# DECAY OF CORRELATIONS

Carlangelo Liverani

University of Rome II

*Dedicated to Micheline Ishay*

Abstract. This paper offers a new unified approach for studying statistical properties of hyperbolic dynamical systems and their small random perturbations. The approach is based on a technique that permits an investigation of the convergence of a semigroup of operators by using an adapted Hilbert metric. An exponential rate of decay of correlations is proven in several classes of examples (including some hyperbolic billiards).

## CONTENT



**I would like to thank P. Boyland, N. Chernov, L. Chierchia, V. Donnay, G. De Martino, C. Golè, J.L. Lebowitz, M. Lyubich, M. Rychlik, I.G. Schwarz, S. Vaienti and especially G. Gallavotti for helpful and enlightening discussions. In addition, I am indebted to P. Collet and, most of all, M. Wojtkowski for introducing me to the magical world of cones. Finally, I thank J. Milnor, director of the Institute for Mathematical Sciences at Stony Brook University, where I was visiting during part of this work, and the Italian C.N.R.-GNFM for providing travel funds.**

## §0 INTRODUCTION.

The decay of correlations or, alternatively, the rate of approach of some initial distribution to an invariant one, is a problem addressed in many fields. In statistical mechanics such problem is discussed via the study of the transfer operator (also called Ruelle-Perron-Frobenius operator) [Ru]. In the field of stochastic differential equations it is possible to obtain general results by PDE techniques [H]. For dynamical systems one can use Markov partitions [Bow].

The Markov partition is a tool whereby a hyperbolic dynamical system can be reduced to (or, better, coded into) a statistical mechanical like system: the "symbolic dynamics." At the moment, the main limitations of this approach come from the need for uniform hyperbolicity and from the necessity to use infinite Markov partitions for discontinuous systems. These limitations obstruct the extension of this technique to more general systems, with some notable exceptions [BSC], [LSV].

As already mentioned, in statistical mechanics the desired results are obtained by studying the transfer operator. Given that the Perron-Frobenius operator is a natural object in the context of dynamical systems, it seems plausible that some advantages could be gained by investigating directly the Perron-Frobenius operator, without previously coding the dynamics. Moreover, adopting such point of view, the investigation of dynamical systems would be more similar to that of random systems, and the study of random perturbations of dynamical systems could be facilitated.

Results in the direction of a direct study of the Perron-Frobenius operator already exist, but they are limited to one dimensional systems; particularly relevant is the investigation of one-dimensional quadratic families of maps and their random perturbations in [Y],[Co], since it is unclear if Markov partitions could be used to study such examples; and [F], [FS2], where one-dimensional expansive maps are studied using an approach similar to the one proposed here.

In this paper I describe a technique, originally due to G. Birkhoff [B1], [B2], that permits a direct study of the Perron-Frobenius operator, and I show that its field of applicability is wider than that of Markov partitions. In essence, it is possible to construct systematically metrics (Hilbert metrics) with respect to which the Perron-Frobenius operator is a contraction. Such contraction allows us to obtain the invariant measure (if not already known) by an elementary, and constructive, fixed point theorem, rather than by some compactness argument (this may please some idiosyncratic people, myself included), and automatically implies an exponential rate for the decay of correlations.

I illustrate such an approach by applying it to several examples. For the sake of brevity and clarity the results are not presented in their full generality. In particular, all the arguments used for two-dimensional smooth maps can be easily extended to the $n$-dimensional case.

Results concerning more general systems (notably general billiards and non-uniformly hyperbolic maps) will be published in separate papers. I also hope that the present exposition will prompt others to apply this method to the many cases where it could yield new results (e.g. dissipative systems, flows, etc.).

The structure of the paper is as follows: section 1 describes the Hilbert metric and its properties. It is a brief review of the subject, intended to provide an easy reference for the reader. Section 2 shows how the technique works in the simplest example: a one dimensional uniformly hyperbolic map. It also mentions other consequences that can be obtained (e.g. Central Limit Theorem type results). Section 3 applies the same technique to the study of random perturbations. As a byproduct, it is possible to explicitly estimate the change of the invariant measure due to random perturbations. In sections 4 I show how to extend the approach to the multidimensional case—the smooth case is treated. Section 5 is the most technical section; in it I show that even the presence of singularities can be accommodated, provided they are not too wild. I do not claim to have the weakest possible conditions on the singularities; more work in this direction is needed. This issue is partially explored in section 6. There I extend the approach to prove exponential decay of correlations in some special billiards. The section is the only one in the paper that is not self contained (it relays partially on [BSC]) and cannot be read independently of the previous sections. Unfortunately, a more detailed investigation of billiards would require considerably more space. Nevertheless, the present result settles a question that has been open for some time notwithstanding thorough investigations (e.g., [BSC], [Ch]). In addition, the numerical simulations (see [CCG], [BD], [FM]) did not provide a clear cut conjecture, as is clearly shown by recent results [GG]. Finally, sections 7 deals with random perturbations of multidimensional systems, and in particular shows which of the results holding in one dimension remain valid. The paper includes two Appendices. The first presents a standard result in the theory of maps with singularities. I do not refer to the literature because the form in which the result is needed here is non standard and (misleadingly) appears to be stronger than the usual statement. In the second Appendix, I prove a result concerning the absolute continuity of the stable foliation for discontinuous systems. More precisely, I show that the Jacobian of the canonical isomorphism between two stable manifolds is Hölder continuous. This result is not surprising but, to my knowledge, is not present in the literature.

A final word: the reader should be aware that, in order to simplify the formulas, no attempt has been made to optimize the constants involved. Many times the value of a constant used in an inequality reflects only my predilection for the numbers 2 and 3.

## §1 OPERATORS AND INVARIANT CONVEX CONES.

This section illustrates some results in lattice theory originally due to Garrett Birkhoff. For more details see [B2], and [N] for a recent overview of the field.

Consider a topological vector space $\mathbb{V}$ with a partial ordering "$\preceq$," that is a vector lattice.[1] We require the partial order to be "continuous," i.e. given $\{f_n\} \in \mathbb{V}$

[1] We are assuming the partial order to be well behaved with respect to the algebraic structure: for each $f, g \in \mathbb{V}$ $f \succeq g \iff f - g \succeq 0$; for each $f \in \mathbb{V}$, $\lambda \in \mathbb{R}^+ \setminus \{0\}$ $f \succeq 0 \implies \lambda f \succeq 0$; for each $f \in \mathbb{V}$ $f \succeq 0$ and $f \preceq 0$ imply $f = 0$ (antisymmetry of the order relation).

$\lim_{n\to\infty} f_n = f$, if $f_n \succeq g$ for each $n$, then $f \succeq g$. We call such vector lattices "integrally closed."[2]

We define the closed convex cone[3] $\mathcal{C} = \{f \in \mathbb{V} \mid f \neq 0,\ f \succeq 0\}$ (hereafter, the term "closed cone" $\mathcal{C}$ will mean that $\mathcal{C} \cup \{0\}$ is closed), and the equivalence relation "$\sim$": $f \sim g$ iff there exists $\lambda \in \mathbb{R}^+ \backslash \{0\}$ such that $f = \lambda g$. If we call $\widetilde{\mathcal{C}}$ the quotient of $\mathcal{C}$ with respect to $\sim$, then $\widetilde{\mathcal{C}}$ is a closed convex set. Conversely, given a closed convex cone $\mathcal{C} \subset \mathbb{V}$, enjoying the property $\mathcal{C} \cap -\mathcal{C} = \emptyset$, we can define an order relation by

$$f \preceq g \iff g - f \in \mathcal{C} \cup \{0\}.$$

Henceforth, each time that we specify a convex cone we will assume the corresponding order relation and vice versa. The reader must therefore be advised that "$\preceq$" will mean different things in different contexts.

It is then possible to define a projective metric $\Theta$ (Hilbert metric), in $\mathcal{C}$, by the construction:

$$\begin{aligned}
\alpha(f,\, g) &= \sup\{\lambda \in \mathbb{R}^+ \mid \lambda f \preceq g\}\\
\beta(f,\, g) &= \inf\{\mu \in \mathbb{R}^+ \mid g \preceq \mu f\}\\
\Theta(f,\, g) &= \log\left[\frac{\beta(f,\, g)}{\alpha(f,\, g)}\right]
\end{aligned}$$

where we take $\alpha = 0$ and $\beta = \infty$ if the corresponding sets are empty.

The importance of the previous constructions is due, in our context, to the following theorem.

**Theorem 1.1.** *Let $\mathbb{V}_1$, and $\mathbb{V}_2$ be two integrally closed vector lattices; $T : \mathbb{V}_1 \to \mathbb{V}_2$ a linear map such that $T(\mathcal{C}_1) \subset \mathcal{C}_2$, for two given closed convex cones $\mathcal{C}_1 \subset \mathbb{V}_1$ and $\mathcal{C}_2 \subset \mathbb{V}_2$. Let $\Theta_i$ be the Hilbert metric corresponding to the cone $\mathcal{C}_i$. Setting $\Delta = \sup_{f,\, g \in T(\mathcal{C}_1)} \Theta_2(f,\, g)$ we have*

$$\Theta_2(Tf,\, Tg) \leq \tanh\left(\frac{\Delta}{4}\right)\Theta_1(f,\, g) \qquad \forall f,\, g \in \mathcal{C}_1$$

*(*$\tanh(\infty) = 1$*).*

*Proof.* The proof is provided for the reader convenience.

Let $f,\, g \in \mathcal{C}_1$, on the one hand if $\alpha = 0$ or $\beta = \infty$, then the inequality is obviously satisfied. On the other hand, if $\alpha \neq 0$ and $\beta \neq \infty$, then

$$\Theta_1(f,\, g) = \ln\frac{\beta}{\alpha}$$

[2]To be precise, in the literature "integrally closed" is used in a weaker sense. First, $\mathbb{V}$ does not need a topology. Second, it suffice that for $\{\alpha_n\} \in \mathbb{R}$, $\alpha_n \to \alpha$; $f,\, g \in \mathbb{V}$, if $\alpha_n f \succeq g$, then $\alpha f \succeq g$. Here we will ignore these and other subtleties: our task is limited to a brief account of the results relevant to the present context.

[3]Here, by "cone," we mean any set such that, if $f$ belongs to the set, then $\lambda f$ belongs to it as well, for each $\lambda > 0$.

where $\alpha f \preceq g$ and $\beta f \succeq g$, since $\mathbb{V}_1$ is integrally closed. Notice that $\alpha \geq 0$, and $\beta \geq 0$ since $f \succeq 0$, $g \succeq 0$. If $\Delta = \infty$, then the result follows from $\alpha Tf \preceq Tg$ and $\beta Tf \succeq Tg$. If $\Delta < \infty$, then, by hypothesis,

$$\Theta_2\left(T(g - \alpha f),\, T(\beta f - g)\right) \leq \Delta$$

which means that there exist $\lambda,\, \mu \geq 0$ such that

$$\begin{aligned} \lambda T(g - \alpha f) &\preceq T(\beta f - g) \\ \mu T(g - \alpha f) &\succeq T(\beta f - g) \end{aligned}$$

with $\ln \frac{\mu}{\lambda} \leq \Delta$. The previous inequalities imply

$$\begin{aligned} \frac{\beta + \lambda\alpha}{1+\lambda} Tf &\succeq Tg \\ \frac{\mu\alpha + \beta}{1+\mu} Tf &\preceq Tg. \end{aligned}$$

Accordingly,

$$\begin{aligned} \Theta_2(Tf,\, Tg) &\leq \ln \frac{(\beta + \lambda\alpha)(1+\mu)}{(1+\lambda)(\mu\alpha + \beta)} = \ln \frac{e^{\Theta_1(f,\, g)} + \lambda}{e^{\Theta_1(f,\, g)} + \mu} - \ln \frac{1+\lambda}{1+\mu} \\ &= \int_0^{\Theta_1(f,\, g)} \frac{(\mu - \lambda)e^{\xi}}{(e^{\xi} + \lambda)(e^{\xi} + \mu)} d\xi \leq \Theta_1(f,\, g) \frac{1 - \frac{\lambda}{\mu}}{\left(1 + \sqrt{\frac{\lambda}{\mu}}\right)^2} \\ &\leq \tanh\left(\frac{\Delta}{4}\right) \Theta_1(f,\, g). \end{aligned}$$

**Remark 1.2.** *If $T(\mathcal{C}_1) \subset \mathcal{C}_2$ then it follows that $\Theta_2(Tf,\, Tg) \leq \Theta_1(f,\, g)$. Hovever, a uniform rate of contraction depends on the diameter of the image being finite.*

In particular, if an operator maps a convex cone strictly inside itself (in the sense that the diameter of the image is finite), then it is a contraction in the Hilbert metric. This implies the existence of a "positive" eigenfunction (provided the cone is complete with respect to the Hilbert metric), and, with some additional work, the existence of a gap in the spectrum of $T$ (see [B1] for details). The relevance of this theorem for the study of invariant measures and their ergodic properties is obvious.

It is natural to wonder about the strength of the Hilbert metric compared to other, more usual, metrics. While in general it depends on the cone, it is nevertheless possible to state an interesting result.

**Lemma 1.3.** *Let $\|\cdot\|$ be a norm in $\mathbb{V}$, and suppose that, for each $f,\, g \in \mathbb{V}$,*

$$-f \preceq g \preceq f \Longrightarrow \|f\| \geq \|g\|.$$

*Then, given $f,\, g \in \mathcal{C} \subset \mathbb{V}$ for which $\|f\| = \|g\|$,*

$$\|f - g\| \leq \left(e^{\Theta(f,\, g)} - 1\right) \|f\|.$$

*Proof.* We know that $\Theta(f,\,g)=\ln\frac{\beta}{\alpha}$, where $\alpha f\preceq g$, $\beta f\succeq g$. This implies that $-g\preceq 0\preceq\alpha f\preceq g$, i.e. $\|g\|\geq\alpha\|f\|$, or $\alpha\leq 1$. In the same manner it follows that $\beta\geq 1$. Hence,

$$\begin{aligned}g-f\preceq&(\beta-1)f\preceq(\beta-\alpha)f\\ g-f\succeq&(\alpha-1)f\succeq-(\beta-\alpha)f\end{aligned}$$

which implies

$$\|g-f\|\leq(\beta-\alpha)\|f\|\leq\frac{\beta-\alpha}{\alpha}\|f\|=\left(e^{\Theta(f,\,g)}-1\right)\|f\|.$$

Many vector lattices satisfy the hypothesis of Lemma 1.3 (e.g. Banach lattices[4]); nevertheless, we will see that some important examples treated in this paper do not.

## §2 ONE DIMENSIONAL MAPS (THE UNIFORMLY HYPERBOLIC CASE).

In this section we study a simple class of uniformly hyperbolic maps using the general approach described in the previous section. The results discussed here are well known [Ru], but this is an ideal starting point for the task at hand.

We consider a special one dimensional map (see Figure 1). Let $I_1$, $I_2$ be two closed intervals such that $I_1\cup I_2=[0,\,1]$, $I_1\cap I_2=\partial I_1\cap\partial I_2$, and $T_i:I_i\to[0,\,1]$, one to one and onto, such that $T_i\in C^{(2)}(I_i)$ and $DT_i\geq\lambda>1$. We consider the map[5]

$$T(x)=\begin{cases}T_1(x) & \text{if } x\in I_1\\ T_2(x) & \text{if } x\in I_2\end{cases}\tag{2.1}$$

We will see how the results in section 1 can be used in order to show that the map (2.1) has a unique invariant measure that is absolutely continuous with respect to the Lebesgue measure. In addition, any absolutely continuous measure, if iterated by the map, will converge to the invariant one at an exponential rate (implying the exponential decay of correlations).

Let us see how the map can be lifted to the measures. Let $f,\,g\in L^2([0,\,1])$

$$\int_0^1 f\circ T(x)g(x)dx=\sum_{i=1}^2\int_0^1 f(x)g\circ T_i^{-1}(x)|D_xT_i^{-1}|dx=\int_0^1 f\widetilde{T}g$$

[4]A Banach lattice $\mathbb{V}$ is defined by the property $\|\,|f|\,\|=\|f\|$ for each $f\in\mathbb{V}$, where $|f|$ is the least upper bound of $f$ and $-f$. For this definition to make sense it is necessary to require that $\mathbb{V}$ is "directed," i.e. any two elements have an upper bound.

[5]The present technique applies to a much wider class of maps (for example any expansive Markov map); in fact, it would be interesting to see the extent to which it can be pushed. Nevertheless, this section has a purely propaedeutic and didactic aim; an attempt to present the results in their full generality would be incompatible with such purpose.

**Figure 1** The map.

The operator $\widetilde{T}$, which describe how the density of the measure $g(x)dx$ is transformed under the iteration of the map $T$, is called Perron-Frobenius operator[6].

The result will follow from the study of the semigroup generated by the operator $\widetilde{T}$. Such a semigroup maps a cone of functions strictly inside itself, and therefore it has a fixed point that corresponds to the density of the unique absolutely continuous invariant measure.

We define the following cone[7]

$$\mathcal{C}_a = \left\{ g \in C^{(0)} \;\middle|\; \forall x,\, y \in [0,\,1] \;\; g(x) > 0;\; \frac{g(x)}{g(y)} \leq e^{a|x-y|} \right\}.$$

**Lemma 2.1.** *Given $\sigma \in (\lambda^{-1},\, 1]$, $\widetilde{T}\mathcal{C}_a \subset \mathcal{C}_{\sigma a}$ provided $a \geq \frac{D}{\sigma - \lambda^{-1}}$, where*

$$D = \sup_{\substack{x \in [0,\,1] \\ i \in \{1,\,2\}}} \left| \frac{D_x^2 T_i^{-1}}{D_x T_i^{-1}} \right| .$$

[6]Other definitions of the Perron-Frobenius operator can be found in the literature, but the differences with the one adopted here are immaterial.

[7]This choice is made in order to simplify the exposition. Essentially, we require the logarithm of the functions in the cone to be Lipschitz, with a Lipschitz constant less than $a$. We could have imposed a condition on the Hölder continuity or the variation of the functions in the cone; this would have produced an estimate for the decay of correlations that is valid for a wider class of functions, and allowed us to treat maps that are $C^{(\alpha)}$, $\alpha > 1$, but not $C^{(2)}$.

*Proof.* Let $g \in \mathcal{C}$, $x \leq y \in [0,\,1]$

$$
\begin{aligned}
\widetilde{T}g(x) &\leq \sum_{i=1}^{2} g(T_i^{-1}(y))|D_xT_i^{-1}|e^{a|T_i^{-1}(x)-T_i^{-1}(y)|} \\
&\leq \sum_{i=1}^{2} g\circ T_i^{-1}(y)|D_yT_i^{-1}|e^{a|T_i^{-1}(x)-T_i^{-1}(y)|+\left|\ln(|D_xT_i^{-1}|)-\ln(|D_yT_i^{-1}|)\right|} \\
&\leq \sum_{i=1}^{2} g\circ T_i^{-1}(y)|D_yT_i^{-1}|e^{a\int_x^y|D_\xi T_i^{-1}|d\xi+\int_x^y\left|\frac{D_\xi^2T_i^{-1}}{D_\xi T_i^{-1}}\right|d\xi} \\
&\leq \widetilde{T}g(y)e^{a\lambda^{-1}|x-y|+D|x-y|} \leq \widetilde{T}g(y)e^{\sigma a|x-y|}.
\end{aligned}
$$

In order to see if we can apply Theorem 1.1, we have to derive the Hilbert metric associated with our cone.

**Lemma 2.2.** *Let $\Theta$ be the Hilbert metric associated to $\mathcal{C}_a$. For each $f,\,g\in\mathcal{C}_a$*

$$
\Theta(f,\,g) = \ln \sup_{\substack{x,\,y\in[0,\,1]\\ u,\,v\in[0,\,1]}} \frac{\left(e^{a|x-y|}g(y)-g(x)\right)\left(e^{a|u-v|}f(v)-f(u)\right)}{\left(e^{a|x-y|}f(y)-f(x)\right)\left(e^{a|u-v|}g(v)-g(u)\right)}
$$

*Proof.* We need to compute $\alpha$ and $\beta$. ¿From the definitions, it follows that $\alpha f \preceq g$ iff

$$
g(x)-\alpha f(x) \geq 0 \qquad \forall x\in[0,\,1]
$$

and

$$
g(x)-\alpha f(x) \leq e^{a|x-y|}\left(g(y)-\alpha f(y)\right)\ \forall x,\,y\in[0,\,1].
$$

Accordingly:

$$
\alpha = \min\left\{\inf_{x\in[0,\,1]}\frac{g(x)}{f(x)};\ \inf_{\substack{x,\,y\in[0,\,1]\\ x\neq y}}\frac{e^{a|x-y|}g(y)-g(x)}{e^{a|x-y|}f(y)-f(x)}\right\}.
$$

However, choosing $x_0\in[0,\,1]$ such that

$$
\frac{g(x_0)}{f(x_0)} = \inf_{x\in[0,\,1]}\frac{g(x)}{f(x)}
$$

then, for each $x\neq x_0$,

$$
\frac{e^{a|x_0-x|}g(x_0)-g(x)}{e^{a|x_0-y|}f(x_0)-f(x)} = \frac{e^{a|x_0-x|}\frac{g(x_0)}{f(x_0)}f(x_0)-\frac{g(x)}{f(x)}f(x)}{e^{a|x_0-y|}f(x_0)-f(x)} \leq \frac{g(x)}{f(x)}.
$$

This means that

$$
\alpha = \inf_{\substack{x,\,y\in[0,\,1]\\ x\neq y}}\frac{e^{a|x-y|}g(y)-g(x)}{e^{a|x-y|}f(y)-f(x)}.
$$

Analogously,

$$
\beta = \sup_{\substack{x,\,y\in[0,\,1]\\ x\neq y}}\frac{e^{a|x-y|}g(y)-g(x)}{e^{a|x-y|}f(y)-f(x)}.
$$

**Lemma 2.3.**
$$\Delta = diam(\mathcal{C}_{\sigma a}) = 2\ln\frac{1+\sigma}{1-\sigma} + 2\sigma a.$$

*Proof.* For each $f,\, g \in \mathcal{C}_{\sigma a}$ we have
$$\begin{aligned}\Theta(f,\, g) &= \ln \sup_{\substack{x,\, y\in[0,\,1]\\ u,\, v\in[0,\,1]}} \frac{\left(e^{a|x-y|} - e^{-\sigma a|x-y|}\right)\left(e^{a|u-v|} - e^{-\sigma a|u-v|}\right) g(y)f(v)}{\left(e^{a|x-y|} - e^{\sigma a|x-y|}\right)\left(e^{a|u-v|} - e^{\sigma a|u-v|}\right) f(y)g(v)}\\ &\le \ln\frac{(1+\sigma)^2}{(1-\sigma)^2}e^{2\sigma a}.\end{aligned}$$

Lemmas 2.1–2.3, together with Theorem 1.1, imply that $\forall f,\, g \in \mathcal{C}_a$
$$\Theta(\widetilde{T}f,\, \widetilde{T}g) \le \Lambda\Theta(f,\, g)$$
with
$$\Lambda = \tanh\left(\frac{\Delta}{4}\right).$$

**Remark 2.4.** *In order to have $\Delta$ as small as possible it is convenient to choose $a = \frac{D}{\sigma-\lambda^{-1}}$. It is then possible to choose the $\sigma = \sigma_*$ that minimizes $\Delta$. Hence, the contraction rate $\Lambda$ depends only on $\lambda$ and $D$.*

Note that, for each $g_1,\, g_2 \in C^{(0)}([0,\,1])$, $-g_1 \preceq g_2 \preceq g_1$ implies $-g_1(x) \le g_2(x) \le g_1(x)$, that is $|g_2(x)| \le g_1(x)$, for each $x \in [0,\,1]$. Consequently, for each $L^p$ norm, $\|g_2\|_p \le \|g_1\|_p$, and Lemma 1.3 applies.

Let $g \in C^{(0)}([0,\,1])$, $\int_0^1 g = 1$, and $g \in \mathcal{C}_{a_*}$ with $a_* = \frac{D}{\sigma_*-\lambda^{-1}}$. Since $\int_0^1 \widetilde{T}^n g = \int_0^1 g = 1$,
$$\|\widetilde{T}^{n+m}g - \widetilde{T}^n g\|_1 \le e^{\Theta(\widetilde{T}^{n+m}g,\widetilde{T}^n g)} - 1 \le e^{\Lambda^{n-1}\Theta(\widetilde{T}(\widetilde{T}^m g),\widetilde{T}g)} - 1 \le e^{\Delta\Lambda^{-1}}\Delta\Lambda^{n-1}.$$
This means that $\{\widetilde{T}^n g\}$ is a Cauchy sequence in $L^1$; in addition, $\{\widetilde{T}^n g\}$ are equicontinuous. Let $\varphi_* = \lim_{n\to\infty}\widetilde{T}^n g$, clearly $\widetilde{T}\varphi_* = \varphi_*$, and $\varphi_* \in \mathcal{C}_{\sigma_* a_*}$, and $\varphi_*$ is not dependent on $g$; i.e., $\varphi_*$ is the density of the unique invariant measure absolutely continuous with respect to Lebesgue.

Up to now, we have shown the existence of the invariant measure, obtained some of its regularity properties, and been able to estimate the decay of correlations for all functions $f \in L^\infty([0,\,1])$, $g \in \mathcal{C}_{a_*}$. The next Theorem gives a slightly stronger result, and shows how to avoid Lemma 1.3.

**Theorem 2.5.** *There exist $K, r \in \mathbb{R}^+$ such that, for each $f \in L^1([0,\,1])$, $g \in C^{(1)}([0,\,1])$ with $\int_0^1 g = 1$*
$$\left|\int_0^1 f\circ T^n g dx - \int_0^1 f\varphi_* dx\right| \le K\|f\|_1(\|g\|_1 + r\|g'\|_\infty)\Lambda^n$$

*Proof.* Consider $g \in \mathcal{C}_{a_*}$, normalized so that $\int_0^1 g = 1$ (i.e., $g$ can be thought as the density of a measure). Then,
$$\left|\int_0^1 f(\widetilde{T}^n g - \varphi_*)dx\right| \le \|f\|_1\left\|\frac{\widetilde{T}^n g}{\varphi_*} - 1\right\|_\infty \|\varphi_*\|_\infty.$$

**Lemma 2.6.** *If $\mathcal{C}_+ = \{g \in C^{(0)}([0,1]) \mid g(x) \geq 0, \quad \forall x \in [0,1]\}$, and $\Theta_+$ is the corresponding Hilbert metric, then, for each $g_1, g_2 \in \mathcal{C}_a$,*

$$\Theta_+(g_1, g_2) \leq \Theta(g_1, g_2)$$

*Proof.* Since $\mathcal{C}_+ \supset \mathcal{C}_a$, the identity is a map from $C^{(0)}([0,1])$ to itself that maps $\mathcal{C}_a$ into $\mathcal{C}_+$. The result follows then from Theorem 1.1.

A simple computation yields

$$\Theta_+(g_1, g_2) = \ln \sup_{x, y \in [0,1]} \frac{g_1(x) g_2(y)}{g_1(y) g_2(x)}.$$

Using the previous facts, and the trivial equality

$$\frac{(\widetilde{T}^n g)(x)}{\varphi_*(x)} = \frac{\widetilde{T}^n g(x) \varphi_*(y)}{\widetilde{T}^n g(y) \varphi_*(x)} \frac{\widetilde{T}^n g(y)}{\varphi_*(y)},$$

we have

$$e^{-\Theta_+(\widetilde{T}^n g, \varphi_*)} \frac{\widetilde{T}^n g(y)}{\varphi_*(y)} \leq \frac{\widetilde{T}^n g(x)}{\varphi_*(x)} \leq e^{\Theta_+(\widetilde{T}^n g, \varphi_*)} \frac{\widetilde{T}^n g(y)}{\varphi_*(y)}$$

for each $x, y \in [0,1]$. Because $\int_0^1 (\widetilde{T}^n g - \varphi_*) = 0$, and remembering that both functions are continuous, there exists $y_n \in [0,1]$: $\widetilde{T}^n g(y_n) = \varphi_*(y_n)$. Using the previous inequalities, with $y = y_n$, we obtain, for each $x \in [0,1]$,

$$e^{-\Theta_+(\widetilde{T}^n g, \varphi_*)} \leq \frac{\widetilde{T}^n g(x)}{\varphi_*(x)} \leq e^{\Theta_+(\widetilde{T}^n g, \varphi_*)}$$

and

$$\left\| \frac{\widetilde{T}^n g}{\varphi_*} - 1 \right\|_\infty \leq e^{\Theta_+(\widetilde{T}^n g, \varphi_*)} - 1 \leq e^{\Lambda^{n-1} \Theta(\widetilde{T} g, \varphi_*)} - 1 \leq e^{\Lambda^{-1}\Delta} \Delta \Lambda^{n-1}.$$

This estimate shows that for each $f \in L^1([0,1])$, $g \in \mathcal{C}_{a_*}$

$$\left| \int_0^1 f \circ T^n g dx - \int_0^1 f \varphi_* dx \int_0^1 g dx \right| \leq K \|f\|_1 \|g\|_1 \Lambda^n$$
$$K = e^{\Lambda^{-1}\Delta} \Delta \Lambda^{-1} \|\varphi_*\|_\infty.$$

Let us now consider $g \in C^{(0)}([0,1])$, $g \geq 0$, and piecewise differentiable with $\|g'\|_\infty < \infty$. If we define $g_\rho = g + \rho \varphi_*$, $\rho \in \mathbb{R}^+$, for each $x > y$ we have

$$\begin{aligned} g_\rho(x) &\leq e^{\ln[g(x) + e^{\sigma_* a_* (x-y)} \rho \varphi_*(y)] - \ln[g(y) + \rho \varphi_*(y)]} g_\rho(y) \\ &\leq e^{\int_y^x \frac{|g'(\xi) + \rho \sigma_* a_* e^{\sigma_* a_* (\xi - y)} \varphi_*(y)|}{g(\xi) + \rho e^{\sigma_* a_* (\xi - y)} \varphi_*(y)} d\xi} g_\rho(y) \leq e^{\left[\frac{\|g'\|_\infty}{\varphi_*(y)} \rho^{-1} + \sigma_* a_*\right](x-y)} g_\rho(y). \end{aligned}$$

Choosing,

$$\rho = \sup_{x \in [0,1]} \frac{\|g'\|_\infty}{\varphi_*(x) a_* (1 - \sigma_*)}$$

we have $g_\rho \in \mathcal{C}_{a_*}$; hence, the decay of correlations for $g_\rho$ implies the decay of correlations for $g$. The result for arbitrary $g \in C^{(1)}([0,1])$ follows since any smooth function can be written as the difference of two piecewise differentiable positive functions.

Theorem 2.5 immediately yields

**Corollary 2.7 (Mixing).** *For each* $f \in L^\infty([0,1])$, $g \in L^1([0,1])$, $\int_0^1 g = 1$, *we have*

$$\lim_{n\to\infty} \left| \int_0^1 (f \circ T^n)\, g - \int_0^1 f\varphi_* \right| = 0.$$

Another corollary of Theorem 2.5 is the so called Central Limit Theorem [Ke].

**Corollary 2.8 (CLT).** *For each* $f \in C^1([0,1])$, $\int_0^1 f\varphi_* = 0$, *the distribution of the random variable*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n f \circ T^n$$

*converges to a Gaussian law.*

## §3 ONE DIMENSIONAL MAPS (RANDOM PERTURBATIONS).

In this section we will investigate random perturbations of the example discussed in the previous section. Again, the argument presented holds for more general examples (at least for random perturbations of Markov maps), but I will ignore this, since my primary goal is to present the method.

We will consider two types of questions: first, establishing results about the invariant measure and the rate of mixing; second, estimating the difference between the invariant measure in the deterministic case and the invariant measure in the stochastic case.

Let $\phi : [0,1] \times \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous function such that $\int_0^\infty \phi(x,\xi)d\xi = 1$ for each $x \in [0,1]$, $\phi(x,\xi) \ge \varphi(\xi) > 0$ for each $x \in [0,1]$, $\varphi$ monotone decreasing, and $\phi_+ = \sup \phi < \infty$. Letting $z_\varepsilon(x) = \varepsilon^{-1} \int_0^1 \phi(x, \varepsilon^{-1}|x-\xi|)d\xi$, we define the following Markov process

$$\widehat{T}_\varepsilon f(x) = \frac{1}{\varepsilon z_\varepsilon(Tx)} \int_0^1 \phi(Tx, \varepsilon^{-1}|Tx-y|) f(y) dy.$$

The corresponding P-F operator is given by

$$\widetilde{T}_\varepsilon g(x) = \varepsilon^{-1} \int_0^1 z_\varepsilon(Ty)^{-1} \phi(Ty, \varepsilon^{-1}|Ty-x|) g(y) dy.$$

We want to study the operator $\widetilde{T}_\varepsilon$.

¿From a purely probabilistic point of view, it makes sense to consider the cone $\mathcal{C}_+$ (see Lemma 2.6). Clearly, $\widetilde{T}_\varepsilon \mathcal{C}_+ \subset \mathcal{C}_+$. In addition,

$$\widetilde{T}_\varepsilon g(x) \ge \varphi(\varepsilon^{-1}) \|g\|_1,$$

and, by choosing $\varepsilon$ so small that $\varepsilon z_\varepsilon(x) \ge \frac{1}{2}$ for each $x \in [0,1]$,

$$\widetilde{T}_\varepsilon g(x) \le 2\phi_+ \|g\|_1.$$

These estimates imply that the diameter $\Delta$ of $\widetilde{T}_\varepsilon \mathcal{C}_+$ is bounded by

$$\Delta \leq 2\ln\frac{\|g\|_+}{\|g\|_-} \leq 2\ln\frac{\phi_+}{\varphi(\varepsilon^{-1})}$$

so we have a unique smooth invariant measure, which is mixing.

The previous proof is extremely simple, especially because we have completely obliterated the dynamics. But we paid a heavy price: we have very little information about the invariant measure, and the estimate for the decay of correlations has an unreasonably bad dependence upon $\varepsilon$.

If we want better results we have to take into account explicitly the dynamics of the underlying deterministic system. We will do so in a slightly different model.

First we will identify $[0, 1]$ with the unit circle $S^1$. This is done to eliminate the boundary effects. In fact, the presence of the boundary tends to change considerably the derivative of the density of the invariant measure in a small neighborhood of the boundary, although it does not change substantially its $C^{(0)}$ norm. This would force us to consider a more complex cone taking such behavior into account (of course, any cone that works in our approach must contain the invariant density), making our life harder, without substantially adding to the argument. From now on, $|\cdot|$ will mean the distance in $S^1$; nevertheless we will continue writing the set as $[0, 1]$.

Let us require $\phi(x, \xi) \geq 0$, $\phi(x, \xi) = 0$ for each $\xi \geq 1$, $x \in [0, 1]$ (this hypothesis is not essential, but we use it to simplify the argument), and

$$\frac{\phi(x, \xi)}{\phi(y, \xi)} \leq e^{\theta|x-y|}$$

(a sort of weak translation invariance).

This time we will use the cone $\mathcal{C}_a$ (see beginning of §2). For $x \in S^1$,

$$\begin{aligned}\widetilde{T}_\varepsilon g(x) &= \varepsilon^{-1}\int_0^1 \phi(Ts, \varepsilon^{-1}|Ts - x|)g(s)ds \\ &= \sum_{i=1}^{2}\int_0^1 \phi(s, \varepsilon^{-1}|s - x|)g(T_i^{-1}s)|D_sT_i^{-1}|ds \\ &= \int_0^1 \phi(s, \varepsilon^{-1}|s - x|)(\widetilde{T}g)(s)ds.\end{aligned}$$

It is important to notice that, since we are now on a circle, we can assume, without loss of generality, that $x$ is far away from 0 or 1 (which, on the circle, are the same arbitrary point). From the previous section we know that, for $a$ sufficiently big, there exist $\sigma < 1$, such that $\widetilde{T}g \in \mathcal{C}_{\sigma a}$. Consequently, for each $y \in [0, 1]$ such that $|x - y| \ll 1$, we have

$$\begin{aligned}\widetilde{T}_\varepsilon g(x) &= \int_0^1 \phi(s, \varepsilon^{-1}|(s - x + y) - y|)(\widetilde{T}g)(s)ds \\ &\leq \int_0^1 \phi(s - y + x, \varepsilon^{-1}|(s - y|)(\widetilde{T}g)(s - y + x)ds \leq e^{(\sigma a+\theta)|x-y|}\widetilde{T}_\varepsilon g(y).\end{aligned}$$

This implies that, for $a$ large enough, there exists $\sigma' < 1$ such that $\widetilde{T}_\varepsilon \mathcal{C}_a \subset \mathcal{C}_{\sigma' a}$, which suffices to obtain a unique smooth invariant measure and an exponential rate for the decay of correlations. Let us call $\varphi_*^\varepsilon$ the invariant measure associated to $\widetilde{T}_\varepsilon$, and let $\varphi_*^0 = \varphi_*$ be the invariant measure associated with the deterministic map). The proof of the next theorem is based on an argument in [Ry].

**Theorem 3.1.** *For each stochastic process defined by a function $\phi$ with the above mentioned properties, there exists a constant $K_0 \in \mathbb{R}^+$ such that*

$$\|\varphi_* - \varphi_*^\varepsilon\|_\infty \leq K_0 \varepsilon \ln \varepsilon^{-1}.$$

*Proof.* We start by noticing that, for each function $g \in \mathcal{C}_a$,

$$\left|\widetilde{T}_\varepsilon g(x) - \widetilde{T} g(x)\right| \leq (e^{a\varepsilon} - 1)\widetilde{T}_\varepsilon g(x).$$

Let $u(x) = (\widetilde{T}\varphi_*^\varepsilon)(x) - \varphi_*^\varepsilon(x)$. On the one hand,

$$\|u\|_\infty = \|\widetilde{T}\varphi_*^\varepsilon - \widetilde{T}_\varepsilon \varphi_*^\varepsilon\|_\infty \leq (e^{a\varepsilon} - 1)\|\varphi_*^\varepsilon\|_\infty \leq e^{2a}\varepsilon,$$

and

$$\|u'\|_\infty \leq 2e^{2a}.$$

On the other hand

$$(1\!\!1 - \widetilde{T})(\varphi_* - \varphi_*^\varepsilon) = \widetilde{T}\varphi_*^\varepsilon - \varphi_*^\varepsilon = u.$$

Since, $\int_0^1 u = 0$ Theorem 2.5 implies that

$$\|\widetilde{T}^n \left(\widetilde{T} u\right) \|_\infty \leq K\Lambda^n(\|u\|_1 + r\|u'\|_\infty).$$

Hence, we can write

$$\varphi_*^\varepsilon - \varphi_* = \sum_{n=0}^{\infty} \widetilde{T}^n u.$$

¿From the latter equality we have

$$\|\varphi_*^\varepsilon - \varphi_*\|_\infty \leq Ne^{2a}\varepsilon + \sum_{n=N}^{\infty} \Lambda^n(\varepsilon + 2r)e^{2a} \leq Ne^{2a}\varepsilon + (1 - \Lambda)^{-1}\Lambda^N(\varepsilon + 2r)e^{2a}$$

which yields the result by choosing $\Lambda^N = \varepsilon$.

As a consequence of the above results we can reach some conclusions about the use of computers to compute Lyapunov exponents. It is commonly thought that the simulation of a dynamical system by a computer introduces some randomness into the system due to roundoff errors. This is obviously false from a rigorous point of view (since the algorithms run by computers are deterministic), but it is likely to be a reasonable approximation.

If we consider such a model, then two natural questions are: how good are the computed results, and which length of trajectories is it reasonable to consider in the computations.

Let $\Omega = [0, 1]^{\mathbb{N}}$ be the space of the trajectories of the random system, and $\mathbb{E}$ the expectation induced by the measure; what we can evaluate by using a computer is

$$\lambda_n(\omega) = \frac{1}{n}\sum_{i=0}^{n} \ln |D_{(\sigma^i\omega)_0}T|$$

where $\omega \in \Omega$ is the trajectory, and $\sigma(\omega)_i = \omega_{i+1}$. By ergodicity,

$$\lambda_\varepsilon = \lim_{n\to\infty} \lambda_n(\omega) = \int_0^1 \ln |D_z T| \varphi_*^\varepsilon(z) dz.$$

If $\lambda$ is the Lyapunov exponent of the deterministic system, then

$$|\lambda - \lambda_\varepsilon| \leq K_0 \varepsilon \ln \varepsilon^{-1}.$$

In addition, by the CLT,

$$\mathbb{E}\left(|\lambda_n(\omega) - \lambda_\varepsilon|\right) \leq \frac{\text{constant}}{\sqrt{n}}.$$

Accordingly, if $\varepsilon$ is the precision with which we are carrying out the computations, it is unreasonable to expect a precision better than $\varepsilon \ln \varepsilon^{-1}$, or to hope for an improvement of the result by looking at trajectories longer than $(\varepsilon \ln \varepsilon^{-1})^{-2}$.

## §4 TWO DIMENSIONS (THE SMOOTH UNIFORMLY HYPERBOLIC CASE).

In the rest of the paper we will see how the approach presented in section two can be implemented in more general situations.

Clearly the first possible generalization is the multidimensional case. For simplicity, we will confine ourselves to the two dimensional area–preserving case. Hence, this section is dedicated to symplectic smooth uniformly hyperbolic two dimensional maps.

The symplectic nature of the maps implies the existence of a special invariant measure (i.e. the symplectic volume), which simplifies our approach. The higher dimensional conservative case can be treated in the same fashion. On the contrary, in the case of dissipative systems the present technique should still yield results, but more care would be needed in its application.

The main idea is that in one direction the P-F operator acts as in the one dimensional case, i.e. it makes functions smoother, but in the other direction the P-F operator increases oscillations. To overcome this phenomenon we integrate the functions on short curves in the latter direction. Under the action of the P-F operator, this is equivalent to averaging the functions on longer and longer curves. Hence, also in the second direction a regularization is taking place.[8]

The previous argument will be made precise by Lemma 4.2. The Hilbert metric used in this section is weaker than any $L^p$ norm (in particular Lemma 1.3 does not hold); nonetheless, according to Theorem 4.9, it suffices to control the decay of correlation for Hölder continuous functions.

We will start by describing the class of maps under consideration and by recalling some standard constructions. We will then introduce the cones of functions used in the proof.

**The map.**

Let $\mathcal{M}$ be a compact smooth symplectic manifold and $T : \mathcal{M} \to \mathcal{M}$ a symplectomorphysm. Let there exist a bundle of cones $\{\mathcal{C}(x)\}_{x\in\mathcal{M}}$ in the tangent bundle (i.e., $\mathcal{C}(x) \subset \mathcal{T}_x\mathcal{M}$) strictly invariant (i.e., $D_xT^{-1}\mathcal{C}(x) \subset \operatorname{int} \mathcal{C}(T^{-1}x) \cup \{0\}$) and continuous (i.e., the "edges" of the cones vary with continuity). Then the map $T$ is uniformly hyperbolic and its stable and unstable foliations are continuous (conversely, for such maps, it is easy to construct a family of continuous strictly invariant cones with the required properties; for more details see [LW]).

It follows that there exists a smooth metric on $\mathcal{M}$ such that, calling $v^s(x)$, $v^u(x)$ the stable and unstable normalized vectors at $x$

$$
\begin{aligned}
\|D_xTv^s(x)\| &\leq \lambda^{-1} \\
\|D_xTv^u(x)\| &\geq \lambda
\end{aligned}
$$

for some $\lambda > 1$. Furthermore, the Riemannian volume is equal to the symplectic one. We will, from now on, use this metric. In addition, we will suppose that

$$
\begin{aligned}
&\inf_{v\in\mathcal{C}(x)} \|D_xT^{-1}v\| \geq \lambda \\
&\sup_{\substack{v,\, w\in\mathcal{C}(x) \\ \|v\|=\|w\|=1}} \|v-w\| \leq \frac{1}{2}.
\end{aligned}
$$

(This can always be achieved, eventually by choosing a smaller $\lambda$ and by replacing $\mathcal{C}(x)$ by $D_{T^nx}T^{-n}\mathcal{C}(T^nx)$ for a large enough $n$).

Further, we assume that the map enjoys the mixing property.[9] In the present case the Perron-Frobenius operator $\widetilde{T} : L^1(\mathcal{M}) \to L^1(\mathcal{M})$ is given by

$$
\widetilde{T}g = g \circ T^{-1}.
$$

**Cones of functions.**

[8] The necessity to take averages comes from the fact that, in the multidimensional case, measures converge to the invariant one only weakly, contrary to what we have seen in the one dimensional case.

[9] The previous assumptions imply that the map is locally ergodic [BG],[LW], [KB], and mixing on each ergodic component [GO], [P], but several ergodic components could still be present.

Due to the presence of both contracting and expanding directions the approach of the previous sections must be slightly modified. To overcome this difficulty we will, in some sense, quotient away (or, better yet, "integrate away") the unwanted direction through the introduction of suitable test functions. This is achieved by introducing a class of closed, connected and differentiable curves

$$\Gamma_\delta = \{\gamma \subset \mathcal{M} \mid \delta \leq \text{length}(\gamma) \leq 2\delta;\ \forall x \in \gamma,\ \gamma'(x) \subset \mathcal{C}(x);\ |\kappa_\gamma(x)| \leq H\}$$

where $\gamma'(x)$ is the unitary tangent vector to $\gamma$ at $x$, $\kappa_\gamma(x)$ is the curvature of $\gamma$ at $x$, and $\delta \in (0, 1)$.[10] On each of these curves, the measure $m_\gamma$, induced by the Riemannian structure (arc-length), is naturally defined. The plan is to perform integrals on $\gamma$, with respect to a large class of measures. We proceed to characterize such class by specifying the properties of their densities with respect to $m_\gamma$.

Let us define the cones of "density" functions. Given any connected smooth curve $\gamma$, let

$$\widetilde{\mathcal{D}}_a(\gamma) = \left\{ f \in C^{(0)}(\gamma) \mid f > 0;\ \frac{f(x)}{f(y)} \leq e^{a d(x,\, y)^\nu} \right\}$$

where $d$ means the distance along $\gamma$, and $\nu \in (0, \frac{1}{2}]$. In other words, we are considering positive functions whose logarithm has a bounded modulus of $\nu$-Hölder continuity.

The reader will certainly notice the similarity with the cones used in the one dimensional case. Let us remember that

$$\rho_\gamma(f_1,\, f_2) = \ln\left[\sup_{x,\, y,\, u,\, v \in \gamma} \frac{\left[e^{ad(x,\, y)^\nu} f_1(x) - f_1(y)\right]\left[e^{ad(u,\, v)^\nu} f_2(u) - f_2(v)\right]}{\left[e^{ad(x,\, y)^\nu} f_2(x) - f_2(y)\right]\left[e^{ad(u,\, v)^\nu} f_1(u) - f_1(v)\right]}\right]$$

is the Hilbert metric in $\widetilde{\mathcal{D}}_a(\gamma)$.

The set of functions we will consider is

$$\mathcal{D}_a(\gamma) = \{f \in \widetilde{\mathcal{D}}_a(\gamma) \mid \rho_\gamma(1,\, f) \leq R\}.$$

For later purposes we also define

$$\mathcal{D}_+(\gamma) = \{f \in C^{(0)}(\gamma) \mid f \geq 0\},$$

together with its Hilbert metric

$$\rho_+(f_1,\, f_2) = \ln\left[\sup_{x,\, y \in \gamma} \frac{f_1(x) f_2(y)}{f_1(y) f_2(x)}\right].$$

Of course, by Theorem 1.1, $\rho_+ \leq \rho_\gamma$ (since the inclusion of $\widetilde{\mathcal{D}}_a(\gamma)$ in $\mathcal{D}_+(\gamma)$ is a contraction).

[10] To fix a length scale, let us assume $\text{diam}(\mathcal{M}) = 1$.

It is now possible to introduce the cone of functions used in the proof. We introduce two fundamental quantities relative to a given function $g \in C^{(0)}(\mathcal{M})$

$$
\begin{aligned}
|||g|||_+ &= \sup_{\substack{\gamma\in\Gamma_\delta \\ f\in\mathcal{D}_a(\gamma)}} \frac{\int_\gamma fg}{\int_\gamma f} \\
|||g|||_- &= \inf_{\substack{\gamma\in\Gamma_\delta \\ f\in\mathcal{D}_a(\gamma)}} \frac{\int_\gamma fg}{\int_\gamma f}.
\end{aligned}
$$

(Here, as in the following, the integration is done with respect to the measure $m_\gamma$).

The cone of functions is defined as follows:

$$
\mathcal{C}_{b,\,c} = \Bigg\{ g \in C^{(0)}(\mathcal{M}) \;\Bigg|\; \forall\gamma\in\Gamma_\delta,\ \forall x\in\gamma,\ \forall f,\ f_1,\ f_2 \in \mathcal{D}_a(\gamma)\ \int_\gamma gf > 0; \\
\frac{f_2(x)\int_\gamma gf_1}{f_1(x)\int_\gamma gf_2} \le e^{b\rho_\gamma(f_1,\,f_2)};\ \frac{\|D^u g\|_\infty}{|||g|||_-} \le c \Bigg\}
$$

where, by $D^u$ we mean the derivative in the unstable direction. We require that $D^u g$ is well defined, apart from, at most, countably many points, on each unstable manifold (the reader may feel free to think that $g$ is everywhere differentiable: we will use our weaker assumption only at the very end of Theorem 4.9, and not in an essential way).

The choice of piecewise smooth functions, in the unstable direction, is done only to simplify the exposition. More general functions can easily be considered by defining larger cones.[11] As anticipated, once we integrate along the "problematic" direction, we can consider a cone quite similar to the previous ones.

**Contraction of the Hilbert metric.**

The reader may wonder about how big the cone $\mathcal{C}_{b,\,c}$ is. The next Lemma shows that it is far from empty.

**Lemma 4.1.** *For each $b > 1$ the cone $\mathcal{C}_{b,\,c}$ is a closed convex cone with non empty interior in the $C^{(1)}$ topology.*

*Proof.* The closure and the convexity can be easily checked by direct computation. The third property follows since the open set $\{g \in C^{(1)}(\mathcal{M}) \mid g > p \ge 0;\ \|Dg\|_\infty <$

[11] If $\gamma$ and $\widetilde{\gamma}$ are two curves in $\Gamma_\delta$, obtained one from the other by "translation" along the unstable direction and $\phi$ is the canonical isomorphism from $\widetilde{\gamma}$ to $\gamma$ along the unstable direction, one can replace the last condition in $\mathcal{C}_{b,c}$ by

$$
\frac{\int_\gamma gf}{\int_{\widetilde{\gamma}} gf\circ\phi} \le e^{c\tau^\sigma}
$$

for some $\sigma \in (0,\,1)$ and where $\tau$ is the distance between $\gamma$ and $\widetilde{\gamma}$.

$cp\}$ is contained in $\mathcal{C}_{b,\,c}$. In fact,

$$\begin{aligned} f_2(x)\int_\gamma gf_1 =& f_1(x)\int_\gamma g(y)f_2(y)\frac{f_2(x)f_1(y)}{f_1(x)f_2(y)}dm_\gamma(y)\leq e^{\rho_+(f_1,\,f_2)}f_1(x)\int_\gamma gf_2\\ \leq& e^{\rho_\gamma(f_1,\,f_2)}f_1(x)\int_\gamma gf_2. \end{aligned}$$

By now, it is clear that we need the cone to be invariant under the action of the P-F operator. Lemma 4.2 shows that this is the case.

**Lemma 4.2.** *For $a$, $b$ and $H$ large enough, $\widetilde{T}\mathcal{C}_{b,\,c}\subset\mathcal{C}_{\chi b,\,\chi c}$ for some $\chi<1$.*

*Proof.* We check the first property first. Let $\gamma\in\Gamma_\delta$, $x\in\gamma$ and $f_1,\,f_2\in\mathcal{D}_a(\gamma)$ then, for each $g\in\mathcal{C}_{b,\,c}$,

$$\int_\gamma(\widetilde{T}g)f_1=\int_\gamma g\circ T^{-1}f_1=\int_{T^{-1}(\gamma)}gf_1\circ T\left|\det(D_xT\big|_{T^{-1}\gamma})\right|.$$

By $D_xT\big|_{T^{-1}\gamma}$ we mean the restriction of $D_xT$ to a map from the direction $(T^{-1}\gamma)'$ to the direction $\gamma'$. Moreover, we use $\det(\cdot)$, in spite of the fact that, in our case, is superfluous, to emphasize the multidimensional nature of the argument.

**Sub-Lemma 4.3.** *For each $\gamma\in\Gamma_\delta$, there exists a collections of curves $\{\gamma_i\}\in\Gamma_\delta$, such that $\gamma_i\cap\gamma_j=\partial\gamma_i\cap\partial\gamma_j$ and $\cup_i\gamma_i=T^{-1}\gamma$, provided $H$ has been chosen large enough.*

*Proof.* We start by subdividing $T^{-1}(\gamma)$ in connected pieces of length $\delta$ plus, at most, one piece longer than $\delta$ but shorter than $2\delta$. Since $\gamma'(x)\in\mathcal{C}(x)$, and $D_xT^{-1}\mathcal{C}(x)\subset\mathcal{C}(T^{-1}x)$, we have $\left(T^{-1}\gamma\right)'(T^{-1}x)\in\mathcal{C}(T^{-1}x)$. We are left with the condition on the curvature.

Let $\widetilde{\gamma}(s)$ and $\widetilde{\gamma}_1(s_1)$ be the curves $\gamma$ and $T^{-1}\gamma$ parametrized by arc-length. If $\eta$ is the unit vector in the direction $\widetilde{\gamma}_1''$, then $\langle\eta,\,DT^{-1}\gamma'\rangle=0$ and $|\kappa_{T^{-1}\gamma}(s_1)|=|\langle\eta,\,\frac{d}{ds_1}\widetilde{\gamma}_1'(s_1)\rangle|$. Since $\frac{ds_1}{ds}=\|DT^{-1}\gamma'\|$, it follows

$$\begin{aligned} |\kappa_{T^{-1}\gamma}(s_1)|\leq&\lambda^{-1}|\langle\eta,\,\frac{d}{ds}\frac{D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}'(s)}{\|D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}'(s)\|}\rangle|\leq\lambda^{-2}|\langle\eta,\,\frac{d}{ds}D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}'(s)\rangle|\\ \leq&\lambda^{-2}(c_2+|\langle\eta,\,D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}''(s)\rangle|) \end{aligned}$$

where $c_2$ depends only on the second derivative of $T$. To conclude the estimate, let us fix $s$ and choose explicit orthogonal coordinates at $\widetilde{\gamma}(s)$ and $T^{-1}\widetilde{\gamma}(s)$. Namely, we take the axes in the directions $\widetilde{\gamma}'$, $\widetilde{\gamma}''$, and $\widetilde{\gamma}_1'$, $\widetilde{\gamma}_1''$, respectively. In these coordinates $DT^{-1}$ has the form

$$DT^{-1}=\begin{pmatrix}a&b\\0&c\end{pmatrix}$$

with $a=\|D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}'(s)\|\geq\lambda$. In addition, since the Riemannian volume equals the symplectic one, $\det(DT^{-1})=1$, i.e., $a=c^{-1}$. Using such a representation yields

$$|\langle\eta,\,D_{\widetilde{\gamma}(s)}T^{-1}\widetilde{\gamma}''(s)\rangle|\leq(0\ 1)\begin{pmatrix}a&b\\0&a^{-1}\end{pmatrix}\begin{pmatrix}0\\|\kappa_\gamma(s)|\end{pmatrix}=|\kappa_\gamma(s)|a^{-1},$$

or

$$|\kappa_{T^{-1}\gamma}(s_1)| \leq \lambda^{-2}(c_2 + \lambda^{-1}H) \leq H.$$

It is then natural to write

$$\int_\gamma (\widetilde{T}g) f_1 = \sum_i \int_{\gamma_i} g f_1 \circ T |\det(DT\big|_{\gamma_i})|.$$

To continue we need to analyze the maps $\widehat{T}_i : C^{(0)}(\gamma) \to C^{(0)}(\gamma_i)$ defined by

$$(\widehat{T}_i f)(x) = f \circ T(x) \left|\det(D_x T\big|_{\gamma_i})\right| \qquad \forall x \in \gamma_i. \tag{4.1}$$

¿From now on, with a slight abuse of notations, we will drop the subscript "$i$", since the range of $\widehat{T}$ is always clear from the context.

**Sub-Lemma 4.4.** *$\widehat{T}(\widetilde{\mathcal{D}}_a(\gamma)) \subset \widetilde{\mathcal{D}}_{\sigma a}(\gamma_i)$ for some $\sigma < 1$, provided $a > c_3\delta^{1-\nu}(1 - \lambda^{-\nu})^{-1}$, where $c_3$ depends only on the system and $H$ (see definition of $\Gamma_\delta$).*

*Proof.* For each $f \in \widetilde{\mathcal{D}}_a(\gamma)$, $x,\, y \in \gamma_i \in \Gamma_\delta$

$$\frac{(\widehat{T}f)(x)}{(\widehat{T}f)(y)} \leq e^{ad(Tx,\, Ty)^\nu + \mu(c_2+\mu H)d(x,\, y)} \leq e^{(a\lambda^{-\nu} + c_3\delta^{1-\nu})d(x,\, y)^\nu},$$

where $\mu$ is the maximal expansion coefficient.

**Sub-Lemma 4.5.** *If $R$ is sufficiently large, then $\widehat{T}(\mathcal{D}_a(\gamma)) \subset \mathcal{D}_a(\gamma_i)$; in addition for each $f_1,\, f_2 \in \mathcal{D}_a(\gamma)$*

$$\rho_{\gamma_i}(\widehat{T}f_1,\, \widehat{T}f_2) \leq \xi \rho_\gamma(f_1,\, f_2)$$

*for some $\xi < 1$.*

*Proof.* The constant function 1 is an element of $\mathcal{D}_a(\gamma_i)$; for each $f \in \widetilde{\mathcal{D}}_{\sigma a}(\gamma_i)$ we have

$$\begin{aligned}
\rho_{\gamma_i}(1,\, f) &= \ln\left[\sup_{x,\, y,\, u,\, v \in \gamma_i} \frac{\left[e^{ad(x,\, y)^\nu} - 1\right]\left[e^{ad(u,\, v)^\nu} f(u) - f(v)\right]}{\left[e^{ad(x,\, y)^\nu} f(x) - f(y)\right]\left[e^{ad(u,\, v)^\nu} - 1\right]}\right] \\
&\leq \ln\left[\sup_{x,\, y,\, u,\, v \in \gamma_i} \frac{\left[e^{ad(x,\, y)^\nu} - 1\right]\left[e^{a(1+\sigma)d(u,\, v)^\nu} - 1\right] f(v)}{\left[e^{a(1-\sigma)d(x,\, y)^\nu} - 1\right]\left[e^{ad(u,\, v)^\nu} - 1\right] f(y)}\right] \\
&\leq \ln\left[\frac{(1+\sigma)}{(1-\sigma)} e^{a\sigma 2^\nu \delta^\nu}\right] \leq R.
\end{aligned}$$

Hence, $\widehat{T}(\mathcal{D}_a(\gamma)) \subset \mathcal{D}_a(\gamma_i)$. In addition, since the diameter of $\widehat{T}_i(\widetilde{\mathcal{D}}_a(\gamma))$ in $\widetilde{\mathcal{D}}_a(\gamma_i)$ is less than $R$, the second part of the result follows from Theorem 1.1.

Accordingly,

$$\int_\gamma \widetilde{T} g f = \sum_i \int_{\gamma_i} g \widehat{T} f \geq 0.$$

Next, let $x \in \gamma$, and choose points $z_i \in T\gamma_i$;

$$\begin{aligned}
f_2(x)\int_\gamma (\widetilde{T}g)f_1 =&\sum_i \frac{f_2(x)}{\widehat{T}f_2(z_i)}\widehat{T}f_2(z_i)\int_{\gamma_i} g\widehat{T}f_1\\
\leq&\sum_i e^{b\rho_{\gamma_i}(\widehat{T}f_1,\,\widehat{T}f_2)}\frac{f_2(x)}{\widehat{T}f_2(z_i)}\widehat{T}f_1(z_i)\int_{\gamma_i} g\widehat{T}f_2\\
\leq& e^{\xi b\rho_\gamma(f_1,\,f_2)}\sup_{z,\,z'\in\gamma}\frac{f_2(z)f_1(z')}{f_1(z)f_2(z')}f_1(x)\sum_i\int_{\gamma_i} g\widehat{T}f_2\\
\leq& e^{[\xi b\rho_\gamma(f_1,\,f_2)+\rho_+(f_1,\,f_2)]}f_1(x)\sum_i\int_{T\gamma_i}\widetilde{T}gf_2\\
\leq& e^{(\xi b+1)\rho_\gamma(f_1,\,f_2)}f_1(x)\int_\gamma(\widetilde{T}g)f_2
\end{aligned}$$

which yields the wanted result for $b$ large enough. The third of the conditions is easily checked after noticing that

$$\frac{\int_\gamma(\widetilde{T}g)f}{\int_\gamma f}\geq\sum_i\frac{\int_{\gamma_i}\widehat{T}_i f}{\int_\gamma f}|||g|||_- = |||g|||_-$$

implies $|||\widetilde{T}g|||_- \geq |||g|||_-$.

We know now that the cone is mapped inside itself, yet this does not necessarily means that the diameter of the image if finite. To check this last property we start with a preliminary result.

**Lemma 4.6.** *Setting $K = \frac{(1+\chi)(1+b)}{(1-\chi)(b-1)}$, for each $g \in \mathcal{C}_{\chi b,\chi c}$,*

$$\Theta(g,\,1)\leq\ln\left[K\frac{|||g|||_+}{|||g|||_-}\right]$$

*(from now on $\Theta$ designates the Hilbert metric in $\mathcal{C}_{b,\,c}$).*

*Proof.* For each $\lambda > 0$, such that $g-\lambda\in\mathcal{C}_{\chi b,\chi c}$, a direct computation shows

$$\begin{aligned}
\lambda\leq&\inf_{\substack{\gamma\in\Gamma_\delta\\ f\in\mathcal{D}_a(\gamma)}}\frac{\int_\gamma gf}{\int_\gamma f}=\alpha_0,\\
\lambda\leq&\inf_{\substack{\gamma\in\Gamma_\delta\\ f_1,\,f_2\in\mathcal{D}_a(\gamma)}}\frac{e^{b\rho_\gamma(f_1,\,f_2)}f_1\int_\gamma gf_2 - f_2\int_\gamma gf_1}{e^{b\rho_\gamma(f_1,\,f_2)}f_1\int_\gamma f_2 - f_2\int_\gamma f_1}=\alpha_1,\\
\lambda\leq&\frac{c|||g|||_- - \|D^u g\|_\infty}{c}=\alpha_2,
\end{aligned}$$

which implies $\alpha = \min\{\alpha_0,\,\alpha_1,\,\alpha_2\}$ (see §1 for a definition of $\alpha$ in this context). In addition,

$$\begin{aligned}
\alpha_0 =& |||g|||_-,\\
\alpha_1\geq&\inf_{\substack{\gamma\in\Gamma_\delta\\ f_1,\,f_2\in\mathcal{D}_a(\gamma)}}\frac{\left[e^{b\rho_\gamma(f_1,\,f_2)}-e^{b\chi\rho_\gamma(f_1,\,f_2)}\right]f_1\int_\gamma gf_2}{\left[e^{b\rho_\gamma(f_1,\,f_2)}-e^{-\rho_\gamma(f_1,\,f_2)}\right]f_1\int_\gamma f_2}\\
\geq&|||g|||_-\inf_{t\in\mathbb{R}^+}\frac{e^{bt}-e^{\chi bt}}{e^{bt}-e^{-t}}\geq\frac{(1-\chi)b}{b+1}|||g|||_-,
\end{aligned}$$

and

$$\alpha_2 \geq (1-\chi)|||g|||_-.$$

This shows that $\alpha \geq \frac{(1-\chi)b}{b+1}|||g|||_-$. Moreover, a similar computation shows that $\beta \leq \frac{(1+\chi)b}{b-1}|||g|||_+$ and concludes the proof, since $\Theta(g,1) = \ln\frac{\beta}{\alpha}$.

It is interesting (and helpful) to notice that, if we consider the cone

$$\mathcal{C}_+ = \left\{ g \in C^{(0)}(\mathcal{M}) \;\middle|\; g \not\equiv 0;\ \int_\gamma fg \geq 0 \ \ \forall \gamma \in \Gamma_\delta,\ f \in \mathcal{D}_a(\gamma) \right\}$$

ant its Hilbert metric $\Theta_+$, then

$$\Theta_+(g,1) = \ln \frac{|||g|||_+}{|||g|||_-}$$

In this language, Lemma 4.6 states: for each $g \in \mathcal{C}_{\chi b,\,\chi c}$, $\Theta(g,1) \leq \ln K + \Theta_+(g,1)$.

Up to now we have let the scale $\delta$ be arbitrary. In the next step it is instead essential to compare averages on curves having some fixed scale. In the present context we could simply fix $\delta$ to be some appropriate value $\delta_0$. Nevertheless, it is inspiring to note that one can relate averages on any small scale $\delta$ with averages on some fixed scale $\delta_0$.

Let $\delta_0 > 0$ be fixed, define

$$|||g|||_+^0 = \sup_{\substack{\gamma \in \Gamma_{\delta_0} \\ f \in \mathcal{D}_a(\gamma)}} \frac{\int_\gamma gf}{\int_\gamma f}$$

$$|||g|||_-^0 = \inf_{\substack{\gamma \in \Gamma_{\delta_0} \\ f \in \mathcal{D}_a(\gamma)}} \frac{\int_\gamma gf}{\int_\gamma f}.$$

**Lemma 4.7.** *For each $\delta \leq \delta_0$, there exists $N_0 \in \mathbb{N}$ such that, for each $g \in \mathcal{C}_{b,c}$, $n \geq N_0$,*

$$|||\widetilde{T}^n g|||_+ \leq |||g|||_+^0$$
$$|||\widetilde{T}^n g|||_- \geq |||g|||_-^0,$$

*provided $a > \frac{c_3}{1-\lambda^{-1}}\delta_0^{1-\nu}$.*

*Proof.* For each $\gamma \in \Gamma_\delta$, $f \in \mathcal{D}_a(\gamma)$ and $n \geq \frac{\ln(2\delta_0\delta^{-1})}{\ln\lambda} = N_0$ we have $|T^{-n}\gamma| > 2\delta_0$, we can then divide $T^{-n}\gamma$ in curves $\{\gamma_i\} \subset \Gamma_{\delta_0}$. The condition on $a$ insures that $\widehat{T}^n f \in \mathcal{D}_a(\gamma_i)$.

$$\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^n g f = \frac{1}{\int_\gamma f}\sum_i \int_{\gamma_i} g\widehat{T}^n f \leq \frac{1}{\int_\gamma f}\sum_i \int_{\gamma_i} \widehat{T}^n f\, |||g|||_+^0 = |||g|||_+^0$$

that is, $|||\widetilde{T}^n g|||_+ \leq |||g|||_+^0$. The other inequality is proven in the same way.

**Diameter of the image.**

**Figure 2** Square neighborhood.

While up to now we did not use the mixing property, it will be necessary to do so in what follows. The main implication of mixing, in our context, is of a geometric nature: it asserts that given a finite number of sets of positive measure, some image of each one will intersect all the others. Consequently, given two curves that intersect such sets, the images of one curve will eventually get close to the other.

To make this argument precise we will construct a covering of $\mathcal{M}$, made up of sets that are well behaved with respect to the curves in $\Gamma_{\delta_0}$. The construction of the covering may seem a bit technical, but its main point is to insure that each curve in $\Gamma_{\delta_0}$ is contained in some element of the covering, and that the relevant geometric properties of the iterates of any curve can be described solely in terms of the covering.

Consider $z \in \mathcal{M}$. There exists a Cartesian system of coordinates in a neighborhood of $z$ such that the axes are parallel to $W^s(z)$ and $W^u(z)$, respectively. By the continuity of the cone family, there exists a smaller neighborhood of $z$ where each curve parallel to one of the axes is either in $\mathcal{C}$ (in the stable direction) or in its complement (unstable direction). Let $\widehat{Q}(z)$ be a square (in the Cartesian coordinates), with center in $z$, enjoying the above properties.

We can use the Cartesian structure to identify all the tangent space at $w \in \widehat{Q}(z)$ with the tangent space at $z$, and to use the Riemannian metric at $z$ to define a Cartesian metric in $\widehat{Q}(z)$. We then consider some smaller neighborhood $\widehat{Q}_1(z)$ such that, calling $d_c$ the Cartesian distance and $d$ the Riemannian one, $\frac{3}{4}d \leq d_c \leq \frac{5}{4}d$, and such that $\bigcup_{w \in \widehat{Q}_1(z)} \mathcal{C}(w) \subset \{|\xi| \leq \frac{1}{2}|\eta|\} = \widehat{\mathcal{C}}$; this means that we can consider a unique cone $\widehat{\mathcal{C}}$ in the neighborhood $\widehat{Q}_1(z)$. We finally consider the square neighborhood $Q(z)$, centered at $z$, but with size half the size of $\widehat{Q}_1(z)$ (see [LW] for more detail on such construction).

Since $\mathcal{M}$ is compact, there are finitely many $z_i \in \mathcal{M}$ such that $\bigcup_i Q(z_i) = \mathcal{M}$. We choose $\delta_0$ to be $\frac{1}{4}$ the size of the smallest of the $Q(z_i)$.

Given a square $G \subset \widehat{Q}_1(z_i)$, for some $i$, of size $4\delta_0$, it is easy to see that every curve in $\Gamma_{\delta_0}$ that intersects a concentrical square of size $\frac{\delta_0}{4}$ will be completely contained in $G$. We will call such concentrical square the "core" of $G$ (see [LW, §1-3] for a more extensive discussion of the concept of core and its use). We also define the "central strip" of a square, as a strip, parallel to $W^u(z)$, of size $\frac{\delta_0}{4}$, and containing the core.

In each $\widehat{Q}_1(z_i)$ we construct squares $G$ such that the union of their cores covers $Q(z_i)$; the collection of such squares is a covering of $\mathcal{M}$ with the following properties

(1) $\bigcup_{G\in\mathcal{G}} \operatorname{core}(G) = \mathcal{M}$;
(2) $\forall\gamma \in \Gamma_{\delta_0}$, $G \in \mathcal{G}$ if $\gamma \cap \operatorname{core}(G) \neq \emptyset$, then $\gamma \subset G$;
(3) $\forall\gamma \in \Gamma_{\delta_0}$ there exists $G \in \mathcal{G}$: $\gamma \cap \operatorname{core}(G) \neq \emptyset$, and the part of $\gamma$ that intersects the central strip of $G$ consists of $\gamma$ minus two pieces, each of length larger than $\frac{\delta_0}{4}$.

We will say that $\gamma$ "intersects properly" $G$ if the last of the latter properties is satisfied.

Such covering is the main ingredient in the following fundamental Lemma.

**Lemma 4.8.** *Let $\delta \leq \delta_0$, then there exist $N_* \in \mathbb{N}$ and $\Delta \in \mathbb{R}^+$ such that*

$$\textit{diameter}\ (\widetilde{T}^{N_*}\mathcal{C}_{b,\,c}) \leq \Delta < \infty,$$

*$N_* \leq N + N_0$, where $N$ depends only on $\delta_0$ and $T$.*

*Proof.* Since the map under consideration is mixing there exists $N \in \mathbb{N}$ such that, for each $G$, $G' \in \mathcal{G}$,

$$T^{-N}\operatorname{core}(G) \cap \operatorname{core}(G') \neq \emptyset.$$

Without loss of generality we can impose $N > n_1$ ($n_1$ to be chosen later). Given any function $g \in \mathcal{C}_{b,c}$, we choose $\gamma_* \in \Gamma_{\delta_0}$ and $f_* \in \mathcal{D}_a(\gamma_*)$ such that

$$\frac{\int_{\gamma_*} g f_*}{\int_{\gamma_*} f_*} \geq \frac{1}{2}|||g|||_+^0.$$

Let us call $G_*$ an element of $\mathcal{G}$ properly intersected by $\gamma_*$. Given any $\gamma \in \widehat{\Gamma}_{\delta_0}$, there will be an element $G$ of $\mathcal{G}$ that is properly intersected by $\gamma$. This implies that there exists $w \in \operatorname{core}(G) \cap T^N\operatorname{core}(G_*)$; by definition, $W^u_{\text{loc}}(w) \cap \gamma \neq \emptyset$, and consists of only one point $w_1$, moreover $d(T^N(w_1), T^N(w)) \leq \lambda^{-N}\frac{\delta_0}{4}$. Hence, there exists $\gamma_1 \subset T^{-N}\gamma$ such that it is the "translate"[12], along the unstable direction, of $\gamma_*$. In fact, by construction, $W^u(T^{-N}w) \cap \gamma_* \neq \emptyset$, and, if we cut $T^{-N}\gamma$ at $T^{-N}w_1$, each of the two resulting pieces is longer than $\lambda^N\frac{\delta_0}{8} \geq \lambda^{n_1}\frac{\delta_0}{8} \geq 3\delta_0$ (where we have chosen

[12]By this we mean that for each $z \in \gamma_*$ $W^u_{\text{loc}}(z) \cap \gamma_1 \neq \emptyset$ and consists of a single point; in addition, for each $w \in \gamma_1$ $W^u_{\text{loc}}(w) \cap \gamma_* \neq \emptyset$ and consists of a single point. See Appendix II for more details.

$n_1$ such that $\lambda^{n_1} \geq 24$). Hence, $T^{-N}\gamma\backslash\gamma_1$ can be divided into pieces belonging to $\Gamma_{\delta_0}$, and

$$\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^N gf \geq \frac{1}{\int_\gamma f}\int_{\gamma_1} g\widehat{T}^N f.$$

Let $\phi$ be the canonical isomorphism (see Appendix II) from $\gamma_*$ to $\gamma_1$ and $J_\phi$ its Jacobian.

$$\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^N gf \geq \frac{1}{\int_\gamma f}\int_{\gamma_*} g\circ\phi\left[(\widehat{T}^N f)\circ\phi\right]J_{\phi^{-1}}$$

¿From the results recalled in Appendix II it follows that, for $R$, $a$ large and $\delta_0$ small enough, $\widetilde{f} = (\widehat{T}^N f)\circ\phi\, J_{\phi^{-1}} \in \mathcal{D}_a(\gamma_*)$. Hence, for some $z\in\gamma_*$,

$$\begin{aligned}
\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^N gf \geq & \frac{1}{\int_\gamma f}\int_{\gamma_*} g\widetilde{f} - \frac{\delta_0\|D^u g\|_\infty}{\int_\gamma f}\int_{\gamma_*}\widetilde{f}\\
\geq & \frac{1}{\int_\gamma f}\left[\frac{\widetilde{f}(z)}{f_*(z)}\int_{\gamma_*} f_* g e^{-2bR} - c\delta_0 |||g|||_-\int_{\gamma_*}\widetilde{f}\right]\\
\geq & \frac{1}{\int_\gamma f}\left[\frac{\widetilde{f}(z)}{2f_*(z)}|||g|||_+^0 e^{-2bR}\int_{\gamma_*} f_* - c\delta_0|||g|||_-\int_{\gamma_*}\widetilde{f}\right]\\
\geq & \left[\frac{e^{-2(b+1)R}}{2}|||g|||_+^0 - c\delta_0|||g|||_-\right]\frac{\int_{\gamma_1}\widehat{T}^N f}{\int_\gamma f}.
\end{aligned}$$

On the one hand, if $\frac{e^{2(b+1)R}}{4}|||g|||_+^0 \leq c\delta_0|||g|||_-$, we have immediately

$$\frac{|||\widetilde{T}^{N+N_0}g|||_+}{|||\widetilde{T}^{N+N_0}g|||_-} \leq \frac{4c\delta_0}{e^{2(b+1)R}}.$$

On the other hand, if $\frac{e^{2(b+1)R}}{4}|||g|||_+^0 \geq c\delta_0|||g|||_-$, we choose an arbitrary point $\zeta \in T^N\gamma_1 \subset \gamma$, and we have

$$\begin{aligned}
\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^N gf \geq & \frac{\int_{T^N\gamma_1}\frac{f(y)}{f(\zeta)}dm(y)}{\int_\gamma \frac{f(y')}{f(\zeta)}dm(y')}\frac{e^{-2(b+1)R}}{4}|||g|||_+^0\\
\geq & e^{-2^\nu a\delta_0^\nu}\frac{|T^N\gamma_1|}{|\gamma|}\frac{e^{-2(b+1)R}}{4}|||g|||_+^0 \geq e^{-2^\nu a\delta_0^\nu}\mu^{-N}\frac{e^{-2(b+1)R}}{4}|||g|||_+^0
\end{aligned}$$

where $\|DT\|\leq\mu$. ¿From the above chain of inequalities follows:

$$\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^N gf \geq e^{-2^\nu a\delta_0^\nu}\mu^{-N}\frac{e^{-2(b+1)R}}{4}|||\widetilde{T}^N g|||_+^0.$$

This last inequality yields

$$|||\widetilde{T}^N g|||_-^0 \geq \frac{e^{-2^\nu a\delta_0^\nu}\mu^{-N}e^{-2(b+1)R}}{4}|||\widetilde{T}^N g|||_+^0,$$

which implies

$$\frac{|||\widetilde{T}^{N+N_0}g|||_+}{|||\widetilde{T}^{N+N_0}g|||_-} \leq 4e^{2^\nu a\delta_0^\nu}\mu^N e^{2(b+1)R}.$$

Thanks to Lemma 4.8 we can apply Theorem 1.1.

Summarizing, there exists $\Lambda_0 \in (0,\, 1)$ such that, for each $g_1,\, g_2 \in \mathcal{C}_{b,\,c}$

$$\Theta(\widetilde{T}^n g_1,\, \widetilde{T}^n g_2) \leq \Lambda_0^{\overline{\frac{n}{N+N_0}}}\Theta(g_1,\, g_2).$$

Let $\Lambda = \Lambda_0^{(N+N_0)^{-1}}$.

**Decay of correlations.**

**Theorem 4.9.** *There exist $D > 0$, and $r > 0$ such that, for each $g,\, f \in C^{(1)}(\mathcal{M})$*

$$\left| \int_{\mathcal{M}} f\widetilde{T}^n g - \int_{\mathcal{M}} f \int_{\mathcal{M}} g \right| \leq D\|g\|_*\|f\|_*\Lambda^n$$

*with $\|h\|_* = \int_{\mathcal{M}} |h| + r\|h'\|_\infty$.*

*Proof.* The basic idea behind this result is to partition $\mathcal{M}$ into curves belonging to $\Gamma_\delta$. Unfortunately, such a simple idea it is not trivially implementable. Of course, such partitions exist (e.g. Markov partitions) but it is not immediately obvious which constraints they force on $\delta$. In order to avoid the problem entirely, we content ourselves with a much coarser partition and use a dynamical argument to refine it.

We have already constructed a covering $\{Q(z_i)\}$ of $\mathcal{M}$. From such covering we can obtain a partition $\mathcal{P} = \{P_i\}$ of $\mathcal{M}$, such that $Q(z_i) \supset P_i$ and $\partial P_i$ is piecewise smooth.

Next, for each $\delta_1 < \delta$, we can partition $Q(z_i)$ into squares of size $\delta_1$, with sides parallel to the axes. We will collect into $\mathcal{P}_+^i(\delta_1)$ all the squares $p$ such that $p \cap P_i \neq \emptyset$, and into $\mathcal{P}_-^i(\delta_1)$ all the squares $p$ such that $p \subset P_i$; let $\mathcal{P}_+(\delta_1) = \cup_i \mathcal{P}_+^i(\delta_1)$, $\mathcal{P}_+(\delta_1) = \cup_i \mathcal{P}_+^i(\delta_1)$. Clearly, for each $\varepsilon > 0$, there exists $\delta_1 \in \mathbb{R}^+$ such that

$$\begin{aligned} &\sum_{p\in\mathcal{P}_+} m(p) \leq 1+\varepsilon \\ &\sum_{p\in\mathcal{P}_-} m(p) \geq 1-\varepsilon. \end{aligned} \tag{4.2}$$

This two series of "quasi-partitions" of $\mathcal{M}$ will be sufficient to prove:

**Lemma 4.10.** *If $f \in C^{(0)}(\mathcal{M})$, $f \in \widetilde{\mathcal{D}}_{\sigma a/2}(\gamma)$ for each $\gamma \in \Gamma(\delta)$, and $g \in \mathcal{C}_{b,c}$, then*

$$\int_{\mathcal{M}} f|||g|||_- \leq \int_{\mathcal{M}} fg \leq \int_{\mathcal{M}} f|||g|||_+.$$

*Proof.* For each $\varepsilon > 0$, choose $\delta_1$ such that (4.2) holds, then, for each $n_1 > 0$,

$$\int_{\mathcal{M}} fg = \int_{\mathcal{M}} \widetilde{T}^{n_1}(fg) \geq \sum_{p\in\mathcal{P}_-} \int_p \widetilde{T}^{n_1}(fg).$$

By construction, each $p$ is naturally partitioned into curves $\gamma_s$ belonging to $\Gamma_{\delta_1}$. If $\widetilde{h}_p$ is the smooth density of the invariant measure expressed in the local Cartesian coordinate $(\eta,\, \xi)$ of the square $Q(z_i)$, then, by applying Fubini Theorem,

$$\int_{\mathcal{M}} fg \geq \sum_{p\in\mathcal{P}_-} \int d\xi \int_{\gamma_\xi} \widetilde{h}_p \widetilde{T}^{n_1}(fg) = \sum_{p\in\mathcal{P}_-} \int d\xi \int_{T^{-n_1}\gamma_\xi} \widehat{T}^{n_1}(\widetilde{h}_p) fg.$$

We choose $n_1$ large enough such that, for each $p$, $f\widehat{T}^{n_1}\widetilde{h}_p \in \widetilde{\mathcal{D}}_{\sigma a}(\gamma) \subset \mathcal{D}_a(\gamma)$, and $|T^{-n_1}\gamma_\xi| > \delta$ for each $\xi$,

$$\begin{aligned}\int_{\mathcal{M}} fg &\geq \sum_{p\in\mathcal{P}_-} \int d\xi \int_{T^{-n_1}\gamma_\xi} \widehat{T}^{n_1}(\widetilde{h}_p) f |||g|||_- = \sum_{p\in\mathcal{P}_-} \int d\xi \int_{\gamma_\xi} \widetilde{h}_p \widetilde{T}^{n_1} f |||g|||_- \\ &\geq \left( \int_{\mathcal{M}} f - \int_{\mathcal{M}\setminus \bigcup\limits_{p\in\mathcal{P}_-} p} \widetilde{T}^{n_1} f \right) |||g|||_- \geq \left( \int_{\mathcal{M}} f - \varepsilon \|f\|_\infty \right) |||g|||_-.\end{aligned}$$

This yields the first inequality, since $\varepsilon$ is arbitrary. The second inequality is proven in the same way.

The previous Lemma implies

$$\int_{\mathcal{M}} f\widetilde{T}^n g \leq \frac{|||\widetilde{T}^n g|||_+}{|||\widetilde{T}^n g|||_-} \int_{\mathcal{M}} f \int_{\mathcal{M}} g \leq e^{\Theta(\widetilde{T}^n g,\, 1)} \int_{\mathcal{M}} f \int_{\mathcal{M}} g$$

and

$$\int_{\mathcal{M}} f\widetilde{T}^n g \geq e^{-\Theta(\widetilde{T}^n g,\, 1)} \int_{\mathcal{M}} f \int_{\mathcal{M}} g.$$

The preceding inequalities shows that there exists a constant $D$ such that the Theorem holds for the functions under consideration.

Next, we will consider $f,\, g \in C^{(0)}(\mathcal{M})$, $f \geq 0$, $g \geq 0$, $f$, $g$ almost everywhere differentiable with $\|f'\|_\infty < \infty$, and $\|g'\|_\infty < \infty$; in addition, for each stable curve $\gamma$ $f$ is piecewise differentiable along $\gamma$, and $g$ is piecewise differentiable along each unstable manifold. Let $f_\rho = f + \rho$, and $g_\xi = g + \xi$ ($\rho,\, \xi \in \mathbb{R}^+$), then

$$\frac{f(x)}{f(y)} \leq e^{\int_x^y \frac{\|f'\|_\infty}{\rho}} \leq e^{\frac{\|f'\|_\infty}{\rho} d(x,\, y)^\nu \delta^{1-\nu}}.$$

It follows then that, choosing $\rho = \frac{2\|f'\|_\infty \delta^{1-\nu}}{\sigma a}$, we have $f \in \mathcal{D}_{\frac{\sigma a}{2}}(\gamma)$ for each $\gamma \in \Gamma_\delta$. We can perform a similar computation for $g_\xi$; it yields $g_\xi \in \mathcal{C}_{b,\, c}$, provided $\xi = \frac{\|g'\|_\infty}{c}$. This implies

$$\begin{aligned}\left| \int_{\mathcal{M}} f\widetilde{T}^n g - \int_{\mathcal{M}} f \int_{\mathcal{M}} g \right| &= \left| \int_{\mathcal{M}} f_\rho \widetilde{T}^n g_\xi - \int_{\mathcal{M}} f_\rho \int_{\mathcal{M}} g_\xi \right| \leq D\Lambda^n \int_{\mathcal{M}} f_\rho \int_{\mathcal{M}} g_\xi \\ &\leq D\Lambda^n \left[ \int_{\mathcal{M}} f + \delta^{1-\nu} \frac{2\|f'\|_\infty}{a\sigma} \right] \left[ \int_{\mathcal{M}} g + \frac{\|g'\|_\infty}{c} \right].\end{aligned}$$

Again, the proof is concluded since any smooth function can be written as the difference of positive functions with the above properties.

**Remark 4.11.** *An interesting feature of the proof presented in this section is the possibility of computing explicitly the constants $D$ and $\Lambda$ by only looking at a finite number of iterations of the map. It would be interesting to perform such computations (eventually by some numerical method) in specific examples to verify if the present strategy can produce realistic estimates. By contrast, the attempt to implement numerically the proof based on Markov partitions is by no means trivial* [Ga].

**Review**

Let us review the strategy of the proof before going to the next section. The proof consists, essentially, of five steps. The first is the specification of an invariant cone of functions in terms of smoothness in the unstable direction and averages in the stable direction. The second is the reduction of the estimate of the diameter of the images to the estimate of the ratio of the maximal and minimal average. The third (not particularly essential in this context but essential in the following) is the reduction of the estimate of the averages on the scale $\delta$ to averages on some fixed scale $\delta_0$. The fourth is the estimate on the averages on the scale $\delta_0$ by using mixing first and the absolute continuity second. The fifth (and last) step consists of showing that the contraction in the Hilbert metric suffices to imply the exponential decay of correlation.

In the next section we will carry out the same five steps in a more general context. I advise the reader to refer to this section if some confusion arises in the next, more technical, one.

## §5 TWO DIMENSIONS (THE NON-SMOOTH UNIFORMLY HYPERBOLIC CASE).

In this section we will consider a class of maps that is similar to, but more restrictive than, the one considered in [LW] and [KS].

More precisely, let $\mathcal{M} \subset \mathbb{R}^2$ be a connected compact set on which the map $T$ is defined: $T : \mathcal{M} \to \mathcal{M}$ .[13] For simplicity, we will assume $\mathcal{M}$ to be a square, but we could as well require that $\partial\mathcal{M}$ consists of a finite number of smooth lines intersecting transversally. We will call a set with the above more general property a "symplectic box" (see [LW] for a general definition). The symplectic box $\mathcal{M}$ is partitioned in two ways into unions of an equal number of symplectic boxes

$$\mathcal{M} = \mathcal{M}_1^+ \bigcup \cdots \bigcup \mathcal{M}_m^+ = \mathcal{M}_1^- \bigcup \cdots \bigcup \mathcal{M}_m^-.$$

Two boxes of one partition can overlap at most on their boundaries, i.e.,

$$\mathcal{M}_i^\pm \bigcap \mathcal{M}_j^\pm \subset \partial\mathcal{M}_i^\pm \bigcap \partial\mathcal{M}_j^\pm, \quad i, j = 1, \ldots, m.$$

[13]The case in which the map is defined on finitely many disconnected sets can be treated in the same way.

The map $T$ is defined separately on each of the symplectic boxes $\mathcal{M}_i^+$, $i = 1, \dots, m$. It is a symplectomorphism (i.e. it preserves the area and it is smooth) of the interior of each $\mathcal{M}_i^+$ onto the interior $\mathcal{M}_i^-$, $i = 1, \dots, m$ and a homomorphism of $\mathcal{M}_i^+$ onto $\mathcal{M}_i^-$, $i = 1, \dots, m$. We will assume that $T$ is $C^{(2)}$ in each box up to and including the boundary.[14]

We will say that $T$ is a (discontinuous) symplectic map of $\mathcal{M}$. Formally $T$ is not well defined on the set of points which belong to the boundaries of several plus-boxes: it has several values. We adopt the convention that the image of a subset of $\mathcal{M}$ under $T$ contains all such values.

Let us introduce the singularity sets $\mathcal{S}^+$ and $\mathcal{S}^-$.

$$\mathcal{S}^\pm = \{p \in \mathcal{M} \mid p \text{ belongs to at least two of the boxes } \mathcal{M}_i^\pm, i = 1, \dots, m\}.$$

The plus-singularity set $\mathcal{S}^+$ is a closed subset, and $T$ is continuous on its complement. Similarly $T^{-1}$ is continuous on the complement of $\mathcal{S}^-$.

It is the case that $\mathcal{S}^+ \bigcup \partial\mathcal{M}$ is the union of all the boundaries of the plus-boxes and $\mathcal{S}^- \bigcup \partial\mathcal{M}$ is the union of all the boundaries of the minus-boxes, i.e.,

$$\mathcal{S}^\pm \bigcup \partial\mathcal{M} = \bigcup_{i=1}^m \partial\mathcal{M}_i^\pm.$$

Note that most of the points in the boundary $\partial\mathcal{M}$ of $\mathcal{M}$ do not belong to $\mathcal{S}^-$ or $\mathcal{S}^+$.

Our assumptions imply that the singularity sets $\mathcal{S}^\pm$ and the union of boundaries $\bigcup_{i=1}^m \partial\mathcal{M}_i^\pm$ is the finite union of smooth lines intersecting transversally.

We assume $T$ to be uniformly hyperbolic, and that at each point $x \in \partial\mathcal{M}_i^\pm$ the stable and unstable direction be uniformly transversal to $\partial\mathcal{M}_i^\pm$.

We assume that $T$ is mixing.[15]

In addition, we chose $\zeta \in [\sqrt{3} - 1, 1]$ such that

$$\lambda^{-\zeta}\mu^{1-\zeta} < 1 \tag{5.1}$$

where $\lambda$ and $\mu$ are the minimal and maximal expansion coefficients of $T$, respectively (see §4 for a more precise definition of $\lambda$ and $\mu$).

Next, we have a condition specific to the present context, which does not play any role in the study of weaker statistical properties (like mixing and ergodicity). We assume that there exists $n_0 \in \mathbb{N}$, such that

$$M(n_0)\lambda^{-\zeta n_0} < \frac{1}{25}, \tag{5.2}$$

$M(n)$ being the maximal number of smooth lines belonging to $\bigcup_{i=0}^n T^i\mathcal{S}^-$ that intersect at one point.[16]

[14] This is a serious restriction: in particular, billiards do not satisfy it. See §6 for a discussion of one example that does not satisfy such restrictions.

[15] The observations in footnote 9 apply here as well.

[16] The choice of 25 is completely arbitrary. The point is that if $M^n\lambda^{-\zeta n} < 1$, for some $n$, then there exists some $n_0$ that satisfies (5.2).

**Remark 5.1.** *Note that condition 5.2 can always be fulfilled if $M(n)$ has a polynomial growth. But it is not necessary to be able to control $M(n)$ for each $n$, since 5.2 is a condition concerning some fixed power of $T$.*

Henceforth, we will let $n_0 = 1$, since this can always be achieved, eventually replacing $T$ by $T^{n_0}$.

Finally, we assume that there exists a family of cones in the tangent of $\mathcal{M}$ (not to be confused with the cones of functions used later). First we have a cone family $\mathcal{C}_-(x) \in \mathcal{T}_x\mathcal{M}$, continuous on $\text{int}(\mathcal{M})$. We assume that $\mathcal{C}_-(x)$ contains the stable, but not the unstable, direction, and that if $x \in \partial\mathcal{M}$, then one edge of the cone is tangent to $\partial\mathcal{M}$.[17]

We define $\mathcal{C}_+(x)$ to be the complement of $\mathcal{C}_-(x)$. In addition, we will consider $\widehat{\mathcal{C}}_-(x) = DT^{-l}\mathcal{C}_-(T^l x)$, where $l$ is chosen so that the width of $\widehat{\mathcal{C}}_-(x)$ is less than some $\epsilon \in \mathbb{R}^+$, and $DT^{-1}\widehat{\mathcal{C}}_-(x) \subset \mathcal{C}_-(T^{-1}x)$ (which implies $DT^{-l-1}\widehat{\mathcal{C}}(x) \subset \widehat{\mathcal{C}}(T^{-l-1}x)$; i.e., $\widehat{\mathcal{C}}$ is a strictly invariant cone family for some fixed power of $T$).

If the reader wishes to keep in mind a concrete example of the type of systems the results of this section apply to, she can consider the following: $\mathcal{M} = \mathbb{T}^2$, $T(x, y) = F(A(x, y))$ where

$$A = \begin{pmatrix} 1 & a \\ a & 1 + a^2 \end{pmatrix}$$

with $a \notin \mathbb{N}$ (which implies that the linear part is discontinuous on the torus); $T$ is area preserving; $F : \mathbb{T}^2 \to \mathbb{T}^2$ is a $C^2$ function sufficiently close to the identity so that the hyperbolicity is not spoiled (see [LW] for a discussion of the structure of the symplectic boxes formalism in this case).

**Definition 5.2.** *A curve $\gamma$ is said to be "stable (unstable)" if it is smooth, and $\gamma'(x) \in int(\mathcal{C}_-(x))$ ($\gamma'(x) \in int(\mathcal{C}_+(x))$ ) for each $x \in \gamma$.*

*A curve $\gamma$ is said to be "(strictly) weakly stable" if it is connected; it consists of finitely many smooth pieces; $\gamma'(x) \in \mathcal{C}_-(x)$ ($\gamma'(x) \in \widehat{\mathcal{C}}_-(x)$) for each $x \in \gamma$ for which $\gamma'(x)$ exists; and for each unstable curve $\gamma^u$ shorter than some fixed $\delta_0$, if $x \in \gamma^u \cap \gamma$, then we have $\gamma \bigcap \gamma^u = \{x\}$.*

### Cones of Functions

We will use the same notations of the previous section. The only difference is due to the presence of the discontinuities, which forces us to choose a more complicated

[17] This condition amounts to little more than transversality of $\partial\mathcal{M}$ with respect to the unstable direction, together with uniform transversality of the stable and unstable direction at each point at which they are both defined. In fact, in almost all the known examples the proof of the hyperbolicity of the system starts exactly with such a cone family.

cone instead than $\mathcal{C}_{bc}$.

$$\begin{aligned}
\Gamma_\delta =&\{\gamma \subset \mathcal{M} \mid \delta \leq \text{length}(\gamma) \leq 2\delta;\ \text{weakly stable};\ |\kappa_\gamma| \leq H\}\\
\widehat{\Gamma}_\delta =&\{\gamma \subset \mathcal{M} \mid \delta \leq \text{length}(\gamma) \leq 2\delta;\ \text{strictly weakly stable};\ |\kappa_\gamma| \leq H\}
\end{aligned}$$

$$\begin{aligned}
\mathcal{C}_{b,\,c,\,d,\,L}(\delta) =&\Bigg\{ g \in C^{(0)}\left(\mathcal{M}\setminus \bigcup_{n=0}^{\infty} T^n\mathcal{S}^-\right) \bigcap L^1(\mathcal{M}) \;\Bigg|\; \forall \gamma \in \Gamma_\delta;\ f,\, f_1,\, f_2 \in \mathcal{D}_a(\gamma);\\
& x \in \gamma;\ I \subset \gamma \text{ connected},\quad \int_\gamma gf > 0;\ \frac{f_2(x)\int_\gamma f_1 g}{f_1(x)\int_\gamma f_2 g} \leq e^{b\rho_\gamma(f_1,\, f_2)};\\
& \frac{\|D^u g\|_\infty}{|||g|||_-} \leq c;\quad \frac{\left|\int_I gf\right|}{f(x)} \leq d\delta^{1-\zeta}|I|^\zeta |||g|||_-;\quad \frac{|||g|||_+}{|||g|||_-} \leq L\Bigg\}
\end{aligned}$$

It was necessary to weaken the regularity requirements on $g$ since, given $g \in C^{(0)}$, $g \circ T^{-1}$ may be discontinuous; our requirement is the maximal invariant regularity one can ask for. In addition, although $g$ may be discontinuous on a dense set, it makes sense to talk about the derivative in the unstable direction. In fact, the local unstable manifold $W^u_{\text{loc}}(x)$ is well defined at almost every point [KS], and, by definition, enjoys the property $W^u_{\text{loc}}(x) \cap \cup_{n=0}^{\infty} T^n\mathcal{S}^- = \emptyset$. Hence, we can define $D^u_x g$ to be the derivative along $W^u_{\text{loc}}(x)$. Finally, we have a condition concerning integrals on arbitrarily short curves: since such pieces are unavoidable, a condition of this type seems necessary; the last condition turns out to be needed for technical reasons.

**Contraction of the Hilbert Metric**

Of course, our aim is to prove the following:

**Theorem 5.3.** *There exists $m \in \mathbb{N}$, $b,\, c,\, d,\, L \in \mathbb{R}^+$, $\delta_s \in \mathbb{R}^+$, and $\chi \in (0,\,1)$ such that, for $\delta \leq \delta_s$,*

$$\widetilde{T}^m \mathcal{C}_{b,\,c,\,d,\,L}(\delta) \subset \mathcal{C}_{\chi b,\,\chi c,\,\chi d,\,\chi L}(\delta).$$

Indeed, we shall see that this is sufficient to obtain a contraction of the Hilbert metric (Lemma 5.17). Nevertheless, the route to Theorem 5.3 (at least the one found by the author) is not very direct. Several intermediate steps shall be needed.

The next Lemma is analogous to Lemma 4.2. It is however less satisfactory: it does not provide us with an invariant cone. The condition that fails pertains the ratio of the averages. Yet, in the previous section we were able to control effectively this ratio only at the very end, by an independent argument. It is then reasonable to content oneself temporarily with the following result, and to proceed with crossed fingers.

**Lemma 5.4.** *There exists $N(\delta)$, $\lim_{\delta\to 0} N(\delta) = \infty$ and $A \in \mathbb{R}^+$, such that, if $L > 1$, $8L < d$, $b$ large enough (see 5.8), and $2^\nu\delta^\nu a < \ln 2$, then, for each $n \in \{A \ln d,\, \dots,\, N(\delta)\}$,*

$$\widetilde{T}^n \mathcal{C}_{b,\,c,\,d,\,L}(\delta) \subset \mathcal{C}_{\chi b,\,\chi c,\,\chi d,\,\tau_n^2 L}(\delta)$$

*where $\chi < 1$, and $\tau_n^{-1} = 1 - 3d\lambda^{-n}M^n \geq 1/\sqrt{2}$.*

*Proof.* Not surprisingly all the difficulties arise because $T^{-n}\gamma$ may contain many, eventually very short, connected pieces. The main tool to overcome such problems follows:

**Sub-Lemma 5.5.** *There are $\delta_0 \in \mathbb{R}^+$ and $N : \mathbb{R}^+ \to \mathbb{N}$, such that, for each $\delta \le \delta_0$, $n \le N(\delta)$, and $\gamma \in \Gamma_\delta$, $T^{-n}\gamma$ consists, at most, of $M^n$ pieces.*

*Proof.* Property (5.2) implies that there exists $\delta_0$ such that $\mathcal{S}^-$ intersects at most $M$ times a segment in the stable direction of length $\delta_0$, and $M\lambda^{-\zeta} < 1$. Given $\delta \le \delta_0$ and $\gamma \in \Gamma_\delta$, $T^{-1}\gamma$ will consists, at most of $M$ pieces. If $\mu = \|DT^{-1}\|_\infty$, then the length of such pieces will be, at most, $\mu\delta$. The result follows then by choosing $N(\delta) = \left[\frac{\ln \delta_0 - \ln \delta}{\ln \mu}\right]$.

The first step consists in comparing $|||g|||_-$ and $|||\widetilde{T}^n g|||_-$, for $n \le N(\delta)$. Given $\gamma \in \Gamma_\delta$, $T^{-n}\gamma$ consists, at most, of $M^n$ pieces $\{\widetilde{\gamma}_i\}$. We divide such pieces into two collections. The first, $\mathcal{J}_-$, will contain all the maximal connected pieces of length less than $\delta$. The second collection, $\mathcal{J}_+$, will consist of connected pieces of $T^{-n}\gamma$ belonging to $\Gamma_\delta$. In addition, we require $\left(\bigcup_{\widetilde{\gamma}\in\mathcal{J}_-}\widetilde{\gamma}\right) \cup \left(\bigcup_{\widetilde{\gamma}\in\mathcal{J}_+}\widetilde{\gamma}\right) = T^{-n}\gamma$. For each $\widetilde{\gamma} \in \mathcal{J}_-$ we choose some $x_{\widetilde{\gamma}} \in \widetilde{\gamma}$. Using such notations we can write

$$
\begin{aligned}
\frac{1}{\int_\gamma f}\int_\gamma \widetilde{T}^n g f =& \frac{1}{\int_\gamma f}\int_{\bigcup\limits_{\widetilde{\gamma}\in\mathcal{J}_+}\widetilde{\gamma}} g\widehat{T}^n f + \frac{1}{\int_\gamma f}\int_{\bigcup\limits_{\widetilde{\gamma}\in\mathcal{J}_-}\widetilde{\gamma}} g\widehat{T}^n f \\
\ge& \frac{1}{\int_\gamma f}\sum_{\widetilde{\gamma}\in\mathcal{J}_+}\int_{\widetilde{\gamma}} g\widehat{T}^n f - \frac{1}{\int_\gamma f}\sum_{\widetilde{\gamma}\in\mathcal{J}_-} f(T^n x_{\widetilde{\gamma}}) d\lambda^{-n}\delta^{1-\zeta}|\widetilde{\gamma}|^\zeta |||g|||_- \\
\ge& \left\{\left(1 - \sum_{\widetilde{\gamma}\in\mathcal{J}_-}\frac{\int_{\widetilde{\gamma}}\widehat{T}^n f}{\int_\gamma f}\right) - dM^n\lambda^{-n}e^{a2^\nu\delta^\nu}\right\}|||g|||_- \\
\ge& \left\{(1 - 2\lambda^{-n}M^n) - 2dM^n\lambda^{-n}\right\}|||g|||_- \\
\ge& \left(1 - 3d(M\lambda^{-1})^n\right)|||g|||_-.
\end{aligned}
$$

Hence, if $n \ge \frac{2\ln d}{\ln M^{-1}\lambda} > \frac{\ln 3d}{\ln M^{-1}\lambda}$, we have

$$\int_\gamma \widetilde{T}^n g f \ge 0.$$

In addition, for $A > 0$ sufficiently large, given $n \in \{A\ln d, \ldots, N(\delta)\}$ we get

$$|||\widetilde{T}^n g|||_- \ge \tau_n^{-1}|||g|||_- \ge \frac{1}{\sqrt{2}}|||g|||_-. \tag{5.3}$$

A similar computation yields

$$|||\widetilde{T}^n g|||_+ \le \tau_n |||g|||_+. \tag{5.4}$$

We are now ready to verify the remaining conditions. We start by analyzing integrals over short curves.

Let $n \leq N(\delta)$ and consider $I \subset \gamma \in \Gamma_\delta$, $I$ connected. The image $T^{-n}I$ can be divided into many connected pieces that are either in $\Gamma_\delta$ or are shorter than $\delta$; we call $\mathcal{J}_+$ the collection of the first and $\mathcal{J}_-$ the collection of the second. Given $f \in \mathcal{D}_a(\gamma)$ and $x \in I$

$$\frac{\left|\int_I \widetilde{T}^n gf\right|}{f(x)} \leq \frac{1}{f(x)}\left\{\sum_{J\in\mathcal{J}_-}\left|\int_J g\widehat{T}^n f\right| + \sum_{J\in\mathcal{J}_+}\int_J g\widehat{T}^n f\right\}.$$

For each $J \in \mathcal{J}_-$, choose $x_J \in J$ such that $\left|\det(D_{x_J}T^n\big|_J)\right| = \inf_{z\in J}\left|\det(D_z T^n\big|_J)\right|$, then

$$\begin{aligned}\frac{\left|\int_I \widetilde{T}^n gf\right|}{f(x)} \leq &\frac{1}{f(x)}\left\{\sum_{J\in\mathcal{J}_-} f(T^n x_J)\left|\det(D_{x_J}T^n\big|_J)\right| d|J|^\zeta\delta^{1-\zeta}|||g|||_-\right.\\ &\left.+\sum_{J\in\mathcal{J}_+}\int_J \widehat{T}^n f|||g|||_+\right\}.\end{aligned}$$

Since,

$$|J|\left|\det(D_{x_J}T^n\big|_J)\right| \leq \int_{T^nJ}\frac{\left|\det(D_{x_J}T^n\big|_J)\right|}{\left|\det(D_{T^{-n}y}T^n\big|_J)\right|}dm(y) \leq |T^nJ|, \tag{5.5}$$

it follows that

$$\begin{aligned}\frac{\left|\int_I \widetilde{T}^n gf\right|}{f(x)} \leq &\left\{\sum_{J\in\mathcal{J}_-} d\frac{f(T^n x_J)}{f(x)}\lambda^{(\zeta-1)n}|T^nJ|^\zeta\delta^{1-\zeta} + L\sum_{J\in\mathcal{J}_+}\int_{T^nJ}\frac{f(y)}{f(x)}dm(y)\right\}\\ &\times|||g|||_- \leq de^{a2^\nu\delta^\nu}\left\{\lambda^{(\zeta-1)n}\delta^{1-\zeta}\left[\sum_{J\in\mathcal{J}_-}|T^nJ|\right]^\zeta\left[\sum_{J\in\mathcal{J}_-}1\right]^{1-\zeta}\right.\\ &\left.+Ld^{-1}\left|\bigcup_{J\in\mathcal{J}_+}T^nJ\right|^\zeta\delta^{1-\zeta}\right\}|||g|||_-\end{aligned}$$

which implies:

$$\frac{\left|\int_I \widetilde{T}^n gf\right|}{f(x)} \leq dK_1\delta^{1-\zeta}|I|^\zeta|||g|||_- \tag{5.6}$$

where $K_1$ can be taken to be $4\max\{(\lambda^{-1}M)^{(1-\zeta)n},\, Ld^{-1}\}$, provided $\delta$ is sufficiently small.

Choosing $\delta$ small and $A$ large enough, it is possible to have $K_1 \leq 1/2$, provided that

$$L \leq \frac{d}{8}. \tag{5.7}$$

**Figure 3** Extending $\gamma$.

Collecting 5.3 and 5.6 we get

$$\frac{\left|\int_I \widetilde{T}^n g f\right|}{f(x)} \leq \chi d\delta^{1-\zeta}|I|^{\zeta}|||\widetilde{T}^n g|||_-,$$

provided $n \in \{A \ln d, \ldots, N(\delta)\}$.

Moreover,

$$0 \leq \frac{\|D^u \widetilde{T}^n g\|_\infty}{|||\widetilde{T}^n g|||_-} \leq \chi c.$$

Finally, when we look at the ratios of different averages, difficulties again originate from the possible existence of short pieces. The next result resolves these difficulties.

**Sub-Lemma 5.6.** *For each $\gamma \in \Gamma_\delta$, $f_1$, $f_2 \in \mathcal{D}_a(\gamma)$, $I \subset \gamma$ connected, $|I| < \delta$, $x \in I$, and $g \in \mathcal{C}_{b,c,d,L}(\delta)$ holds*

$$\frac{\int_I g f_1}{f_1(x)} \leq e^{b\rho_\gamma(f_1,\, f_2)} \frac{\int_I g f_2}{f_2(x)} + 8e^{b\rho_\gamma(f_1,\, f_2)}\delta L b\rho_\gamma(f_1,\, f_2)|||g|||_-.$$

*Proof.* We introduce a curve $\gamma_e \in \Gamma_\delta$ such that $\operatorname{int}(I) \bigcap \operatorname{int}(\gamma_e) = \emptyset$ but $\overline{\gamma} = I \bigcup \gamma_e$ is connected and $\overline{\gamma} \in \Gamma_\delta$.[18] In addition, given $f \in \mathcal{D}_a(\gamma)$ we modify it to a function $\widetilde{f} \in \mathcal{D}_a(\overline{\gamma})$ by the definition: for each $z \in \gamma_e$ we set $\widetilde{f}(z) = f(y)$, where $\{y\} = I \bigcap \gamma_e$ is the point in $I$ closest to $z$. For simplicity, hereafter, we will identify $f$ with

[18] Such a curve always exists. The only delicate point arises if either one or both of the ends of $\gamma$ are closer than $\delta$ to $\partial\mathcal{M}$. If one of the edges is close to $\partial M$, then we can clearly prolong $\gamma$ at one side, provided the curve $\gamma_e$ cannot meet a corner of $\partial\mathcal{M}$ in that direction. In this last case it is easy to convince oneself that $\gamma$ can be prolonged at the other end, since $\mathcal{T}\partial\mathcal{M} \subset \mathcal{C}_-$.

$\widetilde{f}$, since this can be done without any ambiguity as long as we are interested in integrating on $I$ only. Also, note that $\rho_I(f_1, f_2) = \rho_{\widetilde{\gamma}}(\widetilde{f}_1, \widetilde{f}_2) \leq \rho_\gamma(f_1, f_2)$.

$$
\begin{aligned}
\frac{\int_I g f_1}{f_1(x)} =& \frac{\int_{\widetilde{\gamma}} g f_1}{f_1(x)} - \frac{\int_{\gamma_e} g f_1}{f_1(x)} \leq e^{b\rho_{\widetilde{\gamma}}(f_1, f_2)} \frac{\int_{\widetilde{\gamma}} g f_2}{f_2(x)} - e^{-b\rho_{\widetilde{\gamma}}(f_1, f_2)} \frac{\int_{\gamma_e} g f_2}{f_2(x)} \\
\leq& \left(e^{b\rho_\gamma(f_1, f_2)} - e^{-b\rho_\gamma(f_1, f_2)}\right) \frac{\int_{\gamma_e} g f_2}{f_2(x)} + e^{b\rho_\gamma(f_1, f_2)} \frac{\int_I g f_2}{f_2(x)} \\
\leq& 4\delta \left(e^{b\rho_\gamma(f_1, f_2)} - e^{-b\rho_\gamma(f_1, f_2)}\right) |||g|||_+ + e^{b\rho_\gamma(f_1, f_2)} \frac{\int_I g f_2}{f_2(x)} \\
\leq& e^{b\rho_\gamma(f_1, f_2)} \frac{\int_I g f_2}{f_2(x)} + 8\delta e^{b\rho_\gamma(f_1, f_2)} \rho_\gamma(f_1, f_2) bL |||g|||_-,
\end{aligned}
$$

where we have used the fact that the modified functions are constant on $\gamma_e$.

**Remark 5.7.** *It is possible to assume the previous relation in our definition of the cone. This would remove the necessity of being able to extend a weakly stable curve, and therefore slightly weaken our conditions on the cone family. However, I do not know of any concrete example in which such weakening could be useful.*

Let $\gamma \in \Gamma_\delta$ and $\mathcal{J}$ be the collection of the connected pieces of $T^{-n}\gamma$. In $\mathcal{J}_-$ we put the pieces shorter than $\delta$, and we collect the others (after sub-dividing them in pieces belonging to $\Gamma_\delta$) in $\mathcal{J}_+$.

$$
\begin{aligned}
\int_\gamma \widetilde{T}^n g f_1 \leq& \sum_{\widetilde{\gamma} \in \mathcal{J}} \frac{\widehat{T}^n f_1(x_{\widetilde{\gamma}})}{\widehat{T}^n f_2(x_{\widetilde{\gamma}})} \int_{\widetilde{\gamma}} g \widehat{T}^n f_2 e^{\xi b\rho_\gamma(f_1, f_2)} \\
&+ \sum_{\widetilde{\gamma} \in \mathcal{J}_-} \xi \widehat{T}^n f_1(x_{\widetilde{\gamma}}) 8\delta e^{\xi b\rho_\gamma(f_1, f_2)} bL \rho_\gamma(f_1, f_2) |||g|||_-
\end{aligned}
$$

where we have used the fact that the analogues of Lemmas 4.4 and 4.5 hold in this case (providing us with $\xi < 1$).

$$
\begin{aligned}
\frac{\int_\gamma \widetilde{T}^n g f_1}{f_1(x)} \leq& \Bigg[ \sum_{\widetilde{\gamma} \in \mathcal{J}} \frac{f_1(T^n x_{\widetilde{\gamma}}) f_2(x)}{f_2(T^n x_{\widetilde{\gamma}}) f_1(x)} \int_{\widetilde{\gamma}} g \widehat{T}^n f_2 \\
&+ 16\lambda^{-n} M^n \delta bL \rho_\gamma(f_1, f_2) |||\widetilde{T}^n g|||_- \sum_{\widetilde{\gamma} \in \mathcal{J}_-} \frac{f_2(x) f_1(T^n x_{\widetilde{\gamma}})}{f_1(x)} \Bigg] \frac{e^{\xi b\rho_\gamma(f_1, f_2)}}{f_2(x)}.
\end{aligned}
$$

Let us call $\widetilde{\mathcal{J}}_- \subset \mathcal{J}_-$ the set of curves $\widetilde{\gamma}$ such that $\int_{\widetilde{\gamma}} g\widehat{T}^n f_2 < 0$, then

$$\begin{aligned}
\frac{\int_\gamma \widetilde{T}^n g f_1}{f_1(x)} &\leq \Bigg[ \sum_{\widetilde{\gamma}\in\mathcal{J}} e^{\rho_\gamma(f_1,\, f_2)} \int_{\widetilde{\gamma}} g\widehat{T}^n f_2 + \sum_{\widetilde{\gamma}\in\widetilde{\mathcal{J}}_-} \left[e^{\rho_\gamma(f_1,\, f_2)} - e^{-\rho_\gamma(f_1,\, f_2)}\right] \left|\int_{\widetilde{\gamma}} g\widehat{T}^n f_2\right| \\
&\quad + 16\lambda^{-n} M^n \delta b L \rho_\gamma(f_1,\, f_2) |||\widetilde{T}^n g|||_- \frac{f_2(x) f_1(T^n x_{\widetilde{\gamma}})}{f_1(x)}\Bigg] \frac{e^{\xi b \rho_\gamma(f_1,\, f_2)}}{f_2(x)} \\
&\leq \left[1 + 16bd\rho_\gamma(f_1,\, f_2) M^n \lambda^{-n}\right] e^{(\xi b+1)\rho_\gamma(f_1,\, f_2)} \frac{\int_\gamma \widetilde{T}^n g f_2}{f_2(x)} \\
&\leq e^{\left[\xi + b^{-1} + 16 M^n \lambda^{-n} d\right] b\rho_\gamma(f_1,\, f_2)} \frac{\int_\gamma \widetilde{T}^n g f_2}{f_2(x)} \\
&\leq e^{\chi b \rho_\gamma(f_1,\, f_2)} \frac{\int_\gamma \widetilde{T}^n g f_2}{f_2(x)}
\end{aligned}$$

where we have set $b$ large enough, and

$$b^{-1} + \xi + 16(M\lambda^{-1})^n d \leq \chi < 1. \tag{5.8}$$

Let us choose $n_* = A \ln d$, and $\delta$ so small that $N(\delta) \geq 2A \ln d$.

**Averages on different scales**

Next, we need estimates that allow us to compare the average of $\widetilde{T}^n g$ on the scale $\delta$ with the average of $g$ on a fixed scale $\delta_0$; for the time being, the only assumption on $\delta_0$ is $a \in \left(\frac{c_4 \delta_0^{1-\nu}}{1-\lambda^{-\nu}},\, 2^{-\nu} \ln 2\delta_0^{-\nu}\right)$.

We define

$$\begin{aligned}
|||g|||_+^0 &= \sup_{\substack{\gamma\in\widehat{\Gamma}_{\delta_0} \\ f\in\mathcal{D}_a(\gamma)}} \frac{\int_\gamma gf}{\int_\gamma f} \\
|||g|||_-^0 &= \inf_{\substack{\gamma\in\widehat{\Gamma}_{\delta_0} \\ f\in\mathcal{D}_a(\gamma)}} \frac{\int_\gamma gf}{\int_\gamma f}.
\end{aligned}$$

**Lemma 5.8.** *There exists* $N_-(\delta) \leq \frac{N(\delta)\delta^\rho}{16 d \delta_0^\rho} = N_+(\delta)$ *such that, for each* $g \in \widetilde{T}^{[A \ln d]} \mathcal{C}_{b,c,d,L}(\delta)$*, and* $n \in \{N_-(\delta), \ldots, N_+(\delta)\}$ *either*

$$\frac{|||\widetilde{T}^n g|||_+}{|||\widetilde{T}^n g|||_-} \leq \frac{1}{2} \frac{|||g|||_+}{|||g|||_-},$$

*or*

$$\begin{aligned}
|||\widetilde{T}^n g|||_+ &\leq 4 |||g|||_+^0 \\
|||\widetilde{T}^n g|||_- &\geq \frac{1}{2} |||g|||_-^0.
\end{aligned}$$

*Proof.* We start by noticing that, although Lemma 5.4 did not provide us with an invariant cone, we can control a huge number of images.

**Sub-Lemma 5.9.** *There exists $\rho \in \mathbb{R}^+$ such that, given $n \in \{A \ln d, \ldots, \frac{N(\delta)\delta_0^\rho}{16d\delta^\rho}\}$,*

$$\widetilde{T}^n \mathcal{C}_{b,\, c,\, d,\, L} \subset \mathcal{C}_{\chi b,\, \chi c,\, \chi d,\, 4L}$$

*provided that $32L \leq d$. In addition, for each $g \in \mathcal{C}_{b,c,d,L}$*

$$\begin{aligned} |||\widetilde{T}^n g|||_+ &\leq 2|||g|||_+ \\ |||\widetilde{T}^n g|||_- &\geq \frac{1}{2}|||g|||_-. \end{aligned}$$

*Proof.* For each $n \in \{A \ln d, \ldots, \left(\frac{\delta_0}{\delta}\right)^\rho \frac{N(\delta)}{16d}\}$, we write $n = \frac{kN(\delta)}{2} + m$, where we require $m \in \{A \ln d, \ldots, N(\delta)\}$. According to Lemma 5.4 (eq. 5.3, 5.4), for $l \in \{A \ln d, \ldots, N(\delta)\}$

$$\begin{aligned} |||\widetilde{T}^l g|||_- &\geq (1 - 3d\lambda^{-l}M^l)|||g|||_- = \tau_l^{-1}|||g|||_- \\ |||\widetilde{T}^l g|||_+ &\leq \tau_l |||g|||_+ \end{aligned}$$

Moreover, $A$ has been chosen such that $\tau_{A \ln d}^{-1} \geq \frac{1}{\sqrt{2}}$. Setting $\tau = \tau_{N(\delta)/2}$, all we need to check is that

$$\tau^{2(k-1)} \leq 2.$$

The result follows, since

$$\tau^2 \leq e^{6d\lambda^{-\frac{N(\delta)}{2}} M^{\frac{N(\delta)}{2}}},$$

and, defining $\rho = \frac{\ln \lambda M^{-1}}{2 \ln \mu}$,

$$\tau^2 \leq e^{6d\left(\frac{\delta}{\delta_0}\right)^\rho}$$

which implies the Lemma.

The idea of the proof is to iterate $\gamma \in \Gamma_\delta$ until the majority of the pieces produced are longer than $\delta_0$. To be more precise we will iterate pieces of the images in order to obtain long connected pieces, but we will stop iterating them once they become longer than $\delta_0$.

Let $\gamma_0 = \gamma$; let $n_i \in \mathbb{N}$ be the first integer for which, given $\gamma_i \subset \gamma$, $T^{-n_i}\gamma_i$ contains a connected piece longer than $\delta_0$,[19] we will collect such long pieces in $\mathcal{J}_i$, after subdividing them in pieces belonging to $\widehat{\Gamma}_{\delta_0}$.[20] The sets $\gamma_i \subset \gamma$ are defined by induction: $\gamma_{i+1} = \gamma_i \setminus \left(\bigcup_{\widetilde{\gamma} \in \mathcal{J}_i} T^{n_i}\widetilde{\gamma}\right)$. In addition, we will collect all the connected pieces belonging to $T^{-n_i}\gamma_{i+1}$ (by construction such pieces are all shorter than $\delta_0$) into two groups: $\mathcal{J}_i^+$ will contain the pieces longer than $\delta$, and $\mathcal{J}_i^-$ the shorter pieces; $\widetilde{\mathcal{J}}_i = \mathcal{J}_i^+ \cup \mathcal{J}_i^-$.

[19] It is easy to convince oneself that, if $\gamma_i$ contains an interval, then $n_i$ exists finite.

[20] Since, $n_1$ must be larger than $N(\delta)$, choosing $\delta$ small enough, we can insure $N(\delta) \geq l$ (see the definition of $\widehat{\mathcal{C}}_-$ and $\widehat{\Gamma}$ at the beginning of the section).

By construction $n_{i+1} > n_i$, and $T^{-n_i}\gamma_i$ consists, at most, of $M^{n_i}$ connected pieces. For each $i_* > 0$, and $m \geq n_{i_*}$, provided $\left(\frac{\delta}{\delta_0}\right)^{-\rho} \frac{N(\delta)}{16d} \geq m$,

$$\begin{aligned}\frac{\int_\gamma \widetilde{T}^m g f}{\int_\gamma f} &= \sum_{i=0}^{i_*} \sum_{\widetilde{\gamma} \in \mathcal{J}_i} \frac{\int_{\widetilde{\gamma}} \widetilde{T}^{m-n_i} g \widehat{T}^{n_i} f}{\int_\gamma f} + \sum_{\widetilde{\gamma} \in \widetilde{\mathcal{J}}_{i_*}} \frac{\int_{\widetilde{\gamma}} \widetilde{T}^{m-n_{i_*}} g \widehat{T}^{n_{i_*}} f}{\int_\gamma f} \\ &\leq \sup_{n \in \{0,\dots,m\}} |||\widetilde{T}^n g|||_+^0 + \frac{2}{|\gamma|} |||\widetilde{T}^{m-n_{i_*}} g|||_+ \sum_{\widetilde{\gamma} \in \mathcal{J}_{i_*}^+} |T^{n_{i_*}} \widetilde{\gamma}| \\ &+ 2d(\lambda^{-1}M)^{n_{i_*}} |||\widetilde{T}^{m-n_{i_*}} g|||_- .\end{aligned}$$

To conclude the argument, we have to choose $i_*$ large enough, so that the unwanted terms in the previous equation are negligible but not beyond the applicability of Sub-Lemma 5.9 $\left(\text{i.e., } n_{i_*} \leq \left(\frac{\delta}{\delta_0}\right)^{-\rho} \frac{N(\delta)}{16d}\right)$.

Since $n_{i_*} \geq N(\delta)$, it follows that $2d(\lambda^{-1}M)^{n_{i_*}} \leq \frac{1}{8}$, provided $\delta$ is sufficiently small. Let $i_*$ be the first integer $i$ for which $\sum_{\widetilde{\gamma} \in \widetilde{\mathcal{J}}_i} |T^{n_i}\widetilde{\gamma}| = |\gamma_{i+1}| \leq \frac{|\gamma|}{128L}$. This means that $|\gamma_{i_*}| > \frac{\delta}{128L}$; in addition, $T^{-k}\gamma_{i_*}$ will consist, at most, of $M^k$ pieces, shorter than $\delta_0$, for each $k < n_{i_*}$. Let us call $\mathcal{J}_k^*$ the collection of such pieces.

$$\frac{\delta}{128L} \leq \sum_{\widetilde{\gamma} \in \mathcal{J}_k^*} \frac{|T^k \widetilde{\gamma}|}{|\widetilde{\gamma}|} |\widetilde{\gamma}| \leq \sum_{\widetilde{\gamma} \in \mathcal{J}_k^*} \frac{|T^k \widetilde{\gamma}|}{|\widetilde{\gamma}|} \delta_0 \leq \left(\lambda^{-1} M\right)^k \delta_0.$$

The previous computation shows that, setting $k_0 = \frac{\ln(128L\delta_0\delta^{-1})}{\ln \lambda M^{-1}}$, $n_{i_*} \leq k_0 + 1$.

Using our choice of $i_*$ and the previous estimates, we conclude that, for each $m \in \{k_0+1, \dots, N_+(\delta)\}$,

$$\frac{\int_\gamma \widetilde{T}^m g f}{\int_\gamma f} \leq \sup_{n \in \{0,\dots,m\}} |||\widetilde{T}^n g|||_+^0 + \frac{1}{4} |||\widetilde{T}^m g|||_- .$$

In the same manner one can prove:

$$\frac{\int_\gamma \widetilde{T}^m g f}{\int_\gamma f} \geq (1 - 2\lambda^{-1}M) \inf_{n \in \{0,\dots,m\}} |||\widetilde{T}^n g|||_-^0 - \frac{1}{4} |||\widetilde{T}^m g|||_- .$$

To obtain a useful inequality it is necessary to estimate $|||\widetilde{T}^n g|||_\pm^0$. This is done by the following.

**Sub-Lemma 5.10.** *For each $g \in \widetilde{T}^{[A \ln d]} \mathcal{C}_{b,c,d,L}(\delta)$, holds*

$$\begin{aligned}|||\widetilde{T}^n g|||_+^0 &\leq |||g|||_+^0 + \frac{1}{8} |||g|||_+ \\ |||\widetilde{T}^n g|||_-^0 &\geq \frac{3}{4} |||g|||_-^0,\end{aligned}$$

*for each* $n \leq N_+(\delta)$.

*Proof.* The proof is by induction. Let $\gamma \in \widehat{\Gamma}_{\delta_0}$, if $\mathcal{J}_1$ contains the connected pieces of $T^{-1}\gamma$ longer than $\delta_0$, $\mathcal{J}_2$ the one shorter than $\delta_0$ but longer than $\delta$, $\mathcal{J}_3$ the shorter pieces, and $x_{\widetilde{\gamma}} \in \widetilde{\gamma}$ is an arbitrary point, then

$$\begin{aligned}\frac{\int_\gamma \widetilde{T}gf}{\int_\gamma f} &\leq \sum_{\widetilde{\gamma}\in\mathcal{J}_1}\frac{\int_\gamma g\widehat{T}f}{\int_\gamma f} + \sum_{\widetilde{\gamma}\in\mathcal{J}_2}\frac{\int_{\widetilde{\gamma}}\widehat{T}f}{\int_\gamma f}|||g|||_+ + \sum_{\widetilde{\gamma}\in\mathcal{J}_3}\frac{\widehat{T}f(x_{\widetilde{\gamma}})}{\int_\gamma f}\delta d|||g|||_- \\ &\leq |||g|||_+^0 + (\lambda^{-1}M)\left[2 + 2d\frac{\delta}{\delta_0}\right]|||g|||_+ \leq |||g|||_+^0 + 3(\lambda^{-1}M)|||g|||_+\end{aligned}$$

provided $\delta$ is sufficiently small.

Our induction hypothesis is

$$|||\widetilde{T}^n g|||_+^0 \leq |||g|||_+^0 + 3|||g|||_+\sum_{i=1}^n(\lambda^{-1}M)^i.$$

We use a construction similar to the one used previously. We collect in $\mathcal{J}_1$ all the pieces of $T^{-1}\gamma$ longer than $\delta_0$, in $\mathcal{J}_2$ all the pieces $\widetilde{\gamma}$ of $T^{-2}\gamma$ longer than $\delta_0$, provided $T\widetilde{\gamma}$ was not contained in an element of the collection $\mathcal{J}_1$, and so on. The connected pieces $\widetilde{\gamma}$ of $T^{-n}\gamma$ that are shorter than $\delta_0$, and such that $T^i\widetilde{\gamma}$, $i \in \{1, \dots, n\}$, was always contained in maximal connected pieces shorter than $\delta_0$, are collected in two sets: $\mathcal{J}_+$ contains the pieces longer than $\delta$, and $\mathcal{J}_-$ the shorter ones.

$$\begin{aligned}\frac{\int_\gamma \widetilde{T}^{n+1}gf}{\int_\gamma f} &\leq \sum_{i=0}^n\sum_{\widetilde{\gamma}\in\mathcal{J}_{i+1}}\frac{\int_{\widetilde{\gamma}}\widetilde{T}^{n-i}g\widehat{T}^{i+1}f}{\int_\gamma f} + \sum_{\widetilde{\gamma}\in\mathcal{J}_+}\frac{\int_{\widetilde{\gamma}}g\widehat{T}^{n+1}f}{\int_\gamma f} + \frac{2(M\lambda^{-1})^{n+1}d\delta}{\delta_0}|||g|||_- \\ &\leq \sum_{i=0}^n\sum_{\widetilde{\gamma}\in\mathcal{J}_{i+1}}\frac{\int_{\widetilde{\gamma}}\widehat{T}^{i+1}f}{\int_\gamma f}\left\{|||g|||_+^0 + 3|||g|||_+\sum_{j=1}^n(\lambda^{-1}M)^j\right\} \\ &\quad + 3(\lambda^{-1}M)^{n+1}|||g|||_+ \leq |||g|||_+^0 + 3|||g|||_+\sum_{i=1}^{n+1}(\lambda^{-1}M)^i\end{aligned}$$

which implies

$$|||\widetilde{T}^n g|||_+^0 \leq |||g|||_+^0 + \frac{3(\lambda^{-1}M)}{1-\lambda^{-1}M}|||g|||_+ \leq |||g|||_+^0 + \frac{1}{8}|||g|||_+.$$

The second inequality is proven in a similar way. In fact,

$$\begin{aligned}\frac{\int_\gamma \widetilde{T}gf}{\int_\gamma f} &\geq \sum_{\widetilde{\gamma}\in\mathcal{J}_1}\frac{\int_\gamma g\widehat{T}f}{\int_\gamma f} - 2d\lambda^{-1}M\frac{\delta}{\delta_0}|||g|||_- \geq (1-2\lambda^{-1}M)|||g|||_+^0 - \frac{M\lambda^{-1}}{2}|||g|||_- \\ &\geq (1-2\lambda^{-1}M)|||g|||_+^0 - M\lambda^{-1}|||\widetilde{T}g|||_-,\end{aligned}$$

since, for each function $\widetilde{g} \in \mathcal{C}_{b,c,d,L}$, $|||\widetilde{g}|||_- \le |||\widetilde{g}|||^0_-$, we have

$$\frac{\int_\gamma \widetilde{T} g f}{\int_\gamma f} \ge \frac{1 - 2\lambda^{-1}M}{1 + \lambda^{-1}M} |||g|||^0_+.$$

This time, we will use the following induction hypothesis

$$|||\widetilde{T}^n g|||^0_- \ge h_n |||g|||^0_-$$

where $h_1 = \frac{1-2\lambda^{-1}M}{1+\lambda^{-1}M}$ and $h_{n+1} = \frac{1-2(\lambda^{-1}M)^{n+1}}{1+(\lambda^{-1}M)^{n+1}} h_n$.

$$\frac{\int_\gamma \widetilde{T}^{n+1} g f}{\int_\gamma f} \ge (1 - 2(\lambda^{-1}M)^{n+1}) h_n |||g|||^0_- - \frac{(\lambda^{-1}M)^{n+1}}{2} |||g|||_-.$$

This concludes the argument, since

$$h_n = \prod_{i=1}^{n} \frac{1 - 2(\lambda^{-1}M)^i}{1 + (\lambda^{-1}M)^i} \ge e^{-4\sum_{i=1}^{n}(\lambda^{-1}M)^i} \ge \frac{3}{4}.$$

¿From the previous Lemma we obtain immediately

$$\frac{5}{4} |||\widetilde{T}^n g|||_- \ge (1 - \frac{2}{25}) \frac{3}{4} |||g|||^0_-$$

which implies the last inequality in the statement of the Lemma.

To conclude the argument let us consider two cases. First suppose that

$$|||\widetilde{T}^m g|||_+ \le \frac{1}{4} |||g|||_+.$$

This would immediately imply the first possibility in the statement. Second, suppose that

$$|||\widetilde{T}^m g|||_+ \ge \frac{1}{4} |||g|||_+.$$

The above inequality, together with the inequalities obtained before, yields

$$|||\widetilde{T}^m g|||_+ \le |||g|||^0_+ + \frac{3}{4} |||\widetilde{T}^m g|||_+,$$

which conclude the proof.

### Comparison of averages

The only task left is to compare the value of the integral on different $\gamma$'s. Due to the presence of the singularities, this is much more complicated that in the smooth case. Nonetheless, the basic idea is the same. The only difference is that we will break up a curve in $\widehat{\Gamma}_{\delta_0}$ into pieces belonging to $\widehat{\Gamma}_{\delta_1}$, with $\delta_0 \ge \delta_1 \ge \delta$, and we will compare such pieces. The essential feature of this decomposition is that $\delta_1$ can be chosen smaller than $\delta_0$, depending on the parameters of the cone.

To carry out this plan we need to introduce appropriate neighborhoods of $\mathcal{S}^-$.

**Definition 5.11.**
$$\mathcal{S}_\varepsilon^- = \{z \in \mathcal{M} \mid \ dist(z, \mathcal{S}^-) \leq \varepsilon\}$$
$$\Sigma_n = \bigcup_{i=0}^{n} T^i\left(\mathcal{S}^-_{\lambda^{-i}\delta_1}\right).$$

*Moreover, we will say that two curves $\gamma_1, \gamma_2 \in \widehat{\Gamma}_\varepsilon$ are the translate of each other in the unstable direction if each sufficiently long unstable manifold intersecting one curve intersects the other as well.*

We are now able to state our main comparison Lemma.

**Lemma 5.12.** *There exists $m_* \in \mathbb{N}$ such that, if two curves $\gamma_1, \gamma_2 \in \widehat{\Gamma}_{\delta_1}$ are the translate of each other in the unstable direction, their distance is less than $\delta_1/2$, and $\gamma_1 \bigcap \Sigma_{m_*} = \emptyset$, then*

$$\frac{1}{2}\int_{\gamma_2} g - \frac{e^{-bR}}{256}|||g|||_-\delta_1 \leq \int_{\gamma_1} g \leq 2\int_{\gamma_2} g + \frac{e^{-bR}}{256}|||g|||_-\delta_1$$

*Proof.* The Lemma is a consequence of the results stated in Appendix II.

$$\begin{aligned}\int_{\gamma_1} g &= \int_{D(\phi)} g + \int_{\gamma_1\setminus D(\phi)} g \leq \int_{R(\phi)} g\circ\phi^{-1}J_\phi + D\delta_1(\lambda^{-\zeta}\mu^{1-\zeta})^{m_*}|||g|||_- \\ &\leq \int_{R(\phi)} gJ_\phi + \|D^u g\|_\infty \frac{\delta_1}{2}\int_{R(\phi)} J_\phi + D\delta_1(\lambda^{-\zeta}\mu^{1-\zeta})^{m_*}|||g|||_-\end{aligned}$$

Here we have used Lemma II.2. To use effectively the properties of $g$ it is necessary for the domain of integration to be connected, while $R(\phi)$ it is not. But, from Lemma II.5 and II.6 it follows that both $\phi$ and $J_\phi$ can be extended to continuous functions defined on all $\gamma_2$,[21] and the extensions satisfies the same bounds (i.e., the extension of $\phi$ is Liptchitz and the extension of $J_\phi$ is Hölder continuous). Let us use a tilde to identify such extensions.

$$\int_{\gamma_1} g \leq \int_{\gamma_2} g\widetilde{J}_\phi + c|||g|||_-\delta_1^2 + 2D\delta_1(\lambda^{-\zeta}\mu^{1-\zeta})^{m_*}|||g|||_-.$$

Lemma II.6 shows that for $a$ large and $\delta_1$ small enough, $\widetilde{J}_\phi \in \mathcal{D}_a(\gamma_2)$, therefore

$$\int_{\gamma_1} g \leq e^{b\rho_{\gamma_2}(1,\widetilde{J}_\phi)}\int_{\gamma_2} g + \left(2D(\lambda^{-\zeta}\mu^{1-\zeta})^{m_*} + c\delta_1\right)|||g|||_-\delta_1.$$

The result follows by choosing $\delta_1$ sufficiently small, and $m_*$ sufficiently large, with respect to $b$ and $c$.

To conclude this part of the argument, as we did in §4, we need to construct an appropriate covering of $\mathcal{M}$. We will use such covering to control the images of a given curve.

[21] For example, one can take the extension to be linear where the functions are not defined.

We start with the same collection of squares $Q(z_i)$ constructed in §4–"Diameter of the image".[22] As in the previous section given any square we have a notion of "core," although a more stringent one. In the previous section was called "core of a square" a concentrical square of size $\frac{1}{4}$, here we will call "weak-core" a concentrical square of size $\frac{1}{4t}$ ($t > 1$ will be chosen in Sub-Lemma 5.16), while by "core" of a square of size $t\varepsilon$, we mean the set of points in the weak-core that have a stable and unstable manifold longer than $4\varepsilon t$. In addition, we will call a square "$\alpha$–connecting" (for some $\alpha < 1$) if the measure of its core is at least $\alpha$ times the measure of the square.[23]

The key ingredient in the construction of our covering is a result very similar to the "Sinai Theorem," which is the cornerstone of the proof of ergodicity for the systems under consideration. In fact, the proof of the following theorem is almost indistinguishable from the proof of Sinai Theorem in the simple uniformly hyperbolic case (I refer to the version in [LW, §3]). However, since the statement looks stronger, a sketch of a proof is provided in Appendix I.

**Sinai Theorem 5.13.** *For each square $Q$ of size $\delta_0$,and $t > 1$, there exist $\alpha_Q < 1$, $\delta_Q \leq \delta_0$, and $m_Q \in \mathbb{N}$ such that, for each $\alpha \leq \alpha_Q$, $m \geq m_Q$, and $\delta_1 \leq \delta_Q$ we can construct a covering $\mathcal{G}(Q, \delta_1, m)$ made of squares of size $t\delta_1$ with the following properties:*

- i) *the number of squares $G$ for which $\bigcup_{|n|\leq m_Q} T^n\mathcal{S} \cap G = \emptyset$, and $G$ is not $\alpha$-connecting, is less than $\frac{1}{256}K_Q^{-1}$ ($K_Q$ being the maximal number of overlaps in the covering);*
- ii) *for each $\gamma \in \widehat{\Gamma}_{\delta_1}$, $\gamma \subset Q$, there exists $G \in \mathcal{G}(Q, \delta_1, m)$ such that $\gamma \subset G$ and $\gamma$ intersects properly (see §4) $G$.*

In other words, a square $Q$ can be covered by the weak–core of sufficiently small squares whose overwhelming majority contains an abundance of stable fibers.

According to Theorem 5.13, for every square $Q \in \{Q(z_i)\}$ we can construct coverings $\mathcal{G}(Q, \delta_1, m)$. Let us choose $m_0 \geq \max\{\{m_Q\}_{Q\in\mathcal{Q}}, m_*\}$, and $\delta_1 \leq \min\{\{\delta_Q\}_{Q\in\mathcal{Q}}, \mu^{-m_0}\delta_0\}$, and $\alpha = \min\{\alpha_Q\}$, then $\mathcal{G} = \cup_{Q\in\mathcal{Q}}\mathcal{G}(Q, \delta_1, m_0)$ is the covering of $\mathcal{M}$ that will be used in the following.

We are now positioned to make the last step: estimate the ratio of the averages; as in §4, a new element, the mixing property, allows us to close the argument.

**Lemma 5.14.** *There exists $B \in \mathbb{R}$ and $N \in \mathbb{N}$ such that, for $m \in \{N_-(\delta), \dots, N_+(\delta)\}$,*

$$|||\widetilde{T}^m g|||_+ \leq B\mu^N |||\widetilde{T}^m g|||_-.$$

*Proof.* We have assumed the map to be mixing. Hence, there exists $N \in \mathbb{N}$ such that, given any two $\alpha$-connecting $G, G' \in \mathcal{G}$

$$T^{-N}(\mathrm{core}(G)) \cap \mathrm{core}(G') \neq \emptyset.$$

[22] The presence of singularities restrict the possible points that can be chosen as the center of a square: only points with well defined stable and unstable direction are allowed. Nevertheless, almost all points are available, and the squares can still be chosen to be of uniform size.

[23] The definition of core differs from the one given in [LW, Definition 11.1]: there the requirement on the size of the manifolds is absent. Nevertheless, the definition of $\alpha$-connecting is essentially the same.

The strategy of the proof is as follows: first we will show that it is possible to control the average on any curve of size $\delta_0$ via the average on a shorter curve contained in a connecting square (Sub-Lemma 5.15). Then we will see that the mixing property implies that a connected piece belonging to some image of such shorter curve is $\delta_1$-close to any given curve. We will then be in a position to compare any two curves.

**Sub-Lemma 5.15.** *For each $\delta \leq \delta_1$ small enough, $g \in \mathcal{C}_{b,c,d,4L}(\delta)$, $\gamma \in \widehat{\Gamma}_{\delta_0}$, $f \in \mathcal{D}_a(\gamma)$, there exists $\gamma^+$, $\gamma^- \in \widehat{\Gamma}_{\delta_1}$ ($\gamma^+$, $\gamma^- \subset \gamma$), such that*

$$\frac{1}{2}\frac{\int_{\gamma^-} fg}{\int_{\gamma^-} f} \leq \frac{\int_\gamma fg}{\int_\gamma f} \leq \frac{\int_{\gamma^+} fg}{\int_{\gamma^+} f} + \frac{1}{64}|||g|||_+.$$

*In addition, both $\gamma^+$ and $\gamma^-$ intersect properly an $\alpha$–connecting squares belonging to $\mathcal{G}$, and are disjoint from $\bigcup_{|n|\leq m_0} T^n\mathcal{S}$.*

*Proof.* Consider the set $\gamma\backslash(\bigcup_{|n|\leq m_0} T^n\mathcal{S})$. Since $\delta_1 \leq \delta_0\mu^{-m_0}$, it consists of, at most, $2M^{m_0}$ connected curves. We collect in $\mathcal{J}_1$ the ones that are longer than $\beta\delta_1$; in $\mathcal{J}_2$ we collect the curves shorter than $\beta\delta_1$, but longer than $\delta$; finally in $\mathcal{J}_3$ we collect the shorter ones.

$$\frac{\int_\gamma fg}{\int_\gamma f} \leq \sum_{\widetilde{\gamma}\in\mathcal{J}_1} \frac{\int_{\widetilde{\gamma}} fg}{\int_\gamma f} + \sum_{\widetilde{\gamma}\in\mathcal{J}_2} \frac{\int_{\widetilde{\gamma}} f}{\int_\gamma f}|||g|||_+ + \sum_{\widetilde{\gamma}\in\mathcal{J}_3} \frac{f(x_{\widetilde{\gamma}})}{\int_\gamma f} d\delta|||g|||_-.$$

Next we split $\mathcal{J}_1$ into three parts: the curves properly contained in a connecting square, collected in $\mathcal{J}_0$, after subdividing them into curves belonging to $\widehat{\Gamma}_{\delta_1}$; the curves closest to $\bigcup_{|n|\leq m_0} T^n\mathcal{S}$, collected in $\mathcal{J}_4$; and the rest, collected in $\mathcal{J}_5$. Since we have assumed the uniform transversality between $\mathcal{S}^\pm$ and the stable and unstable directions, and since the direction of $\gamma'$, for $\gamma \in \widehat{\Gamma}_\delta$, may differ from the stable direction by, at most, some fixed arbitrary $\epsilon$, we can assume, without loss of generality, that the curves in $\widehat{\Gamma}_\delta$ are uniformly transversal to $\mathcal{S}^\pm$. We can therefore choose $\beta$ so large that, if $z \in \widetilde{\gamma} \in \widehat{\Gamma}_\delta$ and the distance of $z$ from $\mathcal{S}^\pm$, along $\gamma$, is larger than $\beta$, than $d(z, \mathcal{S}^\pm) \geq t\delta_1$. Hence, by construction, the curves in $\mathcal{J}_5$ belong to squares that do not intersect $\bigcup_{|n|\leq m_0} T^n\mathcal{S}$. The Sinai Theorem implies that the total length of the curves in $\mathcal{J}_5$ must be less than $\frac{\delta_0}{256}$. Among the curves in $\mathcal{J}_0$ there exists one on which the average is larger, let us call it $\gamma^+$, then

$$\frac{\int_\gamma fg}{\int_\gamma f} \leq \frac{\int_{\gamma^+} fg}{\int_{\gamma^+} f} + 2\left(\frac{1}{256} + 4\beta M^{m_0}\frac{\delta_1}{\delta_0}\right)|||g|||_+ + 4dM^{m_0}\frac{\delta}{\delta_0}|||g|||_-.$$

Similarly, we have

$$\frac{\int_\gamma fg}{\int_\gamma f} \geq \sum_{\widetilde{\gamma}\in\mathcal{J}_0} \frac{\int_{\widetilde{\gamma}} fg}{\int_{\widetilde{\gamma}} f} - \sum_{\widetilde{\gamma}\in\mathcal{J}_3} \frac{\left|\int_{\widetilde{\gamma}} gf\right|}{\int_{\widetilde{\gamma}} f}.$$

Let $\gamma^-$ be the curve in $\mathcal{J}_0$ on which the average is smaller, then

$$\frac{\int_\gamma fg}{\int_\gamma f} \geq \frac{3}{4}\frac{\int_{\gamma^-} fg}{\int_{\gamma^-} f} - 4dM^{m_0}\frac{\delta}{\delta_0}|||g|||_- \geq \frac{3}{4}\frac{\int_{\gamma^-} fg}{\int_{\gamma^-} f} - 4dM^{m_0}\frac{\delta}{\delta_0}\frac{\int_\gamma fg}{\int_\gamma f}$$

The Lemma follows by choosing $\delta$ and $\delta_1$ sufficiently small.

Given $g \in \mathcal{C}_{b,c,d,4L}(\delta)$, choose $\gamma^* \in \widehat{\Gamma}_{\delta_0}$, and $f_* \in \mathcal{D}_a(\gamma^*)$, such that

$$|||g|||_+^0 \leq 2\frac{\int_{\gamma^*} gf_*}{\int_{\gamma^*} f_*}.$$

Then, there exists $\gamma_+^* \in \widehat{\Gamma}_{\delta_1}$, and an $\alpha$-connecting square $G_* \in \mathcal{G}$ containing it, such that

$$|||g|||_+^0 \leq 2\frac{\int_{\gamma_+^*} gf_*}{\int_{\gamma_+^*} f_*} + \frac{1}{32}|||g|||_+.$$

Furthermore, for each $g \in \mathcal{C}_{b,c,d,4L}(\delta)$, $\gamma \in \widehat{\Gamma}_{\delta_0}$, and $f \in \mathcal{D}_a(\gamma)$, there exists $\gamma_- \in \widehat{\Gamma}_{\delta_1}$, and an $\alpha$-connecting square $G_\gamma \in \mathcal{G}$ such that $\gamma_- \subset G_\gamma$, and

$$\frac{\int_\gamma fg}{\int_\gamma f} \geq \frac{1}{2}\frac{\int_{\gamma_-} fg}{\int_{\gamma_-} f}.$$

We can finally conclude our argument.

**Sub-Lemma 5.16.** *$T^{-N}\gamma_-$ contains a translate of $\gamma_+^*$.*

*Proof.* Let $w \in \operatorname{core}(G_\gamma) \cap T^N\operatorname{core}(G_*)$, then, by definition, there exists a stable manifold $W^s$ and an unstable manifold $W^u$ of size $a\delta_1 t$ (consequently intersecting $G_\gamma$ completely), and $W^s \cap W^u = \{w\}$. This implies $W^u \cap \gamma_- \neq \emptyset$. Let $y = W^u \cap \gamma_-$, and $\widetilde{\gamma}_n$ the connected component of $T^{-n}\gamma_- \cap G_*$ containing $T^{-n}y$.

Since $T^{-n}W^u \cap \mathcal{S}^- = \emptyset$, and since $\mathcal{S}^-$ is uniformly transversal to $W^u$, there exits $\theta \in \mathbb{R}^+$ such that, for each $z \in T^{-n}W^u$,

$$d(z, \mathcal{S}^-) \geq \theta \min\{d(z, T^{-n}w_1), d(z, T^{-n}w_2)\}$$

where $\{w_1, w_2\} = \partial(W^u \cap G_\gamma)$. Let $w_1$ be the point of $\partial(W^u \cap G_\gamma)$ further away from $y$, and $\widetilde{w}_1$ the point of $\partial(W^s(T^{-N}w) \cap G_*)$ further away from $T^{-N}y$. Then, setting $W_1^s = W^s(T^{-N}w) \cap G_*$,

$$\begin{aligned} d(T^{-n}y, \mathcal{S}^-) \geq& \theta d(T^{-n}y, T^{-n}w_1) \\ d(T^{-N+n}y, \mathcal{S}^-) \geq& d(T^{-N+n}w, T^n\widetilde{w}_1) - \theta d(\widetilde{\gamma}_{N-n}, T^nW_1^s). \end{aligned}$$

Next, we choose $n_* \in \{0, \ldots, N\}$ such that, for the first time,

$$d(T^{-n_*}w, T^{N-n_*}\widetilde{w}_1) \geq 4d(\widetilde{\gamma}_{n_*}, T^{N-n_*}W_1^s).$$

Then, for each $n \le n_*$,
$$|\widetilde{\gamma}_n| \ge \theta d(T^{-n}y, T^{-n}w_1).$$

In fact, the latter relation is certainly true for $n = 0$, now suppose that it is true for $n-1$ and let $\widehat{\gamma}_n$ be the connected component of $\widetilde{\gamma}_{n-1} \cap \mathcal{S}^-$ containing $T^{-n}y$, then
$$|\widehat{\gamma}_n| \ge \theta d(T^{-n+1}y, T^{-n+1}w_1)$$

and
$$|\widetilde{\gamma}_n| = |T^{-1}\widehat{\gamma}_n| \ge \lambda^2\theta d(T^{-n}y, T^{-n}w_1).$$

Accordingly,

$$\begin{aligned}|\widehat{\gamma}_{n_*}| &\ge \theta d(T^{-n_*+1}y, T^{-n_*+1}w_1) \ge \theta d(\widetilde{\gamma}_{n_*-1}, T^{N-n_*+1}W_1^s)\\ &\ge \frac{\theta}{4}d(T^{-n_*+1}w, T^{N-n_*+1}\widetilde{w}_1)\end{aligned}$$

and
$$|\widetilde{\gamma}_{n_*}| \ge \frac{\theta}{4}e^{-c_4\lambda^{-n_*}\delta_1 t}d(T^{-n_*}w, T^{N-n_*}\widetilde{w}_1).$$

where we have used the usual distortion estimate and $c_4$ is some fixed constant. ¿From our bounds on the distance between the images of $y$ and $\mathcal{S}^-$ follows that $\widetilde{\gamma}_n \cap \mathcal{S}^- = \emptyset$, hence

$$|\widetilde{\gamma}_n| \ge \frac{\theta}{4}e^{-c_4\lambda^{-n_*}\sum_{i=0}^{n-n_*}\lambda^i\delta_1 t}d(T^{-n}w, T^{N-n}\widetilde{w}_1)$$

and
$$|\widetilde{\gamma}_N| \ge \frac{\theta}{4}e^{-c_4\delta_1 t}\delta_1 t \ge 6\delta_1$$

where we have chosen $t$ sufficiently large ($t \ge 52\theta$) and $\delta_1$ sufficiently small ($e^{-c_4\delta_1 t} \ge \frac{1}{2}$). Moreover, from the argument is clear that dividing $\widetilde{\gamma}_N$ at $T^{-N}y$ each of the two pieces is long, at least, $3\delta_1$. From this follows that $\widetilde{\gamma}_N$ must contain a translate of $\gamma_+^*$.

Let $\gamma_0 \subset T^{-N}\gamma_-$ be a translate of $\gamma_*^+$, then

$$\begin{aligned}\frac{\int_\gamma f\widetilde{T}^N g}{\int_\gamma f} &\ge \frac{1}{2}\frac{\int_{\gamma^-}\widetilde{T}^N gf}{\int_{\gamma^-}f} \ge \frac{1}{2}\frac{\int_{\gamma_0}g\widehat{T}^N f}{\int_{\gamma^-}f} \ge \frac{e^{-bR}}{8\delta_1}\mu^{-N}\int_{\gamma_0}g\\ &\ge \frac{e^{-bR}}{16\delta_1}\mu^{-N}\int_{\gamma_*^+}g - \frac{e^{-2bR}\mu^{-N}}{2048\delta_1}|||g|||_+\\ &\ge \frac{e^{-2bR}}{32}\mu^{-N}\frac{\int_{\gamma_+^*}gf_*}{\int_{\gamma_+^*}f_*} - \frac{e^{-2bR}\mu^{-N}}{2048\delta_1}|||g|||_+\\ &\ge \frac{e^{-2bR}}{64}\mu^{-N}\left(|||g|||_+^0 - \frac{1}{16}|||g|||_+\right) \ge \frac{e^{-2bR}}{64}\mu^{-N}\left(|||g|||_+^0 - \frac{L}{16}|||g|||_-\right).\end{aligned}$$

Accordingly,

$$|||\widetilde{T}^N g|||_-^0 \geq \frac{e^{-2bR}\mu^{-N}}{64}\left(|||g|||_+^0 - \frac{L}{16}|||g|||_-\right).$$

Lemma 5.8 implies that there exists $m \in \{N_-(\delta), \ldots, N_+(\delta)\}$ such that, calling $\widetilde{g} = \widetilde{T}^{[A \ln d]} g$, we have

$$|||\widetilde{T}^{m+N}\widetilde{g}|||_- \geq \frac{1}{2}|||\widetilde{T}^N\widetilde{g}|||_-^0 \geq \frac{e^{-2bR}\mu^{-N}}{128}\left(|||\widetilde{g}|||_+^0 - \frac{L}{16}|||\widetilde{g}|||_-\right).$$

If $|||\widetilde{g}|||_+^0 \leq \frac{3L}{32}|||\widetilde{g}|||_-$, then

$$|||\widetilde{T}^m\widetilde{g}|||_+ \leq 4|||\widetilde{g}|||_+^0 \leq \frac{3L}{8}|||\widetilde{g}|||_- \leq \frac{3L}{4}|||\widetilde{T}^m\widetilde{g}|||_-;$$

otherwise

$$|||\widetilde{T}^{m+N}\widetilde{g}|||_- \geq \frac{e^{-2bR}\mu^{-N}}{128 \times 12}|||\widetilde{T}^{m+N}\widetilde{g}|||_+.$$

### Decay of correlations

Lemma 5.14 was the last piece in the proof of Theorem 5.3: it suffices to choose

$$L > 1536\chi^{-1}\mu^N e^{2bR},$$

and

$$m = A \ln d + N_-(\delta) + N,$$

in order to close the argument.[24]

We announced at the beginning of the section that Theorem 5.3 was sufficient for our purposes; this is shown by the next Lemma.

**Lemma 5.17.** *There exists* $\Delta \in \mathbb{R}^+$*:*

$$diam(\mathcal{C}_{\chi b,\, \chi c,\, \chi d,\, \chi L}(\delta)) \leq \Delta$$

*where the diameter in computed with respect to the Hilbert metric* $\Theta$ *of* $\mathcal{C}_{b,\, c,\, d,\, L}(\delta)$.

*Proof.* This Lemma is similar to Lemma 4.5, only more computations are needed since we are now dealing with a more complex geometry.

Let $g \in \mathcal{C}_{\chi b,\, \chi b,\, \chi d,\, \chi L}$, if $\lambda \preceq g \preceq \mu$, then, in analogy with Lemma 4.5, we have

[24]At this point the reader may wonder if such a choice is possible. In fact, $N$ depends on $\delta_1$, but not on $\delta$. In turn, the choice of $\delta_1$ depends on $c$ and $b$, but not on $d$ and $L$. It is therefore possible to choose al the parameters involved but $L$, $d$ and $\delta$. The next relation forces then a choice of $L$ (and, consequently, of $d$), nonetheless we are still free to choose $\delta$ sufficiently small to satisfy all the necessary requirements.

$$\lambda \leq \inf_{\substack{\gamma\in\Gamma_\delta\\ f\in\mathcal{D}_a(\gamma)}} \frac{\int_\gamma gf}{\int_\gamma f} = \alpha_0$$

$$\lambda \leq \inf_{\substack{\gamma\in\Gamma_\delta\\ f_1,\, f_2\in\mathcal{D}_a(\gamma)}} \frac{e^{b\rho_\gamma(f_1,\, f_2)} f_1\int_\gamma gf_2 - f_2\int_\gamma gf_1}{e^{b\rho_\gamma(f_1,\, f_2)} f_1\int_\gamma f_2 - f_2\int_\gamma f_1} = \alpha_1$$

$$\lambda \leq \frac{c|||g|||_- - \|D^u g\|_\infty}{c} = \alpha_2$$

In addition, in the present context, we have two extra conditions defining the cone that produce the following constraints:

$$\lambda \leq \inf \frac{d\delta^{1-\zeta}|I|^\zeta |||g|||_- - \frac{|\int_I gf|}{f(x)}}{d\delta^{1-\zeta}|I|^\zeta - \frac{\int_I f}{f(x)}} = \alpha_3$$

or

$$\lambda \leq \inf \frac{d\delta^{1-\zeta}|I|^\zeta |||g|||_- + \frac{|\int_I gf|}{f(x)}}{d\delta^{1-\zeta}|I|^\zeta + \frac{\int_I f}{f(x)}} = \alpha_4$$

and

$$\lambda \leq \frac{L|||g|||_- - |||g|||_+}{L-1} = \alpha_5,$$

again $\alpha = \min\{\alpha_i\}$. ¿From Lemma 4.5 we know that $\min\{\alpha_0,\, \alpha_1,\, \alpha_2\} \geq \frac{(1-\chi)b}{b+1}|||g|||_-$, a direct computation shows that

$$\alpha_3 \geq \inf \frac{(1-\chi)d\delta^{1-\zeta}|I|^\zeta}{d\delta^{1-\zeta}|I|^\zeta - \frac{1}{2}|I|}|||g|||_- = (1-\chi)|||g|||_-$$

$$\alpha_4 \geq \inf \frac{d\delta^{1-\zeta}|I|^\zeta}{d\delta^{1-\zeta}|I|^\zeta + 2|I|}|||g|||_- = \frac{d}{d+2}|||g|||_-$$

$$\alpha_5 \geq \frac{(1-\chi)L}{L-1}|||g|||_-.$$

These inequalities, together with the analogous ones for $\beta_i$, show that there exists a constant $C$, depending only on $b$, $c$, $d$, $L$, and $\chi$, such that

$$\Theta(g,\, 1) \leq \ln\frac{\beta}{\alpha} \leq \ln C\frac{|||g|||_+}{|||g|||_-} \leq \ln CL.$$

**Theorem 5.18.** *For each $f,\, g \in C^{(1)}(\mathcal{M})$, there exist $D,\, r > 0$ and $\Lambda < 1$ such that*

$$\left|\int_{\mathcal{M}} (f\circ T^n)g - \int_{\mathcal{M}} f\int_{\mathcal{M}} g\right| \leq D\Lambda^n \|f\|_*\|g\|_*,$$

*where, given $h \in C^{(1)}$, $\|h\|_* = \|h\|_1 + r\|h'\|_\infty$.*

*Proof.* We start with the same coarse partition $\mathcal{P} = \{P_i\}$ used in the smooth case (notice that, in the present context $P_i$ can be chosen to be polygons). It would be possible to perform the same dynamical argument, only the presence of the singularities would make it much more complicated. Yet, since in the present case it has been necessary to introduce a very small scale $\delta$, nothing much is lost by placing some further restrictions on it. We further refine the partitions until it consists only of convex polygons.

We can assume, without loss of generality, that each element of the new partition has a bottom and a top corner, and that the line joining such corners is completely contained in the element and it is a stable line. Consider the partition obtained by dividing each element of such partition into two by the above mentioned line.

Each element of the final partition can be foliated into stable curves in many ways. An easy way to construct a measurable partition is the following. Let $\varphi(t)$, $t \in [-1, 1]$, be the equation of the middle line, and $\zeta(t)$ be the equation of the rest of the boundary; $s\varphi((1-\frac{s}{2})t) + (1-s)\zeta(t)$ for $(s, t) \in [0, 1] \times [-1, 1]$ is a regular partition of the element into weakly–stable curves. By construction, all the weakly–stable curves so generated are longer than some $\delta_* \leq \delta_0$; in addition, it is possible to choose the cone large enough to contain the Jacobian of the corresponding change of variables. Using such a partition the analogue of Lemma 4.9 follows immediately, and the proof of the theorem is concluded by the same argument used in Theorem 4.8

## §6 SINAI BILLIARD.

In the previous sections we have seen how to deal with discontinuities, provided the map behaves nicely near them. Unfortunately, some important examples do not satisfy such conditions. In order to have a more satisfactory theory it would be necessary to allow the first and second derivative of the map to diverge near a line of discontinuity.

A little thought about what we have done in the latter section, in particular concerning the constant use of distortion estimates and the use of an upper bound on the expansion coefficient, shows that a generalization of the present strategy to a larger class of examples is not a trivial matter. It would certainly be necessary to impose some conditions on the singularities of the map but it is unclear which ones. Nonetheless, some extensions are possible. To show how, I will briefly discuss the chief example not covered by the previous conditions: the Sinai Billiard.

The Sinai billiard consists of the motion of a point particle in a two dimensional torus, where the motion is linear, apart from elastic reflection by a finite number of disjoint convex scatterers. To fix the ideas we will assume the scatterers to be circles of radius one. Here we will consider the simplest case: finite horizon, i.e., the scatterers are disposed in such a way that the particle cannot travel more than a fixed distance without colliding. In addition, to further simplify matters, we require that there are no trajectories with more than two consecutive tangent collisions.

The standard way to treat this system is to introduce a Poincaré map [Si], namely the map from immediately after a collision to immediately after the next. If we take as coordinates the distance $s$ of the collision point from a fixed point on the boundary of the scatterer, together with the angle $\varphi$ between the direction of the particle and the tangent to the boundary of the scatterer, the phase space turns out to be the union of squares (or, better yet, cylinders). More precisely each scatterer $S_i$ will have associated with it the cylinder $Q_i = S^1 \times [0, \pi]$. The flow of the particles induces a map $T : Q = \cup_i Q_i \to Q$ which preserves the area element $\sin\varphi ds \wedge d\varphi$; we also introduce the natural metric $\sin\varphi ds^2 + d\varphi^2$.

Due to the existence of both a minimal and a maximal length between consecutive collisions, it is possible to choose a strictly invariant cone family smaller than the usual one (the convergent family). More precisely, we can choose $\mathcal{C}_-(s, \varphi) = \{(ds, d\varphi) \in \mathcal{T}_{(s,\varphi)}Q_i \mid 1 \leq -\frac{d\varphi}{ds} \leq u_0\}$ for some appropriated $u_0$, depending upon the billiard.

The detailed derivation of the properties of the map $T$ would take considerable space; in addition such properties have been discussed at length in many articles. We will therefore refer to [BSC] where the decay of correlations for such (and more general) systems is discussed at length by using a technique akin to Markov partitions. In the aforementioned article, it is shown that the decay of correlations is, at least, sub-exponential.

Setting $(s_1, \varphi_1) = T(s_0, \varphi_0)$, one can write

$$D_{(s_0,\varphi_0)}T = \begin{pmatrix} \frac{\sin\varphi_0}{\sin\varphi_1} + \frac{\tau}{\sin\varphi_1} & \frac{\tau}{\sin\varphi_1} \\ \frac{\sin\varphi_0}{\sin\varphi_1} + \frac{\tau}{\sin\varphi_1} + 1 & \frac{\tau}{\sin\varphi_1} + 1 \end{pmatrix} \tag{6.1}$$

where $\tau$ is the length, in the usual Riemannian metric, of the trajectory between the two collisions.

We will measure the length of a stable curve[25] by the semi-metric $\sin\varphi ds$, since it is uniformly increased, and (for $\varphi \neq 0$ and $\varphi \neq \pi$) is equivalent to the true metric, when restricted to a stable curve. If $v = (1, -u)$ is a tangent vector ($v \in \mathcal{C}_-$ iff $1 \leq u \leq u_0$) at $x = (s, \varphi) \in Q$, then it is easy to compute the expansion $\lambda_x(v)$ of $v$ under $D_xT^{-1}$: $\lambda_x(v) = 1 + \frac{\tau}{\sin\varphi}(1+u)$. Finally, we will assume that the obstacles are sufficiently apart, so that the minimal expansion rate $\lambda$, for vectors in the cone, is larger than 100.[26]

We adopt the same cone definitions as in §5, with the only remark that, in the definition of $\Gamma_\delta$, the curvature $\kappa_\gamma$ is replaced by $\left|\frac{d^2\varphi}{ds^2}\right|$, and the integral on a curve $\gamma$ is now taken with respect to the measure $dm_\gamma = \sin\varphi ds$.

As already mentioned, the problems here arise due to the unboundedness of the first derivative of the map at discontinuous points (i.e., tangent collisions). The key idea to tackle such a problem is found in [BSC] and consists of introducing "homogeneity strips" around singularity lines.

In the present context the singularity lines are images or preimages of $(s, 0)$ or $(s, \pi)$ (i.e., tangency lines) for the various $Q_i$.

[25] Note that, in this section, we do not have the distinction between stable and weakly stable curves (or between $\widehat{\Gamma}_\delta$ and $\Gamma_\delta$). In fact, thanks to the very explicit representation of the phase space and the cone, we can both use a sufficiently small cone and extend stable curves without difficulties.

[26] Again, 100 is a completely arbitrary number. It is chosen just in order to simplify the proof. Using some more care, any value larger than one should be admissible.

**Definition 6.1.** *By "homogeneity strip" we mean a strip between two lines belonging to the set* $\{(s, Dn^{-\rho})\}_{n\in\mathbb{N}}$ *or the set* $\{(s, \pi - Dn^{-\rho})\}_{n\in\mathbb{N}}$ *for each square* $Q_i$ *and some* $\rho > 1$*; in addition, we will call the images,* $\{\mathcal{S}^h_n\}$*, of the previous set of lines under* $T$*, "h-discontinuity lines."*

Clearly, $\mathcal{S}^-$ (the discontinuity set of $T^{-1}$) consists of the images of $\{\varphi = 0\}$ and $\{\varphi = \pi\}$ under $T$ and is therefore the accumulation set of the $\{\mathcal{S}^h_n\}$. The importance of the above definition is illustrated by the following results:

**Lemma 6.2.** *If* $\gamma \in \Gamma_\delta$*,* $f \in \mathcal{D}^\nu_a(\gamma)$ *and* $\widetilde{\gamma} = T^{-1}\big(\gamma \cap (\textit{homogeneity strip})\big) \neq \emptyset$*, then* $\widehat{T}f \in \mathcal{D}^\nu_{\sigma a}(\widetilde{\gamma})$*, provided* $\nu < (1+\rho)^{-1}$*.*[27]

*Proof.* The fact that all the connected pieces $\widetilde{\gamma} \subset T^{-1}\gamma$, of length $\ell$, belong to $\Gamma_\ell$ is proved exactly as Sub-Lemma 4.3, by using formula (6.1). The only fact to check explicitly is the condition on $\varphi''$.

Let us use the representation $(t, \varphi_1(t)))$ for $\gamma$ and $(s_0(t), \varphi_0(t)) = T^{-1}(t, \varphi_1(t))$ for $\widetilde{\gamma}$. A direct computation yields

$$\begin{aligned}\frac{d\varphi_0}{ds_0}(t) &= -\left[1 + \frac{(1-\varphi_1'(t))\sin\varphi_0(t)}{\sin\varphi_1(t) + \tau(1-\varphi_1'(t))}\right]\\ \frac{ds_0}{dt}(t) &= \frac{\sin\varphi_1(t)}{\sin\varphi_0(t)} + \frac{\tau(1-\varphi_1'(t))}{\sin\varphi_0(t)},\end{aligned}$$

accordingly

$$\left|\frac{d^2\varphi_0}{ds_0^2}\right| \leq |\varphi_1''|\frac{\sin\varphi_1}{[\tau(1-\varphi_1') + \sin\varphi_1]^3} + c_3\left|\frac{d\tau}{dt}\right|\sin\varphi_0 + c_4.$$

(Through this lemma $c_i$ will designate constants that depend only on the system and the parameters of the cone.) For each $t_1, t_2 \in \mathbb{R}$ it is clear that

$$|\tau(t_1) - \tau(t_2)| \leq |t_1 - t_2| + |s_0(t_1) - s_0(t_2)|$$

so

$$\left|\frac{d\tau}{dt}\right| \leq 1 + \left|\frac{ds_0}{dt}\right|.$$

Remembering that we have assumed the existence of $\tau_-$ and $\tau_+$, such that $0 < \tau_- \leq \tau \leq \tau_+ < \infty$, follows

$$\left|\frac{d^2\varphi_0}{ds_0^2}\right| \leq \frac{1}{27\tau_-^2}|\varphi_1''| + c_5,$$

which suffice, provided $\tau_- > \frac{1}{\sqrt{27}}$.

In the present case the operator $\widehat{T}$ reads

$$\widehat{T}f(x) = f(Tx)\lambda_{Tx}(\gamma')^{-1}.$$

[27]Here, we call $\mathcal{D}^\nu_a$ the cone that, in the preceding sections, was called $D_a$. The reason being that now $\nu$ plays a much more important role.

Next, let $x,\, y \in \widetilde{\gamma}$ contained in the $n$–th homogeneity strip (i.e., the strip between the $(n-1)$–th and the $n$–th homogeneity line), and let $\lambda$ be the minimal expansion coefficient under $DT^{-1}$. Then, by choosing $D$ sufficiently small, follows

$$\frac{\widehat{T}f(x)}{\widehat{T}f(y)} \leq e^{a\lambda^{-1}n^{-\rho\nu}d(x,\,y)^{\nu}}\frac{\lambda_{Ty}(\gamma')}{\lambda_{Tx}(\gamma')}.$$

Using again the representation $(t,\, \varphi_1(t))$ for $\gamma$, and setting $(t_1,\, \varphi_1(t_1)) = Tx$ and $(t_2,\, \varphi_1(t_2)) = Ty$, we can estimate

$$\frac{\lambda_{Ty}(\gamma')}{\lambda_{Tx}(\gamma')} \leq e^{\left|\int_{t_1}^{t_2}\lambda_{\gamma(\xi)}(\gamma'(\xi))^{-1}\frac{d}{d\xi}\lambda_{\gamma(\xi)}(\gamma'(\xi))d\xi\right|}.$$

A direct computation (see also [BSC] and remember the bound on $\varphi''$ contained in the definition of $\Gamma_\delta$) shows that

$$\left|\frac{d}{d\xi}\lambda_{\gamma(\xi)}(\gamma'(\xi))\right| \leq c_6\left(\frac{1}{\varphi_1(\xi)^2}+\frac{1}{\varphi_1(\xi)\sin\varphi_0(\xi)}\right).$$

Using the above estimate yields

$$\frac{\lambda_{Ty}(\gamma')}{\lambda_{Tx}(\gamma')} \leq e^{c_7\int_{t_1}^{t_2}\left[\frac{1}{\varphi(\xi)}+\frac{1}{\sin\varphi_0(\xi)}\right]d\xi} \leq e^{c_8\left\{n^{\rho}\left[\int_{s_0(t_1)}^{s_0(t_2)}dm_{\widetilde{\gamma}}\right]^{\nu}|t_1-t_2|^{1-\nu}+\int_{s_0(t_1)}^{s_0(t_2)}\frac{1}{\sin\varphi_0}dm_{\widetilde{\gamma}}\right\}}.$$

Since $\gamma'(\xi) \in \mathcal{C}_-(\gamma(\xi))$, by definition, the intersection between $\gamma$ and the $n$–th homogeneity strip must be very short, more precisely $|t_1-t_2| \leq \text{const} \times n^{-\rho-1}$. In addition,

$$\int_{s_0(t_1)}^{s_0(t_2)}\frac{1}{\sin\varphi_0}dm_{\widetilde{\gamma}} = |s_0(t_2)-s_0(t_1)|,$$

but

$$d(x,\,y) \geq c_9\int_0^{|s_0(t_2)-s_0(t_1)|}\xi d\xi = \frac{c_9}{2}|s_0(t_2)-s_0(t_1)|^2.$$

Consequently

$$\frac{\widehat{T}f(x)}{\widehat{T}f(y)} \leq e^{[\lambda^{-1}n^{-\rho\nu}+a^{-1}c_{10}n^{\rho-(\rho+1)(1-\nu)}+c_{10}a^{-1}d(x,\,y)^{\frac{1}{2}-\nu}]ad(x,\,y)^{\nu}}.$$

The above Lemma is the counterpart of Lemmas 4.4 and 4.5. In addition, it shows that it is possible to control the distortion, although on shorter and shorter curves. The next Lemma takes advantage of the fact that the expansion coefficient is unbounded near the singularities to show that the union of the too short pieces is negligible. Although we will not use it directly in the following, it illustrates well the key property on which all the further arguments rest.

**Lemma 6.3.** *If $\gamma \in \Gamma_\delta$, then it is possible to write $T^{-1}\gamma = \bigcup_{\widetilde{\gamma}\in\mathcal{J}_1\cup\mathcal{J}_2} \widetilde{\gamma}$, where $\widetilde{\gamma} \in \mathcal{J}_1$ are connected, longer than $\delta$ and $T\widetilde{\gamma}$ belongs to only one homogeneity strip. In addition, for each $\xi \in [\zeta, 1]$,*

$$\sum_{\widetilde{\gamma}\in\mathcal{J}_2} |T\widetilde{\gamma}|^\xi \le c_1\delta^{\frac{3}{2}\xi}.$$

*Proof.* As we have seen above, a stable curve is expanded by $\lambda n^\rho$, if it falls in the $n$-th homogeneity strip. A maximal piece of $T^{-1}\gamma$ that crosses completely the $n$-th homogeneity strip has a length of at least $n^{-\rho-1}\delta_0$. Hence, if we choose $n_0 = \left(\frac{\delta_0}{\delta}\right)^{\frac{1}{\rho+1}}$ the pieces in $\mathcal{J}_2$ must belong to homogeneity strips with $n$ greater than some $n_1 > n_0$. Accordingly,

$$\sum_{\widetilde{\gamma}\in\mathcal{J}_2} |T\widetilde{\gamma}|^\xi \le \sum_{n>n_1} c_0^{-\xi} n^{-\rho\xi}|\widetilde{\gamma}|^\xi \le \sum_{n>n_0} c_0^{-\xi} n^{-\rho\xi}\delta^\xi.$$

The result then follows by choosing $\rho = \frac{\zeta+2}{\zeta} \ge 3$.

The ideas in Lemmas 6.2 and 6.3 are all that is needed to prove the equivalent of Lemma 5.4.

**Lemma 6.4.** *If $L > 1$, $4L < d$, $2^\nu a\delta^\nu < \ln 2$, and $\rho > 2$, then for each $n \in [A\ln d, \ldots, N(\delta)]$*

$$\widetilde{T}^n\mathcal{C}_{b,c,d,L}(\delta) \subset \mathcal{C}_{\chi b,\chi c,\chi d,\tau_n^2 L}(\delta)$$

*where $\chi < 1$, and $\tau_n = (1 - 3d(4\lambda^{-1})^n) \le \frac{1}{\sqrt{2}}$.*

*Proof.* The proof follows faithfully the path layed down in Lemma 5.4. Except that more care is required in keeping track of the pieces generated iterating a curve, since now there can be countably many.

In fact, given $\gamma \in \Gamma_\delta$, we will divide $\gamma$ in connected pieces belonging to the same homogeneity strip, with the possible exception of the pieces containing the endpoints or intersecting $\mathcal{S}^-$ (in this last case it is certainly possible to readjust a homogeneity line, so that the corresponding breaking point on $\gamma$ coincide with the one determined by $\mathcal{S}^-$, without spoiling any of the above estimates). This means that $T^{-1}\gamma$ can consist, at most, of a countable collections of curves if $\gamma$ intersect an homogeneity region, and consists at most of three curves otherwise (since we have assumed no more than two consecutive tangencies). Since curves that expand very much are very short, it is easy to see that for at least $N(\delta) \sim \ln\delta^{-1}$ iterations no curves become longer than some fixed $\delta_0$. So, if we put into $\mathcal{J}_-$ all the curves shorter than $\delta$ belonging to $T^{-n}\gamma$ and we call $\lambda(\widetilde{\gamma})$ the minimal expansion coefficient of $T^{-n}$ on $T^n\widetilde{\gamma}$ we have

$$\begin{aligned}
\sum_{\widetilde{\gamma}\in\mathcal{J}_-} \lambda(\widetilde{\gamma})^{-1} &\le \sum_{k_1=0}^{\infty}\cdots\sum_{k_n=0}^{\infty} \lambda^{-n}\prod_{j=1}^{n}\left((k_j+1)^{-\rho}+2\right) \\
&= \lambda^{-n}\left(2+\sum_{k=0}^{\infty}(k+1)^{-\rho}\right)^n \\
&\le \left(\lambda^{-1}(3+\frac{1}{2^{\rho-1}(\rho-1)})\right)^n \le (4\lambda^{-1})^n
\end{aligned}$$

which, remembering our assumptions on $\lambda$, is less than one (here we have used our assumption on $\rho$).

Since the above quantity is the only one used in Lemma 5.4 to estimate the influence of short pieces (there it is estimated by $\lambda^{-n}M^n$) the same results follows in the same way.

This implies that it is possible to control a large number of iterates of $\widetilde{T}$: the analogue of Sub-Lemma 5.9.

In Theorem A3.1 of [BSC] it is shown that a curve in $\Gamma_\delta$ will consist overwhelmingly of pieces longer than $\delta_0$ after a number of iterations proportional to $\ln \delta^{-1}$. Hence, we can obtain the analogue of Lemma 5.8.

To obtain the wanted result we need only to be able to compare averages on curves in $\Gamma_{\delta_0}$. This can be done exactly in the same way as in §5, Lemma 5.14. In fact, the arguments used there are completely general and depend only on the results contained the Appendices. But, Appendix II can be generalized to include the present case, as it is explained at the end of the Appendix itself. In addition, Appendix I can be not only generalized to the present context, but it is also possible to ask that a weak-core contains a positive measure of long *homogenous* stable and unstable manifolds (i.e., manifolds whose images fall always in a unique homogeneity strip) instead that simply stable and unstable manifolds (see [BSC, Appendix 1] for details). Such a generalization suffices to prove the corresponding of Sub-Lemma 5.16, and consequently establish the like of Lemma 5.14.[28]

This concludes the discussion of the proof of the equivalent of Theorem 5.3. Since all the rest is equal, the exponential decay of correlations follows.

## §7 TWO DIMENSIONS (RANDOM PERTURBATIONS).

In this section we will study the random perturbations of a two dimensional, area–preserving smooth system. Although the class of perturbations treated is much less general than the one considered in [Ki], the present technique provides an easier approach and new results concerning the dependence of the invariant measure from the size of the perturbation. The reader will easily see the possibility of generalizing the following reasoning.

For simplicity we will assume $\mathcal{M} = \mathbb{T}^2$.

We will assume that the deterministic map $T$ satisfies the hypotheses of §4, moreover, calling $v^u(x)$, $v^s(x)$ the unstable and stable direction at $x$, we assume

$$
\begin{aligned}
|v^u(x) - v^u(y)| &\leq A|x-y|^\tau \\
|v^s(x) - v^s(y)| &\leq A|x-y|^\tau
\end{aligned}
$$

[28] In Lemma 5.14 the maximal expansion coefficient $\mu$ appears only in two places. First, in Lemma 5.12 where we use Lemma II.2. The analogous result for billiards (stated at the end of Appendix II) does not involves the maximal expansion (that is now $\infty$) but forces us to choose $\zeta$ sufficiently close to 1 as well. Second, in the final estimate of the ratio of the averages. Here $\mu^N$ is substituted by $\sup_{\gamma\in\Gamma_{\delta_1}} \frac{|\gamma|}{|T^N\gamma|} < \infty$.

for some $\tau \in (0, 1]$. In fact, the existence of such $\tau$ is quite general, since, for the case under consideration, the stable and unstable foliations are, at least, Hölder continuous.

The random perturbation, in analogy with the one dimensional case, will be defined by the following Markov process.

Let $\phi : \mathbb{T}^2 \times \mathbb{R}^+ \to \mathbb{R}^+$ be a smooth function such that $\phi(x, \xi) = 0$ for each $\xi \geq 1$, $x \in \mathbb{T}^2$, and $2\pi \int_0^1 \phi(x, \xi)\xi d\xi = 1$ for each $x \in \mathbb{T}^2$. Then, for each $f \in L^1(\mathbb{T}^2)$ we define

$$\left(\widehat{T}_\varepsilon f\right)(x) = \frac{1}{\pi\varepsilon^2} \int_{\mathbb{T}^2} \phi(Tx, \varepsilon^{-1}|Tx - y|f(y)dy.$$

In addition, we assume a weak form of translation invariance:

$$\frac{\phi(x, \xi)}{\phi(y, \xi)} \leq e^{\theta|x-y|^\tau}$$

for each $\xi \in \mathbb{R}^+$, $x, y \in \mathbb{T}^2$.

Two types of problems are of interest:

i) ergodic properties of the random process;
ii) distance between the invariant measure of the random and the deterministic processes.

Our claim is that both problems can be addressed with the technique described in this paper. The approach to the first question is reasonably straightforward: one chooses a cone of functions and proves contraction. As we have seen in §3 various choices of cones are permissible. One can simply choose the cone of positive function, since it is possible to shows that, after approximately $\ln \varepsilon^{-1}$ iterations, the kernel of the transfer operator is strictly positive (better, there exists constants $B_-$, $B_+$, independent of $\varepsilon$, that lower and upper bound it), it follows that the correlation between any two continuous functions $f$, $g$ decays as $\Lambda^{\frac{n}{\ln \varepsilon^{-1}}} \|f\|_\infty \|g\|_\infty$.[29] This result is both stronger (the rate of decay it is not proportional to the derivatives of the functions) and weaker (the rate of decay depends upon $\varepsilon$) then the one obtained for the deterministic case. Yet, it is likely that a cone similar to the one used in §4 can yield an $\varepsilon$-independent rate of decay for $C^{(1)}$ functions.

In the present paper we will not discuss point (i) any further; we will instead concentrate on point (ii). Here it is less obvious how to obtain results similar to the one discussed in §3. Nevertheless, something can be done.

The Perron-Frobenius operator reads

$$\widetilde{T}_\varepsilon g(x) = \frac{1}{\pi\varepsilon^2} \int_{\mathbb{T}^2} \phi(y, \varepsilon^{-1}|Ty - x|)g(y)dy.$$

Let

$$\|g\|_+ = \sup_{x\in\mathbb{T}^2} g(x)$$
$$\|g\|_- = \inf_{x\in\mathbb{T}^2} g(x).$$

[29] See [Ki] for similar results.

We designate by $W^s(x,\,t)$ the point obtained by translating $x$ by $t$ along its stable manifold, we then define the cone

$$\mathcal{C}^\tau_{a,\,H}=\left\{f\in C^{(1)}\mid f\geq 0;\ \frac{f(x)}{f(W^s(x,\,t))}\leq e^{at^\tau}\ \text{for}\ t\leq\varepsilon;\ \frac{\|f\|_+}{\|f\|_-}\leq H\right\}.$$

**Lemma 7.1.** *If, for some given constant $c_1$, $\theta+c_1H+\lambda^{-1}a<a$, then there exists $\sigma<1$ such that $\widehat{T}_\varepsilon\mathcal{C}^\tau_{a,\,H}\subset\mathcal{C}^\tau_{\sigma a,\,H}$.*

*Proof.* The only condition that is not obvious is the second one.

$$\begin{aligned}\widehat{T}_\varepsilon f(x)=&\frac{1}{\pi\varepsilon^2}\int_{\mathbb{T}^2}\phi(Tx,\,\varepsilon^{-1}|Tx-y|)f(y)dy\\ \leq&\frac{e^{\theta t_1^\tau}}{\pi\varepsilon^2}\int_{\mathbb{T}^2}\phi(TW^s(x,\,t),\,\varepsilon^{-1}|Tx-y|)f(y)dy\end{aligned}$$

where $t_1\leq\lambda^{-1}t$ is such that $W^s(Tx,\,t_1)=TW^s(x,\,t)$. Since the direction of the stable manifold at $y$ differs from the one at $x$ by, at most, a quantity proportional to $\varepsilon^\tau$, it follows, setting $Tx-y=\xi$,

$$W^s(Tx,\,t_1)+\xi=W^s(y,\,t_1)+\mathcal{O}(\varepsilon^\tau t)$$

that is

$$|\xi|=|TW^s(x,\,t)-W^s(y,\,t_1)|+\mathcal{O}(\varepsilon^\tau t).$$

Hence,

$$\begin{aligned}\widehat{T}_\varepsilon f(x)\leq&\frac{e^{\theta t^\tau}}{\pi\varepsilon^2}\int_{\mathbb{T}^2}\phi\left(TW^s(x,\,t),\,\frac{1}{\varepsilon}|TW^s(x,\,t)-W^s(y,\,t_1)|\right)f(y)+\mathcal{O}(\varepsilon^{\tau-1}t)\|f\|_+\\ \leq&\frac{e^{\theta t^\tau}}{\pi\varepsilon^2}\int_{\mathbb{T}^2}\phi\left(TW^s(x,\,t),\,\varepsilon^{-1}|TW^s(x,\,t)-W^s(y,\,t_1)|\right)f(W^s(y,\,t_1))dy\\ &+\mathcal{O}(\varepsilon^{\tau-1}t)H\|f\|_-.\end{aligned}$$

To conclude we need to perform the change of variable $z=W^s(y,\,t_1)$ and compute its Jacobian.

Let us choose Cartesian coordinates, in a $\varepsilon$-neighborhood of $y$, with axes parallel to $v^s(y)$ and $v^u(y)$, respectively, and write $y=(y_s,\,y_u)$. Then, $\frac{\partial z_s}{\partial y_s}=1$; $\frac{\partial z_s}{\partial y_u}=\mathcal{O}(t^\tau)$ by the conditions imposed on the regularity of the foliations; $\frac{\partial z_u}{\partial y_s}=\mathcal{O}(t^\tau)$; $\frac{\partial z_u}{\partial y_u}=1+\mathcal{O}(t^\tau)$.

In conclusion, the Jacobian of the change of variable is equal to $1+\mathcal{O}(t^\tau)$; hence

$$\widehat{T}_\varepsilon f(x)\leq\left[e^{(\theta+\lambda^{-1}a)t^\tau}+c_1t^\tau H\right]\widetilde{T}_\varepsilon f(W^s(x,\,t))$$

which proves the Lemma.

We will designate by $\widetilde{T}_0$ and $\widehat{T}_0$ the operators corresponding to the deterministic map. The following provides an interesting bound of the difference between the correlations of the random and the deterministic processes.

**Lemma 7.2.** *For each $f,\, g \in C^{(1)}$ and $\varepsilon$ sufficiently small, holds*

$$\left|\int_{\mathbb{T}^2} \widehat{T}_\varepsilon^n fg - \int_{\mathbb{T}^2} \widehat{T}_0^n fg\right| \le nc_3\varepsilon^\tau \|f\|_*\|g\|_*.$$

*where $\|\cdot\|_*$ is defined in Theorem 4.8.*

*Proof.* Choose $f \in \mathcal{C}^\tau_{a,\,H}$ and $g \in \mathcal{C}_{b,c,d,L}$, then

$$\begin{aligned}\left|\int_{\mathbb{T}^2}(\widehat{T}_\epsilon^n f)g - \int_{\mathbb{T}^2}(\widehat{T}_0^n f)g\right| &\le \sum_{k=0}^{n-1}\left|\int_{\mathbb{T}^2}\widehat{T}_0^{n-k-1}\widehat{T}_0\widehat{T}_\epsilon^k fg - \int_{\mathbb{T}^2}\widehat{T}_0^{n-k-1}\widehat{T}_\epsilon\widehat{T}_\epsilon^k fg\right| \\ &\le \sum_{k=0}^{n-1}\left|\int_{\mathbb{T}^2}\widehat{T}_0\widehat{T}_\epsilon^k f\widetilde{T}_0^{-n+k+1}g - \int_{\mathbb{T}^2}\widehat{T}_\epsilon\widehat{T}_\epsilon^k f\widetilde{T}_0^{-n+k+1}g\right|.\end{aligned}$$

Since $f_1 = \widehat{T}_\epsilon^k f \in \mathcal{C}^\tau_{a,\,H}$ (Lemma 7.2) and $g_1 = \widetilde{T}_0^{-n+k+1}g \in \mathcal{C}_{b,c,d,L}$ (Lemma 4.2) we can write the generic term of the previous expression as

$$\begin{aligned}&\left|\frac{1}{\pi\varepsilon^2}\int_{(\mathbb{T}^2)^2} f_1(x)\phi(x,\,\varepsilon^{-1}|x-Ty|)g_1(y) - \int_{\mathbb{T}^2} T_0 f_1 g_1\right| \\ &\qquad = \left|\frac{1}{\pi\varepsilon^2}\int_{(\mathbb{T}^2)^2} f_1(x)\phi(x,\,\varepsilon^{-1}|x-y|)g_1(T^{-1}y) - \int_{\mathbb{T}^2} f_1 g_1\circ T^{-1}\right|.\end{aligned}$$

This means that the problem is reduced to the estimate of

$$\left|\frac{1}{\pi\varepsilon^2}\int_{(\mathbb{T}^2)^2} f(x)\phi(x,\,\varepsilon^{-1}|x-y|)g(y) - \int_{\mathbb{T}^2} fg\right|$$

where $f,\, g$ are generic functions in the respective cones. We introduce the following change of variable:

$$\begin{aligned}\xi &= x-y \\ \eta &= \alpha(x,\,y)\end{aligned} \tag{7.1}$$

where $\alpha(x,\,y) = W^s(x)\cap W^u(y)$.

**Sub-Lemma 7.3.** *Calling $\widehat{J}$ the Jacobian of the change of variable (7.1), for each $|\xi| \le \varepsilon$, and $\varepsilon$ small enough,*

$$|\widehat{J}-1| \le c_1\varepsilon^\tau.$$

*Proof.*

Let us fix a point $x \in \mathbb{T}^2$. Since the Jacobian can be computed in any system of coordinates, we choose Cartesian coordinates with axes parallel to the stable and unstable direction at $x$ respectively. We can then write $x = (x_s,\, x_u)$, and the same for $y,\, \xi,\, \eta$. We will then write the Jacobian matrix $\widetilde{J}$ in this coordinates and compute its determinant.

Clearly, $\frac{\partial \xi_l}{\partial x_l} = 1$ for $l \in \{s, u\}$ while $\frac{\partial \xi_l}{\partial y_l} = -1$. By the definition of $\alpha$ follows $\frac{\partial \eta}{\partial x_s} = 0$. The computation of $\frac{\partial \eta_u}{\partial x_u}$ is more delicate. The derivate, with respect to $x_u$, can be computed by taking the derivative along the unstable manifold of $x$. By the absolute continuity of the unstable foliation it follows that $\alpha(W^u(x, h), y) = W^u(\eta, h')$ with $e^{-c_1\varepsilon} \leq \frac{|h'|}{|h|} \leq e^{c_1\varepsilon}$. Consequently, since the unstable direction at $\eta$ differs from the unstable direction at $x$ by, at most, $\varepsilon^\tau$ we have

$$
\begin{aligned}
\frac{\partial \eta_u}{\partial x_u} =& 1 + \mathcal{O}(\varepsilon^\tau) \\
\frac{\partial \eta_s}{\partial x_u} =& \mathcal{O}(\varepsilon^\tau).
\end{aligned}
$$

By similar arguments one can compute the remaining elements of $\widetilde{J}$, obtaining

$$
\widetilde{J} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + \mathcal{O}(\varepsilon^\tau).
$$

Using the properties of $f$ and $g$ together with the above bound easily yields

$$
\begin{aligned}
&\frac{1}{\pi\varepsilon^2}\int_{(\mathbb{T}^2)^2} f(x)\phi(x, \varepsilon^{-1}|x-y|)g(y) \\
&\qquad =\frac{1}{\pi\varepsilon^2}\int_{|\xi|\leq\varepsilon} d\xi \int_{\mathbb{T}^2} d\eta f(\eta)\phi(\eta, \varepsilon^{-1}|\xi|)g(\eta) + \mathcal{O}(\varepsilon^\tau)\|f\|_\infty\|g\|_\infty \\
&\qquad =\int_{\mathbb{T}^2} fg + \mathcal{O}(\varepsilon^\tau)\|f\|_\infty\|g\|_\infty.
\end{aligned}
$$

The result for general $f$ and $g$ follows then via the same arguments used in Theorem 4.8.

Using the previous Lemma we can easily prove the announced estimate.

**Theorem 7.4.** *Calling $\mu_\varepsilon$ the invariant measure of the random process $\widetilde{T}_\varepsilon$, for each $f \in C^{(1)}$*

$$
\left|\int f d\mu_\varepsilon - \int f\right| \leq A(\ln\varepsilon^{-1})^2\varepsilon^\tau\|f\|_*
$$

*Proof.*

$$
\begin{aligned}
\left|\int f d\mu_\varepsilon - \int f\right| &\leq \left|\int f d\mu_\varepsilon - \int f\widetilde{T}_\varepsilon^n 1\right| + \left|\int f\widetilde{T}_\varepsilon^n 1 - \int f\widetilde{T}_0^n 1\right| \\
&\leq c_2\left[n\varepsilon^\tau + \Lambda^{-\frac{n}{\ln\varepsilon^{-1}}}\right]\|f\|_*
\end{aligned}
$$

the Theorem follows by choosing $n = \frac{\tau}{\ln\Lambda}(\ln\varepsilon^{-1})^2$.

**Appendix I (Sinai Theorem).**

This appendix is dedicated to the proof of Theorem 6.13. Since the argument is reasonably standard, we will provide only a sketch, and refer to [LW] for more details.

We start by introducing a coarse partition of $Q$ via $\mathcal{S}_{m_Q} = \bigcup_{n=0}^{m_Q} (T^n\mathcal{S}^-)\cup(T^{-n}\mathcal{S}^+)$. Clearly $Q\backslash\mathcal{S}_{m_Q}$ is partitioned into finitely many open domains $\{R_i\}$. Since $\widetilde{\mathcal{C}}_-(z) = DT^{-m_Q}\mathcal{C}_-(T^{m_Q}z)$ and $\widetilde{\mathcal{C}}_+(z) = DT^{m_Q}\mathcal{C}_+(T^{-m_Q}z)$ are uniformly continuous in each $R_i$, it is possible to further refine the partitions and obtain a new partition $\{R_i^j\}$ made of elements of uniform size and on which the cone families are almost constant. Next, we choose a point $z_i^j$ in each $R_i^j$ such that $z_i^j \notin \Sigma_\infty$, and we introduce in each $R_j^i$ a cartesian system of coordinates with axes parallel to the stable and unstable directions at $z_j^i$. We can then introduce in $R_j^i$ a family of squares of size $t\delta_1$ with sides parallel to the axes, such that their weak-core covers $R_j^i$. We will consider only the squares whose weak-core do not intersect the boundary of $R_j^i$.

The set $\mathcal{S}_m = \bigcup_{|n|\leq m} T^n\mathcal{S}$ consists of a finite collection of lines. Let us call $\mathcal{R}_m$ the set of points that belong to more than one line. By hypothesis $\mathcal{R}_m$ is a discrete set containing a finite number of points. We choose $\delta_1$ smaller than $\frac{t}{4}$ the smallest distance among points in $\mathcal{R}_m$. If $\widetilde{\mathcal{R}}_m$ is the $2t\delta_1$ neighborhood of $\mathcal{R}_m$, then $\mathcal{S}_m\backslash\widetilde{\mathcal{R}}_m$ consists of a set of disjoint lines, at least $2t\delta_1$ apart. By choosing $\delta_1$ small enough, the total measure of the squares that intersect $\widetilde{\mathcal{R}}_m$ is less than $\frac{\delta}{512\delta_0}$. Moreover, choosing $m_Q$ large and $\alpha$ small enough, it is possible to insure that a line belonging to $\mathcal{S}_m\backslash\mathcal{S}_{m_Q}$ is not sufficient to make a square not $\alpha$-connecting. The result follows then by choosing $m$ sufficiently large and by the usual estimate of the size of the tail $\Sigma_\infty\backslash\Sigma_m$.

**Appendix II (Absolute continuity).**

This section provides some result that are not readily available in the literature.

The main result presented here concerns the map induced by the unstable foliation between two nearby stable curves. It is well known that such map is absolutely continuous, with Jacobian close to the identity, if the two curve are sufficiently nearby [K-S]; I will show that the map is Hölder continuous together with its Jacobian. This result is well known for smooth maps [P], [M] but, to my knowledge, it is new for discontinuous maps, although some partial results along this line exist in the case of billiards [B-S-C].

We will restrict ourselves to the two dimensional case, this will allow us to have easier and more transparent proofs with respect to [K-S]. Although an extension to the multidimensional case should not be out of the reach of the technique used here, the author feels that it would be out of the scopes of his present efforts.

We will consider the class of maps described at the beginning of §6.

Consider two strictly stable curves $\gamma_1$, $\gamma_2$, such that $d(\gamma_1, \gamma_2) \leq \delta_2/2$.

We define $D(\phi) = \{z \in \gamma_1 \mid W^u(z) \cap \gamma_2 \neq \emptyset\}$, were $W^u$ is the unstable manifold of the $z$ of size less or equal to $2\delta_2$,[30] and the map $\phi : D(\phi) \subset \gamma_1 \to \gamma_2$ by

$$\phi(x) = W^u(x) \cap \gamma_2.$$

If $R(\phi)$ is the range of the map (and the domain of $\phi^{-1}$), then we define a measure $\mu$ on $R(\phi)$ by

$$\int_{R(\phi)} f d\mu = \int_{D(\phi)} f \circ \phi dm.$$

¿From the general theory developed in [K-S] follows:

**Theorem II.1.** *The map $\phi$ is absolutely continuous. In addition, calling $J_\phi$ its Jacobian, for each density point $z$ of $R(\phi)$ holds*

$$\frac{d\mu}{dm} = J_\phi(z) = \prod_{j=0}^{\infty} \left| \frac{D_{T^{-j}z} T\big|_{T^{-j}\gamma_2}}{D_{T^{-j}\phi^{-1}(z)} T\big|_{T^{-j}\gamma_1}} \right| \leq e^{D\delta_2}.$$

In order to use effectively the map $\phi$ in our context, we need three extra informations: an estimate of the measure of $\gamma_1 \backslash D(\phi)$ and $\gamma_2 \backslash R(\phi)$; an estimate of the smoothness of $\phi$; and an estimate of the smoothness of $J_\phi$.

Let $\Sigma_n(c, \delta_2) = \cup_{i=0}^n T^i \mathcal{S}^-_{c\lambda^{-i}\delta_2}$. ¿From the construction of the unstable manifold (see [K-S]) follows that $\gamma_1 \backslash \Sigma_\infty(c, \delta_2) \supset D(\phi)$. The first estimates are then accomplished by the following.

**Lemma II.2.** *There exists $D > 0$ such that for each $\gamma \in \widehat{\Gamma}_{\delta_2}$, and $m \in \mathbb{N}$, if $\gamma \cap \Sigma_m(c, \delta_2) = \emptyset$, with $\delta \leq \delta_2 < \delta_0$, $c < 1$, then*

$$m\left(\gamma \backslash D(\phi)\right) \geq Dc\delta_2 \lambda^{-m}$$
$$\frac{1}{f(x)} \left| \int_{\gamma \backslash D(\phi)} gf \right| \leq Dc\delta_2 (\lambda^{-\zeta}\mu^{1-\zeta})^m |||g|||_-,$$

*where $m(\cdot)$ is the Lebesgue measure restricted to $\gamma$, and $g \in \mathcal{C}_{b,c,d,L}(\delta)$, $f \in \mathcal{D}_a(\gamma)$.*

*Proof.* We start by recalling a simple distortion estimate.

**Sub-Lemma II.3.** *If $I \in \widehat{\Gamma}_{\delta_0}$, $J \subset I$, and $T^n I$ is connected, then*

$$A^{-1}\frac{|I|}{|J|} \leq \frac{|T^n I|}{|T^n J|} \leq A\frac{|I|}{|J|}.$$

*Proof.*

$$\frac{|T^n I|}{|T^n J|} = \frac{\int_I |D_x T^n\big|_I|}{\int_J |D_y T^n\big|_I|} \leq e^{c_2\delta_0(1-\lambda^{-1})^{-1}} \frac{|I|}{|J|}.$$

[30] $W^u(z) = \{z\}$ if $z$ does not have an unstable manifold.

Consider $\gamma \in \widehat{\Gamma}_{\delta_*}$, $\delta_* \in [\delta_2, \delta_0]$, $c < 1$. Suppose $\gamma \cap \Sigma_k(c, \delta_2) = \emptyset$, for some $k \geq 0$. Then $T^{-k}\gamma = \gamma_0$ is connected, and $|\gamma_0| \geq \lambda^k \delta_2$.

Let $S_0 = \{J_i\}$ be the collection of the, at most, $M$ connected segments constituting $\gamma_0 \cap \Sigma_0(\lambda^{-k}c, \delta_2)$, and set $\widetilde{\gamma}_0 = \gamma_0 \setminus \cup_{J \in S_0} J$. Consider $T^{-1}\widetilde{\gamma}_0$, it will consist, at most, of $M$ connected pieces. Collect in $\mathcal{J}_1$ the pieces longer than $\delta_0$, after subdividing them, if necessary, in pieces shorter than $2\delta_0$. Collect the remaining pieces in $\mathcal{J}_1^*$. For each piece in $\mathcal{J}_1^*$ repeat the previous construction (i.e., intersect the piece with $\Sigma_0(\lambda^{-(k+1)}c, \delta_2)$, collect the connected pieces of the intersection in $S_1$, consider the preimage of the rest, collect the long pieces in $\mathcal{J}_2$ and the remaining ones in $\mathcal{J}_2^*$, and so forth).

By construction, the pieces in $\mathcal{J}_l^*$ can be at most $M^l$ and are all shorter than $\delta_0$. Hence, for each $\xi \in [\zeta, 1]$,

$$\sum_{\widetilde{\gamma} \in \mathcal{J}_l^*} |T^{-k-l}\widetilde{\gamma}|^\xi \leq M^l \lambda^{-(l+k)\xi} \delta_0^\xi.$$

In addition, we can estimate the pieces that fall too close to a singularity line during the first $l$ iterations:

$$\sum_{i=k}^{l-1} \sum_{\widetilde{\gamma} \in S_{i-k}} |T^{-i}\widetilde{\gamma}|^\xi \leq c^\xi \sum_{i=k}^{l-1} M^{i+1-k} \lambda^{-2i\xi} \delta_2^\xi \leq Bc^\xi \lambda^{-2k\xi} \delta_2^\xi,$$

since $M\lambda^{-\zeta} < 1$. Consequently, calling $\mathcal{J}(\gamma, c)$ the collection of segments that either get to close to the singularity or never grow larger than $\delta_0$, we have

$$\sum_{\widetilde{\gamma} \in \mathcal{J}(\gamma_0, c)} |\widetilde{\gamma}|^\xi \leq Bc^\xi \lambda^{-2k\xi} \delta_2^\xi.$$

We still need to deal with the segments in the collection $\cup_{i=k+1}^\infty T^i \mathcal{J}_{i-k}$.

Let us consider $\overline{\gamma} \in \mathcal{J}_1$, iterating it we will obtain again segments that are to close to singularity, segments that stay short for all times, and segments that grow to size $\delta_0$. We will estimate the first two group while we will add the segment belonging to the last category to the above mentioned collections of long segments.

$$\begin{aligned}
\sum_{\overline{\gamma} \in \mathcal{J}_1} \sum_{\widetilde{\gamma} \in \mathcal{J}(\overline{\gamma}, \lambda^{-k-1}c)} |T^{k+1}\widetilde{\gamma}|^\xi &\leq \sum_{\overline{\gamma} \in \mathcal{J}_1} \sum_{\widetilde{\gamma} \in \mathcal{J}(\overline{\gamma}, \lambda^{-k-1}c)} \frac{|T^{k+1}\widetilde{\gamma}|^\xi}{|T^{k+1}\overline{\gamma}|^\xi} |T^{k+1}\overline{\gamma}|^\xi \\
&\leq A^\xi \sum_{\overline{\gamma} \in \mathcal{J}_1} \sum_{\widetilde{\gamma} \in \mathcal{J}(\overline{\gamma}, \lambda^{-k-1}c)} \frac{|\widetilde{\gamma}|^\xi}{|\overline{\gamma}|^\xi} |T^{k+1}\overline{\gamma}|^\xi \\
&\leq A^\xi Bc^\xi \lambda^{-(k+1)\xi} \delta_2^\xi \delta_0^{-\xi} \sum_{\overline{\gamma} \in \mathcal{J}_1} |T^{k+1}\overline{\gamma}|^\xi \\
&\leq A^\xi Bc^\xi \lambda^{-(k+1)\xi} \delta_2^\xi \delta_0^{-1} \mu^{(k+1)(1-\xi)} \sum_{\overline{\gamma} \in \mathcal{J}_1} |T^{k+1}\overline{\gamma}| \\
&\leq 2A^\xi Bc^\xi \lambda^{-(k+1)\xi} \delta_2^\xi \mu^{(k+1)(1-\xi)}
\end{aligned}$$

Still, we have not analyzed all the pieces that must be estimated. The remaining ones are pieces $\widehat{\gamma}$ with some preimage $T^{-j}\widehat{\gamma}$, $j > k$, contained in a connected component of $T^{-j}\gamma$ of size larger than $\delta_0$. The key observation to conclude the argument is that the last estimate does not depend from the number of pieces in $\mathcal{J}_1$, but only from the fact that all such pieces are longer than $\delta_0$ and that $\mathcal{J}_1 \subset T^{-k-1}\gamma$. This means that we can apply the above estimate iteratively and obtain

$$\sum_{I\subset\Sigma_\infty(c,\,\delta_2)} |I|^\xi \le \sum_{i=k}^{\infty} 2A^\xi Bc^\xi(\lambda^{-\xi}\mu^{(1-\xi)})^i\delta_2^\xi \le D(\lambda^{-\xi}\mu^{(1-\xi)})^k\delta_2^\xi.$$

The Lemma easily follows from the above estimate, and from noticing that $\gamma\backslash D(\phi)$ is a countable collection of segments, each one of which is contained in a segment belonging to $\Sigma_\infty(c,\, \delta_1)$.

The first part of the wanted estimates is completed by noticing the symmetry between $\phi$ and $\phi^{-1}$.

Next, we can study the smoothness of $\phi$.

**Lemma II.4.** *There exists $A \in \mathbb{R}^+$, such that for each $x,\, y \in \gamma_1\backslash\Sigma_\infty(c,\, \delta_2)$*

$$d(\phi(x),\, \phi(y)) \le Ad(x,\, y).$$

*Proof.* Let $x,\, y \in \gamma\backslash\Sigma_\infty(c,\, \delta_2)$, $c \le 1$; suppose

$$d(T^{-n}x,\, T^{-n}y) \le d(x,\, y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}$$

then we have

$$\lambda^n d(x,\, y) \le d(x,\, y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}$$

that is

$$d(T^{-n}x,\, T^{-n}y) \le \lambda^{-n}\delta_2.$$

The above estimate means that the two points cannot be on opposite sides of $\mathcal{S}^-$ (remember that the distance is taken along $\gamma$ or its images), otherwise one of the points would belong to $\Sigma_\infty$ contrary to the hypothesis. Accordingly, another interaction is possible without disconnecting the part of $\gamma$ lying between $x$ and $y$. This shows that it is possible to iterate the two points until

$$d(T^{-n}x,\, T^{-n}y) \ge d(x,\, y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}.$$

¿From the definition follows that

$$T^{-n}\phi(z) = W^u(T^{-n}z) \cap T^{-n}\gamma_2.$$

**Sub-Lemma II.5.** *If $z,\, w \in \gamma_1 \backslash \Sigma_\infty(c,\, \delta_2)$ and $z' = \phi(z)$, $w' = \phi(w)$, then, given $m \in \mathbb{N}$ such that*

$$d(T^{-m}z,\, T^{-m}w) \geq d(z,\, w)^{\frac{1}{2}} \delta_2^{\frac{1}{2}}$$

*and $T^{-m}[z,\, w] \subset T^{-m}\gamma$ is connected, we have*

$$d(T^{-m}z,\, T^{-m}z') \leq \rho^2 d(z,\, w)^{\frac{1}{2}} \delta_2^{\frac{1}{2}}$$
$$d(T^{-m}w,\, T^{-m}w') \leq \rho^2 d(z,\, w)^{\frac{1}{2}} \delta_2^{\frac{1}{2}}$$

*for some $\rho \geq 1$.*

*Proof.* Let us consider the rectangle $Q = [z,\, w,\, w',\, z']$. Since the cone field is continuous the curves $\gamma_i$ are uniformly transversal to the unstable direction. This implies that there exists some fixed $\rho \in \mathbb{R}^+$, such that the area of the rectangle is less than $\rho d(z,\, w)\delta_2$. Therefore, since the area of $T^{-n}Q$ will be larger than $\rho^{-1} d(T^{-m}z,\, T^{-m}w) d(T^{-m}z,\, T^{-m}z')$ the result follows by the area preserving property of $T$.

In conclusion, given $x,\, y \in \gamma_1$, $d(x, y) < \delta_2^{\frac{1}{2}}$, we choose the first $n \in \mathbb{N}$ such that

$$d(T^{-n}x,\, T^{-n}y) \geq d(x,\, y)^{\frac{1}{2}} \delta_2^{\frac{1}{2}}.$$

Then, due to the continuity of the cone field, the fact that $T^{-n}$ is smooth on the rectangle determined by the points $x$, $y$, $\phi(x)$, $\phi(y)$, and uniform hyperbolicity, we have that $T^{-n}[x,\, y]$ and $[T^{-n}\phi(x),\, T^{-n}\phi(y)]$ are essentially parallel curves. Sub-Lemma II.5 implies then that

$$\begin{aligned} d(T^{-n}\phi(x),\, T^{-n}\phi(y)) &\leq d(T^{-n}x,\, T^{-n}y) + 2\rho d(T^{-n}\gamma_1,\, T^{-n}\gamma_2) \\ &\leq d(T^{-n}x,\, T^{-n}y) + 2\rho^3 d(x,\, y)^{\frac{1}{2}} \delta_2^{\frac{1}{2}} \end{aligned}$$

or

$$\frac{d(T^{-n}\phi(x),\, T^{-n}\phi(y))}{d(T^{-n}x,\, T^{-n}y)} \leq 1 + 2\rho^3$$

so, we have

$$B^{-1} \leq \frac{d(T^{-n}\phi(x),\, T^{-n}\phi(y))}{d(T^{-n}x,\, T^{-n}y)} \leq B.$$

Hence,

$$\begin{aligned} \frac{d(\phi(x),\, \phi(y))}{d(x,\, y)} &\leq \frac{\int_{T^{-n}(\phi(x))}^{T^{-n}(\phi(y))} \left| D_z T^n \big|_{T^{-n}\gamma_2} \right|}{\int_{T^{-n}(x)}^{T^{-n}(y)} \left| D_w T^n \big|_{T^{-n}\gamma_1} \right|} \\ &\leq e^{\sum_{j=0}^{n} c_2 \delta_2 \lambda^{-j}} \frac{d(T^{-n}\phi(x),\, T^{-n}\phi(y))}{d(T^{-n}x,\, T^{-n}y)} \leq A. \end{aligned}$$

Finally,

**Lemma II.6.** *There exists $D > 0$ such that, for each $x, y \in \gamma_2 \backslash \Sigma_\infty(2c, \delta_2)$, holds*

$$\frac{J_\phi(x)}{J_\phi(y)} \leq e^{Dd(x,\,y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}}$$

*Proof.* Again, we choose $x, y \in \gamma_2 \backslash \Sigma_\infty(2c, \delta_2)$ such that $d(x, y) \leq \delta_2$, and $n \in \mathbb{N}$ such that, for the first time,

$$d(T^{-n}x, T^{-n}y) \geq d(x, y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}.$$

Setting $z = \phi^{-1}(x)$, $w = \phi^{-1}(y)$, we have

$$\begin{aligned}
\frac{J_\phi(x)}{J_\phi(y)} &= \prod_{j=0}^{\infty} \left| \frac{D_{T^{-j}x}T\big|_{T^{-j}\gamma_2}}{D_{T^{-j}\phi^{-1}(x)}T\big|_{T^{-j}\gamma_1}} \right| \left| \frac{D_{T^{-j}\phi^{-1}y}T\big|_{T^{-j}\gamma_1}}{D_{T^{-j}y}T\big|_{T^{-j}\gamma_2}} \right| \\
&= \prod_{j=0}^{n-1} \left| \frac{D_{T^{-j}x}T\big|_{T^{-j}\gamma_2}}{D_{T^{-j}(y)}T\big|_{T^{-j}\gamma_2}} \right| \left| \frac{D_{T^{-j}\phi^{-1}y}T\big|_{T^{-j}\gamma_1}}{D_{T^{-j}\phi^{-1}(x)}T\big|_{T^{-j}\gamma_1}} \right| \\
&\times \prod_{j=n}^{\infty} \left| \frac{D_{T^{-j}(x)}T\big|_{T^{-j}\gamma_2}}{D_{T^{-j}\phi^{-1}(x)}T\big|_{T^{-j}\gamma_1}} \right| \left| \frac{D_{T^{-j}\phi^{-1}(y)}T\big|_{T^{-j}\gamma_1}}{D_{T^{-j}y}T\big|_{T^{-j}\gamma_2}} \right| \\
&\leq e^{(1+A)\sum_{j=0}^{n-1}\lambda^{-n+j}d(x,\,y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}} e^{2\sum_{j=n}^{\infty}\rho^2\lambda^{n-j}d(x,\,y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}} \\
&\leq e^{D\delta_2^{\frac{1}{2}}d(x,\,y)^{\frac{1}{2}}}.
\end{aligned}$$

**Billiards.**

We conclude with a brief discussion on how to extend the above results to the case of Sinai Billiards (with finite horizon and no breaking points).

The only real change is the necessity to better estimate $D(\phi)$. To do so, some preliminaries are called for.

In §6 we have introduced two strictly invariant cone fields: $\mathcal{C}_- = \{(1, -u) \mid u \in [1, u_0]\}$ and $\mathcal{C}_+ = \{(1, u) \mid u \in [1, u_0]\}$ ($DT^{-1}\mathcal{C}_- \subset \mathcal{C}_-$ and $DT\mathcal{C}_+ \subset \mathcal{C}_+$).

Moreover, if $v \in \mathcal{C}_-(x)$ we will call $\lambda^s(x, v)$ the expansion of $v$ under $DT^{-1}$, and, for $v \in \mathcal{C}_+$, we will call $\lambda^u(x,v)$ the expansion of $v$ under $DT$ (the length of a vector is measured in the seminorm $\sin\varphi ds$ introduced in §6). Also, let us define $\lambda_n^s(x, v) = \prod_{i=0}^{n-1}\lambda^s(T^{-i}x, D_xT^{-i}v)$ and $\lambda_n^s(x) = \prod_{i=0}^{n-1}\lambda^s(T^{-i}x, 0) \leq \inf_{v\in\mathcal{C}_-}\lambda_n^s(x, v)$, we adopt the corresponding definitions for $\lambda_n^u(x, v)$ and $\lambda_n^u(x)$. It is easy to check that $\lambda_n^u(x) = \lambda_n^s(T^nx)$, and that exists $D \in \mathbb{R}^+$ such that $D^n\lambda_n^u(x) \geq \lambda_n^u(x, v)$ for each $v \in \mathcal{C}_+$.

Within the above framework we can redefine $\Sigma_n(c, \delta_2)$ (our new approximation of the complement of $D(\phi)$).

**Definition II.7.** *Let $\mathcal{S}_n(\delta_2) = \left\{z \in \mathbb{T}^2 \mid d(z, \mathcal{S}^-) \leq \delta_2\lambda_n^u(z)^{-\alpha}\right\}$, $\alpha \in (0, \frac{1}{2})$, then*

$$\Sigma_n(c, \delta_2) = \bigcup_{i=0}^{n} T^i\mathcal{S}_i(c\delta_2).$$

The next Lemma shows that the above definition it is not too restrictive.

**Lemma II.8.** *If $\gamma_1, \gamma_2 \in \widehat{\Gamma}(\delta_1)$, $\phi : \gamma_1 \to \gamma_2$ is the canonical isomorphism, and their distance is less than $\delta_2$, then $\gamma_1 \backslash \Sigma_\infty(c, \delta_2) \supset D(\phi)$, provided $\lambda^{\alpha-1} D^\alpha < 1$, $c$ is chosen large enough, and $\delta_2$ is smaller than some fixed value.*

*Proof.* All is necessary to show is that the points in $\gamma_1 \backslash \Sigma_\infty(c, \delta_2)$ have an unstable manifold longer than some constant times $\delta_2$.

To prove such a fact, one can use the usual construction: let $x \in \gamma_1 \backslash \Sigma_\infty(c, \delta_2)$, $T^{-n}x = (s_n, \varphi_n)$, and $\overline{\gamma}_n = \{(s_n + t, \varphi_n + t) \mid t \in [-\delta_2, \delta_2]\}$, then it is well known that the connected component of $T^n \overline{\gamma}_n$ containing $x$ (eventually restricted to a $\delta_2$ neighborhood of $x$) converges to $W^u(x)$. Let us call $\widehat{\gamma}_n$ such a connected component and let us estimate its size. To do so we need first to estimate $\Lambda_-(\widehat{\gamma}_n) = \inf_{z \in \widehat{\gamma}_n} \lambda_n^s(z)$ and $\Lambda_+(\widehat{\gamma}_n) = \sup_{z \in \widehat{\gamma}_n} \lambda_n^s(z)$.

Clearly, there exists $B > 0$ such that $\frac{\Lambda_+(\widehat{\gamma}_0)}{\Lambda_-(\widehat{\gamma}_0)} \le B$, let us suppose that $\frac{\Lambda_+(\widehat{\gamma}_i)}{\Lambda_-(\widehat{\gamma}_i)} \le B$ for each $i < n$, then, if $y \in \widehat{\gamma}_n$

$$
\begin{aligned}
\frac{\lambda_n^s(x)}{\lambda_n^s(y)} &\le e^{c_4 \sum_{i=0}^{n-1} [d(T^{-i}x, T^{-i}y) \Lambda_+(\widehat{\gamma}_i)^\alpha D^{i\alpha} + d(T^{-i}x, T^{-i}y)^{\frac{1}{2}}]} \\
&\le e^{c_4 d(x, y)^{\frac{1}{2}} B^\alpha \sum_{i=0}^{n-1} \Lambda_-(\widehat{\gamma}_i)^{\alpha-1} D^{i\alpha}} \le e^{c_4 B^\alpha d(x, y)^{\frac{1}{2}} \sum_{i=0}^{n-1} (\lambda^{\alpha-1} D^\alpha)^i} \le B,
\end{aligned}
$$

providing we choose $\alpha$ such that $\lambda^{\alpha-1} D^\alpha < 1$.

The latter inequality implies $B\lambda_n^s(x) \ge \Lambda_+(\widehat{\gamma}_n) \ge \Lambda_-(\widehat{\gamma}_n) \ge B^{-1}\lambda_n^s(x)$, which suffice to prove the Lemma.

The rest of the argument goes more or less as in the bounded derivative case.

Lemma II.2 is proven in the same manner, apart from the fact that now we divide $T^n\gamma$ also according to homogeneity strips, as already discussed in §6. In particular, Lemma 6.2 contains the estimates necessary to prove the corresponding of Sub-Lemma II.3; the result follows provided one chooses $1 > \zeta > \frac{1}{1+\alpha^2}$.

The only problem with the extension of Lemma II.4 to the present context is in the distortion estimate. Let us discuss such issue.

Let $\widetilde{\gamma}_n = T^{-n}[x, y] \subset T^{-n}\gamma$, $\gamma \in \widehat{\Gamma}_{\delta_1}$. If $\widetilde{\gamma}_i$ is connected,

$$
d(T^{-i}x, T^{-i}y) \le d(x, y)^{\frac{1}{2}} \delta_2^{\frac{1}{2}},
$$

and for $\Lambda_-(\widetilde{\gamma}_i)$, $\Lambda_+(\widetilde{\gamma}_i)$ (defined as above)

$$
\frac{\Lambda_+(\widetilde{\gamma}_i)}{\Lambda_-(\widetilde{\gamma}_i)} \le B
$$

for each $i < n$, then

$$
\Lambda_-(\widetilde{\gamma}_i)^{-1} \ge d(x, y)^{\frac{1}{2}} \delta_2^{-\frac{1}{2}}.
$$

Hence, for each $z, w \in \widetilde{\gamma}_n$

$$
\frac{\lambda_n^u(z)}{\lambda_n^u(w)} \le e^{c_4 \sum_{i=0}^{n-1} \Lambda_+(\widetilde{\gamma}_i)^\alpha D^{\alpha i} d(T^{-i}x, T^{-i}y)} \le e^{c_4 B^\alpha \sum_{i=0}^{n-1} \Lambda_-(\widetilde{\gamma}_i)^{\alpha-1} D^{\alpha i} \delta_2} \le B.
$$

Accordingly, we can iterate the interval $[x, y]$ until

$$d(T^{-n}x, T^{-n}y) \geq d(x, y)^{\frac{1}{2}}\delta_2^{\frac{1}{2}}$$

without disconnecting it and maintaining uniformly bounded distortion. Lemmas II.5, II.6 follow by choosing $\alpha$ sufficiently small (although, in this case, $J_\phi$ turns out to be Hölder continuous of order strictly less than $\frac{1}{2}$).

## References


[AS] D.V. Anosov, Ya.G.Sinai, *Certain smooth ergodic systems*, Russ. Math. Surv. **22** (1982), 103 – 167.

[BD] J.P. Bouchaud, P. Le Doussal, *Numerical study of a D-dimensional periodic Lorentz gas with universal properties*, Journal of Statistical Physics **41** (1985), 225–248.

[BG] K. Burns, M. Gerber, *Continuous invariant cone families and ergodicity of flows in dimension three*, Ergodic Theory and Dynamical Systems **9** (1989).

[BSC] L. Bunimovich, Ya. Sinai, N. Chernov, *Statistical properties of two-dimensional hyperbolic billiards*, Russian Mathematical Survey **46:4** (1991), 47–106.

[BY] M. Benedicks, L-S. Young, *Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps*, Ergodic Theory and Dynamical Systems **12** (1992), 13–37.

[B1] Garret Birkhoff, *Extension of Jentzsch's Theorem*, Transaction American Mathematical Society **85** (1957), 219–227.

[B2] Garret Birkhoff, *Lattice Theory*, **25** 3rd ed., American Mathematical Society Colloquium Publications, Providence, Rhode Island, 1967.

[Bo] A. Boyarsky, *Randomness implies order*, Journal of Mathematical Analysis Applied **76** (1980), 483–497.

[Bow] R. Bowen, *Equilibrium states and ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics **470** (1975).

[Bu] P.J. Bushell, *Hilbert's metric and positive contraction mappings in a Banach space*, Archives of Rational Mechanics **52** (1973), 330–338.

[CCG] G. Casati, G. Comparin, I. Guarnieri, *Decay of correlations in certain hyperbolic systems*, Phys. Rev. A **26** (1982), 717–719.

[Ch] N.I. Chernov, *Ergodic and statistical properties of piecewise linear hyperbolic automorphism of the 2-torus*, Journal of Statistical Physics **69** (1992), 111–134.

[Co] Collet, *Ergodic properties of some unimodal mappings of the interval*, preprint, Institute Mittag-Leffler (1984).

[CS] N.I.Chernov,Ya.G.Sinai, *Ergodic properties of some systems of* 2*-dimensional discs and* 3*-dimen- sional spheres*, Russ.Math.Surv. **42** (1987), 181 – 207.

[F] P. Ferrero, *Contribution a la theorie des etats d'equilibre en mechanique statistique*, Theses (1981).

[FM] B. Friedman, R. Martin, *Decay of the velocity autocorrelation function for periodic Lorentz gas*, Physics Letters **105 A** (1984), 23-26.

[FS1] P. Ferrero, B. Schmitt, *Produits aléatories d'opérateurs matrices de transfert*, Probability Theory and Related Fields **79** (1988), 227–248.

[FS2] P. Ferrero, B. Schmitt, *On the rate of convergence for some limit ratio theorem related to endomorphisms with a non regular invariant density*, preprint.

[FS2] P. Ferrero, B. Schmitt, *Ruelle's Perron-Frobenius Theorem and Projective Metrics*, Random Fields, Esztergom (Hungary), Colloquia Mathematica Societas János Bolyai **27** (1979).

[Go] A.O. Golosov, *Small random perturbations of dynamical systems*, Transactions Moscow Society **2** (1984), 251–271.

[Ga] G. Gallavotti, private communication.

[GG] P. Garrido, G. Gallavotti, *Billiards correlation functions*, preprint.

[Gor] M.I. Gordin, *The central limit theorem for stationary processes*, Soviet Meth. Dokl. **10, 5** (1969), 1174–1176.

[GO] G. Gallavotti, D. Ornstein, *Billiards and Bernoulli schemes*, Communications in Mathematical Physics **38** (1974), 83–101.

[H] L. Hormander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171.

[Ke] G. Keller, *Un théorem de la limite centrale pour une classe de transformations monotones par morceaux*, Comptes Rendus de l'Académie des Sciences, Série A **291** (1980), 155–158.

[KB] A. Katok in collaboration with K. Burnes, *Infinitesimal Lyapunov functions, invariant cone families and stochastic properties of smooth dynamical systems*, preprint.

[KK] A. Katok, Yu. Kifer, *Random perturbations of transformations of an interval*, Journal D'Analyse Mathématique **47** (1986), 194–237.

[Ki] Yuri Kifer, *Random perturbations of dynamical systems*, Birkhäuser, 1988.

[KS] A. Katok, J.-M. Strelcyn with the collaboration of F. Ledrappier and F. Przytycki, *Invariant manifolds, entropy and billiards; smooth maps with singularities*, Lecture Notes in Math. 1222, Springer-Verlag, 1986.

[LSV] A. Lambert, S. Siboni, S. Vaienti, *Statistical Properties of a non-uniformly hyperbolic map of the interval*, preprint.

[LW] C. Liverani, M.P. Wojtkowski, *Ergodicty in Hamiltonian systems*, To appear in Dynamics Reported.

[M] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer–Verlag, Berlin Heidelberg, 1987.

[N] R.D. Nussbaum, *Hilbert projective metric*, Memoirs American Mathematical Society **391** (1988 pages 1–137).

[P] Ya. B. Pesin, *Lyapunov characteristic exponents and smooth ergodic theory*, Russian Mathematical Surveys **32, 4** (1977), 55–114.

[Ru] D. Ruelle, *Thermodynamic formalism*, Addison-Wesley, New York, 1978.

[Ry] Marek Rychlik, *Regularity of the metric entropy for expanding maps*, Transactions of the Americam Mathematical Society **315, 2** (1989), 833–847.

[Sa] H. Samelson, *On the Perron-Frobenius Theorem*, Michigan Mathematical Journal **4** (1956), 57–59.

[Si] Ya.G.Sinai, *Dynamical systems with elastic reflections*, Russ.Math.Surveys **25** (1970), 137 – 189.

[W1] M.P.Wojtkowski, *Invariant families of cones and Lyapunov exponents*, Erg.Th.Dyn.Syst. **5** (1985), 145 – 161.

[W2] M.P.Wojtkowski, *On uniformly contraction generated by positive matrices*, Random matrices and their applications, Contemporary Mathematics **50**, J.E. Cohen et al. ed., 1986, pp. 109–118.

[Y] L-S. Young, *Decay of correlations for certain quadratic maps*, Communications in Mathematical Physics **146** (1992), 123–138.

[Z] K. Ziemian, *On the Perron-Frobenius operator and limit theorems for some maps of an interval*, Ergodic Theory and Related Topics II, Proceedings, Teubner Textezür Math. **94**, 1987, pp. 206–211.



Liverani Carlangelo, Mathematics Department, University of Rome II, Tor Vergata, 00133 Rome, Italy.

*E-mail address*: liverani@mat.utovrm.it