution is *hyperbolic* in some sense[5] and that time averages are given by particular probability measures called SRB measures; these are the nonequilibrium states which replace the microcanonical *ensemble* of equilibrium statistical mechanics. The SRB states correspond to time averages for a set of positive measure of initial conditions. They are characterized by smoothness along unstable directions or equivalently by a variational principle.[6]

The assumption that the systems of nonequilibrium statistical mechanics are hyperbolic and described by SRB measures is unlikely to be exactly true, but it is reasonable to expect that it is approximately true in the sense that it gives correct physical predictions in the limit of large systems (thermodynamic limit).

Actual physical predictions were obtained only after Gallavotti and Cohen[15] supplemented the hyperbolicity assumption by the *reversibility* assumption. The latter assumes that there is a map $i: M \mapsto M$ such that $i^2 = 1, fi = if^{-1}$.

The *chaotic hypothesis* of Gallavotti and Cohen[15, 16] (see also Gallavotti[13, 14]) states that physically correct results (for nonequilibrium systems in the thermodynamic limit) will be obtained by assuming reversibility and treating the system as if it were hyperbolic (in fact Anosov). An essential role in the inspiration of Gallavotti and Cohen was played by the numerical results and analysis by Evans *et al.*[12]

Example

Consider a Hamiltonian $H(X) = \frac{1}{2}(p, M^{-1}p) + U(q)$ where M is the mass matrix and U is the potential energy. We denote by $f^t X$ (with $f^0 X = X$) the solution of Hamilton's equation $\dot{X} = J\,\partial_X H$. Defining

[5] See Eckmann and Ruelle[11] for definitions and a physically oriented review of dynamical systems.

[6] The approach just indicated to the study of nonequilibrium systems was advocated early in lectures by the present author (G. Gallavotti mentions the date of 1973); for the case of turbulence se ref. 24. Other people familiar with SRB measures would have had similar ideas, but these have started to be useful only with the recent (1995) word of Gallavotti and Cohen.[15, 16] The mathematical study of SRB states was made by Sinai[26] for Asonov diffeomorphisms, Ruelle[23] for Axiom A diffeomorphisms, and Bowen and Ruelle[4] for Axiom A flows. The very nontrivial extension to nonuniformly hyperbolic systems is due to Ledrappier and Young.[22]

$i(p, q) = (-p, q)$, we find that $f'i = if^{-\prime}$, which expresses reversibility. Reversibility is preserved if we introduce an external force $F = (\Phi(q), 0)$, and again if we add a Gaussian thermostat.

Scope of the Paper

In what follows we shall analyze entropy production and its positivity for the three cases outlined earlier: (i) diffeomorphism, (ii) noninvertible map, (iii) map near a nonattracting set. The treatment of these three cases will be somewhat uneven because the existing mathematical results range from detailed in case (i) to rather limited in case (iii). Since the emphasis of this paper is on having the physics straight, we allow the uneven mathematical treatment, but suggested some conjectural extensions of the results that are proved. The possibility of a unified presentation will depend on further progress in the ergodic theory of differentiable dynamical systems.

1. ENTROPY PRODUCTION FOR DIFFEOMORPHISMS

Let M be a compact manifold and $f: M \mapsto M$ a C^1 diffeomorphism. Choosing a Riemann metric on M, let $\rho(dx) = \underline{\rho}(x)\,dx$ be a probability measure with density $\underline{\rho}$ with respect to the Riemann volume element dx. The direct image $\rho_1 = f\rho$ has density $\underline{\rho}_1(x) = \underline{\rho}(f^{-1}x)/J(f^{-1}x)$, where $J(X)$ is the absolute value of the Jacobian of f at x (computed with respect to the Riemann metric). The statistical mechanical entropy associated with $\underline{\rho}$ is

$$S(\underline{\rho}) = -\int dx\, \underline{\rho}(x) \log \underline{\rho}(x)$$

[This means that dx is interpreted as the phase space volume element; if dx is the configuration space volume element, then $S(\underline{\rho})$ is the configurational entropy.] The entropy $S(\underline{\rho})$ will have to be distinguished from the Kolmogorov–Sinai (time) entropy $h(\mu)$ of an f-invariant state μ used below. The entropy associated with $\underline{\rho}_1$ is

$$\begin{aligned} S(\underline{\rho}_1) &= -\int dx\, \underline{\rho}_1(y) \log \underline{\rho}_1(y) \\ &= -\int dy \frac{\underline{\rho}(f^{-1}y)}{J(f^{-1}y)} [\log \underline{\rho}(f^{-1}y) - \log J(f^{-1}y)] \\ &= -\int dx\, \underline{\rho}(x)[\log \underline{\rho}(x) - \log J(x)] \end{aligned}$$

The entropy put into the systems in one time step is thus

$$S(\underline{\rho}_1) - S(\underline{\rho}) = \int dx\, \underline{\rho}(x) \log J(x)$$

This means that the entropy pumped out of the system, or produced by the system, is

$$-\int dx\, \underline{\rho}(x) \log J(x)$$

Let $\underline{\rho}_m$ be the density of the measure $\rho_m = f^m \rho$. If $\underline{\rho}_m$ tends vaguely[7] to μ when $m \to \infty$, the entropy production

$$-[S(\underline{\rho}_{m+1}) - S(\underline{\rho}_m)] = -\int dx\, \rho_m(x) \log J(x)$$

tends to

$$-\int \mu(dx) \log J(x)$$

It is thus natural to take as definition of the entropy production for an arbitrary f-invariant probability measure μ the expression

$$e_f(\mu) = -\int \mu(dx) \log J(x)$$

In the rest of this Section we take μ to be ergodic, so that the Lyapunov exponents are constant (μ-a.e.). The general case is obtained by representing μ as an integral over its ergodic components.

Lemma 1.1. The entropy production $e_f(\mu)$ is independent of the choice of Riemann metric and equal to minus the sum of the Lyapunov exponents of μ with respect to f.

Proof. This follows from the Oseledec multiplicative ergodic theorem in the form given in ref. 11. ▮

[7] The vague topology is the w^*-topology on the space of measures considered as dual of the space of continuous functions. We denote a vague limit by v.lim.

We remind the reader that the Kolmogorov–Sinai entropy $h(\mu)$ is the amount of information produced by f in the state μ (see, for instance, Billingsley.[2] We always have

$$h(\mu) \leqslant \sum \text{positive Lyapunov exponents} \tag{1.1}$$

(this inequality is due to Ruelle; see ref. 11). We call μ an SRB measure[22, 11] if

$$h(\mu) = \sum \text{positive Lyapunov exponents} \tag{1.2}$$

(*Pesin identity*). If f is of class C^2, the above condition is equivalent to μ having conditional probabilities on unstable manifolds absolutely continuous with respect to Lebesgue measure.[22] If f is C^2 and μ has no vanishing Lyapunov exponent, then there is a set of positive Riemann volume of points $x \in M$ with time averages $(1/N) \sum_{k=0}^{N-1} \delta(f^k x)$ tending vaguely to μ (this result is due to Pugh and Shub; see ref. 11).

Theorem 1.2. Let f be a C^1 diffeomorphism and μ an f-invariant probability measure on the compact manifold M.

(a) If μ is an SRB measure then $e_f(\mu) \geqslant 0$.

(b) Let f be $C^{1+\alpha}$ with $\alpha > 0$ and μ be an SRB measure. If μ is singular with respect to dx and has no vanishing Lyapunov exponent, then $e_f(\mu) > 0$.

(c) For every a

$$\text{vol} \left\{ x: \frac{1}{m} \sum_{k=0}^{m-1} \log J(f^k x) \geqslant a \right\} \leqslant e^{-ma} \, \text{vol}\, M$$

In particular, if $\mathcal{V}(\mu) = \{x: \text{v.lim}_{m \to \infty} (1/m) \sum_{k=0}^{m-1} \delta(f^k x) = \mu\}$ and $e_f(\mu) < 0$, then vol $\mathcal{V}(\mu) = 0$.

Proof. We have denoted by vol the Riemann volume in M. In view of the result of Pugh and Shub mentioned above, (a) follows from (c) if f is C^2 and μ has no zero characteristic exponent. Here is a direct proof of (a): if μ is SRB, we have

$$\begin{aligned} e_f(\mu) &= -\sum \text{Lyapunov exponents} \\ &= \left[h(\mu) - \sum \text{positive Lyapunov exponents} \right] \\ &\quad - \left[h(\mu) + \sum \text{negative Lyapunov exponents} \right] \end{aligned}$$

$$
\begin{aligned}
&= \left[h(\mu) - \sum \text{positive Lyapunov exponents w.r.t. } f \right] \\
&\quad - \left[h(\mu) - \sum \text{positive Lyapunov exponents w.r.t. } f^{-1} \right] \\
&\geqslant 0
\end{aligned}
$$

where we have used (1.1) and (1.2).

To prove (b), notice that if μ is SRB and $e_f(\mu) = 0$, then, according to (a),

$$
\begin{aligned}
h(\mu) &= \sum \text{positive Lyapunov exponents} \\
&= -\sum \text{negative Lyapunov exponents}
\end{aligned}
$$

This implies that μ is absolutely continuous with respect to dx [see Ledrappier[21] Corollary (5.6)] if f is a class $C^{1+\alpha}$ and μ has no vanishing Lyapunov exponent.

To prove (c), write

$$
\mathcal{Y}^{+}(m) = \left\{ x : \frac{1}{m} \sum_{k=0}^{m-1} \log J(f^k x) \geqslant a \right\}
$$

We have thus

$$
\begin{aligned}
\operatorname{vol} M \geqslant \operatorname{vol} f^m \mathcal{Y}^{+}(m) &= \int_{\mathcal{Y}^{+}(m)} \prod_{k=0}^{m-1} J(f^k x)\, dx \\
&\geqslant e^{ma} \operatorname{vol} \mathcal{Y}^{+}(m)
\end{aligned}
$$

as announded. ▮

Corollary 1.3. If μ is an SRB measure with respect to both f and f^{-1}, then $e_f(\mu) = 0$.

Proof. We have indeed $e_f(\mu) \geqslant 0$, and $e_{f^{-1}}(\mu) = -e_f(\mu) \geqslant 0$. (As pointed out to the author by Joel Lebowitz, this covers the case of the microcanonical ensemble). ▮

2. ENTROPY PRODUCTION FOR NONINVERTIBLE MAPS

2.1. Standing Assumptions

Let M be a compact Riemann manifold, possibly with boundary. We denote by vol the Riemann volume and by dx the volume element. We assume that a closed set $\Sigma \subset M$ is given, containing the boundary of M, and $f: M\backslash\Sigma \to M$ such that the following properties are satisfied:

(A1) $\operatorname{vol}\Sigma = 0$.

(A2) There are disjoint open sets $D_1,..., D_N$ such that $M\backslash\Sigma = \bigcup_{\alpha=1}^N D_\alpha$, and $f|D_\alpha$ is a homeomorphism to fD_α, absolutely continuous with respect to vol. The Jacobian J of f is continuous in $M\backslash\Sigma$ and satisfies

$$\inf_{x\notin\Sigma} J(x) \geqslant e^{-K} > 0$$

(A3) For all pairs (α, β), fD_α and fD_β are either disjoint or identical.

2.2. Comments

It is convenient to use a map f defined outside of an *excluded* set Σ. In particular this allows discontinuities on Σ. When considering the direct image $f\mu$ of a measure $\mu \geqslant 0$ on M by f, we shall have to assume that $\mu(\Sigma) = 0$. (We have made such an assumption for the measure vol.)

Condition (A3) might seem very strong, but can be arranged to hold under the weaker assumption

$$\operatorname{vol}(fD_\alpha \cap \partial f D_\beta) = 0$$

for all pairs (α, β). Let indeed (D_γ^1) be the family of open sets $\bigcap_{k=1}^N (fD_\alpha)^\sim$, where $(fD_\alpha)^\sim$ is either fD_α or $M\backslash\operatorname{clos} fD_\alpha$ for each α. Let $D_{\alpha,\gamma}^* = D_\alpha \cap f^{-1}D_\gamma^1$ and $\Sigma^* = M\backslash\bigcup_\alpha\bigcup_\gamma D_{\alpha,\gamma}^*$, then (A1)–(A3) hold when Σ, (D_α) are replaced by Σ^*, $(D_{\alpha\gamma}^*)$. When considering the direct image $f\mu$, we shall now have to assume that $\mu(\Sigma^*) = 0$.

2.3. Refining (D_α)

Let $fD_\alpha = D_\gamma^1$. We may write

$$D_\gamma^1 = \Sigma_\gamma^1 \cup D_{\gamma 1}^1 \cup \cdots \cup D_{\gamma n}^1$$

where $\operatorname{vol}\Sigma_\gamma^1 = 0$ and the disjoint open sets $D_{\gamma 1}^1,..., D_{\gamma n}^1$ are small. Writing $D_{\alpha i} = D_\alpha \cap f^{-1}D_{\gamma i}^1$, we may replace (D_α) by a family $(D_{\alpha i})$ of arbitrarily small sets. In other words we may refine the family (D_α) to a new family (D_α^*) (with $\alpha \in \{1,..., N^*\}$ and an excluded set Σ^*) so that (A1)–(A3) still hold and the sets D_α^* are arbitrarily small.

In the study of a measure $\mu \geqslant$ with $\mu(\Sigma) = 0$ we can arrange that $f(\mu)(\Sigma_\gamma^1) = 0$, implying that $\mu(\Sigma^*) = 0$.

2.4. Folding Entropy

Let μ be a positive measure on $M\backslash\Sigma$. [We may also consider μ as a positive measure on M such that $\mu(\Sigma)=0$.] Our assumptions imply that there is a *disintegration* of μ associated with the map f (see Bourbaki[3] paragraph 3). In general this means that we have the integral representation

$$\mu = \int \mu_1(dx)\, \nu_x$$

where $\mu_1 = f\mu$ is the direct image of μ by f, and ν_x is a probability measure with $\nu_x(f^{-1}\{x\})=1$. This representation is essentially unique. Here we may assume that μ_x is atomic (with at most N atoms) and write

$$H(\nu_x) = -\sum_\alpha p_\alpha \log p_\alpha$$

where the p_α are the masses of the atoms of ν_x. [In the general case we would write $H(\nu_x)=+\infty$ if ν_x is nonatomic.] We let now

$$F(\mu) = F_f(\mu) = \int \mu_1(dx)\, H(\nu_x)$$

and call $F(\mu)$ the *folding entropy* of μ with respect to f.

Let again $D_\gamma^1 = fD_\alpha$. By the concavity of $t \mapsto -t\log t$, we have

$$(\mu_1(D_\gamma^1))^{-1} \int_{D_\gamma^1} \mu_1(dx)\, H(\nu_x) \leqslant -\sum_{\alpha:\, \gamma(\alpha)=\gamma} \frac{\mu(D_\alpha)}{\mu_1(D_\gamma^1)} \log \frac{\mu(D_\alpha)}{\mu_1(D_\gamma^1)}$$

Therefore, when (D_α) is replaced by (D_α^*), which consists of smaller and smaller sets, the expression

$$F^*(\mu) = \sum_\gamma \mu_1(D_\gamma^{*1}) \left[-\sum_{\alpha:\, \gamma(\alpha)=\gamma} \frac{\mu(D_\alpha^*)}{\mu_1(D_\gamma^{*1})} \log \frac{\mu(D_\alpha^*)}{\mu_1(D_\gamma^{*1})} \right]$$

tends to $F(\mu) = \int \mu_1(dx)\, H(\nu_x)$ from above.

Proposition 2.1. Let P be the set of probability measures on M with the vague topology and

$$I = \{\mu \in P\colon \mu \text{ is } f\text{-invariant}\}$$

$$P_{\backslash\Sigma} = \{\mu \in P\colon \mu(\Sigma)=0\}$$

$$I_{\backslash\Sigma} = I \cap P_{\backslash\Sigma}$$

(a) The function $F\colon P_{\backslash\Sigma} \mapsto \mathbf{R}$ (with values in $[0, \log N]$) is concave upper semicontinuous (u.s.c.).

(b) The restriction of F to $I_{\backslash\Sigma}$ is affine u.s.c.

Proof. Since $H(\nu_x)$ takes values in $[0, \log N]$, so does F. To prove concavity, we have to estimate F at μ', μ'', and $\mu = (1-t)\,\mu' + t\mu''$, with $\mu', \mu'' \in P_{\backslash\Sigma}$. We may choose (D_α^*) arbitrarily fine so that $\mu'(\Sigma^*) = \mu''(\Sigma^*) = 0$; therefore $F(\mu) = \lim F^*(\mu)$, $F(\mu') = \lim F^*(\mu')$, $F(\mu'') = \lim F^*(\mu'')$. Concavity of F follows from the concavity of $t \mapsto F^*((1-t)\,\mu' + t\mu'')$, or the convexity of

$$t \mapsto \sum_\alpha \left[(1-t)\,u_\alpha + tv_\alpha\right] \log \frac{(1-t)\,u_\alpha + tv_\alpha}{\sum_\beta((1-t)\,u_\beta + tv_\beta)}$$

Since P is metrizable (with the vague topology), we prove upper semicontinuity of F by showing that if $\rho^{(m)}$, $\mu \in P_{\backslash\Sigma}$ and if the sequence $(\rho^{(m)})$ tends to μ, then $F(\mu) \leqslant \lim F(\rho^{(m)})$. We may choose (D_α^*) arbitrarily fine so that $\mu(\Sigma^*) = 0$ and $\rho^{(m)}(\Sigma^*) = 0$ for all m; $F(\mu)$ and $F(\rho^{(m)})$ are thus limits of $F^*(\mu)$ and $F^*(\rho^{(m)})$. Since $\mu(\Sigma^*) = \rho^{(m)}(\Sigma^*) = 0$, F^* is continuous for the vague topology on the set $S = \{\mu\} \cup \{\rho^{(m)}\colon m \in \mathbf{N}\}$, and $F \mid S$ is thus the limit of a decreasing family of continuous functions, hence upper semicontinuous.

This proves (a).

To prove (b) we remark that μ', μ'' are absolutely continuous with respect to $\mu = (1-t)\,\mu' + t\mu''$ (if $t \neq 0, 1$) and let $g' = \delta\mu'/\delta\mu$, $g'' = \delta\mu''/\delta\mu$. If $\mu', \mu'' \in I$, the functions g', g'' are f-invariant. Therefore

$$\mu'(dy) = (g'\mu)(dy) = g'(y)\mu(dy) = \int \mu_1(dx)\; g'(y)\; \nu_x(dy)$$

$$= \int \mu_1(dx)\; g'(x)\; \nu_x(dy) = \int (g'\mu)_1\,(dx)\; \nu_x(dy)$$

$$= \int \mu'_1(dx)\; \nu_x(dy)$$

and similarly for $g''\mu$. Therefore

$$(1-t)\,F(\mu') + tF(\mu'') = (1-t) \int \mu'_1(dx)\; H(\nu_x) + t \int \mu''_1(dx)\; H(\nu_x)$$

$$= \int \mu_1(dx)\; H(\nu_x) = F(\mu)$$

This completes the proof of the proposition. ∎

2.5. Extension

If P denotes the set of positive measures on M (rather than the probability measures), we have

$$F(\mu) \in [0, \log N] \cdot \|\mu\|$$

Apart from that, the above proposition remains true, with the same proof. In fact, since $F(\lambda\mu) = \lambda F(\mu)$ for $\lambda \geqslant 0$, $\mu \geqslant 0$, $\mu(\Sigma) = 0$, the extension of F from probability measures to positive measures is trivial.

2.6. Entropy Production

We define now the *entropy production* $e_f(\mu)$ for a dynamical system (M, f) satisfying our standing assumptions and $\mu \in P_{\backslash\Sigma}$ [i.e., μ is a probability measure such that $\mu(\Sigma) = 0$]. We write

$$e_f(\mu) = F(\mu) - \mu(\log J)$$

This definition will be motivated below, first when μ is defined by a density, then more generally.

Proposition 2.2. (a) $e_f(\mu)$ is independent of the choice of Riemann metric on M.

(b) e_f is concave u.s.c. on $P_{\backslash\Sigma}$, and affine u.s.c. on $I_{\backslash\Sigma}$.

(c) If the probability measures $\rho^{(m)}$ are absolutely continuous with respect to Riemann volume and tend vaguely to $\mu \in P_{\backslash\Sigma}$, we have

$$\limsup_{m \to \infty} e_f(\rho^{(m)}) \leqslant e_f(\mu)$$

Proof. A change of Riemann metric replaces J by $J + \Phi - \Phi \circ f$, so that $\mu(\log J)$ and $e_f(\mu)$ are not changed. This proves (a).

The function $K + \log J$ is $\geqslant 0$ and continuous on $M \backslash \Sigma$. Let (χ_n) be an increasing sequence of continuous functions $M \mapsto [0, 1]$, vanishing on Σ and tending to 1 on $M \backslash \Sigma$. Then $((K + \log J) \cdot \chi_n)$ is an increasing sequence of continuous positive functions tending to $K + \log J$ on on $M \backslash \Sigma$. Therefore

$$\mu \mapsto \mu(K + \log J) = K + \mu(\log J)$$

is affine l.s.c. on $P_{\backslash\Sigma}$, and

$$\mu \mapsto -\mu(\log J)$$

is affine u.s.c. on $P_{\backslash\Sigma}$. Together with Proposition 2.1, this proves (b).

To prove (c) we note that, since $\operatorname{vol}\Sigma = 0$, we have $\rho^{(m)}(\Sigma) = 0$. It suffices then to apply (b). ∎

2.7. Entropy Associated with a Density

Let ρ be a probability measure with density $\underline{\rho}$ with respect to Riemann volume, i.e., $\rho(dx) = \underline{\rho}(x)\,dx$. If dx is interpreted as phase space volume element, the statistical mechanical entropy associated with ρ is

$$S(\underline{\rho}) = -\int dx\, \underline{\rho}(x) \log \underline{\rho}(x)$$

Using the concavity of the log, we have

$$S(\underline{\rho}) = \int \underline{\rho}(x) \log \frac{1}{\underline{\rho}(x)} \leqslant \log \int dx \frac{\underline{\rho}(x)}{\underline{\rho}(x)} = \log \operatorname{vol} M \tag{2.1}$$

so that $S(\cdot)$ takes values in $[-\infty, \log \operatorname{vol} M]$, the value $-\infty$ being allowed.

If ψ_α is the inverse of $f|D_\alpha$, the direct image $\rho_1 = f\rho$ has density

$$\underline{\rho}_1 = \sum_\alpha (\underline{\rho} \circ \psi_\alpha) \cdot (\bar{J} \circ \psi_\alpha)$$

where $\bar{J} = 1/J$, and characteristic functions of the sets fD_α have been omitted. Define

$$p_\alpha = \frac{1}{\underline{\rho}_1(x)} \underline{\rho}(\psi_\alpha x) \cdot (\bar{J} \circ \psi_\alpha)$$

$$\nu_x = \sum_\alpha p_\alpha(x)\, \delta(\psi_\alpha x)$$

where $\delta(x)$ denotes the unit mass at x. Note that $f\nu_x = \delta(x)$. We have the disintegration

$$\rho = \int dx\, \underline{\rho}_1(x)\, \nu_x \tag{2.2}$$

and therefore

$$F(\rho) = \int dx\, \underline{\rho}_1(x)\, H(v_x) \tag{2.3}$$

Note also the identity[8]

$$\begin{aligned} \log \underline{\rho}_1(x) &= -\sum_\alpha p_\alpha(x) \log p_\alpha(x) \\ &\quad + \sum_\alpha p_\alpha(x)[\log \underline{\rho}(\psi_\alpha x) + \log \bar{J}(\psi_\alpha x)] \\ &= H(v_x) + v_x(\log \underline{\rho}) + v_x(\log \bar{J}) \end{aligned}$$

Therefore, using (2.3) and (2.2), we have

$$\begin{aligned} -S(\underline{\rho}_1) &= \int dx\, \underline{\rho}_1(x) \log \underline{\rho}_1(x) \\ &= \int dx\, \underline{\rho}_1(x)\, H(v_x) + \int dx\, \underline{\rho}_1(x)\, v_x(\log \underline{\rho}) \\ &\quad + \int dx\, \underline{\rho}_1(x)\, v_x(\log \bar{J}) \\ &= F(\rho) + \rho(\log \underline{\rho}) + \rho(\log \bar{J}) \end{aligned}$$

and, if $S(\rho) = -\rho(\log \underline{\rho})$ is $\neq -\infty$,

$$-[S(\underline{\rho}_1) - S(\underline{\rho})] = F(\rho) + \rho(\log \bar{J}) \tag{2.4}$$

The right-hand side has values $\leqslant \log N + K$, so that $S(\underline{\rho}_1) \neq -\infty$ when $S(\underline{\rho}) \neq -\infty$.

Proposition 2.3. Let $S(\underline{\rho}) \neq -\infty$.

(a) The entropy production associated with the density $\underline{\rho}$ is

$$-[S(\underline{\rho}_1) - S(\underline{\rho})] = F(\rho) + \rho(\log \bar{J}) = e_f(\rho)$$

[8] We are applying the formula (familiar in equilibrium statistical mechanics)

$$\log \sum_i e^{-U(i)} = -\sum_i p_i \log p_i - \sum_i p_i U(i)$$

with $p_i = e^{-U(i)}/\sum_j e^{-U(j)}$.

(b) If the probability measures $\rho^{(m)}$ are absolutely continuous with respect to Riemann volume and tend vaguely to μ such that $\mu(\Sigma)=0$, we have

$$e_f(\mu) \geqslant \limsup_{m\to\infty}[-S(\underline{\rho}_1^{(m)})+S(\underline{\rho}^{(m)})]$$

Proof. Part (a) follows from (2.4); (b) follows from (a) and Proposition 2.2(c). ∎

2.8. Physical Discussion

The above proposition is our justification to define $e_f(\mu)$ as the entropy production associated with $\mu \in P_{\backslash\Sigma}$. Note that the definition of $e_f(\mu)$ depends only on μ and f and not on the choice of an approximation of μ by absolutely continuous measures $\mu^{(m)}$. However, we only have the inequality

$$e_f(\mu)=F(\mu)+\mu(\log \bar{J}) \geqslant \limsup_{m\to\infty}[F(\rho^{(m)})+\rho^{(m)}(\log \bar{J})]$$

where one might hope for an inequality. The term $\mu(\log \bar{J})$ poses no serious problem in this respect: if we assume that $\log \bar{J}$ is bounded, we have

$$\mu(\log \bar{J})=\lim_{m\to\infty} \rho^{(m)}(\log \bar{J})$$

For the term $F(\mu)$ there might, however, be a discontinuity of F at μ. What this means is that some mass of $\rho^{(m)}$ gets folded more in the limit $\rho^{(m)}\to\mu$. For instance, f might be injective on supp $\rho^{(m)}$ but not on suppp μ; this would give $F(\rho^{(m)})=0$, but possibly $F(\mu)>0$.

Physically one should only think of μ as an idealization of $\rho^{(m)}$ for large m. When the map f "folds together" some mass of μ, it almost folds together the corresponding mass of $\rho^{(m)}$ and, in a coarse-grained description, it thus makes sense to replace $F(\rho^{(m)})$ by $F(\mu)$ and to interpret the latter as the physical folding entropy of our system.

Take $\rho^{(m)}=f^m\rho$ and suppose that $\rho^{(m)}\to\mu\in I_{\backslash\Sigma}$. In the step between time m and time $m+1$, the entropy production is

$$-S(\underline{\rho}^{(m+1)})+S(\underline{\rho}^{(m)})=F(\rho^{(m)})+\rho^{(m)}(\log \bar{J})$$

which we approximate by $F(\mu)+\mu(\log \bar{J})$. This seems to mean that $S(\underline{\mu})$ increases by a fixed amount at each time step, which is absurd since μ does not depend on time. In fact, typically, μ is singular, i.e., its density $\underline{\mu}$ does not exist, and we should write $S(\underline{\mu}) = -\infty$. We shall argue later that the entropy production is positive; the system produces this entropy by having its own entropy $S(\underline{\rho}^{(m)})$ decrease toward $-\infty$ when $m \to \infty$. The entropy produced is absorbed (or transfered to the outside world) by the time evolution f (i.e., by the forces which cause the time evolution).

Let $\rho \mapsto \rho * \theta$ denote the action of a stochastic diffusion operator θ close to the identity operator. Let us replace the time evolution $\rho \mapsto f\rho$ by the "noisy evolution" $\rho \mapsto (f\rho) * \theta$. We assume that this stochastic evolution has a steady state μ_θ tending to μ when $\theta \to$ identity. Here $S(\underline{\mu}_\theta)$ is finite and we can see that the entropy production is due to the diffusion $* \theta$. We may indeed write

$$-S(\underline{\mu}') + S(\underline{\mu}) = S(\underline{\mu}'') - S(\underline{\mu}')$$

where $\underline{\mu}, \underline{\mu}', \underline{\mu}''$ are the densities associated, respectively, with $\mu_\theta, f\mu_\theta$, and $(f\mu_\theta) * \theta = \mu_\theta$. The left hand side in the above formula is our familiar expression for the entropy production, and the right-hand side is the entropy produced by the diffusion. Let $\mu^{(m)}$ be obtained from ρ by the noisy evolution after m time steps. Because $\mu^{(m)}$ is smeared as compared with $\rho^{(m)} = f^m\rho$, we expect that the folding entropy $F(\mu^{(m)})$ will be close to $F(\mu_\theta)$ or $F(\mu)$. This is further justification for our choice of the definition $e_f(\mu)$ for the entropy production.

2.9. Positivity of Entropy Production

The following result, showing that $e_f(\mu) \geqslant 0$ for physically reasonable μ, is close to the results obtained when f is a diffeomorphism. The proof is remarkably simple.

Theorem 2.4. Let μ be a probability measure with density ρ on M. If $S(\rho)$ is finite and if μ is a vague limit of the measures $\rho^{(m)} = (1/m) \sum_{k=0}^{m-1} f^k\rho$ when $m \to \infty$, then $e_f(\mu) \geqslant 0$.

Proof. By Proposition 2.2(c) and 2.2(b), respectively we have

$$e_f(\mu) \geqslant \limsup_{m \to \infty} e_f(\rho^{(m)})$$

$$e_f(\rho^{(m)}) \geqslant \frac{1}{m} \sum_{k=0}^{m-1} e_f(f^k\rho)$$

Using Proposition 2.3(a), we also have

$$\sum_{k=0}^{m-1} e_f(f^k \rho) = -S(\underline{\rho}_m) + S(\underline{\rho})$$

Therefore

$$e_f(\mu) \geqslant \limsup \frac{1}{m} [-S(\underline{\rho}_m) + S(\underline{\rho})]$$

Since $-S(\underline{\rho}_m) \geqslant -\log \text{vol}\, M$ [by (2.1) above] we obtain $e_f(\mu) \geqslant 0$. ∎

2.10. Alternate Approach

Instead of our standing assumption, let us suppose that M is a compact manifold and $f\colon M \to M$ a C^1 map. One may then conjecture that

$$h(\mu) \leqslant F(\mu) + \left|\sum \text{negative Lyapunov exponents}\right|$$

when μ is an f-ergodic probability measure. (If our standing assumptions hold and f is piecewise C^1, with $\mu(\Sigma) = 0$, this can be proved along the lines of Ruelle.[25]) For an SRB state μ we have

$$h(\rho) = \sum \text{positive Lyapunov exponents}$$

and our conjecture implies

$$\begin{aligned} e_f(\mu) &= F(\mu) - \sum \text{positive Lyapunov exponents} \\ &\quad + \left|\sum \text{negative Lyapunov exponents}\right| \\ &\geqslant h(\mu) - \sum \text{positive Lyapunov exponents} = 0 \end{aligned}$$

3. ENTROPY PRODUCTION ASSOCIATED WITH DIFFUSION

Let M be a compact manifold, $f\colon M \to M$ a diffeomorphism, and A a compact f-invariant subset of M. Given a small open neighborhood U of A, we define

$$U_m = \{x : f^k x \in U \text{ for } k = 0, \ldots, m\}$$

Since we do not assume that the set A is attracting, mass will in general leak out of U. i.e., vol $U_m \to 0$ when $m \to \infty$. It is conjectured[19, 11, 17], that in many cases vol $U_m \approx e^{mP}$, and the escape rate from A under f is (up to change of sign)

$$P = P_{Af} = \sup_{\rho \in \partial I_A} \left\{ h(\rho) - \sum \text{positive Lyapunov exponents for } (\rho, f) \right\}$$
$$\leqslant 0$$

where ∂I_A is the set of f-ergodic probability measures with support in A.

If χ_m is the characteristic function of U_m, let $\rho_{[m]}$ and $\underline{\rho}^*_{[m]}$ be given by

$$\rho_{[m]}(dx) = \frac{\chi_m(x)}{\text{vol } U_m} dx$$

$$(f^m \rho_{[m]})(dx) = \underline{\rho}^*_{[m]}(x)\, dx$$

Then we may define the entropy production associated with escape from A as

$$e_A = \lim_{m \to \infty} \frac{1}{m} [S(\underline{\rho}_{[0]}) - S(\underline{\rho}^*_{[m]})] \tag{3.1}$$

if this limit exists.

The next proposition deals with the Axiom A case,[9] which is well understood mathematically. One may conjecture that results obtained in that case hold much more generally, but proofs are lacking at this time.

Proposition 3.1. Let A be a basic set for the C^2 Axiom A diffeomorphism f, and U_m, P, $\rho_{[m]}$, $\underline{\rho}^*_{[m]}$ be as above:

(a) $\lim_{m \to \infty}(1/m) \log \text{vol } U_m = P$.

(b) There is a unique f-ergodic probability measure μ on A such that

$$h(\mu) - \sum \text{positive Lyapunov exponents for } (\mu, f) = P$$

[9] Smale's foundational article[27] is still a convenient introduction to hyperbolic dynamical systems (with the definition of Axiom A diffeomorphisms, basic sets, etc.). For further references see ref. 11.

(c) Define

$$\rho^{(m)} = \frac{1}{m} \sum_{k=0}^{m-1} f^k \rho_{[m]}$$

Then v-lim $\rho^{(m)} = \mu$ when $m \to \infty$.

(d) The limit (3.1) defining e_A exists, and

$$e_A = \lim_{m \to \infty} \frac{1}{m} [S(\underline{\rho}_{[0]}) - S(\underline{\rho}^*_{[m]})] = -P_{Af} - \mu(\log J)$$

(where J is the absolute value of the Jacobian of f).

Proof. Part (a) can readily be extracted from Bowen and Ruelle,[4] where a slightly weaker result is proved (and flows are considered instead of diffeomorphisms).

If J^u denotes the Jacobian in the unstable direction, $\log J^u$ is Hölder continuous on A, and since $f|A$ is topologically transitive, there is a unique equilibrium state μ maximizing $h(\mu) - \mu(\log J^u)$.[23] This proves (b).

The *volume lemma* of ref. 4 establishes a close relation between ρ_m and μ. In fact it follows from ref. 4 that any vague limit of $\rho^{(m)}$ when $m \to \infty$ is absolutely continuous with respect to μ. Such a limit is also f-invariant and, since μ is ergodic, equal to μ. This proves (c).

Since $\underline{\rho}_{[m]}(x) = \chi_m(x)/\text{vol } U_m$, we have

$$S(\underline{\rho}_{[0]} - S(\underline{\rho}_{[m]}) = \log \text{vol } U_m - \log \text{vol } U_0$$

and (a) yields

$$\lim_{m \to \infty} \frac{1}{m} [S(\underline{\rho}_{[0]}) - S(\underline{\rho}_{[m]})] = -P_{Af}$$

We also have

$$S(\underline{\rho}_{[m]}) - S(\underline{\rho}^*_{[m]}) = -\int dx\, \underline{\rho}_m(x) \log \prod_{k=0}^{m-1} J(f^k x)$$

$$= -m \int \rho^{(m)}(dx) \log J(x)$$

hence, using (c),

$$\lim_{m \to \infty} \frac{1}{m} [S(\underline{\rho}_{[m]}) - S(\underline{\rho}^*_{[m]})] = -\mu(\log J)$$

and (d) follows. ∎

In conclusion, the entropy production e_A associated with escape from the Axiom A basic set A under f is

$$e_{Af}(\mu) = -P_{Af} - \mu(\log J)$$

This may be taken as a *definition* of $e_{Af}(\mu)$ for all $\mu \in I_A$ when A is an f-invariant set, f is not necessarily an Axiom A diffeomorphism, and I_A is the set of f-invariant probability measures with support in A. Notice that $e_{Af}(\mu) \neq e_f(\mu)$ unless $P_{Af} = 0$; this corresponds to the fact that e_{Af} and e_f describe different processes of entropy production (they coincide if A is an attracting set). It is readily seen that $e_{Af}(\mu)$ is independent of the choice of Riemann metric. Here again we shall prove positivity of the entropy production.

Proposition 3.2. Let $\mu \in \partial I_A$ satisfy the following extension of the Pesin identity:

$$h(\mu) - \sum \text{positive Lyapunov exponents} = P_{Af}$$

We have then

$$e_{Af}(\mu) \geqslant -P_{Af^{-1}} \geqslant 0 \tag{3.2}$$

Proof. We have indeed

$$\begin{aligned} e_{Af}(\mu) &= -h(\mu) + \sum \text{positive Lyapunov exponents for } (\mu, f) \\ &\quad - \sum \text{Lyapunov exponents for } (\mu, f) \\ &= -h(\mu) - \sum \text{negative Lyapunov exponents for } (\mu, f) \\ &= -\left[h_{f^{-1}}(\mu) - \sum \text{positive Lyapunov exponents for } (\mu, f^{-1}) \right] \\ &\geqslant -P_{Af^{-1}} \end{aligned}$$

and (3.2) follows from $P_{Af} \leqslant 0$. ▮

Remarks.

(a) Proposition 3.2 holds without restriction, but the interpretation of $e_{Af}(\mu)$ as entropy production and of $|P_{Af}|$ as escape rate are guaranteed only in the Axiom A case. For more general situations such interpretations remain conjectural.

(b) In the Axiom A case $e_{Af}(\mu) = 0$ implies that $P_{Af^{-1}} = 0$, i.e., A is an attractor for f^{-1}, and μ is the corresponding SRB measure on A.

ACKNOWLEDGMENTS

I am greatly indebted to Giovanni Gallavotti and Joel Lebowitz for discussion of the basic concepts and issues involved in the present paper, and to Jean-Pierre Eckmann for critical reading of the manuscript.

REFERENCES

1. R. Artuso, E. Aurell, and P. Cvitanović, Recycling of strange sets: I. Cycle expansions, II. Applications, *Nonlinearity* **3**:325–359, 361–386 (1990).
2. P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965).
3. N. Bourbaki, *Eléments de mathématiques. VI. Intégration* (Hermann, Paris, 1959), Chapter 6.
4. R. Bowen and D. Ruelle, The ergodic theory of Axiom A flow, *Invent. Math.* **29**:181–202 (1975).
5. N. I. Chernov, G. L.Eyink, J. L. Lebowitz, and Ya. G. Sinai, Steady-state electrical conduction in the periodic Lorentz gas, *Commun. Math. Phys.* **154**:569–601 (1993).
6. N. I. Chernov and J. L. Lebowitz, Stationary shear flow in boundary driven Hamiltonian systems, *Phys. Rev. Lett.* **75**:2831–2834 (1995).
7. N. I. Chernov and J. L. Lebowitz, In preparation.
8. P. Cvitanović, J.-P. Eckmann, and P. Gaspard, Transport properties of the Lorentz gas in terms of periodic orbits, *Chaos Solitons and Fractals* **4** (1995).
9. J. R. Dorfman, From molecular chaos to dynamical chaos, Lecture notes, Maryland (1995).
10. J. R. Dorfman and P. Gaspard, Chaotic scattering theory of transport and reaction rate coefficients, *Phys. Rev. E* **51**:28 (1995).
11. J. -P. Eckmann and D. Ruelle, Ergodic theory of strange attractors, *Rev. Mod. Phys.* **57**:617–656 (1985).
12. 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).
13. G. Gallavotti, Reversible Anosov diffeomorphisms and large deviations, *Math. Phys. Electronic J.* **1**:1–12 (1995).
14. G. Gallavotti, Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorems, preprint.
15. G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in nonequilibrium statistical mechanics, *Phys. Rev. Lett.* **74**:2694–2697 (1995).
16. G. Gallavotti and E. G. D. Cohen. Dynamical ensembles in stationary states, *J. Stat. Phys.* **80**:931–970 (1995).
17. P. Gaspard and G. Nicolis, Transport properties, Lyapunov exponents, and entropy per unit time, *Phys. Rev. Lett.* **65**:1693–1696 (1990).
18. W. G. Hoover, *Molecular Dynamics* (Springer, Berlin; 1986).
19. H. Kantz and P. Grassberger, Repellers, semi-attractors, and long-lived chaotic transients, *Physica* **17D**:75–86 (1985).
20. J. L. Lebowitz, Boltzmann's entropy and time's arrow, *Phys. Today* **46**(g):32–38 (1993).
21. F. Ledrappier, Propriétés ergodiques des mesures de Sinai, *Publ. Math. IHES* **59**:163–188 (1984).
22. F. Ledrappier and L. S. Young, The metric entropy of diffeomorphisms: I. Characterization of measures satisfying Pesin's formula, II. Relations between entropy, exponents and dimension, *Ann. Math.* **122**:509–539, 540–574 (1985).

23. D. Ruelle, A measure associated with Axiom A attractors, *Am. J. Math.* **98**:619–654 (1976).
24. D. Ruelle, Sensitive dependence on initial condition and turbulent behavior of dynamical systems, *Ann. N. Y. Acad. Sci.* **316**:408–416 (1978)
25. D. Ruelle, An inequality for the entropy of differentiable maps, *Bol. Soc. Bras. Mat.* **9**:83–87 (1978).
26. Ya. G. Sinai, Gibbs measures in ergodic theory, *Usp. Mat. Nauk* **27**(4):21–64 (1972) [*Russ. Math. Surv.* **27**(4):21–69 (1972)].
27. S. Smale, Differentiable dynamical systems *Bull. Am. Math. Soc.* **73**:747–817 (1967).