

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): A. A. PINTO and D. A. RAND

Article Title: RIGIDITY OF HYPERBOLIC SETS ON SURFACES

Year of publication: 2005

Link to published

version: <http://dx.doi.org/10.1112/S0024610704006052>

Publisher statement: None

RIGIDITY OF HYPERBOLIC SETS ON SURFACES

A. A. PINTO AND D. A. RAND

ABSTRACT

Given a hyperbolic invariant set of a diffeomorphism on a surface, it is proved that, if the holonomies are sufficiently smooth, then the diffeomorphism on the hyperbolic invariant set is rigid in the sense that it is C^{1+} conjugate to a hyperbolic affine model.

1. Introduction

In dynamics, rigidity occurs when simple topological and analytical conditions on the model system imply that there is no flexibility and so there is a unique smooth realisation. One can paraphrase this by saying that the moduli space for such systems is a singleton. For example, a famous result of this type due to Arnol'd, Herman and Yoccoz [1, 9, 24] is that a sufficiently smooth diffeomorphism of the circle with an irrational rotation number which satisfies the usual Diophantine condition is C^{1+} conjugate to a rigid rotation. The rigidity depends upon both the analytical hypothesis concerning the smoothness and the topological condition given by the rotation number, and if either are relaxed, then it fails.

The analytical part of the rigidity hypotheses for hyperbolic surface dynamics will be a condition on the smoothness of the holonomy maps along stable and unstable manifolds. Given a diffeomorphism f on a surface with a hyperbolic invariant set Λ (with local product structure and with a dense orbit on Λ), we show that if the holonomies are sufficiently smooth, then the diffeomorphism f is rigid, that is, there is conjugacy on Λ between f and a hyperbolic affine model which has a C^{1+} extension to the surface. A corollary of our result is that if f is C^r and the holonomies of f are C^r with $r - 1$ greater than the Hausdorff dimension along the stable and unstable leaves intersected with Λ , then f is rigid. We allow both the case where $\Lambda = M$ (so that f is Anosov and $M \cong \mathbb{T}^2$ [6, 16]) and the case where Λ is a proper subset (for example a horseshoe, or an attractor with 1-dimensional unstable manifolds such as the Plykin attractor).

Before stating our results we recall some previous rigidity results for surface dynamics. These are about Anosov diffeomorphisms of the torus. In this case the hyperbolic affine model is a hyperbolic toral automorphism. In general, the topological conjugacy between such a system and the corresponding hyperbolic affine model is only Hölder continuous and need not be any smoother. This is the case if there is a periodic orbit of f whose eigenvalues differ from those of the hyperbolic affine model.

Received 30 January 2003; revised 14 May 2004.

2000 *Mathematics Subject Classification* 37D20 (primary), 37E30 (secondary).

The first author was supported by the Gulbenkian Foundation, FCT, PROLYN-ESF, PRAXIS and the Centro de Matemática, Universidade do Porto. The second author was supported by the Wolfson Foundation and the UK Science and Engineering Research Council.

For Anosov diffeomorphisms f of the torus, there are the following results, all of which have the form that if a C^k f has C^r foliations, then f is C^s -rigid, that is, f is C^s -conjugate to the corresponding hyperbolic affine model.

- (1) Area-preserving Anosov diffeomorphisms f with $r = \infty$ are C^∞ -rigid (Avez [2]).
- (2) C^k area-preserving Anosov diffeomorphisms f with $r = 1 + o(t|\log t|)$ are C^{k-3} -rigid (Hurder and Katok [10]).
- (3) C^1 area-preserving Anosov diffeomorphisms f with $r \geq 2$ are C^r -rigid (Flaminio and Katok [7]).
- (4) C^k Anosov diffeomorphisms f ($k \geq 2$) with $r \geq 1 + \text{Lipschitz}$ are C^k -rigid (Ghys [8]).

For the Anosov case, our main result has the following corollary. If f is a C^r Anosov diffeomorphism ($r > 2$) and the holonomies of f are $C^{1+\text{zygmund}}$, then f is C^r -rigid.

1.1. Main theorem

Throughout this paper we consider a C^r diffeomorphism f , with $r > 1$, of a compact surface M which has a topologically transitive hyperbolic invariant subset Λ with local product structure and with one-dimensional local stable and unstable leaves (see the definitions of stable and unstable leaves and local product structure in Subsections 2.3 and 2.5 and [23]).

To state our rigidity result, we have to introduce the notion of a hyperbolic affine model and the definition of $C^{1,\alpha}$ classes of smooth regularities for homeomorphisms on the real line.

DEFINITION 1. A hyperbolic affine model for f on Λ is an atlas \mathcal{A} with the following properties.

- (i) The union of the domains U of the charts $i : U \rightarrow \mathbb{R}^2$ of \mathcal{A} (which are open in M) cover Λ .
- (ii) Any two charts $i : U \rightarrow \mathbb{R}^2$ and $j : V \rightarrow \mathbb{R}^2$ in \mathcal{A} have overlap maps $j \circ i^{-1} : i(U \cap V) \rightarrow \mathbb{R}^2$ with affine extensions to \mathbb{R}^2 (see Figure 1).
- (iii) f is affine with respect to the charts in \mathcal{A} .
- (iv) Λ is a basic hyperbolic set.
- (v) The images of the stable and unstable local leaves under the charts in \mathcal{A} are contained in horizontal and vertical lines.
- (vi) The basic holonomies have affine extensions to the stable and unstable leaves with respect to the charts in \mathcal{A} .

See the definition of basic holonomies in Subsection 2.9.

DEFINITION 2. Let $\theta : I \subset \mathbb{R} \rightarrow J \subset \mathbb{R}$ be a homeomorphism. If $0 < \alpha < 1$, then θ is said to be $C^{1,\alpha}$ if it is differentiable and for all points $x, y \in I$,

$$|\theta'(y) - \theta'(x)| \leq \chi(|y - x|), \tag{1.1}$$

where the positive function $\chi(t)$ is $o(t^\alpha)$, that is, $\lim_{t \rightarrow 0} \chi(t)/t^\alpha = 0$.

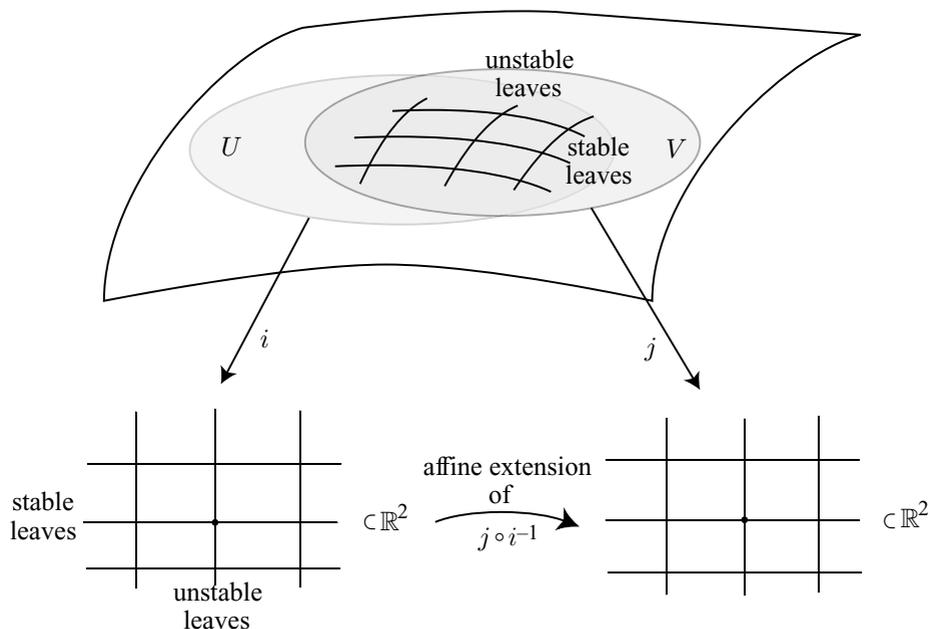


FIGURE 1. Overlap map for two charts in the affine structure for the affine model of f ; the local representation of f in such charts is affine.

The map $\theta : I \rightarrow J$ is said to be $C^{1,1}$ if, for all points $x, y \in I$,

$$\left| \log \theta'(x) + \log \theta'(y) - 2 \log \theta' \left(\frac{x+y}{2} \right) \right| \leq \chi(|y-x|), \tag{1.2}$$

where the positive function $\chi(t)$ is $o(t)$, that is, $\lim_{t \rightarrow 0} \chi(t)/t = 0$. The functions χ are called the *modulus of continuity* of θ .

In particular, a $C^{1+\beta}$ diffeomorphism is $C^{1,\alpha}$ if $0 < \alpha < \beta$, and a $C^{2+\gamma}$ diffeomorphism is $C^{1,1}$ if $\gamma > 0$.

We note that the regularity $C^{1,1}$ (also denoted by $C^{1+\text{zygmund}}$) of a diffeomorphism θ used in this paper is stronger than the regularity $C^{1+\text{Zygmund}}$ (see [15, 22]).

The importance of these $C^{1,\alpha}$ smoothness classes for a homeomorphism $\theta : I \rightarrow J$ follows from the fact that if $0 < \alpha < 1$, then the map θ will distort ratios of lengths of short intervals in an interval $K \subset I$ by an amount that is $o(|I|^\alpha)$, and if $\alpha = 1$, then the map θ will distort the cross-ratios of quadruples of points in an interval $K \subset I$ by an amount that is $o(|I|)$ (see the definition in Section 6). In fact, it is just these distortion properties that we will use in the proofs of our results.

In Subsection 2.10, we introduce the definition of a complete system of holonomies, and we explain how the definition of a $C^{1,\alpha}$ homeomorphism leads to the notion of a $C^{1,\alpha}$ complete system of holonomies. One example of a complete system of holonomies is to take all basic holonomies. However, in general, a complete system of holonomies can be taken to be much smaller, and for many systems such as Anosov systems, one can take a single holonomy. On the other hand, for Smale horseshoes, a complete system of holonomies is countably infinite.

THEOREM 1 (hyperbolic rigidity). *Let HD^s and HD^u be, respectively, the Hausdorff dimension of the intersection with Λ of the stable and unstable leaves of f . If f is C^r with $r - 1 > \max\{HD^s, HD^u\}$, and there is a complete system of holonomies for f in which the stable holonomies are C^{1,HD^s} and the unstable holonomies are C^{1,HD^u} , then the map f on Λ is $C^{1+\gamma}$ conjugate to a hyperbolic affine model for some $\gamma > 0$.*

In Subsection 2.5, we give the definition of a $C^{1+\gamma}$ conjugacy. In assuming that f is C^r with $r - 1 > \max\{HD^s, HD^u\}$ in the previous theorem, we actually only use the fact that f is C^{1,HD^i} along ι -leaves for $\iota \in \{s, u\}$.

If f is an Anosov diffeomorphism, then the hyperbolic affine model always exists and up to affine conjugacy is unique in its topological conjugacy class [6, 7, 14, 16]. In fact, the affine model is, up to affine conjugacy, an Anosov automorphism. We use Theorem 1 to prove the following corollary, which extends the result mentioned above of Ghys [8].

COROLLARY 1 (Anosov rigidity). *If f is a C^r Anosov diffeomorphism of a surface with $r > 2$, and there is a complete system of holonomies for f in which the stable and unstable holonomies are $C^{1,1}$, then f is C^r -conjugate to an affine model.*

If $\Lambda \neq M$, then, up to affine conjugacy, the set of hyperbolic affine models for f is either finite-dimensional or empty. In the case of the well known Smale horseshoes f , as presented in Figure 5, the set of affine models form a two-dimensional set homeomorphic to $\mathbb{R}^+ \times \mathbb{R}^+$. In the case of hyperbolic attractors f with $HD^s < 1$, in Ferreira and the authors [4] showed that there are no affine models for f , and so the stable holonomies can never be smoother than C^{1,HD^s} .

2. $C^{1,HD}$ complete set of holonomies

In this section, we present some basic facts on hyperbolic dynamics that we include for clarity of exposition. We also introduce the definition of a $C^{1,HD}$ complete set of holonomies in Subsection 2.10.

2.1. Interval notation

We also use the notation of interval arithmetic for some inequalities where the following hold.

- (i) If I and J are intervals, then $I + J$, $I \cdot J$ and I/J have the obvious meanings as intervals.
- (ii) If $I = \{x\}$, then we often denote I by x .
- (iii) $I \pm \varepsilon$ denotes the interval consisting of those x such that $|x - y| < \varepsilon$ for all $y \in I$.

Thus $\phi(n) \in 1 \pm \mathcal{O}(\nu^n)$ means that there exists a constant $c > 0$ depending only on explicitly mentioned quantities such that, for all $n \geq 0$, $1 - c\nu^n < \phi(n) < 1 + c\nu^n$.

2.2. Stable and unstable superscripts

Throughout the paper we will use the following notation. We use ι to denote an element of the set $\{s, u\}$ of the stable and unstable superscripts and ι' to denote the

element of $\{s, u\}$ that is not ι . In the main discussion we will often refer to objects which are qualified by ι , such as, for example, an ι -leaf. This is a leaf which is a leaf of the stable lamination if $\iota = s$, or a leaf of the unstable lamination if $\iota = u$. In general, the meaning should be quite clear.

We define the map $f_\iota = f$ if $\iota = u$ or $f_\iota = f^{-1}$ if $\iota = s$.

2.3. Leaf segments

Let d be a fixed Riemannian metric on M . For $\varepsilon > 0$ and $\iota \in \{s, u\}$, if $x \in \Lambda$ then we denote the *local ι -leaves* through x by

$$W^\iota(x, \varepsilon) = \{y \in M : d(f_\iota^{-n}x, f_\iota^{-n}y) \leq \varepsilon, \text{ for all } n \geq 0\}.$$

By the stable manifold theorem [23], these sets are, respectively, contained in the stable and unstable immersed manifolds

$$W^\iota(x) = \bigcup_{n \geq 0} f_\iota^n(W^\iota(f_\iota^{-n}x, \varepsilon)),$$

which are the image of a C^r immersion $\lambda_x^\iota : \mathbb{R} \rightarrow M$. A *full ι -leaf segment* I is defined as a subset of $W^\iota(x)$ of the form $\lambda_x^\iota(I_1)$, where I_1 is an open subinterval in \mathbb{R} . An *ι -leaf segment* is the intersection with Λ of a full ι -leaf segment. The *endpoints* of such a full ι -leaf segment are the points $\lambda_x^\iota(u)$ and $\lambda_x^\iota(v)$, where u and v are the endpoints of I_1 . The *endpoints* of such an ι -leaf segment I are the points of the minimal full ι -leaf segment containing I . A map $u : I \rightarrow \mathbb{R}$ is an *ι -leaf chart* of an ι -leaf segment I if has an extension $\hat{u} : \hat{I} \rightarrow \mathbb{R}$ to a full ι -leaf segment \hat{I} with the following properties: $I \subset \hat{I}$ and \hat{u} is a homeomorphism onto its image.

2.4. Topological and smooth conjugacies

Let $f : M \rightarrow M$ be a C^r diffeomorphism with a hyperbolic basic set Λ . More unusually, we also want to highlight the C^r structure on M in which f is a diffeomorphism. By a *C^r structure* on M , we mean a maximal set of charts with open domains in M such that the union of their domains cover M , and whenever U is an open subset contained in the domains of any two of these charts i and j , then the overlap map $j \circ i^{-1} : i(U) \rightarrow j(U)$ is C^r . We note that by compactness of M , given such a C^r structure on M , there is an atlas consisting of a finite set of these charts which cover M and for which the overlap maps are uniformly bounded in the C^r norm. We denote by \mathcal{C}_f the C^r structure on M in which f is a diffeomorphism. Usually one is not concerned with this as, given two such structures, there is a homeomorphism of M sending one onto the other, and thus, from this point of view, all such structures can be identified. For our discussion it will be important to maintain the identity of the different smooth structures on M .

Let f and g be C^r diffeomorphisms with hyperbolic invariant sets Λ_f and Λ_g , respectively. We say that a map $h : \Lambda_f \rightarrow \Lambda_g$ is a *topological conjugacy* between f and g if there is a homeomorphism $h : \Lambda_f \rightarrow \Lambda_g$ with the following properties.

- (i) $g \circ h(x) = h \circ f(x)$ for every $x \in \Lambda_f$.
- (ii) The pull-back of the ι -leaf segments of g by h are ι -leaf segments of f .

Similarly, we say that a topological conjugacy $h : \Lambda_f \rightarrow \Lambda_g$ is a *C^s conjugacy*, with $1 < s \leq r$, if h has a C^s diffeomorphic extension to an open neighbourhood of Λ_f in the surface M with respect to the C^r structures \mathcal{C}_f and \mathcal{C}_g on M .

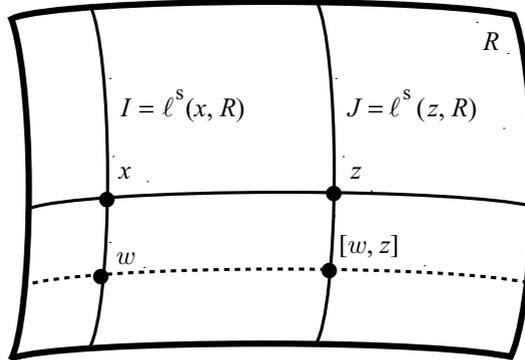


FIGURE 2. A rectangle.

2.5. Rectangles

Since Λ is a hyperbolic invariant set of a diffeomorphism $f : M \rightarrow M$, for $0 < \varepsilon < \varepsilon_0$ there is $\delta = \delta(\varepsilon) > 0$ such that, for all points $w, z \in \Lambda$ with $d(w, z) < \delta$, $W^u(w, \varepsilon)$ and $W^s(z, \varepsilon)$ intersect in a unique point that we denote by $[w, z]$. Since we assume that the hyperbolic set has a *local product structure*, we have $[w, z] \in \Lambda$. Furthermore, the following properties are satisfied: (i) $[w, z]$ varies continuously with $w, z \in \Lambda$, (ii) the bracket map is continuous on a δ -uniform neighbourhood of the diagonal in $\Lambda \times \Lambda$, and (iii) whenever both sides are defined, $f([z, w]) = [f(z), f(w)]$. Note that the bracket map does not really depend on δ provided that it is sufficiently small.

We emphasise that it is a standing hypothesis that all the hyperbolic sets considered here have such a local product structure.

A *rectangle* R is a subset of Λ which is (i) closed under the bracket (that is, $x, y \in R \Rightarrow [x, y] \in R$), and (ii) proper (that is, is the closure of its interior in Λ). This definition imposes that a rectangle always has to be proper, which is more restrictive than the usual one which only insists on the closure condition.

If ℓ^s and ℓ^u are, respectively, stable and unstable leaf segments intersecting in a single point, then we denote by $[\ell^s, \ell^u]$ the set consisting of all points of the form $[w, z]$ with $w \in \ell^s$ and $z \in \ell^u$. We note that if the stable and unstable leaf segments ℓ and ℓ' are closed, then the set $[\ell, \ell']$ is a rectangle. Conversely, in this 2-dimensional situation, any rectangle R has a product structure in the following sense. For each $x \in R$, there are closed stable and unstable leaf segments of Λ , $\ell^s(x, R) \subset W^s(x)$ and $\ell^u(x, R) \subset W^u(x)$ such that $R = [\ell^s(x, R), \ell^u(x, R)]$. The leaf segments $\ell^s(x, R)$ and $\ell^u(x, R)$ are called *stable and unstable spanning leaf segments* for R (see Figure 2). For $\iota \in \{s, u\}$, we denote by $\partial \ell^\iota(x, R)$ the set consisting of the endpoints of $\ell^\iota(x, R)$, and we denote by $\text{int } \ell^\iota(x, R)$ the set $\ell^\iota(x, R) \setminus \partial \ell^\iota(x, R)$. The *interior* of R is given by $\text{int } R = [\text{int } \ell^s(x, R), \text{int } \ell^u(x, R)]$, and the *boundary* of R is given by $\partial R = [\partial \ell^s(x, R), \ell^u(x, R)] \cup [\ell^s(x, R), \partial \ell^u(x, R)]$.

2.6. Markov partitions

A *Markov partition* of f is a collection $\mathcal{R} = \{R_1, \dots, R_k\}$ of rectangles such that the following hold.

- (i) $\Lambda \subset \bigcup_{i=1}^k R_i$.
- (ii) $R_i \cap R_j = \partial R_i \cap \partial R_j$ for all i and j .
- (iii) If $x \in \text{int } R_i$ and $fx \in \text{int } R_j$, then
 - (a) $f(\ell^s(x, R_i)) \subset \ell^s(fx, R_j)$ and $f^{-1}(\ell^u(fx, R_j)) \subset \ell^u(x, R_i)$;
 - (b) $f(\ell^u(x, R_i)) \cap R_j = \ell^u(fx, R_j)$ and $f^{-1}(\ell^s(fx, R_j)) \cap R_i = \ell^s(x, R_i)$.

Condition (b) means that $f(R_i)$ goes across R_j just once. In fact, it follows from condition (a) provided that the rectangles R_j are chosen to be sufficiently small [13]. The rectangles which make up the Markov partition are called *Markov rectangles*.

2.7. Leaf n -cylinders and leaf n -gaps

For $\iota = s$ or u , an ι -leaf primary cylinder of a Markov rectangle R is a spanning ι -leaf segment of R . For $n \geq 1$, an ι -leaf n -cylinder of R is an ι -leaf segment I such that the following hold.

- (i) $f_\iota^n I$ is an ι -leaf primary cylinder of a Markov rectangle M .
- (ii) $f_\iota^n(\ell^{\iota'}(x, R)) \subset M$ for every $x \in I$.

For $n \geq 2$, an ι -leaf n -gap G of R is a pair $\{x, y\}$ of distinct points x and y in a Markov rectangle R such that the following hold.

- (i) There is an embedding $i: \ell^\iota(x, R) \rightarrow \mathbb{R}$ in the topological lamination structure $\{x, y\} = i^{-1}(J)$ for some non-trivial closed interval J in \mathbb{R} .
- (ii) n is the smallest integer such that both of the leaves $f_\iota^{n-1}\ell^{\iota'}(x, R)$ and $f_\iota^{n-1}\ell^{\iota'}(y, R)$ are contained in ι' -rectangles.

We note that an ι -leaf segment I of a Markov rectangle R can be simultaneously an n_1 -cylinder, $(n_1 + 1)$ -cylinder, ..., n_2 -cylinder of R if $f^{n_1}(I), f^{n_1+1}(I), \dots, f^{n_2}(I)$ are all ι -spanning leaf segments. Furthermore, if I is an ι -leaf segment contained in the common boundary of two Markov rectangles R_i and R_j , then I can be an n_1 -cylinder of R_i and an n_2 -cylinder of R_j with n_1 distinct from n_2 . If $G = \{x, y\}$ is an ι -gap of R contained in the interior of R , then there is a unique n such that G is an n -gap. However, if $G = \{x, y\}$ is contained in the common boundary of two Markov rectangles R_i and R_j , then G can be an n_1 -gap of R_i and an n_2 -gap of R_j with n_1 distinct from n_2 . Since the number of Markov rectangles R_1, \dots, R_k is finite, there is $C \geq 1$ such that, in all the above cases for cylinders and gaps, we have $|n_2 - n_1| \leq C$.

We say that a leaf segment K is the i th mother of an n -cylinder or an n -gap J of R if $J \subset K$ and K is a leaf $(n - i)$ -cylinder of R .

By the properties of a Markov partition, for every $n \geq 1$ and every $j \geq 1$, a leaf n -cylinder K of a Markov rectangle R is equal to the union of all leaf $(n + j)$ -cylinders and of all leaf $(n + i)$ -gaps of R contained in K with $i \in \{1, \dots, j\}$.

2.8. Metric on Λ

We say that a rectangle R is an (n_s, n_u) -rectangle if there is $x \in R$ such that, for $\iota = s$ and u , the spanning leaf segments $\ell^\iota(x, R)$ are either an ι -leaf n_ι -cylinder or the union of two such cylinders with a common endpoint.

The reason for allowing the possibility of the spanning leaf segments being inside two touching cylinders is to allow us to regard geometrically very small rectangles intersecting a common boundary of two Markov rectangles as being small in the sense of having n_s and n_u large.

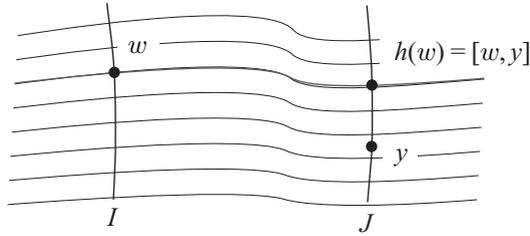


FIGURE 3. Basic stable holonomy from I to J .

If $x, y \in \Lambda$ and $x \neq y$, then $d_\Lambda(x, y) = 2^{-n}$, where n is the biggest integer such that both x and y are contained in an (n_s, n_u) -rectangle with $n_s \geq n$ and $n_u \geq n$. Similarly, if I and J are ι -leaf segments, then $d_\Lambda(I, J) = 2^{-n_\iota}$, where $n_\iota = 1$ and n_ι is the biggest integer such that both I and J are contained in an (n_s, n_u) -rectangle.

2.9. Basic holonomies

Suppose that x and y are two points inside any rectangle R of Λ . Let I and J be two stable leaf segments containing x and y respectively and inside R . Then we define $h : I \rightarrow J$ by $h(w) = [w, y]$. See Figure 3. Such maps are called the *basic stable holonomies*. They generate the pseudo-group of all stable holonomies. We define the basic unstable holonomies similarly.

2.10. C^{1,HD^ι} complete system of holonomies

We are going to prove that if the holonomies are sufficiently smooth, then the system is essentially affine. Rather than considering all holonomies, it is enough to consider a complete set in the following sense.

Suppose that R_i and R_j are Markov rectangles, and $x \in R_i$ and $y \in R_j$. If $\iota = s$ or u , then we say that x and y are ι -holonomically related if (i) there is an ι' -leaf segment $\ell^\iota(x, y)$ such that $\partial\ell^\iota(x, y) = \{x, y\}$, and (ii) there are two distinct spanning ι' -leaf segments $\ell^{\iota'}(x, R_i)$ and $\ell^{\iota'}(y, R_j)$ such that their union contains $\ell^\iota(x, y)$.

For every Markov rectangle $R_i \in \mathcal{R}$, let x_i be a chosen point in R_i . Let $\mathcal{I}^\iota = \{I_i = \ell^\iota(x_i, R_i) : R_i \in \mathcal{R}\}$. A complete system of ι -holonomies $\mathcal{H}^\iota = \{h_\alpha\}$ with respect to \mathcal{I}^ι consists of a minimal set of basic holonomies with the following property. If $x \in I_i$ is holonomically related to $y \in I_j$, where $I_i, I_j \in \mathcal{I}^\iota$, then for some α , either h_α or h_α^{-1} is the holonomy from a neighbourhood of x in I_i to I_j which sends x to y . We call \mathcal{I}^ι the domain of the complete system of ι -holonomies \mathcal{H}^ι . For each ι -leaf segment $I_i \in \mathcal{I}^\iota$, let \hat{I}_i be a full ι -leaf segment such that $I_i = \hat{I}_i \cap \Lambda$, and let $u_i : \hat{I}_i \rightarrow \mathbb{R}$ be a C^r ι -leaf chart of the submanifold structure of \hat{I}_i given by the stable manifold theorem. (For instance, we can consider the chart $u_i \in \mathcal{A}^\iota(\rho)$ as defined in Subsection 3.3.)

DEFINITION 3. A complete system of holonomies \mathcal{H}^ι is C^{1,HD^ι} if, for every holonomy $h_\alpha : I \rightarrow J$ in \mathcal{H}^ι with $I \subset I_i$ and $J \subset I_j$, the map $u_j \circ h_\alpha \circ u_i^{-1}$ and its inverse have a C^{1,HD^ι} diffeomorphic extension to \mathbb{R} such that the modulus of continuity does not depend upon $h_\alpha \in \mathcal{H}^\iota$.

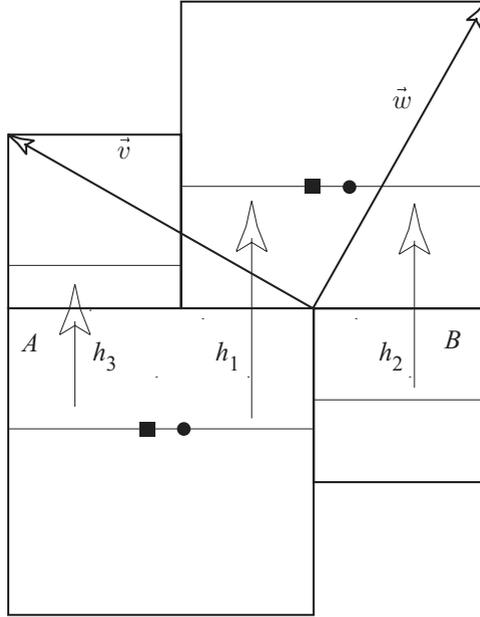


FIGURE 4. Complete set of holonomies $\mathcal{H} = \{h_1, h_2, h_3, h_1^{-1}, h_2^{-1}, h_3^{-1}\}$ for the Anosov map $g : \mathbf{R}^2 \setminus (\mathbf{Z}\vec{v} \times \mathbf{Z}\vec{w}) \rightarrow \mathbf{R}^2 \setminus (\mathbf{Z}\vec{v} \times \mathbf{Z}\vec{w})$ defined by $g(x, y) = (x + y, y)$ and with Markov partition $\mathcal{M} = \{A, B\}$.

For many systems such as Anosov diffeomorphisms and codimension 1 attractors or repellers, there are only a finite number of holonomies in a complete system (see Figure 4). In this case the uniformity hypothesis is redundant. However, for a Smale horseshoe, this is not the case (see Figure 5).

3. HR-structures with C^{1,HD^s} distortion

An HR-structure associates an affine structure to each stable and unstable leaf segment in such a way that these vary Hölder continuously with the leaf and are invariant under f . (The abbreviation HR stands for Hölder ratios.)

3.1. Hölder ratios

An affine structure on a stable or unstable leaf is equivalent to a ratio function $r(I : J)$ which can be thought of as prescribing the ratio of the size of two leaf segments I and J in the same stable or unstable leaf. A ratio function $r(I : J)$ is positive and continuous in the endpoints of I and J . Moreover,

$$r(I : J) = r(J : I)^{-1} \quad \text{and} \quad r(I_1 \cup I_2 : K) = r(I_1 : K) + r(I_2 : K), \quad (3.1)$$

provided that I_1 and I_2 intersect in at most one of their endpoints.

DEFINITION 4. We say that r is an ι -ratio function if (i) for all ι -leaf segments K , $r(I : J)$ ($I, J \subset K$) defines a ratio function on K , (ii) r is invariant under f , that is, $r(I : J) = r(fI : fJ)$ for all ι -leaf segments, and (iii) for every basic ι -holonomy map $h : I \rightarrow J$ between the leaf segment I and the leaf segment J defined with

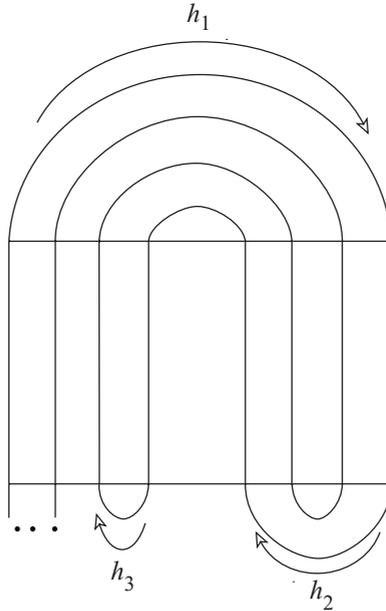


FIGURE 5. The cardinality of the complete set of holonomies $\mathcal{H} = \{h_1, h_2, h_3, \dots\}$ is not finite.

respect to a rectangle R , and for every ι -leaf segment $I_0 \subset I$ and every ι -leaf segment or gap $I_1 \subset I$,

$$\left| \log \frac{r(hI_0 : hI_1)}{r(I_0 : I_1)} \right| \leq \mathcal{O}((d_\Lambda(I, J))^\varepsilon), \tag{3.2}$$

where $\varepsilon \in (0, 1)$ depends upon r and the constant of proportionality also depends upon R , but not on the segments considered.

DEFINITION 5. An *HR-structure* is a pair (r^s, r^u) consisting of a stable and an unstable ratio function.

3.2. Lamination atlas $\mathcal{A}(r)$

Given an ι -ratio function r , we define the embeddings $e : I \rightarrow \mathbb{R}$ by

$$e(x) = r(\ell(\xi, x), \ell(\xi, R)), \tag{3.3}$$

where ξ is an endpoint of the ι -leaf segment I , $\delta(\ell(\xi, x)) = \{\xi, x\}$, and R is a Markov rectangle containing ξ (see Figure 6). For this definition it is not necessary that R contains I . We denote the set of all these embeddings e by $\mathcal{A}(r)$. Combining Proposition 2.5 and Proposition 3.5 of [20], we get the following result.

PROPOSITION 1. Let $h : I \rightarrow J$ be a basic holonomy between ι -leaf segments in a rectangle R . There is $0 < \eta < 1$ such that the holonomy h is $C^{1+\eta}$ with respect to the charts in $\mathcal{A}(r^\iota)$. Furthermore, there is $0 < \beta < 1$ with the property that for all charts $i : I \rightarrow \mathbb{R}$ and $j : J \rightarrow \mathbb{R}$ in $\mathcal{A}(r^\iota)$, there is an affine map $a : \mathbb{R} \rightarrow \mathbb{R}$ such

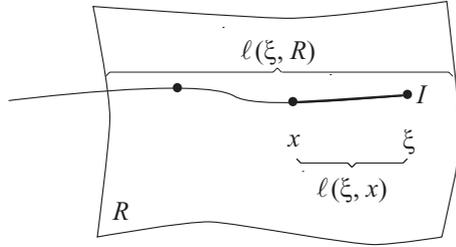


FIGURE 6. Definition of the embedding $e : I \rightarrow \mathbb{R}$.

that $j \circ h \circ i^{-1}$ has a $C^{1+\eta}$ diffeomorphic extension \tilde{h} and

$$\|\tilde{h} - a\|_{C^{1+\eta}} \leq \mathcal{O}((d_\Lambda(I, J))^\beta),$$

where η and β depend upon r^ι and the constant of proportionality also depends upon R .

3.3. Lamination atlas $\mathcal{A}^\iota(\rho)$

Let ρ be a C^{1+} Riemannian metric in the manifold containing Λ . The ι -lamination atlas $\mathcal{A}^\iota(\rho)$ determined by ρ is the set of all maps $e : I \rightarrow \mathbb{R}$, where $I = \Lambda \cap \hat{I}$, with \hat{I} a full ι -leaf segment, such that e extends to an isometry between the induced Riemannian metric on \hat{I} and the Euclidean metric on the reals. We call the maps $e \in \mathcal{A}^\iota(\rho)$ the ι -lamination charts. If I is an ι -leaf segment (or a full ι -leaf segment), then by $|I|_\rho$ we mean the length in the Riemannian metric ρ of the minimal full ι -leaf containing I .

3.4. Realisable ratio functions

By hyperbolicity of f in Λ , there are $0 < \nu < 1$ and $C > 0$ such that, for all ι -leaf segments I and all $m \geq 0$, we get

$$|f_{\nu^m} I|_\rho \leq C\nu^m |I|_\rho. \tag{3.4}$$

(Recall that $f_\iota = f$ if $\iota = u$ and $f_\iota = f^{-1}$ if $\iota = s$.) For every ι -leaf segment I , let us denote $|I|_\rho$ by $|I|$. Using the mean value theorem and the fact that f_ι is C^r with $r > 1$, for all short leaf segments K and all leaf segments I and J contained in K , the ι -realised ratio function r_f^ι given by

$$\begin{aligned} r_f^\iota(I : J) &= \lim_{n \rightarrow \infty} \frac{|f_{\nu^n} I|}{|f_{\nu^n} J|} \\ &= \frac{|f_{\nu^m} I|}{|f_{\nu^m} J|} \prod_{n=m}^{\infty} \left(\frac{|f_{\nu^{n+1}} I|}{|f_{\nu^{n+1}} J|} \frac{|f_{\nu^n} J|}{|f_{\nu^n} I|} \right) \\ &\in \frac{|f_{\nu^m} I|}{|f_{\nu^m} J|} \prod_{n=m}^{\infty} (1 \pm \mathcal{O}(\nu^n |K|^\alpha)) \\ &\subset \frac{|f_{\nu^m} I|}{|f_{\nu^m} J|} (1 \pm \mathcal{O}(\nu^m |K|^\alpha)) \end{aligned} \tag{3.5}$$

is well defined, where $\alpha = \min\{1, r - 1\}$. This construction gives the HR-structure on Λ determined by f . Combining Proposition 3.5 and Proposition 3.7 of [20], we get the following result.

PROPOSITION 2. *The diffeomorphism f determines a unique HR-structure on Λ given by (r_f^s, r_f^u) with the following property. For every ι -leaf segment I , if $e_1 : I \rightarrow \mathbb{R} \in \mathcal{A}(r_f^s)$ and $e_2 : I \rightarrow \mathbb{R} \in \mathcal{A}^\iota(\rho)$, then $e_2 \circ e_1^{-1}$ has a C^r extension to \mathbb{R} . If g is a C^s diffeomorphism on $\hat{\Lambda}$ which is C^{1+} conjugated to f , then it determines the same HR-structure as f , that is, $r_f^\iota(x, y, z) = r_g^\iota(\psi x, \psi y, \psi z)$, where $\psi : \Lambda \rightarrow \hat{\Lambda}$ is the C^{1+} conjugacy between f and g . Conversely, if f and g are topologically conjugate by $\psi : \Lambda \rightarrow \hat{\Lambda}$ and they determine the same HR-structure (that is, $r_f^\iota(x, y, z) = r_g^\iota(\psi x, \psi y, \psi z)$), then f and g are C^{1+} conjugated.*

3.5. C^{1,HD^ι} distortion

Consider an ι -ratio function r^ι and let $h : K \rightarrow K'$ be an ι -basic holonomy. We will consider two distinct cases, (i) (presence of gaps) when the ι -leaf segments have gaps, and (ii) (absence of gaps) when the ι -leaf segments do not have gaps.

Case (i) (presence of gaps): The *ratio distortion* of h in $I \subset K$ with respect to a ratio function r^ι is defined by

$$\text{rd}(h, I) = \sup_{I_0, I_1} \log \frac{r^\iota(hI_0 : hI_1)}{r^\iota(I_0 : I_1)},$$

where the supremum is over all pairs $I_0, I_1 \subset I$ such that I_0 is a leaf n -cylinder and I_1 is either a leaf n -cylinder or a leaf n -gap which has a unique common endpoint with I_0 and $n \geq 1$.

Case (ii) (absence of gaps): Suppose that J_0, J_1 and J_2 are distinct leaf n -cylinders such that J_0 and J_1 have a common endpoint, and J_1 and J_2 also have a common endpoint. Let J be the union of J_0, J_1 and J_2 . Then the *Poincaré length* with respect to a ratio function r^ι is defined by

$$P_{r^\iota}(J_1 : J) = \log \left(\frac{1 + r^\iota(J_1 : J_0)}{r^\iota(J_2 : J)} \right).$$

The *cross-ratio distortion* of h in $I \subset K$ with respect to a ratio function r^ι is defined by

$$\text{crd}(h, I) = \sup_{J_0, J_1, J_2} P_{r^\iota}(hJ_1 : hJ) - P_{r^\iota}(J_1 : J),$$

where the supremum is taken over all such triples J_0, J_1, J_2 with the property that $J \subset I$.

We observe that if $\text{rd}(h, I) = 0$, then h is affine on I , and if $\text{crd}(h, I) = 0$, then h is Möbius with respect to the atlas $\mathcal{A}(r^\iota)$ determined by r^ι . Here, for simplicity of exposition, we give a slightly different definition of cross-ratio distortion from the usual one (see [15]); however, this is equivalent for our purposes.

DEFINITION 6. The ratio function r^ι has $C^{1,\alpha}$ *distortion* with respect to a complete system of holonomies \mathcal{H}^ι if there is a modulus of continuity χ with the following properties.

- (i) $\lim_{t \rightarrow 0} \chi(t)/t^\alpha = 0$, that is, $\chi(t)$ is $o(t^\alpha)$.
- (ii) For every $h : K \rightarrow K'$ contained in \mathcal{H}^ι and for every ι -leaf segment $I \subset K$, let ξ be an endpoint of K and let R be a Markov rectangle containing ξ .
 - (a) If $\alpha < 1$, then the ι -leaf segments have gaps and $|\text{rd}(h, I)| \leq \chi(r^\iota(I, \ell(\xi, R)))$.
 - (b) If $\alpha = 1$, then the ι -leaf segments do not have gaps and $|\text{crd}(h, I)| \leq \chi(r^\iota(I, \ell(\xi, R)))$.

The following lemma gives the essential link between a $C^{1,\alpha}$ complete system of holonomies \mathcal{H}^ι and $C^{1,\alpha}$ distortion of r^ι with respect to \mathcal{H}^ι .

LEMMA 1. *Suppose that $0 < \alpha, \alpha' \leq 1$. Let (r_f^s, r_f^u) be the HR-structure determined by f on Λ . If $r - 1 > \max\{\alpha, \alpha'\}$ and there is a complete system of holonomies \mathcal{H}^ι for f in which the stable holonomies are $C^{1,\alpha}$ and the unstable holonomies are $C^{1,\alpha'}$, then r_f^s has $C^{1,\alpha}$ distortion and r_f^u has $C^{1,\alpha'}$ distortion with respect to \mathcal{H}^ι .*

Proof. For simplicity of notation, we will denote $|J|_\rho$ by $|J|$ for every ι -leaf segment. Let $h : K \rightarrow K'$ be a $C^{1,\alpha}$ holonomy in the ι -complete set of holonomies. Let ξ be an endpoint of K and let R be a Markov rectangle containing ξ . We will prove separately the cases where (i) $0 < \alpha < 1$, and (ii) $\alpha = 1$. For simplicity of notation, we will denote r_f^ι by r . Let $I \subset K$ be an ι -leaf segment. Using inequalities (3.2) and (3.5), we obtain

$$|hI| < \mathcal{O}(r(I, \ell(\xi, R))) \quad \text{and} \quad |I| < \mathcal{O}(r(I, \ell(\xi, R))). \tag{3.6}$$

Case (i): Let I_0, I_1 be disjoint ι -leaf segments contained in $I \subset K$ such that I_0 is a leaf n -cylinder and I_1 is either a leaf n -cylinder or a leaf n -gap which has a common endpoint with I_0 . From inequalities (3.5) and (3.6), we get

$$\frac{r(hI_0 : hI_1)}{r(I_0 : I_1)} \in \frac{|hI_0| |I_1|}{|hI_1| |I_0|} (1 \pm \mathcal{O}((r(I, \ell(\xi, R)))^\beta)), \tag{3.7}$$

where $\beta = \min\{1, r - 1\}$. Since h is $C^{1,\alpha}$, using the mean value theorem we get

$$\frac{|hI_0| |I_1|}{|hI_1| |I_0|} \in (1 \pm o((r(I, \ell(\xi, R)))^\alpha)). \tag{3.8}$$

Noting that $\alpha < \beta$ and putting (3.7) together with (3.8), we obtain

$$\frac{r(hI_0 : hI_1)}{r(I_0 : I_1)} \in (1 \pm o((r(I, \ell(\xi, R)))^\alpha)).$$

Therefore, for every ι -leaf segment $I \subset K$, we have $|\text{rd}(h, I)| \leq o(r(I, \ell(\xi, R)))^\alpha$.

Case (ii): Let J_0, J_1 and J_2 be leaf n -cylinders contained in an ι -leaf segment $I \subset K$ such that J_0 and J_1 have a common endpoint, and J_1 and J_2 also have a common endpoint. Let J be the union of J_0, J_1 and J_2 . Let

$$P_\rho(J_1 \subset J) = \log \left(1 + \frac{|J_1|_\rho |J|_\rho}{|J_0|_\rho |J_2|_\rho} \right). \tag{3.9}$$

Since f_ι is C^r with $r > 2$, from Lemma 3 (see Appendix A), (3.4) and (3.6) we get

$$\begin{aligned} P_\rho(f_\iota^{-(n+1)} J_1 : f_\iota^{-(n+1)} J) - P_\rho(f_\iota^{-n} J_1 : f_\iota^{-n} J) &\in o(\nu^n |J|_\rho) \\ &\subset \pm o(\nu^n r^\iota(J, \ell(\xi, R))). \end{aligned}$$

Therefore,

$$\begin{aligned} P_{r^\iota}(J_1 : J) &= \lim_{n \rightarrow \infty} P_\rho(f_\iota^{-n} J_1 : f_\iota^{-n} J) \\ &= P_\rho(f_\iota^{-m} J_1 : f_\iota^{-m} J) \\ &\quad + \sum_{n=m}^\infty (P_\rho(f_\iota^{-(n+1)} J_1 : f_\iota^{-(n+1)} J) - P_\rho(f_\iota^{-n} J_1 : f_\iota^{-n} J)) \\ &\subset P_\rho(f_\iota^{-m} J_1 : f_\iota^{-m} J) \pm o(\nu^m r^\iota(J, \ell(\xi, R))). \end{aligned}$$

Thus, since h is $C^{1,1}$, and from Lemma 3, we get

$$\begin{aligned} P_r(hJ_1 : hJ) - P_{r^\iota}(J_1 : J) &= \lim_{n \rightarrow \infty} (P_\rho(f_\iota^{-n} hJ_1 : f_\iota^{-n} hJ) - P_\rho(f_\iota^{-n} J_1 : f_\iota^{-n} J)) \\ &\in P_\rho(hJ_1 : hJ) - P_\rho(J_1 : J) \pm o(r(J, \ell(\xi, R))) \\ &\subset \pm o(r(J, \ell(\xi, R))). \end{aligned}$$

Therefore, for every ι -leaf segment $I \subset K$, we have

$$|\text{crd}(h, I)| \leq o(r(I, \ell(\xi, R))). \quad \square$$

4. Fundamental rigidity lemma

Let (r^s, r^u) be the HR-structure determined by f on Λ . Let HD^s and HD^u be, respectively, the Hausdorff dimensions of the stable and unstable leaves intersected with the hyperbolic invariant set Λ .

THEOREM 2 (fundamental rigidity lemma). *If the ι -ratio function r^ι has C^{1,HD^ι} distortion, then all basic holonomies are affine with respect to the atlas $\mathcal{A}(r^\iota)$, that is, they leave r^ι invariant.*

To prove Theorem 2, we have to introduce Markov maps on train-tracks, and we have to use Gibbs states to analyse the properties of the Hausdorff measures on the train-tracks.

4.1. Train-tracks

Roughly speaking, train-tracks are the optimal leaf-quotient spaces on which the stable and unstable Markov maps induced by the action of f on leaf segments are local homeomorphisms.

For each Markov rectangle R , let t'_R be the set of ι' -segments of R . Thus by the local product structure one can identify t'_R with any spanning ι -leaf segment $\ell^\iota(x, R)$ of R .

We form the space \mathbf{B}^ι by taking the disjoint union $\bigsqcup_{R \in \mathcal{R}} t'_R$ (union over all Markov rectangles R of the Markov partition \mathcal{R}) and identifying two points $I \in t'_R$ and $J \in t'_{R'}$ if either (i) the ι' -leaf segments I and J are ι' -boundaries of Markov rectangles and their intersection contains at least a point which is not an endpoint of I or J , or (ii) there is a sequence $I = I_1, \dots, I_n = J$ such that all I_i, I_{i+1} are both identified in the sense of (i). This space is called the ι -train-track and is denoted by \mathbf{B}^ι .

Let $\pi_\iota : \bigsqcup_{R \in \mathcal{R}} R \rightarrow \mathbf{B}^\iota$ be the natural projection sending $x \in R$ to the point in \mathbf{B}^ι represented by $\ell^{\iota'}(x, R)$. A *topologically regular point* I in \mathbf{B}^ι is a point with a unique preimage under π_ι (that is, the preimage of I is not a union of distinct ι' -boundaries of Markov rectangles). If a point has more than one preimage by π_ι ,

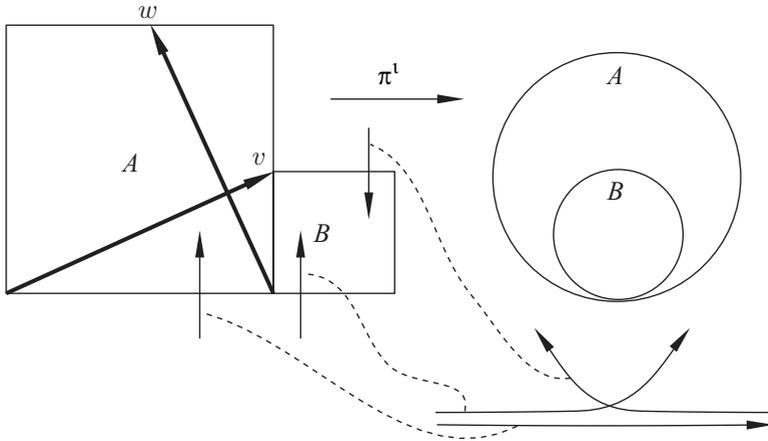


FIGURE 7. (Unstable) train-track for an Anosov diffeomorphism. (The rectangles A and B are the Markov rectangles and the vertical arrows show paths along unstable manifolds from A to A and from B to A . The train-track is represented by the pair of circles, and the curves below it show the smooth paths through the junction of the two circles which arise from the smooth paths between the rectangles A and B along unstable manifolds. Note that there is no smooth path from B to B even though in this representation of the train-track it looks as though there ought to be. This is because there is no unstable manifold running directly from the rectangle B to itself.)

then we call it a *junction*. Since there are only a finite number of ι' -boundaries of Markov rectangles, there are only finitely many junctions (see Figure 7).

We say that I_T is a *train-track segment* if there is an ι' -leaf segment I , not intersecting ι' -boundaries of Markov rectangles, such that $\pi_{\iota'}|I$ is an injection and $\pi_{\iota'}(I) = I_T$. Furthermore, the chart $i : I \rightarrow \mathbb{R}$ in $\mathcal{A}(r^{\iota'})$ determines a *train-track chart* $i_T : I_T \rightarrow \mathbb{R}$ for I_T given by $i_T = i \circ \pi_{\iota'}^{-1}$. We denote by $\mathcal{B}(r^{\iota'})$ the set of all train-track charts for all train-track segments determined by $\mathcal{A}(r^{\iota'})$. Given any train-track charts $i_T : I_T \rightarrow \mathbb{R}$ and $j_T : J_T \rightarrow \mathbb{R}$ in $\mathcal{B}(r^{\iota'})$, the overlap map $j_T \circ i_T^{-1} : i_T(I_T \cap J_T) \rightarrow j_T(I_T \cap J_T)$ is equal to $j_T \circ i_T^{-1} = j \circ h \circ i^{-1}$, where $i = i_T \circ \pi_{\iota'} : I \rightarrow \mathbb{R}$ and $j = j_T \circ \pi_{\iota'} : J \rightarrow \mathbb{R}$ are charts in $\mathcal{A}(r^{\iota'})$, and

$$h : i^{-1}(i_T(I_T \cap J_T)) \rightarrow j^{-1}(j_T(I_T \cap J_T))$$

is a basic ι' -holonomy. By Proposition 1, there is $\eta > 0$ such that, for all train-track charts i_T and j_T in $\mathcal{B}(r^{\iota'})$, the overlap maps $j_T \circ i_T^{-1} = j \circ h \circ i^{-1}$ have $C^{1+\eta}$ diffeomorphic extensions with a uniform bound for their $C^{1+\eta}$ norm. Hence $\mathcal{B}(r^{\iota'})$ is a $C^{1+\eta}$ atlas for the train-track segments in $\mathbf{B}^{\iota'}$.

4.2. Markov maps

The Markov map $m_{\iota'} : \mathbf{B}^{\iota'} \rightarrow \mathbf{B}^{\iota'}$ is the mapping induced by the action of f on leaf segments, that is, it is defined as follows. If $I \in \mathbf{B}^{\iota'}$, $m_{\iota'}I$ is the ι' -leaf segment containing the $f_{\iota'}$ -image of the ι' -leaf segment I . This map is a local homeomorphism because $f_{\iota'}$ sends a short ι' -leaf segment homeomorphically onto a short ι' -leaf segment. Since f on Λ along leaves has affine extensions with respect to the charts in $\mathcal{A}(r^{\iota'})$ and the basic ι' -holonomies have $C^{1+\eta}$ extensions, we get that the Markov maps $m_{\iota'}$ also have $C^{1+\eta}$ extensions with respect to the charts in $\mathcal{B}(r^{\iota'})$ for some $\eta > 0$.

An n -cylinder of m_ι is the projection under π_ι of an ι -leaf n -cylinder, and an n -gap of m_ι is the projection under π_ι of an ι -leaf n -gap. Let us denote by $|I|_i$ the length of the smallest interval containing the image under $i \in \mathcal{B}(r^\iota)$ of a train-track segment I of m_ι . By hyperbolicity of f in Λ , the train-track atlas $\mathcal{B}(r^\iota)$ has bounded geometry, that is, there is $C > 1$ such that if I is an n -cylinder and J is an n -cylinder or an n -gap with a common endpoint with I such that $I \cup J$ is a train-track segment, then $C^{-1} < |I|_i/|J|_i < C$ for all charts $i \in \mathcal{B}(r^\iota)$ whose domains contain $I \cup J$. In particular, there are $0 < \nu < 1$ and $C > 0$ such that

$$|I|_i < C\nu^n \tag{4.1}$$

for all n -cylinders or n -gaps I of m_ι and for all charts $i \in \mathcal{B}(r^\iota)$ whose domains contain I .

We use the following proposition in the proof of the fundamental rigidity lemma. It can be deduced from standard results about Gibbs states such as those in [3], and it also follows from the results proved in [19].

PROPOSITION 3. *There is a unique m_ι -invariant probability measure μ on \mathbf{B}^ι such that, if δ is the Hausdorff dimension of \mathbf{B}^ι , then there is $C \geq 1$ such that*

$$C^{-1} \leq \frac{\mu(I)}{|I|_i^\delta} \leq C$$

for all n -cylinders I , for all $n \geq 1$, and for all train-track charts $i \in \mathcal{B}(r^\iota)$. It follows from this that the Hausdorff δ -measure \mathcal{H}^δ is finite and positive on \mathbf{B}^ι , and μ is absolutely continuous (equivalent) with respect to \mathcal{H}^δ .

Proof of Theorem 2. We shall prove Theorem 2 for the stable holonomies. The unstable result is proved in the same way by replacing f by f^{-1} .

Let $h : I \rightarrow I'$ be a basic stable holonomy in the rectangle R , where I and I' are stable spanning leaves of R and R has the property that every spanning stable and unstable leaf segment of R is either contained inside a single primary cylinder or inside the union of two touching primary cylinders. We shall prove that since there is a complete set of holonomies with C^{1,HD^s} distortion, h has an affine extension with respect to the charts in $\mathcal{A}(r^s)$.

For every $n \geq 1$, the rectangle $f^n R$ is equal to $\bigcup_{j=0}^{m(n)} M_j^n$, where the rectangles $M_j^n = [J_j^n, U_j^n]$ have the following properties (see Figure 8).

- (i) For j equal to 0 and $m(n)$, we have the following.
 - (a) $f^n I = J_0^n$ and $f^n I' = J_{m(n)}^n$.
 - (b) If J_j^n is contained in a single Markov rectangle, then U_j^n is an unstable spanning leaf of this Markov rectangle intersected with $f^n R$.
 - (c) If J_j^n is not contained in a Markov rectangle, then U_j^n is the biggest possible unstable leaf segment in $f^n R$ contained in the union of the unstable boundaries of Markov rectangles and intersecting J_j^n .
- (ii) For $j = 1, \dots, m(n) - 1$, one of the following holds.
 - (a) J_j^n is a stable spanning leaf segment of M_j^n contained in a leaf segment of the domain \mathcal{I}^ι of the complete system of holonomies \mathcal{H}^ι , and U_j^n is an unstable spanning leaf segment of the Markov rectangle containing J_j^n .
 - (b) J_j^n is a stable leaf segment not contained in a single Markov rectangle, and U_j^n is the biggest possible unstable leaf segment contained in the

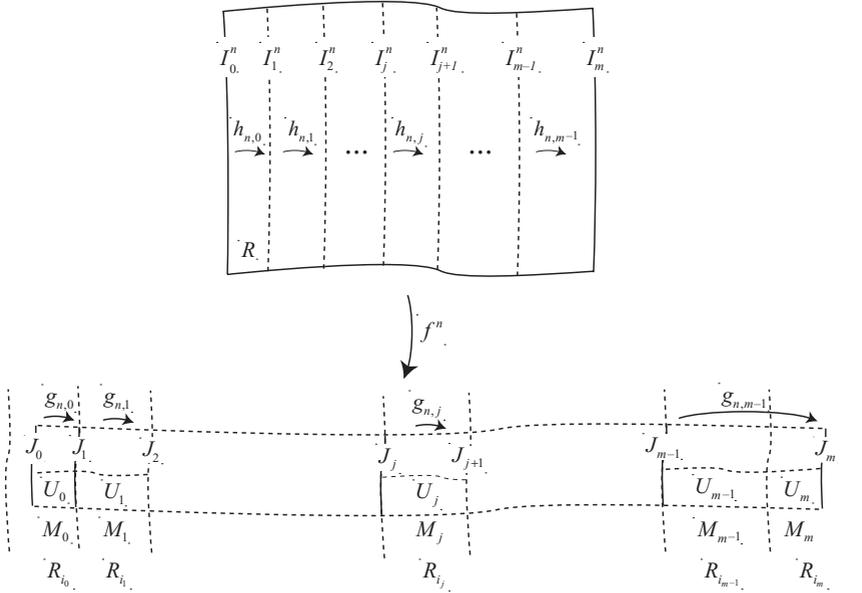


FIGURE 8. The rectangles R and $f^n R$.

- union of the unstable boundaries of Markov rectangles and intersecting J_j^n .
- (iii) M_j^n intersects M_{j+1}^n only along a common stable boundary, and $M_i^n \cap M_j^n = \emptyset$ if $|j - i| \geq 1$.

Let Θ_n^a be the set of $j \in \{1, \dots, m(n) - 1\}$ such that J_j^n and J_{j+1}^n are contained in the domain \mathcal{I}^u , and let Θ_n^b be equal to $\{0, \dots, m(n) - 1\} \setminus \Theta_n^a$. Since the number of Markov rectangles is finite, the cardinality of the set Θ_n^b is uniformly bounded independently of n .

Set $I_j^n = f^{-n} J_j^n$. Then we can decompose h as the composition $h_{n,m-1} \circ \dots \circ h_{n,0}$, where $h_{n,j}$ is the basic holonomy between I_j and I_{j+1} defined by R . Now consider the holonomies $g_{n,j} = f^n \circ h_{n,j} \circ f^{-n} : J_j^n \rightarrow J_{j+1}^n$ and observe that since f is affine in the HR structure, $\text{rd}(h_{n,j}, I_j^n) = \text{rd}(g_{n,j}, J_j^n)$ and $\text{crd}(h_{n,j}, I_j^n) = \text{crd}(g_{n,j}, J_j^n)$. Furthermore, if $j \in \Theta_n^a$, then $g_{n,j}$ belongs to the complete system of holonomies. Let us first consider the case where $HD^s < 1$. By hypotheses, for every $j \in \Theta_n^a$ we have

$$\sum_{j \in \Theta_n^a} |\text{rd}(g_{n,j}, J_j^n)| \leq \sum_{j \in \Theta_n^a} \chi(r(J_j^n, \ell(x_j^n, R_j^n))),$$

where x_j^n is an endpoint of J_j^n , R_j^n is a Markov rectangle containing x_j^n , and the positive function χ is independent of h and $\chi(t) = o(t^{HD^s})$. From inequality (3.2), for every $j \in \Theta_n^b$ we get

$$\sum_{j \in \Theta_n^b} |\text{rd}(h_{n,j}, I_j^n)| \leq \sum_{j \in \Theta_n^b} \mathcal{O}(d_\Lambda(I_j^n, I_{j+1}^n)^\alpha).$$

Therefore,

$$\begin{aligned} |\text{rd}(h, I)| &\leq \sum_{j=0}^{m-1} |\text{rd}(h_{n,j}, I_j^n)| \\ &\leq \sum_{j \in \Theta_n^b} |\text{rd}(h_{n,j}, I_j^n)| + \sum_{j \in \Theta_n^a} |\text{rd}(g_{n,j}, J_j^n)| \\ &\leq \sum_{j \in \Theta_n^b} \mathcal{O}(d_\Lambda(I_j^n, I_{j+1}^n)^\alpha) + \sum_{j \in \Theta_n^a} \chi(r(J_j^n, \ell(x_j^n, R_j^n))). \end{aligned}$$

Now, we note that

$$r(J_j^n, \ell(x_j^n, R_j^n)) \leq \mathcal{O}(|K_j^n|),$$

where $K_j^n = \pi^s(J_j^n)$ is the projection of J_j^n into the train-track \mathbf{B}^s under π_s , and the size $|K_j^n|$ of K_j^n is measured in any chart of the bounded atlas $\mathcal{B}(r^s)$ of \mathbf{B}^s . Therefore,

$$|\text{rd}(h, I)| \leq \sum_{j \in \Theta_n^b} \mathcal{O}(d_\Lambda(I_j^n, I_{j+1}^n)^\alpha) + \sum_{j \in \Theta_n^a} \hat{\chi}(|K_j^n|), \tag{4.2}$$

where $\hat{\chi}$ is a positive function independent of h , and $\hat{\chi}(t) = o(t^{HD^s})$. In the case where $HD_s = 1$, a similar argument gives

$$|\text{crd}_{h,I}| \leq \sum_{j \in \Theta_n^b} \mathcal{O}(d_\Lambda(I_j^n, I_{j+1}^n)^\alpha) + \sum_{j \in \Theta_n^a} C_1 \hat{\chi}(|K_j^n|), \tag{4.3}$$

where $\hat{\chi}$ is a positive function independent of h and $\hat{\chi}(t) = o(t)$. We now show that the right-hand sides of (4.2) and (4.3) tend to zero as n tends to infinity and thus that the left-hand sides are zero. For every $j \in \Theta_n^b$, the distance $d_\Lambda(I_j^n, I_{j+1}^n)$ converges to zero when n tends to infinity, and since the cardinal of Θ_n^b is uniformly bounded independently of n , we get

$$\sum_{j \in \Theta_n^b} \mathcal{O}(d_\Lambda(I_j^n, I_{j+1}^n)^\alpha) \rightarrow 0 \tag{4.4}$$

when n tends to infinity. Now we are going to prove that $\sum_{j \in \Theta_n^a} \chi(|K_j^n|)$ also converges to zero when n tends to infinity. Since R has the property that every spanning stable leaf segment of R is either contained inside a single primary cylinder or inside the union of two touching primary cylinders, we obtain that the train-track segments K_j^n can only intersect in endpoints, and moreover each of them is either contained in an n -cylinder or two adjacent n -cylinders of the Markov map m_s on \mathbf{B}^s . Hence, using (4.1), there is a continuous positive function η with $\eta(0) = 0$ such that

$$\sum_{j \in \Theta_n^a} \chi(|K_j^n|) \leq \eta(\nu^n) \sum_{n\text{-cyls}} |C^n|^{HD^s}, \tag{4.5}$$

where the sum on the right-hand side is over all n -cylinders. By Proposition 3, there is an m_s -invariant probability measure μ and a positive constant C_1 such that

$$\sum_{n\text{-cyls}} |C^n|^{HD^s} \leq C_1 \sum_{n\text{-cyls}} \mu(C^n) \leq C_1. \tag{4.6}$$

Putting together (4.5) and (4.6), we get

$$\sum_{j \in \Theta_n^a} \chi(|K_j^n|) \rightarrow 0 \tag{4.7}$$

when n tends to infinity. If $HD^s < 1$, applying (4.4) and (4.7) to (4.2), we get $rd(h, I) = 0$. Therefore, h is affine on I , which completes the proof for this case. If $HD^s = 1$, applying (4.4) and (4.7) to (4.3), we get $crd(h, I) = 0$. Therefore, h is Möbius on I and extends to a Möbius homeomorphism of the global leaf, where the affine structures of the global leaves are determined by the invariance of the affine structures under iteration by f . Since a Möbius homeomorphism of \mathbb{R} is an affine map, the holonomy maps h are affine. □

5. Existence of affine models

LEMMA 2 (existence of affine models). *If r^s has C^{1,HD^s} distortion and r^u has C^{1,HD^u} distortion, then there is a hyperbolic affine model for g on $\hat{\Lambda}$ which is topological conjugated to f on Λ and is such that the HR-structures are the same, (that is, $r^\iota(x, y, z) = r'_g(\psi x, \psi y, \psi z)$ for $\iota \in \{s, u\}$, where $\psi : \Lambda \rightarrow \hat{\Lambda}$ is the conjugacy between f and g).*

In Lemma 2, (r_s, r_u) is any HR-structure and not necessarily the HR structure determined by f .

Proof of Lemma 2. Let $\{R_1, \dots, R_k\}$ be a Markov partition for f . For every Markov rectangle R_m , we take a rectangle $M_m \supset R_m$ which contains a small neighbourhood of R_m with respect to the distance d_Λ . We construct an orthogonal chart $i_m : M_m \rightarrow \mathbb{R}^2$ as follows. Choose an $x \in M_m$ and let $e_s : \ell^s(x, M_m) \rightarrow \mathbb{R}$ be in $\mathcal{A}(r^s)$ and $e_u : \ell^u(x, M_m) \rightarrow \mathbb{R}$ be in $\mathcal{A}(r^u)$. The orthogonal chart i_m on M_m is now given by $i_m(z) = (e_s([z, x]), e_u([x, z])) \in \mathbb{R}^2$.

Let $\phi_{m,n} : i_m(M_m \cap M_n) \rightarrow i_n(M_m \cap M_n)$ be the map defined by $\phi_{m,n}(x) = i_m \circ i_n^{-1}(x)$. By Theorem 2, the stable and unstable holonomies have affine extensions with respect to the charts in $\mathcal{A}(r^s)$ and $\mathcal{A}(r^u)$. Hence there is a unique affine extension $\Phi_{m,n} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ of $\phi_{m,n}$. This extension sends vertical lines into vertical lines and horizontal lines into horizontal lines.

Let us denote by S_m the rectangle in \mathbb{R}^2 whose boundary contains the image under i_m of the boundary of R_m . For every pair of Markov rectangles R_m and R_n which intersect in a partial side $I_{m,n} = R_m \cap R_n$, let $J_{m,n}$ and $J_{n,m}$ be the smallest line segments containing the sets $i_m(I_{m,n})$ and $i_n(I_{m,n})$ respectively. We call $J_{m,n}$ and $J_{n,m}$ *partial sides*. Hence $J_{m,n} = \Phi(J_{n,m})$. Let $\tilde{R} = \bigsqcup_{m=1}^k S_m / \{\Phi_{m,n}\}$ be the disjoint union of the squares, S_m where we identify two points $x \in J_{m,n}$ and $y \in J_{n,m}$ if $\Phi_{n,m}(x) = y$. Hence \tilde{R} is a topological surface, possibly with boundary. By taking appropriate extensions E_m of the rectangles S_m and using the maps $\Phi_{m,n}$ to determine the identifications along the boundaries, we get a surface $\hat{R} = \bigsqcup_{m=1}^k E_m / \{\Phi_{m,n}\}$ without boundary. The surface \hat{R} has a natural affine atlas that we now describe. If a point z is contained in the interior of E_m , then we take a small open neighbourhood U_z of z contained in E_m and we define a chart $u_z : U_z \rightarrow \mathbb{R}^2$ as being the inclusion of $U_z \cap E_m$ into \mathbb{R}^2 . Otherwise z is contained in a boundary of two, three or four sets E_{m_1}, \dots, E_{m_k} , which we order such that the $J_{m_i, m_{i+1}}$ are partial sides. In this case, for a small open neighbourhood U_z of z , we define the chart $u_z : U_z \rightarrow \mathbb{R}^2$ as follows.

- (i) $u_z|_{(U_z \cap E_{m_k})}$ is the inclusion of $U_z \cap E_{m_k}$ into \mathbb{R}^2 .
- (ii) $u_z|_{(U_z \cap E_j)} = \Phi_{m_{k-1}, m_k} \circ \dots \circ \Phi_{m_j, m_{j+1}}$ for $j \in \{1, \dots, k-1\}$.

Since the maps $\Phi_{m_1, m_2}, \dots, \Phi_{m_{k-1}, m_k}$ and Φ_{m_k, m_1} are affine, we deduce that the set of all these charts form an affine atlas \mathcal{S} on \hat{M} .

Let $\psi : \Lambda \rightarrow \hat{\Lambda}$ be the natural embedding of Λ into \hat{M} , and let $G : \hat{\Lambda} \rightarrow \hat{\Lambda}$ be the map $G = \psi \circ f \circ \psi^{-1}$ conjugate to f . For every $x \in \hat{\Lambda}$, we take charts $u : U \rightarrow \mathbb{R}^2$ and $v : V \rightarrow \mathbb{R}$ in the affine atlas \mathcal{S} such that $x \in U$ and $\hat{f}x \in V$. Since G along leaves and also the holonomies have affine extensions with respect to the charts in $\mathcal{A}(r^s)$ and $\mathcal{A}(r^u)$, the map $v \circ \hat{f} \circ u^{-1}$ has a unique affine extension g_x to \mathbb{R}^2 . These affine extensions determine a unique affine extension g of G to an open set of \hat{M} .

The maps g_x send horizontal lines into horizontal lines and vertical lines into vertical lines. Furthermore, $g_{g^n x} \circ \dots \circ g_x$ contracts horizontal lines exponentially fast and expands vertical lines exponentially fast with respect to any fixed finite set of charts in \mathcal{S} covering \hat{M} . Hence g is hyperbolic on $\hat{\Lambda}$ and the image under these charts of the stable and unstable leaves are contained in horizontal and vertical lines respectively.

Since the holonomies have affine extensions with respect to the charts in $\mathcal{A}(r^s)$ and $\mathcal{A}(r^u)$, they also have affine extensions along leaves with respect to the charts in this affine atlas. By construction of the affine model for g on $\hat{\Lambda}$, we get $r^\iota(x, y, z) = r_g^\iota(\psi x, \psi y, \psi z)$ for $\iota \in \{s, u\}$. \square

6. Proof of the hyperbolic and Anosov rigidity

Here we show how to use the fundamental rigidity lemma and the existence of affine models (Lemma 2) to prove Theorem 1 and Corollary 1.

Proof of Theorem 1. By Proposition 2, f determines on Λ an HR-structure (r^s, r^u) . Let $\iota \in \{s, u\}$. By Lemma 1, r^ι has C^{1, HD^ι} distortion. By Theorem 2, all the ι -basic holonomies are affine with respect to the atlas $\mathcal{A}(r^\iota)$. Hence, by Lemma 2, there is a diffeomorphism g with a hyperbolic invariant set $\hat{\Lambda}$ and a hyperbolic affine model for g on $\hat{\Lambda}$ such that there is a conjugacy between f and g such that $r_f^\iota(x, y, z) = r_g^\iota(\psi x, \psi y, \psi z)$ for $\iota \in \{s, u\}$. By Proposition 2, we get f is C^{1+} conjugated to g . \square

Before proving Corollary 1, we state a proposition due to Journé [12] that we will use in the proof.

PROPOSITION 4. *If f is a continuous function in an open set $V \subset \mathbb{R}^2$ which is C^r along the leaves of two transverse foliations with uniformly smooth leaves, then f is C^r .*

We note that Corollary 1 also follows from the fact that the holonomies and f are affine with respect to the atlases $\mathcal{A}(r^s)$ and $\mathcal{A}(r^u)$ (see the proof of Theorem 1) and [8, Corollary 3.3].

Proof of Corollary 1. If $f : M \rightarrow M$ is a C^r surface Anosov diffeomorphism, then $\Lambda = M$. By [6, 7, 14, 16], there is a unique hyperbolic toral automorphism $\hat{f} : \hat{M} \rightarrow \hat{M}$ topologically conjugate to f . By Theorem 1, there is a C^{1+} conjugacy $\psi : M \rightarrow \hat{M}$ between f and \hat{f} . By Proposition 2, we have $r_f^\iota(x, y, z) = r_{\hat{f}}^\iota(\psi x, \psi y, \psi z)$ for $\iota \in \{s, u\}$. By a somewhat standard blow-down–blow-up argument, we get ψ is

C^r along stable and unstable leaves (see [15, 22]). Hence, by Proposition 4, ψ is C^r . □

Acknowledgements. We are grateful to Dennis Sullivan for a number of very fruitful and useful discussions on this work. We also thank Flávio Ferreira for his helpful comments. An anonymous referee of this paper made some very helpful comments which enabled us to improve the paper. Part of this research was started during a visit by the authors to the IHES and CUNY, and continued at the Isaac Newton Institute, IMPA and Stonybrook. We thank them for their hospitality.

Appendix A. Proof of Lemma 3

In this appendix we prove Lemma 3, which is used in the proof of Lemma 1.

For a complete discussion on the relations between smoothness of diffeomorphisms and cross-ratio distortions, see [15, 22].

Let $\theta : J \rightarrow K$ be either a holonomy h or f_ι , and let J and K be ι -leaf segments. Let $I_0, I_1, I_2 \subset J$ be leaf n -cylinders such that I_0 is adjacent to I_1 , and I_1 is adjacent to I_2 and $I = I_0 \cup I_1 \cup I_2$. Let $\mathcal{A}^\iota(\rho)$ be an ι -lamination atlas induced by a Riemannian metric ρ on the surface, and let $|I'| = |I'|_\rho$ for every ι -leaf segment I' . We define $B(I_0, I_1, I_2)$ and $B_\theta(I_0, I_1, I_2)$ as follows. Let

$$B(I_0, I_1, I_2) = \frac{|I_1||I|}{|I_0||I_2|}$$

$$B_\theta(I_0, I_1, I_2) = \frac{|\theta I_1||\theta I|}{|\theta I_0||\theta I_2|}.$$

We define the cross-ratio distortion $\text{crd}_{\theta,\rho}(I_0, I_1, I_2)$ of θ with respect to $\mathcal{A}^\iota(\rho)$ by

$$\text{crd}_{\theta,\rho}(I_0, I_1, I_2) = \log(1 + B_\theta(I_0, I_1, I_2)) - \log(1 + B(I_0, I_1, I_2)).$$

We note that for $\varepsilon > 0$, a $C^{2+\varepsilon}$ diffeomorphism θ is a $C^{1,1}$ diffeomorphism (see [15]).

LEMMA 3. *Let $\theta : J \subset \mathbb{R} \rightarrow K \subset \mathbb{R}$ be a $C^{1,1}$ diffeomorphism with respect to the atlas $\mathcal{A}(\rho)$. Then*

$$\text{crd}_{\theta,\rho}(I_0, I_1, I_2) \leq o(|I|)$$

for all $n \geq 1$ and for all n -cylinders $I_0, I_1, I_2 \in J$ such that I_0 is adjacent to I_1 , I_1 is adjacent to I_2 , and $I = I_0 \cup I_1 \cup I_2$.

Proof. By [15, theorem, p. 294], we get

$$|B_\theta(I_0, I_1, I_2) - B(I_0, I_1, I_2)| \leq o(|I|B(I_0, I_1, I_2)). \tag{A.1}$$

Therefore,

$$|\text{crd}_{\theta,\rho}(I_0, I_1, I_2)| = \left| \log \left(1 + \frac{B_\theta(I_0, I_1, I_2) - B(I_0, I_1, I_2)}{1 + B(I_0, I_1, I_2)} \right) \right|$$

$$\leq o \left(\frac{|I|B(I_0, I_1, I_2)}{1 + B(I_0, I_1, I_2)} \right) \leq o(|I|). \quad \square$$

References

1. V. I. ARNOL'D, 'Small denominators. I: On the mapping of a circle into itself', *Investijia Akad. Nauk Math.* 25 (1961) 21–96 (Russian), *Transl. Amer. Math. Soc.* 46, 213–284 (English).
2. A. AVEZ, 'Anosov diffeomorphisms', *Topological Dynamics. An International Symposium* (ed. W. Gottschalk and J. Auslander, W. A. Benjamin, New York, 1968) 17–51.
3. R. BOWEN, *Equilibrium states and the ergodic theory of Axiom A diffeomorphisms*, Lecture Notes in Mathematics 470 (Springer, New York, 1975).
4. F. F. FERREIRA, A. A. PINTO and D. A. RAND, 'Non-existence of affine models for attractors on surfaces', manuscript.
5. L. FLAMINIO and A. KATOK, 'Rigidity of symplectic Anosov diffeomorphisms on low dimensional tori', *Ergodic Theory Dynam. Systems* 11 (1991) 427–440.
6. J. FRANKS, 'Anosov diffeomorphisms on tori', *Trans. Amer. Math. Soc.* 145 (1969) 117–124.
7. J. FRANKS, 'Anosov diffeomorphisms', *Global analysis* (ed. S. Smale, American Mathematical Society, Providence, RI, 1970) 61–93.
8. E. GHYS, 'Rigidité différentiable des groupes Fuchsien', *Inst. Hautes Études Sci. Publ. Math.* 78 (1993) 163–185.
9. M. R. HERMAN, 'Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations', *Inst. Hautes Études Sci. Publ. Math.* 49 (1979) 5–233.
10. S. HURDER and A. KATOK, 'Differentiability, rigidity and Godbillon–Vey classes for Anosov flows', *Inst. Hautes Études Sci. Publ. Math.* 72 (1990) 5–61.
11. J. L. JOURNÉ, 'On a regularity problem occurring in connection with Anosov diffeomorphisms', *Comm. Math. Phys.* 106 (1986) 345–352.
12. J. L. JOURNÉ, 'A regularity lemma for functions of several variables', *Rev. Mat. Iberoamericana* 4 (1988) 187–193.
13. R. MAÑÉ, *Ergodic theory and differentiable dynamics* (Springer, Berlin, 1987).
14. A. MANNING, 'There are no new Anosov diffeomorphisms on tori', *Amer. J. Math.* 96 (1974) 422.
15. W. MELO and S. STRIEN, *One-dimensional dynamics*, Modern Surveys in Mathematics (Springer, New York, 1993).
16. S. NEWHOUSE, 'On codimension one Anosov diffeomorphisms', *Amer. J. Math.* 92 (1970) 671–762.
17. A. A. PINTO and D. A. RAND, 'Classifying C^{1+} structures on hyperbolic fractals. 1: The moduli space of solenoid functions for Markov maps on train tracks', *Ergodic Theory Dynam. Systems* 15 (1995) 685–696.
18. A. A. PINTO and D. A. RAND, 'Classifying C^{1+} structures on hyperbolic fractals. 2: Embedded trees', *Ergodic Theory Dynam. Systems* 15 (1995) 969–992.
19. A. A. PINTO and D. A. RAND, 'Existence, uniqueness and ratio decomposition for Gibbs states via duality', *Ergodic Theory Dynam. Systems* 21 (2001) 533–544.
20. A. A. PINTO and D. A. RAND, 'Teichmüller spaces and HR structures for hyperbolic surface dynamics', *Ergodic Theory Dynam. Systems* 22 (2002) 1905–1931.
21. A. A. PINTO and D. A. RAND, 'Smoothness of holonomies for codimension 1 hyperbolic dynamics', *Bull. London Math. Soc.* 34 (2002) 341–352.
22. A. A. PINTO and D. SULLIVAN, 'Dynamical systems applied to asymptotic geometry', manuscript.
23. M. SHUB, *Global stability of dynamical systems* (Springer, 1987).
24. J. C. YOCCOZ, 'Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne', *Ann. Sci. École Norm. Sup.* (4) 17 (1984) 333–359.

A. A. Pinto
 Faculdade de Ciências
 Universidade do Porto
 4000 Porto
 Portugal

aapinto@fc.up.pt

D. A. Rand
 Mathematics Institute
 University of Warwick
 Coventry CV4 7AL
 United Kingdom

dar@maths.warwick.ac.uk