On statistical properties of hyperbolic systems with singularities

Nikolai Chernov[1] and Hong-Kun Zhang[2]

April 20, 2009

Abstract

We study hyperbolic systems with singularities and prove the coupling lemma and exponential decay of correlations under weaker assumptions than previously adopted in similar studies. Our new approach allows us to study the mixing rates of the reduced map for certain billiard models that could not be handled by the traditional techniques. These models include modified Bunimovich stadia, which are bounded by minor arcs, and flower-type regions that are bounded by major arcs.

AMS classification numbers: 37D50, 37A25

1 Introduction

This article is devoted to hyperbolic dynamical systems with singularities. A general class of such systems was introduced in the fundamental work by Katok and Strelcyn [19]. They studied maps $T\colon M \to M$ defined on a Riemannian manifold M such that T is a C^2 diffeomorphisms from an open set $M \setminus \mathcal{S}$ onto its image; the closed set $\mathcal{S} \subset M$ is called the *singularities* of T. Katok and Strelcyn make the following assumptions on $\mathcal{S}$: the derivatives

[1]Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL 35294; chernov@math.uab.edu.

[2]Department of Math. & Statis., University of Massachusetts, Amherst, MA, 01003; hongkun@math.umass.edu

of T can only grow mildly near $\mathcal{S}$ (they are bounded by a negative power of the distance to $\mathcal{S}$), and the ε-neighborhood $U_\varepsilon(\mathcal{S})$ of $\mathcal{S}$ is not too heavy, i.e. $\mu(U_\varepsilon(\mathcal{S})) = \mathcal{O}(\varepsilon^a)$ for some constant $a > 0$; here μ denotes the T-invariant probability measure supported on M. Such assumptions are sufficient for the construction of stable and unstable manifolds, their absolute continuity, and certain formulas for the entropy of T, see [19]. In particular, stable and unstable manifolds $W^s(x)$ and $W^u(x)$ exist at μ-almost every point $x \in M$. Moreover, if $r^s(x)$ and $r^u(x)$ denote the distance from $x \in M$ to the boundaries $\partial W^s(x)$ and $\partial W^u(x)$, respectively, then

$$\mu\bigl(x \in M\colon r^{u,s}(x) < \varepsilon\bigr) \le c\varepsilon^a \tag{1.1}$$

for some constant $c > 0$; here $a > 0$ is as above. Moreover, in the studies of ergodic and statistical properties of T a 'local' version of (1.1) plays an important role; we describe it in the simplest case $\dim M = 2$ (in which case stable and unstable manifolds are one-dimensional). Let $W \subset M$ be a smooth curve uniformly transversal to all stable manifolds crossing it and m_W denote the Lebesgue measure (length) on W. Then the 'local' version of (1.1) reads

$$m_W\bigl(x \in W\colon r^s(x) < \varepsilon\bigr) \le c\varepsilon^a, \tag{1.2}$$

and a similar estimate holds for r^u if W is uniformly transversal to unstable manifolds.

The general studies by Katok and Strelcyn [19] were motivated by billiards; the latter remain the main (if not only) physically interesting class of systems with singularities. In billiards, the invariant measure μ is smooth and has bounded density, and the singularity set $\mathcal{S}$ consists of smooth compact submanifolds in M of codimension one. If the number of the smooth components of $\mathcal{S}$ is finite, which is the case for Sinai billiards with finite horizon [30, 12], then clearly $\mu(U_\varepsilon(\mathcal{S})) = \mathcal{O}(\varepsilon)$, i.e. $a = 1$ in the above formulas. In Sinai billiards with infinite horizon, $\mathcal{S}$ has countably many smooth components, and this implies $\mu(U_\varepsilon(\mathcal{S})) = \mathcal{O}(\varepsilon|\ln\varepsilon|)$, but one can conveniently change metric in M so that again $\mu(U_\varepsilon(\mathcal{S})) = \mathcal{O}(\varepsilon)$; see [12, Section 4.14]. The fact that $a = 1$ in (1.1)–(1.2) for all Sinai billiards is essential in Sinai's studies: his so-called Fundamental Theorem and his proof of ergodicity, see [30] and [12, Chapter 6], work only under the condition $a = 1$.

In the later studies of finer statistical properties of billiards and related models by Bunimovich, Sinai, and Chernov [5, 7, 8] and Young [32], the fact that $a = 1$ in (1.1)–(1.2) played a vital role, too. Even the introduction of additional (secondary) singularities (separating the so called 'homogeneity strips'), which were needed for controlling distortions (see [5, 8] and [12, Section 5.3]), increased the ε-neighborhood of all singularities, but the crucial estimates (1.1)–(1.2) with $a = 1$ remained in place. We see that all the present approaches to the studies of ergodic and statistical properties of hyperbolic systems with singularities are effectively designed for systems with $a = 1$; these include planar Sinai billiards [8], Bunimovich's stadium [21], higher-dimensional Lorentz gases [2], systems of two hard balls of different masses [10], certain abstract multidimensional models [7], and others.

Recently we extended [13] these studies to various non-traditional planar billiards such as a pinball machines in a box, skewed stadia, and flower-type tables. In these models, hyperbolicity is weak in the sense that expansion and contraction are not uniform. Then one needs to find a subset $M_1 \subset M$ such that the first return map $T_1 \colon M_1 \to M_1$ is strongly hyperbolic, i.e. its expansion and contraction are uniform. We call M_1 a *reduced space* and T_1 a *reduced map*; it preserves the measure μ conditioned on M_1. As a rule, the reduced map has more complicated singularities $\mathcal{S}_1 \subset M_1$ than the original map, i.e. $\mathcal{S}_1$ is usually a much larger set than $\mathcal{S} \cap M_1$.

As a result, $\mu(U_\varepsilon(\mathcal{S}_1))$ may be much larger than $\mu(U_\varepsilon(\mathcal{S}))$. However, in all the new models covered in [13], we still have $\mu(U_\varepsilon(\mathcal{S}_1)) = \mathcal{O}(\varepsilon)$, i.e. the reduced map satisfies the estimates (1.1)–(1.2) with $a = 1$. In fact our proofs are based on the following *one-step expansion condition* for unstable curves, which holds for the collision map in dispersing billiards and for the reduced map in the above mentioned weakly hyperbolic billiards. Given an unstable curve W (i.e. a curve whose tangent vectors lie in unstable cones), let us denote by $W_i \subset W$ the connected components of $W \setminus \mathcal{S}$, i.e. the segments of W on which T is smooth, and by Λ_i the (minimal) factor of expansion of W_i under T; then we require that

$$\liminf_{\delta_0 \to 0} \sup_{W \colon |W| < \delta_0} \sum_i \Lambda_i^{-1} < 1, \tag{1.3}$$

where the supremum is taken over unstable curves W of length $< \delta_0$. This condition quite easily leads to (1.1)–(1.2) with $a = 1$, see [13].

On the other hand, there are plenty of physical models where $\mu(U_\varepsilon(\mathcal{S}_1)) = \mathcal{O}(\varepsilon^a)$ only with $a < 1$, and (1.1)–(1.2) only hold for $a < 1$. These include large classes of planar chaotic billiards introduced by Wojtkowski [31], Markarian [20], and Donnay [18], and even more physically interesting gases of hard balls on a torus [26, 27, 28, 29] and in a box [25]. And there are no working methods developed for the studies of statistical properties for these systems. Though we cannot go that far yet, we are making steps in this direction.

In this paper we develop methods that can handle *some* systems with $a < 1$. Our approach is rather general, but we also present two examples – two Bunimovich billiard tables described in [13]. In those tables, there exist arbitrarily short unstable curves $W \subset M$ for which, in terms, of (1.3), $\Lambda_i \sim 1/i$ for all $i \geq i_0 = i_0(\delta_0)$, i.e. the series (1.3) diverges. Thus the methods of [13] do not apply, but our new method presented here works and allows us to fully investigate the mixing rates in this difficult case. We include the analysis of those two billiard tables in the end of our paper, as our main goal is the extension of the existing techniques to some general systems with $a < 1$.

Now we describe our work on some technical level. Young [32, 33] has considered abstract hyperbolic systems and gave sufficient conditions under which the dynamics can be represented in a 'semi-symbolic' way, by a tower map, leading to exponential mixing rates. The tower method is powerful, but constructing a tower in real systems, like billiards, involves fairly hard labor [32]. Chernov [8] somewhat simplified Young's construction reducing it to the verification of a certain set of conditions for just one iterate of the map (without having to deal with its higher powers). This approach was put into a more abstract form in [13] and applied to several classes of billiards.

There is an alternative technique, avoiding the tower representation altogether, based on coupling of the images of probability measures [33]. It is quite explicit and dynamical, and it is formalized in the so-called coupling lemma (see [10] and [12, Chapter 7]). The coupling method leads to a somewhat sharper estimates on correlations and more direct proofs of some limit theorems [12, Chapter 7]. Though generally it is weaker than the tower construction (see [11]), the main element of the coupling lemma – the magnet, see [12, Chapter 7] – can be defined so that it serves as the basis of Young's tower, hence producing the tower (and all its benefits) as well. We use this approach here, i.e. we prove the coupling lemma and additionally construct Young's tower.

Lastly we address our assumptions on the singularities of the map. The construction of stable and unstable manifolds can be done under very general assumptions on $\mathcal{S}$; see [19]. The construction of natural invariant measures (SRB measures) and the studies of their ergodic properties can be done under similar assumptions on $\mathcal{S}$; see [1, 22, 24]. But our studies of fine statistical properties of SRB measures require more restrictive assumptions on $\mathcal{S}$ – we assume that $\mathcal{S}$ is a piecewise smooth set and has a structure somewhat similar to billiard singularities. Though we do not need standard bounds on the complexity of singularities [16, 32], as this property is incorporated into our one-step expansion condition (3.8) generalizing (1.3).

As a last remark, we restrict our studies to 2-dimensional maps, i.e. assume that $\dim M = 2$. This will keep our presentation simpler and cleaner, though our methods extend to higher dimensions, as will be explained in the end of Section 3.

Acknowledgment. We would like to thank the anonymous referees for very helpful remarks and suggestions.

2 Statement of results

Let $T\colon M \to M$ be a C^2 diffeomorphism of a 2-dimensional Riemannian manifold M with singularities $\mathcal{S}$, i.e. T maps $M \setminus \mathcal{S}$ onto $T(M \setminus \mathcal{S})$ diffeomorphically. In our studies $\mathcal{S}$ is a finite or countable union of smooth compact curves. Assume T preserves a probability measure μ.

For any pair of integrable functions (observables) $f, g \in L^1_\mu(M)$, the *correlations* of $f \circ T^n$ and g are defined by

$$\mathcal{C}_{f,g}(n) = \int_M (f \circ T^n)\, g\, d\mu - \int_M f\, d\mu \int_M g\, d\mu, \qquad n \in \mathbb{N}. \tag{2.1}$$

It is well known that $T : M \to M$ is *mixing* if and only if

$$\lim_{n\to\infty} \mathcal{C}_{f,g}(n) = 0, \qquad \forall f, g \in L^2_\mu(M). \tag{2.2}$$

Accordingly, the rate of mixing of T is characterized by the speed of convergence in (2.2) for smooth enough functions f and g (Hölder continuity of f and g is sufficient for smooth maps, and a weaker property – dynamical Hölder continuity defined in Section 3 – is enough for maps with singularities). We say that (T, μ) enjoys exponential decay of correlations if for any pair of dynamically Hölder continuous functions f, g, there exists $b = b(f, g) > 0$ such that any $n \in \mathbb{N}$

$$|\mathcal{C}_{f,g}(n)| \le C_{f,g} e^{-bn}$$

(b only depends on the Hölder exponents of f and g). Otherwise if it can only be proved that $|\mathcal{C}_{f,g}(n)| \leq C_{f,g} n^{-a}$ for some $a > 0$, we say T enjoys polynomial decay of correlations.

In Section 3, we list our assumptions (**H.1-H.5**) on the map T.

Theorem 1. *Under the conditions* (**H.1-H.5**)*, the system* (T, μ) *enjoys exponential decay of correlations.*

We also prove Coupling Lemma, whose statement is rather technical, it is given in Section 7, and construct Young's tower.

Our paper is organized as follows. In Sections 3–8 we deal with general hyperbolic maps with singularities. In Section 3 we list our basic assumptions, **(H.1)-(H.5)**, including the new one-step expansion estimate (3.8). In Section 4 we introduce our main technical tool – unstable curves with regular probability densities on them (standard pairs and standard families). In Section 5 we prove that short unstable curves grow exponentially fast (this important property is formalized in various Growth Lemmas). Then we construct a special rectangle (the "magnet") in Section 6, after which we prove the Coupling Lemma, which directly implies Theorem 1, in Section 8. Lastly, as applications, we present two classes of billiards in Section 9.

3 A general theorem on exponential mixing

In Sections 3–8, we work with abstract hyperbolic maps, $T\colon \Omega \to \Omega$, with singularities. Assuming μ is a T-invariant mixing measure and T has singularities, we give sufficient conditions for exponential mixing. Here Ω denotes a two-dimensional connected compact Riemannian manifold. We first introduce general conditions (**H.1-H.5**), which will be assumed throughout Sections 3–8.

Let d denote the distance in Ω induced by the Riemannian metric ρ. For any smooth curve W in Ω, denote by $|W|$ its length, and by m_W the Lebesgue measure on W induced by the Riemannian metric ρ_W restricted to W. Also let $\upsilon_W = m_W/|W|$ be the normalized (probability) measure on W.

(H.1) Hyperbolicity of T (with uniform expansion and contraction). There exist two families of cones C^u_x (unstable) and C^s_x (stable) in the tangent spaces $\mathcal{T}_x\Omega$, for all $x \in \Omega$, and there exists a constant $\Lambda > 1$, with the following properties:

(1) $DT(C^u_x) \subset C^u_{Tx}$ and $DT(C^s_x) \supset C^s_{Tx}$ whenever DT exists.

(2) $\|D_xT(v)\| \geq \Lambda\|v\|, \forall v \in C^u_x$ and $\|D_xT^{-1}(v)\| \geq \Lambda\|v\|, \forall v \in C^s_x$

(3) These families of cones are continuous on Ω and the angle between C^u_x and C^s_x is uniformly bounded away from zero.

We say that a smooth curve $W \subset \Omega$ is an unstable (stable) curve if at every point $x \in W$ the tangent line $\mathcal{T}_xW$ belongs in the unstable (stable) cone C^u_x (C^s_x). As usual, a curve $W \subset \Omega$ is an unstable (resp. stable) manifold if $T^{-n}(W)$ is an unstable (resp. stable) curve for all $n \geq 0$ (resp. ≤ 0).

(H.2) Singularities and smoothness. Let $\mathcal{S}_0$ be a closed subset in Ω, such that $M := \Omega \setminus \mathcal{S}_0$ is a dense set in Ω. We put $\mathcal{S}_{\pm 1} = T^{\mp 1}\mathcal{S}_0$. We make the following assumptions:

(1) $T : M \setminus \mathcal{S}_1 \to M \setminus \mathcal{S}_{-1}$ is a C^2 diffeomorphism.

(2) $\mathcal{S}_0 \cup \mathcal{S}_1$ is a finite or countable union of smooth, compact curves in Ω.

(3) Curves in $\mathcal{S}_0$ are transversal to stable and unstable cones. Every smooth curve in $\mathcal{S}_1$ (resp. $\mathcal{S}_{-1}$) is a stable (resp. unstable) curve. Every curve in $\mathcal{S}_1$ terminates either inside another curve of $\mathcal{S}_1$ or on $\mathcal{S}_0$.

(4) There exists $\beta \in (0, 1)$ and $c > 0$ such that for any $x \in M \setminus \mathcal{S}_1$

$$\|D_xT\| \leq c\, d(x, \mathcal{S}_1)^{-\beta}. \tag{3.1}$$

The last condition is standard in [19]. In all billiards this condition holds with $\beta = 1/2$; see [12].

Remark 1. In dispersing billiards, there are natural (primary) singularities, where the map T fails to be smooth, and additional (secondary) singularities – the boundaries of the homogeneity strips; the latter are added to $\partial\Omega$ in order to guarantee the distortion bounds (3.6); see [12, Chapter 5] for a detailed description. This makes unstable manifolds homogeneous manifolds.

Remark 2. For convenience we assume that the lengths of unstable/stable manifold are uniformly bounded (by a constant, C_M). This is not a restrictive assumption, as one can always partition Ω into finitely many domains in which the unstable manifolds have bounded length, and include the boundaries of those domains in the singularity set.

Whenever we say 'the singularity of T', we refer to $\mathcal{S}\colon = \mathcal{S}_1$. Denote $\mathcal{S}_{\pm n} = \cup_{k=0}^{n-1} T^{\mp k}\mathcal{S}_{\pm 1}$ and $\mathcal{S}_{\pm\infty} = \cup_{k=0}^{\infty} T^{\mp k}\mathcal{S}_{\pm 1}$. For each $n \in \mathbb{N}$, $\mathcal{S}_n$ is a union of stable curves and $\mathcal{S}_{-n}$ is a union of smooth unstable curves. Note that $\mathcal{S}_{\pm\infty}$ could be a dense set in M. Let ξ^n be the partition of M into the connected components of $M \setminus \mathcal{S}_n$. We denote $\xi^s = M \setminus \mathcal{S}_\infty$ and $\xi^u = M \setminus \mathcal{S}_{-\infty}$. Note that any unstable manifold W^u is a connected component in ξ^u, usually with two end points in $\mathcal{S}_{-\infty}$.

Definition 1. *For every $x, y \in M$, define $\mathbf{s}_+(x,y)$, the forward* separation time *of x, y, to be the smallest integer $n \geq 0$ such that x and y belong to distinct elements of ξ^n. Similarly we define the backward separation time $\mathbf{s}_-(x,y)$. A function $f\colon M \to \mathbb{R}$ is said to be dynamically Hölder continuous, if there are $\vartheta_f \in (0,1]$ and $C_f > 0$ such that for any x and y lying on one unstable manifold W^u*

$$|f(x) - f(y)| \leq C_f \vartheta_f^{\mathbf{s}_+(x,y)} \tag{3.2}$$

and for any x and y lying on one stable manifold W^s

$$|f(x) - f(y)| \leq C_f \vartheta_f^{\mathbf{s}_-(x,y)}. \tag{3.3}$$

We denote by $\mathcal{H}^{\pm}_{\vartheta_f}$ the space of functions that satisfy only (3.3) or (3.2) with fixed ϑ_f.

Remark 3. Note that by uniform hyperbolicity, any ordinary Hölder continuous function is automatically forward and backward dynamically Hölder continuous. but on the other hand, a dynamically Hölder continuous function can be only piecewise continuous. For example, if A is a union of some unstable manifolds, then the characteristic function χ_A is backward dynamically Hölder continuous.

(H.3) Regularity of smooth unstable curves. We assume that there is a T-invariant class of unstable curves $W \subset M$ that are regular in the following sense:

(1) **Bounded curvature.** The curvature of W is uniformly bounded from above by a positive constant B.

(2) **Distortion bounds of T.** There exist $\gamma \in (0,1)$ and $C_T > 1$ such that for any regular unstable curve $W \subset M$ and any $x, y \in W$,

$$|\ln \mathcal{J}_W(x) - \ln \mathcal{J}_W(y)| \leq C_T d(x,y)^\gamma \tag{3.4}$$

where $\mathcal{J}_W(x) = |D_x T|_{T_x W}|$ denotes the Jacobian of T at $x \in W$.

(3) **Absolute continuity.** Let W_1, W_2 be two regular unstable curves close to each other. Denote

$$W_i' = \{x \in W_i \,:\, W^s(x) \cap W_{3-i} \neq \emptyset\}, \quad i = 1, 2.$$

The map $\mathbf{h} : W_1' \to W_2'$ defined by sliding along stable manifolds is called the *holonomy* map. Assume $\mathbf{h}_* m_{W_1'} \prec m_{W_2'}$. Furthermore, $\mathbf{h}$ satisfies the distortion bound:

$$|\ln \frac{\mathcal{J}\mathbf{h}(y)}{\mathcal{J}\mathbf{h}(x)}| \le C_T \vartheta^{\mathbf{s}_+(x,y)}, \qquad \forall x, y \in W_1', \tag{3.5}$$

where $\mathcal{J}\mathbf{h}$ is the Jacobian of $\mathbf{h}$.

We will only consider regular unstable curves.

REMARK 4. Note that (3.6) and the uniform hyperbolicity implies that for any smooth curve in $W \subset M$, any $x, y \in W$, if $0 \le k < \mathbf{s}_+(x, y)$, then

$$\left| \ln \frac{\mathcal{J}_W^k(x)}{\mathcal{J}_W^k(y)} \right| \le C_{\mathbf{r}} \vartheta^{\mathbf{s}_+(x,y)-k}, \tag{3.6}$$

where $\mathcal{J}_W^k(x)$ is the Jacobian of T^k at $x \in W$ and $\vartheta = \Lambda^{-\gamma}$, $C_{\mathbf{r}} = C_T/(1-\vartheta)$. Furthermore, notice for any $n \in \mathbb{N}$ and any unstable curve W in one connected component in ξ^n, the expansion factor is almost constant on W by choosing n large.

(H.4) SRB measure. μ is a Sinai-Ruelle-Bowen (SRB) measure. This means that for any unstable manifold W^u the conditional measure μ_{W^u} on W^u induced by μ is absolutely continuous with respect to m_{W^u}. We also assume that μ is mixing.

We note that under our other assumptions, the existence and *finitude* of SRB measures can be derived by standard arguments (the finitude means that there are finitely many ergodic SRB measures, and each of them is mixing up to a cyclic permutation). But we do not pursue the goals of constructing SRB measures and establishing their ergodic properties, see e.g. [1]. Since μ might not be the unique SRB measure, from now on, whenever we pick initial points (or stable/unstable curves) we take them from the basin of μ automatically without emphasizing.

Remark 5. The density function $\rho_{W^u} = d\mu_{W^u}/dm_{W^u}$ satisfies

$$\frac{\rho_{W^u}(y)}{\rho_{W^u}(z)} = \lim_{n\to\infty} \frac{\mathcal{J}^{-n}(y)}{\mathcal{J}^{-n}(z)} \tag{3.7}$$

for any $y, z \in W^u$. This is a standard formula in ergodic theory, see [12] page 105. For any unstable manifold $W^u \subset M$, the unique probability density ρ_{W^u} satisfying (3.7) is called the u-SRB density, and the corresponding probability measure μ_{W^u} on W^u is called the u-SRB measure.

For any $m \in \mathbb{N}$, the partition ξ^m induces an index set M/ξ^m. Denote V_α as the connected component in $T^m W$ with index $\alpha \in M/\xi^m$ and $W_\alpha = T^{-1}V_\alpha$. Next comes our main assumption.

(H.5) One-step expansion. There exists $q \in (0, 1]$ such that

$$\liminf_{\delta_0\to 0} \sup_{W\colon |W|<\delta_0} \sum_{\alpha\in M/\xi^1} \left(\frac{|W|}{|V_\alpha|}\right)^q \cdot \frac{|W_\alpha|}{|W|} < 1, \tag{3.8}$$

where the suppremum is taken over all unstable curves $W \subset M$.

Remark 6. For any index subset A, we define $\hat\lambda(A) = T_*\upsilon_W(V_\alpha \,|\, \alpha \in A)$, then $(W/\xi^1, \hat\lambda)$ is a probability space. Thus the inequality in (3.8) can be written as follows:

$$\liminf_{\delta_0\to 0} \sup_{W\colon |W|<\delta_0} \int_{\alpha\in W/\xi^1} \left(\frac{|W|}{V_\alpha}\right)^q d\hat\lambda(\alpha) < 1 \tag{3.9}$$

This has a clear intuitive meaning: if we regard $\zeta_q(V_\alpha) := |V_\alpha|^{-q}$ as a measurement of the length of the curve V_α; then (3.9) says that the average of this quantity decreases at each iteration; in that sense the components of the images of short unstable curves grow, on average.

Remark 7. Our assumption **(H.5)** can be extended to multidimensional hyperbolic systems, even though the 'size' $|W|$ of an unstable manifold of dimension ≥ 2 has no clear meaning. Instead, we define

$$\zeta(W) = \int_W [\,\mathrm{dist}(x, \partial W)]^{-q}\, d\upsilon_W,$$

where υ_W again denotes the normalized Lebesgue measure on W, and replace (3.8) with

$$\liminf_{\delta_0\to 0} \sup_{W\colon \mathrm{diam}(W)<\delta_0} [\zeta(W)]^{-1} \int_{M/\xi^1} \zeta(V_\alpha)\, d\hat\lambda(\alpha) < 1.$$

Now most of the results obtained in this paper carry over to higher dimensional case, but we do not pursue this goal here, as we do not have specific applications yet. We hope to do it in a separate paper.

4 Standard families of unstable curves

By uniform hyperbolicity (**H.1**), T^n expands any unstable curve $W \subset M$ at least by a factor Λ^n. At the same time, $T^n(W)$ gets broken by singularities into pieces. In this process, arbitrarily short pieces may appear, and the total number of pieces may grow exponentially with n or become infinite. Thus the hyperbolicity of T only guarantees exponential growth of unstable curves in a local sense. Accordingly, we do not expect a uniform growth for every component in $T^n(W)$, but still hope there is a certain growth at least on average at each step.

From now on we denote by W unstable curves, by C various large constants and by c small constants.

Definition 2. *Fix $C_{\mathbf{r}}$ as defined in (3.6). A probability measure ν on Ω supported on an unstable curve W is called regular, if ν is absolutely continuous with respect to the Lebesgue measure υ_W, such that the density function f satisfies*

$$|\ln f(x) - \ln f(y)| \leq C_{\mathbf{r}} \vartheta^{\mathbf{s}_+(x,y)}. \tag{4.1}$$

In that case (W, ν) is called a standard pair.

If ν is a regular probability measure supported on $W \subset M \setminus \xi^n$, for some $n \in \mathbb{N}$, then it is equivalent to the probability measure υ_W induced by the Lebesgue measure m_W in the following sense:

$$e^{-C_{\mathbf{r}}\vartheta^n} \upsilon_W(A) \leq \nu(A) \leq e^{C_{\mathbf{r}}\vartheta^n} \upsilon_W(A). \tag{4.2}$$

It follows from (3.2) that

$$e^{-C_{\mathbf{r}}\vartheta^n} \leq \frac{\min_{x\in W} f(x)}{\max_{x\in W} f(x)} \leq \frac{\max_{x\in W} f(x)}{\min_{x\in W} f(x)} \leq e^{C_{\mathbf{r}}\vartheta^n},$$

which implies

$$e^{-C_{\mathbf{r}}\vartheta^n} |W|^{-1} \leq f(x) \leq e^{C_{\mathbf{r}}\vartheta^n} |W|^{-1}.$$

Clearly, as n gets larger, the regular measure ν becomes almost uniform on W. From now on, we denote

$$c_1 = e^{C_{\mathbf{r}}} \tag{4.3}$$

Then for any regular measure ν on W,

$$c_1^{-1} \le \frac{\nu(A)}{\upsilon_W(A)} \le c_1, \qquad \forall A \subset W. \tag{4.4}$$

Note that for any standard pair (W, ν), $T_*\nu$ is a measure supported on $TW = \cup_{\alpha \in W/\xi^1} V_\alpha$. We extend the transformation T on (W, ν) as follows: the image of the standard pair (W, ν) under T can be viewed as a collection of pairs $\{(V_\alpha, \nu_\alpha) : \alpha \in W/\xi^1\}$, where ν_α is the conditional measure of $T_*\nu$ on the smooth component V_α. Now we extend our definition of standard pairs:

Definition 3. *Let $\{(W_\alpha, \nu_\alpha)\}$, $\alpha \in \mathcal{A}$ be a (countable or uncountable) family of standard pairs. We call it a standard family if there exists a probability factor measure λ on $\mathcal{A}$, which defines a measure $\nu_{\mathcal{G}}$ supported on $\mathcal{W} = \{W_\alpha \,|\, \alpha \in \mathcal{A}\}$ by*

$$\nu_{\mathcal{G}}(B) = \int_{\alpha \in \mathcal{A}} \nu_\alpha(B \cap W_\alpha)\, d\lambda(\alpha) \qquad \forall B \subset \Omega. \tag{4.5}$$

The measure $\nu_{\mathcal{G}}$ can be regarded as the 'weighted sum' or a 'convex sum' of the measures ν_α on individual standard pairs. For simplicity, we denote a standard family by $\mathcal{G} = (\mathcal{W}, \nu_{\mathcal{G}})$.

Lemma 1. *If (W, ν) is a standard pair, then $T(W, \nu)$ is a standard family.*

Proof. Let (W, ν) be any standard pair. The density function of $T_*\nu$ can be written as

$$f_1(x) = \frac{f(T^{-1}(x))}{\mathcal{J}(T^{-1}(x))}, \quad \forall x \in T(W).$$

Note that for any x, y belong to the same smooth component $V_\alpha \subset T(W)$,

$$\begin{aligned}
&|\ln f_1(x) - \ln f_1(y)| \\
&\qquad \le |\ln f(T^{-1}x) - \ln f(T^{-1}y)| + |\ln \mathcal{J}(T^{-1}(x)) - \ln \mathcal{J}(T^{-1}(y))| \\
&\qquad \le (C_{\mathbf{r}} + C_T)\vartheta^{\mathbf{s}_+(T^{-1}x, T^{-1}y)} \\
&\qquad = (C_{\mathbf{r}} + C_T)\vartheta^{\mathbf{s}_+(x,y)+1} \le C_{\mathbf{r}}\vartheta^{\mathbf{s}_+(x,y)}.
\end{aligned}$$

This implies for each $\alpha \in W/\xi^1$, (V_α, ν_α) is a standard pair, where ν_α is the conditional measure of $T_*\nu$ on V_α. Furthermore, for any A in the index set, let $\lambda(A) = T_*\nu(V_\alpha \,|\, \alpha \in A)$, then $T_*\nu$ satisfies (4.5). Accordingly, $T(W,\nu)$ is a standard family. $\square$

In general, if $\mathcal{G} = (\mathcal{W}, \nu_{\mathcal{G}})$ is a standard family with a factor measure λ and satisfies (4.5) then $T^n_*\nu_{\mathcal{G}}$ induces a standard family with $T^n\mathcal{G} := (T^n\mathcal{W}, T^n_*\nu_{\mathcal{G}})$, where for $n \geq 0$

$$T^n_*\nu_{\mathcal{G}}(B \cap T^n\mathcal{W}) := \int_{\alpha\in\mathcal{A}} T^n_*\nu_\alpha(B \cap T^n\mathcal{W}) d\lambda(\alpha), \qquad \forall B \subset \Omega. \tag{4.6}$$

Let $\mathfrak{F}$ denote the collection of all standard families $\mathcal{G}$ in M. For any $p \in (0, q]$, define a characteristic function $\mathcal{Z}_p$ on $\mathfrak{F}$, such that for any $\mathcal{G} \in \mathfrak{F}$,

$$\mathcal{Z}_p(\mathcal{G}) = \int_{\mathcal{A}} |W_\alpha|^{-p}\, d\lambda(\alpha). \tag{4.7}$$

Let $\mathfrak{F}_p$ denote those $\mathcal{G} \in \mathfrak{F}$ such that $\mathcal{Z}_p(\mathcal{G}) < \infty$.

Lemma 2. *Let $p \in (0, q]$, and $n \geq 0$.*
(1) If $\mathcal{G} \in \mathfrak{F}_p$ consists of a single curve W, denote $\lambda_n(A) = T^n_*\nu_{\mathcal{G}}(V_\alpha \,|\, \alpha \in A)$ for any smooth component $V_\alpha \subset T^nW$ and $A \subset W/\xi^n$ then

$$\mathcal{Z}_p(T^n\mathcal{G}) = \int_{W/\xi^n} \mathcal{Z}_p(\mathcal{G}_\alpha) d\lambda_n(\alpha).$$

(2) If $\mathcal{G} \in \mathfrak{F}_p$ consists of $\mathcal{W} = \cup_{\alpha\in\mathcal{A}} W_\alpha$ with $\lambda(A) = \nu_{\mathcal{G}}(W_\alpha,\, \alpha \in A)$, then

$$\mathcal{Z}_p(T^n\mathcal{G}) := \int_{\mathcal{A}} \mathcal{Z}_p(T^n\mathcal{G}_\alpha) d\lambda(\alpha).$$

Proof. (1) follows directly from the definition of $\mathcal{Z}_p$. So it is enough to show (2). If $n = 0$, it follows from the definition that

$$\mathcal{Z}_p(\mathcal{G}) = \int_{\mathcal{A}} \frac{1}{|W_\alpha|^p} d\lambda(\alpha) = \int_{\mathcal{A}} \mathcal{Z}_p(\mathcal{G}_\alpha) d\lambda(\alpha).$$

If $n > 0$, then for each smooth curve W_α in $\mathcal{W}$ we introduce the measure λ^n_α on the index space W_α/ξ^n such that for any $\beta_\alpha \in W_\alpha/\xi^n$,

$$\lambda^n_\alpha(\beta_\alpha) := T^n_*\nu_\alpha(V_{\beta_\alpha}),$$

where V_β is a smooth component in $T^n(W_\alpha)$. Denote $\mathcal{A}_n = \{W_\alpha/\xi^n : \alpha \in \mathcal{A}\}$ and define a measure λ^n on $\mathcal{A}_n$, such that for any $\beta_\alpha \in W_\alpha/\xi^n$ with $\alpha \in \mathcal{A}$, we have $\lambda^n(\beta_\alpha) = \lambda^n_\alpha(\beta_\alpha)\lambda(\alpha)$. Then $T^n_*\nu$ is a weighted sum of regular measures $\{\nu_{\beta_\alpha} : \alpha \in \mathcal{A}, \beta_\alpha \in W_\alpha/\xi^n\}$, where ν_{β_α} is the measure $T^n_*\nu_\alpha$ conditioned on V_{β_α}. Note that

$$\begin{aligned}\mathcal{Z}_p(T^n\mathcal{G}) &= \int_{\mathcal{A}_n} |V_{\beta_\alpha}|^{-p}\, d\lambda^n(\beta_\alpha) \\ &= \int_{\mathcal{A}} \int_{W_\alpha/\xi^n} |V_{\beta_\alpha}|^{-p}\, d\lambda^n_\alpha(\beta_\alpha)\, d\lambda(\alpha) \\ &= \int_{\mathcal{A}} \mathcal{Z}_p(T^n\mathcal{G}_\alpha) d\lambda(\alpha).\end{aligned}$$

□

The value of $\mathcal{Z}_p(T^n\mathcal{G})$ characterizes, in a certain way, the "average size" of the smooth components in $T^n(\mathcal{W})$, the larger they are the smaller $\mathcal{Z}_p(T^n\mathcal{G})$ is. For example, if $\mathcal{G} \in \mathfrak{F}_p$ is supported on $W \subset M \setminus \mathcal{S}_n$, then for any $k = 1, ..., n$,

$$\mathcal{Z}_p(T^k_*\nu) \leq \Lambda^{-pk}\mathcal{Z}_p(\mathcal{G}).$$

Notice T^kW is smooth for any $k = 1, ..., n$, so for any $p \in (0, q]$,

$$\mathcal{Z}_p(T^k_*\nu) = \frac{1}{|T^kW|^p} = \frac{|T^{k-1}W|^p}{|T^kW|^p} \cdots \frac{|W|^p}{|TW|^p} \cdot \frac{1}{|W|^p} \leq \Lambda^{-pk}\mathcal{Z}_p(\mathcal{G}).$$

Lemma 3. *There exist $\delta_0 > 0$ and $\theta \in (0, 1)$ such that for any standard pair $\mathcal{G} = (W, \upsilon_W)$,*
(1) if $|W| < \delta_0$, then $\mathcal{Z}_q(T\mathcal{G}) \leq \theta\mathcal{Z}_q(\mathcal{G})$;
(2) if $|W| > \delta_0$, then $\mathcal{Z}_q(T\mathcal{G}) \leq 4\theta\delta_0^{-q}$.

Proof. By (3.8), there exists $\delta_0 > 0$ such that

$$\theta := \sup_{W:|W|<\delta_0} \int_{W/\xi^1} \Big(\frac{|W|}{|V_\alpha|}\Big)^q d\hat{\lambda}(\alpha) < 1. \tag{4.8}$$

This implies for any $|W| < \delta_0$,

$$\mathcal{Z}_q(T\mathcal{G}) = \int_{W/\xi^1} \frac{1}{|V_\alpha|^q} d\hat{\lambda}(\alpha) \leq \theta\mathcal{Z}_q(\mathcal{G}). \tag{4.9}$$

On the other hand, if $|W|\geq\delta_0$, we divide (W,υ_W) into $k=[|W|/\delta_0]+1$ pieces $\{(W_1,\upsilon_1),...,(W_k,\upsilon_k)\}$ with $|W_i|\in[\delta_0/2,\delta_0)$, and υ_i being the conditional measure of υ_W on W_i. Clearly, each $\mathcal{G}_i:=(W_i,\upsilon_i)$ is a standard pair, for any $i=1,...,k$. Then by the first statement, we have

$$\mathcal{Z}_q(T\mathcal{G}_i)\leq\theta\mathcal{Z}_q(\mathcal{G}_i)=\frac{\theta}{|W_i|^q}\leq\theta\Big(\frac{2}{\delta_0}\Big)^q. \tag{4.10}$$

Note that the set $\{V_\alpha\,|\,\alpha\in W_i/\xi^1,\,i=1,...,k\}$ contains more short pieces than $T(W)$. It follows that

$$\begin{aligned}\mathcal{Z}_q(T\mathcal{G})&\leq\sum_{i=1}^k\sum_{\alpha\in W_i/\xi^1}\frac{1}{|V_\alpha|^q}T_*\upsilon_W(V_\alpha)\\&=\sum_{i=1}^k\sum_{\alpha\in W_i/\xi^1}\frac{1}{|V_\alpha|^q}T_*\upsilon_{W_i}(V_\alpha)\cdot\upsilon_W(W_i)\\&\leq\frac{\delta_0}{|W|}\sum_{i=1}^k\mathcal{Z}_q(T\mathcal{G}_i)\\&\leq\frac{\delta_0}{|W|}\theta\sum_{i=1}^k\mathcal{Z}_q(\mathcal{G}_i)\leq4\theta\delta_0^{-q}.\end{aligned}$$

This proves the second statement. $\square$

Next we will show that the value of $\mathcal{Z}_q(T^n\mathcal{G})$ decreases exponentially in n until it becomes small enough. This will imply that in any standard family, small unstable manifolds grow exponentially in size on average.

Lemma 4. *There exists $C>0$ such that for any standard pair $\mathcal{G}=(W,\upsilon_W)\in\mathfrak{F}_q$, and any $n\geq0$, one has*

$$\mathcal{Z}_q(T^n\mathcal{G})\leq\theta^n\mathcal{Z}_q(\mathcal{G})+C \tag{4.11}$$

Proof. We first prove the following formula

$$\mathcal{Z}_q(T^n\mathcal{G})\leq\theta^n\mathcal{Z}_q(\mathcal{G})+C_1(\theta+...+\theta^n), \tag{4.12}$$

where C_1 is a uniform constant. The formula (4.12) can be proved by induction on n. If $n=1$, it follows from Lemma 3 that

$$\mathcal{Z}_q(T\mathcal{G})\leq\theta\mathcal{Z}_q(\mathcal{G})+C_1\theta, \tag{4.13}$$

where $C_1 = 4\delta_0^q$. Assume that (4.12) is proved for some $n \geq 1$. Then we apply it to each component $V_\alpha \subset T(W)$ with conditional measure ν_α on W_α and obtain

$$\mathcal{Z}_q(T^n(V_\alpha, \nu_\alpha)) \leq \theta^n \mathcal{Z}_q(V_\alpha, \nu_\alpha) + C_1(\theta + ... + \theta^n).$$

By Lemma 2,

$$\begin{aligned}
\mathcal{Z}_q(T^{n+1}(\mathcal{G})) &= \int_{W/\xi^1} \mathcal{Z}_q(T^n(V_\alpha, \nu_\alpha)) d\lambda(\alpha) \\
&\leq \theta^n \int_{W/\xi^1} \mathcal{Z}_q(V_\alpha, \nu_\alpha) d\lambda(\alpha) + C_1(\theta + ... + \theta^n) \\
&= \theta^n \int_{W/\xi^1} \frac{1}{|V_\alpha|^q} d\lambda(\alpha) + C_1(\theta + ... + \theta^n) \\
&= \theta^n \mathcal{Z}_q(T\mathcal{G}) + C_1(\theta + ... + \theta^n).
\end{aligned}$$

By (4.13), we get

$$\mathcal{Z}_q(T^{n+1}\mathcal{G}) \leq \theta^{n+1} \mathcal{Z}_q(\mathcal{G}) + (\theta + ... + \theta^{n+1}) C_1.$$

Combining this with Lemma 3 gives (4.11) for $C = C_1/(1-\theta) + 1$. $\square$

The above results can be extended to any unstable curve W equipped with any regular measure ν and any standard family $\mathcal{G} \in \mathfrak{F}_q$.

Lemma 5. *There exists $C_{\mathbf{z}} > 0$, such that for any standard family $\mathcal{G} \in \mathfrak{F}_q$ and $n \geq 0$,*

$$\mathcal{Z}_q(T^n \mathcal{G}) \leq c_1 \theta^n \mathcal{Z}_q(\mathcal{G}) + C_{\mathbf{z}}, \tag{4.14}$$

where c_1 is defined as in (4.4).

Proof. If $\mathcal{G} = (W, \nu) \in \mathfrak{F}_q$, then it follows from Lemma 4 and (4.2) that

$$\lambda(A) = T_* \nu(V_\alpha \,|\, \alpha \in A) \leq e^{C_{\mathbf{r}} \vartheta} T \upsilon_W(V_\alpha \,|\, \alpha \in A),$$

where A is a subset in the index set. This implies

$$\mathcal{Z}_q(T^n(W, \nu)) \leq e^{C_{\mathbf{r}} \vartheta^n} \theta^n (\mathcal{Z}_q(W, \nu) + C). \tag{4.15}$$

Let $\mathcal{W} = \{W_\alpha \mid \alpha \in \mathcal{A}\}$. We first apply (4.15) to each W_α with conditional measure ν_α:

$$\mathcal{Z}_q(T^n(W_\alpha, \nu_\alpha)) \leq e^{C_{\mathbf{r}}\vartheta^n}(\theta^n \mathcal{Z}_q(W_\alpha, \nu_\alpha) + C) \tag{4.16}$$

Note that by Lemma 2,

$$\mathcal{Z}_q(T^n\mathcal{G}) = \int_{\mathcal{A}} \mathcal{Z}_q(T^n(W_\alpha, \nu_\alpha))d\lambda(\alpha).$$

It follows from (4.16) that

$$\mathcal{Z}_q(T^n\mathcal{G}) \leq c_1\theta^n \mathcal{Z}_q(\mathcal{G}) + C_{\mathbf{z}},$$

where $C_{\mathbf{z}} = Ce^{C_{\mathbf{r}}}$. □

We see that if $\mathcal{Z}_q(\mathcal{G})$ is very large, the sequence $\{\mathcal{Z}_q(T^n\mathcal{G}), n \in \mathbb{N}\}$ will decrease exponentially fast until it reaches a certain threshold. In particular, for any $p \in (0, q)$, (3.8) also holds for q replaced by p. Thus Lemmas 3-5 are still valid if we replace q by p. This implies that even if $\mathcal{Z}_p(\mathcal{G})$ is very large for some $p \in (0, q]$, eventually, $\mathcal{Z}_p(T^n\mathcal{G})$ will be under control and less than certain fixed constant.

5 Growth Lemmas

Growth Lemmas show that expansion always prevails over fragmentation. Here the Growth Lemmas follow from our one-step expansion (3.8).

Let $\mathcal{G} = (\mathcal{W}, \nu_{\mathcal{G}})$ be a standard family. To get a better control of $T^n\mathcal{G}$, we need to estimate the size of "bad" points in $\mathcal{W}$ whose images under T^n get too close to the singular set $\mathcal{S}_{-n}$. For any $\varepsilon > 0$, $n \in \mathbb{N}$, define

$$B_{\varepsilon,n}(\mathcal{W}) := \{x \in \mathcal{W} \colon |V_\alpha(x)| < 2\varepsilon, \alpha \in \mathcal{W}/\xi^n\}. \tag{5.1}$$

Clearly, the set $B_{\varepsilon,n}(\mathcal{W})$ contains points in $\mathcal{W}$ whose T^n images are contained in short unstable curves. Let

$$\mathrm{F}_n(\varepsilon) = \nu_{\mathcal{G}}(B_{\varepsilon,n}(\mathcal{W})), \qquad n \in \mathbb{N} \tag{5.2}$$

be the distribution of $B_{\varepsilon,n}(\mathcal{W})$ in the probability space $T^n\mathcal{G}$. In fact we will see, for "typical" regular standard families $\mathcal{G}$, $\mathrm{F}_n(\varepsilon)$ will decay exponentially in n. Given any standard pair $\mathcal{G} = (\mathcal{W}, \nu_{\mathcal{G}})$, for any $x \in \mathcal{W}$, denote $r_{\mathcal{W}}(x)$ or $r_{\mathcal{G}}(x)$ as the shortest distance from x to $\partial\mathcal{W}$ measured along $\mathcal{G}$.

Lemma 6. *There exists $c > 0$ such that for any $p \in (0, q]$, and standard family $\mathcal{G} \in \mathfrak{F}_p$, any $\varepsilon > 0$ and $n \geq 0$, we have*

$$\nu_{\mathcal{G}}(r_{\mathcal{G}_n}(x) < \varepsilon) \leq (c_1\theta^n \mathcal{Z}_q(\mathcal{G}) + C_{\mathbf{z}})\varepsilon^p, \tag{5.3}$$

Proof. Denote by $\mathcal{A}_\varepsilon \subset \mathcal{W}/\xi^n$, such that for any $\alpha \in \mathcal{A}_\varepsilon$, $|V_\alpha| < 2\varepsilon$. Thus

$$\begin{aligned}\mathcal{Z}_p(T^n\mathcal{G}) = \int_{\mathcal{W}/\xi^n} |V_\alpha|^{-p} d\lambda^n(\alpha) > \int_{\mathcal{A}_\varepsilon} |V_\alpha|^{-p} d\lambda^n(\alpha)\\ \geq \varepsilon^{-p} \int_{\mathcal{A}_\varepsilon} d\lambda^n(\alpha) = \varepsilon^{-p}\nu_{\mathcal{G}}(B_{\varepsilon,n}(\mathcal{W})).\end{aligned}$$

Note that for any $n \geq 0$, the set $\{x \in \mathcal{W} : r_{\mathcal{G}_n}(x) < \varepsilon\}$ contains two parts. One is $B_{\varepsilon,n}(\mathcal{W})$, and the other is

$$D_n := T^{-n}\{y \in V_\alpha : d_{V_\alpha}(y, \partial V_\alpha) < \varepsilon, \alpha \in \mathcal{A}_\varepsilon^c\},$$

where $d_{V_\alpha}(\,,\,)$ is the distance measured along V_α. Since for any $\alpha \in \mathcal{A}_\varepsilon^c$, $|W_\alpha| \geq 2\varepsilon$. Thus the measure of D_n is bounded by

$$\begin{aligned}\nu_{\mathcal{G}}(D_n) \leq c_1 \int_{(2\varepsilon,\infty)} \frac{2\varepsilon}{s} d\mathrm{F}_n(s)\\ \leq 2c_1\varepsilon \int_{(\varepsilon,\infty)} |s|^{p-1}|s|^{-p} d\mathrm{F}_n(s) = 2c_1\varepsilon^p \mathcal{Z}_p(T^n\mathcal{G}),\end{aligned}$$

where c_1 is the constant defined in (4.3) and F_n is the distribution defined as in (5.2). This implies

$$\nu_{\mathcal{G}}(r_{\mathcal{G}_n}(x) < \varepsilon) \leq \nu_{\mathcal{G}}(B_{\varepsilon,n}(\mathcal{W})) + \nu_{\mathcal{G}}(D_n) \leq c\varepsilon^p \mathcal{Z}_p(T^n\mathcal{G}) \tag{5.4}$$

where $c = 1 + 2c_1$. □

Motivated by the above lemma, we introduce the notion of *proper* families whose characteristic functions are uniformly bounded. Let

$$C_{\mathrm{q}} > \frac{C_{\mathbf{z}}}{1-\theta}, \tag{5.5}$$

be a very large constant. Given a standard family $\mathcal{G}$, if $\mathcal{Z}_q(\mathcal{G}) < C_{\mathrm{q}}$ we say $\mathcal{G}$ is a proper standard family. Denote by $\mathfrak{F}_q^* = \mathcal{Z}_q^{-1}(0, C_{\mathrm{q}}]$ the set of all proper standard families in M. Similarly, for $p \in (0, q)$, denote $\mathfrak{F}_p^* = \mathcal{Z}_p^{-1}(0, C_{\mathrm{q}} + 1]$ as the set of all proper standard families in M with respect to p. Note that for $p < q$, $\mathfrak{F}_q^* \subset \mathfrak{F}_p^*$.

Corollary 1. *For any $p \in (0, q]$ and $\mathcal{G} \in \mathfrak{F}_p$, there exists $N = N(\mathcal{G})$, such that for any $n > N$, $T^n\mathcal{G} \in \mathfrak{F}_p^*$. In particular, there exists $\tilde{\delta} > 0$, for any unstable curve $W \subset M$ with length $> \tilde{\delta}$ and regular measure ν on W, $(W, \nu) \in \mathfrak{F}_q^*$.*

Corollary 2. *For any $p \in (0, q]$, there exists $c > 0$ such that for any $\mathcal{G} \in \mathfrak{F}_p^*$, any $\varepsilon > 0$ and $n \geq 0$,*

$$\nu_{\mathcal{G}}(r_{\mathcal{G}_n}(x) < \varepsilon) \leq c\varepsilon^p. \tag{5.6}$$

Next we state the Growth Lemmas similar to that of [12, Chapter 7], which follows from Lemma 5 and Lemma 6.

Lemma 7 (First Growth Lemma)**.** *There exist constants $c > 0, C > 0$ such that for any $\varepsilon \in (0, 1)$ and any standard pair $\mathcal{G} = (W, \nu)$, we have*

$$\nu(r_{\mathcal{G}_n}(x) < \varepsilon) \leq c\,\nu(r_{\mathcal{G}}(x) < \theta^n \varepsilon^q) + C\varepsilon^q \tag{5.7}$$

Lemma 8 (Second Growth Lemma)**.** *There exist constants $c > 0, \chi > 0$ such that for any standard pair $\mathcal{G} = (W, \nu)$, $\mathcal{G}_n \in \mathfrak{F}_q^*$ for any $n > \chi|\ln|W||$. Furthermore, for any $\varepsilon \in (0, 1)$,*

$$\nu(r_{\mathcal{G}_n}(x) < \varepsilon) \leq c\varepsilon^q, \qquad \forall n > \chi|\ln|W|| \tag{5.8}$$

The Growth Lemmas imply that a standard pair (W, ν) will eventually be proper in $\mathfrak{F}_q$ after a certain number of iterations. But for arbitrary standard family $\mathcal{G}$, its images might not belong to $\mathfrak{F}_q^*$. It all depends on the distributions of short unstable manifolds in $\mathcal{G}$. By choosing $p < q$, we still might guarantee that $T^n\mathcal{G}$ belongs to $\mathfrak{F}_p^*$ for all large n. As we will see that the existence of such a $p \in (0, q]$ is good enough for the proof of the Coupling Lemma. The Growth Lemmas guarantee that that for any large n, the n-th image of any standard pair (W, ν) is a proper family. Thus the set of points on T^nW which come too close to the singularity set has very small measure. We still need to prove the corresponding global estimates.

Note that for any $x \in M \setminus \mathcal{S}_{-\infty}$, $W^u(x)$ exists and connects x with $\mathcal{S}_{-\infty}$, see [12] page 93-95, so we define $d^u(x, \mathcal{S}_{-\infty})$ as the distance from x to $\mathcal{S}_{-\infty}$ along $W^u(x)$. Let $U^u_\varepsilon(\mathcal{S}_{-\infty})$ be the ε-neighborhood of $\mathcal{S}_{-\infty}$ in the d^u metric. Denote by $r_n(x) = d^u(T^nx, \mathcal{S}_{-\infty})$, and $r^u(x) = d^u(x, \mathcal{S}_{-\infty})$. Similarly, we define $d^s(x, \mathcal{S}_\infty)$ and $U^s_\varepsilon(\mathcal{S}_\infty)$. Given a standard family $\mathcal{G} = (\mathcal{W}, \nu)$, if $\mathcal{W}$ is made of maximal unstable manifolds, then the shortest distance from x to $\partial\mathcal{W}$ measured along $\mathcal{W}$ is $r(x)$. This implies $r_{\mathcal{G}_n}(x) = r_n(x)$, for any $n \geq 0$ and any $x \in \mathcal{W}$.

Lemma 9. *There exists $\tilde{c} > 0$ such that for any $x \in M \setminus \mathcal{S}_\infty$,*

$$r^s(x) \geq \inf_{n \geq 0} \tilde{c}\Lambda^n r_n(x). \tag{5.9}$$

Furthermore, there exists a uniform constant $c = c(T) > 0$ such that for any proper standard pair $\mathcal{G} = (W, \nu)$,

$$\nu(r^s(x) < \varepsilon) \leq c\varepsilon^q. \tag{5.10}$$

Proof. For every $x \in M \setminus \mathcal{S}_\infty$ there exists $\hat{C} > 0$, such that

$$r^s(x) \geq \inf_{n \geq 0} \hat{C}\Lambda^n d^s(T^n x, \mathcal{S}_1). \tag{5.11}$$

Since both stable and unstable manifolds have uniformly bounded curvature, there exist $\tilde{c}, \hat{c} > 0$, such that for any $x \in M$,

$$\tilde{c}d^u(x, \mathcal{S}_{-1}) \leq \hat{c}d(x, \mathcal{S}_{-1}) \leq \hat{C}d(x, \mathcal{S}_1) \leq \hat{C}d^s(x, \mathcal{S}_1).$$

Hence

$$r^s(x) \geq \inf_{n \geq 0} \tilde{c}\Lambda^n r_n(x). \tag{5.12}$$

Therefore

$$\nu(r^s(x) < \varepsilon) \leq \sum_{n=0}^{\infty} \nu(r_n(x) < \tilde{c}^{-1}\Lambda^{-n}\varepsilon).$$

Due to the Second Growth Lemma,

$$\nu(r_n(x) < \tilde{c}^{-1}\Lambda^{-n}\varepsilon) \leq C_1\Lambda^{-np}\varepsilon^q.$$

Summing over all $n \geq 0$, we prove (5.10). □

The above lemma recovers the fact that on any long unstable manifold W^u, a majority of points $y \in W^u$ have long stable manifolds stretching far beyond W^u on both sides of W^u. Denote by $\mathcal{W}^u$ the family of all maximal unstable manifolds in $M \setminus \mathcal{S}_\infty$.

Lemma 10. *For any small $\kappa > 0$, define*

$$N_\kappa = \cap_{n=0}^{\infty}\left(x \in M \,|\, r_n(x) \geq \frac{\kappa}{\Lambda^n}\right). \tag{5.13}$$

Then there exists $d > 0$ such that for any proper family $\mathcal{G} = (\mathcal{W}, \nu) \in \mathfrak{F}_p^*$ with $\mathcal{W} \subset \mathcal{W}^u$

$$\nu_{\mathcal{G}}(N_\kappa) > d. \tag{5.14}$$

Proof. Note that for any $x \in (N_\kappa)^c$, there exists $n > 0$ such that $r_n(x) < \kappa\Lambda^{-n}$. By Corollary 2, there exists $c > 0$ such that

$$\nu_{\mathcal{G}}(r_n(x) < \kappa\Lambda^{-n}) \leq c\Lambda^{-np}.$$

Thus $\nu_{\mathcal{G}}(N_\kappa) \geq d$, where $d = 1 - c\Lambda^{-q}$. □

Lemma 11. *There exists $c > 0$ such that for any small $\varepsilon > 0$ one has*

$$\mu(U^u_\varepsilon(\mathcal{S}_{-\infty})) \leq c\varepsilon^q, \qquad \mu(U^s_\varepsilon(\mathcal{S}_{\infty})) \leq c\varepsilon^q \tag{5.15}$$

Proof. Let W^u be a maximum unstable manifold with length larger than $\tilde{\delta}$ contained in the basin of μ. According to the Second Growth Lemma, there exists $c > 0$ such that for any $n > 0$,

$$\upsilon_{W^u}(r_n(x) < \varepsilon) \leq c\varepsilon^q.$$

On the other hand, for each $n \geq 1$,

$$T^n_* \upsilon_{W^u}(U^u_\varepsilon(\mathcal{S}_{-\infty})) \leq \upsilon_{W^u}(r_n(x) < \varepsilon) \leq c\varepsilon^q.$$

Let

$$\eta_n = \frac{1}{n}\sum_{k=0}^{n-1} T^k_* \upsilon_{W^u}.$$

Since (T, μ) is mixing, so there exists a subsequence η_{m_n} that converges to μ i.e. $\lim_{n\to\infty} \eta_{m_n} = \mu$. For detailed reference, see [23, 22, 24, 1]. This implies

$$\mu(U^u_\varepsilon(\mathcal{S}_{-\infty})) < c\varepsilon^q.$$

The second inequality easily follows from Lemma 9. □

Lemma 12. *Let $\mathcal{W}^u$ be the collection of all unstable manifolds in M. Then $\mathcal{E} = (\mathcal{W}^u, \mu)$ is a standard proper family in $\mathfrak{F}_p$, for any $p < q$.*

Proof. Let ξ^u be the measurable partition of $(M, \mathcal{B}, \mu)$ into smooth unstable manifolds. Then $\mathcal{A} = M/\xi^u$ is the index set, and for any $\alpha \in \mathcal{A}$, μ_α is the conditional measure of μ on the unstable manifold W_α. Furthermore, λ is the factor measure associated with the partition. Note ν_α has the following properties. First, by (3.7) and the distortion bound, for each $\alpha \in \mathcal{A}$, (W_α, ν_α)

is a standard pair. Second, for every $B \in \mathcal{B}$, $\nu_\alpha(B \cap W_\alpha)$ is a measurable function of α, and

$$\mu(B) = \int_\alpha \nu_\alpha(B \cap W_\alpha) d\lambda(\alpha).$$

Thus $\mathcal{E} = (\mathcal{W}^u, \mu)$ is a standard family. Note that for any fixed $\delta > 0$ and $p < q$,

$$\mathcal{Z}_p(\mathcal{E}) = \int |W_\alpha|^{-p} d\lambda(\alpha) = \int_0^\delta s^{-p} d\mathrm{F}_0(s) + \int_\delta^\infty s^{-p} d\mathrm{F}_0(s),$$

where $\mathrm{F}_0(\varepsilon)$ denote the measure of short unstable manifolds with length $|W| < \varepsilon$. The second integral is bounded by δ^{-p}, so it suffices to show that the first integral is bounded. According to Lemma 11, for any $\varepsilon > 0$,

$$\int_\varepsilon^\infty \frac{\varepsilon}{s} d\mathrm{F}_0(s) \leq \mu(U_\varepsilon^u(\mathcal{S}_{-\infty})) \leq c\varepsilon^q,$$

which implies that

$$\int_\varepsilon^{2\varepsilon} \frac{1}{s^p} d\mathrm{F}_0(s) = \int_\varepsilon^{2\varepsilon} \frac{s^{1-p}}{s} d\mathrm{F}_0(s) \leq c_1 \varepsilon^{q-p}.$$

Accordingly, we have

$$\begin{aligned}
\int_0^\delta \frac{1}{s^p} d\mathrm{F}_0(s) &= \sum_{m=0}^\infty \int_{\frac{\delta}{2^{m+1}}}^{\frac{\delta}{2^m}} \frac{1}{s^p} d\mathrm{F}_0(s) \\
&\leq c_2 \sum_{m=0}^\infty 2^{(q-p)m} < c_3.
\end{aligned}$$

By choosing δ small enough, we see that for any $p < q$, $\mathcal{E} \in \mathfrak{F}_p$. □

Growth Lemmas guarantee that for each individual unstable manifold W^u, no matter how short it is, eventually $(T^n W^u, \mu_{W^u})$ will be proper in $\mathfrak{F}_q^*$. In addition, Lemma 11 indicates that $\mathcal{E}$ should belong to $\mathfrak{F}_q$, but we have not achieved this goal. Instead, we will see that $\mathcal{E} \in \mathfrak{F}_p^*$ for $p < q$, and this will be good enough for our purpose. From now on, we fix $p < q$ and only consider proper families in $\mathfrak{F}_p^*$.

6 Construction of the "magnet"

In this section, we will first construct a family of stable manifolds crossing an unstable manifold. It will serve as a "magnet" along which measures will be coupled. For any $x \in M$, let $W^{u,s}(x)$ denote the maximal smooth unstable/stable manifold of x. In this section whenever we say a smooth unstable manifold, we really mean a *maximum* smooth unstable manifold with length less than C_M, see Remark 2. For any $\kappa > 0$, let $W^{u,s}_{\kappa}(x)$ denote portion of the smooth unstable/stable manifold centered at x with length 2κ. Denote

$$N_{\tilde{\delta}} = \cap_{n=0}^{\infty}\left(x \in M \mid r_n(x) \geq \frac{2\tilde{\delta}}{\Lambda^n}\right). \tag{6.1}$$

By (5.12), for any $x \in N_{\tilde{\delta}}$, both $W^s(x)$ and $W^u(x)$ exist with $r^s(x) \geq 2\kappa$ and $r^u(x) \geq 2\tilde{\delta}$, where $\kappa = \tilde{c}\tilde{\delta}$.

Note that even if $W^s(x)$ has length $> 2\kappa$, x may not belong to $N_{\tilde{\delta}}$, since $T^n x$ may approach to $\mathcal{S}_{-\infty}$ much faster than $\tilde{\delta}/\Lambda^n$. On the other hand, according to Lemma 10 and Lemma 12, there exists $d > 0$ such that

$$\mu(N_{\tilde{\delta}}) > d.$$

Furthermore, $x \in N_{\tilde{\delta}}$, we have

$$\upsilon_{W^u(x)}(N_{\tilde{\delta}}) > d, \tag{6.2}$$

where we used the fact that any unstable manifold has length less that C_M.

For any $x \in N_{2\tilde{\delta}}$, let

$$\Gamma^s(x) = \{W^s(y) \mid y \in W^u_{\tilde{\delta}}(x) \cap N_{\tilde{\delta}}\}.$$

Note that Γ^s is the collection of all maximal stable manifolds along $W^u_{\tilde{\delta}}(x) \cap N_{\tilde{\delta}}$ which stick on both sides of $W^u(x)$ by at least κ. As the length of stable manifolds in $\Gamma^s(x)$ maybe very irregular, so we need to chop off a portion to get our magnet.

Let $\mathcal{U}(x)$ be a "rectangular" shaped region such that $W^s_{\kappa/2}(x) \subset \mathcal{U}(x)$ and $W^u_{\tilde{\delta}/2}(x) \subset \mathcal{U}(x)$. And the boundary $\partial\mathcal{U}(x)$ is made of 2 unstable manifolds with length $\tilde{\delta}$ and 2 stable manifolds with length κ. Accordingly, the region $\mathcal{U}(x)$ can be viewed as a "rectangle" centered at x with dimensions $\tilde{\delta} \times \kappa$.

We say that an unstable manifold W^u *fully u-crosses* $\Gamma^s(x)$, if W^u meets every stable manifold in $\Gamma^s(x)$. Let $\Gamma^u(x)$ be the collection of all maximal unstable manifolds $W^u(y)$ that fully u-cross $\Gamma^s(x)$, with $y \in W^s(x) \cap \mathcal{U}(x)$. Define

$$\mathcal{R}(x) = \Gamma^s(x) \cap \Gamma^u(x) \cap \mathcal{U}(x) \cap N_{\tilde{\delta}/4}. \tag{6.3}$$

According to (6.2) and the distortion bound of the stable holonomy map $\mathbf{h}$, there exists $d_0 > 0$ such that for any $W \in \Gamma^u$,

$$\upsilon_W(W \cap \mathcal{R}(x)) > d_0. \tag{6.4}$$

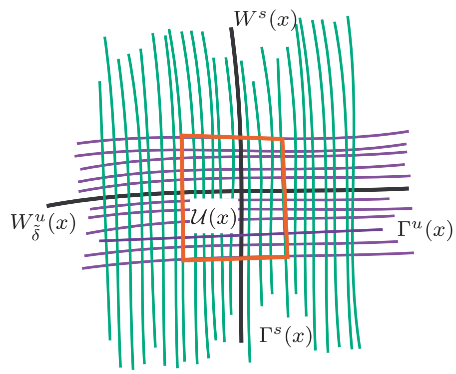

Figure 1: $\mathcal{R}(x) = \mathcal{U}(x) \cap \Gamma^u(x) \cap \Gamma^s(x) \cap N_{\tilde{\delta}/4}$

We fix $x_0 \in N^u_{2\tilde{\delta}}$, and define our magnet by $\mathcal{R}^* := \mathcal{R}(x_0)$ and $\Gamma^{s/u} = \Gamma^{s/u}(x_0)$. The magnet will be used to couple regular measures supported on unstable manifolds in Γ^u.

Lemma 13. *There exists $n_0 > 0$, $d_1 > 0$, for any standard pair (W^u, ν) with $|W^u| > 4\tilde{\delta}$ and $n \geq n_0$,*

$$\nu(T^{-n}\mathcal{R}^* \cap W^u) \geq d_1. \tag{6.5}$$

Proof. Let $\mathcal{H}(\Omega)$ be the set of all closed unstable curves on Ω equipped with the topology induced by the Hausdorff metric. Then $\mathcal{H}(\Omega)$ is compact and complete. Thus the set of all closed unstable curves of length larger or equal to $\tilde{\delta}$ is a compact set in $\mathcal{H}(\Omega)$. This implies that there exists $m_0 > 0$ such that $\mathcal{U}(x_1)$, ..., $\mathcal{U}(x_{m_0})$ is a cover of $\{x \in M \,|\, |W^u(x)| \geq 4\tilde{\delta}\}$. Accordingly, any unstable manifolds with length longer than $4\tilde{\delta}$ must be fully u-across $\mathcal{R}(x_i)$, for some $i \leq m_0$. Now by the mixing property, there exist $n_0 > 0$, and $d_1 > 0$ such that for any $n \geq n_0$,

$$\mu(T^n\mathcal{R}(x_i) \cap \mathcal{R}^*) > d_1.$$

According to assumption (**H3**), both T and $\mathbf{h}$ satisfy the distortion bound. Thus for any unstable manifold W^u with length longer than $4\tilde{\delta}$ and a regular measure ν on W^u, there exists $\hat{d}_2 \leq d_1$ such that for $n \geq n_0$,

$$\nu(W^u \cap T^{-n}\mathcal{R}^*) \geq \hat{d}_2. \tag{6.6}$$

Note that here we used the fact that $|W^u| < C_M$ by the remarks under (**H.2**). Accordingly, by (6.4), we get (6.5). □

Lemma 14. *Fix $p \in (0, q)$. There exists a positive constant $d_0 \in (0, 1)$ such that for any $\mathcal{G} = (\mathcal{W}, \nu_\mathcal{G}) \in \mathfrak{F}_p^*$ and $n \geq n_0$ we have $\nu_\mathcal{G}\,(\mathcal{W} \cap T^{-n}\mathcal{R}^*) > d_0$.*

Proof. For any $p \in (0, q)$ and $\mathcal{G} = (\mathcal{W}, \nu_\mathcal{G}) \in \mathfrak{F}_p^*$, it follows from Corollary 2 that $\nu_\mathcal{G}(B_{2\tilde{\delta},n}(\mathcal{W})) \leq c(2\tilde{\delta})^p$. This implies that for any $n > n_0$,

$$\begin{aligned}\nu_\mathcal{G}(\mathcal{W} \cap T^{-n}\mathcal{R}^*) &\geq \nu_\mathcal{G}(\mathcal{W} \cap T^{-n}\mathcal{R}^*, |W_\alpha| > 4\tilde{\delta}) \\ &\geq d_1(1 - c(2\tilde{\delta})^p).\end{aligned}$$

This Lemma says that if $\mathcal{G} \in \mathfrak{F}_p^*$, then for any $n \geq n_0$, at least d_0 fraction of $T^n\mathcal{G}$ has base points returning to $\mathcal{R}^*$ properly.

7 Coupling Lemma

In this section, we fix $p \in (0, q]$, and only consider standard pairs or families in $\mathfrak{F}_p$ built on regular unstable manifolds.

Let (W, ν_W) be a standard pair, and define

$$\hat{W} = \{(x, t) \,|\, x \in W,\, t \in [0, 1]\}.$$

Then $\hat{W}$ is a rectangle based on W. We equip $\hat{W}$ with a probability measure $\hat{\nu}$, such that for any $(x, t) \in \hat{W}$,

$$d\hat{\nu}_W(x, t) = d\nu_W(x)dt. \tag{7.1}$$

Note that the map T^n defined on W can be extended to $\hat{W}$ with $T^n(x,t) := (T^n(x),t)$. Let $\mathcal{G} = \{(W_\alpha, \nu_{W_\alpha}), \alpha \in \mathcal{A}\}$ be a standard family with probability measure $\nu_{\mathcal{G}}$. Then the rectangles based on $\mathcal{W}$ are

$$\hat{\mathcal{G}} = \{(x,t)\colon x \in \cup_\alpha W_\alpha,\, t \in [0,1]\}.$$

Again we define the probability measure $\hat{\nu}_{\hat{\mathcal{G}}}$, such that $d\hat{\nu}_{\hat{\mathcal{G}}}(x,t) = d\nu_{\mathcal{G}}(x)dt$.

Lemma 15. *Given any two families $\mathcal{G}, \mathcal{E} \in \mathfrak{F}_p$, there exists a measure preserving map (the coupling map) $\Theta : \hat{\mathcal{G}} \to \hat{\mathcal{E}}$ with*

$$\Theta(x,t) = (y,s), \quad \Theta_* \hat{\nu}_{\hat{\mathcal{G}}} = \hat{\nu}_{\hat{\mathcal{E}}}$$

and a coupling time function Υ defined on $\hat{\mathcal{G}}$, such that $T^{\Upsilon(x,t)}x$ and $T^{\Upsilon(x,t)}y$ lie on the same stable manifold. Furthermore, the coupling time function $\Upsilon : \hat{\mathcal{G}} \to \mathbb{N}$ has exponential tail bound:

$$\hat{\nu}_{\hat{\mathcal{G}}}(\Upsilon > n) \leq C_\Upsilon \vartheta_\Upsilon^n, \tag{7.2}$$

where C_Υ is a positive constant, and $\vartheta_\Upsilon \in (0, 2\Lambda^{-p})$.

The main idea of the proof of this lemma is that we first fix a special subset $\mathcal{R}$ with hyperbolic structure as we constructed in last section. Then we try to match the measures $\hat{\nu}_{\hat{\mathcal{G}}}$ and $\hat{\nu}_{\hat{\mathcal{E}}}$ according to the simultaneous, proper returns of the unstable manifolds (bases of regions) to $\mathcal{R}$. At each coupling time Υ, a fraction of both measures are matched and pumped out of the system. Thus the total measure remaining at time n is an upper bound for $|T_*^n \hat{\nu}_{\hat{\mathcal{G}}} - T^n \hat{\nu}_{\hat{\mathcal{E}}}|$. In next section we will show that the speed of coupling leads to the rate of decay of correlations. Now we consider an ideal case. At n_0, both regions $\hat{\mathcal{G}}$ and $\hat{\mathcal{E}}$ have at least d_0 fraction whose base images under T^{n_0} properly return $\mathcal{R}^s$. If we couple $d = d_0/2$ fraction, then both regions remain $1-d$ fraction of total measure. Assume the two new families still satisfy the assumptions, then we couple another d fraction of measure after another n_0 iterations. In this way, after n_0 iterations, we will couple all points on both regions except $(1-d)^n$ fraction of each measure. This would give us the exponential tail bound for the coupling time function, which will lead to the exponential decay of correlations.

However, the real situation is that after each coupling, the two remaining density functions do not satisfy distortion bound any more. To guarantee distortion bound, we need to cut the regions into pieces at the place they got coupled. Thus the base of each region contains tons of arbitrarily short unstable manifolds, which may need arbitrarily long time to recover. In other words, the corresponding two new families are not proper anymore. More precisely, notice for any unstable manifold W that fully u-crosses $\mathcal{R}^s$, $W \cap \mathcal{R}^s$ is a closed nowhere dense (Cantor-like) set on W. After coupling, the remaining set will be a countable union of very short unstable manifolds. Next we need to estimate the recovery time for the remaining manifolds.

The proof of the Coupling Lemma follows from a similar argument as in [9, 10] and [12] Section 7.12-7.15, which we will not repeat except the following technical changes due to our general assumption on singularities. Let V be a connected component (or a gap) of $(W \setminus \mathcal{R}^*) \cap \mathcal{U}$. Then we define the rank of V as the first moment when the image of V is split into pieces.

Lemma 16. *Let V be a connected component (or a gap) of $(W \setminus \mathcal{R}^*) \cap \mathcal{U}$ of rank n. There there exists $C > 0$ such that*

$$|T^{n-1}V| \geq C\Lambda^{\frac{n}{\beta-1}}, \tag{7.3}$$

where $\beta \in (0,1)$ was given in (3.1).

Proof. Since rank $V = n$, so $T^n(V)$ gets split for the first time. Let x be any one of the end points of V. By (6.3), we know $x \in \mathcal{R}^* \subset N_{\tilde{\delta}/4}$. This implies that for any $m \geq 1$,

$$r_m(x) \geq \tilde{\delta}/(2\Lambda^m). \tag{7.4}$$

Let V_0 be the smooth component in T^nV that contains T^nx. By (7.4),

$$|V_0| \geq r_n(x) \geq \tilde{\delta}/(2\Lambda^n).$$

It follows from the regularity of unstable curves that there exists $c_1 > 0$ such that for the middle point $y \in T^{-1}V_0$,

$$|V_0| \leq c_1 \mathcal{J}_{T^{-1}V_0}(y)|T^{-1}V_0|$$

By our assumption on singular curves, there exist $c_2, c_3 > 0$ such that

$$= d(y, \mathcal{S}_1) \geq c_2 d(y, \partial T^{-1}V_0) \geq c_3|T^{-1}V_0|$$

Accordingly to (3.1), there exists $c > 0$ such that

$$c|V_0| \leq |T^{-1}V_0|^{1-\beta}$$

which implies that for some $C > 0$,

$$|T^{-1}V_0| \geq (c|V_0|)^{1/(1-\beta)} \geq C\Lambda^{\frac{n}{\beta-1}}.$$

$\square$

It follows from Corollary 1 that there exists $\hat{n}_0 = \hat{n}_0(V)$ such that for any $m > \hat{n}_0$, $T^m(V, \nu_V)$ will be a proper standard family again, where ν_V is the conditional measure of a regular measure ν_W on V. Accordingly, we define recovery time function $\mathbf{r_p}(x) = \hat{n}_0(V) + n_0$ on V. Clearly, the recovery function is constant on each gap and for any $n \geq \mathbf{r_P}$, $T^n(V, \nu_V)$ has at least d_0 fraction properly return to the magnet $\mathcal{R}^*$. Next we estimate the tail bound of the recovery time function.

Lemma 17. *For any unstable manifold W with $|W| \geq 4\tilde{\delta}$ that fully u-crosses $\mathcal{R}^*$, for all $n \geq 1$,*

$$\nu_W\big(x \in W \setminus W_\kappa \,|\, \mathbf{r_p}(x) > n\big) \leq \text{ const}\,\Lambda^{-qn}. \tag{7.5}$$

Proof. For any x in the interior of a gap V with rank n, there exists $m \geq n$ such that $r_m(x) < \tilde{c}\Lambda^{-m}\kappa$. It follows that

$$V \subset \cup_{m \geq n}\{r_m(x) < \tilde{c}\Lambda^{-m}\kappa\}.$$

Since (W, ν_W) is proper, by (5.8), for any $m \geq 0$, $\nu_W(r_m(x) < \varepsilon) \leq c\varepsilon^q$. This implies that there exists $C > 0$ such that $\nu_W(V) \leq C\Lambda^{-qn}$. Summing over all V with stopping time greater than n, we get

$$\nu_W\big(\mathbf{r_p} > n\big) \leq \text{ const}\,\Lambda^{-qn}|W|.$$

Thus for any $n > 1$,

$$\nu_W\big(\mathbf{r_p} > n\big) \leq \text{ const}\,\Lambda^{-qn}. \tag{7.6}$$

This completes our proof of Coupling Lemma. In addition, we can now construct Young's tower by using the magnet $\mathcal{R}^*$ as its base. The Markovness of the returns would be guaranteed by the formula (6.1), in the same way as it was done in [8, 32]. The exponential tail bound follows by exactly the same argument as the proof of Coupling Lemma.

8 Proof of the main theorem

Notice that the rate of decay of correlations is actually the speed of convergence of a random distribution to the equilibrium state (more precisely, SRB measure). Accordingly, it is enough to study the rate of convergence in $|T_*^n \nu - \mu|$, where ν is absolutely continuous with respect to the SRB measure μ. Let $\mathcal{G}$ be a proper family, then any observable g on $\mathcal{G}$ can be naturally extended to the region $\hat{\mathcal{G}}$ as $g(x,t) = g(x)$. For brevity, we use the notation $\langle g \rangle = \int_M g\, d\mu$. The following theorem follows from the Coupling Lemma by a similar argument as in [9, 10] and [12], pages 175–177.

Theorem 2 (Equidistribution)**.** *Let $\mathcal{G}$ be a proper standard family. For any dynamically Hölder continuous function $f \in \mathcal{H}^-_{\vartheta_f} \cap L_\infty(M,\mu)$ and $n \geq 0$*

$$\left| \int_M f \circ T^n \, d\nu_{\mathcal{G}} - \int_M f \, d\mu \right| \leq B_f \theta_f^n \tag{8.1}$$

where $B_f = 2C_\Upsilon \big(K_f + \|f\|_\infty\big)$ and $\theta_f = \big[\max\{\vartheta_\Upsilon, \vartheta_f\}\big]^{1/2} < 1$.

Theorem 3 (Exponential decay of correlations)**.** *For every pair of dynamically Hölder continuous functions $f \in \mathcal{H}^-_{\vartheta_f} \cap L_\infty(M,\mu)$, $g \in \mathcal{H}^+_{\vartheta_g} \cap L_\infty(M,\mu)$ and $n \geq 0$*

$$\big|\langle g \cdot (f \circ T^n)\rangle - \langle f \rangle \langle g \rangle\big| \leq B_{f,g}\, \theta_{f,g}^n \tag{8.2}$$

where

$$\theta_{f,g} = \big[\max\big\{\vartheta_\Upsilon, \vartheta_f, \vartheta_g, e^{-p/\varkappa}\big\}\big]^{1/4} < 1, \tag{8.3}$$

where

$$B_{f,g} = C_0\big(K_f\|g\|_\infty + K_g\|f\|_\infty + \|f\|_\infty\|g\|_\infty\big), \tag{8.4}$$

and $C_0 = C_0(\mathcal{D}) > 0$ is a constant.

The above results can be extended to variables made at multiple times. Let $f_0, f_1, \dots, f_k \in \mathcal{H}^-_{\vartheta_f, C_f}$, and $\|f_i\|_\infty = \|f\|_\infty$, $i = 1, \dots, k$. Consider the product $\tilde{f} = f_0 \cdot (f_1 \circ T) \cdots (f_k \circ T^k)$.

Lemma 18. *Let $\mathcal{G}$ be a proper family in $\mathfrak{F}_p^*$. Then there exists $B_{\tilde{f}} > 0$ such that for any $n \geq 0$,*

$$\Big| \int_M \tilde{f} \circ T^n d\nu_{\mathcal{G}} - \int_M \tilde{f} d\mu \Big| \leq B_{\tilde{f}} \vartheta_f^n. \tag{8.5}$$

Furthermore, let $g_0, g_1, ..., g_k \in \mathcal{H}^+_{\vartheta_g}$, and $\|g_i\|_\infty = \|g\|_\infty$, $i = 1, ..., k$. Consider the product $\tilde{g} = g_0 \cdot (g_1 \circ T) \cdots (g_k \circ T^k)$. Then we can estimate the correlations between observables $\tilde{f}$ and $\tilde{g}$ as we did in Theorem 3

Theorem 4. *There exists $B_{\tilde{f},\tilde{g}} > 0$, for all $n \geq 0$,*

$$\left|\langle \tilde{g} \cdot (\tilde{f} \circ T^n) \rangle - \langle \tilde{f} \rangle \cdot \langle \tilde{g} \rangle \right| \leq B_{\tilde{f},\tilde{g}} \theta^n_{f,g},$$

where $\theta_{f,g}$ is the same as in (8.2).

9 Applications to billiards

To illustrate our method, we apply it to two classes of billiards for which previous methods failed. Since these examples play a secondary role, we only sketch the arguments.

First we recall standard definitions, see [4, 5, 6, 8]. A 2D flat billiard is a dynamical system where a point moves freely at unit speed in a domain $Q \subset \mathbb{R}^2$ and bounces off its boundary ∂Q by the law of elastic reflection. We assume that $\partial Q = \cup_i \Gamma_i$ is a finite union of piecewise smooth curves, such that each smooth component $\Gamma_i \subset \partial Q$ is either convex inward (dispersing), or flat, or convex outward (focusing). Following Bunimovich, see [3] and [12, Chapter 8], we assume that every focusing component Γ_i is an arc of a circle such that there are no points of ∂Q on that circle or inside it, other than the arc Γ_i itself. Under these assumptions the billiard dynamics is hyperbolic, ergodic, and mixing.

Let $\mathcal{M} = \partial Q \times [-\pi/2, \pi/2]$ be the *collision space*, which is a standard cross-section of the billiard system. Canonical coordinates on $\mathcal{M}$ are r and φ, where r is the arc length parameter on ∂Q and $\varphi \in [-\pi/2, \pi/2]$ is the angle of reflection. The collision map $\mathcal{F} \colon \mathcal{M} \to \mathcal{M}$ takes an inward unit vector at ∂Q to the unit vector after the next collision, and preserves smooth measure $d\hat{\mu} = c \cdot \cos\varphi \, dr \, d\varphi$ on $\mathcal{M}$, here c is a normalization constant. Furthermore, $\partial\mathcal{M} \cup \mathcal{F}^{-1}(\partial\mathcal{M})$ is the singular set of $\mathcal{F}$.

For billiards with focusing boundary components, the expansion and contraction (per collision) may be weak during long series of successive reflections along certain trajectories. To study the mixing rates, one needs to find and remove the spots in the phase space where expansion (contraction) slows down. Such spots come in several types and are easy to identify, for example, see [13] and [12, Chapter 8]. Traditionally, we denote

$$\partial Q = \partial^0 Q \cup \partial^{\pm} Q,$$

where ∂Q^0 is the union of flat boundaries, ∂Q^- contains focusing boundaries and ∂Q^+ corresponds to dispersing boundaries. The collision space can be naturally divided into focussing, dispersing and neutral parts:

$$\mathcal{M}_0 = \{(r,\varphi) \,:\, r \in \partial^0 Q\}, \qquad \mathcal{M}_\pm = \{(r,\varphi) \,:\, r \in \partial^\pm Q\}.$$

Let

$$\bar{M} = \{x \in \mathcal{M}_- :\ \pi(x) \in \Gamma_i,\, \pi(\mathcal{F}^{-1}x) \in \Gamma_j,\, j \neq i\} \cup \mathcal{M}_+, \tag{9.1}$$

Note that $\bar{M}$ only contains the *first* collisions with the focusing arcs (the collisions with the straight lines are skipped altogether) and all collisions on dispersing boundaries. The map F preserves the measure μ conditioned on $\bar{M}$, which we denote by $\mu = [\hat{\mu}(\bar{M})]^{-1}\hat{\mu}$. Furthermore, F has uniform expansion and contraction, since we skipped all collisions too close to the "bad spots" in the collision space. But it has a larger singular set than the original map. Let $\mathcal{S}_0 = \partial \bar{M}$, then $\mathcal{S} = \mathcal{S}_0 \cup F^{-1}\mathcal{S}_0$ is the singular set of F.

Now we turn to two specific classes. The first is a non-smooth stadium. It is a convex domain Q bounded by two parallel straight segments and two minor arcs (i.e., arcs smaller than a semicircle) with radius $\mathbf{r}_3$ and $\mathbf{r}_4$. Let Q satisfy the standard Bunimovich assumptions [3], i.e. the complement of each arc in ∂Q to a full circle crosses both straight segments in ∂Q, but does not cross the other arc; see Fig. 1 (left). As demonstrated in [13], the reduced map F fails condition (1.3).

Theorem 5. *For this type of the stadia, the correlations for the reduced billiard map* $F\colon \bar{M} \to \bar{M}$ *decay exponentially.*

The detailed analysis of Bunimovich stadia was presented in [3, 5] and [12, Chapter 8], we only focus on the novel parts here, see Fig. 1 (left).

The space $\bar{M}$ of the reduced map F is shown in Fig. 1 (right). The map F has two types of infinite sequences of singularity curves, as shown on Fig. 1. The first type accumulates near the top and bottom vertices x_1, x_2 of the parallelograms, they are generated by trajectories nearly "sliding" along the circular arcs. Denote these curves as $\mathbf{S}_1 := \{S_{k,n} \,:\, k = 1, 2, n \in \mathbb{N}\}$. And let

$\mathcal{S}_0 = \partial\bar{M} \cup \mathbf{S}_1$, $\mathcal{S}_{\pm 1} = F^{\mp 1}\mathcal{S}_0$. Denote $M = \bar{M}\backslash\mathcal{S}_0$. Then $F : M\backslash\mathcal{S}_1 \to M\backslash\mathcal{S}_{-1}$ is a C^2 diffeomorphism, which follows from the smoothness of $\mathcal{F}$ and $\mathcal{S}_0$.

The second type singular curves accumulate near the other two vertices x_3, x_4 on the line $\varphi = \pm\varphi_0$, they are generated by trajectories experiencing many bounces off the two straight sides of the stadium. Note that $\mathcal{S}$ consists of countably many curves $\{S_{n,k}, n = 0, 1, 2, ...\}$, $k = 1, 2, 3, 4$, accumulating near singular points $x_k \in M$. For $n \in \mathbb{N}$, denote $M_{n,k}$ as the connected region bounded by the adjacent curves $S_{n,k}, S_{n-1,k}$ in $M \setminus \mathcal{S}_1$. $M_{n,k}$ is called a n-cell. Points in $M_{n,k}$, $k = 3, 4$ experience exactly n reflections off the straight sides before landing on the opposite arc of ∂Q. By geometric calculation as explained in [15], one of the long boundaries of $M_{n,3}$ can be approximated by the line segment

$$r = (\varphi - \varphi_0)\mathbf{r}_3 + \frac{l}{2n\cos\varphi_3},$$

where l is the length of the flat boundary segment and φ_k is the angle of x_k. Furthermore, the slanted line through x_3 has equation $r = 2\mathbf{r}_3(\varphi - \varphi_3)$. Geometric structure in the vicinity of x_4 is similar.

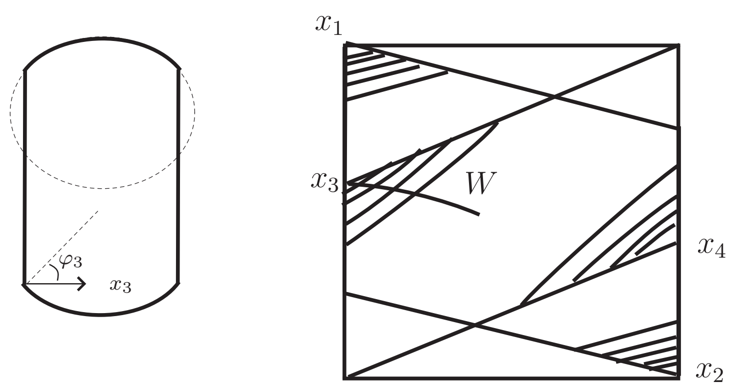

Figure 2: Discontinuity manifolds of stadium

It is also shown in [13, 12] that the expansion factor in $M_{n,k}$ satisfies

$$\mathcal{J}(x) \geq \begin{cases} c_1 n^{\frac{3}{2}}, & \text{if } k = 1, 2; \\ c_1 n, & \text{if } k = 3, 4, \end{cases} \tag{9.2}$$

where $c_1 > 0$ is a constant. We note that for $k = 1, 2$, the reduced map has an expansion n during the sliding process and another expansion at least of order $\sqrt{n}$ before the trajectory landing on the opposite arc as explained in Exercise 8.36 in [12].

To prove the exponential decay of correlations for F, we need to verify (**H.3**) and one-step expansion condition. For $j = 1, 2$, the expansion factors are strong enough so that the series of their reciprocals converges, $\sum n^{-3/2} < \infty$. Thus they satisfy the 'old' one-step expansion condition (1.3), so they can be easily handled as in [13]. For $k = 3, 4$, the series of the reciprocals of the expansion factors diverges, $\sum n^{-1} = \infty$, and this is where our new one-step expansion condition (3.8) is used.

For a typical unstable manifold W, we need to estimate the expansion factor on $W \cap M_{n,k}$. It is enough to consider $k = 3$. Assume $x \in W_n := W \cap M_{n,k}$, then the free path of x is $\tau_n(x) = 2n\mathbf{r}_3 \cos\varphi_3$. Although $\mathbf{r}_3$ might be different from $\mathbf{r}_4$, it follows from the existence of two parallel boundaries that

$$\mathbf{r}_3 \cos\varphi_3 = \mathbf{r}_4 \cos\varphi_4.$$

Let $x = (r_1, \varphi_1) \in W_n$, and $Fx = (r_1, \varphi_1) \in V_n := FW_n$. It is easy to verify, see [15]

$$\Lambda(x) \geq \frac{2n\cos\varphi_0}{\cos\varphi_2} \geq n$$

A direct geometric calculation shows that $M_{n,k}$ has dimensions $l(2n^2\mathbf{r}_k \cos\varphi_k)^{-1}$ in the stable direction and $(n\mathbf{r}_k \cos\varphi_k)^{-1} l\sqrt{\mathbf{r}_k^2 + 1}$ in the unstable direction. Accordingly, for any unstable curve W that crosses finitely or infinitely many of $M_{n,3}$ and get cut into pieces W_n with length $l(2n^2\mathbf{r}_3 \cos\varphi_3)^{-1}$, the minimal expansion factor is n Thus,

$$\sum_{n=m_0}^{\infty} \Big(\frac{|W|}{|W_n|}\Big)^q \frac{|W_n|}{|W|} = \sum_{n=m_0}^{\infty} \Big(\frac{|W_n|}{|V_n|}\Big)^q \Big(\frac{|W_n|}{|W|}\Big)^{1-q}$$
$$\leq \sum_{n=m_0}^{\infty} \Lambda_n^{-q} (\frac{|W_n|}{|W|})^{1-q} \leq m_0^{1-q} \sum_{n=m_0}^{\infty} n^{-2+q} < \frac{1}{1-q}$$

(Here we used the fact that $|W| = \sum_{n=m_0}^{\infty} |W_n|$). Although the right hand side above is uniformly bounded, it is not < 1 as required by (3.8). To resolve the problem, we can select an equivalent norm or consider high iterations of F as explained in [13] to get (3.8). Thus by Theorem 1, the return map $F\colon M \to M$ has an exponential mixing rate.

Our second class of billiards is made by Bunimovich tables [3, 12] whose focusing boundaries contain major arcs (i.e. arcs greater than a semicircle).

Such arcs add a new type of 'bad spots' where the hyperbolicity is weak, see [13] and [12, Chapter 8]. For simplicity, we assume that the major arcs are less than 240°, to prevent even further types of bad spots. Also we assume that the boundary components are either focusing or dispersing, and they intersect each other transversally (do not make cusps).

Theorem 6. *For the above Bunimovich-type billiard tables with major arcs, the correlations for the reduced billiard map $F : \bar{M} \to \bar{M}$ decay exponentially.*

In this class of billiards, the trouble is caused by long series of collisions occurring at one focusing arc where $|\varphi|$ is near 0. In this case the trajectory is close to a periodic orbit running along a diameter of the corresponding arc; we call such series *diametrical* (they can occur only on a focusing arc larger than half-circle).

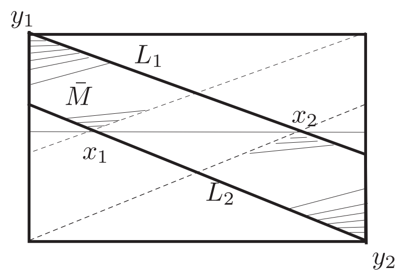

Figure 3: Singularities in the collision space of billiards with major arcs

The space $\bar{M}$ of the reduced map and the structure of singularities is shown on Figure 3. The singularity set $\mathcal{S} \cap \bar{M}$ of the map F consists of two types of infinite sequences of curves. There are four infinite sequences of almost parallel straight segments running between the nearby sides of $\bar{M}$ and converging to x_1, x_2, y_1, y_2, see Fig. 3. Curves of the first type accumulate near the top and bottom vertices y_1, y_2, they are generated by trajectories nearly "sliding" along the circular arcs. Curves of the second type accumulate near two points x_1 and x_2 (where $\bar{M}$ intersects the line $\varphi = 0$). These singularities are generated by trajectories experiencing arbitrary many collisions with the large arc while running almost along its diameter. We only concentrate on singularities near x_1, and denote by $\{S_n\}$ the sequence of singularity curves converge to x_1. Notice that the point x_1 has a trajectory of period 2. Thus points $x \in \bar{M}$ have trajectories running near a diameter of Γ_1, they hit Γ_1 on the opposite side and then come back, so that the points x, $\mathcal{F}^2(x)$, $\mathcal{F}^4(x), \dots$ are close to each other. Then the two sequences $\{\mathcal{F}^{2m}(x)\}$ and $\{\mathcal{F}^{2m+1}(x)\}$ move slowly along the arc Γ_1 until one of them finds an opening in Γ_1 and escapes. So we let M_m denote the m-cell of x_1 bounded by S_m, S_{m+1}, $\partial\hat{M}$ and $\partial F(\hat{M})$. It is easy to show the two slanted boundaries L_1, L_2 of the parallelogram $\hat{M}$ have equations $r = (\pi + \omega - 2\varphi)\mathbf{r}$ and $r = (\omega - 2\varphi)\mathbf{r}$, S_m has equation $r = (\omega - 4m\varphi)\mathbf{r}$. Accordingly, for any unstable curve W that crosses finitely many or infinitely many of M_n's, the minimal expansion factor on $W \cap M_n$ satisfies is $\sim 4n\mathbf{r}$. Thus for any $q \in (0,1)$,

$$
\begin{aligned}
&\sum_{n=m_0}^{\infty} \left(\frac{|W|}{|F(W \cap M_{n,k})|}\right)^q \frac{|W \cap M_{n,k}|}{|W|} \\
&\leq \sum_{n=m_0}^{\infty} \Lambda_n^{-q} \left(\frac{|W \cap M_{n,k}|}{|W|}\right)^{1-q} < \frac{m_0^{1-q}}{(4\mathbf{r})^q} \sum_{n=m_0}^{\infty} n^{-2+q} < \frac{1}{(1-q)(4\mathbf{r})^q}.
\end{aligned}
$$

We use the fact that $|W|$ is $(4m_0\mathbf{r})^{-1}$ and $|W \cap M_{n,k}|$ is of order $(4n^2\mathbf{r})^{-1}$ in the above inequality. Although the right hand side is uniformly bounded, but maybe not less than one. In this case, we can pick an equivalent norm or consider high iterations of F as explained in [13] to get (3.8). Thus by Theorem 1, the return map $F\colon \bar{M} \to \bar{M}$ has exponential mixing rate.

References

[1] Alves, A., Bonatti, C., and Viana M., *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), 351–398.

[2] Bálint, P., and Toth, I. P., *Exponential decay of correlations in multidimensional dispersing billiards*, preprint.

[3] Bunimovich, L. A., *On the ergodic properties of nowhere dispersing billiards*, Comm. Math. Phys., **65** (1979), 295–312.

[4] Bunimovich, L. A., Sinai, Ya. G. and Chernov, N., *Markov partitions for two-dimensional hyperbolic billiards,* Russian Math. Surveys, **45** (1990) 105–152.

[5] Bunimovich L. A., Sinai, Ya. G. and Chernov, N., *Statistical properties of two-dimensional hyperbolic billiards,* Russian Math. Surveys, **46** (1991) 47–106.

[6] Chernov, N., *Entropy, Lyapunov exponents and mean-free path for billiards*, J. Statist. Phys., **88** (1997), 1–29.

[7] Chernov, N., *Statistical properties of piecewise smooth hyperbolic systems in high dimensions*, Discr. Cont. Dynam. Syst. **5** (1999), 425–448.

[8] Chernov, N., *Decay of correlations in dispersing billiards*, J. Statist. Phys. **94** (1999), 513–556.

[9] Chernov, N., *Advanced statistical properties of dispersing billiards,*J. Statist. Phys. **122** (2006), 1061–1094.

[10] Chernov, N. and Dolgopyat, D., *Brownian Motion - I*, Memoirs of AMS, to appear.

[11] Chernov, N. and Dolgopyat, D., *Hyperbolic billiards and statistical physics*, Proc. ICM (Madrid, Spain, 2006), Vol. II, Euro. Math. Soc., Zurich, 2006, pp. 1679-1704.

[12] Chernov, N. and Markarian, R., *Chaotic Billiards*, Mathematical Surveys and Monographs, **127**, AMS, Providence, RI, 2006.

[13] Chernov, N. and Zhang, H.-K., *Billiards with polynomial mixing rates.* Nonlineartity, **4** (2005), 1527–1553

[14] Chernov, N. and Zhang, H-K., *Improved estimates for correlations in billiards*, Communications in Mathematical Physics, **277** (2008), 305–321.

[15] Chernov, N. and Zhang, H-K., *Regularity of Bunimovich Stadia.* Regular and Chaotic Dynamics, **3** (2007), 335–356.

[16] Demers, M. and Liverani, C., *Stability of statistical properties in two-dimensional piecewise hyperbolic maps*, Transactions of the AMS, 360 (2008), 4777-4814.

[17] Dolgopyat D., *Limit Theorems for partially hyperbolic systems*, Transactions of the AMS, **356** (2004), 1637–1689.

[18] V. Donnay, *Using integrability to produce chaos: billiards with positive entropy*, Comm. Math. Phys. **141** (1991), 225–257.

[19] Katok, A., and Strelcyn, J.-M. (with the collaboration of F. Ledrappier & F. Przytycki), *Invariant manifolds, entropy and billiards; smooth maps with singularities*, Lect. Notes Math., **1222**, Springer, New York (1986).

[20] R. Markarian, *Billiards with Pesin region of measure one*, Comm. Math. Phys. **118** (1988), 87–97.

[21] Markarian, R., *Billiards with polynomial decay of correlations*, Erg. Th. Dynam. Syst. **24** (2004), 177–197.

[22] Ya. Pensin *Dynamical Systems With Generalized Hyperbolic Attractors: Hyperbolic, Ergodic and Topological Properties*, Ergod. Theory and Dyn. Syst., **12** (1992), 123–152.

[23] Pesin Ya. B and Sinai Ya. G. *Hyperbolicity and Stochasticity of Dynamical Systems*, Soviet Scientific Reviews Section C, Mathematical Physics Review **4** (1981), 53–115

[24] E. Sataev, *Invariant measures for hyperbolic maps with singularities*, Russ. Math. Surv. **47**, 191–251

[25] N. Simányi, *Ergodicity of hard spheres in a box*, Ergod. Th. Dynam. Syst. **19** (1999), 741–766.

[26] N. Simányi, *The complete hyperbolicity of cylindric billiards*, Ergod. Th. Dyn. Sys. **22** (2002), 281–302.

[27] N. Simányi, *Proof of the Boltzmann-Sinai ergodic hypothesis for typical hard disk systems*, Invent. Math. **154** (2003), 123–178.

[28] N. Simányi, *Proof of the ergodic hypothesis for typical hard ball systems*, Ann. H. Poincaré **5** (2004), 203–233.

[29] Ya. G. Sinai and N. I. Chernov, *Ergodic properties of some systems of two-dimensional discs and three-dimensional spheres* , Russian Math. Surveys 42 (1987), 181–207.

[30] Sinai, Ya. G., *Dynamical systems with elastic reflections*. Ergodic properties of diepersing billiards, Russian Math. Surveys, **25** (1970), 137–189.

[31] M. Wojtkowski, *Principles for the design of billiards with nonvanishing Lyapunov exponents*, Comm. Math. Phys. **105** (1986), 391–414.

[32] Young, L.-S., *Statistical properties of systems with some hyperbolicity including certain billiards*. Ann. Math., **147** (1998) 585–650.

[33] Young, L.-S., *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.