A fluctuation theorem in a random environment

F. Bonetto, G. Gallavotti, G. Gentile

Mathematics, Georgia Tech., Atlanta

I.N.F.N. Roma 1, Fisica Roma1

Dipartimento di Matematica, Università di Roma Tre, 00146 Roma, Italy

Abstract. *A simple class of chaotic systems in a random environment is considered and their shadowing properties are studied. As an example of application the fluctuation theorem is extended under the assumption of reversibility.*

1. Random chaos

A chaotic system in random environment is defined by

(1) an *environment* which can be in states $\boldsymbol{\omega}$ belonging to a space Ω and varying in time, the time evolution being given by a map τ on Ω;

(2) a τ–invariant probability distribution P on Ω describing the statistical properties of the environment evolution;

(3) a family of maps $x \to f_{\boldsymbol{\omega}}(x)$ of a manifold $\mathcal{M}$ (*phase space*) into itself.

The space of the random events Ω will be supposed a space of sequences of finitely many symbols $\boldsymbol{\omega} = \{\omega_j\}_{j=-\infty}^{\infty}$, for instance a sequence of spins $\omega_j = \pm 1$, with the usual metric $d(\boldsymbol{\omega}, \boldsymbol{\omega}') = \sum_j |\omega_j - \omega'_j| 2^{-|j|}$, and with τ the shift to the left: $(\tau\boldsymbol{\omega})_j = \omega_{j+1}$. A "reflection" operation will be defined on Ω as $\boldsymbol{\omega} \longleftrightarrow \boldsymbol{\omega}^T$, with $(\boldsymbol{\omega}^T)_j = \omega_{-j}$. The probability distribution P will be a τ–invariant and reflection–invariant mixing process on Ω, for instance a Bernoulli shift or a Markov process with symmetric transition probabilities or a suitably symmetric general Gibbs distribution on Ω. Reflection–invariant means that $P(E) \equiv P(E^T)$ for every Borel set E, where E^T is the image of E under the reflection $\boldsymbol{\omega} \longleftrightarrow \boldsymbol{\omega}^T$. The manifold $\mathcal{M}$ will be a torus ^{m} and $f_{\boldsymbol{\omega}}$ will be supposed to be close to a map f which is a "linear hyperbolic torsion" of ^{m}, independent of the environment, *i.e.* close to a map which is defined by the action of a hyperbolic matrix with integer entries and determinant ± 1 on the torus.

For instance the point $x \in \mathcal{M}$ can be visualized as a pair of clock arms $x = \varphi = (\varphi_1, \varphi_2)$

frantically moving as

$$\varphi \to f\varphi = (\varphi_1 + \varphi_2, \varphi_1). \tag{1.1}$$

The map in (1.1) is a "linear hyperbolic torsion" of ^{2} and is a paradigmatic example of Anosov map.

The actual $f_{\boldsymbol{\omega}}$ will be small perturbations of f. For a small ε, we will take them of the form

$$f_{\boldsymbol{\omega}}(x) = fx + \varepsilon\psi_{\boldsymbol{\omega}}(x), \tag{1.2}$$

where $\psi_{\boldsymbol{\omega}}$ is periodic on ^{m}, analytic in a domain independent of $\boldsymbol{\omega}$ and Hölder continuous in $\boldsymbol{\omega}$ (we say that a function $g(\boldsymbol{\omega}, x)$ on $\Omega \times {}^m$ is Hölder continuous of modulus $C > 0$ and exponents $0 \le \beta < 1, \beta' > 0$ if $|g(\boldsymbol{\omega}, x) - g(\boldsymbol{\omega}', x')| \le C\,(|x - x'|^\beta + 2^{-\beta'\, n(\boldsymbol{\omega},\boldsymbol{\omega}')})$, where $n(\boldsymbol{\omega}, \boldsymbol{\omega}')$ is the maximum integer such that $\omega_i = \omega'_i$ for $|i| < n(\boldsymbol{\omega}, \boldsymbol{\omega}')$). For instance one can take $\psi_{\boldsymbol{\omega}}(x) = \omega_0\psi(x)$ with ψ a trigonometric polynomial.

Therefore we are led to consider the dynamical system $(\Omega \times \mathcal{M}, \mathcal{F})$, with

$$\mathcal{F}(\boldsymbol{\omega}, x) = (\tau\,\boldsymbol{\omega},\ f_{\boldsymbol{\omega}}(x)). \tag{1.3}$$

In other words at every instant t the "next coin is flipped", *i.e.* the next spin state is observed, and the system point x_t is moved by $fx_t + \varepsilon\psi_{\boldsymbol{\omega}}(x_t)$. We shall denote $\xi = (\boldsymbol{\omega}, x)$ the points in $\Omega \times \mathcal{M}$. We shall consistently note by roman letter f (possibly with labels) maps of $\mathcal{M}$ and by calligraphic letters $\mathcal{F}$ (possibly with labels) the corresponding random maps.

According to standard terminology if μ_0 denotes the normalized volume measure on $\mathcal{M}$ ("Liouville measure") the system will be said to possess a well defined statistics if, for almost all $(\boldsymbol{\omega}, x) \in \Omega \times \mathcal{M}$ in the $P \times \mu_0$-distribution sense, the limit

$$\lim_{T\to\infty} \frac{1}{T} \sum_{j=0}^{T-1} F(f^{*j}_{\boldsymbol{\omega}}(x)) = \int_{\mathcal{M}} \mu(\mathrm{d}y) F(y) \tag{1.4}$$

exists for all continuous "observables" F on $\mathcal{M}$ (hence for many others), and is independent of $\xi = (\boldsymbol{\omega}, x)$, thus defining a probability distribution μ which will be called the *statistics* of the motion or the SRB distribution; here $f^{*j}_{\boldsymbol{\omega}}$, $j \in$, is the x-component of the map $\mathcal{F}^j(\boldsymbol{\omega}, x)$, *i.e.* the composition of j maps $f_{\tau^{j-1}\boldsymbol{\omega}} \circ \ldots \circ f_{\boldsymbol{\omega}}$ if $j > 0$ and $f^{-1}_{\tau^j\boldsymbol{\omega}} \circ \ldots \circ f^{-1}_{\tau^{-1}\boldsymbol{\omega}}$ if $j < 0$, while $f^{*0}_{\boldsymbol{\omega}} = 1$. Here, and when appropriate below, 1 is used to denote *any* identity map.

More ambitiously one could look for a distribution $\mu_{\rm srb}$ of the form

$$\mu_{\rm srb}({\rm d}\boldsymbol{\omega}\,{\rm d}x) = P({\rm d}\boldsymbol{\omega})\mu_{\boldsymbol{\omega}}({\rm d}x), \tag{1.5}$$

with $\lim_{T\to\infty}\frac{1}{T}\sum_{j=0}^{T-1}F(\mathcal{F}^{*j}(\boldsymbol{\omega},x))$ existing for all continuous observables F on $\Omega\times\mathcal{M}$, apart from a 0-probability set with respect to the distribution $P\times\mu_0$, and having the form $\int_{\Omega\times\mathcal{M}}F(\boldsymbol{\omega}',y)\mu_{\rm srb}({\rm d}\boldsymbol{\omega}'{\rm d}y)\stackrel{\rm def}{=}\langle F\rangle_{\rm srb}$; when existing, $\mu_{\rm srb}$ is also called the SRB distribution and it is related to (1.4) by $\mu({\rm d}x)=\int_\Omega\mu_{\boldsymbol{\omega}}({\rm d}x)\,P({\rm d}\boldsymbol{\omega})$.

Systems more general than the above have been considered in [2]: there the quantity

$$\sigma_{\boldsymbol{\omega}}(x) = -\log\left|\det\frac{\partial f_{\boldsymbol{\omega}}(x)}{\partial x}\right| \tag{1.6}$$

is introduced and called *entropy production* rate. Furthermore it is proved that $\sigma_+\stackrel{\rm def}{=}\langle\sigma\rangle_{\rm srb}\ge 0$. Here we shall mainly consider the case $\sigma_+>0$, which is the generic case for ε small.

Particular attention will be given to *reversible* systems, namely systems for which there is an ε–independent smooth map $\mathcal{I}:\mathcal{I}(\boldsymbol{\omega},x)=(\boldsymbol{\omega}^T,Ix)$, with $(\boldsymbol{\omega}^T)_j=\boldsymbol{\omega}_{-j}$ in our cases, such that $\mathcal{I}^2=\pm1$ and $\mathcal{I}\circ\mathcal{F}^k=\mathcal{F}^{-k}\circ\mathcal{I}$ for some integer k; by smooth here we mean analytic in x. Examples of such systems are rather simple to give in the non–random case, see below for a few examples.

It can be remarked that the dynamical system (1.3) is a "skew product map" as the dynamics of the random generator of the $\boldsymbol{\omega}$'s is independent of the dynamics of the x's but controls it. Furthermore if the dynamics of the generator of the $\boldsymbol{\omega}$'s is a Markov process then the stationary distribution is also a Markov process.

Our aim in this paper is to discuss a method for constructing the properties of the stationary distribution, with all details necessary to treat concrete applicatons like the analysis of the large deviations and the corresponding fluctuation theorem. We follow the general ideas of the "cluster expansion" techniques, which were introduced in [3–7], but we develop them by following the tree expansion method already used in [1], see also [8].

2. Fluctuation theorem

In the non–random case, *i.e.* if $f_{\boldsymbol{\omega}}=f$ is independent of $\boldsymbol{\omega}$, and for f a general Anosov map it has been shown that a time reversal symmetry can be translated into certain relations between the probabilities of the "large fluctuations" of the time averages of the dimensionless observable $\sigma_{\boldsymbol\omega}(x)/\sigma_+$ (see after (1.6) for the definition of σ_+). Namely consider the observable

$$p(\xi) \stackrel{\rm def}{=} \frac{1}{T}\sum_{j=0}^{T-1} \frac{\sigma_{\tau^j\boldsymbol\omega}(f^{*j}_{\boldsymbol\omega}(x))}{\sigma_+}, \tag{2.1}$$

which is $\boldsymbol\omega$–independent if $f_{\boldsymbol\omega} \equiv f$ and which is introduced in a more general form for later reference; and call $\pi_T(p \in \Delta)$ the probability that it takes value in an interval Δ evaluated in the stationary SRB distribution (which for Anosov maps exists). Then $\pi_T(p \in \Delta) = e^{T\max_{p\in\Delta}\zeta(p)+O(1)}$ where $\zeta(p)$ is an analytic function of p defined in the interval $(-p^*, p^*)$, for some $p^* \geq 1$. The time reversal symmetry implies

$$\zeta(-p) = \zeta(p) - p\,\sigma_+, \qquad |p| < p^*, \tag{2.2}$$

which is called the *fluctuation theorem*, and was proven in [9–11].

The above statements concern non–random maps (*i.e.* $f_{\boldsymbol\omega} = f$ is $\boldsymbol\omega$-independent). In the present note we work in the following frame (which summarizes the notions introduced so far).

DEFINITION 1. *(i) Let* $\Omega \stackrel{\rm def}{=} \{1, 2, \ldots, n\}^{\mathbb Z}$ *be a space of bilateral sequences with* n *symbols per site, and let* $P(d\boldsymbol\omega)$ *be a probability distribution on* Ω *which is a short range Gibbs distribution (i.e. it has exponentially decaying potentials, for instance a stationary Bernoulli scheme or a stationary Markov process) and which is reflection invariant, i.e. invariant under the operation* $\boldsymbol\omega \longleftrightarrow \boldsymbol\omega^T$ *with* $(\omega^T)_j = \omega_{-j}$.
(ii) A "simple random map" $\Omega \times {}^m$ *will be a map of the form* $\mathcal F(\boldsymbol\omega, x) = (\tau\boldsymbol\omega, f_{\boldsymbol\omega}(x))$ *with* $f_{\boldsymbol\omega}(x) = fx + \varepsilon\psi_{\boldsymbol\omega}(x)$*, where* f *is a hyperbolic linear toral automorphism on* ^{m} *and the function* $\psi_{\boldsymbol\omega}(x)$ *is periodic and real analytic in* x*, with analyticity domain independent of* $\boldsymbol\omega$*, and Hölder continuous in* $\boldsymbol\omega$ *with modulus* C *and exponent* β'.
(iii) A ε*–independent map* $\mathcal I$*,* $\mathcal I(\boldsymbol\omega, x) = (\boldsymbol\omega^T, Ix)$ *such that* Ix *is analytic in* x*,* $\mathcal I^2 = \pm 1$ *and* $\mathcal I \circ \mathcal F^k = \mathcal F^{-k} \circ \mathcal I$ *for some integer* k *and for all* ε*, is called a "time reversal symmetry" for* $\mathcal F$ *and* $\mathcal F$ *is called a "reversible" random map.*

Examples of simple reversible maps can be obtained easily. Consider the paradigmatic example in which for $\varepsilon = 0$ the map f is the map of ^{2} in (1.1) and let $\mathcal F$ be any simple random map on $\Omega \times {}^2$; let $\mathcal T$ be the map $\mathcal T(\boldsymbol\omega, x, \boldsymbol\eta, y)$ defined on $\Omega \times {}^2 \times \Omega \times {}^2$ by

$$\mathcal T(\boldsymbol\omega, x, \boldsymbol\eta, y) = (\mathcal F(\boldsymbol\omega, x), \mathcal F^{-1}(\boldsymbol\eta, y)). \tag{2.3}$$

Then a time reversal symmetry for $\mathcal{T}$ (*i.e.* a smooth isometry anticommuting with time and squaring to the identity) can be defined as $\mathcal{I}(\boldsymbol{\omega}, x, \boldsymbol{\eta}, y) = (\boldsymbol{\eta}, y, \boldsymbol{\omega}, x)$.

The latter example is somewhat artificial: in applications time reversal should be a built-in symmetry so that its checking should be immediate. This is often the case in non–random systems. Note that in numerical simulations realizing the system $\mathcal{T}$ out of a simulation realizing a system $\mathcal{F}$ would be easy.

The above is a formal description of the structures so far discussed. In this paper we shall prove the following result.

THEOREM 1. *(i) Let $\mathcal{F}(\boldsymbol{\omega}, x) = (\tau\boldsymbol{\omega}, f_{\boldsymbol{\omega}}(x))$ be a simple random map with $f_{\boldsymbol{\omega}}(x) = fx + \varepsilon\psi_{\boldsymbol{\omega}}(x)$ (see Definition 1). Then for ε small enough the* SRB *distribution $\mu_{\rm srb}$, introduced in (1.5), exists and is unique.*
(ii) If furthermore $\mathcal{F}$ is reversible and the entropy production, i.e. (1.6), is such that $\sigma_+ \stackrel{\rm def}{=} \int \sigma_{\boldsymbol{\omega}}(x)\mu_{\rm srb}({\rm d}\boldsymbol{\omega}\,{\rm d}x) > 0$, relation (2.2) is satisfied by the large deviations of the random variable $p(\xi)$ in (2.1).

A proof could be attempted by trying to fall back on the already existing proofs of extensions of the fluctuation theorem to stochastic processes; the first one in [12] has been followed and widely extended in [13, 14].

In the first examples mentioned at the beginning of §1, in which the $\boldsymbol{\omega}$'s are a sequence of ± 1 with independent distribution or with finite range coupling, the process $\mu_{\rm srb}({\rm d}\boldsymbol{\omega}\,{\rm d}x)$ should be the stationary distribution generated by $P({\rm d}\boldsymbol{\omega})\mu_0({\rm d}x)$ by an evolution that is (essentially) a Markov process and the analysis in §2 of [13] should in principle be applicable. In the present case the phase space is a continuum but the results in [13] nevertheless apply formally, at least if one is willing to consider formally ratios of delta functions and interpret them as suitable Jacobian determinants.

To discuss the example of a general Gibbs distribution, in which the distribution of the $\boldsymbol{\omega}$ is far from Markovian, along the lines of [13] is harder. But one could consider the more general approach in [14]. However in this case one would have to prove that the SRB distribution is a "space-time Gibbs state", which is the basic object studied in [14].

In all cases an explicit determination of the stationary state seems *necessary* in order to check the assumptions and to define and compute the quantity $e(\lambda)$ of the quoted references. In other words the existing literature provides strong arguments (particularly in the

Markovian cases) for the validity of a relation like (2.2): but some work to check it remains to be done, on a case by case basis, and a substantial further work is necessary if one wishes to compute the function $\zeta(p)$ or, at least, some of its properties.

We shall also obtain (applying ideas and techniques in [1, 15]) the following result.

THEOREM 2. *For ε small enough and under the same assumptions of Theorem 1 (i), the following statements hold.*
(i) The SRB distribution $\mu_{\rm srb}$ is a short range Gibbs distribution;
(ii) one has, generically, $\sigma_+ > 0$ for $\varepsilon \neq 0$.

In §3 we construct a Hölder continuous homeomorphism conjugating $\mathcal{F}_0$ with $\mathcal{F}$ ("shadowing", Lemma 1). In §5 we derive the relation (which is not a conjugation) between the tangent map to $\mathcal{F}_0$ with the tangent map to $\mathcal{F}$ ("overshadowing", Lemma 2) preparing it by heuristic remarks in §4. The recognition that the SRB distribution is a Gibbs state with short range potential is in §6 (Lemma 3), allowing us to prove Theorem 2. In §7 the time reversal symmetry is used for the first time and from it a symmetry property of the short range potential is derived (Lemma 4); the proof of Theorem 1, which includes a fluctuation theorem, is then completed by collecting all the previous results. Several remarks throughout the text and in the concluding §8 address open problems and related conjectures.

Remark. The restriction to perturbations of Anosov maps generated by a linear torsion of a torus is due to the use made of the flatness and parallelism properties of the stable (respectively unstable) planes for such torsions. However by applying the methods in [1] we think that the same results could be extended to the general case in which perturbations of generic analytic Anosov maps of tori of arbitrary dimension are considered: this is a conjecture, see also comments in §8.

3. Decoupling and shadowing

The key for fulfilling the program set in §2, *i.e.* the proof of the two theorems, is to show the following result.

LEMMA 1. *Under the assumptions of Theorem 1 (i), there is a unique map H_ε of the form*

$$(\boldsymbol{\omega}', x') = H_\varepsilon(\boldsymbol{\omega}, x) \stackrel{\rm def}{=} (\boldsymbol{\omega}, x + h_{\boldsymbol{\omega}}(x)), \tag{3.1}$$

with $h_{\boldsymbol{\omega}}(x)$ *analytic in* ε *(but not in* x*, in general) such that if* $\mathcal{F}_0(\boldsymbol{\omega}, x) = (\tau\boldsymbol{\omega}, fx)$ *one has*

$$\mathcal{F} \circ H_\varepsilon = H_\varepsilon \circ \mathcal{F}_0, \tag{3.2}$$

and, for every fixed $\boldsymbol{\omega}$*,* $H_\varepsilon(\boldsymbol{\omega}, \cdot)$ *is a Hölder continuous homeomorphism of* ^{m}*, hence* H_ε *is also a Hölder continuous homeomorphism of* $\Omega \times {}^m$*.*

Remark. This means that there is a change of variables turning the perturbed map into the unperturbed one, *i.e.* the perturbed map $\mathcal{F}$ can be "conjugated" to the unperturbed one $\mathcal{F}_0$.

Proof. For simplicity the proof will be presented in the simplest case in which $m = 2$ and f is given by (1.1). The general case requires adding a few labels that enumerate a basis in the stable and unstable planes of the map f. Consider first that $\psi_{\boldsymbol{\omega}}$ is a trigonometric polynomial of degree N. The relation (3.2) becomes, see (1.2),

$$fh_{\boldsymbol{\omega}}(x) - h_{\tau\boldsymbol{\omega}}(fx) = -\varepsilon\psi_{\boldsymbol{\omega}}(x + h_{\boldsymbol{\omega}}(x)), \tag{3.3}$$

where f is a 2×2 matrix.

We look for a solution which is analytic in ε, *i.e.* $h_{\boldsymbol{\omega}}(x) = \varepsilon h^{(1)}_{\boldsymbol{\omega}}(x) + \varepsilon^2 h^{(2)}_{\boldsymbol{\omega}}(x) + \cdots$, with $h^{(k)}_{\boldsymbol{\omega}}(x)$ ε–independent functions. For instance the equation for the first order is

$$f\, h^{(1)}_{\boldsymbol{\omega}}(x) - h^{(1)}_{\tau\boldsymbol{\omega}}(fx) = -\psi_{\boldsymbol{\omega}}(x). \tag{3.4}$$

Call $v_\pm$ the two normalized eigenvectors of f relative to the eigenvalues $\lambda_\pm$ and let λ the inverse of the largest one ($\lambda = \frac{1}{2}(\sqrt{5} - 1)$), so that $\lambda_+ = \lambda^{-1}, \lambda_- = -\lambda$. Note that $\lambda < 1$.

The functions $\psi_{\boldsymbol{\omega}}, h_{\boldsymbol{\omega}}$ can be split into two components along the vectors $v_\pm$:

$$\psi_{\boldsymbol{\omega}}(x) = \psi_{\boldsymbol{\omega},+}(x)\, v_+ + \psi_{\boldsymbol{\omega},-}(x)\, v_-, \qquad h_{\boldsymbol{\omega}}(x) = h_{\boldsymbol{\omega},+}(x)\, v_+ + h_{\boldsymbol{\omega},-}(x)\, v_-, \tag{3.5}$$

and the equation (3.4) for $h^{(1)}_{\boldsymbol{\omega},\pm}$ gives

$$\lambda_+ h^{(1)}_{\boldsymbol{\omega},+}(x) - h^{(1)}_{\tau\boldsymbol{\omega},+}(fx) = \psi_{\boldsymbol{\omega},+}(x), \qquad \lambda_- h^{(1)}_{\boldsymbol{\omega},-}(x) - h^{(1)}_{\tau\boldsymbol{\omega},-}(fx) = \psi_{\boldsymbol{\omega},-}(x), \tag{3.6}$$

which can be solved *uniquely* by simply setting

$$h^{(1)}_{\boldsymbol{\omega},\alpha}(x) = -\sum_{p \in \mathbb{Z}_\alpha} \alpha \lambda_\alpha^{-|p+1|_\alpha} \psi_{\tau^p\boldsymbol{\omega},\alpha}(f^p x), \tag{3.7}$$

where $\alpha = \pm$, $_+ = [0, \infty) \cap$, $_- = (-\infty, 0) \cap$ and the inequality $\lambda < 1$ ensures convergence.

Hence the equations for $h^{(k)}_\pm$ become

$$
\begin{aligned}
h^{(k)}_{\boldsymbol{\omega},\alpha}(x) = -\sum_{s=0}^{\infty} \frac{1}{s!} \sum_{\substack{k_1+\cdots+k_s=k-1,\, k_i\ge 0\\ \alpha_1,\ldots,\alpha_s=\pm}} \sum_{p\in\mathbb{Z}_\alpha} \alpha\lambda_\alpha^{-|p+1|\alpha} \cdot \\
\cdot \Big(\prod_{j=1}^{s} (v_{\alpha_j} \cdot \partial_x) \Big) \psi_{\tau^p\boldsymbol{\omega},\alpha}(f^p x) \cdot \Big(\prod_{j=1}^{s} h^{(k_j)}_{\tau^p\boldsymbol{\omega},\alpha_j}(f^p x) \Big).
\end{aligned}
\tag{3.8}
$$

This *uniquely* determines, by recursion, $h^{(k)}_{\boldsymbol{\omega},\alpha}$ and can be written via a graphical representation as in Figure 1, where

(a) the "graph elements" consisting of an arrow emerging from a bullet and carrying the labels α, k or α_i, k_i represent $h^{(k)}_{\boldsymbol{\omega},\alpha}(x)$ or $h^{(k_i)}_{\boldsymbol{\omega},\alpha_i}(x)$, and will be called "endlines";

(b) the "graph element" consisting in the small black circle carrying a label p into which the s lines merge represents the result of the operation $-\frac{1}{s!}\alpha\lambda_\alpha^{-|p+1|\alpha}\Big(\prod_{j=1}^{s}(v_{\alpha_j}\cdot\partial_x)\Big)\psi_{\tau^p\boldsymbol{\omega},\alpha}(f^p x)$, and will be called a "node";

(c) the graph on the right represents the product of the quantities represented by its graph elements (s endlines and 1 node).

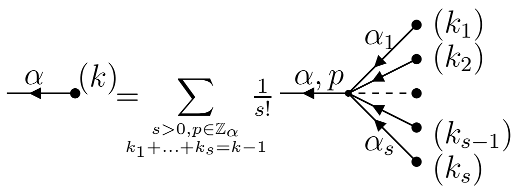

FIGURE 1. Graphical interpretation of (3.8) for $k > 1$.

Remarks. (1) Note that the graphical representation is simply a way of visualizing (3.8): the graph contains all labels present in (3.8) and necessary to represent the functions of $(\boldsymbol{\omega}, x)$ which intervene in the formula.

(2) Graphs are of great help when large expressions with many labels have to be studied: in many cases they allow us to see immediately the validity of general bounds on quantities of interest. They are of wide use in perturbation theory in quantum mechanics, both non–relativistic and relativistic, and in classical mechanics. Their power is exploited in studying convergence of series in which cancellations or resummations have to be exhibited in order to show convergence, as for the renormalized series in quantum field theory or in Hamiltonian stability in classical mechanics.

(3) The present case is particularly simple because no cancellations will be needed and the graphical representation will only be used to obtain "graph-by-graph" bounds.

Representing again, in the same way, the graph elements that appear on the r.h.s. one obtains an expression for $h^{(k)}_{\boldsymbol{\omega},\alpha}(x)$ in terms of *trees*, oriented "toward the root".

We recall briefly the notion of tree (by referring to [8] for further details). A graph is a collection of points (nodes) and lines connecting all of them, and it is called planar if it can be drawn in a plane without lines crossing. A tree is a planar graph containing no closed loops. An oriented tree is a tree with a special node v_0. We can add a further line ℓ_0 connecting v_0 to a further point, which is not a node and which will be called the root of the tree: the line ℓ_0 will be called the root line. The tree so modified will be called a rooted tree. We shall say that the tree is labeled if the nodes and the lines carry some labels. In the following by tree we shall always mean a labeled rooted tree. A rooted tree is a partially ordered set: each line can be imagined to carry an arrow pointing toward the root.

The labels will be defined as follows. A tree θ with k nodes will carry on each line ℓ a pair of labels α_ℓ, p_ℓ, with $p_\ell \in$ and $\alpha_\ell \in \{-,+\}$, and on each node v a pair of labels (not shown in Figure 1) α_v, p_v, with $\alpha_v = \alpha_{\ell_v}$ and $p_v \in {}_{\alpha_v}$ such that

$$p(v) = p_{\ell_v} = \sum_{w \succeq v} p_v, \tag{3.9}$$

where the sum is over the nodes following v (*i.e.* over the nodes w, denoted $w \succeq v$, along the path connecting v to the root) and ℓ_v denotes the line exiting from the node v. To each tree we shall assign, given $\boldsymbol{\omega}$, a *value* given by the product of all the quantities represented by its graph elements:

$$\mathrm{Val}(\theta) = \prod_{v \in V(\theta)} \frac{-\alpha_v}{s_v!} \lambda_{\alpha_v}^{-|p_v+1|\alpha_v} \cdot \Big(\prod_{j=1}^{s_v} \partial_{\alpha_{v_j}} \Big) \psi_{\tau^{p(v)}\boldsymbol{\omega},\alpha_v}(f^{p(v)}x), \tag{3.10}$$

where $\partial_\alpha \stackrel{\text{def}}{=} v_\alpha \cdot \partial_x$, $V(\theta)$ is the set of nodes in θ, and the nodes $v_1, \ldots, v_{s_v}$ are the s_v nodes preceding v (if v is a top node then the derivatives are simply missing). If $\Theta_{k,\alpha}$ denotes the set of all trees with k nodes and with label α associated with the root line, then one has

$$h_{\boldsymbol{\omega},\alpha}(x) = \sum_{k=1}^{\infty} \varepsilon^k h^{(k)}_{\boldsymbol{\omega},\alpha}(x), \qquad h^{(k)}_{\boldsymbol{\omega},\alpha}(x) = \sum_{\theta \in \Theta_{k,\alpha}} \mathrm{Val}(\theta), \tag{3.11}$$

and the "only" problem left is to estimate the radius of convergence of the above formal power series.

The estimates are most conveniently done by studying the Fourier transform of the function $h_{\boldsymbol{\omega},\alpha}(x)$. This is easily done, again, graphically. Since the value of a tree θ is a product of functions of x associated with the nodes of θ then its Fourier tranform is obtained simply by considering the Fourier transform of the functions associated with each node v evaluated at some $\nu_v \in {}^2$ and taking their convolution. Hence it is enough to attach a label $\nu_v \in {}^2$ to each node and define the *momentum* that "flows" on the line ℓ_v as

$$\nu_{\ell_v} \stackrel{\rm def}{=} \sum_{w \preceq v} \nu_w, \tag{3.12}$$

where the sum is over the nodes preceding v. Then in (3.11) we can write

$$h^{(k)}_{\boldsymbol{\omega},\alpha}(x) = \sum_{\nu \in \mathbb{Z}^2} {\rm e}^{i\nu\cdot x}\, h^{(k)}_{\boldsymbol{\omega},\alpha,\nu}, \tag{3.13}$$

and the convolution necessary to produce the ν-th Fourier component of $h^{(k)}$ can be written immediately as (note that $f = f^T$ for f as in (1.1))

$$\begin{aligned} h^{(k)}_{\boldsymbol{\omega},\alpha,\nu} &= \sum_{\theta\in\Theta_{k,\nu,\alpha}} {\rm Val}(\theta), \\ {\rm Val}(\theta) &= \Big(\prod_{v\in V(\theta)} \frac{-\alpha_v}{s_v!} \lambda_{\alpha_v}^{-|p_v+1|\alpha_v} \psi_{\tau^{p(v)}\boldsymbol{\omega},\alpha_v,f^{-p(v)}\nu_v} \Big) \cdot \prod_{\substack{v\in V(\theta)\\ v\neq v_0}} (i f^{-p(v')}\nu_{v'}\cdot v_{\alpha_v}), \end{aligned} \tag{3.14}$$

where $\Theta_{k,\nu,\alpha}$ denotes the set of all trees with k nodes and with labels ν and α associated with the root line, v' denotes the node immediately following v (*i.e.* such that ℓ_v enters v'), v_0 is the special node of θ, *i.e.* the node immediately preceding the root, and we have redefined the tree value in order to take into account also the new Fourier labels.

This completes the description of an algorithm leading to the explicit expression of the Taylor coefficients $h^{(k)}_{\boldsymbol{\omega}}$ for $h_{\boldsymbol{\omega}}$ and we can proceed to bound them. Actually we not only want to show that the power series for $h_{\boldsymbol{\omega}}$ converges but we also want to prove that the resulting function of x is quite regular, namely that it is Hölder continuous with exponent $\beta < 1$ which can be prefixed at the prize of taking ε small enough.

Since we have at hand an expression for the Fourier transform of $h^{(k)}$ we can achieve a proof of convergence and of Hölder continuity with exponent β simply by showing that $\sum_k \sum_\nu |\nu|^\beta |h^{(k)}_{\alpha,\nu}| < \infty$ (a sufficient condition for Hölder continuity of exponent $\beta < 1$ of a function $g(x)$ is that $\sum_\nu |g_\nu|\,|\nu|^\beta < \infty$ because of the obvious inequality $|{\rm e}^{i\nu\cdot(x-y)} - 1| \le 2|\nu|^\beta|x-y|^\beta$ for $x, y \in {}^m$).

The only problem is given by the presence of the factor $|\nu|^\beta$. Consider first the case $\beta = 0$. Recall that we are assuming that $\psi_{\boldsymbol{\omega}}(x)$ is a trigonometric polynomial of degree N, *i.e.* $\psi_{\boldsymbol{\omega},\alpha,\nu}\neq 0$ only for $|\nu|\le N$. Then there are only $(2N+1)^2<(3N)^2$ possible choices for each ν_v, given $p(v)$, such that $|f^{-p(v)}\nu_v|\le N$. Hence fixed θ, $\{\alpha_v\}_{v\in V(\theta)}$ and $\{p_v\}_{v\in V(\theta)}$ the remaining sum of products in (3.14) is bounded by (if $\lambda\equiv\lambda_+^{-1}\equiv-\lambda_-$)

$$(3N)^{2k}N^k\Psi^k\prod_{v\in V(\theta)}\frac{\lambda^{|p_v+1|}}{s_v!},\qquad \Psi=\max_{\substack{|\nu|\le N\\ \alpha=\pm}}\max_{\boldsymbol{\omega}\in\Omega}|\psi_{\boldsymbol{\omega},\alpha,\nu}|, \tag{3.15}$$

having bounded by N^k the last product in (3.14). The sum over the $\{p_v\}_{v\in V(\theta)}$ is a multiple geometric series bounded by $(1/(1-\lambda))^k$.

The combinatorial factors arising from the Taylor expansion leading to (3.8) (see the $s!^{-1}$ factors in (3.8) or the $s_v!^{-1}$ factors in (3.15)) can be bounded by 1 in the case of trigonometric polynomials. We shall see that, on the contrary, they will play a role in the case of analytic perturbations.

We regard two trees as distinct if they cannot be superposed by pivoting the lines around the nodes that they enter. The number of distinct trees with k lines is bounded by the number of random walks of $2k$ steps, hence by 2^{2k}. (Note that, unlike in [8], we are not numbering the lines of the trees: this would lead to a new combinatorial factor $1/k!$ instead of $\prod_v(1/s_v!)$ and to the bound $k!2^{2k}$ on the number of trees with numbered lines).

In conclusion, for $\beta=0$ the conjugating function H_ε exists and it is uniformly continuous and uniformly bounded with a uniformly summable Fourier transform, for ε (even complex) in $|\varepsilon|<\varepsilon_0(0)\stackrel{\rm def}{=}(3N)^{-3}\Psi^{-1}2^{-4}(1-\lambda)$, where an extra factor 2^{-1} has been inserted in order to obtain uniform bounds.

Given a tree θ, taking $\beta>0$ requires estimating $|\nu|^\beta$: we bound it by $\sum_v|\nu_v|^\beta$. Then from $|f^{-p(v)}\nu_v|\le N$ (otherwise the tree value vanishes), we infer that $|\nu_v|\le\lambda^{-|p(v)|}BN$, where $B\ge 1$ is a suitable constant. The sum $\sum_v|\nu_v|^\beta$ is over k terms which can be estimated separately so that we can write $\sum_v|\nu_v|^\beta\le k|\nu_{\overline{v}}|^\beta$ where $|\nu_{\overline{v}}|=\max_v|\nu_v|$. This can be taken into account by multiplying (3.15) by an extra factor $(BN)^\beta\lambda^{-\beta|p(\overline{v})|}\le BN\lambda^{-\beta\sum_v|p_v|}$. Therefore if $\beta<1$ the sum $\sum_\nu|\nu|^\beta|h^{(k)}_{\boldsymbol{\omega},\alpha,\nu}|$ can be studied as in the case $\beta=0$ but the result is modified into

$$\varepsilon_0(\beta)=(3N)^{-3}\Psi^{-1}(1-\lambda^{1-\beta})2^{-4}. \tag{3.16}$$

This shows that H_ε is analytic in ε in the disk with radius $\varepsilon_0(\beta)$. Furthermore, since in (3.16) we inserted (for simplicity) an extra factor 2^{-1} in excess of the result obtained by the procedure described, the Hölder modulus is also uniformly bounded by a suitable function

$C(\beta)$ of β. Note that $\varepsilon_0(\beta) \to 0$ as $\beta \to 1$ so that the function $h_{\boldsymbol{\omega}}$ cannot be shown to be Lipschitz continuous for ε small (and in general it is not). The map H_ε is a Hölder continuous map of ^{2} at fixed $\boldsymbol{\omega}$ and we have shown that (3.2) is an identity between the Taylor coefficients of the two sides which are also holomorphic for small $|\varepsilon|$: hence (3.2) holds. The hyperbolicity of f and (3.2) imply that $H_\varepsilon(\boldsymbol{\omega}, x) = H_\varepsilon(\boldsymbol{\omega}, x')$ if and only if $x = x'$ since in $\mathcal{F}_0$ the evolution of the noise $\boldsymbol{\omega}$ and of the point x are independent, so that $H_\varepsilon(\boldsymbol{\omega}, \cdot)$ is a homeomorphism for each $\boldsymbol{\omega}$ and therefore H_ε is also a homeomorphism as its action on $\boldsymbol{\omega}$ is trivially the identity.

Finally we discuss what changes in the above discussion if $\psi_{\boldsymbol{\omega}}(x)$ is analytic in x. In that case one has $|\psi_{\boldsymbol{\omega},\alpha,\nu}| \le \overline{\Psi}\, e^{-\kappa|\nu|}$, for some $\boldsymbol{\omega}$–independent constants $\overline{\Psi}$ and κ. We have still to study the series $\sum_\nu |\nu|^\beta |h^{(k)}_{\boldsymbol{\omega},\alpha,\nu}|$. For $\beta = 0$ we can perform the sum over the Fourier labels by using that for each node v there is a factor $e^{-\kappa|\nu'_v|}\overline{\Psi}$ times a factor $|\nu'_v|^{s_v}/s_v!$, with $\nu'_v \stackrel{\text{def}}{=} f^{-p(v)}\nu_v$. The sums over the ν_v's can be bounded by a factor $D_\kappa^{s_v}$ per node, with $D_\kappa = 2/\kappa$ (this leaves a factor $e^{-\kappa|\nu_v|/2}$ per node that in this case we neglect), hence they produce an overall factor D_κ^k, because $\sum_{v\in V(\theta)} s_v = k-1$. The bound on the number of trees proceeds as before, and gives 2^{2k}. If $\beta \ne 0$, by reasoning as in the case of trigonometric polynomials, we obtain an extra factor $k|\nu_{\overline{v}}|^\beta \le k\,\lambda^{-|p(\overline{v})|\beta} B|\nu'_{\overline{v}}|^\beta$. The factor $|\nu'_{\overline{v}}|^\beta$ can be bounded together with the remaining factor $e^{-\kappa|\nu'_{\overline{v}}|/2}$ by $\widetilde{D} \stackrel{\text{def}}{=} \sum_{\nu\in\mathbb{Z}} e^{-\kappa|\nu|/2}|\nu|^\beta$,while the extra k is harmless for the convergence analysis. Since the factor $\lambda^{-|p(\overline{v})|\beta}$ can be dealt with exactly as in the trigonometric polynomial case, the end result is simply convergence for $|\varepsilon|$ smaller than $\varepsilon_0(\beta) = (D_\kappa \overline{\Psi})^{-1}(1-\lambda^{1-\beta})2^{-4}$, which replaces (3.16). ∎

Remark. The result can also be interpreted as a "shadowing theorem". Consider the "noisy" trajectory $k \to \mathcal{F}^{*k}(\boldsymbol{\omega}, x)$ as a perturbation of a noiseless one. Then a noisy trajectory starting at $(\boldsymbol{\omega}, x)$ will remain forever close to the trajectory of the point $H_\varepsilon^{-1}(\boldsymbol{\omega}, x)$ *evolving under the noiseless motion.* This is usually called a *shadowing property.*

4. Heuristic considerations

In this section we develop the heuristic basis on which Lemma 2 in §5 and Lemma 3 in §6 are developed.

The notion of hyperbolicity or chaoticity can be extended, [16], to random systems, for a general map $\mathcal{F}$ acting on $\Omega \times \mathcal{M}$, as follows:

(a) at every point $\xi = (\boldsymbol{\omega}, x) \in \Omega \times \mathcal{M}$ the plane $W(x)$ tangent to $\mathcal{M}$ contains two planes $W^s_{\boldsymbol{\omega}}(x), W^u_{\boldsymbol{\omega}}(x)$, called respectively the contracting and expanding planes, with $W(x) = W^s_{\boldsymbol{\omega}}(x) \oplus W^u_{\boldsymbol{\omega}}(x)$, with positive dimensions d_s, d_u and which are *covariant*, *i.e.* $\mathcal{F} W^a_{\boldsymbol{\omega}}(x) = W^a_{\tau\boldsymbol{\omega}}(f_{\boldsymbol{\omega}}(x))$;

(b) for some $C,\ \rho > 0$

$$\begin{aligned} \|\partial_x f^{*k}_{\boldsymbol{\omega}}(x) v\| &\le C e^{-k\rho} \|v\|, \qquad v \in W^s_{\boldsymbol{\omega}}(x), \\ \|\partial_x f^{*(-k)}_{\boldsymbol{\omega}}(x) v\| &\le C e^{-k\rho} \|v\|, \qquad v \in W^u_{\boldsymbol{\omega}}(x), \end{aligned} \tag{4.1}$$

for all $k > 0$; $f^{*k}_{\boldsymbol{\omega}}$ is defined after (1.4) and the sizes $\|\cdot\|$ of the vectors are evaluated in the metrics at the points to which they are applied (*i.e.* $f^{*k}_{\boldsymbol{\omega}}(x)$ or x or $f^{*(-k)}_{\boldsymbol{\omega}}(x)$);

(c) $W^a_{\boldsymbol{\omega}}(x)$ are continuous in $\boldsymbol{\omega}, x$.

In the example (1.3) with $\varepsilon = 0$ one has $\partial_x(f^{*(\pm k)}_{\boldsymbol{\omega}}) = f^{\pm k}$ so that properties (a,b,c) are satisfied, trivially, because the map considered is an Anosov map.

It is also natural to consider the logarithms of the Jacobian determinants at time N along the unstable and stable manifolds of $\mathcal{F}$, as a measure of the expansion and contraction under the action of $\mathcal{F}$. They will simply be

$$J^a_{\boldsymbol{\omega},N}(x) = \log \left| \det \Big(\frac{\partial f^{*(a\,N)}_{\boldsymbol{\omega}}(x)}{\partial x} \Big)_a \right|, \qquad a = u, s, \tag{4.2}$$

where $(a\,N)$ means N if $a = u$ and $-N$ if $a = s$ while the label a appended to the Jacobian matrix means that it is regarded as acting as a map from the plane $W^a_{\boldsymbol{\omega}}(x)$ to the plane $W^a_{\tau^N\boldsymbol{\omega}}(f^{*N}_{\boldsymbol{\omega}}(x))$.

The above Jacobian determinants are *not* intrinsic geometric objects: as well as the constants $C,\ \rho$ they depend on the coordinates used on the manifold.

The volume measure $\mu_0(\mathrm{d}x)$ on $\mathcal{M}$ can be visualized, fixed $\boldsymbol{\omega} \in \Omega$, in terms of the above expansions and contractions: imagine to fix, for every point $x \in \mathcal{M}$, two "surface elements" located with center in x and lying on the unstable or on the stable planes through x: the surface elements will all have the same infinitesimal surface and we call them $\Delta^u_{\boldsymbol{\omega}}(x)$ and $\Delta^s_{\boldsymbol{\omega}}(x)$, respectively. Then we can build a tiny "parallelepiped" through x by considering the surface elements through x

$$\delta_u(N,x) = f^{*(-N)}_{\tau^N\boldsymbol{\omega}}(\Delta^u_{\tau^N\boldsymbol{\omega}}(f^{*N}_{\boldsymbol{\omega}}(x))), \qquad \delta_s(N,x) = f^{*N}_{\tau^{-N}\boldsymbol{\omega}}(\Delta^s_{\tau^{-N}\boldsymbol{\omega}}(f^{*(-N)}_{\boldsymbol{\omega}}(x))), \tag{4.3}$$

and by drawing through each point $y \in \delta_s(N,x)$ the surface element $\delta_u(N,y)$ and through each point of $y \in \delta_s(N,x)$ the surface element $\delta_u(N,y)$ and collecting the intersections. This is a well defined set, and will be called $\delta_{\boldsymbol\omega,N}(x)$. It is a set close to a parallelepiped with volume exponentially small as $N \to \infty$ and, to leading order in N, given by

$$\mathrm{vol}\,\delta_{\boldsymbol\omega,N}(x) = \mathrm{e}^{-J^u_{\boldsymbol\omega,N}(x)}\,\mathrm{e}^{J^s_{\boldsymbol\omega,N}(x)}\gamma_{\boldsymbol\omega,N}(x), \tag{4.4}$$

with

$$\gamma_{\boldsymbol\omega,N}(x) \stackrel{\text{def}}{=} \sin\alpha_{\boldsymbol\omega}(x) \prod_{a=u,s} |\Delta^a_{\tau^{(a\,N)}\boldsymbol\omega}(f^{*(a\,N)}_{\boldsymbol\omega}(x))|, \tag{4.5}$$

where $\alpha_{\boldsymbol\omega}(x)$ equals the angle between the stable and unstable planes $W^a_{\boldsymbol\omega}(x)$ at x, the symbol $|\Delta|$ denotes the area of the surface element Δ, and $(a\,N)$ has the meaning explained after (4.2). The function $\gamma_{\boldsymbol\omega,N}(x)$ is uniformly bounded away from 0 and ∞. Hence the ratio of the volumes of two such parallelepipeds, centered at x and y, can be computed by the ratio between the contraction factors in (4.3) evaluated at the two points (because the basic surface elements $\Delta^a_{\boldsymbol\omega}(x)$ have all the same size).

A consequence is that, if we attribute to each such parallelepiped a measure proportional to

$$e^{-J^u_{\boldsymbol\omega,2N}(f^{*(-N)}_{\boldsymbol\omega}(x))}, \tag{4.6}$$

then we expect that in the limit as $N \to \infty$ the probability distribution has a limit $\mu_{\boldsymbol\omega}(\mathrm{d}x)$ such that $P(\mathrm{d}\boldsymbol\omega)\mu_{\boldsymbol\omega}(\mathrm{d}x) \stackrel{\text{def}}{=} \mu(\mathrm{d}\boldsymbol\omega\mathrm{d}x)$ is an invariant distribution for the system $(\Omega \times \mathcal{M}, \mathcal{F})$ which should describe the statistical properties of almost all data initially chosen with the distribution $P(\mathrm{d}\boldsymbol\omega)\mu_0(\mathrm{d}x)$: *i.e.* $\mu(\mathrm{d}\boldsymbol\omega\mathrm{d}x)$ *should be the SRB distribution.* Likewise replacing in (4.6) N with $-N$ and J^u with $-J^s$ one should obtain the SRB distribution for the backwards motion (*i.e.* for the map $\mathcal{F}^{-1}$).

5. Overshadowing

Therefore we look for an explicit algorithm to construct a useful representation of (4.6) or, what is the same, for the function $J_{\boldsymbol\omega,1}(x)$ because

$$J^u_{\boldsymbol\omega,2N}(f^{*(-N)}_{\boldsymbol\omega}(x)) = \sum_{j=-N}^{N} J^u_{\boldsymbol\omega,1}(f^{*j}_{\boldsymbol\omega}(x)), \tag{5.1}$$

by the chain differentiation rule for composition of functions and by the multiplication rules of determinants of matrices. Note that $J^u_{\boldsymbol\omega,1}$ is the logarithm of the Jacobian determinant at time 1 – or simply Jacobian determinant – along the unstable manifold. We shall use the notations of §3, where the conjugation H_ε, transforming the perturbed map $\mathcal{F}(\boldsymbol{\omega}, x) = (\tau\boldsymbol{\omega}, fx + \varepsilon\psi_{\boldsymbol{\omega}}(x))$ of the torus $\Omega \times {}^m$ into a noiseless map $\mathcal{F}_0(\boldsymbol{\omega}, x) = (\tau\boldsymbol{\omega}, fx)$, has been derived. The homeomorphism $(\boldsymbol{\omega}', x') \longleftrightarrow (\boldsymbol{\omega}, x + h_{\boldsymbol{\omega}}(x)) \equiv H_\varepsilon(\boldsymbol{\omega}, x)$ can be used not only to construct the dynamics but also the stable and unstable manifolds of each point. The latter manifolds through $(\boldsymbol{\omega}, x + h_{\boldsymbol{\omega}}(x))$ are given by parametric equations of the form

$$\gamma_\alpha(t) = (\boldsymbol{\omega}, h_{\boldsymbol{\omega}}(x + tv_\alpha)) \qquad t \in , \quad \alpha = \pm, \tag{5.2}$$

where $t \to x + t\,v_\alpha$ is symbolically a parametrization of the unstable or stable manifold for the unperturbed map f of ^{2} into itself, respectively if $\alpha = +$ or $\alpha = -$. The actual meaning is $tv_\alpha \stackrel{\text{def}}{=} \sum_{i=1}^{d_\alpha} t_i v_\alpha^{(i)}$ if $v_\alpha^{(i)}$ is a base on the unstable or stable plane, of dimension d_α, tangent to the torus at x, and $t = (1, \ldots, t_{d_\alpha}) \in {}^{d_\alpha}$.

The construction of §3 gives a Hölder continuous conjugation with a prefixed exponent $\beta < 1$ for a perturbation strength ε that is suitably small, depending on how close β is to 1. However, in general, the conjugation is not differentiable: therefore (5.2) cannot be used to compute the derivatives appearing in the Jacobians $J^a_{\boldsymbol{\omega},1}(x)$ and the simple parametrization (5.2) is not very useful.

Instead of constructing the stable an unstable manifolds parameterized so that the required derivatives appearing in (5.1) can be computed (or just shown to exist) we remark that all we need are the expansion coefficients in the stable and unstable directions of any point $(\boldsymbol{\omega}, x)$. For this purpose it is useful to introduce a few more definitions.

Definition 2. *Let $\widehat{\Theta}$ be the (non–compact) space $\Omega \times {}^m \times {}^m$ and define on it the dynamical system $(\widehat{\Theta}, \widehat{\mathcal{F}}_0)$ with*

$$\widehat{\mathcal{F}}_0(\boldsymbol{\omega}, x, v) = (\tau\boldsymbol{\omega}, fx, fv), \tag{5.3}$$

where v is a tangent vector to ^{m} in x and fv denotes the action of the derivative of the map $x \to fx$ on the tangent vector v.
(ii) A "simple" perturbation of the map $\widehat{\mathcal{F}}_0$ is any map on $\Omega \times {}^m \times {}^m$ defined by

$$\widehat{\mathcal{F}}(\boldsymbol{\omega}, x, v) = (\tau\boldsymbol{\omega}, fx + \varepsilon\psi_{\boldsymbol{\omega}}(x), fv + \varepsilon(v \cdot \partial_x)\psi_{\boldsymbol{\omega}}(x)), \tag{5.4}$$

where the third term specifies how the tangent vector v at x is tranformed by the map $x' = fx + \varepsilon\psi_{\boldsymbol{\omega}}(x)$ into a tangent vector at x'.

Remarks. (1) Since the map f is locally linear and constant there is no point in distinguishing between the constant matrix $\partial_x fx$ and the map f so that we indulge, here and below, in the abuse of notation implicit in (5.3) and (5.4).
(2) The dynamical system $(\widehat{\Theta}, \widehat{\mathcal{F}}_0)$ fails to be an Anosov system not only because of the noise (which acts trivially, however) but also because the space $\Omega \times {}^m \times {}^m$ is not compact.
(3) Nevertheless we can still try to find an isomorphism between the two systems. Success would mean the possibility of decoupling from the noise *also* the evolution of the tangent vectors, *i.e.* of the infinitesimal displacements: not only we could match individual trajectories in the perturbed and in the unperturbed system, as done in §3, but we could even achieve "shadowing" of infinitesimally close pairs of trajectories which would split apart at the same rates as the unperturbed ones. Not surprisingly this turns out to be in general *impossible* (as it will be implicit in what follows).

If the property in the last remark is impossible the notion can be usefully weakend and one is led, in trying to compare infinitesimally close pairs of trajectories, to the following definition of "overshadowing".

Definition 3. *(i) The map $\overline{\mathcal{F}}_0$ (acting on the same space $\Omega \times {}^m \times {}^m$ but different from the above $\widehat{\mathcal{F}}_0$), given by*

$$\overline{\mathcal{F}}_0(\boldsymbol{\omega}, x, v) = (\tau\boldsymbol{\omega}, fx, fv + \Gamma_{\boldsymbol{\omega}}(x)v), \tag{5.5}$$

where $\Gamma_{\boldsymbol{\omega}}(x)$ is a matrix which leaves invariant the expanding and contracting planes of f and is Hölder continuous in $(\boldsymbol{\omega}, x)$, is called "similar" to $\widehat{\mathcal{F}}_0$.
(ii) The map $\widehat{\mathcal{F}}$ of $\Omega \times {}^m \times {}^m$ is "overshadowed" by $\overline{\mathcal{F}}_0$ if there exists a transformation

$$\widehat{H}_\varepsilon \,:\, (\boldsymbol{\omega}, x, w) \longleftrightarrow ((\boldsymbol{\omega}, x + h_{\boldsymbol{\omega}}(x)), w + K_{\boldsymbol{\omega}}(x)w), \tag{5.6}$$

with $K_{\boldsymbol{\omega}}(x)$ Hölder continuous, which satisfies the conjugation equation $\widehat{\mathcal{F}} \circ \widehat{H}_\varepsilon = \widehat{H}_\varepsilon \circ \overline{\mathcal{F}}_0$.

In guessing (5.6), advantage is taken of the already known conjugation H_ε between $\mathcal{F}$ and $\mathcal{F}_0$, from the analysis of §3, giving us the function $h_{\boldsymbol{\omega}}(x)$.

Remarks. (1) Let $\mathcal{L}_{\boldsymbol{\omega}}(x) = f + \Gamma_{\boldsymbol{\omega}}(x)$ be functions with values in the $m \times m$ matrices. Then the matrix $\Gamma_{\boldsymbol{\omega}}(x)$ has to consist of two blocks of respective dimension $d_u \times d_u$ and $d_s \times d_s$, if d_u and d_s are the respective dimensions of the stable and unstable planes of f. Then the conjugation equation after (5.6) is equivalent to the following equation:

$$\partial_x f_{\boldsymbol{\omega}}(H_\varepsilon(\boldsymbol{\omega},x))\mathcal{K}_{\boldsymbol{\omega}}(x)v = \mathcal{K}_{\tau\boldsymbol{\omega}}(fx)\mathcal{L}_{\boldsymbol{\omega}}(x)v \tag{5.7}$$

and to (3.3) for H_ε, which is an equation independent of (5.7) (already solved in §3). Here we have set $\mathcal{K}_{\boldsymbol{\omega}}(x) = 1 + K_{\boldsymbol{\omega}}(x)$.

(2) Let $v_+^{(i)}$, $i = 1, \ldots, d_u$ be a basis in the d_u-dimensional expanding plane for f and let $v_-^{(i)}$ be a basis in the d_s-dimensional contracting plane of f. Then the vectors $w_{\boldsymbol{\omega},\pm}^{(i)}(x) = \mathcal{K}_{\boldsymbol{\omega}}(x)v_\pm^{(i)}$ satisfy

$$\partial_x f_{\boldsymbol{\omega}}(H_{\boldsymbol{\omega}}(x))w_{\boldsymbol{\omega},\pm}^{(i)}(x) = (\lambda_{\boldsymbol{\omega},\pm}(x))_{i,s}w_{\tau\boldsymbol{\omega},\pm}^{(s)}(fx), \tag{5.8}$$

where $\lambda_{\boldsymbol{\omega},\pm}(x)$ are the transposed of the mentioned blocks of $\mathcal{L}_{\boldsymbol{\omega}}(x)$ regarded as matrices on the bases selected in the expanding and contracting planes of f, and repeated indices mean implicit summation (to abridge notations).

(3) The conjugation defined in (5.6) is the right one to look at for the stable and unstable directions. Its solution gives the stable and unstable directions as functions of the unperturbed point $H_\varepsilon^{-1}(\boldsymbol{\omega},x)$. Note that (5.6) is not a conjugation of $\widehat{\mathcal{F}}$ to $\widehat{\mathcal{F}}_0$, as it conjugates $\widehat{\mathcal{F}}$ to $\overline{\mathcal{F}}_0$.

(4) Equation (5.7) does not determine $\mathcal{K}_{\boldsymbol{\omega}}(x)$ uniquely. Indeed, if $l_{\boldsymbol{\omega},\pm}(x)$ are two non–zero matrices of respective dimensions $d_u \times d_u$ and $d_s \times d_s$, then if $\lambda_{\boldsymbol{\omega},\pm}(x)$, $w_{\boldsymbol{\omega},\pm}^{(i)}(x)$ solve (5.8), also

$$\overline{\lambda}_{\boldsymbol{\omega},\pm}(x) = l_{\boldsymbol{\omega},\pm}(x)\,\lambda_{\boldsymbol{\omega},\pm}(x)\,l_{\tau\boldsymbol{\omega},\pm}(fx)^{-1}, \qquad \overline{w}_{\boldsymbol{\omega},\pm}^{(i)}(x) = (l_{\boldsymbol{\omega},\pm}(x))_{i,s}w_{\boldsymbol{\omega},\pm}^{(s)}(x) \tag{5.9}$$

solve it. To fix this ambiguity *we will require* that the diagonal blocks of $K_{\boldsymbol{\omega}}(x)$, on both bases $v_+^{(i)}$ and $v_-^{(i)}$ are equal to 0, *i.e.* the matrix $K_{\boldsymbol{\omega}}(x)$ is completely off–diagonal with respect to the expanding and contracting planes of f.

(5) Therefore the dimensions of $\lambda_{\boldsymbol{\omega},+}(x)$ and $\lambda_{\boldsymbol{\omega},-}(x)$ are, repectively, $d_u \times d_u$ and $d_s \times d_s$, while the matrices $K_{\boldsymbol{\omega}}(x)$ consist of off-diagonal blocks of dimension $d_u \times d_s$ and $d_s \times d_u$.

LEMMA 2. *(i) Under the assumptions of Theorem 1 (i), there exist two Hölder continuous matrices $\Gamma_{\boldsymbol{\omega}}(x), K_{\boldsymbol{\omega}}(x)$, analytic in ε for ε small enough, such that the map $\widehat{\mathcal{F}}$ in (5.4) is overshadowed by a map $\overline{\mathcal{F}}_0$ similar to $\widehat{\mathcal{F}}_0$ with a transformation of the form (5.6).*

(ii) The matrices $\Gamma_{\boldsymbol{\omega}}(x)$ can be taken diagonal with respect to the decomposition of ^{m} into expanding and contracting planes for f, while the matrices $K_{\boldsymbol{\omega}}(x)$ can be taken off-diagonal. Therefore the first consist of two blocks $\lambda_{\boldsymbol{\omega},\pm}(x)$ of dimensions $d_u \times d_u$ and $d_s \times d_s$, if d_u and

d_s are the dimensions of the stable and unstable planes for f ($d_u + d_s \equiv m$), and the second is a matrix with two blocks $k_{\boldsymbol{\omega},\pm}(x)$ of dimensions $d_u \times d_s$ and $d_s \times d_u$.
(iii) The Jacobian determinant $\exp J^u_{\boldsymbol{\omega},1}(x)$, along the unstable manifold of $\mathcal{F}$, is $\det \lambda_{\boldsymbol{\omega},+}(x)$ and, therefore, it is analytic in ε and Hölder continuous in $(\boldsymbol{\omega}, x)$.

Proof. Also in this case we give explicitly the proof in the case of $m = 2$, f given by (1.1) and $\psi_{\boldsymbol{\omega}}$ a trigonometric polynomial (adaptation to the general case proceeds as in §3). This will make the matrices $\lambda_{\boldsymbol{\omega},\pm}(x)$ simply scalars and $K_{\boldsymbol{\omega}}(x)$ will be a 2×2 matrix with off-diagonal entries $k_{\boldsymbol{\omega},\pm}(x)$ only, because in this case the stable and unstable planes are one–dimensional. To simplify the rather heavy notations we denote

$$\begin{aligned} &\psi_\alpha(\boldsymbol{\omega}, x) \equiv \psi_{\boldsymbol{\omega},\alpha}(x), \quad H(\boldsymbol{\omega}, x) \equiv (\boldsymbol{\omega}, x + h_{\boldsymbol{\omega}}(x)), \\ &\Gamma(\boldsymbol{\omega}, x) \equiv \Gamma_{\boldsymbol{\omega}}(x), \quad K(\boldsymbol{\omega}, x) \equiv K_{\boldsymbol{\omega}}(x), \quad \xi = (\boldsymbol{\omega}, x), \end{aligned} \tag{5.10}$$

and in the following we will not write explicitly the subscripts $\boldsymbol{\omega}$. The equation that the matrices $K(\xi), \Gamma(\xi)$ have to satisfy is a transcription of (5.7)

$$\begin{aligned} &(fK(\xi) - K(\mathcal{F}_0\xi)f)_{ij} = \\ &= -\varepsilon\partial_{x_j}\psi_i(H(\xi)) - \varepsilon\partial_{x_k}\psi_i(H(\xi))K(\xi)_{kj} + \Gamma(\xi)_{ij} + (K(\mathcal{F}\xi)\Gamma(\xi))_{ij}, \end{aligned} \tag{5.11}$$

where ∂_x denotes a derivative of ψ with respect to its argument and repeated indices mean implicit summation.

We write the above matrix equation on the basis in which f and Γ are diagonal, *i.e.* on the basis formed by the two eigenvectors $v_\pm$ of f in which the matrices K and Γ have been assumed to take the form

$$\Gamma(\xi) = \begin{pmatrix} \gamma_+(\xi) & 0 \\ 0 & \gamma_-(\xi) \end{pmatrix}, \qquad K(\xi) = \begin{pmatrix} 0 & k_+(\xi) \\ k_-(\xi) & 0 \end{pmatrix}. \tag{5.12}$$

If $\alpha = \pm$, $\beta = -\alpha$ and $\partial_\alpha \stackrel{\text{def}}{=} v_\alpha \cdot \partial_x$, (5.11) becomes for $i = j$

$$0 = -\varepsilon\partial_\alpha\psi_\alpha(H(\xi)) - \varepsilon K_{\beta,\alpha}(\xi)\partial_\beta\psi_\alpha(H(\xi)) + \gamma_\alpha(\xi), \tag{5.13}$$

and, for $i \neq j$ and if $\lambda_+ = \lambda^{-1}, \lambda_- = -\lambda$ are the eigenvalues of f (with $\lambda = (\sqrt{5}-1)/2$),

$$\begin{aligned} &\lambda_\alpha K_{\alpha,\beta}(\xi) - \lambda_\beta K_{\alpha,\beta}(\mathcal{F}_0\xi) = \\ &\quad - \varepsilon\partial_\beta\psi_\alpha(H(\xi)) - \varepsilon K_{\alpha,\beta}(\xi)\partial_\alpha\psi_\alpha(H(x)) + K_{\alpha,\beta}(\mathcal{F}_0\xi)\gamma_\beta(\xi). \end{aligned} \tag{5.14}$$

If $\alpha' = \frac{\alpha-1}{2}$, and $\alpha'' = \frac{\alpha+1}{2}$ we can rewrite the equations (5.13) and (5.14) as

$$\gamma_\alpha(\xi) = \varepsilon\, \partial_\alpha \psi_\alpha(H(\xi)) + \varepsilon\, k_\beta(\xi) \partial_\beta \psi_\alpha(H(\xi)), \tag{5.15}$$

and, respectively,

$$\begin{aligned} k_\alpha(\xi) + \lambda^2 k_\alpha(\mathcal{F}_0^\alpha \xi) = \alpha\lambda\, (- \varepsilon \partial_\beta \psi_\alpha(H(\mathcal{F}_0^{\alpha'} \xi)) - \\ - \varepsilon k_\alpha(\mathcal{F}_0^{\alpha'} \xi) \partial_\alpha \psi_\alpha(H(\mathcal{F}_0^{\alpha'} \xi)) + k_\alpha(\mathcal{F}_0^{\alpha''} \xi) \gamma_\beta(\mathcal{F}_0^{\alpha'} \xi)). \end{aligned} \tag{5.16}$$

These equations are in a form suitable for a recursive solution in powers of ε. For instance the first order is

$$\begin{aligned} &\gamma_\alpha^{(1)}(\xi) = \partial_\alpha \psi_\alpha(\xi), \qquad \alpha = \pm, \\ &k_+^{(1)}(\xi) + \lambda^2 k_+^{(1)}(\mathcal{F}_0 \xi) = -\lambda\, \partial_- \psi_+(\xi), \qquad k_-^{(1)}(\xi) + \lambda^2 k_-^{(1)}({\mathcal{F}_0}^{-1} \xi) = \lambda\, \partial_+ \psi_-({\mathcal{F}_0}^{-1} \xi), \end{aligned} \tag{5.17}$$

which has the solution

$$\begin{aligned} &\gamma_\alpha^{(1)}(\xi) = \partial_\alpha \psi_\alpha(\xi) \qquad \alpha = \pm \\ &k_+^{(1)}(\xi) = -\lambda \sum_{q=0}^{\infty} (-1)^q \lambda^{2q} \partial_- \psi_+(\mathcal{F}_0^q \xi), \qquad k_-^{(1)}(\xi) = \lambda \sum_{q=0}^{\infty} (-1)^q \lambda^{2q} \partial_+ \psi_-({\mathcal{F}_0}^{-(q+1)} \xi). \end{aligned} \tag{5.18}$$

The equations for $\gamma_\alpha^{(k)}(\xi)$ and $k_\alpha^{(k)}(\xi)$ can be represented in graph form by suitably modifying the similar representation derived for $h_{\boldsymbol{\omega},\alpha}^{(k)}(x)$ in §3. Figure 2 below corresponds to Figure 1 above and translates into a graphical representation (5.15) and (5.16). This time new symbols have to be introduced, but in the end one has just to check that all the graph elements contain enough labels to identify terms appearing in (5.15) and (5.16).

We denote $\overset{\alpha}{\longrightarrow}\!\circ_{(1)}$ the first line of (5.18) and $\overset{\alpha}{\longrightarrow}\!\square_{(1)}$ the second line of (5.18) and for $k > 1$ we have the graphical representation of Figure 2 – the reader will recognize in the first line a pictorial rewriting of (5.15) and in the second a rewriting of (5.16), – where

(a) $\alpha = \pm$ and $\beta = -\alpha$,

(b) all the lines have to imagined to carry arrows (not drawn) pointing toward the root,

(c) the line carrying a label α and emerging from a circle or a square with label (k) denotes $\gamma_\alpha^{(k)}$ or $k_\alpha^{(k)}$, respectively.

(d) the wavy line emerging from a bullet with label (p) and ending in a small circle and carrying a pair of labels γ, δ, represents $[\partial_\gamma \psi_\delta \circ H]^{(p)}$, the p-th order in the power expansion in ε of $(\partial_\gamma \psi_\delta) \circ H$,

(e) the small square in the node closest to the root into which a wavy line arrives will carry a label $q = 0, 1, 2 \dots$ (see the graphs in the second line) and it expresses that the functions $(\partial_\gamma \psi_\delta) \circ H$, k_α and γ_β must be computed at $\mathcal{F}_0^{q_\alpha} \xi$, with $q_\alpha = \alpha' + \alpha q$ in all cases except for the argument of k_α in the third graph, where $q_\alpha = \alpha'' + \alpha q$,

(f) a summation over $q = 0, 1, \dots$, $p = 1, \dots, k-1$ and a multiplication by $-\alpha(-1)^q \lambda^{1+2q}$ is understood to be performed over the nodes represented as small squares, while a trivial factor 1 is associated with the nodes represented as small circles.

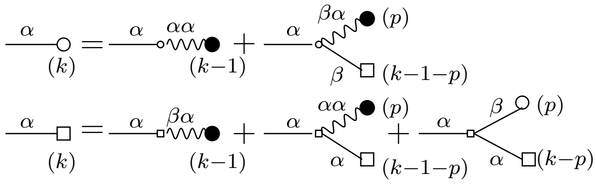

FIGURE 2. Graphical interpretation of (5.15) and (5.16) for $k > 1$.

In this case too we can continue the expansion until in the r.h.s. of Figure 2 all top nodes of the graph are either squares or circles carrying a label (1), *i.e.* they represent a first order contribution to Γ (circle) or to K (square), or bullets representing $\partial_\alpha \psi_\beta$. With each wavy line ℓ we associate two labels $\alpha'_\ell, \alpha_\ell$ and with each non-wavy line we assoiciate only one label α_ℓ. With each node represented as a small square we associate two labels $\alpha_v = \alpha_{\ell_v}$ and $q_v \ge 0$, and we define q_{α_v} as in item (e) above, while with each node represented as a small circle we associate only the label $\alpha_v = \alpha_{\ell_v}$ and we set $q_{\alpha_v} = 0$. There are constraints on the possible choices of the labels $\alpha'_\ell, \alpha_\ell, \alpha_v$, which can be easily deduced from (5.15) and (5.16) – or equivalently from Figure 2. With each top node we associate a label $\alpha_v = \alpha_{\ell_v}$ if the node is a either a circle or a square, and a pair of labels $\alpha'_v = \alpha'_{\ell_v}$ and $\alpha_v = \alpha_{\ell_v}$ is the node is a bullet. Finally for each node represented as a circle of a square we define $q(v) \stackrel{\rm def}{=} \sum_{w \preceq v} q_{\alpha_v}$: then each top node v represents either $k^{(1)}_{\alpha_v}(\mathcal{F}_0^{q(v)}\xi)$ or $\gamma^{(1)}_{\alpha_v}(\mathcal{F}_0^{q(v)}\xi)$ or $\left[\partial_{\alpha'_v} \psi_{\alpha_v}(H(\mathcal{F}_0^{q(v)}\xi))\right]^{(p)}$.

Of course the latter quantities can themselves be represented by the tree expansion discussed in §3: if we do so then we obtain a full expansion in powers of ε in which the wavy lines with label p are replaced by a tree with p nodes.

The rule to construct the value of each tree graph is easily read from (5.11) and from the rules discussed above, and in §3, to build the value of trees representing $h_{\boldsymbol{\omega}}$. It is again a product of factors associated with the graph elements. Therefore the estimates are immediate: one just has to imagine the trees developed until the endlines represent first order contributions and count how many labels pertain to each node, repeating the procedure described in the case of §3.

The estimate of the k–th order contribution is given by (3.15) with an extra factor N^k to take into account the extra derivatives due to the (wavy) lines with two labels. Also the counting of the trees has to be modified, see Appendix A2, but the end result will be that Γ and K are expressed by convergent series in ε for $|\varepsilon| < \varepsilon_0(\beta)$, where $\varepsilon_0(\beta)$ can be taken of the form (3.16) with a different numerical factor and with N replaced by N^2.

Hölder continuity of exponent $\beta < 1$ will hold for $|\varepsilon| < \varepsilon_0(\beta)$ for all the functions and property (iii) is also implicit in the construction.

The case of higher dimensional toral maps f is not different. Again a few more labels have to be introduced. Everything remains essentially the same as soon as one becomes convinced that the number of "free parameters", which are the $2d_ud_s$ matrix elements of K, are just enough "free parameters" to make equations (5.15) and (5.16) for Γ and K solvable. The values of $\Gamma^{(k)}, K^{(k)}$ should be regarded as necessary parameters to solve equations(5.15) and (5.16) which, otherwise, if we tried to set $\Gamma^{(k)}, K^{(k)} = 0$ would not be solvable because the r.h.s. of such relations does not necessarily vanish. Replacing $\lambda_\pm$ with matrices does not lead to problems because the matrices are independent of $(\boldsymbol{\omega}, x)$ and in the estimates the negative powers of λ_+ and the positive ones of λ_- will still be exponentially decreasing. ■

6. Construction of the SRB distribution

Consider the partition $\mathcal{E}_0$ of $\Omega \times \mathcal{M}$ by the sets $E_{\omega,\sigma} = C_\omega \times P_\sigma$ where $\mathcal{C} = \{C_1, \ldots, C_m\}$ is the partition of Ω according to the value of the symbol ω_0 and $\mathcal{P}_0 = \{P_1, \ldots, P_n\}$ is a Markovian pavement of $\mathcal{M}$ for the map unperturbed map f [8]. Given a point $\xi_0 \in \Omega \times \mathcal{M}$ let $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ be its "history" on the partition $\mathcal{E}_0$ for the map $\mathcal{F}_0$. This means that $\mathcal{F}_0^j \xi_0 \in E_{\omega_j,\sigma_j}$ for all $j \in$. In terms of the sequences $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ the motion is just a translation τ such that $\tau(\boldsymbol{\omega}, \boldsymbol{\sigma})_j = (\boldsymbol{\omega}, \boldsymbol{\sigma})_{j+1}$, i.e. we denote by τ both the translation of $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and its restrictions to $\boldsymbol{\omega}$ or $\boldsymbol{\sigma}$.

The partition $\mathcal{E} = H_\varepsilon(\mathcal{E}_0)$ will be called "Markovian for the perturbed system" because it generates a symbolic dynamics controlled by a transitive compatibility matrix M' (with entries $M'_{\omega\sigma,\omega'\sigma'} \equiv M_{\sigma,\sigma'}$, see Appendix A1 for definitions). A correspondence can be generated between points in $\Omega \times \mathcal{M}$ by associating two points ξ and ξ_0 if the first has history $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ on $\mathcal{E}$ under the action of $\mathcal{F}$ identical to the history of ξ_0 on the partition $\mathcal{E}_0$ under the action of the map $\mathcal{F}_0$: we denote $\xi = X_\varepsilon(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and $\xi_0 = X_0(\boldsymbol{\omega}, \boldsymbol{\sigma})$. The correspondence so defined is identical to the correspondence between ξ and ξ_0 defined by $\xi = H_\varepsilon(\xi_0)$. In §5 it has been proved (cf. Lemma 2) that there exists $\varepsilon_0(\beta)$ such that the Jacobian determinant $\exp J^u_{\boldsymbol{\omega},1}(x) = \det \lambda_{\boldsymbol{\omega},+}(x)$ is defined and holomorphic in ε in the disk $|\varepsilon| < \varepsilon_0(\beta)$. As a function of $(\boldsymbol{\omega}, x)$ it is Hölder continuous. We can write $J^u_{\boldsymbol{\omega},1}(x)$ in terms of the history $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ by setting $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) \equiv J^u_{\boldsymbol{\omega},1}(x)$, with $(\boldsymbol{\omega}, x) = X_\varepsilon(\boldsymbol{\omega}, \boldsymbol{\sigma})$. We define the *expansion rate* $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ as the logarithm of the ratio of the image of a volume element on the unstable manifold at a point $\xi = X_\varepsilon(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and the volume itself: this is a quantity depending on the local coordinates used, like $J^u_{\boldsymbol{\omega},1}$, *as well as on the metric used.* The following result shows that it has a simple relation with $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$.

LEMMA 3. *The function $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ is Hölder continuous with some exponent $0 < \beta < 1$ and modulus $C(\beta)$. Its difference from $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ has the form $g(\boldsymbol{\omega}, \boldsymbol{\sigma}) - g(\tau(\boldsymbol{\omega}, \boldsymbol{\sigma}))$ with g a Hölder continuous function.*

Proof. We again discuss the case $m = 2$ and f given by (1.1) and in this case the Jacobian matrix restricted to the unstable manifold is simply a scalar $\lambda_{\boldsymbol{\omega},+}(x) \equiv \lambda_+ + \gamma_{\boldsymbol{\omega},+}(x)$. Remark that if $(\boldsymbol{\omega}, x) = X_\varepsilon(\boldsymbol{\omega}, \boldsymbol{\sigma})$ then the unstable direction at x, which is a vector tangent to $\mathcal{M}$, will be $w_{\boldsymbol{\omega},+}(x) = v_+ + K_{\boldsymbol{\omega}}(x) v_+$. The expansion rate at $(\boldsymbol{\omega}, x)$ will be, by (5.8) and (4.2),

$$A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) = \log\Big((\lambda_+ + \gamma_{\boldsymbol{\omega},+}(x)) \frac{|w_{\tau\boldsymbol{\omega},+}(f_{\boldsymbol{\omega}}(x))|}{|w_{\boldsymbol{\omega},+}(x)|}\Big) = A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) + g(\tau(\boldsymbol{\omega}, \boldsymbol{\sigma})) - g(\boldsymbol{\omega}, \boldsymbol{\sigma}), \qquad (6.1)$$

where $|w_{\boldsymbol{\omega},+}(x)| = \sqrt{1 + k_{\boldsymbol{\omega},-}(x)^2}$ and $g(\boldsymbol{\omega}, \boldsymbol{\sigma}) \stackrel{\text{def}}{=} \log |w_{\boldsymbol{\omega},+}(x)|$.

The functions $\lambda_+ + \gamma_{\boldsymbol{\omega},+}(x)$, $k_{\boldsymbol{\omega},-}(x)$ and $|w_{\boldsymbol{\omega},+}(x)|$ are analytic in ε and Hölder continuous in $\xi = (\boldsymbol{\omega}, x)$ with a uniformly bounded modulus $C(\beta)$ if $\beta < 1$ and ε is small enough. Since $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ fixed means ξ fixed, we see that $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ is analytic, as well as $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$, in ε at fixed $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and Hölder continuous in $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and therefore in ξ. ∎

It is now possible to conclude the proof of Theorem 2.

Proof of theorem 2 (conclusion). The two functions $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ and $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ generate two short range potential on the one–dimensional "spin system" whose states are the sequences $(\boldsymbol{\omega}, \boldsymbol{\sigma})$, in the sense of [14]. An identical argument holds for the function $A^0_{\boldsymbol{\omega},s}(\boldsymbol{\sigma}) \stackrel{\text{def}}{=} J^s_{\boldsymbol{\omega},1}(x)$, with $(\boldsymbol{\omega}, x) = X_\varepsilon(\boldsymbol{\omega}, \boldsymbol{\sigma})$ and the corresponding *contraction rate* $A_{\boldsymbol{\omega},s}(\boldsymbol{\sigma})$ (whose definition is worked out in the obvious way), see Appendix A1 below or [8, Prop.4.3.1].

Suppose that the probability distribution P on $\boldsymbol{\omega}$ has a short range potential energy density $B(\boldsymbol{\omega})$. Standard results on one–dimensional Gibbs distributions – see Appendix A1 below, where for convenience of the reader *the key definitions and results of the classical Ruelle's theory, used in the rest of this section, are quoted and referred to a single bibliographic source*, – imply that the "short range" Gibbs distribution ν with formal energy function

$$\sum_{k=-\infty}^{0} (A^0_{\tau^k\boldsymbol{\omega},s}(\tau^k\boldsymbol{\sigma}) + B(\tau^k\boldsymbol{\omega})) + \sum_{k=0}^{\infty}(A^0_{\tau^k\boldsymbol{\omega},u}(\tau^k\boldsymbol{\sigma}) + B(\tau^k\boldsymbol{\omega})) \tag{6.2}$$

is *well defined* and it has the property that, if $C^N_{\boldsymbol{\omega},\boldsymbol{\sigma}}$ is defined as the "cylinder set" $C^N_{\boldsymbol{\omega},\boldsymbol{\sigma}} \stackrel{\rm def}{=} \{\boldsymbol{\omega}',\boldsymbol{\sigma}' \,|\, \omega'_i = \omega_i,\, \sigma'_i = \sigma_i,\, \forall |i| \le N\}$, then

$$D^{-1} \le \frac{\nu(C^N_{\boldsymbol{\omega},\boldsymbol{\sigma}})}{{\rm vol}(C^N_{\boldsymbol{\omega},\boldsymbol{\sigma}})} \le D \tag{6.3}$$

for a suitable constant $D > 0$. This property is called the "absolute continuity" of the stable and unstable manifolds, and its proof is the same as in the classical Anosov case because the presence of $\boldsymbol{\omega}$ is just an index that has to be carried along possibly applying to it translations, see also [8, Prop.4.3.2]; the heuristic basis for it has been discussed in §4, see (4.6).

Furthermore the Gibbs distribution with energy function $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ is defined by a potential differing from the one of $A_{u,\boldsymbol{\omega}}(\boldsymbol{\sigma})$ by what is often called a Hölder continuous "cocycle", *i.e.* by $g(\boldsymbol{\omega},\boldsymbol{\sigma}) - g(\tau(\boldsymbol{\omega},\boldsymbol{\sigma}))$ in Lemma 4, which, therefore, generates the *same* Gibbs distribution. Hence the distribution ν is absolutely continuous with respect to the volume distribution, see Appendix A1 or [8, App. 6.4]

The distribution ν is *not* translation–invariant because its Gibbs potential in the "far future" is governed by $A^0_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ while in the "far past" it is governed by $A^0_{\boldsymbol{\omega},s}(\boldsymbol{\sigma})$, see (6.2). But short range Gibbs distributions (whether invariant or not) enjoy exponentially mixing properties: hence the distribution ν, and consequently the volume distribution, will be very close on the evolution of the cylinders, for very large times, to the invariant Gibbs distributions with (formal) potentials

$$\sum_{k=-\infty}^{+\infty} A_{\tau^k\boldsymbol{\omega},a}(\tau^k\boldsymbol{\sigma}), \qquad \begin{cases} \text{as } p \to +\infty \text{ if } a = u, \\ \text{as } p \to -\infty \text{ if } a = s, \end{cases} \tag{6.4}$$

respectively. Therefore the distribution with potential in (6.4) with $a = u$ is the SRB distribution for the forward evolution while for $a = s$ the distribution is the SRB distribution for the backward evolution and item (i) of Theorem 2 is proved.

The positivity of σ_+ is a genericity result. Let $\overline{\sigma}_{\boldsymbol{\omega}}(x) = (\sigma_{\boldsymbol{\omega}}(x)\varepsilon^{-1})_{\varepsilon=0}$, and assume for simplicity that f is diagonalizable, with diagonal elements $f_1, \ldots, f_N$, so that

$$\overline{\sigma}_{\boldsymbol{\omega}}(x) = -\sum_{\nu\in\mathbb{Z}}\sum_{j=1}^{m} i\nu_j \psi_{\boldsymbol{\omega},j,\nu} f_j^{-1} e^{i\nu\cdot x} \stackrel{\text{def}}{=} \sum_{\nu\in\mathbb{Z}^m} q_{\boldsymbol{\omega},\nu} e^{i\nu\cdot x}. \tag{6.5}$$

Then a sufficient and generic condition for a nonzero value of σ_+ is that $\overline{\sigma}_{\boldsymbol{\omega}}(0) \neq 0$ or, more generally, that the sum $\sum_k \sigma_{\tau^k\boldsymbol{\omega}}(f^k x) \neq 0$ for some $\mathcal{F}$-periodic point $(\boldsymbol{\omega}, x)$, when the sum over k is extended over the period, see [8, App.6.4]. See Appendix A3 below for some examples in the more difficult case of genericity under the presence of a time reversal symmetry. ∎

7. Time reversal symmetry

The result of §6 allows us to conclude that the SRB distribution $\mu_{\rm srb}$ will be a Gibbs state for the energy function $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ in the sense of [14].

We now *assume* the existence of an ε–independent analytic time reversal symmetry $\mathcal{I}$: $\mathcal{I}(\boldsymbol{\omega}, x) = (\boldsymbol{\omega}^T, Ix)$ and $I \circ f^k = f^{-k} \circ I$ for some integer k. In the following k will be taken $\equiv 1$ without loss of generality (if $k > 1$ the analysis could be performed by making use of f^k instead of f).

We shall also suppose that the Markovian pavement $\mathcal{P}_0$ is "reversible", *i.e.* such that for all $\sigma = 1, \ldots, n$ one has $IP_\sigma = P_{\sigma^T}$ for some σ^T. This is not restrictive because it can be achieved by considering the new pavement $\mathcal{E}_0 = I\mathcal{P}_0 \cap \mathcal{P}_0$. The uniqueness of the homeomorphism H_ε (see Lemma 2) and the analyticity of I, hence of $\mathcal{I}$, assure that H_ε and $\mathcal{I}$ commute (because another conjugating homeomorphism would be $\mathcal{I} \circ H_\varepsilon \circ \mathcal{I}$ *and it would be analytic in* ε: hence $\mathcal{I} \circ H_\varepsilon \circ \mathcal{I} \equiv H_\varepsilon$). Therefore also $\mathcal{E}$ is invariant under the action of $\mathcal{I}$: the map $\mathcal{I}$ applied to an element of the partition $\mathcal{E}$ gives another element of the partition.

The pavement $\mathcal{E} = H_\varepsilon(\mathcal{E}_0)$ considered in §6 will generate, under the action of $\mathcal{F}$, histories $(\boldsymbol{\omega}, \boldsymbol{\sigma})$ of points $\xi = (\boldsymbol{\omega}, x)$ such that the history of $\mathcal{I}\xi$ is $(\boldsymbol{\omega}^T, \boldsymbol{\sigma}^T)$, where $\boldsymbol{\sigma}^T = \{\sigma_{-j}\}_{j=-\infty}^{\infty}$.

The above analysis implies that the SRB distribution $\mu_{\rm srb}$ is a probability distribution with a Gibbs weight generated by a function $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) + B(\boldsymbol{\omega})$, as shown in §6. The functions $A_{\boldsymbol{\omega},a}$, $a = u, s$, are related to each other as follows.

LEMMA 4. *Under the assumptions of Theorem 1, one has*

$$A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) = -A_{\tau^{-1}\boldsymbol{\omega}^T,s}(\tau^{-1}\boldsymbol{\sigma}^T), \tag{7.1}$$

if τ is the translation of a generic sequence.

Proof. This is simply a consequence of the supposed isometric nature of the time reversal $\mathcal{I}(\boldsymbol{\omega},x)=(\boldsymbol{\omega}^T,Ix)\stackrel{\rm def}{=}(\boldsymbol{\omega}',x')$ and its smoothness. Let $\xi=(\boldsymbol{\omega},x)=X_\varepsilon(\boldsymbol{\omega},\boldsymbol{\sigma})$ and call $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma})$ the expansion rate and $A_{\boldsymbol{\omega},s}(\boldsymbol{\sigma})$ the contraction rate at ξ. Then $\mathcal{I}$ maps the stable manifold at ξ into the unstable one at $\mathcal{I}\xi$ and the forward expansion rate at ξ into the backward expansion rate at $\mathcal{I}\xi$, because $\mathcal{I}$ is isometric and smooth. ∎

Proof of Theorem 1 (conclusion). We have now all the elements to conclude the proof of (2.2), under the assumptions of Theorem 1, because it is now reduced, thanks to Lemma 2 and 4, to the proof of validity of (2.2) in the case in which the SRB distribution is a short range Gibbs distribution with a potential energy $A_{u,\boldsymbol{\omega}}(\boldsymbol{\sigma})+B(\boldsymbol{\omega})$ with the property (7.1) and $B(\boldsymbol{\omega})\equiv B(\boldsymbol{\omega}^T)$ which is the analytic form of the assumed reversibility of the random noise (see (i) of Definition 1). This was done first in [9] and a more formal mathematical proof is repeated in [10, Secs. 3,4].

In other words the presence of the noise just increases the number of symbols necessary to describe motions by a symbolic dynamics. The number of symbols remains however finite and the potentials for the future and past evolutions remain finite range. Hence the assumptions in [10, Secs. 3,4] hold in the present case too so that if time reversal is assumed, as in Theorem 1, the proof of the fluctuation relation is reduced to the usual one for Anosov maps and (2.2) follows, completing the proof of Theorem 1. ∎

Remarks on time reversal. (1) The genericity of $\sigma_+>0$ is here obtained only for maps satisfying only the assumptions in (i) of Theorem 1. It is likely, but not proved as far as we know, that $\sigma_+>0$ also for generic time reversible systems. Therefore it is important to show that such systems at least exist; examples are provided in Appendix A3.
(2) It is manifest that the assumption of ε–independence on $\mathcal{I}$ can be weakened into analyticity in ε uniformly in $\boldsymbol{\omega}$, keeping of course the uniform analyticity in x. The independence has been used at the beginning of this section to insure the existence of a reversible partition: this was based upon the commutation property between H_ε and $\mathcal{I}$. If $\mathcal{I}=\mathcal{I}_\varepsilon$ depends on ε,x analytically and uniformly in $\boldsymbol{\omega}$ the same construction can be performed by using that if H_ε is a conjugating homeomorphism then $\mathcal{I}_\varepsilon\circ H_\varepsilon\circ\mathcal{I}_0=H_\varepsilon$, which is all that is necessary to insure that the partition considered in §6 is reversible.

(3) Time reversal symmetry is a symmetry that should be inherited from the microscopic time reversal symmetry of purely Hamiltonian dynamics where it is independent of the perturbation size. For this reason we have supposed $\mathcal{I}$ to be ε–independent, which has also allowed us to point out the commutation property between conjugation and time reversal, which is interesting on its own. The time reversal symmetry, or just its analyticity, might be destroyed by the thermostats, or better, by the models that are introduced to describe them. Therefore it is worth considering the case in which $\mathcal{I}$ exists for a fixed value of ε and is just "somewhat smooth". In fact the analyticity of $\mathcal{I}$ as a function of ε, x is not really necessary: if there is a map $\mathcal{I}$ with the properties of being an isometry with $\mathcal{I}^2 = 1$ and anticommuting with $\mathcal{F}$, *i.e.* $\mathcal{I} \circ \mathcal{F} = \mathcal{F}^{-1} \circ \mathcal{I}$, for some fixed ε small enough so that Theorem 2 applies then one defines $\mathcal{E} = \mathcal{I} H_\varepsilon(\mathcal{E}_0) \wedge H_\varepsilon(\mathcal{E}_0)$ where $\mathcal{E}_0$ is as in §6. Then if $E_{\omega,\sigma,\sigma'}$ are the elements of the partition $\mathcal{E}$ one has $\mathcal{I} E_{\omega,\sigma,\sigma'} = E_{\omega,\sigma',\sigma}$ and, fixed ω_0 each $E_{\omega_0,\sigma,\sigma'}$ is a parallelepiped with axes contained in the manifolds generated by $W^s_\omega(x)$ and $W^u_\omega(x)$, if $\xi = (\boldsymbol{\omega}, x) \in E_{\omega_0,\sigma,\sigma'}$ (see (a) in §4), so that the volume of $\cap_{j=-N}^{N} \mathcal{F}^{*j} E_{\omega_j,\sigma_j,\sigma'_j}$ is proportional, to leading order as $N \to \infty$, to the exponential of (6.2) in the sense that (6.3) holds. This means that the SRB distribution is a Gibbs distribution with energy function $A_{\boldsymbol{\omega},u}(\boldsymbol{\sigma}) + B(\boldsymbol{\omega})$ satisfying (7.1) and this is all that is necessary to prove (2.2) for given ε small enough.

(4) The results of the previous sections provide motivation for the following further considerations on time reversal. Under the assumptions that we have posed upon $\mathcal{F}$ it is possible to show that if the noiseless system $\mathcal{F}_0$ is reversible a "kind" of time reversal symmetry will continue to hold *without extra assumptions* when the noisy perturbation is not neglected, provided the noise also has a reversibility property.

We supposed, since the very beginning in §1, reversibility of the stationary probability distribution P on Ω in the sense that the operation θ defined by $\theta\boldsymbol{\omega} = \boldsymbol{\omega}^T$ with $(\omega^T)_j = \omega_{-j}$ ("history reversal") conserves the distribution P:

$$P(E^T) = P(E), \tag{7.2}$$

where $E^T = \{\boldsymbol{\omega}^T : \boldsymbol{\omega} \in E\}$, for all (Borel-)sets E.

Suppose that the unperturbed system evolution on $\mathcal{M}$, that we assume to be an Anosov map f, is reversible in the sense that there is a map $I_0 : \mathcal{M} \to \mathcal{M}$ with the property $I_0 \circ f^k = f^{-k} \circ I_0$ for some k and $I_0^2 = \pm 1$, or more generally $I_0^{k'} = 1$ for some k' even. The above assumption is satisfied in the simplest cases treated in the previous sections because the time reversal operation $I_0(\varphi_1,\varphi_2)=(\varphi_2,-\varphi_1)$ has the property $I_0\circ f^2=f^{-2}\circ I_0$, if f is given by (1.1).

The problem is that it is not clear that a time reversal symmetry holds when there is a perturbation, even if small. But, if ε is small, the time reversal symmetry $\mathcal{I}_0(\boldsymbol{\omega},x)=(\boldsymbol{\omega}^T,I_0x)$ is not "broken" by the perturbation and the noisy system admits a time reversal symmetry $\mathcal{I}_\varepsilon$. The latter can be immediately written as

$$\mathcal{I}_\varepsilon \stackrel{\rm def}{=} H_\varepsilon\circ\mathcal{I}_0\circ H_\varepsilon^{-1}. \tag{7.3}$$

However for the same reasons for which the conjugacy between $\mathcal{F}$ and $\mathcal{F}_0$ was not smoother than Hölder–continuous it will turn out that $\mathcal{I}_\varepsilon$ is also not necessarily smoother than Hölder–continuous.

In terms of symbolic dynamic the time reversal $\mathcal{I}_\varepsilon$ acts on a point whose symbolic history is $(\boldsymbol{\omega},\boldsymbol{\sigma})$ by transforming it into the point with symbolic history $(\boldsymbol{\omega}^T,\sigma^T)$, where $(\boldsymbol{\omega}^T)_j=\omega_{-j}$, $(\boldsymbol{\sigma}^T)_j=\sigma^T_{-j}$, and σ^T is defined assuming that the unperturbed time reversal $\mathcal{I}_0$ transforms the element P_σ of the Markov pavement for f into another element P_{σ^T} of the same Markov pavement. It is therefore clear that $\mathcal{I}_\varepsilon^{k'}=1$ if $I_0^{k'}=1$.

The lack of smoothness of the new time reversal implies that the fluctuation theorem can be proved only in cases in which the perturbed system also admits a smooth time reversal. However it would be interesting to understand which implications on the rate of the large fluctuations has a strong property like (7.3). For instance to what extent is it possible to use it (and its moderate smoothness) to study the deviations from the fluctuation relation and *estimate them*?

8. Structural stability

Consider the case in which the map $x\to fx$ is just an Anosov map of a general manifold $\mathcal{M}$, still assuming analytic regularity. If the unperturbed Anosov map f is *close to some hyperbolic linear torsion f_0 of a torus* it will be possible conjugate analytically (in ε) the systems $\mathcal{F}$ to $\mathcal{F}_0$ simply because we can think that the perturbation $\varepsilon\psi_{\boldsymbol{\omega}}$ contains a non–random part which is the perturbation of f_0 into f. The same can be said of maps that can be analytically conjugated to hyperbolic torsions. Of course to extend the results on the fluctuation relation, however, the strong assumption about time reversal has to be always added.

More interesting and subtle is the further question of how much the above analysis depends on the assumption that we are perturbing a system close to one $\mathcal{F}_0$ in which a reversible noise and a linear torsion of a torus f do not interact.

Under the only assumption that f is analytic and Anosov, hence *not necessarily close to a linear hyperbolic map of a torus*, the equations for H will have to be written in coordinates and possibly several coordinate charts will have to be used. An apparently unsurmountable difficulty may seem that the directions $v_\pm$ will now depend on x and therefore the solution algorithm will simply fail. A more attentive study of the equations shows, nevertheless, that the equations are *solvable* order by order since at every order the curvature of the manifolds systematically gives contributions of higher order. See [1] for details, showing the existence of a convergent power series in ε for H. Since that is the main difficulty, it is a natural conjecture that the analysis that we have performed in detail for the case of the random perturbations of linear hyperbolic maps of tori carries to the last case.

Appendix A1. One–dimensional Gibbs distributions

Here we collect a few basic elements of the theory of "Gibbs distributions" in 1-dimension [8, Sec.5.1].

DEFINITION A1. *(i) A "compatibility matrix" M is a $n\times n$ ($n\in$) matrix with entries 0 or 1 and with the "transitivity" property: for some p one has $(M^p)_{i,j}>0$ for every $i,j=1,\ldots,n$. The space $\Omega_M=\{1,2,\ldots,n\}^{\mathbb{Z}}_M$ is the space of the bilateral sequences $\boldsymbol{\omega}$ such that $M_{\omega_i,\omega_{i+1}}\equiv 1$. Such sequences will be called "M–compatible". Likewise if $X=\{x,x+1,\ldots,x+q\}$ is a finite "interval" in a sequence $\boldsymbol{\omega}_X\in\{1,2,\ldots,n\}^{\mathbb{Z}}$ is called M–compatible if $M_{\omega_{x+i},\omega_{x+i+1}}=1$ for $i=0,\ldots,q-1$.*
(ii) A "potential function" Φ on $\Omega=\{1,2,\ldots,n\}^{\mathbb{Z}}$ associates with every finite subset $X\subset$ an every "configuration $\boldsymbol{\omega}_X\in\{1,\ldots,n\}^X$ in X" a "potential energy" $\Phi_X(\boldsymbol{\omega}_X)\in$.
(iii) A potential function has "short range" if there exist $\kappa,B>0$ such that

$$\|\Phi\|=\sup_{z\in\mathbb{Z}}\sum_{X\ni z}\max_{\boldsymbol{\omega}_X}|\Phi_X(\boldsymbol{\omega}_X)|\,e^{\kappa\,\mathrm{diam}(X)}=B. \tag{A1.1}$$

(iv) Φ is "translation invariant" if the translation τ of the lattice by one unit is such that $\Phi_X(\boldsymbol{\omega}_X)=\Phi_{\tau X}(\boldsymbol{\omega}_X)$ for all $X\subset$ and for all configurations $\boldsymbol{\omega}_X$.

(v) A potential Φ is "translation invariant at $s\infty$", $s=\pm$, if there exists a tranlation invariant Φ_s and $\kappa, B>0$ such that

$$\sum_{z\in\mathbb{Z}}\sum_{X\ni z}\max_{\omega_X}|\Phi_X(\omega_X)-\Phi_{s,X}(\omega_X)|\,e^{\kappa\,\mathrm{diam}(X)}e^{\kappa\,s\,z}=B<\infty, \tag{A1.2}$$

and a potential which is translation invariant at $+\infty$ and at $-\infty$ will be called a "semitranslationally invariant" potential.
(vi) If Φ is a short range potential, $\Lambda\subset$ is finite, and $\omega\in\Omega$ the "energy of ω in Λ" is

$$U_\Lambda(\omega)=\sum_{X\cap\Lambda\neq\emptyset}\Phi_X(\omega_X). \tag{A1.3}$$

(vii) Let $\omega_X+\omega_Y$ denote the configuration $\omega_{X\cup Y}$ which coincides with ω_X on X and with ω_Y on Y. A probability distribution μ on Ω_M is a "Gibbs distribution" with short range potential Φ and "hard core" M if for every interval $\Lambda\subset$ and every sequence $\omega\in\Omega_M$ the conditional probability $\mu_\Lambda(\omega_\Lambda;\omega_{\mathbb{Z}/\Lambda})$ that a sequence in Ω_M randomly chosen with probability μ coincides with ω_Λ on Λ given that it coincides with $\omega_{\mathbb{Z}/\Lambda}$ on $/\Lambda$ is ("DLR equation"):

$$\mu_\Lambda(\omega_\Lambda;\omega_{\mathbb{Z}/\Lambda})=Z_\Lambda(\omega_{\mathbb{Z}/\Lambda})^{-1}e^{-U_\Lambda(\omega_\Lambda+\omega_{\mathbb{Z}/\Lambda})} \tag{A1.4}$$

for all $(\omega_\Lambda;\omega_{\mathbb{Z}/\Lambda})$ for which $\omega'_\Lambda+\omega_{\mathbb{Z}/\Lambda}$ is compatible, i.e. it is in Ω_M. Here the normalization factor $Z_\Lambda(\omega_{\mathbb{Z}/\Lambda})$ is $\sum^*_{\omega'_\Lambda}e^{-U_\Lambda(\omega'_\Lambda+\omega_{\mathbb{Z}/\Lambda})}$ where the $*$ means that the sum is restricted to the ω'_Λ which are compatible with $\omega_{\mathbb{Z}/\Lambda}$.
(viii) The "energy density at $z\in$ " of a short range potential Φ is the function on Ω

$$A_z(\omega,\Phi)=\sum_{X\ni z}\frac{\Phi_X(\omega_X)}{|X|} \tag{A1.5}$$

and the functions A_z are uniformly Hölder continuous with exponent $\beta'<\beta$ when Φ has short range in the sense of (ii) above.

In the Physics literature the above setting would be often considered as describing a "spin system" in 1–dimension with short reange interaction. This is a class of systems that are so well understood to be considered trivial: remarkably it is nevertheless very useful to study chaotic systems as shown by the following theorem and proposition.

The main theorem on Gibbs states with short range interactions is as follows [8, Sec.5,6].

THEOREM A1. *(i) If Φ is a short range potential and M is a compatibility matrix there is a unique Gibbs state μ on Ω_M with potential Φ*

(ii) The Gibbs distribution μ mixes at exponential rate all Hölder continuous functions: i.e. there exist constants $C = C(F,G), \kappa = \kappa(F,G)$ such that

$$\Big|\int_{\Omega_M} F(\boldsymbol\omega)G(\tau^r\boldsymbol\omega)\mu(d\boldsymbol\omega) - \Big(\int_{\Omega_M} F(\boldsymbol\omega)\mu(d\boldsymbol\omega)\Big)\Big(\int_{\Omega_M} G(\tau^r\boldsymbol\omega)\mu(d\boldsymbol\omega)\Big)\Big| \le C e^{-\kappa r}. \qquad (A1.6)$$

(iii) If the potential Φ is translation invariant at $s\infty$, $s = \pm$ with asymptotic value $\Phi_{s\infty}$ then

$$\Big|\int_{\Omega_M} F(\boldsymbol\omega)G(\tau^r\boldsymbol\omega)\mu(d\boldsymbol\omega) - \Big(\int_{\Omega_M} F(\boldsymbol\omega)\mu(d\boldsymbol\omega)\Big)\Big(\int_{\Omega_M} G(\boldsymbol\omega)\mu_s(d\boldsymbol\omega)\Big)\Big| \le C e^{-\kappa r}. \qquad (A1.7)$$

for $C, \kappa > 0$ suitable functions of F, G. Here μ_s is the Gibbs state generated by the potential $\Phi_{s\infty}$.

(iv) μ is a Gibbs distribution with potential Φ on Ω_M if and only if the conditional probability in (A1.4) satisfies, if A_0 is the function defined in A_ξ (A1.5) evaluated $\xi = 0$,

$$\frac{\mu_\Lambda(\boldsymbol\omega'_\Lambda;\boldsymbol\omega_{\mathbb{Z}/\Lambda})}{\mu_\Lambda(\boldsymbol\omega''_\Lambda;\boldsymbol\omega_{\mathbb{Z}/\Lambda})} = \exp\Big(-\sum_{k=-\infty}^{+\infty}\{A_0(\tau^k\boldsymbol\omega') - A_0(\tau^k\boldsymbol\omega'')\}\Big) \qquad (A1.8)$$

for all $\boldsymbol\omega' = \boldsymbol\omega'_\Lambda + \boldsymbol\omega_{\mathbb{Z}/\Lambda}, \boldsymbol\omega'' = \boldsymbol\omega''_\Lambda + \boldsymbol\omega_{\mathbb{Z}/\Lambda}$ in Ω_M.

Remarks. (1) If the hard core is not trivial, *i.e.* $M_{i,j} \not\equiv 1$, not all sequences are compatible and the values $\Phi_X(\boldsymbol\omega_X)$ of the potentials Φ on incompatible $\boldsymbol\omega_X$, *i.e.* on $\boldsymbol\omega_X$ which are not substrings of a sequence $\boldsymbol\omega \in \Omega_M$, are completely irrelevant: hence they can be set $= 0$. The 1–dimensionality of the lattice which labels the sequences elements is absolutely essential for the above results.

(2) The result (i) is classical and a proof can be found in [8, Prop.5.1.1] for the existence in the case of translation invariant Φ and in [8, Prop.5.2.1] for the uniqueness. The proofs however work unchanged if the more general non–translation invariant cases. The cases that are of interest in the theory of Anosov systems (hence in this work) require the result only for potentials that are translation invariant at $+\infty$ and $-\infty$ simoultaneously and proofs of the results (ii) and (iii) (also classical) can be found in [8, Prop.6.3.3] and in [8, Cor.6.3.1].

(3) The equivalence result (iv) is also a classical remark, see Remark (2) to [8, Def.5.1.2].

(4) Any Hölder continuous function $A(\boldsymbol\omega)$ on Ω_M can be represented as energy of a short range potential. This is immediate [8, Prop.4.3.1].

(5) Any transitive Anosov system on a manifold V admits a Markov partition, [8, Sec.4], *i.e.* a partition which codes the points $x \in V$ into their history on the partition establishing a correspondence between the points and a set of symbolic sequences Ω_M for some transitive compatibility matrix M. Every sequence in Ω_M represents a single point in V and the coding is $1-1$ outside a set of zero volume in V.

PROPOSITION A1. *(i) The normalized Lebesgue measure on a manifold V on which acts a transitive Anosov map f is coded by the symbolic dynamics $x \longleftrightarrow \boldsymbol{\omega}(x)$ associated with f and with a Markov partition into a probability distribution which is a Gibbs distribution with a potential Φ.*
(ii) The potential Φ is translation invariant at $-\infty$ and at $+\infty$ and $\Phi_{+\infty}$ is generated by the Hölder continuous function $A_+(\boldsymbol{\omega}) = \log|\det(\frac{\partial f(x)}{\partial x})_+|$ where $(\frac{\partial f(x)}{\partial x})_+$ is the Jacobian matrix of the restriction of the map f to the unstable manifold through x and $\boldsymbol{\omega}$ is the symbolic representation of x. Likewise $\Phi_{-\infty}$ is generated by the analogous function $A_-(\boldsymbol{\omega}) = -\log|\det(\frac{\partial f(x)}{\partial x})_-|$ relative to the stable manifold.

Remark. The combination of Proposition A1 and Theorem A1 shows the interest of Gibbs distributions to represent the statistics under the evolution of initial data x by f when the data are chosen with a probability distribution which is absolutely continuous with respect to the volume; a proof can be found in [8, Prop.4.3.2]

Appendix A2. Details on overshadowing bounds

Comparing the trees in Figure 2 and those in Figure 1 we see that the new trees, once they are fully developed – so that the labels (k_v) of the top nodes, marking the order in ε, are all (1) – have three kinds of nodes: namely $\square, \circ, \bullet$.

The number of lines in a tree contributing to the formation of the k-th order is not necessarily k (unlike the case of the construction of H): this is due to the last graph in Figure 2 whose order in ε has to be read from the two incoming lines. Looking at the structure of the the graphs it follows that the maximum number of lines that a tree graph contributing to the k-th order in ε is $\le 2k$ (as one can easily show by induction).

Furthermore for a node v which is either $\square$ or $\circ$ both the label α_v and the labels $(\alpha'_\ell, \alpha_\ell)$ or α_ℓ of the entering lines ℓ are uniquely determined by the label α_{ℓ_v} (see Figure 2). Thus for each of them we have at most three possible choices (in fact two for $\circ$ and three for $\square$).

On the contrary for each • we have two possible choices, as $\alpha_v \in \{\pm\}$. This means that the factor 2^k in §3 due to the sum over the labels α_v has to be replaced by a factor 3^{2k}. Hence in the worst case the number of labels cannot exceed $3^{2k} \cdot 2^{4k}$ while in the case for H it was 2^{3k} (recall that in that case the number of nodes was k).

The nodes v with wavy lines give a contribution proportional to a derivative of $\partial\psi$ of the order equal to the number of lines incident on the node while in the case of H the number of derivatives acting on ψ was always equal to the number of lines that enter the node, *hence* 1 *less*. This means that when considering the Fourier transform there will be an extra factor proportional to a component of ν_v per node (in the worst cases): this will give in the bounds an extra factor N. Since each node involving a derivative of ψ raises by one unit the order in ε to which the tree value gives a contribution (since in the map ψ is multiplied by ε) we conlude that to order k there will be an extra factor N^k, so that the factor N^{2k} of §3 has to replaced with N^{3k}. (We note that a more careful analysis would show that in fact again the same bound N^{2k} as before could be obtained, but we neglect such an improvement as we are not looking at all for optimal bounds.)

Therefore the conclusion is that convergence follows for ε smaller than the value in (3.16) with N replaced by N^2 and the constant Ψ suitable increased by $3^2 \cdot 2$.

Appendix A3. A reversible dissipative system

The example in (2.3) can be used for this purpose: consider the map acting on $(\boldsymbol{\varphi},\boldsymbol{\varphi}') \in {}^4$ $({}^4 \equiv [0,2\pi]^2 \times [0,2\pi]^2)$ defined by $(\boldsymbol{\varphi},\boldsymbol{\varphi}') \to (F\boldsymbol{\varphi}, F^{-1}\boldsymbol{\varphi}')$ with $F\boldsymbol{\varphi} = (\varphi_1+\varphi_2+\varepsilon\sin\varphi_2, \varphi_1)$ if $\boldsymbol{\varphi} = (\varphi_1,\varphi_2)$, $\boldsymbol{\varphi}' = (\varphi_1',\varphi_2')$: a direct calculation shows that the quantity called $\overline{\sigma}_{\boldsymbol{\omega}}$ in (6.5) is, in this case, $\overline{\sigma}_{\boldsymbol{\omega},\boldsymbol{\eta}}(\boldsymbol{\varphi},\boldsymbol{\varphi}') = -\cos\varphi_2 + \cos(\varphi_1'-\varphi_2') + O(\varepsilon)$. The point $((0,0),(\pi,\pi))$ has period 3 and the sum of the values of $\overline{\sigma}_{\boldsymbol{\omega},\boldsymbol{\eta}}$ on its orbit is -4 *hence not zero*, so that $\sigma_+ > 0$ for ε small enough, as remarked at the end of §6, [8, App.6.4]. Any further perturbation of the form $\varepsilon\gamma\psi_{\boldsymbol{\omega},\boldsymbol{\eta}}(\boldsymbol{\varphi},\boldsymbol{\varphi}')$ will also have a $\overline{\sigma}_{\boldsymbol{\omega},\boldsymbol{\eta}}(\boldsymbol{\varphi},\boldsymbol{\varphi}')$ with positive average if γ is small enough: thus providing us with many examples of reversible, random and non–random, maps analytic in

φ, φ'.

---

[1] F. Bonetto, P. Falco, and A. Giuliani, *Analyticity of the SRB measure of a lattice of coupled Anosov diffeomorphisms of the torus*, Journal of Mathematical Physics **45** (2004), 3282–3309.

[2] D. Ruelle, *Positivity of entropy production in the presence of a random thermostat*, Journal of Statistical Physics **86** (1997), 935–951.

[3] J. Bricmont and A. Kupiainen, *Coupled analytic maps*, Nonlinearity **8** (1995), 379–396.

[4] M. Jiang and A.E. Mazel, *Uniqueness and exponential decay of correlations for some two-dimensional spin lattice systems*, Journal of Statistical Physics **82** (1996), 797–821.

[5] M. Jiang and Ya.B. Pesin, *Equilibrium measures for coupled map lattices: existence, uniqueness and finite-dimensional approximations*, Communications in Mathematical Physics **193** (1998), 675–711.

[6] J. Bricmont and A. Kupiainen, *High temperature expansions and dynamical systems*, Communications in Mathematical Physics **178** (1996), 703–732.

[7] J. Bricmont and A. Kupiainen, *Infinite-dimensional SRB measures*, Physica D **103** (1997), 18–33.

[8] G. Gallavotti, F. Bonetto, and G. Gentile, *Aspects of the ergodic, qualitative and statistical theory of motion*, Springer Verlag, Berlin, 2004.

[9] G. Gallavotti and E. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics*, Physical Review Letters **74** (1995), 2694–2697.

[10] G. Gallavotti and E. Cohen, *Dynamical ensembles in stationary states*, Journal of Statistical Physics **80** (1995), 931–970.

[11] G. Gallavotti, *Reversible Anosov diffeomorphisms and large deviations*, Mathematical Physics Electronic Journal (MPEJ) **1** (1995), 1–12.

[12] J. Kurchan, *Fluctuation theorem for stochastic dynamics*, J. Physics A **31** (1998), 3719–3729.

[13] J. Lebowitz and H. Spohn, *A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics*, Journal of Statistical Physics **95** (1999), 333–365.

[14] C. Maes, *The fluctuation theorem as a Gibbs property*, Journal of Statistical Physics **95** (1999), 367–392.

[15] F. Bonetto, A. Kupiainen, and J. Lebowitz, *Absolute continuity of projected SRB measures of*

coupled Arnold cat map lattices, Ergodic theory and dynamical systems **25** (2005), 59–88.

[16] P.-D. Liu, *Dynamics of random transformations: smooth ergodic theory*, Ergodic theory and dynamical systems **21** (2001), 1279–1319.

REVTEX & EQALIGNNO & FANCYHDR