Chaotic Hypothesis and Universal Large Deviations Properties

Giovanni Gallavotti

Abstract. Chaotic systems arise naturally in Statistical Mechanics and in Fluid Dynamics. A paradigm for their modelization are smooth hyperbolic systems. Are there consequences that can be drawn simply by assuming that a system is hyperbolic? here we present a few model independent general consequences which may have some relevance for the Physics of chaotic systems.

Keywords and Phrases: Chaotic hypothesis, Anosov maps, Reversibility, Large deviations, Chaos

§1. Chaotic motions.*

A typical system exhibiting chaotic motions is a gas in a box whose particles interact via short range forces with a repulsive core, *e.g.* a hard core. No hope to ever be able to solve the evolution equations.

In the very simple case of pure hard cores it has been possible to prove, mathematically at least in some cases, that the system is ergodic, [Si1], [Sz], but ergodicity in itself is only a beginning of the qualitative theory of the motion. A similar situation arises in Fluid Mechanics: is a qualitative theory of Turbulence possible as, clearly, there are hopes to be able, in the near future, to prove an existence–uniqueness theorem but there is no hope for exact solutions of Navier Stokes equations?

Equilibrium Statistical Mechanics is a brilliant example of a very successful quantitative theory derived from a comprehensive qualitative hypothesis, the *ergodic hypothesis*. The key to its success is a *general* expression for the probability distribution μ on phase space M providing us with the *statistics* of the motions corresponding to given values of the macroscopic parameters determining the state of the system.

The statistics μ is defined in terms of the time evolution map S via the relation:

$$\lim_{T\to\infty}\frac{1}{T}\sum_{j=0}^{T-1}F(S^jx)=\int_M F(y)\mu(dy) \tag{1.1}$$

* Expanded text of the talk at the ICM98 in Berlin, 26 August 1998.

for all smooth *observables* F and for almost all, in the sense of volume measure on M, initial data $x \in M$. In Equilibrium Statistical Mechanics the distribution μ is identified with the uniform distribution on the surface of constant energy (the macroscopic state of the system being detemined by the volume V of the container box and by the energy U), which is an obviously invariant distribution by Liouville's theorem of Hamiltonian Mechanics: this is a necessary consequence of the ergodic hypothesis.

The success of Equilibrium Statistical Mechanics can be traced back to the fact that the ergodic hypothesis provides us with a concrete general, *model independent*, expression for the statistics of the motions. An expression that can be used to derive relations among *time averages* of various observables without even dreaming of ever being able to actually compute any of such averages.

The Boltzmann's *heat theorem*, the positivity of compressibility and specific heat are simple, but great, examples of such relations. They are relations which hold for any model, provided one makes the ergodicity hypothesis, see [Ga1]. A classical argument that can be used to derive the heat theorem (*i.e.* the second law of Thermodynamics) from ergodicity is provided us by Boltzmann, see Appendix A2 and [Ga2].

Consider a mechanical system: viewing its phase space as a discrete set of points the ergodic hypothesis says that motion is a one cycle permutation of the points. Given a initial datum with energy U and with volume V we define *temperature* the time average of kinetic energy $T = \langle K \rangle$ and *pressure* the time average of the derivative of the potential φ with respect to the volume V (note that the force acting on the particles consists of the internal pair forces *and* of the force that the walls exercize upon the particles which depends on the position of the walls, hence it does change when the volume varies). Here and below $\langle F \rangle$ will denote the time average of the observable F.

A general elementary property of a system whose motion on each energy surface is a single periodic motion is that if one calls $p = \langle \partial_V \varphi \rangle$ then:

$$\frac{dU + p\, dV}{T} = \ exact \tag{1.2}$$

which means that if the energy U and a parameter V on which the potential depends (it will be the volume in our case) are varied by dU and dV respectively *then the differential in (1.2) is exact.*

An elementary classical calculation shows that p, see Appendix A2, in the case of a gas in a box, has the meaning of average force exercized per unit surface on the walls of the container as a consequence of the particles collisions: thus we see that the ergodic hypothesis plus a general, trivial, identity among the averages of suitable mechanical quantities yields a relation ("equality of cross derivatives") holding without free parameters.

The reason why such relation is physically relevant for macroscopic systems is that the time necessary for the averages defining T, p to be reached within a good approximation by the finite time averages of $K, \partial_V \varphi$ is not the unobservable *recurrence time* (*i.e.* the superastronomic time for the system to complete a single tour of the energy surface U) but it is a much shorter physically observable time (whose theory is also due to Boltzmann being the essence of the Boltzmann's equation) because the quantities $K, \partial_V \varphi$ have an essentially constant value on the energy surface if the number of particles is large (so that the average of such observables "stabilizes" very rapidly compared to the recurrence times).

To summarize: a simple hypothesis allows us to find the statistics of the motions of an equilibrium system: this implies simple parmeterless relations among averages of physically relevant quantities (*i.e.* $\partial_V \frac{1}{T} = \partial_U \frac{p}{T}$) which are observable in large systems because such quantities average very quickly compared to the recurrence times (being practically constant on the surface of given energy if the system is large).

Thus a natural question arises: is there anything analogous in Non Equilibrium Statistical Mechanics? and in developed Turbulence?

The first problem is "what is the analogue of the uniform Liouville's distribution?". This is a really non trivial question that, once answered, will possibly allow us to try to find relations between time averages of mechanical quantities. The nontriviality is due to the fact that as soon as a system is out of equilibrium, *i.e.* nonconservative forces act upon it, dissipation is necessary in order to be able to reach a stationary state. But this means that *any* model used will be necessarily described by an evolution equation which will have a nonzero divergence: so that phase space will necessarily contract, in the average, and the statistics of the motion will be concentrated on a set of zero Lebesgue volume, see [Ru3].

Ruelle's proposal in the early 1970's was that one should regard such systems as *hyperbolic* so that there would be a unique stationary distribution describing the statistics of almost all initial data (chosen with the uniform distribution on phase space), [Ru1]. The ideas of Krylov, [Kr79], inspired Sinai in his development of the theory of Anosov systems via Markov partitions and, see [Si2], in conceiving complex mechanical systems as hyperbolic, and Ruelle's new ideas and his principle emerged, profiting of the important technical and conceptual achievements of Sinai.

This principle has been interpreted in [GC] as the following:

Chaotic hypothesis: A chaotic mechanical system can be regarded for practical purposes as a topologically mixing Anosov system.

This means that the closure of the attractor is a smooth surface on which the evolution is a Anosov system: of course assuming Axiom A instead of Anosov would be more natural, particularly in few degrees of freedom systems, [Ru1]. However I prefer to formulate the hypothesis in terms of Anosov system as fractality of the closure of the attractor seems to be of little relevance in systems with large number of degrees of freedom occurring in Statistical Mechanics.

The locution *practical purposes* is deliberately ambiguous as we know that even in Equilibrium Statistical Mechanics the corresponding ergodic hypothesis may fail while its consequences, at least some of them, will not (like the heat theorem in a free gas or in a harmonic chain).

The above physical discussion serves as a quick motivation of the mathematical question: *are there general properties shared by mechanical systems that are transitive or mixing Anosov systems?.*

In the next sections I provide some affirmative answer in the class of *time reversible Anosov maps* and of *weakly interacting chains of Anosov maps*. Recall: a *time reversal* symmetry for a dynamical system (M,S) is *any* isometric diffeomorphism I such that:

$$I^2=1,\qquad I\,S=S^{-1}\,I \tag{1.3}$$

Examples in Hamiltonian mechanical systems are the velocity reversal, or the composition of the velocity reversal and the parity symmetry, or the composition of the velocity reversal, parity symmetry and charge conjugation symmetry. In general a time reversal may be a symmetry quite different from the naive one that can be imagined, see [BG].

Hamiltonian systems on which further anholonomic constraints are imposed via Gauss' principle of *least constraint* often generate systems which show a time reversal symmetry, see Appendix A1, thus providing the simplest examples.

§2. Time reversible dissipative Anosov systems. Fluctuation theorem.

We now study a C^∞, topologically mixing, Anosov system (M,S) on a compact manifold M.

Let M be a d–dimensional, C^∞, compact manifold and let S be a C^∞, mixing (transitive would suffice) Anosov diffeomorphism, [AA], [Si1]. If W^u_x, W^s_x denote the *unstable* or *stable* manifold at $x\in M$, we call $W^{u,\delta}_x, W^{s,\delta}_x$ the connected parts of W^u_x, W^s_x containing x and contained in the sphere with center x and radius δ. Let d_u, d_s be the *dimensions* of W^u_x, W^s_x: $d=d_u+d_s$. We shall take δ always smaller than the smallest curvature radius of W^u_x, W^s_x for $x\in M$. Transitivity implies that W^u_x, W^s_x are dense in M for all $x\in M$.

The map S can be regarded, locally near x, either as a map of M to M or of W^u_x to W^u_{Sx}, or of W^s_x to W^s_{Sx}. The *Jacobian matrices* of the "three" maps will be $d\times d$, $d_u\times d_u$ and $d_s\times d_s$ matrices denoted respectively $\partial S(x)$, $\partial S(x)_u$, $\partial S(x)_s$. The absolute values of the respective determinants will be denoted $\Lambda(x)$, $\Lambda_u(x)$, $\Lambda_s(x)$ and are Hölder continuous functions, strictly positive (in fact $\Lambda(x)$ is C^∞), [Si1], [AA], [Ru4]. Likewise one can define the Jacobians of the n–th iterate of S; they are denoted by appending a label n to $\Lambda, \Lambda_u, \Lambda_s$ and are related to the latter by the differentiation chain rule:

$$\begin{aligned}\Lambda_n(x)=\prod_{j=0}^{n-1}\Lambda(S^jx),\quad \Lambda_{u,n}(x)=\prod_{j=0}^{n-1}\Lambda_u(S^jx),\\ \Lambda_{s,n}(S^jx)=\prod_{j=0}^{n-1}\Lambda_{s,n}(S^jx),\qquad \Lambda_n(x)=\Lambda_{u,n}(x)\,\Lambda_{s,n}(x)\chi_n(x)\end{aligned} \tag{2.1}$$

and $\chi_n(x)=\frac{\sin\alpha(S^nx)}{\sin\alpha(x)}$ is the ratio of the sines of the *angles* $\alpha(S^nx)$ and $\alpha(x)$ between W^u and W^s at the points S^nx and x. Hence $\chi_n(x)$ is bounded above and below in terms of a constant $B>0$: $B^{-1}\le\chi_n(x)\le B$, for all x (by the transversality of W^u and W^s).

We can define the *forward* and *backward statistics* or "SRB *distributions*" μ_+, μ_- of the volume measure μ_0 via the limits:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} F(S^{\pm k}x) = \int_{\mathcal{C}} \mu_\pm(dy)F(y) \equiv \mu_\pm(F) \tag{2.2}$$

which *exist for all smooth functions F on M and for all but a set of zero volume of initial points x*, see [Si1].

Therefore it is the probability distribution μ_+ that is the *statistics* μ of the motions (almost surely with respect to the volume measure μ_0 on M), see (1.1): *it plays the role of the Gibbs distribution, or microcanoncial ensemble, of equilibrium Statistical Mechanics.* Hence we are looking for general properties of μ_+, independent of the system considered, if possible.

Let $\Lambda(x) = |\det \partial S(x)|$; let $\mu_\pm$ be the forward and backward statistics of the volume measure μ_0 (*i.e.* the SRB distributions for S and S^{-1}).

Definition: The system (M,S) is dissipative if:

$$-\int_M \mu_\pm(dx)\log\Lambda^{\pm 1}(x) = \overline{\eta}_\pm > 0 \tag{2.3}$$

Remarks: 1) Existence of a time reversal symmetry I, see (1.3), implies $\overline{\eta}_+ = \overline{\eta}_-$ and $\Lambda(x) = \Lambda^{-1}(I\,x)$; furthermore $I\,W^u_x = W^s_{I\,x}$ and the dimensions of the stable and unstable manifolds d_s, d_u are equal: $d_u = d_s$ and $d = d_u + d_s$ is even.
2) if $\Lambda_u(x), \Lambda_s(x)$ denote the absolute values of the Jacobian determinants of S as a map of W^u_x to W^u_{Sx} and of W^s_x to W^s_{Sx}, then $\Lambda_u(x) = \Lambda_s(I\,x)^{-1}$.
3) If a system (M,S) is dissipative then the system (M',S') with $M' = M\times M$ and $S'(x,y) = (Sx, S^{-1}y)$ provides us with an example, setting $I(x,y) = (y,x)$, of a dynamical system in the general class of "reversible" Anosov maps considered in §1. It is remarkable that for Anosov systems it is $\overline{\eta}_\pm \ge 0$, see [Ru3].

From now on only reversible dissipative Anosov dynamical systems (M,S) will be considered: it is for such systems that it will be possible to derive general model independent properties.

Definition: The "dimensionless entropy production rate" or the "phase space contraction rate" at $x\in M$ and over a time τ is the function $\varepsilon_\tau(x)$:

$$x \to \varepsilon_\tau(x) = \frac{1}{\overline{\eta}_+\tau}\sum_{j=-\tau/2}^{\tau/2-1}\log\Lambda^{-1}(S^j x) = \frac{1}{\overline{\eta}_+\tau}\log\overline{\Lambda}_\tau^{-1}(x) \tag{2.4}$$

with $\overline{\Lambda}_\tau(x) \overset{def}{=} \prod_{-\tau/2}^{\tau/2-1}\Lambda(S^j x)$. Hence (see (2.2)) it is, with μ_0–probability 1:

$$\langle\varepsilon_\tau\rangle_+ = \lim_{T\to+\infty}\frac{1}{T}\sum_{j=0}^{T-1}\varepsilon_\tau(S^j x) \equiv \int_M \mu_+(dy)\varepsilon_\tau(y) = 1 \tag{2.5}$$

From the general theory of Anosov systems, [Si1], it follows that the μ_+–probability that $p = \varepsilon_\tau(x)$ is in the interval $[p-\delta, p+\delta]$ can be written as $\max_{q\in[p-\delta,p+\delta]} e^{\tau\overline{\zeta}(q)}$ for some suitably chosen function $\overline{\zeta}(p)$ and up to a factor bounded by $B^{\pm 1}$ with $0 < B < +\infty$. This is a deep result of Sinai that holds because the statistics μ_+ can be regarded as a Gibbs distribution and one can use the large deviation theory for such distributions: see Appendix A3 below for details. Then the following theorem holds, see [GC]:

Fluctuation theorem: The "large deviation function" $\overline{\zeta}(p)$ is analytic in an interval $(-p^*, +p^*)$ with $p^* \geq 1$ and verifies the relation:

$$\frac{\overline{\zeta}(p) - \overline{\zeta}(-p)}{p\overline{\eta}_+} = 1 \qquad |p| < p^* \tag{2.6}$$

i.e. the odd part of $\overline{\zeta}(p)$ is in general linear and its slope is equal to the average entropy creation rate.

What one really checks, see [Ga3], is the existence of $p^* \geq 1$ such that the SRB distribution μ_+ verifies:

$$p - \delta \leq \lim_{\tau\to\infty} \frac{1}{\overline{\eta}_+\tau} \log \frac{\mu_+(\{\varepsilon_\tau(x) \in [p-\delta, p+\delta]\})}{\mu_+(\{\varepsilon_\tau(x) \in -[p-\delta, p+\delta]\})} \leq p + \delta \tag{2.7}$$

for all p, $|p| < p^*$ and for any $\delta > 0$.

The above theorem was first informally proved in [GC] where its interest for nonequilibrium statistical mechanics was pointed out. The theorem can be regarded as a large deviation result for the probability distribution μ_+. Although I think that the physical interest of the theorem far outweighs its mathematical aspects it is useful to see a formal proof. A proof is reproduced in Appendix A3 below: it is taken from [Ga3].

The relation (2.6) has been tested numerically in several cases: it was in fact discovered in a numerical experiment, see [ECM2], and tested in other experiments, see [BGG], [BCL], [LLP]. Why does one need to test a theorem? the reason is that in concrete cases not only it is not known whether the system is Anosov but, in fact, it is usually clear that it is not, see [RT]. Hence the test is necessary to check the Chaotic Hypothesis which says that the failure of the Anosov property should be irrelevant for "practical purposes".

Another interesting aspect, that cannot be treated here for limitations of time, of the above theorem is that it can be interpreted as an extension to non zero forcing (*i.e.* $\overline{\eta}_+ > 0$) of the Green–Kubo relations: see [Ga6].

§3. Fluctuations in large systems.

An important drawback of the above fluctuation theorem, besides the reversibility assumption which is not verified in many important cases, is that it can be practically verified, for physical as well as mathematical reasons, only in (relatively) small systems.

In fact the logarithm of the entropy creation rate distribution $\tau\,\overline{\zeta}(p)$ is, usually, not only proportional to τ, *i.e.* to the time interval over which the entropy creation fluctuation is observed, but *also to the spatial extension of the system*, *i.e.* to the number of degrees of freedom; so that it is extremely unlikely that observing p in a large system one can see a value p which is appreciably different from 1 (note that the normalizing constant $\overline{\eta}_+$ in (2.4) is so chosen that the average of p in the stationary state is 1).

For this reason in macroscopic (or just "large") systems the phase space contraction rate is essentially constant (and its physical interpretation is of strength of the friction) much as the density is constant in gases at equilibrium. Therefore one can hope to see entropy creation rate fluctuations only if one can define a *local entropy creation rate* $\eta_{V_0}(x)$ associated with a microscopic region V_0 of space.

I now discuss, heuristically, why one should expect that a *local entropy creation rate can be defined, at least in some cases, and verifies a local version of the fluctuation law* (2.6). This is discussed in a special example, see [Ga7], as in general one can doubt that a local version of the fluctuation law holds, see [BCL].

The special example that we select is the *chain* of weakly coupled Anosov maps, well studied in the literature, [PS]. The system has a translation invariant spatial structure, *i.e.* it is a chain (or a lattice) of weakly interacting chaotic (mixing Anosov) system. This can be described as follows.

Let (M', S') be a dynamical system whose phase space M' is a product of $2N+1$ identical analytic manifolds $\overline{M}_0$: $M' = \overline{M}_0^{2N+1}$ and $S' : M' \to M'$ is a small perturbation of a product map $\overline{S}_0 \times \ldots \times \overline{S}_0 \stackrel{def}{=} \tilde{S}_0$ on M'. We assume that $(\overline{M}_0, \overline{S}_0)$ is a mixing Anosov systems. The size N (an integer) will be called the "spatial size" of the system.

For $x, y, z \in \overline{M}_0$ let $F_\varepsilon(x,z,y)$ be analytic and such that $z \to F_\varepsilon(x,z,y)$ is a map, of $\overline{M}_0$ into itself, ε*–close to the identity* and ε–analytic for $|\varepsilon|$ small enough. We suppose that, if $\underline{x} = (x_{-N}, \ldots, x_N) \in M'$:

$$(S'\,\underline{x}\,)_i = F_\varepsilon(x_{i-1}, x_i, x_{i+1}) \circ S_0 x_i \tag{3.1}$$

where $x_{\pm(N+1)}$ is *identified* with $x_{\mp N}$ (*i.e.* we regard the chain as periodic); we call such a dynamical system a *chain of interacting Anosov maps* coupled by nearest neighbors. It is a special example of the class of maps considered in [PS].[1]

It is difficult, maybe even impossible, to construct a (non trivial) reversible system of the above form: we therefore (see [Ga3]) consider the system (M, S) where $M = M' \times M'$ and define $S_0 \stackrel{def}{=} \tilde{S}_0 \times (\tilde{S}_0)^{-1}$ and $S \stackrel{def}{=} S' \times (S')^{-1}$, called hereafter the *free evolution* and the *interacting evolution*, respectively. So that the system can be considered as time reversible with a time reversal map $I(\,\underline{x}\,,\,\underline{y}\,) = (\,\underline{y}\,,\,\underline{x}\,)$. Note that the inverse map to (3.1) does not have the same form. The map S is, however, still in the class considered in [PS] because it can be written

[1] In the paper [PS] it is assumed that *also* $\overline{S}_0$ (hence S_0) is close to the identity, *e.g.* within ε: such condition does not seem necessary for the purposes of the present paper, hence it will not be assumed.

as $S(\,\underline{x}\,,\,\underline{y}\,)_i = \big(S(\,\underline{x}\,,\,\underline{y}\,)_{i1}, S(\,\underline{x}\,,\,\underline{y}\,)_{i2}\big)$ with:

$$\begin{aligned} S(\,\underline{x}\,,\,\underline{y}\,)_{i1} =& F_\varepsilon(x_{i-1},x_i,x_{i+1})\circ S_0\,x_i \\ S(\,\underline{x}\,,\,\underline{y}\,)_{i2} =& G_{\varepsilon,i}(\,\underline{y}\,)\circ S_0^{-1}y_i \end{aligned} \tag{3.2}$$

where G has "short range", *i.e.* $|G_\varepsilon(\,\underline{y}\,)_i - G_\varepsilon(\,\underline{y}\,')_i|$ is of order ε^k if $\underline{y}$ and $\underline{y}\,'$ coincide on the sites j with $|j-i|\le k$. By definition the system (M,S) is "*reversible*", *i.e.* the volume preserving diffeomorphism I verifies (1.3) above.

Therefore the points of the phase space M will be $(\,\underline{x}\,,\,\underline{y}\,) = (x_{-N},y_{-N},\ldots,x_N,y_N)$: however, to simplify notations, we shall denote them by $\underline{x} = (x_{-N},\ldots,x_N)$, with x_j denoting, of course, a *pair* of points in $\overline{M}_0$.

If ε is small enough the interacting system will still be hyperbolic, *i.e.* for every point $\underline{x}$ it will be possible to define a stable and an unstable manifolds $W^s_{\underline{x}}, W^u_{\underline{x}}$, [PS], so that the key notion of "Markov partition", [Si1], will make sense and it will allow us to transform, following the work [PS], the problem of studying the statistical properties of the dynamics of the system into an equivalent, but much more familiar, problem in equilibrium statistical mechanics of lattice spin systems interacting with short range forces. The reader will recognize below that this method is the natural extension to chains of the method used in Appendix A3 to study a single Anosov system.

The main notion that we want to introduce for our chain is the notion of *local entropy creation rate* $\eta_{V_0}(\,\underline{x}\,)$, the entropy creation rate inside a fixed finite set $V_0\subset[-N,N]$ of Anosov systems among the $2N+1$ composing the chain.

Definition: Fixed a point $\underline{x} = (\ldots,x_{\ell-1},x_\ell,x_{\ell+1},\ldots)$ *consider the map (3.1) as a map of* $\underline{x}_{\,V_0} \overset{def}{=} (x_j)_{j\in V_0} = (x_{-\ell},\ldots,x_\ell)$ *into:*

$$\underline{x}\,'_{V_0} = S(\ldots,x_{-\ell-1},\underline{x}_{\,V_0},x_{\ell+1},\ldots)_{V_0} \tag{3.3}$$

defined by (3.1) for $i\in[-\ell,\ell]$*. We call "local entropy production rate" associated with the "space like box"* $V_0=[-\ell,\ell]$ *at the phase space point* $\underline{x} = (\ldots,x_{\ell-1},x_\ell,x_{\ell+1},\ldots)$ *the quantity* $\eta_{V_0}(\,\underline{x}\,)$ *equal to minus the logarithm of the determinant of the* $2(2\ell+1)\times 2(2\ell+1)$ *Jacobian matrix of the map (3.3).*

Given a finite region V_0 centered at the origin and a time interval T_0, let η_+ denote the average density of entropy creation rate, *i.e.* $\eta_+ = \lim_{V_0,T_0\to\infty}\frac{1}{|T_0|}\frac{1}{|V_0|}\sum_{j=0}^{|T_0|-1}\eta_{V_0}(S^jx)$, then we set:

$$p = \frac{1}{\eta_+|V|}\sum_{j=-\frac12|T_0|}^{\frac12|T_0|}\eta_{V_0}(S^jx),\qquad V=V_0\times T_0 \tag{3.4}$$

where $\eta_{V_0}(x)$ denotes the entropy creation rate in the region V_0.

Calling $\pi_V(p)$ the probability distribution of p in the stationary state μ_+, *i.e.* in the SRB distribution, *and assuming that the system is a weakly coupled chain of Anosov systems* I shall show, heristically, that:

Proposition: It is $\pi_V(p) = e^{\zeta(p)|V|+O(|\partial V|)}$ *where* $|\partial V|$ *denotes the size of the boundary of the space–time region* V *and* $\zeta(p)$ *is a function analyticin* $p \in (-p^*, p^*)$ *for some* $p^* \ge 1$*. And:*

$$\begin{aligned} &\frac{\zeta(p)-\zeta(-p)}{p\,\eta_+} = 1, \qquad |p| < p^* \\ &\overline{\zeta}(p) = r\,\zeta(p), \qquad \overline{\eta}_+ = r\,\eta_+ \end{aligned} \tag{3.5}$$

where r *is the total "volume"* $(2N+1)$ *of the system, i.e. the "global" and "local" distributions are trivially related if appropriately normalized.*

Note that this implies that if V_0 is an interval of length $L = |V_0|$ and if $H = |T_0|$ then the relative size of the error and of the leading term will be, for some length R, of order $(L+H)R$ compared to order LH. Hence a relative error $O(H^{-1} + L^{-1})$ is made by using simply $\zeta(p)$ to evaluate the logarithm of the probability of p as defined by (3.4)).

The interest of the above statements lies in their independence on the total size $2N+1$ of the systems and the relevance of the above proposition for concrete applications should be clear.

It means that the fluctuation theorem leads to observable consequences if one looks at the far more probable microscopic fluctuations of the local entropy creation rate. One can test the relation (3.5) in a small region V_0 even when the system is very large: in such regions the entropy creation rate fluctuations will be frequent enough to be observable and carefully measurable. These fluctuations behave, therefore, just as ordinary density fluctuations at equilibrium: also the latter are not macroscopically observable but they are easily observable in small volumes.

The key results for the analysis leading to the above proposition are the papers [GC], [Ga3] and, mainly, [PS]: the latter paper provides us with a deep analysis of chains of Anosov systems and it contains, I believe, all the ingredients necessary to make the analysis mathematically rigorous: however I do not attempt at a mathematical proof here. The analysis is presented in Appendix A4 below.

Other types of fluctuation theorems (concerning non SRB distributions) had been previously found, see [ES]; extensions to stochastic systems have been recently discussed, see [Ku], [LS].

ACKNOWLEDGMENTS: *I have profited of stimulating discussions with F. Perroni, who also helped with numerical tests of the above ideas, with F. Bonetto and D. Ruelle. This work is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460; partially supported also by Rutgers University and CNR-GNFM.*

Appendix A1: The Gauss' minimal constraint principle.

Let $\varphi(\underline{\dot x}\,, \underline{x}\,) = 0$, $\underline{x}\; = \{\underline{\dot x}_{\,j}, \underline{x}_{\,j}\}$ be a constraint and let $\underline{R}\,(\underline{\dot x}\,, \underline{x}\,)$ be the constraint reaction and $\underline{F}\,(\underline{\dot x}\,, \underline{x}\,)$ the active force, see also Appendix A1 of [Ga3].

Consider all the possible accelerations $\underline{a}$ compatible with the constraints and a given initial state $\underline{\dot x}\,, \underline{x}\,$. Then $\underline{R}$ is *ideal* or *verifies the principle of minimal constraint* if the actual accelerations $\underline{a}_{\,i} = \frac{1}{m_i}(\underline{F}_{\,i} + \underline{R}_{\,i})$ minimize the *effort*:

$$\sum_{i=1}^{N} \frac{1}{m_i}(\underline{F}_{\,i} - m_i\,\underline{a}_{\,i})^2 \;\longleftrightarrow\; \sum_{i=1}^{N}(\underline{F}_{\,i} - m_i\,\underline{a}_{\,i})\cdot\delta\,\underline{a}_{\,i} = 0 \qquad (A1.1)$$

for all possible variations $\delta\,\underline{a}_{\,i}$ compatible with the constraint φ. Since all possible accelerations following $\underline{\dot x}\,, \underline{x}$ are such that $\sum_{i=1}^{N}\partial_{\underline{\dot x}_{\,i}}\varphi(\underline{\dot x}\,, \underline{x}\,)\cdot\delta\,\underline{a}_{\,i} = 0$ we can write:

$$\underline{F}_{\,i} - m_i\,\underline{a}_{\,i} - \alpha\,\partial_{\underline{\dot x}_{\,i}}\varphi(\underline{\dot x}\,, \underline{x}\,) = \underline{0} \qquad (A1.2)$$

with α such that $\frac{d}{dt}\varphi(\underline{\dot x}\,, \underline{x}\,) = 0$, *i.e.* :

$$\alpha = \frac{\sum_i(\underline{\dot x}_{\,i}\cdot\partial_{\underline{x}_{\,i}}\varphi + \frac{1}{m_i}\underline{F}_{\,i}\cdot\partial_{\underline{\dot x}_{\,i}}\varphi)}{\sum_i m_i^{-1}(\partial_{\underline{\dot x}_{\,i}}\varphi)^2} \qquad (A1.3)$$

which is the analytic expression of the Gauss' principle, see [LA].

Note that if the constraint is even in the $\underline{\dot x}_{\,i}$ then α is odd in the velocities: therefore if the constraint is imposed on a system with Hamiltonian $H = K + V$, with K quadratic in the velocities and V depending only on the positions, and if on the system act other purely positional forces (conservative or not) then the resulting equations of motion are reversible if time reversal is simply defined as velocity reversal.

The gaussian principle has been somewhat overlooked in the Physics literature in Statistical Mechanics: its importance has been only recently brought again to the attention, see the review [HHP]. A notable, though ancient by now, exception is a paper of Gibbs, [Gi], which develops variational formulas which he relates to the Gauss principle of least constraint.

Appendix A2. Heat theorem for monocyclic systems. Evaluation of the average $\langle\partial_V\varphi\rangle$.

Consider a 1–dimensional system with potential $\varphi(x)$ such that $|\varphi'(x)| > 0$ for $|x| > 0$, $\varphi''(0) > 0$ and $\varphi(x) \xrightarrow[x\to\infty]{} +\infty$ (in other words a 1–dimensional system in a confining potential). There is only one motion per energy value (up to a shift of the initial datum along its trajectory) and all motions are periodic so that the system is *monocyclic*. Assume also that the potential $\varphi(x)$ depends on a parameter V.

One defines *state* a motion with given energy E and given V. And:

U = total energy of the system $\equiv K + \varphi$
T = time average of the kinetic energy K

V = the parameter on which φ is suposed to depend
p = $-$ time average of $\partial_V\varphi$

A state is parameterized by U, V and if such parameters change by dU, dV respectively we define:

$$dL = -pdV, \qquad dQ = dU + pdV \tag{A2.1}$$

then:

Theorem (Helmholtz): *the differential* $(dU + pdV)/T$ *is exact.*

In fact let:

$$S = 2\log\int_{x_-(U,V)}^{x_+(U,V)} \sqrt{K(x;U,V)}dx = 2\log\int_{x_-(U,V)}^{x_+(U,V)} \sqrt{U-\varphi(x)}dx \tag{A2.2}$$

($\frac{1}{2}S$ is the logarithm of the action), so that:

$$S = \frac{\int (dU - \partial_V\varphi(x)dV)\frac{dx}{\sqrt{K}}}{\int K\frac{dx}{\sqrt{K}}} \tag{A2.3}$$

and, noting that $\frac{dx}{\sqrt{K}} = \sqrt{\frac{2}{m}}dt$, we see that the time averages are given by integrating with respect to $\frac{dx}{\sqrt{K}}$ and dividing by the integral of $\frac{1}{\sqrt{K}}$. We find therefore:

$$dS = \frac{dU + pdV}{T} \tag{A2.4}$$

Boltzmann saw that this was not a simple coincidence: his interesting (and healthy) view of the continuum (which he probably never really considered more than a convenient artifact, useful for computing quantities describing a discrete world) led him to think that in some sense *monocyclicity was not a strong assumption.*

In general one can call *monocyclic* a system with the property that there is a curve $\ell \to x(\ell)$, parameterized by its curvilinear abscissa ℓ, varying in an interval $0 < \ell < L(E)$, closed and such that $x(\ell)$ covers all the positions compatible with the given energy E.

Let $x = x(\ell)$ be the parametric equations so that the energy conservation can be written:

$$\frac{1}{2}m\dot{\ell}^2 + \varphi(x(\ell)) = E \tag{A2.5}$$

then if we suppose that the potential energy φ depends on a parameter V and if T is the average kinetic energy, $p = -\langle\partial_V\varphi\rangle$ it is, for some S:

$$dS = \frac{dE + pdV}{T}, \qquad p = -\langle\partial_V\varphi\rangle, \quad T = \langle K\rangle \tag{A2.6}$$

where $\langle\cdot\rangle$ denotes time average.

The above can be applied to a gas in a box. Imagine the box containing the gas to be covered by a piston of section A and located to the right of the origin at distance L: so that $V = AL$.

The microscopic model for the pistion will be a potential $\overline{\varphi}(L-\xi)$ if $x = (\xi,\eta,\zeta)$ are the coordinates of a particle. The function $\overline{\varphi}(r)$ will vanish for $r > r_0$, for some r_0, and diverge to $+\infty$ at $r = 0$. Thus r_0 is the width of the layer near the piston where the force of the wall is felt by the particles that happen to roam there.

Noting that the potential energy due to the walls is $\varphi = \sum_j \overline{\varphi}(L-\xi_j)$ and that $\partial_V\varphi = A^{-1}\partial_L\varphi$ we must evaluate the time average of:

$$\partial_L\varphi(x) = -\sum_j \overline{\varphi}'(L-\xi_j) \tag{A2.7}$$

As time evolves the particles with ξ_j in the layer within r_0 of the wall will feel the force exercized by the wall and bounce back. Fixing the attention on one particle in the layer we see that it will contribute to the average of $\partial_L\varphi(x)$ the amount:

$$\frac{1}{\text{total time}} 2\int_{t_0}^{t_1} -\overline{\varphi}'(L-\xi_j)dt \tag{A2.8}$$

if t_0 is the first instant when the point j enters the layer and t_1 is the instante when the ξ–compoent of the velocity vanishes "against the wall". Since $-\overline{\varphi}'(L-\xi_j)$ is the ξ–component of the force, the integral is $-2m|\dot{\xi}_j|$ (by Newton's law), provided $\dot{\xi}_j > 0$ of course.

The number of such contributions to the average per unit time are therefore given by $\rho_{wall}\, A \int_{v>0} 2mv\, f(v)\, v\, dv$ if ρ_{wall} is the density (average) of the gas near the wall and $f(v)$ is the fraction of particles with velocity between v and $v+dv$. Using the ergodic hypothesis (*i.e.* the microcanonical ensemble) and the equivalence of the ensembles to evaluate $f(v)$ it follows that:

$$p \stackrel{def}{=} \langle\partial_V\varphi\rangle = \rho_{wall}\beta^{-1} \tag{A2.9}$$

where $\beta^{-1} = k_B T$ with T the absolute temperature and k_B the Boltmann's constant. That the (A2.9) yields the correct value of the pressure is well known, see [MP], in Classical Statistical Mechanics; in fact often it is even taken as microscopic definition of the pressure.

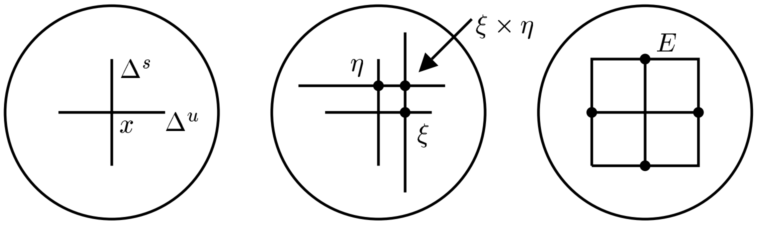
Appendix A3. A proof of the fluctuation theorem.

(A) Description of the SRB statistics.

A set E is a *rectangle* with *center* x and *axes* Δ^u, Δ^s if:
1) Δ^u, Δ^s are two connected surface elements of W^u_x, W^s_x containing x.
2) for any choice of $\xi \in \Delta^u$ and $\eta \in \Delta^s$ the local manifolds $W^{s,\delta}_\xi$ and $W^{u,\delta}_\eta$ intersect in one and only one point $x(\xi,\eta) = W^{s,\delta}_\xi \cap W^{u,\delta}_\eta$. The point $x(\xi,\eta)$ will also be denoted $\xi\times\eta$.

3) the boundaries $\partial\Delta^u$ and $\partial\Delta^s$ (regarding the latter sets as subsets of W^u_x and W^s_x) have zero surface area on W^u_x and W^s_x.
4) E is the set of points $\Delta^u \times \Delta^s$.

Note that *any* $x' \in E$ can be regarded as the center of E because there are Δ'^u, Δ'^s both containing x' and such that $\Delta^u \times \Delta^s \equiv \Delta'^u \times \Delta'^s$. Hence each E can be regarded as a rectangle centered at any $x' \in E$ (with suitable axes). See figure.



The circle is a small neighborhood of x; the first picture shows the axes; the intermediate picture shows the $\times$ operation and $W^{u,\delta}_\eta, W^{s,\delta}_\xi$; the third picture shows the rectangle E with the axes and the four marked points are the boundaries $\partial\Delta^u$ and $\partial\Delta^s$. The picture refers to a two dimensional case and the stable and unstable manifolds are drawn as flat, *i.e.* the Δ's are very small compared to the curvature of the manifolds. The center x is drawn in a central position, but it can be *any* other point of E provided Δ^u and Δ^s are correspondingly redefined. One should meditate on the symbolic nature of the drawing in the cases of higher dimension.

The *unstable boundary* of a rectangle E will be the set $\partial_u E = \Delta^u \times \partial\Delta^s$; the *stable boundary* will be $\partial_s E = \partial\Delta^u \times \Delta^s$. The boundary ∂E is therefore $\partial E = \partial_s E \cup \partial_u E$. The set of the *interior points* of E will be denoted E^0. A *pavement* of M will be a covering $\mathcal{E} = (E_1, \ldots, E_\mathcal{N})$ of M by $\mathcal{N}$ rectangles with pairwise disjoint interiors. The *stable (or unstable) boundary* $\partial_s\mathcal{E}$ (or $\partial_u\mathcal{E}$) of $\mathcal{E}$ is the union of the stable (or unstable) boundaries of the rectangles E_j: $\partial_u\mathcal{E} = \cup_j \partial_u E_j$ and $\partial_s\mathcal{E} = \cup_j \partial_s E_j$.

A pavement $\mathcal{E}$ is called *markovian* if its stable boundary $\partial_s\mathcal{E}$ retracts on itself under the action of S and its unstable boundary retracts on itself under the action of S^{-1}, [Si1], [Bo], [Ru1]; this means:

$$S\partial_s\mathcal{E} \subseteq \partial_s\mathcal{E}, \qquad S^{-1}\partial_u\mathcal{E} \subseteq \partial_u\mathcal{E} \tag{A3.1}$$

Setting $M_{j,j'} = 0$, $j, j' \in \{1, \ldots, \mathcal{N}\}$, if $SE^0_j \cap E^0_{j'} = \emptyset$ and $M_{j,j'} = 1$ otherwise we call C the set of sequences $\underline{j} = (j_k)^\infty_{k=-\infty}$, $j_k \in \{1, \ldots, \mathcal{N}\}$ such that $M_{j_k, j_{k+1}} \equiv 1$. The transitivity of the system (M, S) implies that M is *transitive*: *i.e.* there is a power of the matrix M with all entries positive. The space C will be called the space of the *compatible symbolic sequences*. If $\mathcal{E}$ is a markovian pavement and δ is small enough the map:

$$X : \underline{j} \in C \to x = \bigcap_{k=-\infty}^{\infty} S^{-k} E_{j_k} \in M \tag{A3.2}$$

is continuous and $1-1$ between the complement $M_0 \subset M$ of the set $N = \cup^\infty_{k=-\infty} S^k \partial\mathcal{E}$ and the complement $C_0 \subset C$ of $X^{-1}(N)$. This map is called the *symbolic code* of the points of M: it is a code that associates with each $x \notin N$ a sequence of symbols $\underline{j}$ which are the labels of the rectangles of the pavement that are successively visited by the motion $S^j x$.

The symbolic code X transforms the action of S into the *left shift* ϑ on C: $SX(\underline{j}) = X(\vartheta \underline{j})$. A key result, [Si1], is that it transforms the *volume measure* μ_0 on M into a *Gibbs distribution*, [LR], [Ru2], $\overline{\mu}_0$ on C with formal Hamiltonian:

$$H(\underline{j}) = \sum_{k=-\infty}^{-1} h_-(\vartheta^k \underline{j}) + h_0(\underline{j}) + \sum_{k=0}^{\infty} h_+(\vartheta^k \underline{j}) \tag{A3.3}$$

where, see (2.1):

$$\begin{aligned} h_-(\underline{j}) &= -\log \Lambda_s(X(\underline{j})), \quad h_+(\underline{j}) = \log \Lambda_u(X(\underline{j})), \\ h_0(\underline{j}) &= -\log \sin \alpha(X(\underline{j})) \end{aligned} \tag{A3.4}$$

If F is Hölder continuous on M the function $\overline{F}(\underline{j}) = F(X(\underline{j}))$ can be represented in terms of suitable functions $\Phi_k(j_{-k}, \ldots, j_k)$ as:

$$\overline{F}(\underline{j}) = \sum_{k=1}^{\infty} \Phi_k(j_{-k}, \ldots, j_k), \qquad |\Phi_k(j_{-k}, \ldots, j_k)| \le \varphi e^{-\lambda k} \tag{A3.5}$$

where $\varphi > 0, \lambda > 0$. In particular $h_\pm$ (and h_0) enjoy the property (A3.5) (*short range*).

If $\overline{\mu}_+, \overline{\mu}_-$ are the Gibbs states with formal Hamiltonians:

$$\sum_{k=-\infty}^{\infty} h_+(\vartheta^k \underline{j}), \qquad \sum_{k=-\infty}^{\infty} h_-(\vartheta^k \underline{j}) \tag{A3.6}$$

the distributions $\mu_\pm$ on M, images of $\overline{\mu}_\pm$ via the code X in (A3.2), will be the *forward* and *backward statistics* of the volume distribution μ_0 (corresponding to $\overline{\mu}_0$ via the code X), [Si1]. This means that:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} F(S^{\pm k} x) = \int_M \mu_\pm(dy) F(y) \equiv \mu_\pm(F) \tag{A3.7}$$

for all smooth F and for μ_0–almost all $x \in M$. The distributions $\mu_\pm$ are often called the *SRB distributions*, [ER]; the above statements and (A3.6),(A3.7) constitute the content of a well known theorem by Sinai, [Si1].

An approximation theorem for μ_+ can be given in terms of the *coarse graining* of M generated by the markovian pavement $\mathcal{E}_T = \bigvee_{k=-T}^{T} S^{-k}\mathcal{E}$.[3] If $E_{j_{-T},\ldots,j_T} \equiv \cap_{k=-T}^{T} S^{-k} E_{j_k}$ and $x_{j_{-T},\ldots,j_T}$ is a point chosen in the coarse grain set $E_{j_{-T},\ldots,j_T}$, so that its symbolic sequence is obtained by attaching to the right

[3] Where $\vee$ denotes the operation which, given two pavements $\mathcal{E}, \mathcal{E}'$ generates a new pavement $\mathcal{E} \vee \mathcal{E}'$: the rectangles of $\mathcal{E} \vee \mathcal{E}'$ simply consist of all the intersections $E \cap E'$ of pairs of rectangles $E \in \mathcal{E}$ and $E' \in \mathcal{E}'$.

and to the left of $j_{-T},\ldots,j_T$ arbitrary compatible sequences depending only on the symbols $j_{\pm T}$ respectively. We define the distribution $\mu_{T,\tau}$ by setting:

$$\begin{aligned}&\mu_{T,\tau}(F)\equiv\int_M\mu_{T,\tau}(dx)F(x)=\frac{\sum_{j_{-T},\ldots,j_T}\overline{\Lambda}^{-1}_{u,\tau}(x_{j_{-T},\ldots,j_T})F(x_{j_{-T},\ldots,j_T})}{\sum_{j_{-T},\ldots,j_T}\overline{\Lambda}^{-1}_{u,\tau}(x_{j_{-T},\ldots,j_T})}\\&\overline{\Lambda}_{u,\tau}(x)\stackrel{def}{=}\prod_{k=-\tau/2}^{\tau/2-1}\Lambda_u(S^kx)\end{aligned}\tag{A3.8}$$

Then for all smooth F we have: $\lim_{T\ge\tau/2,\,\tau\to\infty}\mu_{T,\tau}(F)=\mu_+(F)$. Note that equation (A3.8) can also be written:

$$\mu_{T,\tau}(F)=\frac{\sum_{j_{-T},\ldots,j_T}e^{-\sum_{k=-\tau/2}^{\tau/2-1}h_+(\vartheta^k\underline{j}^{\,0})}F(X(\underline{j}^{\,0}))}{\sum_{j_{-T},\ldots,j_T}e^{-\sum_{k=-\tau/2}^{\tau/2-1}h_+(\vartheta^k\underline{j}^{\,0})}}\tag{A3.9}$$

where $\underline{j}^{\,0}\in C$ is the compatible sequence agreeing with $j_{-T},\ldots,j_T$ between $-T$ and T (*i.e.* $X(\underline{j}^{\,0})=x_{j_{-T},\ldots,j_T}\in E_{j_{-T},\ldots,j_T}$) and continued outside as above.

Notation: to simplify the notations we shall write, when T is regarded as having a fixed value, $\underline{q}$ for the elements $\underline{q}=(j_{-T},\ldots,j_T)$ of $\{1,\ldots,\mathcal{N}\}^{2T+1}$; and $E_{\underline{q}}$ will denote $E_{j_{-T},\ldots,j_T}$ and $x_{\underline{q}}$ the above point of $E_{\underline{q}}$.

Remark: Note that the weights in (A3.9) depend on the special choices of the centers $x_{\underline{q}}$ (*i.e.* of $\underline{j}^{\,0}$); but if $x_{\underline{q}}$ varies in $E_{\underline{q}}$ the weight of $x_{\underline{q}}$ changes by at most a factor, bounded above by some $B<\infty$ and below by B^{-1}, for all $T\ge0$, and essentially depending only on the symbols corresponding to the sites close to $\pm T$.

The last formula shows that the forward statistics of μ_0 can be regarded as a Gibbs state for a *short range one dimensional spin chain with a hard core interaction.* The spin at k is the value of $j_k\in\{1,\ldots,\mathcal{N}\}$; the short range refers to the fact that the function $h_+(\underline{j})\equiv\log\Lambda_u(X(\underline{j}))$, ($\Lambda_u(x)$ *being Hölder continuous*), can be represented as in (A3.5) where the Φ_k play the role of "many spins" interaction potentials and the hard core refers to the fact that the only spin configurations $\underline{j}$ allowed are those with $M_{j_k,j_{k+1}}\equiv1$ for all integers k.

(B) A Legendre transform.

First the function (2.4) is converted to a function on the spin configurations $\underline{j}\in C$:

$$\tilde{\varepsilon}_\tau(\underline{j})=\varepsilon_\tau(X(\underline{j}))=\frac{1}{\tau}\sum_{k=-\tau/2}^{\tau/2-1}L(\vartheta^k\underline{j})\tag{A3.10}$$

where $L(\underline{j}\,) \equiv \frac{1}{\overline{\eta}_+} \log \Lambda^{\pm 1}(X(\underline{j}\,))$ has a *short range* representation of the type (A3.5).

The SRB distribution μ_+ is regarded (see above) as a Gibbs state $\overline{\mu}_+$ with short range potential on the space C of the compatible symbolic sequences, associated with a Markov partition $\mathcal{E}$, [Si1], [Ru2]. Therefore, by general large deviations properties of short range Ising systems ([La], [El], [Ol], there is a function $\overline{\zeta}(s)$ real analytic in s for $s \in (-p^*, p^*)$ for a suitable $p^* > 0$, strictly convex and such that if $p < p^*$ and $[p-\delta, p+\delta] \subset (-p^*, p^*)$ we have:

$$\frac{1}{\tau} \log \overline{\mu}_+(\{\tilde{\varepsilon}_\tau(\underline{j}\,) \in [p-\delta, p+\delta]\}) \xrightarrow[\tau\to\infty]{} \max_{s\in[p-\delta,p+\delta]} \overline{\zeta}(s) \tag{A3.11}$$

and the difference between the r.h.s. and the l.h.s. tends to 0 bounded by $D\tau^{-1}$ for a suitable constant D. The function $\overline{\zeta}(s)$ is the Legendre transform of the function $\lambda(\beta)$ defined as:

$$\lambda(\beta) = \lim_{\tau\to\infty} \frac{1}{\tau} \log \int e^{\tau\beta\tilde{\varepsilon}_\tau(\underline{j}\,)}\, \overline{\mu}_+(d\underline{j}\,) \tag{A3.12}$$

i.e. $\lambda(\beta) = \max_{s\in(-p^*,p^*)}(\beta s + \overline{\zeta}(s))$, where the quantity p^* can be taken $p^* = \lim_{\beta\to+\infty} g\beta^{-1}\lambda(\beta)$ and the function $\lambda(\beta)$ is a real analytic, [CO], strictly convex function of $\beta \in (-\infty, \infty)$ and $\beta^{-1}\lambda(\beta) \xrightarrow[\beta\to\pm\infty]{} \pm p^*$, *i.e.* it is asymptotically linear.

The above (A3.11) is a "large deviations theorem" for one dimensional spin chains with short range interactions, [La].

Hence it will be sufficient to prove the following; if $I_{p,\delta} = [p-\delta, p+\delta]$:

$$\frac{1}{\overline{\eta}_+\tau} \log \frac{\overline{\mu}_+(\{\tilde{\varepsilon}_\tau(\underline{j}\,) \in I_{p,\delta\mp\eta(\tau)}\})}{\overline{\mu}_+(\{\tilde{\varepsilon}_\tau(\underline{j}\,) \in I_{-p,\delta\pm\eta(\tau)}\})} \begin{cases} < p + \delta + \eta'(\tau) \\ > p - \delta - \eta'(\tau) \end{cases} \tag{A3.13}$$

with $\eta(\tau), \eta'(\tau) \xrightarrow[\tau\to\infty]{} 0$.

(C) Thermodynamic formalism informations.

In this section X will denote a lattice interval, *i.e.* a set of consecutive integers $X = (x, x+1, \ldots, x+n-1)$: hence it should not be confused with the code X of (A3.2).

Let $\underline{j}_X = (j_x, j_{x+1}, \ldots, j_{x+n-1})$ if $X = (x, x+1, \ldots, x+n-1)$ and n is odd, and call $\overline{X} = x + (n-1)/2$ the *center* of X. If $\underline{j} \in C$ is an infinite spin configuration we also denote $\underline{j}_X$ the set of the spins with labels $x \in X$. The left shift of the interval X will be denoted by ϑ; *i.e.* by the same symbol of the left shift of a (infinite) spin configuration $\underline{j}$.

Let $l_X(\underline{j}_X) = l^{(n)}(j_x, j_{x+1}, \ldots, j_{x+n-1})$, and $h_X^+(\underline{j}_X) = h_+^{(n)}(j_x, j_{x+1}, \ldots, j_{x+n-1})$ be translation invariant, *i.e.* functions such that $l_{\vartheta X}(\underline{j}\,) \equiv l_X(\underline{j}\,)$ and $h_{\vartheta X}^+(\underline{j}\,) = h_X^+(\underline{j}\,)$, and such that the functions $h_+(\underline{j}\,)$, see (2.4), and $L(\underline{j}\,)$, see (A3.10), can be written for suitably chosen constants b_1, b_2, b, b':

$$
\begin{aligned}
&L(\underline{j}\,) = \sum_{\overline{X}=0} l_X(\underline{j}_{\,X}), \qquad && h_+(\underline{j}\,) = \sum_{\overline{X}=0} h_X^+(\underline{j}_{\,X}) \\
&|l_X(\underline{j}_{\,X})| \le b_1 e^{-b_2 n}, && |h_X^+(\underline{j}_{\,X})| \le b e^{-b' n}
\end{aligned}
\tag{A3.14}
$$

Then $\tau\tilde\varepsilon_\tau(\underline{j}\,)$ can be written as $\sum_{\overline{X}\in[-\tau/2,\tau/2-1]} l_X(\underline{j}_{\,X})$.

Hence $\tau\tilde\varepsilon_\tau(\underline{j}\,)$ can be approximated by $\tau\tilde\varepsilon_\tau^M(\underline{j}\,) = \sum^{(M)} l_X(\underline{j}_{\,X})$ where $\sum^{(M)}$ means summation over the sets $X \subseteq [-\frac12\tau-M, \frac12\tau+M]$, while $\overline{X}$ is in $[-\frac12\tau, \frac12\tau-1]$. The approximation is described by:

$$
|\tau\tilde\varepsilon_\tau^M(\underline{j}\,) - \tau\tilde\varepsilon_\tau(\underline{j}\,)| \le b_3 e^{-b_4 M}
\tag{A3.15}
$$

for suitable[4] b_3, b_4 and for all $M \ge 0$. Therefore if $I_{p,\delta} = [p-\delta, p+\delta]$ and $M = 0$ we have:

$$
\mu_+(\{\varepsilon_\tau(x) \in I_{p,\delta}\}) \begin{cases} \le \overline{\mu}_+(\{\tilde\varepsilon_\tau^0 \in I_{p,\delta+b_3/\tau}\}) \\ \ge \overline{\mu}_+(\{\tilde\varepsilon_\tau^0 \in I_{p,\delta-b_3/\tau}\}) \end{cases}
\tag{A3.16}
$$

It follows from the general theory of 1–dimensional Gibbs distributions, [Ru2], that the $\overline{\mu}_+$–probability of a spin configuration which coincides with $\underline{j}_{\,[-\tau/2,\tau/2]}$ in the interval $[-\frac12\tau, \frac12\tau]$,[5] is:

$$
\frac{\Big[e^{-\sum^* h_X^+(\underline{j}_{\,X})}\Big]}{\sum_{\underline{j}'_{[-\tau/2,\tau/2]}} \Big[\,\cdot\,\Big]} P(\underline{j}_{\,[-\tau/2,\tau/2]})
\tag{A3.17}
$$

where $\sum^*$ denotes summation over all the $X \subseteq [-\tau/2, \tau/2-1]$; the denominator is just the sum of terms like the numerator, evaluated at a generic (compatible) spin configuration $\underline{j}\,'_{[-\tau/2,\tau/2]}$; finally P verifies the bound, [Ru2]:

$$
B_1^{-1} < P(\underline{j}_{\,[-\tau/2,\tau/2]}) < B_1
\tag{A3.18}
$$

with B_1 a suitable constant independent of $\underline{j}_{\,[-\tau/2,\tau/2]}$ and of τ (B_1 can be explicitly estimated in terms of b, b'). Therefore from (A3.16) and (A3.17) we deduce for any $T \ge \tau/2$:

$$
\begin{aligned}
&\mu_+(\{\varepsilon_\tau(x) \in I_{p,\delta}\}) \le \overline{\mu}_+(\{\tilde\varepsilon_\tau^0 \in I_{p,\delta+b_3/\tau}\}) \le \\
&\le B_2\, \mu_{T,\tau}(\{\tilde\varepsilon_\tau^0 \in I_{p,\delta+b_3/\tau}\}) \le B_2\, \mu_{T,\tau}(\{\tilde\varepsilon_\tau \in I_{p,\delta+2b_3/\tau}\})
\end{aligned}
\tag{A3.19}
$$

for some constant $B_2 > 0$; and likewise a lower bound is obtained by replacing B_2 by B_2^{-1} and b_3 by $-b_3$.

---

[4] One can check from (A3.14), that the constants b_3, b_4 can be expressed as simple functions of b_1, b_2.

[5] *i.e.* the spin configurations $\underline{j}\,'$ such that $j'_x = j_x$, $x \in [-\frac12\tau, \frac12\tau]$.

Then if $p < p^*$ and $I_{p,\delta} \subset (-p^*, p^*)$ the set of the rectangles $E \in \bigvee_{-T}^{T} S^{-k}\mathcal{E}$ with center x such that $\varepsilon_\tau(x) \in I_{p,\delta}$ is *not empty*, as it follows from the strict convexity and the asymptotic linearity of the function $\lambda(\beta)$ in (A3.12).

We immediately deduce the lemma:

Lemma 1: the distributions μ_+ and $\mu_{T,\tau}$, $T \ge \frac{1}{2}\tau$, verify:

$$\frac{1}{\tau\overline{\eta}_+}\log\frac{\mu_+(\{\varepsilon_\tau(x)\in I_{p,\delta\mp 2b_3/\tau}\})}{\mu_+(\{\varepsilon_\tau(x)\in -I_{p,\delta\pm 2b_3/\tau}\})}\begin{cases} < \frac{\log B_2^2}{\tau\overline{\eta}_+}+\frac{1}{\tau\overline{\eta}_+}\log\frac{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in I_{p,\delta}\})}{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in -I_{p,\delta}\})}\\ > -\frac{\log B_2^2}{\tau\overline{\eta}_+}+\frac{1}{\tau\overline{\eta}_+}\log\frac{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in I_{p,\delta}\})}{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in -I_{p,\delta}\})}\end{cases} \tag{A3.20}$$

for $I_{p,\delta} \subset [-p^*, p^*]$ and for τ so large that $p + \delta + 2b_3/\tau < p^*$.

Hence (A3.13) will follow if we can prove:

Lemma 2: there is a constant $\overline{b}$ such that the approximate SRB distribution $\mu_{T,\tau}$ verifies:

$$\frac{1}{\overline{\eta}_+\tau}\log\frac{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in I_{p,\delta}\})}{\mu_{T,\tau}(\{\tilde\varepsilon_\tau\in -I_{p,\delta}\})}\begin{cases}\le p+\delta+\overline{b}/\tau\\ \ge p-\delta-\overline{b}/\tau\end{cases} \tag{A3.21}$$

for τ large enough (so that $\delta + \overline{b}/\tau < p^* - p$) and for all $T \ge \tau/2$.

The latter lemma will be proved in §4 and it is the only statement that does not follow from the already existing literature.

(D) Time reversal symmetry implications

The relation (A3.20) holds for any choice of the Markov partition $\mathcal{E}$. Note that if $\mathcal{E}$ is a Markov pavement so is $i\mathcal{E}$ (because $iS = S^{-1}i$ and $iW_x^u = W_{ix}^s$); furthermore if $\mathcal{E}_1$ and $\mathcal{E}_2$ are Markov pavements then $\mathcal{E} = \mathcal{E}_1 \vee \mathcal{E}_2$ is also markovian. Therefore:

Lemma 3: there exists a time reversal Markov pavement $\mathcal{E}$, i.e. a Markov pavement such that $\mathcal{E} = i\mathcal{E}$.

This can be seen by taking any Markov pavement $\mathcal{E}_0$ and setting $\mathcal{E} = \mathcal{E}_0 \vee i\mathcal{E}_0$. Alternatively one could construct the Markov pavement in such a way that it verifies automatically the symmetry [G2]. Since the center of a rectangle $E_{\underline{q}} \in \mathcal{E}_T$ can be taken to be any point $x_{\underline{q}}$ in the rectangle $E_{\underline{q}}$ we can and shall suppose that the centers of the rectangles in $\mathcal{E}_T$ have been so chosen that the center of $iE_{\underline{q}}$ is $ix_{\underline{q}}$, *i.e.* the time reversal of the center $x_{\underline{q}}$ of $E_{\underline{q}}$.

For τ large enough the set of configurations $\underline{q} = \underline{j}_{[-T,T]}$ such that $\varepsilon_\tau(x) \in I_{p,\delta}$ for all $x \in E_{\underline{q}}$ is not empty[6] and the ratio in (A3.21) can be written, if $x_{\underline{q}}$ is the center of $E_{\underline{q}} \in \mathcal{E}_T$, as:

[6] Note that $p^* = \sup_x \limsup_{\tau\to+\infty} \varepsilon_\tau(S^{\tau/2}x)$ and let $p \in (-p^* + \delta, p^* - \delta)$; furthermore $\zeta(s)$ is smooth, hence $> -\infty$, for all $|s| < p^*$.

$$\frac{\sum_{\varepsilon_\tau(x_{\underline{q}})\in I_{p,\delta}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})}{\sum_{\varepsilon_\tau(x_{\underline{q}})\in -I_{p,\delta}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})} = \frac{\sum_{\varepsilon_\tau(x_{\underline{q}})\in I_{p,\delta}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})}{\sum_{\varepsilon_\tau(x_{\underline{q}})\in I_{p,\delta}} \overline{\Lambda}^{-1}_{u,\tau}(i\,x_{\underline{q}})} \tag{A3.22}$$

Define $\overline{\Lambda}_{s,\tau}(x)$ as in (A3.8) with s replacing u: then the time reversal symmetry implies that $\overline{\Lambda}_{u,\tau}(x) = \overline{\Lambda}^{-1}_{s,\tau}(ix)$, see remark 2) following definition (B), §2.[7] This permits us to change (A3.22) into:

$$\frac{\sum_{\varepsilon_\tau(x_{\underline{q}})\in I_{p,\delta}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})}{\sum_{\varepsilon_\tau(x_{\underline{q}})\in I_{p,\delta}} \overline{\Lambda}_{s,\tau}(x_{\underline{q}})} \begin{cases} < \max_{\underline{q}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})\overline{\Lambda}^{-1}_{s,\tau}(x_{\underline{q}}) \\ > \min_{\underline{q}} \overline{\Lambda}^{-1}_{u,\tau}(x_{\underline{q}})\overline{\Lambda}^{-1}_{s,\tau}(x_{\underline{q}}) \end{cases} \tag{A3.23}$$

where the maxima are evaluated as $\underline{q}$ varies with $\varepsilon_\tau(x_{\underline{q}}) \in I_{p,\delta}$.

By (2.1) we can replace $\overline{\Lambda}^{-1}_{u,\tau}(x)\overline{\Lambda}^{-1}_{s,\tau}(x)$ with $\overline{\Lambda}^{-1}_{\tau}(x)B^{\pm 1}$, see (A3.8), (2.4); thus noting that by definition of the set of $\underline{q}$'s in the maximum in (A3.23) we have $\frac{1}{\overline{\eta}_+\tau}\log\overline{\Lambda}^{-1}_{\tau}(x_{\underline{q}}) \in I_{p,\delta}$, we see that (A3.21) follows with $\overline{b} = \frac{1}{\overline{\eta}_+}\log B$.

Corollary: the above analysis gives us a concrete bound on the speed at which the limits in (2.6) are approached. Namely the error has order $O(\tau^{-1})$.

This is so because the limit (A3.11) is reached at speed $O(\tau^{-1})$; furthermore the regularity of $\lambda(s)$ in (A3.11) and the size of $\eta(\tau), \eta'(\tau)$ and the error term in (A3.21) have all order $O(\tau^{-1})$.

The above analysis proves a large deviation result for the probability distribution μ_+: since μ_+ is a Gibbs distribution, see (A3.6), various other large deviations theorems hold for it, [DV], [El], [Ol], but unlike the above they are not related to the time reversal symmetry.

APPENDIX A4: HEURISTIC PROOF OF THE LOCAL FLUCTUATION THEOREM.

(A) Markov partitions and symbolic dynamics for the chain.

The reduction of the dynamical nonequilibrium problem of a weakly interacting chain of Anosov maps, see §3, to a short range lattice spin system equilibrium problem is the content of (A), (B) of this appendix, see [Ga7]. This is an extension of the corresponding analysis in Appendix A3 for the case of a single Anosov map: it is necessary to discuss it again in order to exploit the short range nature of the coupling and its weakness in order to obtain results independent on the size N of the chain.

Let $\overline{\mathcal{P}}_0 = (E^0_1, \ldots, E^0_{\mathcal{N}_0})$ be a Markov partition, see [Si1], for the unperturbed "single site" system $(\overline{M}_0 \times \overline{M}_0, \overline{S}_0 \times \overline{S}_0^{-1})$. Then $\overline{\mathcal{P}}_0^{2N+1} = \{E_\alpha\}$,

---

[7] Here it is essential that $\overline{\Lambda}_{u,\tau}(x)$ is the expansion of the unstable manifold between the initial point $S^{-\tau/2}x$ and the final point $S^{\tau/2}x$: *i.e.* it is a trajectory of time length τ, which at its central time is in x.

$\alpha = (\rho_{-N}, \ldots, \rho_N)$ with $E_\alpha = E^0_{\rho_{-N}} \times E^0_{\rho_{-N+1}} \times \ldots \times E^0_{\rho_N}$ is a Markov partition of $(\overline{M}_0^{2(2N+1)}, S_0)$.

The perturbation, *if small enough*, will deform the partition $\overline{\mathcal{P}}_0^{2N+1}$ into a Markov partition $\mathcal{P}$ for (M, S) changing only "slightly" the partition $\overline{\mathcal{P}}_0^{2N+1}$. The work [PS] shows that the above "ε small enough" *mean that ε has to be chosen small but that it can be chosen N–independent*, as we shall always suppose in what follows.

Under such circumstances we can establish a correspondence between points of M that have the same "symbolic history" (or "symbolic dynamics") along $\overline{\mathcal{P}}_0^{2N+1}$ under S_0 and along $\mathcal{P}$ under S; we shall denote it by h; see [PS].

The Markov partition $\overline{\mathcal{P}}_0^{2N+1}$ for S_0 associates with each point $\underline{x} = (x_{-N}, \ldots, x_N)$ a sequence $(\sigma_{i,j})$, $i \in [-N, N], j \in (-\infty, \infty)$ of symbols so that $(\sigma_{i,j})_{j=-\infty}^{\infty}$ is the free symbolic dynamics of the point x_i. We call the first label i of $\sigma_{i,j}$ a "space–label" and the second a "time–label". Not all sequences can arise as histories of points: however (by the definition of h, see above) precisely the same sequences arise as histories of points along $\mathcal{P}_0$ under the free evolution S_0 or along $\mathcal{P}$ under the interacting evolution S.

The map h is Hölder continuous and "short ranged":

$$|h(\underline{x})_i - h(\underline{x}')_i| \le C \sum_j \varepsilon^{|i-j|\gamma'} |x_j - x'_j|^\gamma \tag{A4.1}$$

for some $\gamma, \gamma', C > 0$, [PS], if $|x - y|$ denotes the distance in $\overline{M}_0 \times \overline{M}_0$ (*i.e.* in the single site phase space).

Furthermore the code $\underline{x} \longleftrightarrow \underline{\sigma}$ associating with $\underline{x}$ its "history" or "symbolic dynamics" $\underline{\sigma}(\underline{x})$ along the partition $\mathcal{P}$ under the map S is such that, fixed j:

$$\underline{\sigma}(\underline{x})_i = \underline{\sigma}(\underline{x}')_i \text{ for } |i-j| \le \ell \qquad \Rightarrow \qquad |x_j - x'_j| \le C\varepsilon^{\gamma\ell} \tag{A4.2}$$

The inverse code associating with a history $\underline{\sigma}$ a point with such history will be denoted $\underline{x}(\underline{\sigma})$.

If $\underline{x} = (x_{-N}, \ldots, x_N)$ is coded into $\underline{\sigma}(\underline{x}) = (\underline{\sigma}_{-N}, \ldots, \underline{\sigma}_N) = (\sigma_{i,j})$, with $i = -N, \ldots, N$, and $j \in (-\infty, +\infty)$, the short range property holds also in the time direction. This means that, fixed i_0:

$$\sigma_{i,j} = \sigma'_{i,j} \text{ for } |i - i_0| < k, |j| < p \qquad \Rightarrow \qquad |\underline{x}(\underline{\sigma})_{i_0} - \underline{x}(\underline{\sigma}')_{i_0}| \le C\varepsilon^{\gamma k} e^{-\kappa p} \tag{A4.3}$$

for some $\kappa, \gamma, C > 0$, see lemma 1 of [PS]. The constants $\kappa, \gamma, C, C', B, B' > 0$ above and below should not be thought to be the same even when denoted by the same symbol: however they could be *a posteriori* fixed so that to equal symbols correspond equal values.

By construction the codes $\underline{x} \longleftrightarrow \underline{\sigma}(\underline{x})$ commute with time evolution.

The sequences $(\sigma_{i,j})$ which arise as symbolic dynamics along $\overline{\mathcal{P}}_0$ under the free single site evolution of a point x_i are subject to constraints, that we call "vertical", imposing that $T^0_{\sigma_{i,j},\sigma_{i,j+1}} \equiv 1$ for all j, if $T^0_{\sigma,\sigma'}$ denotes the "compatibility matrix" of the "free single site evolution" (*i.e.* $T^0_{\sigma,\sigma'}=1$ if the $\overline{S}_0\times\overline{S}_0^{-1}$ image of E_σ intersects the interior of $E_{\sigma'}$ and $T^0_{\sigma,\sigma'}=0$ otherwise). We call the latter condition a "compatibility condition" for the spins in the i–th column.

The mixing property of the free evolution immediately implies that a large enough power of the compatibility matrix T^0 has all entries positive. This means that for each symbol σ we can find semiinfinite sequences:

$$\begin{aligned}\sigma_B(\sigma)=&(\ldots,\sigma_{-1},\sigma_0\equiv\sigma),\qquad T^0_{\sigma_{i-1},\sigma_i}=1,\quad\text{for all } i\le 0\\ \sigma_T(\sigma)=&(\sigma\equiv\sigma_0,\sigma_1,\ldots),\qquad T^0_{\sigma_i,\sigma_{i+1}}=1,\quad\text{for all } i\ge 0\end{aligned}\tag{A4.4}$$

and defines two functions σ_B,σ_T, called "compatible extensions", defined on the set $\{1,\ldots,\mathcal{N}_0\}$ of labels of the single site Markov partition $\overline{\mathcal{P}}_0$, with values in the compatible semiinfinite sequences.

In fact there are (uncountably) many ways of performing such compatible extensions "from the bottom" and "from the top" of the symbol σ into semiinfinite compatible sequences of symbols. We imagine to select one pair σ_B,σ_T arbitrarily, once and for all, and call such a selection a "choice of boundary conditions" or "of extensions", on symbolic dynamics, for reasons that should become clear shortly. All this seems unavoidable and it is closely parallel to the corresponding discussion in the analysis of the simpler case of a single Anosov system discussed in Appendix A3, see the discussion preceding (A3.8).

We shall therefore be able to "extend in a standard way" any finite compatible block[8] Q of spins:

$$\underline{\sigma}_Q=(\sigma_{i,j})_{i\in L,j\in K},\qquad L=(a-\ell,a+\ell),\ K=(b-m,b+m)\tag{A4.5}$$

by setting $\sigma_{i,j}=\sigma_B(\sigma_{i,b-n})_{b-n-j}$ for $j<b-n$ and $\sigma_{i,j}=\sigma_T(\sigma_{i,b+n})_{j-b-n}$ for $j>b+n$. Here a,b,ℓ,m are integers.

In the free evolution there are no "horizontal" compatibility constraints; hence it is always possible to extend the finite block $\underline{\sigma}_Q=(\sigma_{i,j})_{i\in L,j\in K}$ to a "full spin configuration" sequence $(\sigma_{i,j})_{i\in[-N,N],j\in(-\infty,\infty)}$, obtained by continuing the columns in the just described standard way, using the boundary extensions σ_B,σ_T, above the top and below the bottom, into a biinfinite sequence and also by extending the spin configuration to the right and to the left to a sequence with spatial labels running over the full spatial range $[-N,N]$. One simply defines $\sigma_{i,j}$ for $i\notin L$ as *any* (but prefixed once and for all) compatible biinfinite sequence of symbols (the same for each column).

The allowed symbolic dynamics sequences for the free dynamics (on $\mathcal{P}_0$) and for the interacting dynamics (on $\mathcal{P}$) *coincide* because the free and the interacting dynamics are conjugated by the map h, [PS]. Therefore the above operations make sense *also* if the sequences are regarded as symbolic sequences of the interacting dynamics, as we shall do from now on.

To conclude: given a "block" $\underline{\sigma}_Q$ of symbols, with space–time labels $(i,j)\in Q=L\times K$, we can associate with it a point $\underline{x}\in M$ whose symbolic dynamics

[8] A block $(\sigma_{i,j})$, $(i,j)\in Q$, is naturally said to be "compatible" if $T^0_{\sigma_{i,j},\sigma_{i,j+1}}=1$ for all $(i,j)\in Q$ such that $(i,j+1)$ is also in Q.

is the above described standard extension of $\underline{\sigma}_Q$. The latter depends only on the values of $\sigma_{i,j}$ for j at the top or at the bottom of Q and, of course, on the boundary conditions σ_B, σ_T chosen to begin with.

(B) Expansion and contraction rates.

Consider the rates of variation of the phase space volume, $\lambda_0(\underline{x})$, or, respectively, of the surface elements of the stable and unstable manifolds $\lambda_s(\underline{x})$ and $\lambda_u(\underline{x})$ at the point $\underline{x}$: they are the logarithms of the Jacobian determinants $\partial S(\underline{x})$, $\partial_{(\alpha)} S(\underline{x})$, $\alpha = s, u$, where $\partial_{(\alpha)}$ denotes the Jacobian of S as a map of $W^\alpha_{\underline{x}}$ to $W^\alpha_{S\underline{x}}$ where $\alpha = u, s$ distinguishes the unstable manifold $W^u_{\underline{x}}$ of $\underline{x}$ or the stable manifold $W^\sigma_{\underline{x}}$ of $\underline{x}$:

$$\lambda_\alpha(\underline{x}) = -\log|\det \partial_{(\alpha)} S(\underline{x})|, \qquad \alpha = 0, u, s \tag{A4.6}$$

where $\partial_{(0)} S(\underline{x}) \stackrel{def}{=} \partial S(\underline{x})$.

A hard technical problem is to represent $\lambda_\alpha(\underline{x})$ in terms of the "symbolic history" of $\underline{x}$ along $\mathcal{P}$, *i.e.* in terms of compatible sequences $\underline{\sigma} = (\sigma_{i,j})$ with $i \in (-N, N)$, $j \in (-\infty, \infty)$. The rates $\lambda_\alpha(\underline{x})$ can be expressed as:

$$\lambda_\alpha(\underline{x}) = -\log\big|\det \frac{\partial S}{\partial \underline{x}}\big|_{W^\alpha(\underline{x})} = \sum_{L \subset [-N,N]} \tilde{\delta}^{(\alpha)}_L(\underline{x}_L) \tag{A4.7}$$

where L is an interval in $[-N, N]$ (with $\pm(N+1)$ identified with $\mp N$), [PS].

For $\alpha = 0$ this can be done by noting that the matrix $J = \frac{\partial S}{\partial x}$ has an almost diagonal structure: $J(\underline{x}) = J_0(\underline{x})(1 + \Delta(\underline{x}))$ where $J_0(\underline{x})$ is the Jacobian matrix of the free motion $J_0(\underline{x}) = \overline{J}_0(x_{-N}) \times \overline{J}_0(x_{-N+1}) \times \ldots \times \overline{J}_0(x_N)$ if $\underline{x} = (x_{-N}, \ldots, x_N)$ and if $D = \big(\prod_{j=-N}^N \det \overline{J}_0(x_j)\big)$:

$$\det J = D \cdot e^{\mathrm{Tr}\,\log(1+\Delta(\underline{x}))} = D \cdot e^{\sum_{k=1}^\infty \frac{(-1)^{k-1}}{k} \mathrm{Tr}\,\Delta(\underline{x})^k} \tag{A4.8}$$

which leads to (A4.7) if one uses that the matrix elements $\Delta_{p,q} = J_0^{-1}(\underline{x})\partial_{x_p}\partial_{x_q} J(\underline{x})$ are essentially local, *i.e.* bounded by $B\,(C\varepsilon)^{|p-q|\gamma}$ for some $\gamma, C, B > 0$ (see (3.1),(3.2), (A4.3)).

For $\alpha = u, s$ (A4.7) can be derived in a similar way using also that:

(1) the stable and unstable manifolds of $\underline{x}$ consist of points $\underline{y}$ which have eventually, respectively towards the future or towards the past, the same history of $\underline{x}$,
(2) they are described in a local system of coordinates around $\underline{x} = (\ldots, x_{-1}, x_0, x_1, \ldots)$ by smooth "short range" functions. Suppose, in fact, that on each factor M_0 we introduce a local system of coordinates (α, β) around the point $x_i \in M_0$, such that the unperturbed stable and unstable manifolds are described locally by graphs $(\alpha, f_s(\alpha))$ or $(f_u(\beta), \beta)$.

The unperturbed stable and unstable manifolds will be smooth graphs $(\alpha_i, f_s(\alpha_i))$ or $(f_u(\beta_i), \beta_i)$ with α_i varying close to $\overline{\alpha}_i$ and β_i close to $\overline{\beta}_i$, with $(\overline{\alpha}_i, \overline{\beta}_i)$ being the coordinates of x_i.

Fixed a point $\underline{x} = (x_{-N}, \ldots, x_N)$ with coordinates $(\overline{\alpha}_i, \overline{\beta}_i)_{i=-N,\ldots,N}$ the perturbed manifolds of the point $\underline{x}$ will be described by smooth (at least C^2 and in fact of any prefixed smoothness if ε is sufficiently small) functions $W^s(\underline{\alpha}), W^u(\underline{\beta})$ of $\underline{\alpha} = (\alpha_i)_{i=-N,N}$ or of $\underline{\beta} = (\beta_i)_{i=-N,N}$ which are "local"; *i.e.* if $\underline{\alpha}$ and $\underline{\alpha}'$ agree on the sites $i-\ell, i+\ell$ or if $\underline{\beta}$ and $\underline{\beta}'$ agree on the sites $i-\ell, i+\ell$ then:

$$\begin{aligned} &||W^u(\underline{\beta})_i - f_u(\beta_i)||_{C^2} < C\varepsilon, \qquad ||W^u(\underline{\beta})_i - W^u(\underline{\beta}')_i||_{C^2} < C\varepsilon^\ell \\ &||W^s(\underline{\alpha})_i - f_s(\alpha_i)||_{C^2} < C\varepsilon, \qquad ||W^s(\underline{\alpha})_i - W^s(\underline{\alpha}')_i||_{C^2} < C\varepsilon^\ell \end{aligned} \tag{A4.9}$$

for some $C > 0$, see [PS] lemmata 1,2. Here the norms in the first column are the norms in C^2 as functions of the arguments $\underline{\beta}$ or respectively $\underline{\alpha}$, while the norms in the second column are C^2 norms evaluated (of course) after identifying the arguments of $\underline{\beta}$ (or $\underline{\alpha}$) and $\underline{\beta}'$ (or $\underline{\alpha}'$) with labels j such that $|i-j| \le \ell$.
(3) If we consider the dependence of the planes tangent to the stable and unstable manifolds $W^s_{\underline{x}}$, $W^u_{\underline{x}}$ at $\underline{x}$ we find that they are Hölder continuous as functions of $\underline{x}$:

$$|(dW^\alpha_{\underline{x}})_i - (dW^\alpha_{\underline{y}})_i| < C\sum_j \varepsilon^{|i-j|\kappa}|x_j - y_j|^\gamma, \qquad \alpha = u, s \tag{A4.10}$$

where $(dW^\alpha_{\underline{x}})_i$ denoted the components relative to the i–th coordinate of $\underline{x}$ of the tangent plane to $W^\alpha_{\underline{x}}$ and $C, \kappa, \gamma > 0$.

The above properties and the Hölder continuity (A4.1), (A4.2), (A4.3) imply that the "horizontal potentials" $\tilde\delta^{(\alpha)}_L(\underline{x}_L)$ in (A4.7) are "short ranged":

$$|\tilde\delta^{(\alpha)}_L(\underline{x}_L)| \le B\,(C\varepsilon)^{(|L|-1)\gamma}, \qquad \alpha = u, s \tag{A4.11}$$

for some $B, C, \gamma > 0$; we denote $|L|$ the number of points in the set L.

We shall use the symbolic representation of $\underline{x} \in M$ to express the rates $\lambda^{(\alpha)}(\underline{x})$. For this purpose let $\underline{x} = (x_i)_{i=-N,N}$ and suppose that such $\underline{x}$ corresponds to the symbolic dynamics sequence $\underline{\sigma} = (\underline{\sigma}_j)_{j=-\infty}^{\infty}$ where $\underline{\sigma}_j = (\sigma_{-N,j}, \ldots, \sigma_{N,j})$. We denote $\underline{\sigma}_L$ the sequence $\underline{\sigma}_L = (\sigma_{i,j})_{i\in L, j=-\infty,\infty}$.

Then $\underline{\sigma}_L$ *does not* determine $\underline{x}_L$ (unless there is no interaction, *i.e.* $\varepsilon = 0$): however the short range property, (A4.3), of the symbolic codes and of the map h conjugating the free evolution and the interacting evolution shows that, if L' is a larger interval containing L and centered around L, then the sequence $\underline{\sigma}_{L'}$ determines each point of $\underline{x}_L$ within an approximation $\le (C\varepsilon)^{(|L'|-|L|)\gamma}$. Hence we can define $\widehat\delta^{(\alpha)}_L(\underline{\sigma}_L)$ so that:

$$\begin{aligned} &\tilde\delta^{(\alpha)}_L(\underline{x}_L) = \sum_{L'\supset L} \widehat\delta^{(\alpha)}_{L'}(\underline{\sigma}_{L'}), \qquad |\widehat\delta^{(\alpha)}_L(\underline{\sigma}_L)| < B'\,(C'\varepsilon^\gamma)^{|L|-1} \\ &\lambda_\alpha(\underline{x}) = \sum_L 2^{|L|}\widehat\delta^{(\alpha)}_L(\underline{\sigma}_L) \end{aligned} \tag{A4.12}$$

for some B', C', γ. This leads to expressing $\lambda_\alpha(\underline{x})$ in terms of the symbolic dynamics of $\underline{x}$ and of the "space–localized" potentials $\widehat\delta^{(\alpha)}_L(\underline{\sigma}_L)$.

Let $Q_n = L \times K$ where $K = [-n, n]$ is a "time–interval" and set

$$\mathcal{L}^{\alpha}_{Q_n}(\underline{\sigma}_{Q_n}) \stackrel{def}{=} \widehat{\delta}^{(\alpha)}_L([\underline{\sigma}_{Q_n}]) - \widehat{\delta}^{(\alpha)}_L([\underline{\sigma}_{Q_{n-1}}]) \tag{A4.13}$$

if $n \geq 1$ and $[\underline{\sigma}_{Q_n}]$ denotes a *standard extension* (in the sense of §3) of $\underline{\sigma}_{Q_n}$; or just set $\mathcal{L}^{\alpha}_{Q_0} \stackrel{def}{=} \widehat{\delta}^{(\alpha)}_L([\underline{\sigma}_{Q_0}])$ for $n = 0$. We define $\mathcal{L}^{\alpha}_Q(\underline{\sigma}_Q)$ for $Q = L \times K$ and K not centered (*i.e.* $K = (a - n, a + n)$, $a \neq 0$) so that it is translation invariant with respect to space time translations (of course the horizontal translation invariance is already implied by the above definitions and the corresponding translation invariance of $\widetilde{\delta}^{(\alpha)}_L$).

The *remarkable property*, consequence of the Hölder continuity of the functions in (A4.6) and of the (A4.3),(A4.12), see [PS], is that for some $\gamma, \kappa, B, C > 0$:

$$|\mathcal{L}^{\alpha}_Q(\underline{\sigma}_Q)| \leq B\,(C\varepsilon^{\gamma})^i\, e^{-\kappa j} \tag{A4.14}$$

if i, j are the horizontal and vertical dimensions of Q.

In this way we define a "space–time local potential" $\mathcal{L}^{(\alpha)}_Q$ which is, by construction, translation invariant and such that, if Λ denotes the box $\Lambda = [-N, N] \times [-M, M]$ the following representations for the rates in (A4.6) hold:

$$-\log|\det \partial_{(\alpha)} S^{2M+1}(S^{-M}\underline{x})| = \sum_{Q \subset \Lambda} \mathcal{L}^{\alpha}_Q(\underline{\sigma}_Q) + O(|\partial\Lambda|) \tag{A4.15}$$

where $O(|\partial\Lambda|)$ is a "boundary correction" due to the fact that in (A4.15) one should really extend the sum over the Q's centered at height $\leq M$ and contained in the infinite strip $[-N, N] \times [-\infty, \infty]$ rather than restricting Q to the region Λ. Hence the remainder in (A4.15) can, in principle, be explicitly written, in terms of the potentials $\mathcal{L}^{(\alpha)}_Q$, in the boundary term form usual in Statistical Mechanics of the 2–dimensional short range Ising model and it can be estimated to be of $O(|\partial\Lambda|)$ by using (A4.14).

(C) Symmetries. SRB states and fluctuations.

Besides the obvious translation invariance symmetry the dynamical system has a *time reversal symmetry*; this is the diffeomorphism I, see (1.3), which *anticommutes* with S and S_0:

$$IS = S^{-1}I, \qquad IS_0 = S_0 I^{-1}, \qquad I^2 = 1 \tag{A4.16}$$

We can suppose that the Markov partition is time reversible, *i.e.* to each element $E_{\underline{\sigma}}$ of the partition $\mathcal{P}$ one can associate an element $E_{\underline{\sigma}'} = IE_{\underline{\sigma}}$ which is *also* an element of the partition. Here we simply use the invariance of the Markov partition property under maps that either commute or anticommute with the evolution S: hence it is not restrictive, see [Ga5],[Ga3], to suppose that for each $\underline{\sigma}$ one can define a $\underline{\sigma}'$ so that $E_{\underline{\sigma}'} = IE_{\underline{\sigma}}$. We shall denote such $\underline{\sigma}'$ as $I\underline{\sigma}$ or also $-\underline{\sigma}$. For

$\varepsilon = 0$, *i.e.* for vanishing perturbation, the map I will act independently on each column of spins of $\underline{\sigma}$. This property remains valid for small perturbations; hence:

$$I\,\underline{\sigma} = \{\sigma'_{i,j}\} = \{-\sigma_{i,-j}\} \stackrel{def}{=} -\,\underline{\sigma}^{\,I} \tag{A4.17}$$

i.e. time reversal simply reflects the spin configuration corresponding to a phase space point and changes "sign" of each spin.

The functions $\lambda_\alpha(\underline{x})$ and their "potentials" $\mathcal{L}^\alpha_Q(\underline{\sigma}_Q)$ verify, as a consequence, if $Q = [-\ell,\ell]\times[-k,k]$ is a centered rectangle:

$$\lambda_\alpha(I\,\underline{x}) = -\lambda_{\alpha'}(\underline{x}), \qquad \mathcal{L}^\alpha_Q(\underline{\sigma}_Q) = -\mathcal{L}^{\alpha'}_Q(-\underline{\sigma}^{\,I}_Q) \tag{A4.18}$$

where $\alpha' = s$ if $\alpha = u$ and $\alpha' = u$ if $\alpha = s$, $\alpha' = 0$ if $\alpha = 0$. The above symmetries will be translated into remarkable properties of the SRB distribution.

The *"local entropy production rate"* associated with the "space like box" $V_0 = [-\ell,\ell]$ at the phase space point $\underline{x} = (\ldots, x_{\ell-1}, x_\ell, x_{\ell+1},\ldots)$ has been defined in §3 in therms of the Jacobian matrix of the map S. We can likewise consider the corresponding Jacobian determinants of the restriction of the map S to the stable and unstable manifolds of $\underline{x}$. Such determinants will depend not only from x_i, $i\in V_0$, and on the nearest neighbors variables $x_{\pm\ell}$ but *also* on the other ones x_k with $|k| > \ell+1$: however their dependence from the variables with labels $|k|>\ell$ is exponentially damped as $\varepsilon^{(|k|-\ell)\gamma}$, by (A4.14). Thus we can define $\eta^s_{V_0}, \eta^u_{V_0}$ in a way completely analogous to $\eta^0_{V_0}$ in (3.3).

If we look at the average phase space variation rates $\eta^0_{V_0}, \eta^s_{V_0}, \eta^u_{V_0}$ between the time $-\vartheta$ and ϑ we can find, via a power expansion like the one in (A4.8) along the lines leading from (A4.8) to (A4.15), a mathematical expression as:

$$\eta^\alpha_{V_0}(\underline{x}) \simeq \sum_Q{}^{*}\mathcal{L}^\alpha_Q(\underline{\sigma}_Q) \tag{A4.19}$$

where the $\sum^*_Q Q$ runs over rectangles Q centered at 0–time $Q = [a-\ell, a+\ell]\times[-k,k]$ with $[a-\ell, a+\ell]\subseteq V_0$. This could be taken as an alternative *definition* of $\eta^\alpha_{V_0}$, as it is a rather natural expression. For our purposes, if $V = V_0\times[-\vartheta,\vartheta]$, one needs to note that (A4.19) holds at least in the sense that:

$$\frac{1}{V_0\cdot(2\vartheta+1)}\sum_{j=-\vartheta}^{\vartheta}\eta^{(\alpha)}_{V_0}(S^j\,\underline{x}) = \frac{1}{V_0\cdot(2\vartheta+1)}\sum_{Q\subset V}\mathcal{L}^\alpha_Q(\underline{\sigma}_Q) + \frac{O(|\partial V|)}{|V|} \tag{A4.20}$$

i.e. expression (A4.19) can be used to compute the average local entropy creation rate in the space–time region V *up to boundary corrections* $O(|\partial V|)$ (that can be neglected for the purposes of the following discussion).

We now study the SRB distribution μ: denoting by $\langle F\rangle_+$ the average value with respect to μ of the observable F we can say, see [Si1], [PS], that if $\Lambda = [-N,N]\times[-T,T]$:

$$\langle F\rangle_+ = \lim_{T\to\infty}\frac{\sum_{\underline{\sigma}} F(\underline{\sigma})e^{\sum_{Q\subset\Lambda}\mathcal{L}^u_Q(\underline{\sigma}_Q)}}{\sum_{\underline{\sigma}} e^{\sum_{Q\subset\Lambda}\mathcal{L}^u_Q(\underline{\sigma}_Q)}} \tag{A4.21}$$

We want to study the properties of the fluctuations of:

$$p=\frac{1}{V\eta_+}\sum_{Q\subset V}\mathcal{L}^u_Q(\underline{\sigma}_Q),\qquad \text{if}\ \ \eta_+=\lim_{V\to\infty}\frac{1}{V}\sum_{Q\subset V}\langle\mathcal{L}^u_Q\rangle_+ \tag{A4.22}$$

for which we expect a distribution of the form $\pi_V(p)=\mathit{const}\,e^{V\zeta(p)+O(\partial V)}$. The SRB distribution gives to the event that p is in the interval dp the probability $\pi_V(p)dp$ with:

$$\pi_V(p)=\mathit{const}\sum_{at\ fixed\ p}e^{\sum_{Q\subset\Lambda}\mathcal{L}^u_Q(\underline{\sigma}_Q)} \tag{A4.23}$$

and (defining implicitly U^u):

$$\begin{aligned}\sum_{Q\subset\Lambda}\mathcal{L}^u_Q(\underline{\sigma}_Q)&=\sum_{Q\subset V}\mathcal{L}^u_Q(\underline{\sigma}_Q)+\sum_{Q\subset\Lambda/V}\mathcal{L}^u_Q(\underline{\sigma}_Q)+O(|\partial V|\,\kappa^{-1})\stackrel{def}{=}\\ &\stackrel{def}{=}U^u_V(\underline{\sigma}_V)+U^u_{\Lambda/V}(\underline{\sigma}_{\Lambda/V})+O(|\partial V|\,\kappa^{-1})\end{aligned} \tag{A4.24}$$

with $\kappa>0$, having used the "short range" properties (A4.14) of the potential.

In the sums in (A4.21) we would like to sum over $\underline{\sigma}_V$ and over $\underline{\sigma}_{\Lambda/V}$ as if such spins were independent labels. This is not possible because of the vertical compatibility constraints. However the mixing property supposed on the free evolution implies that the compatibility matrix T^0 raised to a large power R has positive entries. Hence if we leave a gap of width R above and below V we can regard as independent labels the labels $\sigma_{i,j}$ with i in the space part V_0 of the region $V=V_0\times[-\vartheta,\vartheta]$ and with $|j|>\vartheta+R$, by a distance $\ge R$ above or below the region V. Denoted $V+R\stackrel{def}{=}V_0\times[-\vartheta-R,\vartheta+R]$ remark that:

$$\sum_{Q\subset\Lambda}\mathcal{L}^u_Q(\underline{\sigma}_Q)=U^u_V(\underline{\sigma}_V)+U^u_{\Lambda/(V+R)}(\underline{\sigma}_{\Lambda/(V+R)})+O(|\partial V|\,(R+\kappa^{-1})) \tag{A4.25}$$

Hence, proceeding as in [GC1], we change the sum over (the dummy label) $\underline{\sigma}$ in the denominator to a sum over $-\underline{\sigma}^I$ and using $\mathcal{L}^u_{Q^I}(-\underline{\sigma}^I_Q)=-\mathcal{L}^s_Q(\underline{\sigma}_Q)$:

$$\frac{\pi_V(p)}{\pi_V(-p)}=\frac{\sum_{at\ fixed\ p}e^{\sum_{Q\subset V}\mathcal{L}^u_Q(\underline{\sigma}_Q)}e^{U^u_{\Lambda/(V+R)}(\underline{\sigma}_{\Lambda/(V+R)})}}{\sum_{at\ fixed\ p}e^{\sum_{Q\subset V}-\mathcal{L}^s_Q(\underline{\sigma}_Q)}e^{U^u_{\Lambda/(V+R)}((-\underline{\sigma}^I)_{\Lambda/(V+R)})}}\,e^{O(|\partial V|)} \tag{A4.26}$$

with the summation being over the spin configurations in the "whole space–time" Λ, subject to the specified constraint of having the same value for p, *i.e.* the same average local entropy creation rate in the space–time region V. The latter expression becomes, since the labels $\underline{\sigma},-\underline{\sigma}^I$ (respectively in the numerator and denominator of (A4.26)) are independent *dummy labels*:

$$\frac{\sum_{at\ fixed\ p}e^{\sum_{Q\subset V}\mathcal{L}^u_Q(\underline{\sigma}_Q)}Z(\Lambda/(V+R))}{\sum_{at\ fixed\ p}e^{\sum_{Q\subset V}-\mathcal{L}^s_Q(\underline{\sigma}_Q)}Z(\Lambda/(V+R))}\,e^{O(|\partial V|)} \tag{A4.27}$$

so that by the (A4.20), (A4.22) and since the symmetry relations above imply the relation $\sum_{Q\subset V}(\mathcal{L}^u_Q(\underline{\sigma}_Q)+\mathcal{L}^s_Q(\underline{\sigma}_Q)) = V\,\eta_+\,p$, up to corrections of size $O(|\partial V|\kappa^{-1})$ we find, (note the repetition of the comparison argument given in [GC]):

$$\frac{\pi_V(p)}{\pi_V(-p)} = e^{\eta_+\,V\,p}\;e^{O(|\partial V|)} \tag{A4.28}$$

yielding a *local fluctuation law*, *i.e.* the first of (3.5). The second line of (3.5) is a (simple) consequence of the above analysis but we do not discuss it here.

References.

[AA] Arnold, V., Avez, A.: *Ergodic problems of classical mechanics*, Benjamin, 1966.

[BCL] Bonetto, F., Chernov, N., Lebowitz, J.: , preprint, 1998.

[BG] Bonetto, F., Gallavotti, G.: *Reversibility, coarse graining and the chaoticity principle*, Communications in Mathematical Physics, 189, 263–276, 1997.

[BGG] Bonetto, F., Gallavotti, G., Garrido, P.: *Chaotic principle: an experimental test*, Physica D, 105, 226–252, 1997

[Bo] Bowen, R.: *Markov partitions for Axiom A diffeomorphisms*, American Journal of Mathematics, 92, 725–747, 1970.

[CO] Cassandro, M., Olivieri, E.: *Renormalization group and analyticity in one dimension. A proof of Dobrushin's theorem*, Communications in mathematical physics, 80, 255–269, 1981.

[DV] Donsker, M.D., Varadhan, S.R.S.: *Asymptotic evolution of certain Markov processes expectations for large time*, Communications in Pure and Applied Mathematics, 28, 279–301, 1975; 29, 389–461, 1976; 36, 182–212, 1983.

[ECM2] Evans, D.J.,Cohen, E.G.D., Morriss, G.P.: *Probability of second law violations in shearing steady flows*, Physical Review Letters, 71, 2401–2404, 1993.

[El] Ellis, R.S.: *Entropy, large deviations and statistical mechanics*, New York, Springer Verlag, 1985.

[ER] Eckmann, J.P., Ruelle, D.: *Ergodic theory of strange attractors*, Reviews in Modern Physics, 57, 617–656, 1985.

[ES] Evans, D.J., Searles, D.J.: *Equilibrium microstates which generate second law violating steady states*, Physical Review E, 50E, 1645–1648, 1994.

[Ga1] Gallavotti, G.: *Chaotic dynamics, fluctuations, nonequilibrium ensembles*, Chaos, 8, 384–392, 1998.

[Ga2] Gallavotti, G.: *Statistical Mechanics*, Springer–Verlag, in print.

[Ga4] Gallavotti, G.: *Chaotic hypothesis: Onsager reciprocity and fluctuation–dissipation theorem*, Journal of Statistical Phys., 84, 899–926, 1996.

[Ga3] Gallavotti, G.: *Reversible Anosov maps and large deviations*, Mathematical Physics Electronic Journal, MPEJ, (http:// mpej.unige.ch), 1, 1–12, 1995.

[Ga5] Gallavotti, G.: *New methods in nonequilibrium gases and fluids*, Proceedings of the conference *Let's face chaos through nonlinear dynamics*, U. of Maribor, 24

June– 5 July 1996, ed. M. Robnik, in print in Open Systems and Information Dynamics, Vol. 5, 1998 (also in mp_arc@ math. utexas.edu #96-533 and chao-dyn #9610018).

[Ga6] Gallavotti, G.: *Extension of Onsager's reciprocity to large fields and the chaotic hypothesis*, Physical Review Letters, 77, 4334–4337, 1996.

[Ga7] Gallavotti, G.: *A local fluctuation theorem*, in chao-dyn 9808005.

[GC] Gallavotti, G., Cohen, E.G.D: *Dynamical ensembles in nonequilibrium statistical mechanics*, Physical Review Letters, 74, 2694–2697, 1995. And *Dynamical ensembles in stationary states*, Journal of Statistical Physics, 80, 931–970, 1995.

[Gi] Gibbs, J. W.: *On the fundamental formula of dynamics*, American J. of Mathematics, Vol.II, 49–64, 1879. See also Volume II, Part 2 of Gibbs' *Collected Works*.

[HHP] Holian, B.L., Hoover, W.G., Posch. H.A.: *Resolution of Loschmidt's paradox: the origin of irreversible behavior in reversible atomistic dynamics*, Physical Review Letters, 59, 10–13, 1987.

[Kr] Krylov, N.S.: *Works on the foundations in statistical physics*, Princeton University Press, 1979, (english translations of selected papers).

[Ku] Kurchan, J.: *Fluctuation theorem for stochastic dynamics*, Journal of Physics A, 31, 3719–3729, 1998.

[LA] Levi-Civita, T., Amaldi, U.: *Lezioni di Meccanica Razionale*, Zanichelli, Bologna, 1927 (reprinted 1974), volume 3.

[La] Lanford, O.: *Entropy and equilibrium states in classical statistical mechanics*, ed. A. Lenard, Lecture notes in Physics, Springer Verlag, vol. 20, p. 1–113, 1973.

[LLP] Lepri. S., Livi, R., Politi, A. *Energy transport in anharmonic lattices close and far from equilibrium*, preprint, cond-mat@xyz. lanl. gov #9709156, to appear on Physica D.

[LR] Lanford, O., Ruelle, D.: *Observables at infinity and states with short range correlations in statistical mechanics*, Communications in Mathematical Physics, 13, p. 194–215, 1969.

[LS] Lebowitz, J.L., Spohn, H.: *The Gallavotti-Cohen Fluctuation Theorem for Stochastic Dynamics*, preprint, Rutgers University, June 1998.

[MP] Marchioro, C., Presutti, E.: *Thermodynamics of particle systems in presence of external macroscopic fields*, Communications in Mathematical Physics, 27, 146–154, 1972. And 29, 265–284, 1973.

[MR] Morriss, G.P., Rondoni, L.: *Applications of periodic orbit theory to $N-$particle systems*, Journal of Statistical Physics, 86, 991, 1997.

[Ol] Olla, S.: *Large deviations for Gibbs random fields*, Probability Theory and related fields, 77, 343–357, 1988.

[PS] Pesin, Y.B., Sinai, Y.G.: *Space–time chaos in chains of weakly inteacting hyperbolic mappimgs*, Advances in Soviet Mathematics, 3, 165–198, 1991.

[RT] Rom-Kedar, V., Turaev, D.: *Big islands in dispersing billiard–like potentials*, preprint, Weizmann Institute, April 2, 1998.

[Ru1] Ruelle, D.: *A measure associated with axiom A attractors*, American Journal of Mathematics, 98, 619–654, 1976. And *Sensitive dependence on initial conditions and turbulent behavior of dynamical systems*, Annals of the New York Academy of Sciences, 356, 408–416, 1978.

[Ru2] Ruelle, D.: *Statistical mechanics of a one dimensional lattice gas*, Communications in Mathematical Physics, 9, 267–278, 1968.

[Ru3] Ruelle, D.: *Positivity of entropy production in nonequilibrium statistical mechanics*, Journal of Statistical Physics, 85, 1–25, 1996. And: *Entropy production in nonequilibrium statistical mechanics*, Communications in Mathematical Physics, 189, 365–371, 1997.

[Ru4] Ruelle, D.: *Elements of differentiable dynamics and bifurcation theory*, Academic Press, 1989.

[Si1] Sinai, Y.: *Gibbs measures in ergodic theory*, Russian Mathematical Surveys, 27, 21–69, 1972 and *Lectures in ergodic theory*, Lecture notes in Mathematics, Princeton U. Press, Princeton, 1977.

[Si2] Sinai, Y.G.: *Development of Krylov's ideas*, in Krylov, N.S.: *Works on the foundations in statistical physics*, Princeton University Press, 1979, p. 239–281.

Giovanni Gallavotti
Fisica, Universitá La Sapienza
Roma, Italy
giovanni@ipparco.roma1.infn.it