ON DYNAMICS OF MOSTLY CONTRACTING DIFFEOMORPHISMS.

DMITRY DOLGOPYAT

ABSTRACT. Mostly contracting diffeomorphisms are the simplest examples of robustly nonuniformly hyperbolic systems. This paper studies the mixing properties of mostly contracting diffeomorphisms.

1. INTRODUCTION.

This paper treats a class of partially hyperbolic systems with non-zero Lyapunov exponents.

Before stating our result let us recall some recent work motivating our research. In recent years there were several advances in understanding of statistical properties of weakly hyperbolic dynamical systems.

On one hand L.–S. Young developed quantitative Pesin theory in [38, 39]. Among other things she proved that if a diffeomorphism f has a Pesin set Λ such that the distribution of the return time to Λ has an exponentially decaying tail and if f has no discrete spectrum then it is exponentially mixing. This theory was applied to a number of examples in the above mentioned papers as well as in [3, 10].

On the other hand M. Grayson, C. Pugh and M. Shub showed ([13, 27, 28]) that partial hyperbolicity can give raise to a good ergodic behavior in a robust way. Further examples of systems satisfying their criteria can be found in [7, 8, 16, 20, 37].

These results lead to the natural question if there is an open set of (partially hyperbolic) systems satisfying the conditions of Young's theory with uniform bounds. This question was addressed in a number of papers [1, 2, 4, 32, 36]. Our paper also fits into this framework.

Let us give a few definitions. Let f be a diffeomorphism of a smooth manifold X and let ν be an ergodic f–invariant measure. We call ν an *SRB–measure* for f if there is a subset $Y(\nu) \subset X$ of positive Lebesgue measure such that for almost all $y \in Y$ for any continuous function A

$$\frac{1}{n} \sum_{j=0}^{n-1} A(f^j x) \to \nu(A).$$

$Y(\nu)$ is called *basin of attraction* of ν. Certainly the question of existence of SRB–measures and their dependence on

parameters is quite important in smooth ergodic theory. We say that f has a global attractor if there is only one SRB–measure whose basin is all of X. (Our use of the word *attractor* follows that of [35, 27]. More precisely, ν or supp(ν) should be called a stochastic attractor because it describes the statistical properties of large iterations of f. For a more topological approach see [19].) Let $\mathcal{S}$ be a subset of Diff$_r(X)$, $r > 1$ endowed with some topology (think of $\mathcal{S}$ as a parameter space) and let $f \in \mathcal{S}$. We call f *statistically stable in $\mathcal{S}$* if any diffeomorphism g in some neighborhood of f in $\mathcal{S}$ has a finite number of SRB–measures $\nu_1(g), \nu_2(g) \ldots \nu_k(g)$, the maps $g \to \nu_j(g)$ are continuous and the union of basins of $\nu_j(g)$ is all of X. If $k = 1$ we call f *strongly statistically stable.* Below we deal with the case when $\mathcal{S} = \text{Diff}_2(X)$ with uniform C^2–topology.

In this note we provide some sufficient conditions for statistical stability as well as for other good statistical properties. Our main results are the following.

Theorem I. *Let f be partially hyperbolic dynamically coherent u-convergent mostly contracting diffeomorphism of a three-dimensional manifold X. Then*

(a) f has a global attractor ν;

(b) for any $\gamma > 0$ there are constants $C, \zeta < 1$ such that if $A, B \in C^\gamma(X)$ then for positive n

$$\left| \int B(x)A(f^n x)dx - \nu(A) \int B(x)dx \right| \leq C\zeta^n ||A||_\gamma ||B||_\gamma,$$

$$\left| \int B(x)A(f^n x)d\nu(x) - \nu(A)\nu(B) \right| \leq C\zeta^n ||A||_\gamma ||B||_\gamma;$$

(c) f has non-zero Lyapunov exponents.

Remark. Parts (a) and (c) of this theorem were established before in [4] for a larger class of systems. It follows from (c) and the results of [11, 21] that the system (X, f, ν) is a Bernoulli shift.

Remark. In fact, we prove more than (a). Namely we show that the image under f^n of any unstable leaf becomes equidistributed. In [12] we proved that diffeomorphisms having this property satisfy many classical limit theorems of probability theory.

Theorem II. *Let f be as in Theorem I. If in addition f is stably dynamically coherent then f is strongly statistically stable. More precisely, there exists a neighborhood $\mathcal{O}(f) \subset \text{Diff}_2(X)$ such that any $g \in \mathcal{O}$ satisfies the conditions of Theorem I and the constants C, ζ in Theorem I(b) can be chosen uniformly in $\mathcal{O}(f)$. In particular if*

ν_g *is the SRB measure for* g *then for any* $\gamma > 0$ *the map* $g \to \nu_g$
$\mathcal{O}(f) \to (C^\gamma(X))^*$ *is Holder continuous.*

See Sections 2, 3 for the definition of the terms appearing in the formulation of this theorem. In Section 4 we show that for mostly contracting diffeomorphisms the second forward Lyapunov exponent of almost every point is negative and prove large deviation estimates for the exceptional set. This is done by certain submartingale estimates more common in the theory of stochastic differential equations. In Section 5 we recall the construction of u-Gibbs measures [26] and show that in our situation they are SRB measures. The uniqueness of SRB measures is treated in Section 6–9. In Section 6 we recall the coupling method of L.–S. Young. In Section 7 we describe coupling algorithm for our system. The properties of this algorithm are studied in Sections 8 and 9. The proofs of the main theorems are completed in Section 10. In Sections 11 and 12 we discuss some examples. Final remarks and some open questions are presented in Section 13.

Remark. Independently and slightly earlier A. Castro [9] proved a result similar to our Theorem I. However, because of some technical assumptions in his paper it is not clear if his result can be applied to the examples of Section 12.

2. Partial hyperbolicity.

In this and the next sections we describe the properties of f which appear in the statement of Theorem I. As it was mentioned in the introduction f is a diffeomorphism of X^3. We also assume that f is partially hyperbolic and stably dynamically coherent. Thus the tangent bundle of X is the sum of three continuous one dimensional subbundles E_u, E_c and E_s such that

$$e^{\lambda_1} \le (df|E_s) \le e^{\lambda_2}, \tag{1}$$

$$e^{\lambda_3} \le (df|E_c) \le e^{\lambda_4}, \tag{2}$$

$$e^{\lambda_5} \le (df|E_u) \le e^{\lambda_6} \tag{3}$$

where $\lambda_1 \le \lambda_2 < \lambda_3 \le \lambda_4 < \lambda_5 \le \lambda_6$ and $\lambda_2 < 0$, $\lambda_5 > 0$.

E_u and E_s are always integrable so they are tangent to f invariant foliations: unstable (W^u) and stable (W^s). *Dynamical coherence* means that E_c, $E_c \oplus E_u$ and $E_c \oplus E_s$ are also are tangent to f–invariant foliations which are called central (W^c), center–unstable (W^{cu}) and center–stable (W^{cs}) respectively, and that W^c subfoliates both W^{cu} and W^{cs}. (In fact only unique integrability of $E_c \oplus E_s$ is used in the paper.) Stable dynamical coherence means that any g close to f is also dynamically coherent. The openness of these conditions was studied in [14]. Namely partial hyperbolicity is open. It is unknown if dynamical coherence is open but if the center foliation of f is C^1 then f is stably dynamically coherent.

Let $\mathcal{V}$ be the set of all unstable curves of lengths between 1 and 2. $\mathcal{V}$ is *a Markov family* in the sense that $\forall V \in \mathcal{V}$ there is a finite set $\{\tilde{V}_j\}$ of elements of $\mathcal{V}$ such that $fV = \bigcup_j \tilde{V}_j$. More generally for any unstable curve U of length greater than one there is a finite set $\{V_j\}$ such that $V_j \in \mathcal{V}$ and

$$U = \bigcup_j V_j. \tag{4}$$

We call (4) *Markov decomposition* of U. We will use (4) for $U = f^n V$ where $V \in \mathcal{V}$, $n > 0$.

We call f *u-convergent* if $\forall \varepsilon \ \exists n > 0 \ \forall V_1, V_2 \in \mathcal{V} \ \exists x_j \in V_j$ such that $d(f^n x_1, f^n x_2) \le \varepsilon$. Later on we show that u-convergence is open among mostly contracting diffeomorphisms.

3. Mostly contracting systems.

The assumptions of Section 2 are routine partial hyperbolicity assumptions. The next property guarantees that f is non-uniformly hyperbolic. In order to formulate it we need to recall the definition of canonical densities on W^u. We would like to study SRB–measure for f. *A priori* we do not know if it exists, but if it does then its conditional densities on W^u are given by [26]. Fix a Riemann structure on X. It induces a metric on W^u–fibers. Let V be an interval inside W^u. For $z_1, z_2 \in V$ let

$$\rho(z_1, z_2) = \prod_{j=0}^{\infty} \frac{(df^{-1}|E_u)(f^{-j} z_1)}{(df^{-1}|E_u)(f^{-j} z_2)}. \tag{5}$$

Fix some $z_0 \in V$ and let $\rho_V(z) = C\rho(z, z_0)$ where $C = (\int_V \rho(z, z_0) dz)^{-1}$. Since $\rho(z, z_0') = \rho(z, z_0)\rho(z_0, z_0')$ this definition actually does not depend on z_0. Also if $W = fV$, $y = fz$ then

$$\rho_V(z)dz = C_V \rho(z, z_0)dz = C_V \rho(z, z_0)(df^{-1}|E_u)(y)dy =$$

$$\tilde{C}_V \rho(z, z_0)\frac{(df^{-1}|E_u)(y)}{(df^{-1}|E_u)(y_0)}dy = C_W \rho(y, y_0)dy.$$

Thus if A is continuous then

$$\int_V A(fz)\rho_V(z)dz = \int_{fV} A(y)\rho_{fV}(y)dy.$$

Our last assumption is the following. There is a positive constant α_0 such that for any $V \in \mathcal{V}$

$$\int_V \rho_V(x)\left(\ln(df|E_c)(x)\right)dx \leq -\alpha_0 < 0 \tag{6}$$

We call f *mostly contracting* if some positive power of it satisfies (6). In the proof of the Theorem I we assume, as we may, that f itself satisfies (6).

Remark. Condition (6) is C^2–open. Indeed by standard theory ([14]) the map CENTER$(f) = (df|E_c)(x)$ is continuous: $\mathrm{Diff}_2(X) \to C(X)$. Let $V(x, f, t)$ denote f–unstable curve of length t centered at x. Let $V(x, f, t)(\tau)$ be the arclength parameterization of τ. Denote $\mathbf{T} = \{(t, \tau) : 1 \leq t \leq 2, 0 \leq \tau \leq t\}$. Then the map DENSITY$(f) = \rho_{V(x,f,t)}(V(x, f, t)(\tau))$ is continuous: $\mathrm{Diff}_2(X) \to C(X \times \mathbf{T})$ since ρ is a uniform limit of continuous functions. CENTER is also continuous in C^1-topology, but DENSITY is not because convergence in (5) may fail to be uniform in f (cf. [23]). Thus it is unclear if mostly contractiveness is C^1–open.

4. Large deviations.

Here we prove

Theorem 4.1. $\exists C_1, s > 0, \theta_1 < 1$ *such that* $\forall V \in \mathcal{V}$ $\forall n > 0$

$$\int_V \rho_V(x)\left(df^n|E_c\right)^s(x)dx \leq C_1 \theta_1^n. \tag{7}$$

The proof consists of a number of lemmas.

Lemma 4.2.

$$\forall n > 0 \quad \int_V \rho_V(x)\ln(df^n|E_c)(x)dx \leq -n\alpha_0$$

Proof. We have

$$\int_V \rho_V(x) \ln(df^n|E_c)(x)dx =$$

$$\int_V \rho_V(x) \ln(df|E_c)(x)dx + \int_{fV} \rho_{fV}(y) \ln(df^{n-1}|E_c)(y)dy.$$

Let $fV = \bigcup_j V_j$ be an almost Markov decomposition. Then the second term equals

$$\sum_j c_j \int_{V_j} \rho_{V_j}(y) \ln(df^{n-1}|E_c)(y)dy$$

where $c_j = \int_{f^{-1}V_j} \rho_V(x)dx$. By induction

$$\int_{V_j} \rho_{V_j}(y) \ln(df^{n-1}|E_c)(y)dy \le -(n-1)\alpha_0.$$

Summation over j proves the lemma. ∎

The following distortion bound is standard (see, for example [4], Lemma 3.3).

Proposition 4.3. *There is a constant C so that $\forall n > 0$ $\forall V \in \mathcal{V}$ $\forall x_1, x_2 \in f^{-n}V$*

$$|\ln(df^n|E_c)(x_1) - \ln(df^n|E_c)(x_2)| \le C.$$

If A is a continuous function and U is a piece of unstable manifold we write $||A||_U = \max_{x \in U} |A(x)|$.

Corollary 4.4. *There exists α_1 such that if n is large enough then for any $V \in \mathcal{V}$ for any Markov decomposition $f^n V = \bigcup_j V_j$ the following holds. Let $U_j = f^{-n}V_j$, $c_j = \int_{U_j} \rho_V(x)dx$. Then*

$$\sum_j c_j \ln ||(df^n|E_c)||_{U_j} \le -\alpha_1.$$

Changing if necessary $f \to f^n$ we can assume that this is true for $n = 1$. Under this assumption we have

Lemma 4.5. *If s is small enough there is a constant $\theta_2(\gamma) < 1$ such that under the conditions of the previous corollary*

$$\sum_j c_j ||df|E_c||^s_{U_j} \le \theta_2.$$

Proof. Regard the LHS as a function $r(s)$. Then $r(0) = 1$, $\frac{dr}{ds}(0) \le -\alpha_1$, $|\frac{d^2r}{ds^2}(0)| \le \text{Const.}$ ∎

Repeating the argument of Lemma 4.2 we obtain

Corollary 4.6. *For any $n > 0$ there is a Markov decomposition $f^n V = \bigcup_j V_j$ such that if $U_j = f^{-n} V_j$, $c_j = \int_{U_j} \rho_V(x) dx$ then*

$$\sum_j c_j \|df^n|E_c\|_{U_j}^s \le \theta_2^n. \tag{8}$$

Proof. By induction. Suppose that (8) is valid for all $n \le n_0 - 1$. Let $fV = \bigcup_j V_j$ be a Markov decomposition. By inductive assumption $\forall j$ there is a Markov decomposition $f^{n_0-1} V_j = \bigcup_k V_{jk}$ satisfying (8). Let $U_j = f^{-1} U_j$, $U_{jk} = f^{-n_0} V_{jk}$, $b_j = \int_{U_j} \rho_V(x) dx$, $c_{jk} = \int_{U_{jk}} \rho_{U_j}(x) dx$. Then $f^{n_0} V = \bigcup_{jk} V_{jk}$ is Markov and

$$\sum_{jk} \int_{f^{-n_0} V_{jk}} \rho_V(x) \|(df^{n_0}|E_c)\|_{U_{jk}}^s dx =$$

$$\sum_{jk} b_j c_{jk} \|(df^{n_0}|E_c)\|_{U_{jk}}^s \le$$

$$\sum_j b_j \|(df|E_c)\|_{U_j}^s \sum_k c_{jk} (\|df^{n_0-1}|E_c\|_{fU_{jk}})^s \le$$

$$\sum_j b_j \|(df|E_c)\|_{U_j}^s \theta_2^{n_0-1} \le$$

$$\theta_2^{n_0}. \qquad \blacksquare$$

This Corollary proves Theorem 4.1 since for any Markov decomposition $f^n V = \bigcup_j V_j$ LHS of (8) majorates LHS of (7).

5. Transfer operator.

Now we recall the general method of the construction of SRB–measures for partially hyperbolic systems ([26]). SRB measures are obtained as forward iteration limits of suitable measures. Here we describe the set of the initial measures. Fix some R. Let $E_1(R)$ be the set of measures of the form

$$l(A) = \int_V A(z) e^{G(z)} \rho_V(z) dz$$

where $V \in \mathcal{V}$, $l(1) = 1$ and $|G(z_1) - G(z_2)| \le R d^\gamma(z_1, z_2)$. (Those three conditions also guarantee that G is uniformly bounded.) Let $E_2(R)$ be the convex hall of $E_1(R)$ and $E(R)$ be the closure of $E_2(R)$. The family $\{E(R)\}$ is continuous from above in the sense that $E(R_0) = \bigcap_{R > R_0} E(R)$. (This follows from the fact that $E_1(R_0) = \bigcap_{R > R_0} E_1(R)$.)
    Let $\mathcal{T}(l)(A) = l(A \circ f)$.

Proposition 5.1. $\mathcal{T} : E(R) \to E(Re^{-\lambda_5 \gamma})$.

Proof. If $l \in E_1(R)$. Then

$$(\mathcal{T}l)(A) = \int_V e^{G(z)} A(fz)\rho_V(z)dz = \int_{fV} e^{(G \circ f^{-1})(y)} A(y)\rho_{fV}(y)dy.$$

Let $fV = \bigcup_j V_j$, be a Markov decomposition then $\mathcal{T}l = \sum c_j l_j$, where

$$l_j(A) = \int_{V_j} e^{(G \circ f^{-1})(y)} A(y)\rho_{V_j}(y)dy.$$

Also $|(G \circ f^{-1})(y_1) - (G \circ f^{-1})(y_2)| \leq Re^{-\lambda_5\gamma}d^\gamma(y_1, y_2)$. Thus $\mathcal{T}$: $E_1(R) \to E_2(Re^{-\lambda_5\gamma})$. $\blacksquare$

Since $E(R)$ is a convex compact set Proposition 5.1 implies that there is an f–invariant measure in $E(R)$ which moreover belongs to $\bigcap_{R>0} E(R) = E(0)$ (this is also proven in [26]).

Proposition 5.2. *Any f-invariant measure $\nu \in E(0)$ has two negative Lyapunov exponents.*

Proof. By (6) for any $l \in E(0)$ $l(\ln(df|E_c)) < -\alpha_0$. $\blacksquare$

Proposition 5.2 and Lemma 13 of [26] guarantee that ν satisfies the conditions of Theorem 3 of [27] and so it is a SRB measure. (Another proof of this fact is given in Section 10.)

6. Coupling argument.

We now pass to the uniqueness of ν. It is established via the coupling argument of [39]. We want to show that for large n for any $l_1, l_2 \in E(R)$ $\mathcal{T}^n(l_1)$ is close to $\mathcal{T}^n(l_2)$. First we consider the case when l_1 and l_2 are in $E_1(0)$, say $l_j(A) = \int_{V_j} A(x)\rho_{V_j}(x)dx$. The idea is to divide $f^n V_j$ into small pieces and pair the pieces of $f^n V_1$ to $f^n V_2$ so that the members of the pair are very close to each other. However since f^n gives different weights to different pieces of $f^n V_j$ it is more convenient to regard unstable curves as 1-chains so that the heavier ones can be split into several pieces each one being paired to a different partner.

Let us now give a formal statement. Denote $Y_j = V_j \times [0,1]$. Equip Y_j with the measure $dm_j(x,t) = \rho_{V_j}(x)dxdt$. The heart of the coupling method is the following result the proof of which occupies the next three sections.

Lemma 6.1. *There is a measure preserving map $\tau : Y_1 \to Y_2$, a function $R : Y_1 \to \mathbb{N}$ and constants $C_1, C_2 > 0, \rho_1 < 1, \rho_2 < 1$ such that*
    (A) If $(x_2, t_2) = \tau(x_1, t_1)$ then for $n \geq R(x_1, t_1)$

$$d(f^n x_1, f^n x_2) \leq C_1 \rho_1^{n-R}; \tag{9}$$

    (B) $m_1(R > N) \leq C_2 \rho_2^N$.

Let $||l||_\gamma$ denote the norm of l as the element of $(C^\gamma(X))^*$.

Corollary 6.2. $\exists C_3 > 0, \rho_3 < 1$ *such that* $\forall n > 0$ $\forall l_1, l_2 \in E(0)$ $||\mathcal{T}^n(l_1 - l_2)||_\gamma \le C_3\rho_3^n$.

Proof. It suffices to verify this for $l_j \in E_1(0)$. We have

$$(\mathcal{T}^n l_j)(A) = \int_{Y_j} A(f^n x_j)dm_j(x_j, t_j).$$

Let $(x_2, t_2) = \tau(x_1, t_1)$. Then

$$|\mathcal{T}^n(l_1 - l_2)(A)| \le \int_{Y_1} |A(f^n x_1) - A(f^n x_2)|dm_1(x_1, t_1).$$

Let $Z(n) = \{y : R(y) < \frac{n}{2}\}$ then

$$|\mathcal{T}^n(l_1 - l_2)(A)| \le$$

$$\int_{Z(n)} |A(f^n x_1) - A(f^n x_2)|dm_1(x_1, t_1) + 2||A||_0 m_1(Y_1 \backslash Z(n)) \le$$

$$||A||_\gamma \left((C_1\rho_1^{\frac{n}{2}})^\gamma + 2C_2\rho_2^{\frac{n}{2}}\right). \quad \blacksquare$$

Let ν be some f–invariant measure in $E(0)$. Substituting in Corollary 6.2 $l_2 = \nu$ we obtain

Corollary 6.3. ν *is the only* f–*invariant measure* ν *in* $E(0)$. *Moreover* $\forall l \in E(0)$ $\forall A \in C^\gamma(X)$, $\forall n > 0$

$$\left|\int A(f^n x)dl(x) - \nu(A)\right| \le C_3\rho_3^n||A||_\gamma.$$

7. Coupling algorithm.

Here we define τ and R described in Lemma 6.1. Let $\mathcal{Y}$ be the set of rectangles $Y = V \times I$, $V \in \mathcal{V}$, $I \subset [0,1]$. Let $D > 2$ be a constant defined below (see (13). Let $\tilde{\mathcal{Y}}$ be the set of rectangles $Y = V \times I$, where the length of V is between $\frac{1}{D}$ and D and $I \subset [0,1]$. We write $f(x,t) = (fx, t)$. For $Y_1 \in \mathcal{Y}$, $Y_2 \in \tilde{\mathcal{Y}}$ such that $\text{mes}(Y_1) = \text{mes}(Y_2)$ we give an algorithm defining τ and R. This algorithm will depend on three positive parameters K, λ and ϵ. We require λ be so close to 0 that $\lambda < -\lambda_2$ (recall (1)) and $e^{-\lambda s} > \theta_1$, where s and θ_1 are the constants from Theorem 4.1. By this Theorem and Proposition 4.3 if K is large enough then

$$q_1 = \max_{V \in \mathcal{V}} \text{mes}(U(V)) < 1, \tag{10}$$

where $U(V) = \{x \in V : \exists n > 0, y \in V : d(f^n x, f^n y) \le 2$ and $(df^n|E_c)(y) \ge Ke^{-\lambda n}\}$. Take K so large that (10) is satisfied. Write

$E_{cs} = E_s \oplus E_c$. By partial hyperbolicity there is a constant K' such that $\forall j > 0 \; \forall x$

$$||(df^j|E_{cs})(x)|| \leq K'|(df^j|E_c)(x)|. \tag{11}$$

Set $\tilde{K} = \max(KK', 1)$. Since $(df|E_c)(x)$ is Holder continuous there exists $\delta > 0$ such that if $d(x, y) < \delta$ then $|(df|E_{cs})(y)| \leq e^{\frac{\lambda}{2}}|df|E_{cs}(x)|$. Let

$$\epsilon \leq \frac{\delta}{2\tilde{K}}. \tag{12}$$

Now D is defined by the requirement that if V_1, V_2 are unstable curves, $V_1 \in \mathcal{V}$, V_2 is the image of V_1 under center–stable projection and $\forall x \in V_1 \; d(x, px) < \delta$ then

$$\frac{1}{D} \leq \text{length}(V_2) < D. \tag{13}$$

Our algorithm will work recursively. During the first run we define the map between subsets P_j^∞ of Y_j. For points where τ is not defined we define a stopping time $s(y)$ such that the set $P_j^n = \{y \in Y_j : s(y) = n\}$ will be of the form $f^{-n}(\bigcup_k Y_{jnk})$, $Y_{1nk} \in \mathcal{Y}$, $Y_{2nk} \in \tilde{\mathcal{Y}}$ and $\text{mes}(P_1^n) = \text{mes}(P_2^n)$. Then we can use our algorithm again to couple P_1^n to P_2^n. More precisely since $\text{mes}(P_1^n) = \text{mes}(P_2^n)$ we can chop each Y_{jnk} into several pieces so that the resulting collections $\{\bar{Y}_{jnl}\}$ satisfy $\bigcup_k Y_{jnk} = \bigcup_l \bar{Y}_{jnl}$ and $\text{mes}(\bar{Y}_{1nl}) = \text{mes}(\bar{Y}_{2nl})$. Let $f^{-n}\bar{Y}_{jnl} = U_{jnl} \times I_{jnl}$. Denote $c_{jnl} = \bigcup_{U_{jnl}} \rho_{V_j}(x)dx$. Let Δ_{jnl} be the map $\Delta_{jnl}(x, t) = (f^n x, r_{jnl}(t))$ where r_{jnl} is the affine isomorphism between I_{jnl} and $[0, c_{jnl}|I_{jnl}|]$. (This rescaling is necessary to make Δ's measure preserving.) We now call our algorithm recursively to produce maps $\tau_{nl} : \Delta(f^{-n}\bar{Y}_{1nl}) \to \Delta(f^{-n}\bar{Y}_{2nl})$ and $R_{nl} : \Delta(f^{-n}\bar{Y}_{1nl}) \to \mathbb{N}$ satisfying the conditions of Lemma 6.1. We when set

$$\tau(x, t) = \begin{cases} \tau_{first \ run}(x, t) & if \quad (x, t) \in P_1^\infty \\ \Delta_{2nl}^{-1} \circ \tau_{nl} \circ \Delta_{1nl} & if \quad (x, t) \in f^{-n}\bar{Y}_{1nl}, \end{cases}$$

$$R(x, t) = \begin{cases} R_{first \ run}(x, t) & if \quad (x, t) \in P_1^\infty \\ n + R_{nl}(\Delta_{1nl}(x, t)) & if \quad (x, t) \in f^{-n}\bar{Y}_{1nl}. \end{cases}$$

Let us now describe the first run of our algorithm. By rescaling we may suppose that $Y_j = V_j \times [0, 1]$. By u-convergence there is n_0 and curves $\bar{V}_j$ on distance at least 1 from $\partial(f^n V_j)$ such that $V_1 \in \mathcal{V}$, $\bar{V}_2 = p\bar{V}_1$ and $\forall x \in \bar{V}_1 \; d(x, px) \leq \epsilon$. (Here p means center-stable holonomy.) Let $\hat{c}_j = \int_{f^{-n_0}\bar{V}_j} \rho_{V_j}(x)dx$. Let $(\bar{t}_1, \bar{t}_2) = (1, \frac{\hat{c}_2}{\hat{c}_1})$ if $\hat{c}_2 \leq \hat{c}_1$ and $(\bar{t}_1, \bar{t}_2) = (\frac{\hat{c}_1}{\hat{c}_2}, 1)$ if $\hat{c}_1 \leq \hat{c}_2$. Define $\bar{Y}_j = \bar{V}_j \times [0, \bar{t}_j]$. Let $s(y) = n_0$ for points of $Y_j \backslash f^{-n_0}\bar{Y}_j$.

We now proceed to define P_j^n inductively for $n > n_0$. Let $Q_j^{n-1} = Y_j \backslash \bigcup_{m=n_0}^{n-1} P_j^m$. We assume by induction that $f^{n-1} Q_j^{n-1} = \bigcup_k Z_{jk(n-1)}$ where $Z_{jk(n-1)} = V_{jk(n-1)} \times [0, t_{jk(n-1)}]$,

$$\mathrm{mes}(f^{-(n-1)} Z_{1k(n-1)}) = \mathrm{mes}(f^{-(n-1)} Z_{2k(n-1)}), \qquad (14)$$

$V_{1k(n-1)} \in \mathcal{V}$ and $V_{2k(n-1)} = p(V_{1k(n-1)})$ and $d(x, px) \le r_{n-1}$ where

$$r_n = \tilde{K} \epsilon e^{-\frac{\lambda n}{2}} \qquad (15)$$

Take a Markov decomposition $f V_{1k(n-1)} = \bigcup_l V_{1lkn}$ and let $V_{2lkn} = p(V_{1lkn})$. We note that the fact that $V_{1lkn} \in \mathcal{V}$, (12), (13) and (15) guarantee that then $V_{2lkn} \in \tilde{\mathcal{Y}}$. Let $\beta_{lkn} = \max\limits_{x \in f^{-n} V_{1lkn}} (df^n | E_c)(x)$. If $\beta_{lkn} > K e^{-\lambda n}$ let $s(y) = n$ on $f^{-n} V_{jlkn} \times [0, t_{jk(n-1)}]$. Otherwise let $\tilde{Z}_{jlkn} = V_{jlkn} \times [0, t_{jk(n-1)}]$. In general $\mathrm{mes}(f^{-n} \tilde{Z}_{1lkn}) \ne \mathrm{mes}(f^{-n} \tilde{Z}_{2lkn})$. So cutoff the top of the larger rectangle so that the adjusted ones satisfy $\mathrm{mes}(f^{-n} Z_{1lkn}) = \mathrm{mes}(f^{-n} Z_{2lkn})$ and let $s(y) = n$ on $\tilde{Z}_{jlkn} \backslash Z_{jlkn}$. Now $P_j^\infty = Y_j \backslash \bigcup_n P_1^n$ is a union of vertical intervals of the form $\{(x, [0, t(x)])\}$ where x varies over some positive measure Cantor set R_1. For points of P_1^∞ let $\tau(x, t) = (px, \frac{t(px)}{t(x)} t)$, $R(x, t) = n_0$.

Four things have to be proven:

$-\tau$ is defined on a set of whole measure in Y_1;

$-\tau$ is measure preserving;

$-\tau$ satisfies condition (A) of Lemma 6.1;

$-\tau$ satisfies condition (B) of Lemma 6.1.

The second and the third properties of τ are verified in Section 8. The first and the fourth properties are verified in Section 9.

8. Convergence of images.

Here we prove the property (A) of Lemma 6.1. Let $W_r^*(x)$ denote the ball centered at x of radius r inside W^* with induced Rimannian metric. Let $(x, t) \in P_1^\infty$. Let $x_0 = f^{n_0} x$. Then $\forall j \ge 0\ (df^j | E_c)(x_0) \le K e^{-\lambda j}$, so it suffices to show the following.

Lemma 8.1. *(Cf. [2], Lemma 2.7) If $x_0 \in X$ and $n > 0$ are such that $\forall 0 \le j \le n\ (df^j | E_c)(x_0) \le K e^{-\lambda j}$ then $\forall 0 \le j \le n$*

$$f^j W_\epsilon^{cs}(x_0) \subset W_{r_j}(f^j x_0) \qquad (16)$$

where r_j is given by (15).

Proof. We proceed by induction. For $j = 0$ (16) is true by (12). Suppose that (16) holds for $0 \le j < j_0$. Then $\forall y \in W_\epsilon^{cs}(x)\ \forall 0 \le j < j_0$

$d(f^j y, f^j x_0) \leq \delta$. Hence

$$|(df^{j_0}|E_{cs})|(y) = \prod_{j=0}^{j_0-1} |(df|E_c)(f^j y)| \leq \prod_{j=0}^{j_0-1} \left[e^{\frac{\lambda}{2}}|(df|E_c)(f^j x_0)| \right] \leq K e^{-\frac{j_0\lambda}{2}}.$$

By (11) $||(df^j|E_{cs})(y)|| \leq \tilde{K} e^{-\frac{j_0\lambda}{2}}$ and so $f^j W_\varepsilon^{cs}(x_0) \subset W_{r_j}(f^j x_0)$ as claimed. ∎

Note that the proofs of Lemma 8.1 and Theorem 4.1 do not use u–convergence. Hence we get the following result which will be used in Section 10.

Corollary 8.2. *Assume that f satisfies all the conditions of Theorem I except possibly u–convergence. There are constants $q_1, \epsilon > 0$ such that for any pair of unstable curves V_1, V_2 such that $V_1 \in \mathcal{V}$, $V_2 = p(V_1)$ and $\forall x \in V$ $d(x, px) < \epsilon$*

$$\text{mes}(\{x \in V_1 : d(f^n x, f^n px) \to 0\}) \geq q_1.$$

We now continue with the proof of Lemma 6.1.

Corollary 8.3. *τ is measure preserving.*

Proof. By the recursive structure of our algorithm it suffices to show that $\tau : P_1^\infty \to P_2^\infty$ is measure preserving. Denote by R_j the base of P_j^∞. It follows from Lemma 8.1 by standard Pesin theory (see [24], Section 3 or [27], Section 4) that $p : R_1 \to R_2$ is absolutely continuous. We want to compute its Jacobian $J(x)$. Let $t_n(x)$ denote the height of the rectangle containing $(x, 0)$ and $W_n(x)$ denote its base. By absolute continuity for almost all $x \in R_1$ x is a density point of R_1 and px is a density point of R_2. For such points

$$J(x) = \lim_{n \to \infty} \frac{\int_{pW_n(x)} 1_{R_2}(y)\rho_{V_2}(y)dy}{\int_{W_n(x)} 1_{R_1}(y)\rho_{V_1}(y)dy}$$

Since R_1, R_2 have large densities in $W_n(x)$ and $pW_n(x)$ respectively we can drop indicator functions. So

$$J(x) = \lim_{n \to \infty} \frac{\int_{pW_n(x)} \rho_{V_2}(y)dy}{\int_{W_n(x)} \rho_{V_1}(y)dy} = \lim_{n \to \infty} \frac{t_n(x)}{t_n(px)} = \frac{t(x)}{t(px)}$$

where the second equality follows by (14). Thus τ is measure preserving. ∎

9. Coupling time.

Here we prove the part (B) of Lemma 6.1. We begin with some information about one run of our algorithm.

Lemma 9.1. *There are constants $q, C_0 > 0, \rho_0 < 1$ such that for any pair Y_1, Y_2*

(1) $\frac{\text{mes}(P_1^\infty)}{\text{mes}(Y_1)} \geq q$;

(2) $\frac{\text{mes}(P_1^n)}{\text{mes}(Y_1)} \leq C_0 \rho_0^n$.

Proof. We begin with (2). (y,t) can belong to P_1^n for two reasons. The first, $\exists \tilde{x}$ such that $d(f^n y, f^n \tilde{x}) \leq 2$ and $(df^n|E_c)(\tilde{x}) > K e^{-\lambda n}$. By Proposition 4.3 $(df^n|E_c)(\tilde{x}) > K^* e^{-\lambda n}$. So the measure of such points is exponentially small by Theorem 4.1 and our choice of λ. The second reason is that $f^n(y,t) \in \tilde{Z}_{1lkn} \backslash Z_{1lkn}$. Let

$$\kappa_n(y) = \frac{\text{mes}(f^{-n}\tilde{Z}_{1lkn} \backslash Z_{1lkn})}{\text{mes}(f^{-n}Z_{1kln})} = \frac{\text{mes}(\tilde{Z}_{1lkn} \backslash Z_{1lkn})}{\text{mes}(Z_{1kln})}.$$

Lemma 9.2. *There are constants $C_4 > 0, \rho_4 < 1$ such that*

$$\kappa_n(y) \leq C_4 \rho_4^n. \tag{17}$$

Proof. We may suppose that $\text{mes}(\tilde{Z}_{1lkn}) > \text{mes}(\tilde{Z}_{2lkn})$, since otherwise $Z_{1kln} = \tilde{Z}_{1kln}$. Then

$$\frac{\text{mes}(\tilde{Z}_{1lkn} \backslash Z_{1lkn})}{\text{mes}(Z_{1kln})} = \frac{\text{mes}(V_{1lkn})\text{mes}(V_{2k(n-1)})}{\text{mes}(V_{2lkn})\text{mes}(V_{1k(n-1)})} - 1. \tag{18}$$

Now by inductive assumption we have both on $V_{1k(n-1)}$ and on V_{1lkn}

$$d(x, px) \leq \tilde{K}\varepsilon e^{-\frac{\lambda}{2}(n-1)}. \tag{19}$$

In the proof below C_* will denote various constants which depend on f but not on n or Y_j. Likewise ρ_* will denote various constants which are less then 1. Let x_0 be the center of $V_{1k(n-1)}$ and $\tilde{x}_0 = px_0$. By Holder continuity of unstable foliation $\exists C_5 > 0, \rho_5 < 1$ such that

$$\left| \frac{\rho(x_0, x)}{\rho(\tilde{x}_0, px)} - 1 \right| \leq C_5 \rho_5^n.$$

Divide $V_{1k(n-1)}$ into subintervals σ_m of size $e^{-\frac{\lambda n}{4}}$ and let $\tilde{\sigma}_m = p\sigma_m$. Then

$$\int_{V_{1k(n-1)}} \rho(x_0, x)dx = \sum_m \rho(x_0, x_m)d_m + O(\rho_6^n),$$

where x_m is any point on σ_m and d_m is the distance between the end-points of σ_m. Likewise

$$\int_{V_{1k(n-1)}} \rho(x_0, x)dx = \sum_m \rho(\tilde{x}_0, px_m)\tilde{d}_m + O(\rho_6^n)$$

where $\tilde{d}_m$ is the distance between the endpoints of $\tilde{\sigma}_m$. Now by (19) and the triangle inequality $|\frac{d_m}{\tilde{d}_m} - 1| \leq C_6 e^{-\frac{\tilde{\lambda}n}{2}}$. Hence

$$\left| \frac{\mathrm{mes}(V_{1k(n-1)})}{\mathrm{mes}(V_{2k(n-1)})} - 1 \right| \leq C_7 \rho_7^n.$$

Similarly

$$\left| \frac{\mathrm{mes}(V_{1kln})}{\mathrm{mes}(V_{2kln})} - 1 \right| \leq C_7 \rho_7^n.$$

The last two inequalities together with (18) prove the lemma. ∎
Assertion (2) of Lemma 9.1 now follows from (17) by summation over k and l.

Now let $(x, t) \in P_1^\infty$. Let $\kappa_j(x)$ be the relative measure of points cut off the top of the rectangle containing (x, t) on the jth step. Thus $t(x) = \prod_{j=n_0}^\infty (1 - \kappa_j(x))$. By (17) this series converges uniformly, hence $t(x)$ is uniformly bounded from below. But the measure of the base of P_1^∞ is also uniformly bounded (see (10)). This proves assertion (1) of Lemma 9.1. ∎
Now represent $R(y) = \sum_{j=1}^{k(y)} s_j(y)$, where $s_j(x)$ is the stopping time of the jth run of our algorithm. Let T_k be the set where τ is not defined after k runs of our algorithm, $U_k = T_{k-1} \backslash T_k$. Denote $S_k(x) = \sum_{j=1}^k s_j(x)$ and consider generating functions $\varphi_k(Y_1, \xi) = \frac{1}{\mathrm{mes}(Y_1)} \int_{T_k} \xi^{S_k}(x)dm(y)$, $\psi_k(Y_1, \xi) = \frac{1}{\mathrm{mes}(Y_1)} \int_{U_k} \xi^{S_k}(x)dm(y)$. Lemma 9.1 says that the radius of convergence of φ_1 is strictly greater than 1 and $\varphi_1(1) \leq 1 - q$. We need the following generalization.

Lemma 9.3. $\exists \delta_0, \bar{q}, C > 0$ such that if $0 \leq \xi \leq 1 + \delta_0$ then
  (1) $\varphi_{k+1}(\xi) \leq (1 - \bar{q})\varphi_k(\xi)$;
  (2) $\psi_{k+1} \leq C\varphi_k(\xi)$.

Proof. (1) Take some $y \in T_k$. Assume that after k runs of our algorithm $f^{S_k(y)} \in Y_1^k(y)$ and $(k+1)$-st run couples Y_1^k to some Y_2^k. Let us compare the contributions of Y_1^k to φ_k and φ_{k+1}. (That is if $Y_1^k = \Delta^{(k)}\bar{Y}_1^k$ where $\bar{Y}_1^k \subset Y_1$, $\Delta^{(k)}(x, t) = (f^{S_k(y)}x, at + b)$ and we compare

$$I_k = \frac{1}{\mathrm{mes}(Y_1)} \int_{\bar{Y}_1^k} \xi^{S_k(x)}dm_1 = \xi^{S_k(y)} \frac{\mathrm{mes}(\bar{Y}_1^k)}{\mathrm{mes}(Y_1)}$$

and $I_{k+1} = \frac{1}{\text{mes}(Y_1)} \int_{\bar{Y}_1^k} \xi^{S_{k+1}(x)} dm_1$.) Their ratio equals $r(\xi) = \frac{I_{k+1}}{I_k} = \varphi_1(Y_1^k, \xi)$. By Lemma 9.1 $r(\xi)$ is uniformly convergent and $r(1) \leq 1-q$. So there is $\delta_0, \bar{q}$ such that for $\xi < 1 + \delta_0$ $r(\xi) \leq 1 - \bar{q}$. This proves (1).

$$(2)\ \psi_{k+1}(\xi) = \xi^{n_0} \int_{U_{k+1}} \xi^{S_k(y)} dm(y) \leq \xi^{n_0} \int_{T_k} \xi^{S_k(y)} dm(y) \leq (1+\delta_0)^{n_0} \varphi_k(\xi). \quad \blacksquare$$

Now for $\xi \leq 1 + \delta_0$

$$\int_{Y_1} \xi^{R(y)} dm(y) = \sum_{k=1}^{\infty} \psi_k(\xi) \leq C \sum_{k=1}^{\infty} (1 - \bar{q})^{k-1} \varphi_1(\xi) \leq \frac{C}{\bar{q}} \varphi_1(\xi) < \infty.$$

This shows that $m(R > n) \leq \frac{\text{Const}}{(1+\delta_0)^n}$ and in particular $m(R = \infty) = 0$. These facts complete the proof of Lemma 6.1.

10. PROOF OF THE MAIN RESULTS.

PROOF OF THEOREM I. Consider $l \in E(R)$. By Proposition 5.1 there exists $\tilde{l} \in E(0)$ such that for all $N > 0$

$$||\mathcal{T}^{N+\frac{n}{2}} l - \mathcal{T}^N \tilde{l}|| \leq \text{Const} e^{\frac{-\lambda_5 n}{2}}.$$

Hence

$$||\mathcal{T}^n l - \nu|| \leq \text{Const} e^{\frac{-\lambda_5 n}{2}} + ||\mathcal{T}^{\frac{n}{2}} \tilde{l} - \nu|| \leq C_8 \rho_8^n.$$

It follows from [12] that ν is a global attractor for f. Let $B \in C^{\gamma}(X)$ then, for large R, $B \cdot$Lebesgue and $B \cdot \nu$ are in $E(R)$, provided that γ is sufficiently close to 1 (see [26, 12]). This together with Proposition 5.2 proves Theorem I for γ close to 1. The result for general γ is proved by approximation of A and B by smooth functions. $\blacksquare$

PROOF OF THEOREM II. By the remark at the end of Section 3 there exists a C^2 neighborhood $\mathcal{O}_1(f)$ such that any $g \in \mathcal{O}_1(f)$ is mostly contracting. Also, given ε there is another C^2 neighborhood $\mathcal{O}_2(f)$ such that $\forall g \in \mathcal{O}_2(f)\ \exists n_0$ such that $\forall V_1, V_2 \in \mathcal{V}\ \exists U_j \subset V_j$ such that $g^{n_0} U_1 \in \mathcal{V}$, $U_2 = pU_1$ and $\forall x \in U_1\ d(g^{n_0} x, g^{n_0} px) < \varepsilon$. By Corollary 8.2 g is u–convergent. Constants C, ζ from part (b) of Theorem I can be chosen uniformly for g near f since they depend only on Holder data of invariant foliations and the constant α_0 in (6). Thus

$$\nu_g(A) - \nu_f(A) = \int [A(g^n x) - A(f^n x)] dx + O(||A||_{\gamma} \zeta^n) =$$

$$||A||_{\gamma}(O(\zeta^n) + O((K^n d(f,g))^{\gamma} + \zeta^n)) =$$
$$||A||_{\gamma} O(K^n d(f,g))^{\gamma} + \zeta^n).$$

Taking n so that $(\frac{\zeta}{K^{\gamma}})^n = d(f,g)^{\frac{\gamma}{2}}$ we obtain the result needed. $\blacksquare$

11. Examples of u-convergent diffeomorphisms.

Here we give several conditions sufficient for u-convergence.

(a) Suppose that f has 3-leg accessibility property in the sense that there exists R such that $\forall V_1, V_2 \in \mathcal{V}$ $\exists x_1 \in V_1$, $x_2 \in V_2$ such that $x_2 \in W^s(x_1)$ and $d_s(x_1, x_2) \leq R$ where d_s means the distance along the stable leaf. Then $d_s(f^n x_1, f^n x_2) \leq e^{\lambda_2 n} R$ so f is u-convergent.

(b) Assume that W^u is minimal. Thus given ε there exists R such that for any two unstable curves V_1, V_2 of length at least R there are $x_j \in V_j$ such that $d(x_1, x_2) \leq \varepsilon$. So, f is u-convergent.

(c) To formulate this condition suppose that the fibers of W^c are circles. Suppose that f satisfies the following condition: for any two unstable curves $V_1 \in \mathcal{V}$ and $V_2 \in W^{cu}(V_1) : V_2 = p_c(V_1)$ (where p_c denotes the center holonomy) the following inequality holds

$$\int_{V_1} \rho_{V_1}(x) \ln\left(\frac{d(fx, fpx)}{d(x, px)}\right) dx \leq -\alpha_0 < 0. \tag{20}$$

(Note that letting here V_2 tend to V_1 we obtain (6) so (20) is a stronger assumption).

Proposition 11.1. *there are sets $U_j \subset V_j$ such that $\mathrm{mes}(V_j \backslash U_j) = 0$, $U_2 = p(U_1)$ and for all $x \in U_1$ $d(f^n x, f^n px) \to 0$.*

Proof. Let $U_1 = \{x \in V_1 : d(f^n x, f^n px) \to 0\}$. Repeating the arguments of Corollary 8.2 we get that there is a constant $\tilde{q}$ (depending only on f but not on V_1, V_2) such that

$$\mathrm{mes}(U_1) \geq \tilde{q}\mathrm{mes}(V_1). \tag{21}$$

Now considering Markov decompositions $f^n V_1 = \bigcup_l V_{ln}$ and applying (21) to each V_{ln} we find that $V_1 \backslash U_1$ has no density points and so $\mathrm{mes}(V_1 \backslash U_1) = 0$. Interchanging V_1 and V_2 we get $\mathrm{mes}(V_2 \backslash U_2) = 0$. ∎
Now since W^u and W^{cs} are transverse foliations there exists $R > 0$ such that $\forall x_1, x_2 \in X$ there is $x_3 \in X$ such that $x_3 \in W^{cs}(x_1) \bigcup W^u(x_2)$ and $d_{cs}(x_1, x_3) < R$, $d_u(x_1, x_3) < R$. Iterating this construction forward we find that $\forall \delta$ $\exists \tilde{R}$ such that $\forall x_1, x_2$ $\exists y_1, y_2$ such that $y_1 \in W^s_\delta(x_1)$ and $y_2 \in W^c(y_1) \bigcap W^u_{\tilde{R}}(x_2)$. Now if δ is small enough then by Corollary 8.2 there exist sets $U_1 \subset W^u_1(U_1)$, $\tilde{U} \subset W^u_1(y_1)$ such that $\mathrm{mes}(U_1) > 0$, $\mathrm{mes}(\tilde{U}_1) > 0$ $\tilde{U}_1 = p_{cs}(U_1)$ and $\forall x \in U_1$ $d(f^n x, f^n p_{cs} x) \to 0$. But $y_2 \in W^u_{\tilde{R}}(x_2)$. By compactness there is $R^* > 0$ such that $U_2 \subset W_{R^*}(x_2)$. Thus $\exists z_1 \in W^u_1(x_1), z_2 \in W^u_{R^*}(x_2)$ such that $d(f^n z_1, f^n z_2) \to 0$. Now take $V_1, V_2 \in \mathcal{V}$. There exists n_0 such that the lengths of $f^{n_0} V_j$ is greater than $2R^*$ and so $\exists z_j \in f^{n_0} V_j :$ such that $d(f^n z_1, f^n z_2) \to 0$. Therefore

f is u–convergent. Hence (20) implies both mostly contractiveness and u-convergence.

Similarly we can consider the case when the leaves of f are non-compact and require that (20) is satisfied for $d(V_1, V_2) \leq R$ there $R = R(f)$ is a large constant.

Remark. If fibers of W^c are circles and (20) is satisfied then one can show using Proposition 11.1 that W^c–holonomy restricted to W^{cu} is absolutely continuous. Let us examine underlying geometric picture more closely since it will allow the reader to appreciate better the idea behind the proof of Lemma 6.1. Indeed choose an orientation of W^c and let V_1, V_2 be two unstable curves in the same center–unstable leaves such that $V_2 = p(V_1)$ V_2 is ε–close to V_1 and is on the right of V_1. Then there are subsets $U_j \subset V_j$ such that $U_2 = p(U_1)$, $\mathrm{mes}(V_j/U_j) = 0$ and $\forall x \in U_1\ d(f^n x, f^n p x) \to 0$. The geometry of U_1 is however quite complicated. In fact there is a Cantor set $U_1^{(1)} \subset U_1$ of measure $1 - O(\varepsilon^C)$ such that for $x \in U_1^{(1)}$ $f^n p x$ is always on the right from $f^n x$. Each gap of $U_1^{(1)}$ contains a positive measure Cantor set $U_1^{(2)}$ such that for $x \in U_1^{(2)}$ $f^n(px)$ is on the left from $f^n x$. In turn each gap of $U_1^{(2)}$ contains a positive measure Cantor set $U_1^{(3)}$ such that for $x \in U_1^{(2)}$ $f^n(px)$ is on the right from $f^n x$ and so on.

If fibers of W^c are lines then it is conceivable that W^c is not absolutely continuous inside W^{cu}. Instead Lemma 6.1 allows us to construct a map $\pi : V_1 \to V_2$ which is absolutely continuous and such that $\pi x \in W^{cs}(x)$. However even if V_2 is very close to V_1 still πx sometimes will be different from the naive projection along the W^c–fibers. The fact that the fibers of W^{cs} are dense and so for each x there is a countable number of candidates for πx is really essential to this proof.

12. EXAMPLES OF MOSTLY CONTRACTING DIFFEOMORPHISMS.

(a) Let $T : \mathbb{T}^2 \to \mathbb{T}^2$ be a linear Anosov diffeomorphism. Take $A : \mathbb{T}^2 \to SL_2(\mathbb{R})$. Assume that the image $A(\mathbb{T}^2)$ generates $SL_2(\mathbb{R})$. Let $S_{n_1,n_2}(x) = A(T^{n_1 n_2}x) \ldots A(T^{n_2}x)A(x)$, $M_{n_1,n_2}(x)v = \frac{S_{n_1,n_2}(x)v}{||S_{n_1,n_2}(x)v||}$. Define $f_{n_1,n_2} : \mathbb{T}^2 \times \mathbb{P}^1 \to \mathbb{T}^2 \times \mathbb{P}^1$ by $f_{n_1 n_2}(x, v) = (T^{n_1 n_2}x, M_{n_1,n_2}v)$. We claim that $\exists \bar{n}_2$ such that $\forall n_2 > \bar{n}_2\ \exists \bar{n}_1$ such that $\forall n_1 > \bar{n}_1\ f_{n_1,n_2}$ satisfy (20). Let $d(v_1, v_2) = \mathrm{Area}(v_1, v_2)$. We have

$$d(f_{n_1,n_2}(x, v_1), f_{n_1,n_2}(x, v_2)) =$$
$$\mathrm{Area}(M_{n_1,n_2}(x)v_1, M_{n_1,n_2}(x)v_2) =$$
$$\frac{\mathrm{Area}(S_{n_1,n_2}(x)v_1, S_{n_1,n_2}(x)v_2)}{||S_{n_1,n_2}(x)v_1||\,||S_{n_1,n_2}(x)v_2||} =$$

$$\frac{\text{Area}(v_1, v_2)}{||S_{n_1,n_2}(x)v_1||\,||S_{n_1,n_2}(x)v_2||}$$

since $S \in SL_2(\mathbb{R})$. The reader will have no difficulties to show that for fixed n_2 and large n_1 f_{n_1,n_2} is partially hyperbolic, its unstable manifolds are graphs of functions $\Gamma_n : W^u(x_0) \to \mathbb{P}^1$ and $||d\Gamma_n|| \to 0$ as $n \to \infty$. Now the Riemann structure of $\mathbb{T}^2$ is non-degenerate on the leaves of W^u and with respect to this structure $\rho((x_1, \Gamma_n(x_1)), (x_2, \Gamma_n(x_2))) = \rho(x_1, x_2)$. Let m be the distribution of $A(x)$ with respect to Lebesgue measure and m_n be the n-th convolution power of m. Take now $x_0 \in \mathbb{T}^2$, $v_1, v_2 \in \mathbb{P}^1$ and let $\Gamma_n^j(x)$ be the function defining the unstable manifold through (x_0, v_j). Then

$$\int_V \rho(z) \ln \left(\frac{d(f_{n_1,n_2}(x, \Gamma_n^1(x)), f_{n_1,n_2}(x, \Gamma_n^2(x)))}{d(v_1, v_2)} \right) dz \to \qquad (22)$$

$$-\frac{1}{|V|} \int_V \mathbb{E}_{m_{n_2-1}} \left(\ln ||AM(x)\Gamma_n^1(x)|| + \ln ||AM(x)\Gamma_n^2(x)|| \right) dx$$

as $n_1 \to \infty$. But by [5], Theorem A3.6 for all v $\frac{1}{n}\mathbb{E}_{m_n} \ln ||Av|| \to \lambda_+ > 0$ where λ_+ is the positive exponent of m. So for large n_2, $n_1 \gg n_2$ the expression (22) is negative.

(b) This example is similar to (a). Let $f_n : \mathbb{T}^2 \times S^1 \to \mathbb{T}^2 \times S^1$ be given by $f_n(x, y) = (T^n x, g_n(x)y)$ where the distribution of g_n converges to that of time 1 map of the stochastic differential equation

$$dy = \Phi(y)dw(t), \qquad (23)$$

$dw(t)$ being the white noise. Then for large n f_n is mostly contracting. This follows from the fact that $\xi = \frac{\partial y}{\partial y_0}(t)$ satisfies

$$d \ln \xi = \frac{d\Phi}{dy}dw - \frac{1}{2}\left(\frac{d\Phi}{dy}\right)^2 dt$$

and so

$$\mathbb{E}_{\mathbf{P}} \ln \xi(1) = -\frac{1}{2}\int_0^1 \mathbb{E}_{\mathbf{P}}\left(\frac{d\Phi}{dy}(t)\right) dt < 0.$$

where $\mathbf{P}$ denotes the stationary distribution of process defined by (23). (Of course since the distribution of ξ is not compactly supported some restrictions should be imposed on the rate of convergence of g_n to the distribution of the solution of (23). We leave it as an exercise to the reader to write down the explicit estimates.)

Remark. This example shows that many phenomena occurring in stochastic differential equation can also take place in deterministic systems. More research is needed in this direction.

(c) The above examples are essentially dissipative, since where both f and f^{-1} are mostly contracting. Here we describe a conservative example which is a slight modification of the one given in [32].

Let $T : \mathbb{T}^2 \to \mathbb{T}^2$ be as before. Consider $f_0 : \mathbb{T}^3 \to \mathbb{T}^3$ given by $f_0(x, \theta) = (T(x), \theta + \tau(x))$ where τ is such that f_0 is Bernoullian (this condition discards an infinite codimension submanifold in the space of skewing functions). Let $h : \mathbb{T}^3 \to \mathbb{T}^3$ be a volume preserving diffeomorphisms close to identity. (dh maps small cones $\mathcal{K}_u$ around E_u and $\mathcal{K}_{cu}$ around $E_u \wedge E_c$ into themselves. Likewise df^{-1} preserves cones $\mathcal{K}_s$ and $\mathcal{K}_{cs}$. We want dh map E_u into $\mathcal{K}_u$ and so on.) Let $f_n = f_0^n \circ h \circ f_0^n$. Choose vectorfields $e_u \in E_u$, $e_c \in E_c$ and $e_s \in E_s$ so that $df(e_u) = \lambda e_u$, $df(e_c) = e_c$, $df(e_s) = \frac{1}{\lambda} e_s$. Suppose that in this basis $dh(x) = (A_{ij}(x))$. f_n has an unstable vector of the form $v_u = e_u + \alpha(x)e_c + \beta(x)e_s$. Let $df_n(x)v_u = r_1^{(n)}(x)v_u$. The direct calculation shows that $r_1^{(n)}(x) = \lambda^{2n} A_{11}(f_0^n x)(1 + O(\frac{1}{\lambda^n}))$. Similarly choose the central vector v_0 so that $v_u \wedge v_0 = e_u \wedge e_0 + \dots$ and let $df_n(x)(v_u \wedge v_0) = r_2^{(n)}(x)(v_u \wedge v_0)$. Then $r_2^{(n)}(x) = \lambda^{2n}(A_{11}A_{22} - A_{12}A_{21})(f_0^n x)(1 + O(\frac{1}{\lambda^n}))$. Again one can show that unstable manifolds of f_n are close to unstable manifolds of f_0 and so they are transversal to W^{cs}. Let

$$\mathcal{J}(h) = \int_{\mathbb{T}^3} \ln\left(\frac{A_{11}A_{22} - A_{12}A_{21}}{A_{11}}(x)\right) dx.$$

We have

$$\int_{V_n} \rho_{V_n}(x) \ln(df_n|E_c)(x)dx \approx$$

$$\int_{V_n} \rho_{V_n}(x) \ln\left(\frac{A_{11}A_{22} - A_{12}A_{21}}{A_{11}}\right)(f_0^n x)dx$$

Since f_0 is mixing we obtain (cf. [18] or [15] Section 20.6) that

$$\int_{V_n} \rho_{V_n}(x) \ln\left(\frac{A_{11}A_{22} - A_{12}A_{21}}{A_{11}}\right)(f_0^n x)dx \approx$$

$$\mathcal{J}(h) \int_{V_n} \rho_{V_n}(x)dx \approx \mathcal{J}(h).$$

So, if $\mathcal{J}(h) < 0$ then f_n is mostly contracting for large n. Similarly if $\mathcal{J}(h) > 0$ then f_n^{-1} is mostly contracting for large n. (A more symmetric expression for $\mathcal{J}$ is

$$\mathcal{J} = \int_{\mathbb{T}^3} [\ln A_{33}(h^{-1}) - \ln A_{11}(h)]dx.)$$

To show that $\mathcal{J}$ is not identically zero one can use the following argument of [32]. Let $\mathrm{Diff}^*(\mathbb{T}^3)$ be the space of volume preserving C^{3-} diffeomorphisms which preserve $W^{cs}(f_0)$. Then $\mathcal{J}|_{\mathrm{Diff}^*(\mathbb{T}^3)}$ is a C^2-functional and calculating its first two derivatives at identity one can prove that $\mathcal{J} \not\equiv 0$. We refer the reader to [32] for more details.

(d) A similar example can be given with f_0 being a time one map of the geodesic flow on unit tangent bundle over a negatively curved surface.

(e) In [4] several examples are given of the systems having the following property

(*) $\forall V \in \mathcal{V}$ there is a subset $U \subset V$ of positive measure such that for all $x \in U$ the forward Lyapunov exponent of E_c is negative.

The next proposition is essentially proven in [4] even though it is not stated where. For the convenience of the readers we sketch their arguments below.

Proposition 12.1. *f satisfies (*) $\Leftrightarrow$ it is mostly contracting.*

Proof. In view of Theorem 4.1 we only have to show that if f satisfies (*) then it is mostly contracting. Given $x_0 \in X$ choose V containing x_0 in its interior. By (*) $\exists K(x_0), \lambda(x_0)$ such that the set $L(x_0) = \{x \in V : (df^n|E_c)(x) \le Ke^{-\lambda n}\}$ has positive measure. Given δ there is $\varepsilon > 0$ and a positive measure subset $\tilde{L} \subset L$ such that

(a) $\forall x \tilde{L}$ $W^{cs}_\varepsilon(x)$ belong to the weak stable manifold of x;

(b) If $\tilde{V} \in \mathcal{V}$, $d(\tilde{V}, V) \le \varepsilon$ the the center stable holonomy $p : V \to \tilde{V}$ is absolutely continuous on $\tilde{L}$ and its Jacobian $J(x)$ satisfies $|J(x)-1| \le \delta$.

Let y be a density point of $\tilde{L}(x_0)$. It follows that if I is small enough interval about y then $\forall n$

$$\int_I \rho_I(x) \ln(df^n|E_c)(x)dx \le -\frac{\lambda n}{2} + K_1.$$

Let $C = [I, W^{cs}_\varepsilon(y)]$ ($[\cdot, \cdot]$ denotes (u,cs)-local product). Then if δ, ε are small enough then for any unstable slice J of C

$$\int_J \rho_J(x) \ln(df^n|E_c)(x)dx \le -\frac{\lambda n}{4} + K_2$$

Call $T = [V(x_0), W^{cs}_{\varepsilon(x_0)}(y(x_0))]$ the trap associated with x_0. Call C the core of T. By compactness X is covered by a finite number of traps $\{T_j\}$. Now take any $V \in \mathcal{V}$. Given $m > 0$ let $V_1(m)$ be the set of points which visit some C_j before time m and let $V_2(m) = V \backslash V_1(m)$. We have

$$\mathcal{I}_n = \int_V \rho_V(x) \ln(df^n|E_c)(x)dx =$$

$$\left(\int_{V_1(m)} + \int_{V_2(m)}\right) \rho_V(x) \ln(df^n|E_c)(x)dx = I + II.$$

$$I \leq \left[-\frac{\lambda n}{4} + K(m)\right] \operatorname{mes}(V_1(m)),$$

$$II \leq \operatorname{Const} n \operatorname{mes}(V_2(m))$$

and $\operatorname{mes}(V_2(m)) \leq \operatorname{Const}\theta^m$ for some $\theta < 1$ (the proof of this last inequality is similar to that of Lemma 9.1.) So, for large n, $\mathcal{I}_n$ is negative.
∎

13. Conclusions.

Here we repeat what we have said in the introduction adding more technical details.

13.1.   Here we relate our results with those of [38]. Let K, λ be as in Section 7. For $V \in \mathcal{V}$ let $x(V)$ be the center of V and

$$L(V) = \{x \in V : (df^n|E_c)(x) \leq Ke^{-\lambda n}\}.$$

Let $\Lambda_\delta(V) = [L, W_\delta^{cs}(x(V))]$. Then we have essentially shown that if δ is small enough then $\Lambda_\delta(V)$ satisfies the conditions of Theorem 2 of [38]. We did not deduce our result from [38] but rather repeated some of her arguments in Sections 6–9 in order to show that u-convergence guarantees the absence of the discrete spectrum. Now suppose that f satisfies all the conditions of Theorem I except u-convergence. Then the conclusion can be false (consider, for example, the double covering of f from 12(a) corresponding to $\pi : S^1 \to \mathbb{P}^1$). However we can still say something. Namely by [4] there is a finite number $\nu_1 \ldots \nu_k$ of SRB measures and the union of their basins is the whole of X. Let $\Omega = \bigcup_j \operatorname{supp}(\nu_j)$. Let ϵ be as in Section 7. Choose a finite disjoint set $V_1 \ldots V_m$ which is ϵ-dense in Ω. Take $\delta \ll \epsilon$. Choose small subintervals $U_j \subset V_j$ such that $\Lambda_\delta(U_l)$ are disjoint. Then the arguments of Sections 8 and 9 show that $\Lambda_\delta = \bigcup_{l=1}^m \Lambda_\delta(U_l)$ satisfy Theorem 2 of [38] except maybe f^n is not ergodic for some n. It then follows from the analysis of [38] that $\forall j \, \exists n_j$ so that $\nu_j = \frac{1}{n_j}\sum_{l=0}^{n_j} \nu_{jl}$ and (f^{n_j}, ν_{jl}) is exponentially mixing.

Question. What happens for g close to f? Can the maps $g \to \nu_j(g)$ be chosen continuously?

13.2. It seems that the assumption that f is dynamically coherent can be relaxed (it is however satisfied in all the known examples). In fact, we only used it in Section 8. Let (Ω_1, m_1) and (Ω_2, m_2) be probability spaces. Call the map $\tau : \Omega_1 \to \Omega_2$ ε–*measure preserving* if $\exists A_j \subset \Omega_j$ such that $m_j(A_j) \leq \varepsilon$ and $\tau|_{\Omega_1 \setminus A_1}$ is absolutely continuous and the Jacobian satisfies $|J(x) - 1| \leq \varepsilon$. For our arguments it suffices to know that if $d(V_1, V_2) \leq \varepsilon$ then there is ε^β–measure preserving map $p : V_1 \to V_2$ such that for $x \in V_1 \setminus A_1$ $f^n(x)$ and $f^n(px)$ converge exponentially fast. This, in turn seems to follow from the Pesin theory. However, the proof without the dynamical coherence assumption would be much more complicated.

13.3. **Question.** Let f be as in Theorem I. Is the map $g \to \nu_g$ actually smooth? An easier problem is the following. Assume that $W^c(f)$ is absolutely continuous. Is the map $g \to \nu_g$ differentiable at f? (See [31] for additional discussion.)

13.4. **Question.** How common is (6) among partially hyperbolic diffeomorphisms of three manifolds? In particular is the set $\{f : f$ or f^{-1} is mostly contracting $\}$ dense?

13.5. Let X be a three dimensional manifold. Consider the space $\mathcal{S}$ of partially hyperbolic ergodic volume preserving diffeomorphisms with two negative Lyapunov exponents.
Question. How often are elements of $\mathcal{S}$ mostly contracting? What is the rate of mixing for elements of $\mathcal{S}$?

According to the general scheme proposed in [38] one has to locate a 'bad set' of f and see how long an orbit can stay near it. Analysis of [4] shows that the bad set here is the the set of points whose forward orbits never fall into any trap described in Section 12(e). So a way to attack this problem is to obtain more information about the geometry of this set. For example, can it have the Hausdorff dimension equal to three?

13.6. For $f \in \mathcal{S}$ we have the following elegant characterization due to [4]
   f *is mostly contracting if and only if $W^u(f)$ is minimal.*
   (Minimality implies mostly contractiveness by ([4], Theorem B). The converse implication is easier. See, for example [18].)
Question. How often W^u is minimal? What can be said if it is not?

13.7.   Another natural condition to consider if $\dim W^c = 1$ is

$$\forall V \in \mathcal{V} \quad \int_V \rho_V(x) \ln(df|E_c)(x)dx \geq \alpha_0 > 0.$$

In this case the results similar to ours were obtained in [1], [2]. In fact, [1], [2], [4] do not assume that $\dim(W^c) = 1$ but only that all its Lyapunov exponents have the same sign.

Question. Can a similar theory be developed in case $(f|E_u)$ has both positive and negative Lyapunov exponents?

13.8.   In the example 12(a) f is a skew extension over Anosov base. Similar construction can be made with Axiom A attractors.

Question. What can be said for general Axiom A diffeomorphisms? In particular, call f entropy stable if any g near f has an unique measure of maximal entropy μ_g and $\mu_g \to \mu_f$ as $g \to f$. How large is the set of entropy stable diffeomorphisms? Are examples of Section 12 entropy stable?

13.9.   In [7, 8, 16, 20, 28] a number of examples is given of ergodic systems which remain ergodic after a small volume preserving perturbations.

Question. What happens if we allow non-volume preserving perturbations?

13.10.   Finally, let us remark that the questions we asked are special cases of some general conjectures about statistical properties of a generic dynamical systems. See [22].

References

[1] Alves J. *SRB measures for nonhyperbolic systems with multidimensional expansion* IMPA Thesis, 1997.

[2] Alves J., Bonatti C. & Viana M., *SRB measures for partially hyperbolic diffeomorphisms: the expanding case*, preprint.

[3] Benedics M. & Young L.–S. *Sinai–Bowen–Ruelle measures for certain Henon maps.* Inv. Math. **112** (1993) 541-576.

[4] Bonatti C. & Viana M. *SRB measures for partially hyperbolic systems those central direction is mostly contracting*, to appear in Israel Math. J.

[5] Bougerol P. & Lacroix J. *Products of random matrices with applications to Schrodinger operators* Birkhauser, Boston-Basel-Stuttgart, 1985.

[6] Brin M. & Pesin Ya. B. *Partially hyperbolic dynamical systems* Math. USSR-Izvestiya **8** (1974) 177-218.

[7] Burns K., Pugh C. & Wilkinson A. *Stable ergodicity and Anosov flows*, preprint.

[8] Burns K. & Wilkinson A. *Stable ergodicity of skew products*, preprint.

[9] Castro A. *Backwards inducing and decay of correlations for certain partially hyperbolic maps whose central direction is mostly contracting* IMPA Thesis, 1998.

[10] Chernov N. I. *Decay of correlations and dispersing billiards*, J. Stat. Phys. **94** (1999) 513–556.

[11] Chernov N. I. & Haskell C. *Nonuniformly hyperbolic K-systems are Bernoulli* Erg. Th. & Dyn. Sys. **16** (1996) 19–44.

[12] Dolgopyat D. *Limit theorems for partially hyperbolic systems*, preprint

[13] Grayson M., Pugh C. & Shub M. *Stably ergodic diffeomorphisms* Ann. Math. **140** (1994) 295-329.

[14] Hirsh M., Pugh C. & Shub M. *Invariant manifolds* Lect. Notes in Math. **583** Springer-Verlag, Berlin, 1977.

[15] Katok A. & Hasselblatt B. *Introduction to the modern theory of dynamical systems* Encyclopedia Math., Appl. **54** (1998) Cambridge University Press.

[16] Katok A.& Kononenko A.*Cocycle stability for partially hyperbolic systems* Math. Res. Lett. **3** (1996) 191-210.

[17] Margulis G. A. *Certain measures that are connected with U-flows on compact manifolds* Func. Anal., Appl. **4** (1970) 62–76.

[18] Margulis G. A. PhD Thesis, Moscow State University, 1970.

[19] Milnor J. *On the concept of attractor* Comm. Math. Phys. **99** (1985) 177–195 and **102** (1985) 517–519.

[20] Nitica V. & Torok A. *An open dense set of stably ergodic diffeomorphisms in a neighbourhood of a non-ergodic one,* to appear in Topology.

[21] Ornstein D. & Weiss B. *On Bernoulli nature of systems with some hyperbolic structure* Erg. Th. & Dyn. Sys. **18** (1998) 441–456.

[22] Palis J. *Global view on dynamics* to appear in Asterisque.

[23] Palis J., Pugh C. C. & Robinson R. C. *Nondifferentiability of invariant foliations* Lecture Notes Math. **468** (1975) 234–240 Springer, Berlin.

[24] Pesin Ya. B. *Families of invariant manifolds that correspond to nonzero characteristic exponents* Izv. Akad. Nauk SSSR **40** (1976) 1332–1379.

[25] Pesin Ya. B. *Characteristic Lyapunov exponents, and smooth ergodic theory* Uspehi Mat. Nauk **32** (1977) 55–112.

[26] Pesin Ya. B. & Sinai Ya. G. *Gibbs measures for partially hyperbolic attractors* Erg. Th. & Dyn. Sys. **2** (1982) 417–438.

[27] Pugh C. & Shub M. *Ergodic attractors* Trans. AMS **312** (1989) 1-54.

[28] Pugh C. & Shub M.*Stably ergodic dynamical systems and partial hyperbolicity* J. Complexity **13** (1997) 125-179.

[29] Pugh C. & Shub M. *Stable ergodicity and Julienne quasi-conformality*, preprint.

[30] Ruelle D. *A measure associated with Axiom A attractors* Amer. J. Math. **98** (1976) 619-654.

[31] Ruelle D. *Differentiation of SRB states* Comm. Math. Phys. **187** (1997) 227-241.

[32] Shub M. & Wilkinson A. *Pathological foliations and removable zero exponents* Inv. Math (1999).

[33] Sinai, Ya. G. *Classical dynamic systems with countably-multiple Lebesgue spectrum. II* Izv. Akad. Nauk SSSR Ser. Mat. **30** (1966) 15-68.

[34] Sinai Ya. G. *Gibbs measures in ergodic theory* Uspehi Mat. Nauk **27** (1972) 21-64.

[35] Sinai Ya. G. *Stochasticity of dynamical systems* in Nonlinear waves, Edited by A. V. Gaponov–Grekhov Nauka, Moscow (1979) 192-211.
[36] Viana M. *Multidimensional nonhyperbolic attractors* Publ. IHES **85** (1997) 63-96.
[37] Wilkinson A. *Stable ergodicity of time–one map of a geodesic flow* PhD Thesis, University of California at Berkeley, 1995.
[38] Young L.–S. *Statistical properties of dynamical systems with some hyperbolicity including certain billiards*, Ann. Math. **147** (1998) 585–650.
[39] Young L.–S. *Recurrence times and rates of mixing*, Isr. J. Math **110** (1999) 153–188.

DEPARTMENT OF MATHEMATICS, PENN STATE UNIVERSITY, STATE COLLEGE, PA 16802

E-mail address: dolgop@math.psu.edu