Ergod. Th. & Dynam. Sys. (2002), **22**, 129–152
DOI: 10.1017/S0143385702000056

Anosov diffeomorphisms and coupling

XAVIER BRESSAUD† and CARLANGELO LIVERANI‡

† *Instituto de Matematica e Estatistica, Universidade de São Paulo, Rua do Matão 1010, 05508-900 São Paulo SP, Brazil*
(e-mail: bressaud@ime.usp.br)

‡ *Dipartimento di Matematica, Universita' di Roma 'Tor Vergata', via della ricerca Scientifica, 00133 Rome, Italy*
(e-mail: liverani@axp.mat.uniroma2.it)

(*Received* 30 *June* 2000 *and accepted in revised form* 19 *October* 2000)

Abstract. We prove the existence of a Sinai–Ruelle–Bowen (SRB) measure and the exponential decay of correlations for smooth observables for mixing Anosov $C^{1+\alpha}$ diffeormorphisms on a d-dimensional ($d \geq 2$) Riemannian manifold. The novelty lies in the very simple method of proof. We construct explicitly a coupling between two initial densities so that under the action of the diffeomorphism, both components get closer and closer. The speed of this convergence can be explicitly estimated and is directly related to the speed of decay of the correlations. The existence of the SRB measure and its properties readily follow.

1. *Introduction*

The study of convergence to equilibrium in dynamical systems has witnessed an exciting revival in recent years. This is due to the introduction of new approaches that have supplemented the original one based on Markov partitions [**2**]. In particular, the approach based on quasi-compactness of the Perron–Frobenius operator has shed a new light on expanding maps [**1**]; the use of projective metrics has provided a general constructive tool that can be applied to a large variety of situations [**12, 13, 17, 20, 21**]; a specialized type of tower construction has introduced a new, more flexible, way of coding dynamical systems [**22**]; new insights into the properties and structure of the transfer operator for flows have finally allowed results for a vast class of Anosov flows to be obtained [**5, 6, 7**]; the use of random perturbations has proven an easy tool to obtain rough estimates on the rate of convergence to equilibrium [**9, 14, 15**].

In this paper we explore yet another approach. An approach well known in other fields but only recently introduced in this context by [**3, 4, 23**]—coupling. This is a technique used to compare stochastic processes that has yielded very important results in probability [**11**], statistical mechanics [**10**] and abstract ergodic theory [**18**], to mention just a few.

In this paper we apply it to Anosov diffeomorphisms. The result that we obtain is the exponential decay of correlations for a large (as far as we know larger than usual) class of functions. Of course, this result is far from new, nevertheless it should be emphasized that the method yields explicit estimates on the rate of convergence to equilibrium (unlike other methods)†. The main difference from the Lai-Sang Young approach is that the technique is applied directly to the system without the need to code it beforehand into an expanding tower.

The aim is twofold. On the one hand, we want to introduce the coupling method as a general tool to study the decay of correlations. On the other hand, we have tried to simplify as much as possible the method in order to understand exactly the technical conditions needed to make it work. This is essential preliminary work if one wants to apply coupling techniques to situations with polynomial decay of correlations which is certainly the next challenge in the field.

1.1. *Coupling.* Let X and Y be two random variables valued in $\mathcal{M}$ distributed according to smooth densities with respect to the natural measure on $\mathcal{M}$. The task is to construct a coupling $(\tilde{X}, \tilde{Y})$ of these random variables such that, in the mean, the distance between $T^n\tilde{X}$ and $T^n\tilde{Y}$ decays to zero. Since the dynamic is deterministic, the only possibility is to couple the initial distributions.

A coupling of two random variables is a joint distribution, that is a distribution on $\mathcal{M} \times \mathcal{M}$ with the marginals given by the two random variables. The simpler example is *independent coupling* (simply take the product distribution). If the random variables have the same law, another simple coupling is the *diagonal coupling* ($\tilde{X} = \tilde{Y}$ with probability one). Intermediate cases present some 'correlation' between the two random variables.

Let us think of discrete random variables for one moment. A coupling is a way of filling an array with a constraint on the sums of the lines and of the columns. Given two distributions $(\mu_i)_{i=1,\dots,n}$ and $(\nu_i)_{i=1,\dots,n}$, it is an array $(p_{i,j})_{i,j=1,\dots,n}$ with non-negative entries satisfying $\sum_j p_{i,j} = \mu_i$ and $\sum_i p_{i,j} = \nu_j$. Let us see how to construct a coupling with a strong correlated part. We can fill first the diagonal of the array in an arbitrary way imposing only that the value of each entry is less than the constrained values of the associated row and column (this will be the 'correlated part'). For example, choose a third distribution $(\eta_i)_{i=1,\dots,n}$ and an $\varepsilon > 0$ small enough so that for all i, $\varepsilon\eta_i < \min\{\mu_i, \nu_i\}$ and put the values $\varepsilon\eta_i$ on the diagonal‡. The total amount of mass 'used' is ε. The remainder can (among other possibilities) be coupled 'independently',

$$p_{i,j} = \chi_{\{i=j\}}\varepsilon\eta_i + (1-\varepsilon)^{-1}(\mu_i - \varepsilon\eta_i)(\nu_j - \varepsilon\eta_j).$$

Such an array is indeed a coupling of μ and ν. One can realize this coupling as follows. Drop a Bernoulli coin $(\varepsilon, 1-\varepsilon)$. If you get heads, then choose the two random variables according to the distribution η (in this case they take the same value). If you get

† As mentioned earlier, the only other technique that provides explicit bounds is the one based on Hilbert metrics [**21**]. Such an approach seems to yield convergence in a stronger sense but the present approach could prove more flexible.

‡ To be optimal, one could choose η so that $\varepsilon\eta_i = \min\{\mu_i, \nu_i\}$.

tails, then choose both random variables independently according to the 'remaining' joint distribution, $(1-\varepsilon)^{-2}(\mu_i - \varepsilon\eta_i)(\nu_j - \varepsilon\eta_j)$.

This last point of view easily generalizes to the case of continuous random variables with densities. Consider two density functions, g and h, bounded away from 0. Let f be a third density function and $\varepsilon > 0$ such that $\varepsilon f(x) < \min\{g(x), h(x)\}$. Now, drop a Bernoulli coin $(\varepsilon, 1-\varepsilon)$. If you get heads, then choose the two random variables according to the distribution with density f. If you get tails, then choose both random variables independently according to the 'remaining' joint distribution, whose density on the square is given by $F(x, y) = (1-\varepsilon)^{-2}(g(x) - \varepsilon f(x))(h(y) - \varepsilon f(y))$.

Back to our case, the general idea is to try to make points match whenever possible. Yet, since the map is a diffeomorphism, two points which do not match will never match. Nonetheless, two points that are on the same stable manifold will come closer and closer under the action of the diffeomorphism. Thus, it should suffice to couple points lying in the same stable manifold. A first obstacle stems from the highly non-local nature of the stable manifold which makes it hard to control the speed at which points get closer. To overcome this we will consider only local stable manifolds and proceed inductively.

Fix some $\delta > 0$. First, we decide to couple points that are in the same local stable manifold of size δ. Then, we couple points that will be in the same local stable manifold after one iteration of the diffeomorphism, and so on. What happens is that, at each step, we can couple a certain fraction of the mass that was not yet coupled. Finally, every point is coupled with points that will be in the same local stable manifold after some time (and hence that were in the same stable manifold since the beginning). Technically, the procedure consists in a sequence of couplings $(\widetilde{X}_n, \widetilde{Y}_n)$ of $T^n X$ and of $T^n Y$, constructed inductively, keeping track of the mass coupled at each step. In the limit all the mass is coupled. More precisely, for the coupling at step n the probability that the two components are in different parts of stable leaves of size δ is bounded by $(1-\varepsilon)^n$, decaying exponentially fast with n.

1.2. *Plan of the paper.* In §2 we state precisely the result and introduce some basic notation. Section 3 is devoted to the precise definition of the space of densities that will be used in the rest of the paper. In §4 we define special averages that will be used to actually construct the wanted coupling. In addition, some standard and less standard, but essential, parts of the theory of Anosov diffeomorphisms are recalled and stated. Section 5 investigates the properties of the densities when iterated via the dynamics. Such properties based on the results of §4 are the reason why the approach ultimately works. §6 introduces the coupling and details its properties. In §7 we have the proof of the main theorem and its corollary.

Finally, Appendix A contains the proof of some technical facts that are used in the paper and is added for completeness.

2. *Statement of the result*

Let $\mathcal{M}$ be a d-dimensional ($d \geq 2$) Riemannian manifold endowed with the Riemannian volume m. We consider a transitive (and hence mixing) Anosov $\mathcal{C}^{(1+\alpha)}(\mathcal{M}, \mathcal{M})$

diffeomorphism T on $\mathcal{M}$ by which we mean that T is a diffeomorphism of $\mathcal{M}$ whose differential D_xT at point $x \in \mathcal{M}$ depends α-Hölder continuously on x and is such that there exists an invariant splitting of the tangent bundle $\mathcal{T}_x\mathcal{M} = E^s(x) \bigoplus E^u(x)$ into a stable and an unstable direction, that is with $\|DT|_{E^s}\|_\infty < 1$ and $\|DT^{-1}|_{E^u}\|_\infty < 1$. We denote by d_s and d_u the dimensions of the stable and unstable subspaces, respectively.

We denote by D the distortion constant,

$$e^{-Dd(x,y)^\alpha} \le \frac{|\det(D_xT)|}{|\det(D_yT)|} \le e^{Dd(x,y)^\alpha}, \tag{2.1}$$

where $d(\cdot,\cdot)$ is the Riemannian distance. We set

$$\begin{aligned} \|DT|_{E^s}\|_\infty = \lambda_-^{-1}; \quad \|DT|_{E^u}\|_\infty = \mu_+; \\ \|DT^{-1}|_{E^s}\|_\infty = \lambda_+; \quad \|DT^{-1}|_{E^u}\|_\infty = \mu_-^{-1}. \end{aligned} \tag{2.2}$$

2.1. *Stable and unstable foliations.* For all $x \in \mathcal{M}$, we denote by $W^s(x)$ and $W^u(x)$ the global stable and unstable manifolds (such manifolds are $\mathcal{C}^{(1+\alpha)}$, e.g. [**8**]). For all $x \in \mathcal{M}$ and all $y \in W^s(x)$, we denote by $d^s(x, y)$ the distance measured along the leaves $W^s(x)$ induced by the Riemannian metric on the leaf considered as a submanifold in $\mathcal{M}$. In addition, for each $x, y \in \mathcal{M}$ define $d^s(x, y) = \infty$ if they do not belong to the same stable manifold. The corresponding distance for the unstable manifolds will be denoted by d^u. In the same spirit, m^s will denote the restriction of the Riemannian volume to the stable manifolds. For all $\delta > 0$ and all $x \in \mathcal{M}$, we will denote by $W^s_\delta(x)$ the ball of radius δ centered at x in $W^s(x)$,

$$W^s_\delta(x) = \{y \in \mathcal{M} \mid d^s(x, y) < \delta\}.$$

2.2. *Observables.* Fix $\delta > 0$ and $\beta_s \in (0, 1)$. For all real functions $f : \mathcal{M} \to \mathbb{R}$, we set

$$|f|_s = \sup_{d^s(x,y)\le\delta} \frac{|f(x) - f(y)|}{d^s(x, y)^{\beta_s}}$$

and

$$\|f\|_s = \|f\|_\infty + |f|_s. \tag{2.3}$$

We consider the following subset of the Borel measurable functions $\mathcal{B}(\mathcal{M}, \mathbb{R})$,

$$\mathcal{C}_s = \{f \in \mathcal{B}(\mathcal{M}, \mathbb{R}) \mid \|f\|_s < +\infty\}.$$

2.3. *Statement.* We are interested in the convergence to equilibrium, that is in the speed with which an initial measure, absolutely continuous with respect to the Riemannian measure, converges toward the SRB measure. As is well known, only reasonably smooth initial measures yield fast convergence. To state the result we must then introduce, for the densities of the initial measure, Hölder norms $\|\cdot\|_u$ defined in analogy with (2.3) by using the unstable distance instead of the stable one,

$$\|f\|_u = \|f\|_1 + \sup_{d^u(x,y)\le\delta} \frac{|f(x) - f(y)|}{d^u(x, y)^\alpha}. \tag{2.4}$$

Note the L^1-norm instead of the L^∞-norm to take advantage of the obvious duality.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

THEOREM 2.1. *There exists a $\theta < 1$, computable, and a constant C such that, for each $\beta_s \in (0, 1]$, for all $f \in \mathcal{C}_s$ and $g, h \in \mathcal{C}^{(\alpha)}(\mathcal{M}, \mathbb{R})$ with $\int_{\mathcal{M}} h\, dm = \int_{\mathcal{M}} g\, dm = 1$,*

$$\left| \int_{\mathcal{M}} f \circ T^n g\, dm - \int_{\mathcal{M}} f \circ T^n h\, dm \right| \leq C \|f\|_s \max\{\|g\|_u, \|h\|_u\} \theta^{n\beta_s}. \tag{2.5}$$

From this result it is easy to deduce the existence of an SRB measure for the system as well as the exponential decay of the correlations for this measure. This is the content of the following corollary.

COROLLARY 2.1. *There is a unique T-invariant measure μ such that*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k x} = \mu, \quad m\text{-a.s.}, \tag{2.6}$$

where δ_x denotes the Dirac mass at point x. In addition, there is a $\theta < 1$ and a constant C such that, for all $f \in \mathcal{C}_s$ and all $h \in \mathcal{C}^{(\alpha)}(\mathcal{M}, \mathbb{R})$,

$$\left| \int_{\mathcal{M}} f \circ T^n h\, d\mu - \int_{\mathcal{M}} f\, d\mu \int_{\mathcal{M}} h\, d\mu \right| \leq C \|f\|_s \|h\|_u \theta^n. \tag{2.7}$$

In fact, the other standard properties of the SRB measure can also be easily obtained; we do not insist on them since the present method does not seem to add any particular insight on such issues.

3. *Densities*

To prove these theorems we find it necessary to specify more precisely the regularity requirements for the smoothness of the densities along the unstable manifold. It turns out that in two dimensions or, more generally, when the foliations are smooth, this can be done in a rather naïve way by defining smoothness with respect to the metric d^u as already done in (2.4)†. Nonetheless, in the more general situation in which the foliations are only Hölder it is not immediately obvious how to proceed. As we will see later the problem arises because if $d^u(x, y) \leq \varepsilon$, then the maximal distance of the associated local stable manifolds $W^s_\delta(x)$ and $W^s_\delta(y)$ can be of order ε^τ which is not enough for our needs. To overcome this we introduce a different notion of 'distance' between points, which, in fact, is not a distance (it does not satisfy the triangle inequality and it is degenerate) but is nevertheless well suited to satisfy our needs.

Let $\varphi : \mathbb{R}^+ \to [0, 1]$ be a smooth positive monotone function such that, for some $\delta > 0$,

$$\varphi(x) = 0\ \forall x \in [0, \delta] \quad \text{and} \quad \varphi(x) = 1\ \forall x \geq 2\delta.$$

Definition 3.1. Fix $\nu > 1$. Given two points $x, y \in \mathcal{M}$ we define their *separation* $s(x, y)$ by

$$s(x, y) = \sum_{n=0}^{\infty} \nu^{-n} \varphi(d(T^n x, T^n y)). \tag{3.1}$$

† To pursue this point of view it is also necessary to impose an *a priori* bound on the measure of small balls in the stable manifold; this is a natural condition since the invariant measure is typically singular when restricted to the stable manifold and it is characterized by some Hausdorff dimension.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

A direct computation shows that if $\mu_- < \nu$ this separation is a Lipschitz one, while if $\mu_- > \nu$, it is a Hölder one of exponent $\log \nu / \log \mu_-$. The relation between the separation and the usual distance as well as some useful properties of the separation are clarified by the next lemma.

LEMMA 3.1. *There exists $c_0 = c_0(\delta, \nu) \in \mathbb{R}^+$ such that for each two points $x, y \in \mathcal{M}$, if $d(x, y) \le \delta$, then*

$$s(Tx, Ty) = \nu s(x, y).$$

If $x \in W^u_\delta(y)$, then

$$c_0^{-1} d^u(x, y)^{\alpha_+} \le s(x, y) \le c_0 d(x, y)^{\alpha_-},$$

with $\alpha_- = \ln \nu / \ln \mu_+$ and $\alpha_+ = \ln \nu / \ln \mu_-$. In addition, if $y \in B_{\delta/4}(x)$,

$$c_0^{-1} s(x, y) \le \sup_{\substack{z \in W^s_{\delta/4}(x) \\ w \in W^s_{\delta/4}(y)}} s(z, w) \le c_0 s(x, y).$$

The proof of the lemma is by direct computation. The idea is that s is essentially a continuous version of the usual discrete separation. For completeness we provide the details in Appendix A.

To define the smoothness of a function along the unstable directions we can then define

$$|f|_\sigma = \sup_{d^u(x,y) \le \delta} \frac{|f(x) - f(y)|}{s(x, y)}. \tag{3.2}$$

The relation with the previously mentioned more conventional definitions of smoothness

$$|f|_{u,\beta} = \sup_{d^u(x,y) \le \delta} \frac{|f(x) - f(y)|}{d^u(x, y)^\beta}$$

(which characterizes Hölder continuous functions) is explicitly stated in the following immediate consequence of Lemma 3.1.

COROLLARY 3.1. *Let $\nu \le \mu_-$, then $\alpha_-, \alpha_+ \in (0, 1]$ and*

$$|f|_{u,\alpha_-} \le |f|_\sigma \le |f|_{u,\alpha_+}.$$

We can finally define the class of densities we are interested in

$$\mathcal{C}_u(a) = \left\{ h \in \mathcal{C}^0(\mathcal{M}, \mathbb{R}) \;\middle|\; h > 0; \frac{h(x)}{h(y)} \le e^{as(x,y)}, \forall x, y : d^u(x, y) \le \delta \right\}, \tag{3.3}$$

where $\mathcal{C}^0(\mathcal{M}, \mathbb{R})$ is the space of continuous functions.

4. *Mean on the stable leaves*

4.1. *Holonomy map.* Given two manifolds U, V transversal to the stable foliation, with a distance less than δ, we define the (stable) holonomy map $\phi = \phi^s_{U,V} : U \to V$ by $\phi(z) = W^s_\delta(z) \cap V$†. We define the unstable holonomies symmetrically.

† Of course, the domain of ϕ is given exactly by the $z \in U$ for which $W^s_\delta(z) \cap V \ne \emptyset$.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

Given $z \in U$, let $E_U(z) = \mathcal{T}_z U$. Since $E_U(z)$ is transversal to $E^s(z)$ we can define the map $L_{E_U(z)} : E^u(z) \to E^s(z)$ by asking that $L_{E_U(z)} v$ be the unique vector $w \in E^s(z)$ such that $v + w \in E_U(z)$. Let $\gamma_U : U \to \mathbb{R}$ be

$$\gamma_U(z) = \|L_{E_U(z)}\|,$$

clearly this quantity measures the 'angle' between $E_U(z)$ and $E^u(z)$. Similar definitions can be given for manifolds transversal to the unstable distribution.

PROPOSITION 4.1. *There exists $\tau \in (0, \alpha]$ such that the stable and unstable distributions are τ-Hölder continuous. In addition, the holonomy maps are τ-Hölder continuous and absolutely continuous as well. Finally, calling $J\phi_{U,V}$ the Jacobian of the holonomy, the Jacobian is also τ-Hölder continuous; more precisely,*

$$\frac{J\phi_{UV}(z)}{J\phi_{UV}(\xi)} \leq e^{-\text{const.}\, d(z,\xi)^\tau} \quad \text{if } z, \xi \in U;$$

and it is close to the identity, namely

$$e^{-\text{const.}\,\{d(z,\phi(z))^\tau + \gamma_U(z) + \gamma_V(z)\}} \leq J\phi_{UV}(z) \leq e^{\text{const.}\{d(z,\phi(z))^\tau + \gamma_U(z) + \gamma_V(z)\}}.$$

Proof. These are classical results. See, for example, [**8**] for the Hölder continuity of the distributions and of the holonomies and [**16**] for the bounds on the Jacobian. □

4.2. *The local product structure.* Given $x_0 \in \mathcal{M}$, we want to introduce local coordinates, adapted to the hyperbolic structure, in a neighborhood of x_0. Given $\delta_0 > 0$ small enough, we start by considering the manifolds $W^s_{\delta_0}(x_0)$ and $W^u_{\delta_0}(x_0)$ endowed with the restriction of the Riemannian metric on $\mathcal{M}$. Since they are $\mathcal{C}^{(1+\alpha)}$, we can introduce two $\mathcal{C}^{(1+\alpha)}$ systems of coordinates $\psi^s_{x_0} : \mathbb{R}^{d_s} \to W^s_{\delta_0}(x_0)$ and $\psi^u_{x_0} : \mathbb{R}^{d_u} \to W^u_{\delta_0}(x_0)$†. We can then define the map $\psi_{x_0} : \mathbb{R}^{d_s} \times \mathbb{R}^{d_u} \to B_{\delta_0}(x_0)$ by

$$\psi_{x_0}(\xi, \eta) = W^u_{\delta_0}(\psi^s_{x_0}(\xi)) \cap W^s_{\delta_0}(\psi^u_{x_0}(\eta)).$$

Note that in the coordinates η, ξ the stable and unstable foliations consist of the linear parallel subspaces $\{\xi = a\}$ and $\{\eta = b\}$ respectively. In addition, note that, by Proposition 4.1, ψ_{x_0} is a Hölder, and hence continuous, change of coordinates. We consider the Lebesgue measure $m_L(d\xi, d\eta) = d\xi\, d\eta$. It is also natural to consider that the Riemannian volume m induces the measure $\bar{m} = \psi^{-1*}_{x_0} m$ in $\mathbb{R}^{d_s} \times \mathbb{R}^{d_u}$. It turns out that $\bar{m}$ is absolutely continuous with respect to the Lebesgue measure and that the Jacobian is a rather nice function as the next proposition more precisely states.

PROPOSITION 4.2. *There exist $c, M, \delta_0 > 0$ such that, for each $x_0 \in \mathcal{M}$, one can construct a measurable function $\rho_{x_0} : \mathbb{R}^d \to \mathbb{R}^+$ for which, given any function $f \in L^1(\mathcal{M}, \mathbb{R})$ supported in $B_{\delta_0}(x_0)$,*

$$\int_{\mathcal{M}} f(y) m(dy) = \int_{\mathbb{R}^d} f \circ \psi_{x_0}(\xi, \eta) \rho_{x_0}(\xi, \eta)\, d\xi\, d\eta$$

† If the manifold were $\mathcal{C}^2$, one could use the standard exponential mappings, identifying, isometrically, $E^u(x_0)$ with $\mathbb{R}^{d_u}$ and $E^s(x_0)$ with $\mathbb{R}^{d_s}$.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

holds. This function has the following properties:

$$e^{-c|\eta-\eta'|^\tau} \le \frac{\rho_{x_0}(\xi,\eta)}{\rho_{x_0}(\xi,\eta')} \le e^{c|\eta-\eta'|^\tau}, \quad \forall \xi \in \mathbb{R}^{d_s},\ \eta,\eta' \in \mathbb{R}^{d_u},$$

and

$$\frac{1}{M} \le \rho_{x_0} \le M.$$

This result is more or less known to experts in the field, yet we were not able to locate a clear cut reference to it. Due to this sorry state of affairs we provide a complete proof in Appendix A.

4.3. *Stable averages.* We introduce an average on the local stable manifolds $W^s_\delta(x)$, namely $\mathbb{A}_\delta : \mathcal{C}^0(\mathcal{M},\mathbb{R}) \to \mathcal{C}^0(\mathcal{M},\mathbb{R})$ defined by

$$\mathbb{A}_\delta f(x) = Z_\delta(x) \int_{W^s_\delta(x)} f(z) m^s(dz)$$

where

$$Z_\delta(x)^{-1} = \int_{W^s_\delta(x)} m^s(dz) = m^s(W^s_\delta(x)).$$

The following lemma is the key fact we will use in studying the conditional expectation we will define shortly.

LEMMA 4.1. *For all $\delta > 0$, sufficiently small, there exists $\varepsilon_0 = \varepsilon_0(\delta,a) > 0$ such that, for all $h \in \mathcal{C}_u(a)$,*

$$\varepsilon_0 \int_{\mathcal{M}} h \le \mathbb{A}_\delta h \le \frac{1}{\varepsilon_0} \int_{\mathcal{M}} h. \tag{4.1}$$

Proof. If $\delta < \delta_0$, then we can write $\mathbb{A}_\delta h$ in some local coordinates and use regularity of the observable to compare it to the integral on an open set of positive volume. In particular, we can choose any $\mathcal{C}^{(1+\alpha)}$ coordinate systems $\tilde\psi : \mathbb{R}^{d_s} \times \mathbb{R}^{d_u} \to \mathcal{M}$ such that $\tilde\psi(0) = x$, $\tilde\psi(\{(\xi,0)\}) = W^s_{\text{loc}}(x)$; $\tilde\psi(\{(0,\eta)\}) = W^u_{\text{loc}}(x)$ and $\|D\psi\|_\infty + \|D\psi^{-1}\|_\infty \le M_1$ for some M_1 independent of x†. Let $\tilde\psi(W^s_\delta(x)) := A(x)$ and $\tilde\psi(W^u_\delta(x)) := B(x)$, then const.$^{-1}\delta^{d_u} \le \int_{B(x)} d\eta \le$ const. δ^{d_u}, for some constant independent of x, and

$$\begin{aligned}
\int_{W^s_\delta(x)} h &= \int_{A(x)} h \circ \psi_x(\xi,0) |\det(D_\xi \tilde\psi)(\xi,0)|\, d\xi \\
&\le \frac{e^{M_1(d+d_s)} e^{ac_0\delta^{\alpha+}}}{\text{const.}\, C\delta^{d_u}} \int_{A(x)\times B(x)} h \circ \tilde\psi(\xi,\eta) |\det(D\tilde\psi)(\xi,\eta)|\, d\eta\, d\xi \\
&\le \text{const.} \int_{\tilde\psi(A(x)\times B(x))} h \\
&\le \text{const.} \int_{\mathcal{M}} h.
\end{aligned} \tag{4.2}$$

† If in doubt about existence, see the footnote on p. 147 in Appendix A for the explicit construction of such a coordinate systems.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

For further reference note that the same type of reasoning shows that there exists $c_1 > 1$, independent of $x \in \mathcal{M}$, such that

$$\int_{W^s_{c_1\delta}(x)} h(z) m^s(dz) \geq \text{const.} \int_{B_{2\delta}(x)} h, \tag{4.3}$$

where, as usual, $B_\delta(x)$ is the ball of radius δ centered at x.

The proof of the opposite inequality is slightly more sophisticated.

It is well known that there exists N_0 such that, for each $\varepsilon > 0$, one can cover $\mathcal{M}$ with balls of radius ε in such way that each ball overlaps with less than N_0 other balls of the cover. From now on we will only consider covers with such a property.

Let $\{B_\varepsilon(x_i)\}$ be one such cover. For each function $h \geq 0$ there exists i such that

$$\frac{1}{m(B_\varepsilon(x_i))} \int_{m(B_\varepsilon(x_i))} h \geq N_0^{-1} \int_{\mathcal{M}} h.$$

To see this, just suppose the opposite, then

$$\int_{\mathcal{M}} h \leq \sum_j \int_{m(B_\varepsilon(x_j))} h < N_0^{-1} \sum_j m(B_\varepsilon(x_i)) \int_{\mathcal{M}} h \leq \int_{\mathcal{M}} h$$

which is a contradiction. Let us fix ε such that $c_1\varepsilon < \delta_0$ and $\{B_\varepsilon(x_i)\}$ a cover with the earlier property.

It is a standard consequence of topological mixing (or transitiveness) and the Anosov property that there is a L_ε such that any piece of stable manifold of (inner) diameter L_ε is at a distance less than ε from any point of $\mathcal{M}$. Let us fix n such that $\lambda_-^n\delta > 2L_\varepsilon$. We write

$$\begin{aligned}\int_{W^s_\delta(x)} h &= \int_{T^{-n}W^s_\delta(x)} h \circ T^n |\det(DT^n|_{E^s})| \\ &\geq \lambda_+^{-nd_s} \int_{T^{-n}W^s_\delta(x)} h \circ T^n.\end{aligned}$$

Since $h \circ T^n$ is positive, we can choose $B_\varepsilon(x_i)$ in the cover such that

$$\frac{1}{m(B_\varepsilon(x_i))} \int_{B_\varepsilon(x_i)} h \circ T^n \geq \frac{1}{N_0} \int_{\mathcal{M}} h \circ T^n. \tag{4.4}$$

Since the inner diameter of $T^{-n}W^s_\delta(x)$ is larger than $2L_\varepsilon$, there exists a part $W^s_{c_1\varepsilon}(z) \subset T^{-n}W^s_\delta(x)$, of inner diameter $c_1\varepsilon$, centered on a point $z \in W^u_\varepsilon(x_i)$. Remember that our choice of c_1 is such that $\psi_z^{-1}(B_{2\varepsilon}(z)) \subset C_{c_1\varepsilon}(z) := \psi_z^{-1}(W^u_{c_1\varepsilon}(z)) \times \psi_z^{-1}(W^s_{c_1\varepsilon}(z))$. Accordingly, remembering (4.3) and (4.4),

$$\begin{aligned}\int_{W^s_\delta(x)} h \geq \lambda_+^{-nd_s} \int_{W^s_{c_1\varepsilon}(z)} h \circ T^n &\geq \lambda_+^{-nd_s} \int_{B_{2\varepsilon}(z)} h \circ T^n \\ &\geq \text{const.} \int_{B_\varepsilon(x_i)} h \circ T^n \\ &\geq \text{const.} \int_{\mathcal{M}} h.\end{aligned}$$

Both estimates hold in particular for $h = 1$ and hence for Z_δ. So the ratios defining $\mathbb{A}_\delta h$ are uniformly bounded. □

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

Remark 4.1. Note that the only dynamical property used in Lemma 4.1 (apart from the existence and regularity of the stable and unstable manifolds) is the transitiveness.

For the purposes of the next section it is necessary to define an averaging operator that preserves regularity along the unstable manifold. This is the case for the operator $\mathbb{A}_\delta$ only if the difference between λ_- and λ_+ is sufficiently small†. To deal with the general case it is necessary to define a different (more local) average. The definition of such an operator and its properties are the focus of the rest of the section.

Pick any $\bar{x} \in \mathcal{M}$ and consider

$$R = \bigcup_{z \in W^u_\delta(\bar{x})} W^s_{\delta(z)}(z). \tag{4.5}$$

where $\text{const.}^{-1}\delta \leq \delta(z) \leq \text{const.}\,\delta$, for some constant independent of $\bar{x}$, and $\int_{\psi_{\bar{x}}^{-1}(W^s_{\delta(z)}(z))} d\xi = \delta^{d_s}$.

By construction R is foliated by stable manifolds and, provided δ is small enough, there exists $c_* \in (0, 1)$ such that

$$B_{c_*\delta}(\bar{x}) \subset R \subset B_{c_*^{-1}\delta}(\bar{x}) \subset B_{\delta_0}(\bar{x}).$$

Next, consider the σ-algebra $\mathcal{F}$ associated to the partition $\{\{W^s_{\delta(z)}(z)\}_{z\in W^u_\delta(\bar{x})}, R^c\}$.

Finally, choose a smooth function Φ such that

$$\Phi \geq 0; \int_{\mathcal{M}} \Phi = 1; \operatorname{supp}\Phi \subset B_{c_*\delta}(\bar{x}).$$

Define

$$\mathbb{A}_\Phi f(x) = \Phi(x)\mathbb{E}(\Phi f \mid \mathcal{F})(x), \tag{4.6}$$

where $\mathbb{E}(\cdot \mid \mathcal{F})$ denotes the conditional expectation given $\mathcal{F}$, with respect to the probability measure m. Note that, using the local coordinates around $\bar{x}$ introduced in §4.2, we can write a nice version of this conditional expectation:

$$\mathbb{A}_\Phi f(\psi_{\bar{x}}(\xi,\eta)) = \Phi(\psi_{\bar{x}}(\xi,\eta)) \int_{\{\|\xi'\|\leq\delta\}} (\Phi f)\circ\psi_{\bar{x}}((\xi',\eta)) \frac{\rho_{\bar{x}}(\xi',\eta)}{\rho_{\bar{x}}(\xi,\eta)\delta^{d_s}}\, d\xi'. \tag{4.7}$$

In the following, we shall omit the $\psi_{\bar{x}}$ if no confusion arises. Clearly, if f has support disjoint from R, then $\mathbb{A}_\Phi f = 0$. Other instrumental properties of $\mathbb{A}_\Phi$ are summarized by the following lemma.

LEMMA 4.2. *Assume $\nu \leq \mu_-^\tau$, so that $\tau \geq \alpha_+$, then the following statements hold:*
- $\mathbb{E}(g\mathbb{A}_\Phi f) = \mathbb{E}(f\mathbb{A}_\Phi g)$;
- *there exists $\varepsilon_1 = \varepsilon_1(a,\delta) \in \mathbb{R}$ such that, for all $f, g \in \mathcal{C}_u(a)$,*

$$\|\mathbb{A}_\Phi f\|_\infty \leq \varepsilon_1^{-1}\int_{\mathcal{M}} f \quad \textit{and} \quad \int_{\mathcal{M}} g\mathbb{A}_\Phi f \geq \varepsilon_1 \int_{\mathcal{M}} f \int_{\mathcal{M}} g;$$

† There are two sources of trouble. On the one hand, if the foliation is only Hölder, then the image, under the holonomy, of a ball, in the unstable manifold, is no longer a ball. We know how to handle this mismatch only if the difference between λ_+ and λ_- is small enough. On the other hand, if $y \in W^s_\delta(x)$, then the stable distance, measured along the holonomy, between $W^u_\delta(x)$ and $W^u_\delta(y)$ can be of order $d(x,y)^\beta$, where β also depends on $\lambda_-\lambda_+^{-1}$ and can be <1.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

- *there exists* $\kappa_1 = \kappa_1(a, \delta) > 0$ *such that, for all* $f \in \mathcal{C}_u(a)$,

$$|\mathbb{A}_\Phi f|_\sigma \le \kappa_1 \int_{\mathcal{M}} f.$$

Proof. The first statement follows directly by the properties of the conditional expectation. The second assertion follows from Lemma 4.1. In fact, formula (4.7) shows that $\mathbb{A}_\Phi$ is constructed via stable averages; thus the same arguments as in Lemma 4.1 can be used to obtain upper and lower bounds for such averages (formulae (4.2) and (4.3)).

For the last part of the lemma we must compute

$$\Phi(x)\mathbb{E}(\Phi f \mid \mathcal{F})(x) - \Phi(y)\mathbb{E}(\Phi f \mid \mathcal{F})(y).$$

This is most easily done by using in R the coordinates introduced just before Proposition 4.2. Let (ξ_1, η_1) be the coordinates of x and (ξ_2, η_2) the coordinates of y. Thus we have

$$\begin{aligned}
&|\Phi(x)\mathbb{E}(\Phi f \mid \mathcal{F})(x) - \Phi(y)\mathbb{E}(\Phi f \mid \mathcal{F})(y)| \\
&\quad\le \text{const.}\, s(x, y)\Phi(y)\mathbb{E}(\Phi f \mid \mathcal{F})(y) \\
&\quad\quad + \Phi(x) \int_{\{\|\xi\| \le \delta\}} |\Phi(\xi, \eta_1) f(\xi, \eta_1)\rho_{\bar{x}}(\xi, \eta_1) - \Phi(\xi, \eta_2) f(\xi, \eta_2)\rho_{\bar{x}}(\xi, \eta_2)|\, d\xi.
\end{aligned}$$

But, if $z = \psi(\xi, \eta_1)$ and $w = \psi(\xi, \eta_2)$,

$$\begin{aligned}
|\Phi f \rho_{\bar{x}}(z) - \Phi f \rho_{\bar{x}}(w)| &\le \|\Phi\|_\infty \,\text{const.}(as(z, w) + cd(z, w)^\tau) f\rho_{\bar{x}}(z) \\
&\quad + f\rho_{\bar{x}}(w)|\Phi|_\sigma s(z, w) \\
&\le \text{const.}(f\rho_{\bar{x}}(z) + f\rho_{\bar{x}}(w))s(z, w).
\end{aligned}$$

We can now conclude since, by the third statement of Lemma 3.1, we can control the distance of the two stable manifolds. The integral $\int_{\mathcal{M}} f$ appears by Lemma 4.1. □

Remark 4.2. Concerning possible attempts to generalize the present scheme: note that the last two points of Lemma 4.2 are the only points in our construction where the properties of the holonomies (and their Jacobians) play a role.

5. *Regularity properties of the densities*

5.1. *Losing regularity.* The following lemma gives an estimate on the regularity of $g(1 - \varepsilon\mathbb{A}_\Phi h)$ given that g and h are in $\mathcal{C}_u(a)$ and ε is small enough.

LEMMA 5.1. *Choose a real number* a_0. *For all* g *and* h *in* $\mathcal{C}_u(a)$, *with* $\int_{\mathcal{M}} h = 1$,

$$g(1 - \varepsilon\mathbb{A}_\Phi h) \in \mathcal{C}_u(a + a_0) \tag{5.1}$$

provided $\varepsilon < \min\{\varepsilon_1, a_0\kappa_1^{-1}(1 - \varepsilon\varepsilon_1^{-1})\}$.

Proof. It follows directly by Lemma 4.2. We have

$$\begin{aligned}
g(y)(1 - \varepsilon\mathbb{A}_\Phi h(y)) &\le g(x)e^{as(x,y)}(1 - \varepsilon\mathbb{A}_\Phi h(x)) \left(1 + \varepsilon\frac{\mathbb{A}_\Phi h(x) - \mathbb{A}_\Phi h(y)}{1 - \varepsilon\mathbb{A}_\Phi h(x)}\right) \\
&\le g(x)(1 - \varepsilon\mathbb{A}_\Phi h(x))e^{as(x,y)} \left(1 + \varepsilon\frac{|\mathbb{A}_\Phi h|_\sigma}{1 - \varepsilon\|\mathbb{A}_\Phi h\|_\infty} s(x, y)\right)
\end{aligned}$$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

$$\leq g(x)(1 - \varepsilon \mathbb{A}_\Phi h(x)) e^{as(x,y)} \left(1 + \varepsilon \frac{\kappa_1}{1 - \varepsilon \varepsilon_1^{-1}} s(x,y)\right)$$

$$\leq g(x)(1 - \varepsilon \mathbb{A}_\Phi h(x)) \exp\left[\left(a + \varepsilon \frac{\kappa_1}{1 - \varepsilon \varepsilon_1^{-1}}\right) s(x,y)\right]. \qquad \square$$

5.2. *Recovering regularity.* The diffeomorphism T has regular derivatives. Recall that D is the distortion constant (cf. (2.1)).

The Perron–Frobenius operator is given by†

$$\mathcal{L}g(x) = |\det(D_x T^{-1})| g \circ T^{-1}(x).$$

It satisfies the following lemma.

LEMMA 5.2. *Assume $\nu \leq \mu_-^\alpha$, so that $\alpha \geq \alpha_+$, then*

$$\mathcal{L}(\mathcal{C}_u(a)) \subset \mathcal{C}_u(\nu^{-1} a + c_0 D). \tag{5.2}$$

Proof. Clearly, for all $g \in \mathcal{C}_u(a)$, $\mathcal{L}g > 0$. For $x \in \mathcal{M}$ and $y \in W^u(x)$, $d(x,y) \leq \delta$, by Lemma 3.1,

$$\begin{aligned}\mathcal{L}g(x) &= |\det(D_x T^{-1})| g(T^{-1}x) \\ &\leq |\det(D_y T^{-1})| e^{Dd(x,y)^\alpha} g(T^{-1}y) e^{as(T^{-1}x, T^{-1}y)} \\ &\leq \mathcal{L}g(y) e^{(a\nu^{-1} + c_0 D)s(x,y)}.\end{aligned} \qquad \square$$

6. *Coupling*

We fix a $\delta > 0$ small enough and $\nu \leq \mu_-^\tau$, where τ is given by Proposition 4.1. We choose a_0 and set $a = (\nu^{-1} a_0 + c_0 D)/(1 - \nu^{-1})$. Let g and h be in $\mathcal{C}_u(a)$ with $\|g\|_1 = \|h\|_1 = 1$. We consider two independent random variables X and Y, on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, valued in $\mathcal{M}$ and distributed according to, respectively, $g\,dm$ and $h\,dm$. That is, $\mathbb{P}(X \in A) = \int_A g(x)\,dm(x)$ and $\mathbb{P}(Y \in A) = \int_A h(x)\,dm(x)$.

We set‡

$$\theta > \max\left\{\lambda_-^{-1}, 1 - \varepsilon_1^2, 1 - \frac{a_0 \varepsilon_1^2}{a_0 + \kappa_1 \varepsilon_1}\right\}.$$

Our key estimate consists in the following statement.

† Remember that if the measure μ on $\mathcal{M}$ is absolutely continuous with respect to the Riemannian volume m with density g, then the evolved measure $T_*\mu$, defined as usual by $T_*\mu(f) = \mu(f \circ T)$, is absolutely continuous with respect to m as well and its density is given by $\mathcal{L}g$.

‡ Since the choice of a_0 is arbitrary, the last term is not harmful. It should also be remarked that the second term can be improved. To do so consider n disjoint sets R_i in which to couple (rather than just one set R). To each set will be associated its operator $\mathbb{A}_{\Phi_i}$. Define $\mathbb{A}_\Phi = \sum_i \mathbb{A}_{\Phi_i}$. Clearly $\int \mathbb{A}_\Phi hg \geq n\varepsilon_1 \int g$ while $\mathbb{A}_\Phi h \leq \varepsilon_1^{-1} \int h$. Following the proof of Proposition 6.1 yields the fact that $1 - \varepsilon_1^2$ can be substituted with $1 - n\varepsilon_1^2$. We did not pursue this possibility because, most likely, the estimate we obtain is anyhow not optimal and if one wants to improve the bounds realistically, then more work is certainly needed.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

PROPOSITION 6.1. *There exists a constant C such that, for all $n \geq 0$, one can construct a coupling $(\widetilde{X}_n, \widetilde{Y}_n)$ of $T^n X$ and $T^n Y$ such that*

$$\mathbb{E}[d_\delta^s(\widetilde{X}_n, \widetilde{Y}_n)] \leq C\theta^n,$$

where $d_\delta^s(x, y) = \min\{d^s(x, y), \delta\}$.

Proof. Given two density functions $h, g \in \mathcal{C}_u(a)$ the basic idea to construct a coupling between the corresponding two random variables (measures) is to introduce the auxiliary random variable τ on Ω with values in $\{0, 1\}$, independent of X and Y, having the distribution†

$$\mathbb{P}(\tau = 0) = \varepsilon,$$

where $\varepsilon = 1 - \theta$. We set

$$\delta_{\Phi,h,g}(A \times B) = \frac{\int_{\mathcal{M}} \chi_A g \mathbb{A}_\Phi(h\chi_B)}{\int_{\mathcal{M}} g \mathbb{A}_\Phi h}. \tag{6.1}$$

We can then define the coupling $(\hat{X}, \hat{Y})$ by

$$\mathbb{P}(\hat{X} \in A; \hat{Y} \in B \mid \tau = 0) = \delta_{\Phi,h,g}(A \times B)$$
$$\mathbb{P}(\hat{X} \in A; \hat{Y} \in B \mid \tau = 1) = (1 - \varepsilon)^{-2}[\mathbb{P}(X \in A) - \varepsilon\mathbb{P}(\hat{X} \in A \mid \tau = 0)] \times [\mathbb{P}(Y \in B) - \varepsilon\mathbb{P}(\hat{Y} \in B \mid \tau = 0)].$$

Note that the distribution of $(\hat{X}, \hat{Y})$, conditioned to the absence of coupling (that is to $\{\tau = 1\}$), is a product measure absolutely continuous with respect to the Riemannian volume. A direct computation shows that this is indeed a coupling (i.e. $\mathbb{P}(\hat{X} \in A) = \mathbb{P}(X \in A)$ and $\mathbb{P}(\hat{Y} \in B) = \mathbb{P}(Y \in B)$). It is now natural to define

$$\widetilde{X}_0 = X; \quad \widetilde{Y}_0 = Y; \quad \widetilde{X}_1 = T\hat{X}; \quad \widetilde{Y}_1 = T\hat{Y}.$$

An obvious computation shows that

$$\mathbb{P}(\widetilde{X}_1 \in A; \widetilde{Y}_1 \in B \mid \tau = 1) = \int_A \tilde{g}\, dm \int_B \tilde{h}\, dm,$$

where

$$\tilde{g} = (1 - \varepsilon)^{-1}\mathcal{L}\left(g\left(1 - \frac{\varepsilon}{\int_{\mathcal{M}} g\mathbb{A}_\Phi h}\mathbb{A}_\Phi h\right)\right),$$

and

$$\tilde{h} = (1 - \varepsilon)^{-1}\mathcal{L}\left(h\left(1 - \frac{\varepsilon}{\int_{\mathcal{M}} h\mathbb{A}_\Phi g}\mathbb{A}_\Phi g\right)\right).$$

This is consistent since, according to Lemma 4.2,

$$\int_{\mathcal{M}} h\mathbb{A}_\Phi g = \int_{\mathcal{M}} g\mathbb{A}_\Phi h \geq \varepsilon_1,$$

† Intuitively if $\tau(\omega)$ is zero then the two points $X(\omega)$ and $Y(\omega)$ will be coupled and if it is one then they are left independent. For example, one can imagine that a coin (loaded accordingly to the distribution $\{\varepsilon, 1-\varepsilon\}$) is tossed in order to decide whether to couple the points or not.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

and

$$1-\varepsilon\frac{\mathbb{A}_\Phi h}{\int_{\mathcal{M}} g\mathbb{A}_\Phi h}\geq 1-\frac{\varepsilon}{\varepsilon_1}\mathbb{A}_\Phi h\geq 1-\frac{\varepsilon}{\varepsilon_1}\varepsilon_1^{-1}>0.$$

Lemmas 5.1 and 5.2 together with our choice of a and ε guarantee that $\tilde{g},\tilde{h}\in\mathcal{C}_u(a)$. Obviously, $\|\tilde{g}\|_1=\|\tilde{h}\|_1=1$.

The result of what we have just stated is that $(\widetilde{X}_1,\widetilde{Y}_1)$ is a coupling of (TX,TY). In addition, $\mathbb{E}(\widetilde{X}_1\mid\tau=1)$ and $\mathbb{E}(\widetilde{Y}_1\mid\tau=1)$ are independent random variables with absolutely continuous distributions in $\mathcal{C}_u(a)$. This is exactly the original situation for the variables (X,Y). It is then clear that one can perform the same coupling again. This leads to the following inductive procedure. We start with the independent coupling of the original random variables. Then, given a coupling $(\widetilde{X}_n,\widetilde{Y}_n)$ of T^nX and T^nY, we construct a new coupling $(\hat{X}_n,\hat{Y}_n)$. If τ_{n-1} was equal to zero, meaning that the points were coupled, we keep them coupled, while if τ_{n-1} was equal to one, we use a random variable τ_n, defined independently of what has been constructed up to this stage, to define the law of $(\hat{X}_n,\hat{Y}_n)$; then we obtain $(\widetilde{X}_{n+1},\widetilde{Y}_{n+1})$ by applying T:

$$\begin{aligned}
&\tau_{-1}=1;\quad \widetilde{X}_0=X;\quad \widetilde{Y}_0=Y;\\
&\mathbb{P}(\widetilde{X}_n\in A,\widetilde{Y}_n\in B\mid\tau_{n-1}=1)=\int_A h_n(x)m(dx)\int_B g_n(x)m(dx);\\
&\mathbb{P}(\tau_n=0\mid\tau_{n-1}=0)=1;\\
&\mathbb{P}(\tau_n=0\mid\tau_{n-1}=1)=\varepsilon;\\
&\mathbb{P}(\hat{X}_n\in A;\hat{Y}_n\in B\mid\tau_{n-1}=0)=\mathbb{P}(\widetilde{X}_n\in A;\widetilde{Y}_n\in B\mid\tau_{n-1}=0);\\
&\mathbb{P}(\hat{X}_n\in A;\hat{Y}_n\in B\mid\tau_n=0;\tau_{n-1}=1)=\delta_{\Phi,h_n,g_n}(A\times B);\\
&\mathbb{P}(\hat{X}_n\in A;\hat{Y}_n\in B\mid\tau_n=1)\\
&\qquad=(1-\varepsilon)^{-2}[\mathbb{P}(\widetilde{X}_n\in A\mid\tau_{n-1}=1)-\varepsilon\mathbb{P}(\hat{X}_n\in A\mid\tau_n=0;\tau_{n-1}=1)]\\
&\qquad\quad\times[\mathbb{P}(\widetilde{Y}_n\in B\mid\tau_{n-1}=1)-\varepsilon\mathbb{P}(\hat{Y}_n\in B\mid\tau_n=0;\tau_{n-1}=1)];\\
&\widetilde{X}_{n+1}=T\hat{X}_n;\quad \widetilde{Y}_{n+1}=T\hat{Y}_n.
\end{aligned}$$

The event $\{\tau_n=1\}$ corresponds to the points that have not yet been coupled at time n; their measure is easily computed

$$\mathbb{P}(\tau_n=1)=(1-\varepsilon)^n.$$

To compute the expectation it is useful to introduce the time of coupling $\bar{\tau}=\sum_{n=0}^{\infty}\tau_n$. Clearly the above computation shows that $\bar{\tau}$ is almost everywhere finite. Thus

$$\begin{aligned}
\mathbb{E}[d_\delta^s(\widetilde{X}_n,\widetilde{Y}_n)]&=\sum_{k=0}^{\infty}\mathbb{E}[d_\delta^s(\widetilde{X}_n,\widetilde{Y}_n)\mid\bar{\tau}=k]\mathbb{P}(\tau_{k+1}=0;\tau_k=1)\\
&=\sum_{k=0}^{n}\mathbb{E}[d_\delta^s(T^{n-k}\widetilde{X}_k,T^{n-k}\widetilde{Y}_k)\mid\bar{\tau}=k]\mathbb{P}(\tau_{k+1}=0;\tau_k=1)\\
&\quad+\sum_{k=n+1}^{\infty}\mathbb{E}[d_\delta^s(\widetilde{X}_n,\widetilde{Y}_n)\mid\bar{\tau}=k]\mathbb{P}(\tau_{k+1}=0;\tau_k=1)
\end{aligned}$$

$$\leq \sum_{k=0}^{n} \lambda_-^{-n+k} \mathbb{E}[d_\delta^s(\widetilde{X}_k, \widetilde{Y}_k) \mid \bar{\tau} = k](1-\varepsilon)^k + \sum_{k=n+1}^{\infty} \delta(1-\varepsilon)^k$$
$$\leq \text{const.}\, \theta^n. \qquad \square$$

Remark 6.1. Alternatively, it is possible to construct a coupling $(\widetilde{X}, \widetilde{Y})$ of X and Y such that for all $n \geq 0$,

$$\mathbb{E}[d_\delta^s(T^n \widetilde{X}, T^n \widetilde{Y})] \leq \text{const.}\, \theta^n.$$

The idea is to introduce a random variable $\bar{\tau}$ with a geometrical law of parameter $1-\varepsilon$ and set

$$\mathbb{P}(\widetilde{X} \in A; \widetilde{Y} \in B \mid \bar{\tau} = n) = \delta_{\Phi,h_n,g_n}(T^n A \times T^n B).$$

It is a coupling since

$$\begin{aligned}
\mathbb{P}(\widetilde{X} \in A) &= \sum_{n\geq 0} \mathbb{P}(\widetilde{X} \in A \mid \bar{\tau} = n)\mathbb{P}(\bar{\tau} = n) \\
&= \sum_{n\geq 0} \delta_{\Phi,h_n,g_n}(T^n A \times \mathcal{M})\varepsilon(1-\varepsilon)^n \\
&= \sum_{n\geq 0} (1-\varepsilon)^n \frac{\varepsilon}{\int_{\mathcal{M}} g_n \mathbb{A}_\Phi h_n} \int_{T^n A} g_n \mathbb{A}_\Phi h_n \\
&= \sum_{n\geq 0} (1-\varepsilon)^n \left(\int_{T^n A} g_n - (1-\varepsilon) \int_{T^{n+1} A} g_{n+1} \right) \\
&= \int_A g_0 = \mathbb{P}(X \in A),
\end{aligned}$$

and its properties follow from previous results.

7. *Proofs of the main statements*

7.1. *Proof of Theorem 2.1.* Let g and h be in $\mathcal{C}_u(a)$ such that $\int_{\mathcal{M}} g\, dm = 1$, $\int_{\mathcal{M}} h\, dm = 1$. We consider two random variables X and Y valued in $\mathcal{M}$ distributed according to, respectively, $g\, dm$ and $h\, dm$. Clearly,

$$\mathbb{E}[f(T^n X)] = \int_{\mathcal{M}} f \circ T^n g\, dm. \tag{7.1}$$

Hence

$$\int_{\mathcal{M}} f \circ T^n g\, dm - \int_{\mathcal{M}} f \circ T^n h\, dm = \mathbb{E}[f(T^n X)] - \mathbb{E}[f(T^n Y)]. \tag{7.2}$$

For all couplings $(\widetilde{X}_n, \widetilde{Y}_n)$ of $T^n X$ and $T^n Y$, we have

$$\int_{\mathcal{M}} f \circ T^n g\, dm - \int_{\mathcal{M}} f \circ T^n h\, dm = \mathbb{E}[f(\widetilde{X}_n) - f(\widetilde{Y}_n)]. \tag{7.3}$$

If $f \in \mathcal{C}_s$, we have

$$\begin{aligned}
\left| \int_{\mathcal{M}} f \circ T^n g\, dm - \int_{\mathcal{M}} f \circ T^n h\, dm \right| &\leq \|f\|_s \mathbb{E}[d_\delta^s(\widetilde{X}_n, \widetilde{Y}_n)^{\beta_s}] \\
&\leq \|f\|_s \mathbb{E}[d_\delta^s(\widetilde{X}_n, \widetilde{Y}_n)]^{\beta_s}.
\end{aligned} \tag{7.4}$$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

This together with Proposition 6.1 concludes the proof of Theorem 2.1 in the special case $g, h \in \mathcal{C}_u(a)$. The general case is obtained by noting that each function g with $\|g\|_u < \infty$ can be seen as the difference of two such positive functions. Then we can restrict to the case $g \geq 0$ and, for each $b > 0$, we have

$$\frac{g(x)+b}{g(y)+b} \leq \frac{g(y)+H_\sigma(g)s(x,y)+b}{g(y)+b} \leq \frac{|g|_\sigma s(x,y)}{b} + 1 \leq e^{|g|_\sigma s(x,y)/b}$$

that is $g + b \in \mathcal{C}_u(a)$ provided we choose

$$b = a^{-1}|g|_\sigma.$$

In conclusion, if $b = a^{-1}\max\{|g|_\sigma, |h|_\sigma\}$ and $\|g\|_1 = \|h\|_1$ the following holds:

$$\begin{aligned}&\left|\int_{\mathcal{M}} f \circ T^n g - \int_{\mathcal{M}} f \circ T^n h\right| \\ &\quad = \left|\int_{\mathcal{M}} f \circ T^n \frac{g+b}{\|g\|_1+b} - \int_{\mathcal{M}} f \circ T^n \frac{h+b}{\|h\|_1+b}\right|(\|g\|_1+b) \\ &\quad \leq \text{const.}\, \|f\|_s \max\{\|h\|_u,\ \|g\|_u\}\theta^{n\beta_s}.\end{aligned}$$

□

Remark 7.1. The $\overline{d}$-distance (associated to a distance d) between two random variables X and Y is the infimum over all couplings $(\widetilde{X}, \widetilde{Y})$ of these random variables of $\mathbb{E}[d(\widetilde{X}, \widetilde{Y})]$. It is a general fact that the speed of decay of correlations can be expressed in terms of the $\overline{d}$-distance (associated to the distance d_δ^s):

$$\left|\int_{\mathcal{M}} f \circ T^n g\, d\mu - \int_{\mathcal{M}} f\, d\mu \int_{\mathcal{M}} g\, d\mu\right| \leq \|f\|_s \overline{d}(T^n X, T^n Y)^{\beta_s}. \tag{7.5}$$

7.2. *Proof of Corollary 2.1.* For simplicity, we write the proof for the case $\beta_s = 1$. We start by proving a weak convergence result to identify μ.

Let $h \in \mathcal{C}_u(a)$, $\|h\|_1 = 1$ and set $d\mu_0 = h\, dm$ and $\mu_n = T^{n*}\mu_0$. We want to show that the sequence $\{\mu_n\}$ is weakly convergent.

Let $f \in \mathcal{C}_s$ then, by Theorem 2.1, for each $n, m \geq n_0$,

$$|\mu_n(f) - \mu_m(f)| = \left|\int_{\mathcal{M}} \mathcal{L}^{n-n_0} hf \circ T^{n_0} - \int_{\mathcal{M}} \mathcal{L}^{m-n_0} hf \circ T^{n_0}\right| \leq \text{const.}\, \theta^{n_0}\|f\|_s,$$

since $\mathcal{L}^k h \in \mathcal{C}_u(a)$. Noting that $\mathcal{C}_s \cap \mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$ is dense in $\mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$, by the usual 3-$\varepsilon$ argument it follows that $\{\mu_n\}$ is a weakly Cauchy sequence from which the claim follows.

In addition, an obvious modification of this argument shows that the limit measure μ does not depend on the function h. Clearly μ is an invariant measure. The next step is to prove that μ satisfies (2.7).

Given $h \in \mathcal{C}^{(\alpha)}(\mathcal{M}, \mathbb{R})$, let us apply Theorem 2.1 to the two functions $h_1 = h\mathcal{L}^k 1$ and $h_2 = \mathcal{L}^k 1 \int_{\mathcal{M}} h\mathcal{L}^k 1\, dm$, then for each $f \in \mathcal{C}_s$ holds

$$\left|\int_{\mathcal{M}} h_1 f \circ T^n - \int_{\mathcal{M}} h_2 f \circ T^n\right| \leq \text{const.}\, \|f\|_s \max\{\|h_1\|_u,\ \|h_2\|_u\}\theta^n \tag{7.6}$$

but $\|h_2\|_u \leq$ const. and $\|h_1\|_u \leq \text{const.}(\|h\|_u + 1)$. We can then take the limit for $k \to \infty$ in (7.6) and the result follows.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

By a standard approximation argument (2.7) implies that μ is mixing, and hence ergodic.

Before we turn to the conclusion of the proof, it is very relevant to notice the following stronger convergence result.

LEMMA 7.1. *If* $h \in \mathcal{C}^{(\alpha)}(\mathcal{M}, \mathbb{R})$ *and* $d\mu_0 = h\,dm$, $\mu_n = T^{n*}\mu_0$, *then*

$$\lim_{n\to\infty} \mu_n(g) = \mu(g) \quad \forall g \in \mathcal{C}_s.$$

Proof. We know already that the $\lim_{n\to\infty} \mu_n(g)$ exists; what is not so obvious is that it equals $\mu(g)$, since g may very well not be a continuous function. To prove this the idea is to approximate g by continuous functions in some not too weak sense, the problem is that we need some control on the $\|\cdot\|_s$ of the approximation. This can be achieved by using the average operator $\mathbb{A}_\delta$, let us now see how.

For each $\delta > 0$, $\nu_\delta(f) = \mu(\mathbb{A}_\delta f)$ defines a Borel measure; moreover, if $f \in \mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$ then $\mathbb{A}_\delta f \in \mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$. By Lusin's theorem, for each $g \in \mathcal{C}_s$ and $\varepsilon > 0$ there exists $g_\varepsilon \in \mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$ and a closed set $K_\delta \subset \mathcal{M}$ such that $g|_{K_\delta} = g_\varepsilon|_{K_\delta}$, $\|g\|_\infty = \|g_\varepsilon\|_\infty$ and $(m + \nu_\delta)(K_\delta^c) \le \varepsilon$.

Let $g_{\delta,\varepsilon} = \mathbb{A}_\delta g_\varepsilon$, then $\|g_{\delta,\varepsilon}\|_s \le (\delta^{-\beta_s} + 1)\|g\|_\infty$ and

$$\begin{aligned}\|g_{\delta,\varepsilon} \circ T^n - g \circ T^n\|_1 &\le \|(\mathbb{A}_\delta g_\varepsilon - \mathbb{A}_\delta g) \circ T^n\|_1 + \|\mathbb{A}_\delta g \circ T^n - g \circ T^n\|_\infty \\ &\le \Lambda^n \|\mathbb{A}_\delta(g_\varepsilon - g)\|_1 + \|g\|_s \delta^{\beta_s} \\ &\le \text{const.}(\Lambda^n \varepsilon + \delta^{\beta_s}),\end{aligned}$$

where $\Lambda = |\det(DT^{-1})|_\infty$. Accordingly,

$$\begin{aligned}|\mu(g) - \mu_n(g)| &\le |\mu(\mathbb{A}_\delta g) - \mu_n(g)| + \delta^{\beta_s}\|g\|_s \\ &\le |\mu(g_{\delta,\varepsilon}) - \mu_n(g)| + (\varepsilon + \delta^{\beta_s})\|g\|_s \\ &\le |\mu_n(g_{\delta,\varepsilon}) - \mu_n(g)| + \mathcal{O}((\delta^{-\beta_s}\theta^n + \varepsilon + \delta^{\beta_s})\|g\|_s) \\ &\le \mathcal{O}(\Lambda^n \varepsilon + \delta^{\beta_s} + \delta^{-\beta_s}\theta^n).\end{aligned}$$

Thus, by choosing first δ small, then n sufficiently large and finally ε sufficiently small, the result follows. □

We are now in a position to conclude the argument. Let

$$A = \left\{ x \in \mathcal{M} \;\middle|\; \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x} = \mu \right\}.$$

Then, clearly $T^{-1}A = A$ and A is not empty. In fact, for $f \in \mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$, the set A_f for which

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x}(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \mu(f)$$

is measurable and $\mu(A_f) = 1$ by Birkhoff's theorem. This, plus the separability of $\mathcal{C}^{(0)}(\mathcal{M}, \mathbb{R})$, implies A measurable and $\mu(A) = 1$†. Finally, note that if $x \in A$ then

† Indeed, if $\{f_n\} \subset C^{(0)}(\mathcal{M}, \mathbb{R})$ is dense, then $A = \bigcap_{n\in\mathbb{N}} A_{f_n}$.

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

$W^s(x) \subset A$. Accordingly, $\chi_A \in \mathcal{C}_s$ hence, by Lemma 7.1,

$$m(A) = m(\chi_A \circ T^n) = \lim_{n\to\infty} m(\chi_A \circ T^n) = \mu(A) = 1. \qquad \square$$

A. *Appendix*

Proof of Lemma 3.1. The first statement is obvious:

$$s(Tx, Ty) = \sum_{n=1}^{\infty} \nu^{-n}\varphi(T^{n+1}x, T^{n+1}y) = \nu s(x, y).$$

To prove the next statement, let us consider $x \in W^u_\delta(y)$, then

$$\mu_+^n d(x, y) \geq d(T^n x, T^n y) \geq \mu_-^n d(x, y). \tag{A.1}$$

Next, let $n_\delta \in \mathbb{N}$ be defined as

$$n_\delta = \inf\{n \in \mathbb{N} \mid d(T^n x, T^n y) > \delta\}$$

and set $n_- = n_\delta$ and $n_+ = n_{2\delta}$. Accordingly,

$$s(x, y) \leq \sum_{n=n_-}^{\infty} \nu^{-n} = \frac{\nu^{-n_-}}{1 - \nu^{-1}} \tag{A.2}$$

and

$$s(x, y) \geq \nu^{-n_+}\varphi(d(T^{n_+}x, T^{n_+}y)) = \nu^{-n_+}. \tag{A.3}$$

On the other hand, from (A.1), if δ is small enough,

$$n_- \geq \frac{\ln[\delta/d(x, y)]}{\ln \mu_+} \quad \text{and} \quad n_+ \leq \frac{\ln[2\delta/d(x, y)]}{\ln \mu_-}$$

from which the result follows with $\alpha_- = \ln \nu/\ln \mu_+$ and $\alpha_+ = \ln \nu/\ln \mu_-$.

To prove the third statement of the lemma, first note that the triangle inequality yields

$$d(T^n x, T^n y) - 2\lambda_-^{-n}\frac{\delta}{4} \leq d(T^n z, T^n w) \leq d(T^n x, T^n y) + 2\lambda_-^{-n}\frac{\delta}{4}.$$

Then note that if $n_* = n_{\delta/2}$, then $n_+ - n_*$ is bounded. So that, since $2\lambda_-^{-n}\delta/4 < \delta/2$,

$$\begin{aligned}
s(z, w) &\leq \sum_{n\geq 0} \nu^{-n}\varphi\left(d(T^n x, T^n y) + 2\lambda_-^{-n}\frac{\delta}{4}\right) \\
&\leq \sum_{n\geq n_*} \nu^{-n}\varphi\left(d(T^n x, T^n y) + 2\lambda_-^{-n}\frac{\delta}{4}\right) \\
&\leq \sum_{n\geq n_*} \nu^{-n}\varphi(d(T^n x, T^n y)) + \|\varphi'\|_\infty \nu^{-n}\lambda_-^{-n}\frac{\delta}{2} \\
&\leq s(x, y) + \text{const.}\, \nu^{n_+ - n_*}\nu^{-n_+} \\
&\leq c_0 s(x, y),
\end{aligned}$$

where we have used (A.3) to conclude. $\square$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

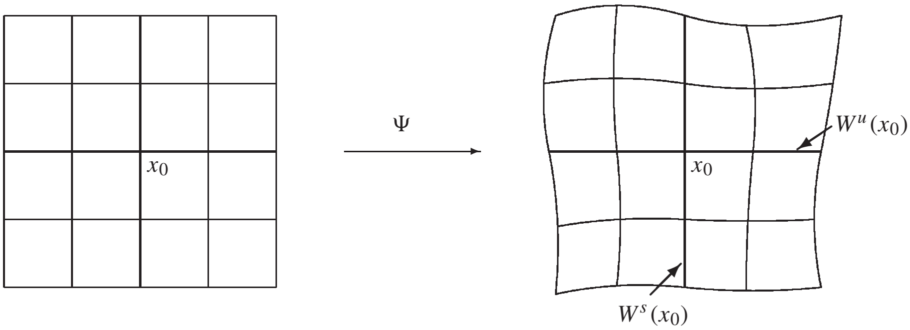

FIGURE A.1. The action of Ψ (all curves on the right are stable and unstable manifolds).

Proof of Proposition 4.2. To prove Proposition 4.2, we start by noting that $\psi^s_{x_0}$ and $\psi^u_{x_0}$ can be extended to a $\mathcal{C}^{(1+\alpha)}$ system of local coordinates on the manifold: $\tilde{\psi}_{x_0} : \mathbb{R}^{d_s} \times \mathbb{R}^{d_u} \to B_\delta(x_0)$ that maps $W^s_\delta(x_0)$ and $W^u_\delta(x_0)$ onto straight subspaces of $\mathbb{R}^d$ (but does not *a priori* send the foliation into a foliation in parallel subspaces)†. The image of the Riemannian volume through this map is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ and its Jacobian is α-Hölder. To simplify this, we shall write m again for this measure as if it was exactly the Riemannian volume (that is we confuse m with its image $\tilde{\psi}^{-1*}m$ in the coordinates) and use m^s and m^u to denote the restrictions of m to subspaces of dimension d_s and d_u. Consequently, we just need to prove absolute continuity and regularity of the Jacobian for the map $\Psi = \tilde{\psi}^{-1}_{x_0} \circ \psi_{x_0} : \mathbb{R}^d \to \mathbb{R}^d$. This means that we (locally) identify the manifold $\mathcal{M}$ with $\mathbb{R}^d$, the stable and unstable manifold around x_0 being straight subspaces. Note that $\Psi(\xi, 0) = (\xi, 0)$ and $\Psi(0, \eta) = (0, \eta)$ by construction (see Figure A.1 for a pictorical representation of the action of the map Ψ).

Our task is then to investigate the regularity of Ψ. That is we want to show that $\Psi^{-1*}m$ is absolutely continuous with respect to the Lebesgue measure and we must compute $d\Psi^{-1*}m/dm_L$.

Let us introduce two sets of families.

$$S_r(x) = \{W^u_\delta(y) \cap W^s_\delta(z) \mid y \in W^s_\delta(x); z \in W^u_\delta(x); \|x - y\| \le r; \|x - z\| \le r\},$$
$$P_r(x) = \{x + v + w \mid v \in E^s(x); w \in E^u(x); \|v\| \le r; \|w\| \le r\},$$

where the tangent space $\mathcal{T}_x\mathcal{M}$ is isometrically identified with $\mathbb{R}^d$. We shall see that it is possible to compare the measure of these sets with the measure of their images

† Consider any chart $\varphi : \mathbb{R}^d \to B_\delta(x) \subset \mathcal{M}$ not too far from an isometry (for example the chart provided by the exponential map). Clearly, provided that δ is chosen small enough ($\delta \le \delta_0$ independent of x_0), there exists $M_0 \in \mathbb{R}^+$, independent of x_0, such that $\|D\varphi\|_\infty + \|D\varphi^{-1}\|_\infty \le M_0$. Then we can define

$$\tilde{\psi}_{x_0}(\xi, \eta) := \varphi(\varphi^{-1} \circ \psi^s_{x_0}(\xi) + \varphi^{-1} \circ \psi^u_{x_0}(\eta)).$$

Since $D_0\psi^s_{x_0}$ and $D_0\psi^u_{x_0}$ are transversal by the Anosov property and $\psi^s_{x_0}, \psi^u_{x_0}$ are $\mathcal{C}^{(1+\alpha)}$, it follows that $D_\xi\psi^s_{x_0}$, $D_\eta\psi^u_{x_0}$ are uniformly transversal, provided $\delta \le \delta_0$ for some δ_0 independent of x_0. Thus $\tilde{\psi}_{x_0}$, is a change of coordinates and there exists $M_1 \in \mathbb{R}^+$, independent of x_0, such that $\|D\tilde{\psi}_{x_0}\|_\infty + \|\tilde{\psi}^{-1}_{x_0}\|_\infty \le M_1$ (we have used the fact that $D_\xi\psi^s_{x_0}$, $D_\eta\psi^u_{x_0}$ are uniformly bounded; see, for example, [**8**]).

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

$\Psi(S_r(x))$. The S_r, we will refer to as *pseudo-rectangles*, are constructed on parts of stable and unstable manifolds which are close to balls (they are intersections of balls in the ambient space and the manifolds) while the P_r, *parallelograms*, are bounded by affine 'approximations' of these parts of manifolds which are real balls of radius r in $\mathbb{R}^{d_s}$ and $\mathbb{R}^{d_u}$. Let us denote $P_r^s(x) = \{x + v \mid v \in E^s(x); \|v\| \leq r\} = \{x + E^s(x)\} \cap B_r(x)$ and $P_r^u(x) = \{x + w \mid w \in E^u(x); \|w\| \leq r\}$.

A first lemma shows that the measure of the S_r is well approximated by the measure of the parallelograms P_r, which we know how to compute in terms of the measure of their faces and the angle between them.

LEMMA A.1. *There exists $c \in \mathbb{R}^+$ such that the following properties hold:*
(1) $P_{r(1-cr^{\tau^2})}(x) \subset S_r(x) \subset P_{r(1+cr^{\tau^2})}(x)$;
(2) *there exists a τ-Hölder function $\theta : \mathcal{M} \to \mathbb{R}^+$ such that*

$$m(S_r(x)) = (1 + \mathcal{O}(r^{\tau^2}))\theta(x) m^s(P_r^s(x)) m^u(P_r^u(x)).$$

Proof. To investigate the shape of the sets $S_r(x)$ it is convenient to introduce normal coordinates with respect to the base point $x \in \mathbb{R}^d$. A stable manifold, in the neighborhood of x, can be represented by $\{x + \xi + A(\xi); \xi \in E^s(x)\}$ where $A : E^s(x) \to E^u(x)$ is a smooth map. The map A associated to $W^s(x)$ clearly has the property $A(0) = 0$ and $D_0 A = 0$. More generally, for $z \in x + E^u(x)$, the manifold $W^s(z)$ is uniquely represented by the map A_z as $\{z + \xi + A_z(\xi); \xi \in E^s(x)\}$. By construction $A_z(0) = 0$ while Proposition 4.1 (the Hölder continuity of the stable distribution) implies

$$\|D_\xi A_z\| \leq \text{const.}\, \|z + \xi + A_z(\xi)\|^\tau,$$

where the constant is independent on x. Proposition 4.1 also shows that to represent points in $S_r(x)$ we need only to consider $|z| \leq r$ and $\|\xi\| < r^\tau$.

For $v \in E^s(x)$, $\|v\| = 1$ we define

$$h(t) = \|A_z(vt)\|.$$

We have

$$\left|\frac{dh}{dt}\right| = \left|\frac{(A_z(tv), D_{tv} A_z v)}{\|A_z(tv)\|}\right| \leq \|D_{tv} A_z\| \leq \text{const.}\, \|z + vt + A_z(tv)\|^\tau$$

so that if $|z| \leq r$, h must satisfy the following differential inequality in the domain $|t| < r^\tau$:

$$\begin{cases} \left|\dfrac{dh}{dt}\right| \leq \text{const.}(r^\tau + h(t))^\tau \\ h(0) = 0. \end{cases}$$

Solving this differential inequality yields

$$h(t) \leq \text{const.}\, t r^{\tau^2} + \mathcal{O}(t^2 r^{-\tau + 2\tau^2}).$$

We conclude that if $|z| \leq r$ and $\|v\| < r^\tau$, we have

$$\|A_z(v)\| \leq \text{const.}\, \|v\| r^{\tau^2}. \tag{A.4}$$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

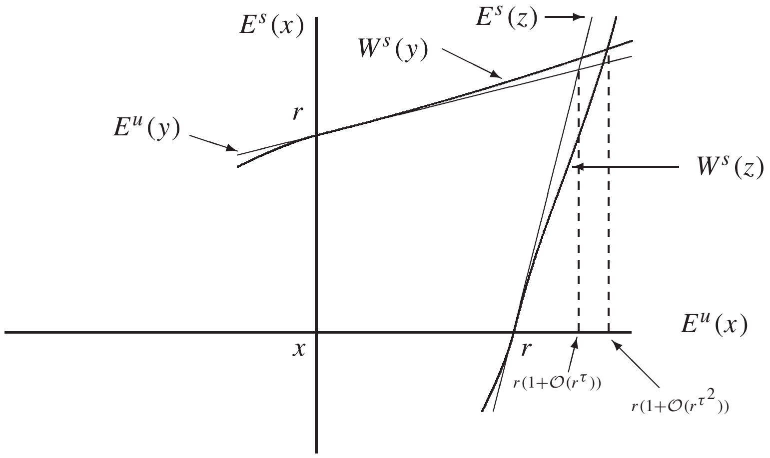

FIGURE A.2. Foliation and its linear approximations ($z = (r, 0)$, $y = (0, r)$).

The same estimate clearly holds for unstable manifolds. The two estimates together immediately imply that (see Figure A.2)

$$P_{r(1-cr^{\tau^2})} \subset S_r \subset P_{r(1+cr^{\tau^2})}.$$

Hence $m(P_{r(1-cr^{\tau^2})}) \leq m(S_r) \leq m(P_{r(1+cr^{\tau^2})})$. But the measure of P_r can be expressed as

$$m(P_r(x)) = \text{const.}\,\theta(x)r^d$$

where, the constant depends only on the dimensions, d_s and d_u, and where Proposition 4.1 shows that, if $\{v_i\}$ is an orthonormal base for $E_s(x)$ and $\{w_i\}$ for $E_u(x)$, then

$$\theta(x) = |\det(v_1 \ldots v_{d_s} w_1 \ldots w_{d_u})|$$

is τ-Holder. Finally,

$$m(S_r(x)) = m(P_r(x))(1 + \mathcal{O}(r^{\tau^2})). \qquad \square$$

The next lemma shows that the Lebesgue measure of the sets

$$C_r(x) = \Psi^{-1}(S_r(\Psi(x)))$$

can be compared to the measure of the corresponding parallelogram $P_r(\Psi(x))$ and hence to the measure of $S_r(\Psi(x)) = \Psi(C_r(x))$.

LEMMA A.2. *There exists a τ-Hölder function $\tilde{\rho}_{x_0}$ such that*

$$\frac{m_L(C_r(\Psi^{-1}(x)))}{m(S_r(x))} = \tilde{\rho}_{x_0}(x)^{-1}(1 + \mathcal{O}(r^{\tau^2})).$$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

Proof. Let $x = \Psi(\xi, \eta)$. The sets $C_r(\xi, \eta)$ are product sets.

$$C_r(\xi, \eta) = C_r^s(\xi, \eta) \times C_r^u(\xi, \eta),$$

where $C_r^s(\xi, \eta) = \{\xi' \in \mathbb{R}^{d_s} \mid \Psi(\xi', \eta) \in S_r(x)\}$ (that is $\Psi^{-1}(W^s(x) \cap S_r(x))$ seen as a subset of $\mathbb{R}^{d_s}$) and $C_r^u(x)$ is similarly defined. Since Ψ is the identity when restricted to the coordinate axis, their measure is exactly

$$\begin{aligned} m_L(C_r) &= m_L(C_r^s) m_L(C_r^u) \\ &= m^s(\Psi(C_r^s \times \{0\})) m^u(\Psi(\{0\} \times C_r^u)). \end{aligned}$$

To compare this measure to the measure of S_r, we shall compare it to the measure of the parallelogram P_r. The images by Ψ of the projection of the faces of C_r on the coordinate axes are approximately images of the faces of P_r through the holonomy map $\phi_u : \{x + E^s(x)\} \to W^s(\Psi(0, \eta))$. More precisely, the estimate (A.4) shows that, if r is small enough,

$$\phi_u(P^s_{r(1-cr^{\tau^2})}) \subset \{\Psi(0, \eta'), \eta' \in C_r^s\} \subset \phi_u(P^s_{r(1+cr^{\tau^2})}).$$

The holonomy map has a Jacobian $\widetilde{J}^u$. This Jacobian is not exactly the one defined in Proposition 4.1 because of the change of coordinate $\widetilde{\psi}_{x_0}$, but together with the regularity of $\widetilde{\psi}_{x_0}$, Proposition 4.1 proves that it is τ-Hölder. We can thus compute the measures by doing a change of variable:

$$\begin{aligned} m_L(C_r^s) &= m^s(\{\Psi(0, \eta'), \eta' \in C_r^s\}) \\ &= \int_{\phi_u^{-1}(\{\Psi(0,\eta'), \eta' \in C_r^s\})} \widetilde{J}^u \\ &= \widetilde{J}^u(x) m^s(\phi_u^{-1}(\{\Psi(0, \eta'), \eta' \in C_r^s\}))(1 + \mathcal{O}(r^\tau)) \\ &= \widetilde{J}^u(x) m^s(P_r(x))(1 + \mathcal{O}(r^{\tau^2})). \end{aligned}$$

Of course, the same estimate holds for the unstable part. Hence, denoting by $\widetilde{J}^s$ the corresponding Jacobian,

$$\frac{m_L(C_r(\Psi^{-1}(x)))}{m(S_r(x))} = \frac{\widetilde{J}^u(x)\widetilde{J}^s(x)}{\theta(x)}(1 + \mathcal{O}(r^{\tau^2})). \tag{A.5}$$

□

We shall now use these estimates to compare the measure $\bar{m} = \Psi^* m$ with the Lebesgue measure.

LEMMA A.3. *The measures $\bar{m}$ and m_L are equivalent.*

Proof. Let us start by proving that $\bar{m}$ is absolutely continuous with respect to m_L. This is equivalent to saying that if, for some measurable set A, $\bar{m}(A) = 0$ then $m_L(A) = 0$. Since, by definition, $\bar{m}(A) = m(\Psi(A))$ this means that, for each $\varepsilon > 0$, it is possible to cover the set $\Psi(A)$ with a collections of pseudo-rectangles S_n such that

$$\sum_n m(S_n) \le \varepsilon.$$

https://doi.org/10.1017/S0143385702000056 Published online by Cambridge University Press

But, by Lemma A.2, $m(S_n) \geq c^{-1} m_L(C_n)$, where $S_n = \Psi(C_n)$. Hence,

$$m_L(A) \leq \sum_n m_L(C_n) \leq c \sum_n m(S_n) \leq c\varepsilon$$

which shows that $m_L(A) = 0$.

Next, let us prove that m_L is absolutely continuous with respect to $\bar{m}$. To prove this we will show that if for a measurable set A, $\bar{m}(A) > 0$, then $m_L(A) > 0$. We will argue by contradiction.

Suppose that there exists a measurable set A such that $\bar{m}(A) > 0$ and $m_L(A) = 0$. Since both $\bar{m}$ and m_L are Borel measures they are regular. Accordingly, for each $\varepsilon > 0$ there exists an open set $U \supset A$ such that

$$m_L(U) \leq \varepsilon,$$

and there exists a compact set $K \subset A$ such that

$$\bar{m}(K) \geq \tfrac{1}{2}\bar{m}(A).$$

Since Ψ is continuous together with its inverse, $\Psi(K)$ is compact and $\Psi(U)$ is open. It is then easy to construct a finite disjoint collection of cubes Γ_n such that $\Gamma_n \subset U$ and

$$m(\cup_n \Gamma_n \cap \Psi(K)) \geq \tfrac{1}{2}\bar{m}(K).$$

By Lemma A.1 it is clear that in each such cube we can fit a pseudo-rectangle S_n such that $m(S_n) \geq c_q m(\Gamma_n)$. Collecting these considerations and remembering $m(S) \leq c m_L(C)$ from Lemma A.2 yields

$$\begin{aligned} m_L(U) &\geq m_L(\cup_n \Psi^{-1}(S_n)) \geq c^{-1} \sum_n m(S_n) \geq c^{-1} c_q \sum_n m(\Gamma_n) \\ &\geq \frac{c_q}{2c}\bar{m}(K) \geq \frac{c_q}{4c}\bar{m}(A) \end{aligned}$$

which leads to the announced contradiction provided ε has been chosen small enough. □

We are now in a position to conclude the proof of Proposition 4.2. In effect, the family $\{S_r(x)\}$ is a family of nicely shrinking sets (in the sense of Rudin [**19**, p. 163]). So we can use the Lebesgue differentiation theorem (see, for example, Rudin [**19**, p. 166]) to conclude that the Radon–Nikodym derivative of $\bar{m}$ with respect to the Lebesgue measure can be computed as inverse of limit of ratios $\psi^* m_L(S_r(x))/m(S_r(x))$ as $r \to 0$. Its value, at all points, is given by Lemma A.2.

The result is obtained by pulling back the Jacobian to the manifold $\mathcal{M}$ through the $C^{(1+\alpha)}$ change of coordinates $\widetilde{\psi}_{x_0}$ to obtain finally ρ_{x_0}. This function inherits the regularity properties of $\widetilde{\psi}_{x_0}$, $\widetilde{J}^s$, $\widetilde{J}^u$ and $\widetilde{\rho}_{x_0}$. □

Acknowledgements. We would like to thank P. Ferrari, A. Galves, F. Ledrappier, L.-S. Young and J. Schmeling for valuable discussions and suggestions. The authors gladly acknowledge the support of the ESF Programme PRODYN; one of us (C.L.) also acknowledges support from the CNR-GNFM.

REFERENCES

[1] V. Baladi. *Dynamical Zeta Functions (NATO Advanced Science Institute Series C Mathematical Physical Sciences, 464)*. Kluwer, Dordrecht, 1995.

[2] R. Bowen *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)*. Springer, New York, Heidelberg, Berlin, 1978.

[3] X. Bressaud, R. Fernandez and A. Galves. Speed of $\overline{d}$-convergence for Markov approximations of chains with complete connections. A coupling approach. *Stochast. Process. Appl.* **83**(1) (1999), 127–138.

[4] X. Bressaud, R. Fernandez and A. Galves. Decay of correlations for non-Hölderian dynamics. A coupling approach. *Electron. J. Prob.* **4**(3) (1999), 19 pp.

[5] D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. Math.* (2), **147**(2) (1998), 357–390.

[6] D. Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergod. Th. & Dynam. Sys.* **18**(5) (1998), 1097–1114.

[7] N. I. Chernov. Markov approximations and decay of correlations for Anosov flows. *Ann. Math.* (2), **147**(2) (1998), 269–324.

[8] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems (Encyclopedia of Mathematics and its Applications, 54)*. Ed. G.-C. Rota. Cambridge University Press, New York, 1995.

[9] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4), **28**(1) (1999), 141–152.

[10] Holley. Remarks on the FKG inequalities. *Comm. Math. Phys.* **36** (1974), 227–231.

[11] T. M. Ligget. *Interacting Particle Systems*. Springer, Berlin, Heidelberg, New York, 1985.

[12] C. Liverani. Decay of correlations. *Ann. Math.* **142** (1995), 239–301.

[13] C. Liverani. Decay of correlations in piecewise expanding maps. *J. Stat. Phys.* **78**(3/4) (1995), 1111–1129.

[14] C. Liverani. Flows, random perturbations and rate of mixing. *Ergod. Th. & Dynam. Sys.* **18**(6) (1998), 1421–1446.

[15] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergod. Th. & Dynam. Sys.* **19**(3) (1999), 671–685.

[16] R. Mañe. *Ergodic Theory and Differentiable Dynamics (A Series of Modern Surveys in Mathematics)*. Springer, Berlin, Heidelberg, New York, 1978.

[17] V. Maume-Deshamps. Propriétés de mélange pour des systèmes dynamiques Markoviens. *Thèse*, Université de Bourgogne, Dijon, 1998.

[18] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.* **4** (1970), 337–352.

[19] W. Rudin. *Real and Complex Analysis (McGraw-Hill Series in Higher Mathematics)*, 2nd edn. McGraw-Hill, New York, 1974.

[20] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel. J. Math.*, **116** (2000), 223–248.

[21] M. Viana. Stochastic dynamics of deterministic systems. *Col. Bras. de Matemática*, notes, 1997.

[22] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.* (2), **147**(3) (1998), 585–650.

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