

SYMBOLIC DYNAMICS FOR SURFACE DIFFEOMORPHISMS  
WITH POSITIVE ENTROPY

OMRI M. SARIG

CONTENTS

<b>Part 0. Introduction and statement of results</b>	2
1.1. Results	2
1.2. Symbolic dynamics	3
1.3. Markov partitions	4
1.4. Comparison to other results in the literature	5
1.5. Overview of the construction of a Markov partition	6
1.6. Notational conventions and standing assumptions	7
<b>Part 1. Chains as pseudo-orbits</b>	8
2. Pesin charts	8
2.1. Non-uniform hyperbolicity	8
2.2. Lyapunov change of coordinates	8
2.3. Pesin Charts	10
2.4. Distortion compensating bounds	11
2.5. $\text{NUH}_\chi^\#(f)$	11
3. Overlapping charts	11
3.1. The overlap condition	12
3.2. The form of $f$ in overlapping charts	14
3.3. Coarse graining	16
4. $\varepsilon$ -chains and an infinite-to-one Markov extension of $f$	18
4.1. Double charts and $\varepsilon$ -chains	18
4.2. Admissible manifolds and the graph transform	20
4.3. A Markov extension	23
4.4. The relevant part of the extension	27
<b>Part 2. Regular chains which shadow the same orbit are close</b>	28
5. The inverse problem for regular chains	28
6. Axes parameters	29
6.1. Staying in windows	29
6.2. Comparison of $\alpha(x_i)$ to $\alpha(y_i)$	30

---

Received by the editors January 17, 2011, Revised: September 2, 2012.

2010 *Mathematics Subject Classification*. Primary 37D25; Secondary 37D35.

*Key words and phrases*. Markov partitions, symbolic dynamics, periodic points.

This work was supported by the NSF (DMS-0400687) and the ESF (ERC-2009-StG 239885).

6.3. Comparison of $R_{x_i}$ to $R_{y_i}$	31
7. Scaling parameters	33
7.1. The $s_\chi$ and $u_\chi$ parameters of admissible manifolds	33
7.2. Comparison of $s_\chi(x_i), u_\chi(x_i)$ to $s_\chi(y_i), u_\chi(y_i)$ .	36
8. Window parameters	38
8.1. $\varepsilon$ -maximality	38
8.2. Comparison of $p_i^{u/s}$ to $q_i^{u/s}$	39
9. Proof of Theorem 5.2	40
<b>Part 3. Markov partitions and symbolic dynamics</b>	42
10. A locally finite countable Markov cover	42
10.1. The cover	42
10.2. Product structure	43
10.3. The symbolic Markov property	44
11. A countable Markov partition	45
11.1. The Bowen–Sinai refinement	45
11.2. Product structure and hyperbolicity	47
11.3. The Markov property	48
12. Symbolic dynamics	50
12.1. A directed graph	50
12.2. The Markov extension	51
12.3. The extension is finite-to-one	54
13. Invariant measures	57
<b>Part 4. Appendix: Proofs of standard results in Pesin Theory</b>	59
References	82

## Part 0. Introduction and statement of results

**1.1. Results.** Let  $M$  be a compact  $C^\infty$  Riemannian manifold of dimension two, and let  $f : M \rightarrow M$  be a  $C^{1+\beta}$  diffeomorphism ( $0 < \beta < 1$ ) with positive topological entropy  $h_{top}(f)$ . Set  $P_n(f) := |\{x \in M : f^n(x) = x\}|$ .

Anatole Katok showed in [K1],[K2] that  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log P_n(f) \geq h_{top}(f)$ , and conjectured that if  $f$  is  $C^\infty$  then  $\limsup_{n \rightarrow \infty} e^{-nh_{top}(f)} P_n(f) > 0$  (see [K3]). We show:

**Theorem 1.1.** *Suppose  $f$  is a  $C^{1+\beta}$  diffeomorphism of a compact smooth surface, and assume  $h_{top}(f) > 0$ . If  $f$  has a measure of maximal entropy, then  $\exists p \in \mathbb{N}$  s.t.  $\liminf_{n \rightarrow \infty, p|n} e^{-nh_{top}(f)} P_n(f) > 0$ .*

This proves Katok’s conjecture, because  $C^\infty$  diffeomorphisms on compact manifolds have measures of maximal entropy (Newhouse [N]).

Jérôme Buzzi has conjectured in [Bu4] that  $f$  admits at most countably many different ergodic measures of maximal entropy. We prove this to be correct:

**Theorem 1.2.** *Suppose  $f$  is a  $C^{1+\beta}$  diffeomorphism of a compact smooth surface. If  $h_{top}(f) > 0$  then  $f$  possesses at most countably many ergodic invariant probability measures with maximal entropy.*

Buzzi also conjectured that if  $f$  is  $C^\infty$ , then the number of different ergodic invariant measures of maximal entropy is finite. This conjecture remains open.

Katok’s conjecture and Buzzi’s conjectures were previously known to hold in the following cases: Hyperbolic automorphisms of the torus [AW], Anosov diffeomorphisms [Si1, Si2], [M], Axiom A diffeomorphisms [B4], [PP], continuous piecewise affine homeomorphisms of affine surfaces [Bu4]. There are also results on non-invertible maps, see [Hof1, Hof2] and [Bu1, Bu5]. A wealth of diffeomorphisms such that  $\limsup_{n \rightarrow \infty} e^{-nh_{top}(f)} P_n(f) = \infty$  can be found in [Kal].

**1.2. Symbolic dynamics.** The proof of Theorems 1.1 and 1.2 is based on a change of coordinates which simplifies the iteration of  $f$ . The idea, which goes back to the work of Hadamard, Birkhoff and Artin on geodesic flows, is to semi-conjugate  $f$  on a large set to the left shift on a *topological Markov shift*. We recall the definition.

Let  $\mathcal{G}$  be a directed graph with a countable collection of vertices  $\mathcal{V}$  s.t. every vertex has at least one edge coming in, and at least one edge coming out. The *topological Markov shift* associated to  $\mathcal{G}$  is the set

$$\Sigma = \Sigma(\mathcal{G}) := \{(v_i)_{i \in \mathbb{Z}} \in \mathcal{V}^{\mathbb{Z}} : v_i \rightarrow v_{i+1} \text{ for all } i\}.$$

We equip  $\Sigma$  with the *natural metric*:  $d(\underline{u}, \underline{v}) := \exp[-\min\{|i| : u_i \neq v_i\}]$ , thus turning it into a complete separable metric space.  $\Sigma$  is compact iff  $\mathcal{G}$  is finite.  $\Sigma$  is locally compact iff every vertex of  $\mathcal{G}$  has finite degree.

The *left shift map*  $\sigma : \Sigma \rightarrow \Sigma$  is defined by  $\sigma[(v_i)_{i \in \mathbb{Z}}] = (v_{i+1})_{i \in \mathbb{Z}}$ .

Let  $\Sigma^\# := \{(v_i)_{i \in \mathbb{Z}} \in \Sigma : \exists u, v \in \mathcal{V} \exists n_k, m_k \uparrow \infty \text{ s.t. } v_{-m_k} = u, v_{n_k} = v\}$ .  $\Sigma^\#$  contains all the periodic points of  $\sigma$ , and by the Poincaré Recurrence Theorem, every  $\sigma$ -invariant probability measure gives  $\Sigma^\#$  full measure.

We say that a set  $\Omega \subset M$  is  $\chi$ -large, if  $\mu(\Omega) = 1$  for every ergodic invariant probability measure  $\mu$  whose entropy is greater than  $\chi$ . We prove:

**Theorem 1.3.** *For every  $0 < \chi < h_{top}(f)$  there exists a locally compact topological Markov shift  $\Sigma_\chi$  and a Hölder continuous map  $\pi_\chi : \Sigma_\chi \rightarrow M$  s.t.  $\pi_\chi \circ \sigma = f \circ \pi_\chi$ ;  $\pi_\chi[\Sigma_\chi^\#]$  is  $\chi$ -large; and s.t. every point in  $\pi_\chi[\Sigma_\chi^\#]$  has finitely many pre-images.*

**Theorem 1.4.** *Denote the set of states of  $\Sigma_\chi$  by  $\mathcal{V}_\chi$ . There exists a function  $\varphi_\chi : \mathcal{V}_\chi \times \mathcal{V}_\chi \rightarrow \mathbb{N}$  s.t. if  $x = \pi_\chi[(v_i)_{i \in \mathbb{Z}}]$  and  $v_i = u$  for infinitely many negative  $i$ , and  $v_i = v$  for infinitely many positive  $i$ , then  $|\pi_\chi^{-1}(x)| \leq \varphi_\chi(u, v)$ .*

**Theorem 1.5.** *Every ergodic  $f$ -invariant probability measure  $\mu$  on  $M$  such that  $h_\mu(f) > \chi$  equals  $\hat{\mu} \circ \pi_\chi^{-1}$  for some ergodic  $\sigma$ -invariant probability measure  $\hat{\mu}$  on  $\Sigma_\chi$  with the same entropy.*

The other direction is trivial: If  $\hat{\mu}$  is an ergodic  $\sigma$ -invariant probability measure on  $\Sigma_\chi$ , then  $\mu := \hat{\mu} \circ \pi_\chi^{-1}$  is an ergodic  $f$ -invariant probability measure on  $M$ , and  $\mu$  has the same entropy as  $\hat{\mu}$  because  $\pi_\chi$  is finite-to-one.

*A remark on the regularity of  $\pi_\chi$ .* Our bound for the Hölder exponent of  $\pi_\chi$  decays to zero as  $\chi \rightarrow 0$ , see the proofs of Proposition 4.15 and Theorem 4.16.

*A remark on  $\chi$ -largeness.* Call a set  $\Omega \subset M$   $\chi$ -very-large, if  $\Omega$  has full measure with respect to every ergodic invariant probability measure with at least one Lyapunov exponent larger than  $\chi$ .

In dimension two, the positive Lyapunov exponent of an ergodic invariant probability measure is bigger than or equal to its metric entropy (“Ruelle’s entropy inequality” [Ru]). Therefore, every  $\chi$ -very-large set is  $\chi$ -large.

As pointed out to the author by Professor L.-S. Young, Professor S. Newhouse, and the referee, the proofs given in this paper actually show that  $\pi_\chi[\Sigma_\chi^\#]$  in Theorem 1.3 is  $\chi$ -very-large, not just  $\chi$ -large. Similarly in Theorem 1.5 one can replace the condition  $h_\mu(f) > \chi$  by the assumption that  $\mu$  has a Lyapunov exponent  $\chi(\mu) > \chi$ .

We explain how to use these results to prove Theorems 1.1 and 1.2. This reduction was already known to Katok and Buzzi [K3],[Bu4]. Write  $\Sigma_\chi = \Sigma(\mathcal{G})$ . By Theorem 1.5, every ergodic measure of maximal entropy  $\mu$  for  $f$  lifts to an ergodic measure of maximal entropy  $\hat{\mu}$  for  $\sigma$ . By ergodicity,  $\hat{\mu}$  is carried by a set  $\Sigma(\mathcal{G}')$  where (1)  $\mathcal{G}'$  is a subgraph of  $\mathcal{G}$ , and (2)  $\mathcal{G}'$  is *irreducible*: for any two vertices  $v_0, v_1$  there exists a path in  $\mathcal{G}'$  from  $v_0$  to  $v_1$ . Since  $\hat{\mu}$  is a measure of maximal entropy for  $\sigma : \Sigma(\mathcal{G}) \rightarrow \Sigma(\mathcal{G})$ , it is also a measure of maximal entropy for  $\sigma : \Sigma(\mathcal{G}') \rightarrow \Sigma(\mathcal{G}')$ .

The irreducibility of  $\mathcal{G}'$  means that  $\sigma : \Sigma(\mathcal{G}') \rightarrow \Sigma(\mathcal{G}')$  is topologically transitive. Gurevich proved in [Gu1, Gu2] that a topologically transitive topological Markov shift  $\Sigma(\mathcal{G}')$  admits at most one measure of maximal entropy, and that such a measure exists iff  $\exists p \in \mathbb{N}$  s.t. for every vertex  $v_0$  in  $\mathcal{G}'$ ,

$$|\{\underline{v} \in \Sigma(\mathcal{G}') : \sigma^n(\underline{v}) = \underline{v}, v_0 = v\}| \asymp \exp[nh_{\max}(\Sigma(\mathcal{G}'))] \text{ as } n \rightarrow \infty \text{ in } p\mathbb{N},$$

where  $h_{\max}(\Sigma(\mathcal{G}')) = \sup\{h_\mu(\sigma) : \mu \text{ a } \sigma\text{-invariant Borel prob. measure on } \Sigma(\mathcal{G}')\}$ , and  $h_\mu(\sigma)$  denotes the metric entropy of  $\mu$  w.r.t.  $\sigma$ . By “ $a_n \asymp b_n$  as  $n \rightarrow \infty$  in  $p\mathbb{N}$ ” we mean that for some  $C > 1$ ,  $C^{-1} \leq a_n/b_n \leq C$  for all  $n \in p\mathbb{N}$  large enough.

Since  $\pi_\chi \circ \sigma = f \circ \pi_\chi$ , the collection  $\{\underline{v} \in \Sigma(\mathcal{G}') : \sigma^n(\underline{v}) = \underline{v}, v_0 = v\}$  is mapped by  $\pi_\chi$  to a collection of points  $x \in M$  s.t.  $f^n(x) = x$ . By Theorem 1.4, the mapping is bounded-to-one, with the number of pre-images bounded by  $\varphi_\chi(v_0, v_0)$ . Thus  $\liminf_{n \rightarrow \infty, p|n} e^{-nh_{\max}(\Sigma(\mathcal{G}'))} P_n(f) > 0$ . By construction,  $h_{\max}(\Sigma(\mathcal{G}')) = h_{\hat{\mu}}(\sigma) = h_\mu(f) = \max\{h_\nu(f) : \nu \text{ } f\text{-inv.}\}$ . The last quantity is equal to  $h_{\text{top}}(f)$  by the variational principle [G]. Theorem 1.1 follows.

This argument also shows that the cardinality of the collection of measures of maximal entropy for  $f$  is bounded by the cardinality of the collection of subgraphs  $\mathcal{G}' \subset \mathcal{G}$  s.t. (1)  $\mathcal{G}'$  is irreducible, (2)  $\Sigma(\mathcal{G}')$  carries a unique measure of maximal entropy, and (3)  $h_{\max}(\Sigma(\mathcal{G}')) = h_{\max}(\Sigma(\mathcal{G}))$ .

Any two such subgraphs are equal, or their sets of vertices are disjoint: Otherwise the shift defined by their union carries at least two measures of maximal entropy, and this contradicts Gurevich’s theorem. It follows that the collection of subgraphs satisfying (1),(2), and (3) is finite or countable. Theorem 1.2 follows.

**1.3. Markov partitions.** As in [AW, Si1, B1], the symbolic description of  $f$  relies on the existence of a *countable Markov partition*. This is a pairwise disjoint collection  $\mathcal{R}$  of Borel sets with the following properties:

- (1) **Covering property:** The union of  $\mathcal{R}$  is  $\chi$ -large.
- (2) **Product structure:** There are  $W^s(x, R), W^u(x, R) \subset R$  ( $x \in R \in \mathcal{R}$ ) s.t.
  - (a)  $W^u(x, R) \cap W^s(x, R) = \{x\}$ .
  - (b)  $\forall x, y \in R, \exists z \in R$  s.t.  $W^u(x, R) \cap W^s(y, R) = \{z\}$ .
  - (c)  $\forall x, y \in R, W^s(x, R)$  and  $W^s(y, R)$  are equal, or they are disjoint. Similarly for  $W^u(x, R), W^u(y, R)$ .
- (3) **Hyperbolicity:** If  $y, z \in W^s(x, R)$ , then  $d(f^n(y), f^n(z)) \xrightarrow[n \rightarrow \infty]{} 0$ . If  $y, z \in W^u(x, R)$ , then  $d(f^{-n}(y), f^{-n}(z)) \xrightarrow[n \rightarrow \infty]{} 0$ .
- (4) **Markov property:** Suppose  $R_1, R_2 \in \mathcal{R}$  and  $x \in R_1, f(x) \in R_2$ , then  $f[W^s(x, R_1)] \subseteq W^s(f(x), R_2)$  and  $f^{-1}[W^u(f(x), R_2)] \subseteq W^u(x, R_1)$ .

We do not ask for the sets  $R$  to be the closure of their interiors.

#### 1.4. Comparison to other results in the literature.

*Markov partitions for diffeomorphisms.* These were previously constructed in the following cases: Hyperbolic toral automorphisms [Be],[AW], Anosov diffeomorphisms [Si1], pseudo-Anosov diffeomorphisms [FS], and Axiom A diffeomorphisms [B1, B2]. This paper treats the general case, in dimension two.

*Katok horseshoes* [K1, K2],[KM]. Katok showed that if a  $C^{1+\beta}$  surface diffeomorphism  $f$  has positive entropy, then for every  $\varepsilon > 0$  there is a compact invariant subset  $\Lambda_\varepsilon$  s.t.  $f : \Lambda_\varepsilon \rightarrow \Lambda_\varepsilon$  has a finite Markov partition, and  $h_{top}(f|_{\Lambda_\varepsilon}) > h_{top}(f) - \varepsilon$ .

Typically,  $\Lambda_\varepsilon$  will have zero measure w.r.t. any ergodic invariant measure with large entropy. This paper constructs a “horseshoe”  $\pi_\chi(\Sigma_\chi)$  with full measure for all ergodic invariant measures with large entropy.

Some differences should be noted: (a) our horseshoe is not compact, (b) its Markov partition is infinite, and (c) the semi-conjugacy  $\pi_\chi$  is not one-to-one as in [KM]. (a) and (b) are unavoidable. I do not know if it is possible to get a semi-conjugacy which is one-to-one on a set of full measure for “nice” measures: the boundaries of the partition elements constructed here could be very large.

Katok’s work also includes the higher dimensional case, with the condition of positive topological entropy replaced by the stronger assumption that there exist ergodic measures without zero Lyapunov exponents with metric entropy arbitrarily close to the topological entropy. We expect a similar generalization of our results.

*Tower extensions* [Ta],[Hof1],[Y]: These are representations of certain maps as infinite-to-one factors of other maps (“towers”) which possess obvious infinite Markov partitions. Such extensions have been used in the study of one-dimensional systems with great success, see e.g. [Hof2],[Bu1], [Bru],[Ke2], [PSZ],[IT],[Z]. For higher dimension, see [Bu4, Bu2, Bu5], [BT], [BY], [Y].

Unlike tower extensions, our coding is finite-to-one. This ensures that any ergodic invariant measure with high entropy can be lifted to the symbolic space (Theorem 1.5, see also (13.1)). For tower extensions proving the existence of a lift is highly non-trivial, and there are very few results in dimension higher than one, see [Ke1], [Bu4], [BT], [PSZ] and references therein.

*Symbolic extensions* [BD],[DN],[BFF]. These are representations of a diffeomorphism as a topological factor of  $\sigma : \Lambda \rightarrow \Lambda$  where  $\Lambda \subset \{1, \dots, N\}^{\mathbb{Z}}$  is closed and shift invariant and  $\sigma$  is the left shift (“subshift”). Burguet has shown that every  $C^2$  surface diffeomorphism has a symbolic extension [Bur]. In the  $C^\infty$  case there are symbolic extensions whose factor maps preserve entropy [Bu1],[BFF]. In lower regularity it is not even always true that  $h_{top}(\sigma) = h_{top}(f)$ .

Unlike symbolic extensions, our symbolic space is not compact. But it is Markovian, and this gives us access to many results which are not true for general subshifts, such as Gurevich’s theory which was essential for Theorems 1.1 and 1.2. Another advantage of our extension is the lifting theorem (Theorem 1.5), which does not seem to be available for general symbolic extensions.

*Markov partitions for billiards.* [BS] and [BSC] construct countable Markov partitions for certain dispersing billiard systems. Their partitions capture sets of full Liouville measure. If our methods could be adapted to handle maps with singularities such as billiards, then one could hope to construct Markov partitions which capture the measure of maximal entropy.

**1.5. Overview of the construction of a Markov partition.** It is useful first to recall Bowen’s construction in the case of an Anosov diffeomorphisms [B4].

Bowen’s idea was to use  $\varepsilon$ -pseudo-orbits. These are sequences of points  $\underline{x} = \{x_i\}_{i \in \mathbb{Z}}$  such that  $d(x_{i+1}, f(x_i)) < \varepsilon$  for all  $i$ . A pseudo-orbit  $\underline{x}$  is said to  $\delta$ -shadow a real orbit  $\{f^i(x)\}_{i \in \mathbb{Z}}$  if  $d(x_i, f^i(x)) < \delta$  for all  $i \in \mathbb{Z}$ . Anosov showed that for every  $\delta$  small enough, there exists an  $\varepsilon > 0$  s.t.

- (A1) Every  $\varepsilon$ -pseudo-orbit  $\underline{x}$   $\delta$ -shadows the real orbit of some unique point  $\pi(\underline{x})$ .
- (A2) “Finite alphabet suffices”: There exists a finite set of points  $A$  such that  $\{\pi(\underline{x}) : \underline{x} \in A^{\mathbb{Z}} \text{ is an } \varepsilon\text{-pseudo-orbit}\}$  is the entire manifold.
- (A3) “Inverse problem”: If two pseudo-orbits  $\underline{x}, \underline{y}$   $\delta$ -shadow the same orbit, then their corresponding coordinates are close,  $d(x_i, y_i) < 2\delta$  for all  $i \in \mathbb{Z}$ .

Since pseudo-orbits are defined in terms of nearest neighbor constraints, one can view the collection of pseudo-orbits in  $A^{\mathbb{Z}}$  as the collection of infinite paths on the graph with set of vertices  $A$ , and edges  $a \rightarrow b$  when  $d(f(a), b) < \varepsilon$ . (A1) and (A2) say that  $f$  is a factor of the topological Markov shift

$$\Sigma := \{\underline{x} \in A^{\mathbb{Z}} : d(x_{i+1}, f(x_i)) < \varepsilon \text{ for all } i \in \mathbb{Z}\}.$$

The factor map is  $\pi$ . It is an infinite-to-one map.

The sets  ${}_0[a] := \{\underline{x} \in \Sigma : x_0 = a\}$  form a natural Markov partition for the left shift on  $\Sigma$ .<sup>1</sup> Their projections  $Z(a) = \{\pi(\underline{x}) : \underline{x} \in \Sigma, x_0 = a\}$  ( $a \in A$ ) would have been natural candidates for a Markov partition, had they not overlapped. Sinai came up with a set-theoretic procedure for refining

$$\mathcal{Z} := \{Z(a) : a \in A\}$$

into a partition without destroying the product structure. This partition is a Markov partition [B4].

Our proof follows a similar strategy. But since Anosov’s theory of pseudo-orbits relies on uniform hyperbolicity and our setting is only non-uniformly hyperbolic, we have to find a substitute for Anosov’s shadowing theory. This problem was previously considered by Krüger & Troubetzkoy [KT], but their construction does not work in our setting.

In part 1, we introduce  $\varepsilon$ -chains as a replacement to  $\varepsilon$ -pseudo-orbits in the non-uniformly hyperbolic setup. Much like a pseudo-orbit, a chain is a sequence of symbols which satisfies nearest neighbor conditions. Each symbol contains partial information on the location of the point and the position and size of its local stable and unstable manifolds. The nearest neighbor conditions are tailored in such a way that the following analogues of parts (A1) and (A2) of Anosov’s theorem hold for a suitable choice of  $\varepsilon$ :

- (A1’) Every  $\varepsilon$ -chain  $\underline{v}$  corresponds to a unique real orbit  $\pi(\underline{v})$ ;
- (A2’) There is a countable set  $A$  of symbols s.t.  $\{\pi(\underline{u}) : \underline{u} \in A^{\mathbb{Z}} \text{ is an } \varepsilon\text{-chain}\}$  is  $\chi$ -large.  $A$  and  $\varepsilon$  depend on  $\chi$ .

As a result, we obtain a representation of  $f$  (restricted to a large invariant set) as a factor of a topological Markov shift.

The next step is to construct  $\mathcal{Z}$  as before and try to apply Sinai’s method to obtain a countable refining partition. Here we run into a serious problem: whereas

<sup>1</sup>The product structure is given by  $W^u(\underline{x}, {}_0[a]) := \{\underline{y} \in \Sigma : y_i = x_i \ (i \leq 0)\}$ ,  $W^s(\underline{x}, {}_0[a]) := \{\underline{y} \in \Sigma : y_i = x_i \ (i \geq 0)\}$ .

Sinai dealt with a finite cover, our cover is infinite, and a general countable cover need not have a countable refining partition. To avoid such pathologies one needs to ensure that  $\mathcal{Z}$  is *locally finite*: Every  $Z \in \mathcal{Z}$  intersects at most finitely many other  $Z' \in \mathcal{Z}$ . This difficulty turns out to be the heart of the matter.

We deal with this issue in part 2. Here we obtain the following analogue of part (A3) of Anosov's theorem:

(A3') If two  $\varepsilon$ -chains  $\underline{v}, \underline{u}$  are "regular" and  $\pi(\underline{u}) = \pi(\underline{v})$ , then  $u_i$  and  $v_i$  are "close" for every  $i \in \mathbb{Z}$  (see §5 for the precise statement).

Unlike (A3), this is not a trivial statement, because the symbols  $u_i, v_i$  contain much more information than mere location. The fact that  $\varepsilon$ -chains satisfy (A3') is the main point of this work.

The alphabet  $A$  from part 1 can be chosen s.t. (a) for every  $u \in A$ , the number of  $v \in A$  "close" to  $u$  is finite, and (b)  $\{\pi(\underline{u}) : \underline{u} \in A^{\mathbb{Z}}, \underline{u} \text{ is a regular } \varepsilon\text{-chain}\}$  has full measure w.r.t. any ergodic invariant probability measure with entropy more than  $\chi$ . As a result, the sets  $Z(v) := \{\pi(\underline{v}) : \underline{v} \in A^{\mathbb{Z}} \text{ is a regular } \varepsilon\text{-chain}\}$  form a locally finite cover  $\mathcal{Z}$  of a large set.

Sinai's refinement procedure can now be safely applied to  $\mathcal{Z}$ . In part 3, we check that the elements of  $\mathcal{Z}$  have the "product structure" and "symbolic Markov properties" needed to push through Bowen's proof that Sinai's refinement is a Markov partition. We also explain how to deduce Theorems 1.3, 1.4, and 1.5. The proofs are modeled on [B4, B3].

Some of the lemmas we need to develop the theory of  $\varepsilon$ -chains are routine modification of well-known results in Pesin Theory. Part 4 collects their proofs.

**1.6. Notational conventions and standing assumptions.** In what follows,  $M$  is a compact  $C^\infty$  Riemannian manifold of dimension two. We assume without loss of generality that  $M$  is orientable (otherwise pass to a finite orientable extension).

Let  $f : M \rightarrow M$  be a  $C^{1+\beta}$  diffeomorphism where  $0 < \beta < 1$ . We assume that the topological entropy of  $f$  is positive, and fix a constant  $0 < \chi < h_{top}(f)$ .

Suppose  $P$  is a property. The statement "for all  $\varepsilon$  small enough  $P$  holds" means " $\exists \varepsilon_0 > 0$  which only depends on  $f, M, \beta$  and  $\chi$  s.t. for all  $0 < \varepsilon < \varepsilon_0$   $P$  holds".

The metric entropy of an  $f$ -invariant measure  $\mu$  is denoted by  $h_\mu(f)$ . The topological entropy of  $f$  is denoted by  $h_{top}(f)$ .

$T_x M$  is the tangent space to  $M$  at  $x$ . The exponential map is denoted by  $\exp_x : T_x M \rightarrow M$ . The Riemannian norm and inner product on  $T_x M$  are denoted by  $\|\cdot\|_x$  and  $\langle \cdot, \cdot \rangle_x$ . Sometimes, we drop the subscript  $x$ . Given two non-zero vectors  $\underline{u}, \underline{v} \in T_x M$ , the angle from  $\underline{u}$  to  $\underline{v}$  is denoted by  $\angle(\underline{u}, \underline{v})$ . This is a signed quantity.

Let  $V$  be a vector space. The zero element in  $V$  is denoted by  $\underline{0}$ . We identify the tangent space to  $V$  at  $\underline{v} \in V$  with  $V$ . Let  $A : V \rightarrow W$  be a linear map between two linear vector spaces  $V, W$ . We identify  $(dA)_{\underline{v}} : T_{\underline{v}} V \rightarrow T_{A\underline{v}} W$  with  $A : V \rightarrow W$ .

Suppose  $a, b, c \in \mathbb{R}$ . We write  $a = b \pm c$  if  $b - c \leq a \leq b + c$ , and  $a = e^{\pm c} b$  if  $e^{-c} b \leq a \leq e^c b$ . Let  $a_n, b_n > 0$ , then  $a_n \sim b_n$  means that  $\frac{a_n}{b_n} \xrightarrow{n \rightarrow \infty} 1$ , and  $a_n \asymp b_n$  means that  $\exists N, c$  s.t.  $\forall n > N$  ( $e^{-c} b_n \leq a_n \leq e^c b_n$ ). Finally,  $a \wedge b := \min\{a, b\}$ .

Some abbreviations: *s.t.* is "such that", *w.r.t* is "with respect to", *i.o.* is "infinitely often", *resp.* is "respectively", and *w.l.o.g* is "without loss of generality".

## Part 1. Chains as pseudo-orbits

### 2. PESIN CHARTS

**2.1. Non-uniform hyperbolicity.** By the variational principle,  $f$  admits ergodic invariant probability measures of entropy larger than  $\chi$  (see [G]). Quite a lot is known about the properties of these measures. We will use the following fact, which follows from Ruelle’s Entropy Inequality [Ru] and the Oseledets Multiplicative Ergodic Theorem [Os] (see [BP]):

**Theorem 2.1** (Oseledets, Ruelle). *Any ergodic invariant probability measure  $\mu$  for  $f$  s.t.  $h_\mu(f) > \chi$  gives full probability to the set  $\text{NUH}_\chi(f)$  of points  $x \in M$  for which for every  $y \in \{f^k(x) : k \in \mathbb{Z}\}$ ,  $T_y M = E^s(y) \oplus E^u(y)$  where*

- (1)  $E^s(y) = \text{span}\{\underline{e}^s(y)\}$ ,  $\|\underline{e}^s(y)\|_y = 1$ ,  $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(df^n)_y \underline{e}^s(y)\|_{f^n(y)} < -\chi$ ;
- (2)  $E^u(y) = \text{span}\{\underline{e}^u(y)\}$ ,  $\|\underline{e}^u(y)\|_y = 1$ ,  $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(df^n)_y \underline{e}^u(y)\|_{f^n(y)} > \chi$ ;
- (3)  $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log |\sin \alpha(f^n(y))| = 0$ , where  $\alpha(y) := \angle(\underline{e}^s(y), \underline{e}^u(y))$ ;
- (4)  $df_y[E^s(y)] = E^s(f(y))$  and  $df_y[E^u(y)] = E^u(f(y))$ .

$\text{NUH}_\chi(f)$  is invariant. Properties (1) and (2) determine the splitting  $E^s \oplus E^u$  uniquely, but the vectors  $\underline{e}^s, \underline{e}^u$  are only determined up to a sign. To fix the sign we use the assumption that  $M$  is orientable to choose a measurable family of positively oriented bases  $(\underline{e}_y^1, \underline{e}_y^2)$  of  $T_y M$  ( $y \in M$ ); then we choose the signs of  $\underline{e}^{s/u}(y)$  so that  $\angle(\underline{e}_y^1, \underline{e}^s(y)) \in [0, \pi)$  and  $(\underline{e}^s(y), \underline{e}^u(y))$  have positive orientation.

$\text{NUH}(f) := \bigcup_{\chi > 0} \text{NUH}_\chi(f)$  is called the *non-uniformly hyperbolic set* of  $f$ , and is  $f$ -invariant. This set has full probability w.r.t. any ergodic invariant probability measure with positive entropy.

The linear spaces  $E^s(x), E^u(x)$  are called, respectively, the *stable* and *unstable* spaces of  $df$ . The numbers

$$\begin{aligned} \log \lambda(x) &:= \lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(df^n)_x \underline{e}^s(x)\|_{f^n(x)} \\ \log \mu(x) &:= \lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(df^n)_x \underline{e}^u(x)\|_{f^n(x)} \end{aligned} \quad (x \in \text{NUH}(f))$$

are called the *Lyapunov exponents* of  $x$ . They are  $f$ -invariant, whence constant a.e. w.r.t. any ergodic invariant measure. The value depends on the measure. On  $\text{NUH}_\chi(f)$ ,  $\log \lambda(x) < -\chi$  and  $\log \mu(x) > \chi$ .

**2.2. Lyapunov change of coordinates.** The splitting  $T_x M = E^s(x) \oplus E^u(x)$  can be used to diagonalize the action of  $df$  on  $\{T_x M : x \in \text{NUH}(f)\}$  (“Oseledets–Pesin Reduction”).

We describe a change of coordinates which achieves this. The construction depends on  $\chi$ . Given  $x \in \text{NUH}_\chi(f)$ , let

$$\begin{aligned} s_\chi(x) &:= \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \|(df^k)_x \underline{e}^s(x)\|_{f^k(x)}^2 \right)^{1/2}; \\ u_\chi(x) &:= \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \|(df^{-k})_x \underline{e}^u(x)\|_{f^{-k}(x)}^2 \right)^{1/2}. \end{aligned}$$

The  $\sqrt{2}$  is needed to make the change of coordinates a contraction, see Lemma 2.5.



**Definition 2.2.** *The Lyapunov change of coordinates (with parameter  $\chi$ ) is the linear map  $C_\chi(x) : \mathbb{R}^2 \rightarrow T_x M$  ( $x \in \text{NUH}_\chi(f)$ ) s.t.  $C_\chi(x)\underline{e}_1 = s_\chi(x)^{-1}\underline{e}^s(x)$ , and  $C_\chi(x)\underline{e}_2 = u_\chi(x)^{-1}\underline{e}^u(x)$ , where  $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .*

Notice that  $C_\chi(x)$  preserves orientation.

**Theorem 2.3** (Oseledets–Pesin Reduction Theorem). *There exists a constant  $C_f$  which only depends on  $f$  s.t. for every  $x \in \text{NUH}_\chi(f)$ ,*

$$C_\chi(f(x))^{-1} \circ df_x \circ C_\chi(x) = \begin{pmatrix} \lambda_\chi(x) & 0 \\ 0 & \mu_\chi(x) \end{pmatrix}$$

where  $C_f^{-1} < |\lambda_\chi(x)| < e^{-\chi}$  and  $e^\chi < |\mu_\chi(x)| < C_f$ .

Pesin’s original construction in [P] is slightly different. He defined  $s_\chi(x)$  and  $u_\chi(x)$  with  $e^{-2k\varepsilon}\lambda(x)^{-2k}$  or  $e^{-2k\varepsilon}\mu(x)^{2k}$  replacing  $e^{2k\chi}$ . His method gives better bounds on  $\lambda_\chi(x)$  and  $\mu_\chi(x)$ , and makes sense on all of  $\text{NUH}(f)$ . Our method can only be guaranteed to work on  $\text{NUH}_\chi(f)$ , but it has the advantage that  $C_\chi(x)$  is not sensitive to the values of  $\lambda(x), \mu(x)$ . This is important, because we want to capture the dynamics of all orbits with exponents bounded away from  $\chi$ , therefore we have to work with points with different Lyapunov exponents.

We need the following definition from linear algebra: suppose  $L : V \rightarrow W$  is an invertible linear map between two finite dimensional vector spaces equipped with inner products, then the *operator norm* of  $L$  is  $\|L\| := \max\{\|Lv\|_W : \|v\|_V = 1\}$ , and the *Frobenius norm* of  $L$  is  $\|L\|_{Fr} := \sqrt{\text{tr}(\Theta^t L^t L \Theta)}$ , where  $\Theta$  is some (any) isometry  $\Theta : W \rightarrow V$ .  $\|L\|_{Fr}$  is well defined,<sup>2</sup> and  $\|L\| \leq \|L\|_{Fr} \leq \sqrt{2}\|L\|$ .<sup>3</sup> One of the advantages of the Frobenius norm is that it has an explicit formula: If  $L$  is represented by the matrix  $(a_{ij})$  w.r.t. to some (any) orthonormal bases for  $V, W$ , then  $\|L\|_{Fr} = \left(\sum_{ij} a_{ij}^2\right)^{1/2}$ .<sup>4</sup>

Some more information on  $C_\chi(x)$  (see the appendix for proofs):

**Lemma 2.4.**  $\|C_\chi(x)^{-1}\|_{Fr} = \sqrt{s_\chi(x)^2 + u_\chi(x)^2} / |\sin \alpha(x)|$ .

**Lemma 2.5.**  $C_\chi(x)$  is a contraction:  $\|C_\chi(x)\begin{pmatrix} \xi \\ \eta \end{pmatrix}\|_x \leq \|\begin{pmatrix} \xi \\ \eta \end{pmatrix}\|$  for all  $\xi, \eta \in \mathbb{R}$ .

**Lemma 2.6.** *There is a  $\chi$ -large invariant set  $\text{NUH}_\chi^*(f) \subset \text{NUH}_\chi(f)$  s.t. for every  $x \in \text{NUH}_\chi^*(f)$ ,*

- (1)  $\lim_{k \rightarrow \pm\infty} \frac{1}{k} \log \|C_\chi(f^k(x))^{-1}\| = 0$ ;
- (2)  $\lim_{k \rightarrow \pm\infty} \frac{1}{k} \log \|C_\chi(f^k(x))\underline{e}_i\|_{f^k(x)} = 0$ , where  $\underline{e}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $\underline{e}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ;
- (3)  $\lim_{k \rightarrow \pm\infty} \frac{1}{k} \log |\det C_\chi(f^k(x))| = 0$ .

<sup>2</sup>Proof:  $\text{tr}(\Theta_2^t L^t L \Theta_2) = \text{tr}[\Theta_2^t \Theta_1 (\Theta_1^t L^t L \Theta_1) (\Theta_2^t \Theta_1)^t] = \text{tr}(\Theta_1^t L^t L \Theta_1)$ .

<sup>3</sup>Proof: Let  $s_1(L) \geq s_2(L)$  denote the singular values of  $L$  (equal by definition to the eigenvalues of  $\sqrt{L^t L}$ ), then  $\|L\| = s_1(L)$ , and  $\|L\|_{Fr} = \sqrt{s_1(L)^2 + s_2(L)^2}$ .

<sup>4</sup>Proof: Let  $\Theta : W \rightarrow V$  be the isometry which maps the base we chose for  $W$  to the base we chose for  $V$ , then  $L\Theta : W \rightarrow W$  is represented w.r.t. the base we chose for  $W$  by the matrix  $(a_{ij})$ . A calculation shows that  $\text{tr}(\Theta^t L^t L \Theta) = \sum a_{ij}^2$ .

**2.3. Pesin Charts.** Having diagonalized the action of the differential of  $f$ , we turn to the action of  $f$  itself. The basic result (due to Pesin [P]) is that  $\text{NUH}_\chi(f)$  has an atlas of charts with respect to which  $f$  is close to a linear hyperbolic map.

Some notation. Let  $\exp_x : T_x M \rightarrow M$  denote the exponential map. We denote the zero vector (in  $T_x M$  or  $\mathbb{R}^2$ ) by  $\underline{0}$ . Balls and boxes are denoted as follows:

$$\begin{aligned} B_\eta(x) &:= \{y \in M : d(x, y) < \eta\} & B_\eta(\underline{0}) &:= \{\underline{v} \in \mathbb{R}^2 : \underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \sqrt{v_1^2 + v_2^2} < \eta\} \\ B_\eta^x(\underline{0}) &:= \{\underline{v} \in T_x M : \|\underline{v}\|_x < \eta\} & R_\eta(\underline{0}) &:= \{\underline{v} \in \mathbb{R}^2 : \underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, |v_1|, |v_2| < \eta\} \end{aligned}$$

Since  $M$  is compact, there exist  $r(M), \rho(M) > 0$  s.t. for every  $x \in M$

$$\exp_x \text{ maps } B_{2r(M)}^x(\underline{0}) \text{ diffeomorphically onto a neighborhood of } B_{\rho(M)}(x). \quad (2.1)$$

We take  $\rho(M)$  so small that  $(x, y) \mapsto \exp_x^{-1}(y)$  is well defined and 2-Lipschitz on  $B_{\rho(M)}(z) \times B_{\rho(M)}(z)$  for all  $z \in M$ , and so small that  $\|(d\exp_x^{-1})_y\| \leq 2$  for all  $y \in B_{\rho(M)}(x)$  (see e.g. [Sp, chapter 9]). Since  $C_\chi$  is a contraction,

$$\Psi_x := \exp_x \circ C_\chi(x) \quad (2.2)$$

maps  $R_{r(M)}(\underline{0})$  diffeomorphically into  $M$ . Since  $C_\chi(x)$  preserves orientation,  $\Psi_x$  preserves orientation.

Let  $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$ , then the linearization of  $f_x$  at  $\underline{0}$  is the linear hyperbolic map  $\begin{pmatrix} \lambda_\chi(x) & 0 \\ 0 & \mu_\chi(x) \end{pmatrix}$ . The question is how large is the neighborhood of  $\underline{0}$  where  $f_x$  can be approximated by its linearization. The size of the neighborhood is known. For reasons that will become clear later, we prefer to define it as a quantity taking values in  $I_\varepsilon := \{e^{-\frac{1}{3}\ell\varepsilon} : \ell \in \mathbb{N}\}$ , where  $\varepsilon$  will be determined later. Set

$$\begin{aligned} Q_\varepsilon(x) &:= \max\{q \in I_\varepsilon : q \leq \tilde{Q}_\chi(x)\} \text{ where} \\ \tilde{Q}_\chi(x) &:= \varepsilon^{3/\beta} (\|C_\chi(x)^{-1}\|_{Fr})^{-12/\beta} \end{aligned} \quad (2.3)$$

**Theorem 2.7** (Pesin). *For all  $\varepsilon$  small enough, and for every  $x \in \text{NUH}_\chi(f)$ ,*

- (1)  $\Psi_x(\underline{0}) = x$  and  $\Psi_x : R_{10Q_\varepsilon(x)}(\underline{0}) \rightarrow M$  is a diffeomorphism onto its image such that  $\|(d\Psi_x)_\underline{u}\| \leq 2$  for every  $\underline{u} \in R_{10Q_\varepsilon(x)}(\underline{0})$ ;
- (2)  $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$  is well defined and injective on  $R_{10Q_\varepsilon(x)}(\underline{0})$  and
  - (a)  $f_x(\underline{0}) = \underline{0}$  and  $(df_x)_\underline{0} = \begin{pmatrix} A(x) & 0 \\ 0 & B(x) \end{pmatrix}$  where  $C_f^{-1} < |A(x)| < e^{-x}$  and  $e^x < |B(x)| < C_f$  (cf. Theorem. 2.3);
  - (b)  $\|f_x - (df_x)_\underline{0}\|_{C^{1+\frac{\beta}{2}}} < \varepsilon$  on  $R_{10Q_\varepsilon(x)}(\underline{0})$ . The  $C^{1+\frac{\beta}{2}}$ -norm of  $r : U \rightarrow \mathbb{R}^2$  on  $U \subset \mathbb{R}^2$  is  $\sup_{\underline{x} \in U} \|r(\underline{x})\| + \sup_{\underline{x} \in U} \|dr_{\underline{x}}\| + \sup_{\underline{x}, \underline{y} \in U, \underline{x} \neq \underline{y}} \frac{\|dr_{\underline{x}} - dr_{\underline{y}}\|}{\|\underline{x} - \underline{y}\|^{\beta/2}}$ .
- (3) The symmetric statement holds for  $f_x^{-1} = \Psi_x^{-1} \circ f^{-1} \circ \Psi_{f(x)}$ .

This is a version of [BP, Theorem 5.6.1]. See the appendix for the proof.

**Definition 2.8.** *Suppose  $x \in \text{NUH}_\chi(f)$  and  $0 < \eta \leq Q_\varepsilon(x)$ . The Pesin chart  $\Psi_x^\eta$  is the map  $\Psi_x : R_\eta(\underline{0}) \rightarrow M$ .*

Some additional information on  $Q_\varepsilon(x)$  (see the appendix for proofs):

**Lemma 2.9.** *The following holds for all  $\varepsilon$  small enough:*

- (1)  $Q_\varepsilon(x) < \varepsilon^{3/\beta}$  on  $\text{NUH}_\chi(f)$ ;
- (2)  $\|C_\chi(f^i(x))^{-1}\|^{12} < \varepsilon^{2/\beta}/Q_\varepsilon(x)$  for  $i = -1, 0, 1$ ;

- (3)  $\{Q_\varepsilon(x) : Q_\varepsilon(x) > t, x \in \text{NUH}_\chi(f)\}$  is finite for all  $t > 0$ ;
- (4)  $\frac{1}{n} \log Q_\varepsilon(f^n(x)) \xrightarrow{n \rightarrow \pm\infty} 0$  on  $\text{NUH}_\chi^*(f)$  (cf. Lemma 2.6);
- (5)  $F^{-1} \leq Q_\varepsilon \circ f / Q_\varepsilon \leq F$  on  $\text{NUH}_\chi(f)$ , where  $F$  is independent of  $\varepsilon$ ;
- (6) there exists a function  $q_\varepsilon : \text{NUH}_\chi^*(f) \rightarrow (0, 1)$  so that  $q_\varepsilon(x) < \varepsilon Q_\varepsilon(x)$  and  $e^{-\varepsilon/3} \leq q_\varepsilon \circ f / q_\varepsilon \leq e^{\varepsilon/3}$  on  $\text{NUH}_\chi^*(f)$ .

**2.4. Distortion compensating bounds.** Our main use of Pesin charts is to analyze local stable and unstable manifolds. First we will use the charts to parameterize the manifolds, and then we will interpret the analytic properties of the parameterizations in terms of the Riemannian metric.

The last step is dangerous, because Pesin charts can distort distances and angles considerably. To see where the distortion comes from, recall that a Pesin chart is given by  $\Psi_x = \exp_x \circ C_\chi(x)$ . The exponential map causes no problems: it is bi-Lipschitz and uniformly smooth. But the linear map  $C_\chi(x)$  can have enormous distortion. We can measure this distortion by  $\|C_\chi(x)^{-1}\|$  (we do not need to worry about  $\|C_\chi(x)\|$  because  $C_\chi(x)$  is contraction). By lemma 2.4,  $\|C_\chi(x)^{-1}\|$  (and therefore the distortion of  $\Psi_x$ ) is large iff

- $s_\chi(x)$  is large (it takes a long time for  $df_x^n$  to contract  $e^s(x)$ ), or
- $u_\chi(x)$  is large (it takes a long time for  $df_x^{-n}$  to contract  $e^u(x)$ ), or
- $|\sin \alpha(x)|$  is small (the stable direction is close to the unstable direction).

So the distortion of  $\Psi_x$  is tied to the quality of hyperbolicity at  $x$ .

For non-uniformly hyperbolic diffeomorphisms, there are no uniform bounds on  $s_\chi(x)$ ,  $u_\chi(x)$  and  $|\sin \alpha(x)|$ . Therefore the distortion of Pesin charts is not bounded.

We will deal with the unbounded distortion of Pesin charts by tying the quality of the estimates we make in Pesin coordinates to the size of  $\|C_\chi(x)^{-1}\|$ : the larger the norm, the stronger the bounds we will require from our parameterized objects. The idea is to make the bounds so strong, that something useful will survive the application of the map  $\Psi_x : R_{Q_\varepsilon(x)}(\mathcal{Q}) \rightarrow M$ . These “distortion compensating bounds” will often take the form

$$\text{distance, error, proximity bound} \leq \text{const } Q_\varepsilon(x)^{\text{some power}} \text{ or } \text{const } \eta^{\text{some power}}$$

where  $x$  is the center of the chart and  $0 < \eta \leq Q_\varepsilon(x)$ .

Since  $Q_\varepsilon(x) \ll \|C_\chi(x)^{-1}\|^{-\text{big power}}$ , this will do the work provided the powers are chosen correctly.

**2.5.  $\text{NUH}_\chi^\#(f)$ .** The set  $\text{NUH}_\chi^*(f)$  constructed in Lemma 2.6 is  $\chi$ -large. By the Poincaré Recurrence Theorem, the set

$$\text{NUH}_\chi^\#(f) := \{x \in \text{NUH}_\chi^*(f) : \limsup_{n \rightarrow \infty} q_\varepsilon(f^n(x)), \limsup_{n \rightarrow \infty} q_\varepsilon(f^{-n}(x)) \neq 0\} \quad (2.4)$$

is  $\chi$ -large. This is the set that we will attempt to cover by a Markov partition.

### 3. OVERLAPPING CHARTS

We would like to replace  $\mathcal{C} := \{\Psi_x^\eta : x \in \text{NUH}_\chi^*(f), 0 < \eta \leq Q_\varepsilon(x)\}$  by a countable collection  $\mathcal{A}$  in such a way that every element of  $\mathcal{C}$  “overlaps” some element of  $\mathcal{A}$  “well”. Later, we will use  $\mathcal{A}$  to construct the set of vertices of a directed graph related to the dynamics of  $f$ .

**3.1. The overlap condition.** We need to compare the maps  $C_\chi(x) : \mathbb{R}^2 \rightarrow T_x M$  for different  $x \in M$ , even though they take values in different spaces. We circumvent the problem as follows. Every  $x \in M$  has an open neighborhood  $D$  of diameter less than  $\rho(M)$  and a smooth map  $\Theta_D : TD \rightarrow \mathbb{R}^2$  s.t.

- (1)  $\Theta_D : T_x M \rightarrow \mathbb{R}^2$  is a linear isometry for every  $x \in D$ ;
- (2) let  $\vartheta_x := (\Theta_D|_{T_x M})^{-1} : \mathbb{R}^2 \rightarrow T_x M$ , then  $(x, \underline{u}) \mapsto (\exp_x \circ \vartheta_x)(\underline{u})$  is smooth and Lipschitz on  $D \times B_2(\mathbb{0})$  with respect to the metric  $d(x, x') + \|\underline{u} - \underline{u}'\|$ ;
- (3)  $x \mapsto \vartheta_x^{-1} \circ \exp_x^{-1}$  is a Lipschitz map from  $D$  into  $C^2(D, \mathbb{R}^2)$ , the space of  $C^2$  maps from  $D$  to  $\mathbb{R}^2$ .

Let  $\mathcal{D}$  be an finite cover of  $M$  by such neighborhoods. Let  $\varepsilon(\mathcal{D})$  be a Lebesgue number for  $\mathcal{D}$ . If  $d(x, y) < \varepsilon(\mathcal{D})$ , then  $x, y$  fall in some element  $D$ . Instead of comparing  $C_\chi(x)$  to  $C_\chi(y)$ , we will compare  $\Theta_D \circ C_\chi(x)$  to  $\Theta_D \circ C_\chi(y)$  (two linear maps from  $\mathbb{R}^2$  to  $\mathbb{R}^2$ ).

**Definition 3.1.** *Two Pesin charts  $\Psi_{x_1}^{\eta_1}, \Psi_{x_2}^{\eta_2}$   $\varepsilon$ -overlap if  $e^{-\varepsilon} < \frac{\eta_1}{\eta_2} < e^\varepsilon$ , and for some  $D \in \mathcal{D}$ ,  $x_1, x_2 \in D$  and  $d(x_1, x_2) + \|\Theta_D \circ C_\chi(x_1) - \Theta_D \circ C_\chi(x_2)\| < \eta_1^4 \eta_2^4$ .*

The overlap condition is symmetric. It is also monotone: if  $\Psi_{x_i}^{\eta_i}$   $\varepsilon$ -overlap, then  $\Psi_{x_i}^{\xi_i}$   $\varepsilon$ -overlap for all  $\eta_i \leq \xi_i \leq Q_\varepsilon(x_i)$  s.t.  $e^{-\varepsilon} < \xi_1/\xi_2 < e^\varepsilon$ . Notice that the overlap requirement is stronger at areas of  $\text{NUH}_\chi(f)$  where  $s_\chi(x)$  or  $u_\chi(x)$  are large or where  $\underline{e}^s(x)$  and  $\underline{e}^u(x)$  are nearly parallel. This is because by construction

$$\eta_i \leq Q_\varepsilon(x_i) \ll \|C_\chi(x_i)^{-1}\|_{Fr}^{-1} = \frac{|\sin \alpha(x)|}{\sqrt{s_\chi(x)^2 + u_\chi(x)^2}}.$$

The following proposition explains what the overlap condition means.

**Proposition 3.2.** *The following holds for all  $\varepsilon$  small. If  $\Psi_{x_1} : R_{\eta_1}(\mathbb{0}) \rightarrow M$  and  $\Psi_{x_2} : R_{\eta_2}(\mathbb{0}) \rightarrow M$   $\varepsilon$ -overlap, then*

- (1)  $\Psi_{x_1}[R_{e^{-2\varepsilon}\eta_1}(\mathbb{0})] \subset \Psi_{x_2}[R_{\eta_2}(\mathbb{0})]$  and  $\Psi_{x_2}[R_{e^{-2\varepsilon}\eta_2}(\mathbb{0})] \subset \Psi_{x_1}[R_{\eta_1}(\mathbb{0})]$ ;
- (2)  $\text{dist}_{C^{1+\frac{\beta}{2}}}(\Psi_{x_i}^{-1} \circ \Psi_{x_j}, \text{Id}) < \varepsilon \eta_i^2 \eta_j^2$  ( $\{i, j\} = \{1, 2\}$ ), where the  $C^{1+\frac{\beta}{2}}$ -distance is calculated on  $R_{e^{-\varepsilon}r(M)}(\mathbb{0})$  and  $r(M)$  is defined in (2.1).

*Remark.* By (2), the greater the distortion of  $\Psi_{x_1}$  or  $\Psi_{x_2}$ , the closer they are to one another. This distortion compensating bound will be used in the sequel to argue that  $\Psi_{f(x)}^{-1} \circ f \circ \Psi_x$  remains close to a linear hyperbolic map if we replace  $\Psi_{f(x)}$  by an overlapping chart  $\Psi_y$  (Proposition 3.4 below).

*Proof.* Suppose  $\Psi_{x_i}^{\eta_i}$   $\varepsilon$ -overlap, and fix some  $D \in \mathcal{D}$  which contains  $x_1$  and  $x_2$  such that  $d(x_1, x_2) + \|\Theta_D \circ C_\chi(x_1) - \Theta_D \circ C_\chi(x_2)\| < \eta_1^4 \eta_2^4$ . Write  $C_i := \Theta_D \circ C_\chi(x_i)$ , then  $\Psi_{x_i} = \exp_{x_i} \circ \vartheta_{x_i} \circ C_i$ .

By the definition of Pesin charts,  $\eta_i \leq Q_\varepsilon(x_i)$ , where  $Q_\varepsilon(x_i)$  is given by (2.3). Lemma 2.5 and the general inequality  $\|\cdot\|_{Fr} \geq \|\cdot\|$  (see page 9) guarantee that

$$\eta_i \leq \varepsilon^{3/\beta} \|C_\chi(x_i)^{-1}\|^{-12/\beta}. \quad (3.1)$$

In particular,  $\eta_i < \varepsilon^{3/\beta}$ .

Our first constraint on  $\varepsilon$  is that it be so small that

$$\varepsilon^{3/\beta} < \frac{\min\{1, r(M), \rho(M)\}}{5(L_1 + L_2 + L_3 + L_4)^3}, \quad (3.2)$$

where  $r(M)$  and  $\rho(M)$  are given by (2.1), and

- (1)  $L_1$  is a common Lipschitz constant for the maps  $(x, v) \mapsto (\exp_x \circ \vartheta_x)(v)$  on  $D \times B_{r(M)}(\mathbb{0})$  ( $D \in \mathcal{D}$ );
- (2)  $L_2$  is a common Lipschitz constant for the maps  $x \mapsto \vartheta_x^{-1} \circ \exp_x^{-1}$  from  $D$  into  $C^2(D, \mathbb{R}^2)$  ( $D \in \mathcal{D}$ );
- (3)  $L_3$  is a common Lipschitz constant for  $\exp_x^{-1} : B_{\rho(M)}(x) \rightarrow T_x M$  ( $x \in M$ );
- (4)  $L_4$  is a common Lipschitz constant for  $\exp_x : B_{r(M)}^x(\mathbb{0}) \rightarrow M$  ( $x \in M$ ).

We assume w.l.o.g. that these constants are all larger than one.

*Part 1.*  $\Psi_{x_1}[R_{e^{-2\varepsilon}\eta_1}(\mathbb{0})] \subset \Psi_{x_2}[R_{\eta_2}(\mathbb{0})]$ .

*Proof.* Suppose  $v \in R_{e^{-2\varepsilon}\eta_1}(\mathbb{0})$ . Lemma 2.5 says that  $C_\chi(x_1)$  is a contraction, therefore  $\|C_1 v\| = \|C_\chi(x_1)v\| \leq \|v\|$ , and  $(x_1, C_1 v), (x_2, C_1 v) \in D \times B_{r(M)}(\mathbb{0})$ . Since  $d(x_1, x_2) < \eta_1^4 \eta_2^4$ ,

$$d(\exp_{x_2} \circ \vartheta_{x_2}[C_1 v], \exp_{x_1} \circ \vartheta_{x_1}[C_1 v]) < L_1 \eta_1^4 \eta_2^4.$$

It follows that  $\Psi_{x_1}(v) \in B_{L_1 \eta_1^4 \eta_2^4}(\exp_{x_2} \circ \vartheta_{x_2}(C_1 v))$ . Call this ball  $B$ .

The radius of  $B$  is less than  $\rho(M)$  because of our assumptions on  $\varepsilon$ . Therefore  $\exp_{x_2}^{-1}$  is well defined and Lipschitz on  $B$ , and its Lipschitz constant is at most  $L_3$ . Writing  $B = \exp_{x_2}[\exp_{x_2}^{-1}(B)]$ , we deduce that

$$\Psi_{x_1}(v) \in B \subset \exp_{x_2}[B_{L_3 L_1 \eta_1^4 \eta_2^4}^{x_2}(\vartheta_{x_2}(C_1 v))] =: \Psi_{x_2}[E],$$

where  $E := C_\chi(x_2)^{-1}[B_{L_3 L_1 \eta_1^4 \eta_2^4}^{x_2}(\vartheta_{x_2}(C_1 v))]$ .

We claim that  $E \subset R_{\eta_2}(\mathbb{0})$ . First note that  $E \subset B_{\|C_\chi(x_2)^{-1}\| L_3 L_1 \eta_1^4 \eta_2^4}(C_2^{-1} C_1 v)$ , therefore if  $\underline{w} \in E$ , then

$$\begin{aligned} \|\underline{w}\|_\infty &\leq \|C_2^{-1} C_1 v\|_\infty + \|C_\chi(x_2)^{-1}\| L_3 L_1 \eta_1^4 \eta_2^4 \\ &\leq \|(C_2^{-1} C_1 - \text{Id})v\|_\infty + \|v\|_\infty + \|C_\chi(x_2)^{-1}\| L_3 L_1 \eta_1^4 \eta_2^4 \\ &\leq \|v\|_\infty + \sqrt{2} \|C_2^{-1}\| \|C_1 - C_2\| \|v\|_\infty + \|C_\chi(x_2)^{-1}\| L_3 L_1 \eta_1^4 \eta_2^4 \\ &\leq e^{-2\varepsilon} \eta_1 + \|C_\chi(x_2)^{-1}\| (\eta_1^4 \eta_2^4 \sqrt{2} e^{-2\varepsilon} \eta_1 + L_3 L_1 \eta_1^4 \eta_2^4) \quad (\because \|C_1 - C_2\| < \eta_1^4 \eta_2^4) \\ &\leq e^{-2\varepsilon} \eta_1 + \|C_\chi(x_2)^{-1}\| \eta_2^4 \cdot [(e^{-2\varepsilon} \sqrt{2} \eta_1 + L_3 L_1) \eta_1^3] \cdot \eta_1 \\ &< e^{-2\varepsilon} \eta_1 + \varepsilon^2 \eta_1, \quad \text{because of (3.1) and (3.2)} \\ &< e^\varepsilon (e^{-2\varepsilon} + \varepsilon^2) \eta_2 < \eta_2, \quad \text{because } \eta_1 < e^\varepsilon \eta_2 \text{ and } 0 < \varepsilon < \frac{1}{5} \text{ by (3.2)}. \end{aligned}$$

It follows that  $E \subset R_{\eta_2}(\mathbb{0})$ . Thus  $\Psi_{x_1}(v) \in \Psi_{x_2}[R_{\eta_2}(\mathbb{0})]$ . Part 1 follows.

*Part 2.* The  $C^{1+\beta/2}$ -distance between  $\Psi_{x_1}^{-1} \circ \Psi_{x_2}$  on  $R_{e^{-\varepsilon}r(M)}(\mathbb{0})$  is less than  $\varepsilon \eta_1$ .

*Proof.* One can show exactly as in the proof of part 1 that  $\Psi_{x_1}[R_{e^{-\varepsilon}r(M)}(\mathbb{0})] \subset \Psi_{x_2}[R_{r(M)}(\mathbb{0})]$ , therefore  $\Psi_{x_1}^{-1} \circ \Psi_{x_2}$  is well defined on  $R_{e^{-\varepsilon}r(M)}(\mathbb{0})$ . We calculate the distance of this map from the identity:

$$\begin{aligned} \Psi_{x_1}^{-1} \circ \Psi_{x_2} &= C_1^{-1} \circ \vartheta_{x_1}^{-1} \circ \exp_{x_1}^{-1} \circ \exp_{x_2} \circ \vartheta_{x_2} \circ C_2 \\ &= C_1^{-1} \circ [\vartheta_{x_1}^{-1} \circ \exp_{x_1}^{-1} + \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} - \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1}] \circ \exp_{x_2} \circ \vartheta_{x_2} \circ C_2 \\ &= C_1^{-1} C_2 + C_1^{-1} \circ [\vartheta_{x_1}^{-1} \circ \exp_{x_1}^{-1} - \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1}] \circ \Psi_{x_2} \\ &= \text{Id} + C_1^{-1}(C_2 - C_1) + C_1^{-1} \circ [\vartheta_{x_1}^{-1} \circ \exp_{x_1}^{-1} - \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1}] \circ \Psi_{x_2}. \end{aligned}$$

The  $C^{1+\beta/2}$ -norm of the second summand is less than  $\|C_1^{-1}\|\eta_1^4\eta_2^4$ . The  $C^{1+\beta/2}$ -norm of the third summand is less than  $\|C_1^{-1}\|\cdot L_2d(x_1, x_2)\cdot L_4^{1+\frac{\beta}{2}}$ . This is less than  $\|C_1^{-1}\|L_2L_4^2\eta_1^4\eta_2^4$ .

It follows that  $\text{dist}_{C^{1+\beta/2}}(\Psi_{x_1}^{-1}\circ\Psi_{x_2}, \text{Id}) < \|C_1^{-1}\|(1+L_2L_4^2)\eta_1^4\eta_2^4$ . This is (much) smaller than  $\varepsilon\eta_1^2\eta_2^2$ , because of (3.1) and (3.2).  $\square$

The following distortion compensating bound is needed in §7 below:

**Lemma 3.3.** *Suppose  $\Psi_{x_1}^{\eta_1}, \Psi_{x_2}^{\eta_2}$   $\varepsilon$ -overlap, then*

$$\frac{s_\chi(x_1)}{s_\chi(x_2)}, \frac{u_\chi(x_1)}{u_\chi(x_2)} \in [e^{-Q_\varepsilon(x_1)Q_\varepsilon(x_2)}, e^{Q_\varepsilon(x_1)Q_\varepsilon(x_2)}].$$

*Proof.* We use the notation of the previous proof.  $\Psi_{x_2}^{-1}\circ\Psi_{x_1}$  maps  $R_{e^{-\varepsilon\eta_1}}(\underline{0})$  into  $\mathbb{R}^2$ . Its derivative at the origin is

$$\begin{aligned} A &:= C_\chi(x_2)^{-1}d(\exp_{x_2}^{-1})_{x_1}C_\chi(x_1) = C_2^{-1}d[\vartheta_{x_2}^{-1}\exp_{x_2}^{-1}]_{x_1}\vartheta_{x_1}C_1 \\ &= C_2^{-1}C_1 + C_2^{-1}[d[\vartheta_{x_2}^{-1}\exp_{x_2}^{-1}]_{x_1} - \vartheta_{x_1}^{-1}]\vartheta_{x_1}C_1 \\ &\equiv C_2^{-1}C_1 + C_2^{-1}(d[\vartheta_{x_2}^{-1}\exp_{x_2}^{-1}]_{x_1} - d[\vartheta_{x_1}^{-1}\exp_{x_1}^{-1}]_{x_1})\vartheta_{x_1}C_1. \end{aligned}$$

Since  $\|d[\vartheta_{x_2}^{-1}\exp_{x_2}^{-1}]_{x_1} - d[\vartheta_{x_1}^{-1}\exp_{x_1}^{-1}]_{x_1}\| < L_2d(x_1, x_2) < L_2\eta_1^4\eta_2^4 < \varepsilon\eta_1^2\eta_2^2$ , and  $\vartheta_{x_1}C_1$  is a contraction, and  $\|A - \text{Id}\| < \text{dist}_{C^1}(\Psi_{x_2}^{-1}\circ\Psi_{x_1}, \text{Id}) < \varepsilon\eta_1^2\eta_2^2$ ,

$$\|C_2^{-1}C_1 - \text{Id}\| < 2\varepsilon\|C_2^{-1}\|\eta_1^2\eta_2^2.$$

Since  $\|C_2\| \leq 1$ , we have that  $\|C_1 - C_2\| < 2\varepsilon\|C_2^{-1}\|\eta_1^2\eta_2^2$ .

Recall that  $s_\chi(x_i)^{-1} = \|C_\chi(x_i)\underline{e}_1\|$  and  $s_\chi(x_i) = \|C_\chi(x_i)^{-1}\underline{e}^s(x_i)\|$ , so

$$\begin{aligned} \left| \frac{s_\chi(x_1)}{s_\chi(x_2)} - 1 \right| &= \left| \frac{s_\chi(x_2)^{-1} - s_\chi(x_1)^{-1}}{s_\chi(x_1)^{-1}} \right| \\ &\leq \|C_\chi(x_1)^{-1}\| \cdot \left| \|C_\chi(x_1)\underline{e}_1\| - \|C_\chi(x_2)\underline{e}_1\| \right| \\ &= \|C_1^{-1}\| \cdot \left| \|C_1\underline{e}_1\| - \|C_2\underline{e}_1\| \right| \\ &\leq \|C_1^{-1}\| \cdot \|C_1 - C_2\| < 2\varepsilon\|C_1^{-1}\|\|C_2^{-1}\|\eta_1^2\eta_2^2 < \varepsilon\eta_1\eta_2. \end{aligned}$$

Similarly  $\left| \frac{u_\chi(x_1)}{u_\chi(x_2)} - 1 \right| < \varepsilon\eta_1\eta_2$ . Since  $\eta_i < Q_\varepsilon(x_i)$ , the lemma follows.  $\square$

**3.2. The form of  $f$  in overlapping charts.** Theorem 2.7 says that  $\Psi_{f(x)}^{-1}\circ f\circ\Psi_x$  is close to a linear hyperbolic map. This remains the case if we replace  $\Psi_{f(x)}$  by some overlapping chart  $\Psi_y$ :

**Proposition 3.4.** *The following holds for all  $\varepsilon$  small enough. Suppose  $x, y \in \text{NUH}_\chi(f)$  and  $\Psi_{f(x)}^\eta$   $\varepsilon$ -overlaps  $\Psi_y^{\eta'}$ , then  $f_{xy} := \Psi_y^{-1}\circ f\circ\Psi_x$  is a well defined injective map from  $R_{10Q_\varepsilon(x)}(\underline{0})$  to  $\mathbb{R}^2$ , and  $f_{xy}$  can be put in the form*

$$f_{xy}(u, v) = (Au + h_1(u, v), Bv + h_2(u, v)), \quad (3.3)$$

where  $C_f^{-1} < |A| < e^{-\chi}$ ,  $e^\chi < |B| < C_f$  (cf. Theorem 2.3),  $|h_i(\underline{0})| < \varepsilon\eta$ ,  $\|\nabla h_i(\underline{0})\| < \varepsilon\eta^{\beta/3}$ , and  $\|\nabla h_i(u) - \nabla h_i(v)\| \leq \varepsilon\|u - v\|^{\beta/3}$  on  $R_{10Q_\varepsilon(x)}(\underline{0})$ .

A similar statement holds for  $f_{xy}^{-1}$ , assuming that  $\Psi_{f^{-1}(y)}^{\eta'}$   $\varepsilon$ -overlaps  $\Psi_x^\eta$ .

*Proof.* We write  $f_{xy} = (\Psi_y^{-1} \circ \Psi_{f(x)}) \circ f_x$  where  $f_x = \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$ , and treat  $f_{xy}$  as a perturbation of  $f_x$ .

By Theorem 2.7, if  $\varepsilon$  is small enough, then  $f_x$  has the following properties:

- (1) It is well-defined, differentiable, and injective on  $R_{10Q_\varepsilon(x)}(\mathbb{0})$ .
- (2)  $f_x(\mathbb{0}) = \mathbb{0}$  and  $(df_x)_{\mathbb{0}} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$  where  $C_f^{-1} < |A| < e^{-\chi}$ ,  $e^\chi < |B| < C_f$ .
- (3) For all  $\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\mathbb{0})$ ,  $\|(df_x)_{\underline{u}} - (df_x)_{\underline{v}}\| \leq 2\varepsilon\|\underline{u} - \underline{v}\|^{\beta/2}$  (because the  $C^{1+\frac{\beta}{2}}$  distance between  $f_x$  and  $(df_x)_{\mathbb{0}}$  on  $R_{10Q_\varepsilon(x)}(\mathbb{0})$  is less than  $\varepsilon$ ).
- (4) For every  $0 < \eta < 10Q_\varepsilon(x)$  and  $\underline{u} \in R_\eta(\mathbb{0})$ ,  $\|(df_x)_{\underline{u}}\| < 3C_f$ , provided  $\varepsilon$  is small enough (because  $\|(df_x)_{\underline{u}}\| \leq \|(df_x)_{\mathbb{0}}\| + \varepsilon\eta^{\beta/2} < 2C_f + \varepsilon$ ).

(2) and (4) imply that  $f_x[R_{10Q_\varepsilon(x)}(\mathbb{0})] \subset B_{30Q_\varepsilon(x)C_f}(\mathbb{0})$ . Since  $Q_\varepsilon(x) < \varepsilon^{3/\beta}$ ,  $f_x[R_{10Q_\varepsilon(x)}(\mathbb{0})] \subset B_{30C_f\varepsilon^{3/\beta}}(\mathbb{0})$ . If  $\varepsilon$  is so small that  $30C_f\varepsilon^{3/\beta} < e^{-\varepsilon}r(M)$ , then  $f_x[R_{10Q_\varepsilon(x)}(\mathbb{0})] \subset R_{e^{-\varepsilon}r(M)}(\mathbb{0})$ .  $R_{e^{-\varepsilon}r(M)}(\mathbb{0})$  is in the domain of  $\Psi_y^{-1} \circ \Psi_{f(x)}$  (Proposition 3.2, part 2), therefore  $f_{xy}$  is well defined, differentiable, and injective on  $R_{10Q_\varepsilon(x)}(\mathbb{0})$ .

Equation (3.3) can be used to define the functions  $h_i(u, v)$ . We check that they satisfy the properties in the statement.

We have  $(h_1(\mathbb{0}), h_2(\mathbb{0})) = f_{xy}(\mathbb{0}) = \Psi_y^{-1}(f(x)) = (\Psi_y^{-1} \circ \Psi_{f(x)})(\mathbb{0})$ , therefore  $\|(h_1(\mathbb{0}), h_2(\mathbb{0}))\| \leq \text{dist}_{C^0}(\Psi_y^{-1} \circ \Psi_{f(x)}, \text{Id}) < \varepsilon\eta^2(\eta')^2 < \varepsilon\eta$ .

We differentiate the identity  $f_{xy} = (\Psi_y^{-1} \circ \Psi_{f(x)}) \circ f_x$  at an arbitrary  $\underline{u} \in R_\eta(\mathbb{0})$ . The result, after some rearrangement is

$$(df_{xy})_{\underline{u}} = [d(\Psi_y^{-1} \circ \Psi_{f(x)})_{f_x(\underline{u})} - \text{Id}](df_x)_{\underline{u}} + [(df_x)_{\underline{u}} - (df_x)_{\mathbb{0}}] + (df_x)_{\mathbb{0}}. \quad (3.4)$$

The norm of the first summand is less than  $3C_f \text{dist}_{C^1}(\Psi_y^{-1} \circ \Psi_{f(x)}, \text{Id})$ , which by Proposition 3.2 is less than  $3C_f\varepsilon\eta^2(\eta')^2 < 3C_f\varepsilon\eta^2$ . The norm of the second summand is less than  $\varepsilon\|\underline{u}\|^{\beta/2} < 2\varepsilon\eta^{\beta/2}$ . The third term is  $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ . Thus

$$\begin{aligned} \left\| \frac{\partial(h_1, h_2)}{\partial(u, v)} \right\| &= \left\| (df_{xy})_{\underline{u}} - \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \right\| < \varepsilon[3C_f + 2]\eta^{\beta/2} \\ &< \varepsilon\eta^{\beta/3} \cdot [3C_f + 2]\eta^{\beta/6} < \varepsilon\eta^{\beta/3} \cdot [3C_f + 2]\sqrt{\varepsilon} \text{ by (3.1)}. \end{aligned}$$

If  $\varepsilon$  is so small that  $[3C_f + 2]\sqrt{\varepsilon} < 1$ , then  $\|\nabla h_i\| < \varepsilon\eta^{\beta/3}$  on  $R_\eta(\mathbb{0})$ . In particular,  $\|\nabla h_i(\mathbb{0})\| < \varepsilon\eta^{\beta/3}$ .

Equation (3.4) also shows that for every  $\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\mathbb{0})$ ,

$$\begin{aligned} \|(df_{xy})_{\underline{u}} - (df_{xy})_{\underline{v}}\| &\leq \|d(\Psi_y^{-1} \circ \Psi_{f(x)})_{f_x(\underline{u})} - d(\Psi_y^{-1} \circ \Psi_{f(x)})_{f_x(\underline{v})}\| \cdot \|(df_x)_{\underline{u}}\| \\ &\quad + \|(df_x)_{\underline{u}} - (df_x)_{\underline{v}}\| \cdot (\|d(\Psi_y^{-1} \circ \Psi_{f(x)})_{f_x(\underline{v})}\| + 1). \end{aligned}$$

By Proposition 3.2,  $\text{dist}_{C^{1+\beta/2}}(\Psi_y^{-1} \circ \Psi_{f(x)}, \text{Id}) < \varepsilon\eta^2(\eta')^2$ , therefore

$$\begin{aligned} \|(df_{xy})_{\underline{u}} - (df_{xy})_{\underline{v}}\| &\leq \varepsilon\eta^2(\eta')^2 \cdot \|f_x(\underline{u}) - f_x(\underline{v})\|^{\frac{\beta}{2}} \cdot 3C_f + 2\varepsilon\|\underline{u} - \underline{v}\|^{\frac{\beta}{2}} (\varepsilon\eta^2(\eta')^2 + 2) \\ &\leq \varepsilon\eta^2 \cdot \sup_{\underline{w} \in R_{10Q_\varepsilon(x)}(\mathbb{0})} \|(df_x)_{\underline{w}}\|^{\frac{\beta}{2}} \cdot \|\underline{u} - \underline{v}\|^{\frac{\beta}{2}} \cdot 3C_f + 5\varepsilon\|\underline{u} - \underline{v}\|^{\frac{\beta}{2}} \\ &\leq \varepsilon((3C_f)^{1+\frac{\beta}{2}}\eta^2 + 5)\|\underline{u} - \underline{v}\|^{\frac{\beta}{2}} \leq \varepsilon((3C_f)^{1+\frac{\beta}{2}}\varepsilon^{6/\beta} + 5)\|\underline{u} - \underline{v}\|^{\frac{\beta}{2}} \\ &\leq 6\varepsilon\|\underline{u} - \underline{v}\|^{\frac{\beta}{2}}, \text{ provided } \varepsilon \text{ is small enough} \end{aligned}$$

$$\begin{aligned} &\leq 6\varepsilon(30Q_\varepsilon(x))^{\beta/6}\|\underline{u}-\underline{v}\|^{\beta/3} < 12\varepsilon^{3/2}\|\underline{u}-\underline{v}\|^{\beta/3} \quad (\because Q_\varepsilon < \varepsilon^{3/\beta}) \\ &\leq \frac{1}{3}\varepsilon\|\underline{u}-\underline{v}\|^{\beta/3}, \text{ provided } \varepsilon \text{ is small enough.} \end{aligned}$$

It follows that  $\|\frac{\partial(h_1, h_2)}{\partial(u, v)}(\underline{u}) - \frac{\partial(h_1, h_2)}{\partial(u, v)}(\underline{v})\| < \frac{1}{3}\varepsilon\|\underline{u}-\underline{v}\|^{\beta/3}$  for all  $\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\underline{0})$ , whence  $\|\nabla h_i(\underline{u}) - \nabla h_i(\underline{v})\| \leq \frac{1}{3}\varepsilon\|\underline{u}-\underline{v}\|^{\beta/3}$  ( $i = 1, 2$ ) for all  $\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\underline{0})$ .  $\square$

**3.3. Coarse graining.** We replace  $\mathcal{C} := \{\Psi_x^\eta : x \in \text{NUH}_\chi^*(f), 0 < \eta \leq Q_\varepsilon(x)\}$  by a ‘‘sufficient’’ countable subset  $\mathcal{A}$ . We remind the reader that  $\text{NUH}_\chi^*$  is defined in Lemma 2.6, and that  $I_\varepsilon = \{e^{-\frac{1}{3}k\varepsilon} : k \in \mathbb{N}\}$ .

**Proposition 3.5.** *The following holds for all  $\varepsilon$  small. There exists a countable collection  $\mathcal{A}$  of Pesin charts with the following properties:*

- (1) Discreteness:  $\{\Psi_x^\eta \in \mathcal{A} : \eta > t\}$  is finite for every  $t > 0$ .
- (2) Sufficiency: For every  $x \in \text{NUH}_\chi^*(f)$  and for every sequence of positive numbers  $0 < \eta_n \leq e^{-\varepsilon/3}Q_\varepsilon(f^n(x))$  in  $I_\varepsilon$  s.t.  $e^{-\varepsilon} \leq \eta_n/\eta_{n+1} \leq e^\varepsilon$ , there exists a sequence  $\{\Psi_{x_n}^{\eta_n}\}_{n \in \mathbb{Z}}$  of elements of  $\mathcal{A}$  s.t. for every  $n$ ,
  - (a)  $\Psi_{x_n}^{\eta_n}$   $\varepsilon$ -overlaps  $\Psi_{f^n(x)}^{\eta_n}$  and  $e^{-\varepsilon/3} \leq Q_\varepsilon(f^n(x))/Q_\varepsilon(x_n) \leq e^{\varepsilon/3}$ ;
  - (b)  $\Psi_{f(x_n)}^{\eta_{n+1}}$   $\varepsilon$ -overlaps  $\Psi_{x_{n+1}}^{\eta_{n+1}}$ ;
  - (c)  $\Psi_{f^{-1}(x_n)}^{\eta_{n-1}}$   $\varepsilon$ -overlaps  $\Psi_{x_{n-1}}^{\eta_{n-1}}$ ;
  - (d)  $\Psi_{x_n}^{\eta'_n} \in \mathcal{A}$  for all  $\eta'_n \in I_\varepsilon$  s.t.  $\eta_n \leq \eta'_n \leq \min\{Q_\varepsilon(x_n), e^\varepsilon \eta_n\}$ .

*Proof.* The general idea is simple: A chart  $\Psi_x^\eta$  is given by a point  $x$ , a matrix  $C_\chi(x)$ , and a real number  $\eta$ . The spaces of points, matrices and real numbers are separable, so all that one needs to do is to find a sufficiently dense discrete subset.

But there is a twist:  $\Psi_x$  does not necessarily depend continuously on  $x$ , because  $x \mapsto C_\chi(x)$  is not necessarily continuous. As a result there is no clear connection between conditions (a), (b), and (c), and we are forced to treat them separately. The following construction will help us to do this. Let

$$X := M^3 \times (0, \infty)^3 \times \text{GL}(2, \mathbb{R})^3,$$

together with the product topology. Next recall the finite open cover  $\mathcal{D}$  of  $M$  from §3.1, and let  $Y \subset X$  denote the collection of all  $(\underline{x}, \underline{Q}, \underline{C}) \in X$  where

- $\underline{x} = (x, f(x), f^{-1}(x))$ ,  $x \in \text{NUH}_\chi^*(f)$ ;
- $\underline{Q} = (Q_\varepsilon(x), Q_\varepsilon(f(x)), Q_\varepsilon(f^{-1}(x)))$  (cf. (2.3));
- $\underline{C} = (\Theta_{D_0} \circ C_\chi(x), \Theta_{D_1} \circ C_\chi(f(x)), \Theta_{D_{-1}} \circ C_\chi(f^{-1}(x)))$ , where  $D_0, D_1, D_{-1} \in \mathcal{D}$  satisfy  $(x, f(x), f^{-1}(x)) \in D_0 \times D_1 \times D_{-1}$ .

Let  $Y_k := \{(\underline{x}, \underline{Q}, \underline{C}) \in Y : x \in \text{NUH}_\chi^*(f), e^{-(k+1)} \leq Q_\varepsilon(x) \leq e^{-(k-1)}\}$  ( $k \in \mathbb{N}$ ).  $Y_k$  is a pre-compact subset of  $X$ . To see this, pick some  $(\underline{x}, \underline{Q}, \underline{C}) \in Y_k$ . The vector  $\underline{x}$  belongs to the compact set  $M^3$ .  $\underline{Q}$  belongs to a compact subset of  $(0, \infty)^3$  because by Lemma 2.9 for each  $i = -1, 0, 1$ ,

$$F^{-1}e^{-(k+1)} \leq Q_\varepsilon(f^i(x)) \leq Fe^{-(k-1)}.$$

$\underline{C}$  belongs to a compact subset of  $\text{GL}(2, \mathbb{R})$ , because (a)  $\Theta_{D_i}$  are isometries; (b)  $\|C_\chi(f^i(x))\| < 1$  (Lemma 2.5); and (c)  $\|C_\chi(f^i(x))^{-1}\| \leq (\varepsilon^{3/\beta} Fe^{k+1})^{\beta/12}$  by (2.3).<sup>5</sup> It follows that  $Y_k$  is a subset of a compact subset of  $M^3 \times (0, \infty)^3 \times \text{GL}(2, \mathbb{R})^3$ .

<sup>5</sup>Here we use the obvious observation that  $\{A \in \text{GL}(2, \mathbb{R}) : \|A\|, \|A^{-1}\| \leq C\}$  is a compact subset of  $\text{GL}(2, \mathbb{R})$  for every  $C > 0$ .



Since  $Y_k$  is pre-compact, it contains a finite set  $Y_{k,m}$  s.t. for every  $(\underline{x}, \underline{Q}, \underline{C}) \in Y_k$  there exists some  $(\underline{y}, \underline{Q}', \underline{C}') \in Y_{k,m}$  such that for every  $|i| \leq 1$ ,

- (1)  $d(f^i(x), f^i(y)) < \frac{1}{2}\varepsilon(\mathcal{D})$  where  $\varepsilon(\mathcal{D})$  is a Lebesgue number of  $\mathcal{D}$ .
- (2)  $d(f^i(x), f^i(y)) + \|\Theta_D \circ C_\chi(f^i(x)) - \Theta_D \circ C_\chi(f^i(y))\| < e^{-8(m+2)}$  for every  $D \in \mathcal{D}$  which contains  $f^i(x)$  and  $f^i(y)$ .
- (3)  $e^{-\varepsilon/3} < Q_\varepsilon(f^i(x))/Q_\varepsilon(f^i(y)) < e^{\varepsilon/3}$ .

Define  $\mathcal{A}$  to be the collection of all Pesin charts  $\Psi_x^\eta$  such that for some  $k, m \in \mathbb{N}$ ,  $x$  is the first coordinate of some element  $(\underline{x}, \underline{Q}, \underline{C}) \in Y_{k,m}$ , and

$$0 < \eta \leq Q_\varepsilon(x), e^{-(m+2)} \leq \eta < e^{-(m-2)}, \text{ and } \eta \in I_\varepsilon = \{e^{-\ell\varepsilon/3} : \ell = 0, 1, 2, \dots\}.$$

*Part 1. Discreteness.*

*Proof.* Suppose  $\Psi_x^\eta \in \mathcal{A}$ . Choose  $k, m \in \mathbb{N}$  s.t.  $x$  is the first coordinate of some  $(\underline{x}, \underline{Q}, \underline{C}) \in Y_{k,m}$ ,  $0 < \eta \leq Q_\varepsilon(x)$ , and  $\eta \in [e^{-m-2}, e^{-m+2}]$ . Since  $Y_{k,m} \subset Y_k$ ,  $Q_\varepsilon(x) \leq e^{-k+1}$ , so  $k \leq |\log Q_\varepsilon(x)| + 1$ . It follows that  $k, m \leq |\log \eta| + 2$ , and so

$$|\{\Psi_x^\eta \in \mathcal{A} : \eta > t\}| \leq \sum_{k, m < |\log t| + 2} |Y_{k,m}| \times |\{\eta \in I_\varepsilon : \eta > t\}|.$$

The last quantity is finite, because  $Y_{k,m}$  are finite.

*Part 2. Sufficiency.*

*Proof.* Suppose  $x \in \text{NUH}_\chi^*(f)$ , and  $\eta_n \in I_\varepsilon$  satisfy  $0 < \eta_n \leq e^{-\varepsilon/3}Q_\varepsilon(f^n(x))$  and  $e^{-\varepsilon} \leq \eta_n/\eta_{n+1} \leq e^\varepsilon$  for all  $n \in \mathbb{Z}$ .

Choose  $m_n, k_n \in \mathbb{N}$  s.t.  $\eta_n \in [e^{-m_n-1}, e^{-m_n+1}]$  and  $Q_\varepsilon(f^n(x)) \in [e^{-k_n-1}, e^{-k_n+1}]$ . Find some element of  $Y_{k_n}$  whose first coordinate is  $f^n(x)$ , and approximate it by some element of  $Y_{k_n, m_n}$  with first coordinate  $x_n$  so that for  $i = -1, 0, 1$ ,

- (A<sub>n</sub>)  $d(f^i(f^n(x)), f^i(x_n)) < \frac{1}{2}\varepsilon(\mathcal{D})$ ;
- (B<sub>n</sub>)  $d(f^i(f^n(x)), f^i(x_n)) + \|\Theta_D \circ C_\chi(f^i(f^n(x))) - \Theta_D \circ C_\chi(f^i(x_n))\| < e^{-8(m_n+2)}$  for every  $D \in \mathcal{D}$  which contains  $f^i(f^n(x)), f^i(x_n)$ ;
- (C<sub>n</sub>)  $e^{-\varepsilon/3} < Q_\varepsilon(f^i(f^n(x)))/Q_\varepsilon(f^i(x_n)) < e^{\varepsilon/3}$ .

*Claim 1.*  $\Psi_{x_n}^{\eta_n} \in \mathcal{A}$  and  $\Psi_{x_n}^{\eta'_n} \in \mathcal{A}$  for all  $\eta'_n \in I_\varepsilon$  s.t.  $\eta_n \leq \eta'_n \leq \min\{e^\varepsilon\eta_n, Q_\varepsilon(x_n)\}$ .

*Proof.* By construction  $x_n$  is the first coordinate of an element of  $Y_{k_n, m_n}$ , and  $\eta_n \in [e^{-m_n-1}, e^{m_n+1}]$ . Since  $\eta_n \leq \eta'_n \leq e^\varepsilon\eta_n$ ,  $\eta'_n \in [e^{-m_n-2}, e^{m_n+2}]$ . It remains to check that  $\eta_n, \eta'_n \leq Q_\varepsilon(x_n)$ . In case of  $\eta'_n$  there is nothing to check. In case of  $\eta_n$ , (C<sub>n</sub>) with  $i = 0$  says that  $Q_\varepsilon(x_n) > e^{-\varepsilon/3}Q_\varepsilon(f^n(x)) \geq \eta_n$ .

*Claim 2.*  $\Psi_{x_n}^{\eta_n}$  and  $\Psi_{f^n(x)}^{\eta_n}$   $\varepsilon$ -overlap.

*Proof.* (A<sub>n</sub>) with  $i = 0$  says that  $d(f^n(x), x_n)$  is smaller than the Lebesgue number of  $\mathcal{D}$ , so there exists  $D \in \mathcal{D}$  s.t.  $f^n(x), x_n \in D$ . (B<sub>n</sub>) with  $i = 0$  says that

$$d(f^n(x), x_n) + \|\Theta_D \circ C_\chi(f^n(x)) - \Theta_D \circ C_\chi(x_n)\| < e^{-8(m_n+2)}.$$

Since  $\eta_n \in [e^{-(m_n+1)}, e^{-(m_n-1)}]$ ,  $e^{-8(m_n+2)} < \eta_n^4 \eta_{n+1}^4$ . Since  $e^{-\varepsilon} \leq \eta_{n+1}/\eta_n \leq e^\varepsilon$ ,  $\Psi_{x_n}^{\eta_n}, \Psi_{f^n(x)}^{\eta_n}$   $\varepsilon$ -overlap.

*Claim 3.*  $\Psi_{f^i(x_n)}^{\eta_n}$   $\varepsilon$ -overlaps  $\Psi_{x_{n+i}}^{\eta_{n+i}}$  for  $i = \pm 1$ .

*Proof.* We do the case  $i = 1$  and leave the case  $i = -1$  to the reader.

Setting  $i = 1$  in  $(A_n)$ , we see that  $d(f(x_n), f(f^n(x))) < \frac{1}{2}\varepsilon(\mathcal{D})$ . Setting  $i = 0$  in  $(A_{n+1})$ , we see that  $d(f^{n+1}(x), x_{n+1}) < \frac{1}{2}\varepsilon(\mathcal{D})$ . It follows that there exists some  $D \in \mathcal{D}$  s.t.  $f(x_n), x_{n+1}, f^{n+1}(x) \in D$ .

By  $(B_n)$  with  $i = 1$  and  $(B_{n+1})$  with  $i = 0$ ,

$$\begin{aligned} d(f(x_n), x_{n+1}) + \|\Theta_D \circ C_\chi(f(x_n)) - \Theta_D \circ C_\chi(x_{n+1})\| &\leq \\ &\leq (d(f(x_n), f(f^n(x))) + \|\Theta_D \circ C_\chi(f(x_n)) - \Theta_D \circ C_\chi(f(f^n(x)))\|) + \\ &\quad + (d(f^{n+1}(x), x_{n+1}) + \|\Theta_D \circ C_\chi(f^{n+1}(x)) - \Theta_D \circ C_\chi(x_{n+1})\|) \\ &\leq e^{-8(m_n+2)} + e^{-8(m_{n+1}+2)} \\ &< e^{-8}(\eta_n^8 + \eta_{n+1}^8) < 2e^{-8}(1 + e^{8\varepsilon})\eta_{n+1}^4\eta_n^4 < \eta_{n+1}^4\eta_n^4. \end{aligned}$$

It follows that  $\Psi_{f(x_n)}^{\eta_{n+1}}$   $\varepsilon$ -overlaps  $\Psi_{x_{n+1}}^{\eta_{n+1}}$ .  $\square$

#### 4. $\varepsilon$ -CHAINS AND AN INFINITE-TO-ONE MARKOV EXTENSION OF $f$

**4.1. Double charts and  $\varepsilon$ -chains.** Recall that  $\Psi_x^\eta$  ( $0 < \eta \leq Q_\varepsilon(x)$ ) stands for the Pesin chart  $\Psi_x : R_\eta(\underline{0}) \rightarrow M$ . An  $\varepsilon$ -double Pesin chart (or just “double chart”) is a pair  $\Psi_x^{p^u, p^s} := (\Psi_x^{p^u}, \Psi_x^{p^s})$ , where  $0 < p^u, p^s \leq Q_\varepsilon(x)$ .

**Definition 4.1.**  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$  means

- $\Psi_y^{q^u \wedge q^s}$  and  $\Psi_{f(x)}^{q^u \wedge q^s}$   $\varepsilon$ -overlap (recall that  $a \wedge b := \min\{a, b\}$ );
- $\Psi_x^{p^u \wedge p^s}$  and  $\Psi_{f^{-1}(y)}^{p^u \wedge p^s}$   $\varepsilon$ -overlap;
- $q^u = \min\{e^\varepsilon p^u, Q_\varepsilon(y)\}$  and  $p^s = \min\{e^\varepsilon q^s, Q_\varepsilon(x)\}$ .

**Definition 4.2.**  $\{\Psi_{x_i}^{p_i^u, p_i^s}\}_{i \in \mathbb{Z}}$  (resp.  $\{\Psi_{x_i}^{p_i^u, p_i^s}\}_{i \geq 0}$ ,  $\{\Psi_{x_i}^{p_i^u, p_i^s}\}_{i \leq 0}$ ) is called an  $\varepsilon$ -chain (resp. positive  $\varepsilon$ -chain, negative  $\varepsilon$ -chain), if  $\Psi_{x_i}^{p_i^u, p_i^s} \rightarrow \Psi_{x_{i+1}}^{p_{i+1}^u, p_{i+1}^s}$  for all  $i$ . We abuse terminology and drop the  $\varepsilon$  in “ $\varepsilon$ -chains”.

Let  $\mathcal{A}$  denote the countable set of Pesin charts which we have constructed in §3.3, and recall that  $I_\varepsilon = \{e^{-k\varepsilon/3} : k \in \mathbb{N}\}$ .

**Definition 4.3.**  $\mathcal{G}$  is the directed graph with vertices  $\mathcal{V}$  and edges  $\mathcal{E}$  where

- $\mathcal{V} := \{\Psi_x^{p^u, p^s} : \Psi_x^{p^u \wedge p^s} \in \mathcal{A}, p^u, p^s \in I_\varepsilon, p^u, p^s \leq Q_\varepsilon(x)\}$ ;
- $\mathcal{E} := \{(\Psi_x^{p^u, p^s}, \Psi_y^{q^u, q^s}) \in \mathcal{V} \times \mathcal{V} : \Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}\}$ .

This is a countable directed graph. Every vertex has finite degree, because of the following lemma, and Proposition 3.5(2):

**Lemma 4.4.** If  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ , then  $e^{-\varepsilon} \leq (q^u \wedge q^s)/(p^u \wedge p^s) \leq e^\varepsilon$ . Therefore for every  $\Psi_x^{p^u, p^s} \in \mathcal{V}$  there are only finitely many  $\Psi_y^{q^u, q^s} \in \mathcal{V}$  s.t.  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$  or  $\Psi_y^{q^u, q^s} \rightarrow \Psi_x^{p^u, p^s}$ .

*Proof.* Since  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ ,  $q^u = \min\{e^\varepsilon p^u, Q_\varepsilon(y)\}$ ,  $p^s = \min\{e^\varepsilon q^s, Q_\varepsilon(x)\}$ ,  $q^s \leq Q_\varepsilon(y)$ , and  $p^u \leq Q_\varepsilon(x)$ . It follows that

$$\frac{q^u \wedge q^s}{p^u \wedge p^s} = \frac{\min\{e^\varepsilon p^u, Q_\varepsilon(y), q^s\}}{\min\{p^u, e^\varepsilon q^s, Q_\varepsilon(x)\}} = \frac{\min\{e^\varepsilon p^u, q^s\}}{\min\{p^u, e^\varepsilon q^s\}}.$$

So  $\frac{q^u \wedge q^s}{p^u \wedge p^s} \leq \frac{\min\{e^\varepsilon p^u, e^{2\varepsilon} q^s\}}{\min\{p^u, e^\varepsilon q^s\}} = e^\varepsilon$ , and  $\frac{q^u \wedge q^s}{p^u \wedge p^s} \geq \frac{\min\{e^\varepsilon p^u, q^s\}}{\min\{e^{2\varepsilon} p^u, e^\varepsilon q^s\}} = e^{-\varepsilon}$ .  $\square$

We establish a connection between the collection of infinite admissible paths on  $\mathcal{G}$  and the set of orbits of  $f$  in  $\text{NUH}_\chi^\#(f)$ . Note that “most” orbits lie in  $\text{NUH}_\chi^\#(f)$ : this set has full measure w.r.t. every  $f$ -ergodic invariant probability measure with entropy greater than  $\chi$ .

**Proposition 4.5.** *For every  $x \in \text{NUH}_\chi^\#(f)$ , there is a chain  $\{\Psi_{x_k}^{p_k^u, p_k^s}\}_{k \in \mathbb{Z}} \subset \Sigma(\mathcal{G})$  s.t.  $\Psi_{x_k}^{p_k^u, p_k^s} \varepsilon$ -overlaps  $\Psi_{f^k(x)}^{p_k^u, p_k^s}$  for all  $k \in \mathbb{Z}$ .*

The proof relies on two simple properties of chains, which we now describe.

Some terminology: Let  $(Q_k)_{k \in \mathbb{Z}}$  be a sequence in  $I_\varepsilon = \{e^{-\ell\varepsilon/3} : \ell \in \mathbb{N}\}$ . A sequence of pairs  $\{(p_k^u, p_k^s)\}_{k \in \mathbb{Z}}$  is called  $\varepsilon$ -subordinated to  $(Q_k)_{k \in \mathbb{Z}}$  if for every  $k \in \mathbb{Z}$ ,  $0 < p_k^u, p_k^s \leq Q_k$ ,  $p_k^u, p_k^s \in I_\varepsilon$ , and

$$p_{k+1}^u = \min\{e^\varepsilon p_k^u, Q_{k+1}\} \text{ and } p_{k-1}^s = \min\{e^\varepsilon p_k^s, Q_{k-1}\}.$$

For example, if  $\{\Psi_{x_k}^{p_k^u, p_k^s}\}_{k \in \mathbb{Z}}$  is a chain, then  $\{(p_k^u, p_k^s)\}_{k \in \mathbb{Z}}$  is  $\varepsilon$ -subordinated to  $\{Q_\varepsilon(x_k)\}_{k \in \mathbb{Z}}$ .

**Lemma 4.6.** *Let  $(Q_k)_{k \in \mathbb{Z}}$  be a sequence in  $I_\varepsilon$ , and suppose  $q_k \in I_\varepsilon$  satisfy  $0 < q_k \leq Q_k$  and  $e^{-\varepsilon} \leq q_k/q_{k+1} \leq e^\varepsilon$  for all  $k \in \mathbb{Z}$ . There exists a sequence  $\{(p_k^u, p_k^s)\}_{k \in \mathbb{Z}}$  which is  $\varepsilon$ -subordinated to  $\{Q_k\}_{k \in \mathbb{Z}}$ , and so that  $p_k^u \wedge p_k^s \geq q_k$  for all  $k$ .*

*Proof.* The following short proof was shown to me by F. Ledrappier. By the assumptions on  $q_k$ ,  $Q_\varepsilon(x_{k-n}), Q_\varepsilon(x_{k+n}) \geq e^{-\varepsilon n} q_k$  for all  $n \geq 0$ , therefore the following definitions make sense:

$$\begin{aligned} p_k^u &:= \max\{t \in I_\varepsilon : e^{-\varepsilon n} t \leq Q_\varepsilon(x_{k-n}) \text{ for all } n \geq 0\}; \\ p_k^s &:= \max\{t \in I_\varepsilon : e^{-\varepsilon n} t \leq Q_\varepsilon(x_{k+n}) \text{ for all } n \geq 0\}. \end{aligned}$$

The sequence  $\{(p_k^u, p_k^s)\}_{k \in \mathbb{Z}}$  is  $\varepsilon$ -subordinated to  $\{Q_\varepsilon(x_k)\}_{k \in \mathbb{Z}}$ .  $\square$

**Lemma 4.7.** *Suppose  $\{(p_n^u, p_n^s)\}_{n \in \mathbb{Z}}$  is  $\varepsilon$ -subordinated to a sequence  $\{Q_n\}_{n \in \mathbb{Z}} \subset I_\varepsilon$ . If  $\limsup_{n \rightarrow \infty} (p_n^u \wedge p_n^s) > 0$  and  $\limsup_{n \rightarrow -\infty} (p_n^u \wedge p_n^s) > 0$ , then  $p_n^u$  (resp.  $p_n^s$ ) is equal to  $Q_n$  for infinitely many  $n > 0$ , and for infinitely many  $n < 0$ .*

*Proof.* We prove the statement for  $p_n^u$ , and leave the statement for  $p_n^s$  to the reader.

$M := \sup Q_n$  is finite, because  $Q_n \in I_\varepsilon$  for all  $n$ . Let  $p_n := p_n^u \wedge p_n^s$ , and define  $m := \frac{1}{2} \min\{\limsup_{n \rightarrow \infty} p_{-n}, \limsup_{n \rightarrow \infty} p_n\}$  and  $N := \lceil \varepsilon^{-1} \log(M/m) \rceil$ .

There exists infinitely many positive (resp. negative)  $n$  s.t.  $p_n > m$ . We claim that for every such  $n$ , there must exist some  $k \in [n, n+N]$  s.t.  $p_k^u = Q_k$ . Otherwise, by  $\varepsilon$ -subordination,

$$p_{n+N}^u = \min\{Q_{n+N}, e^\varepsilon p_{n+N-1}^u\} = e^\varepsilon p_{n+N-1}^u = \cdots = e^{N\varepsilon} p_n^u \geq e^{N\varepsilon} p_n > e^{N\varepsilon} m > M,$$

which is false.  $\square$

We can now prove Proposition 4.5: Suppose  $x \in \text{NUH}_\chi^\#(f)$ , and recall the definition of  $q_\varepsilon(\cdot)$  from Lemma 2.9. Choose  $q_n \in I_\varepsilon \cap [e^{-\varepsilon/3} q_\varepsilon(f^n(x)), e^{\varepsilon/3} q_\varepsilon(f^n(x))]$ . The sequence  $\{q_n\}_{n \in \mathbb{Z}}$  satisfies the assumptions of Lemma 4.6, therefore there exists a sequence  $\{(q_n^u, q_n^s)\}_{n \in \mathbb{Z}}$  that is  $\varepsilon$ -subordinated to  $\{e^{-\varepsilon/3} Q_\varepsilon(f^n(x))\}_{n \in \mathbb{Z}}$  and that satisfies  $q_k^u \wedge q_k^s \geq q_k$ .

Let  $\eta_n := q_n^u \wedge q_n^s$ . As the proof of Lemma 4.4 shows,  $e^{-\varepsilon} \leq \eta_{n+1}/\eta_n \leq e^\varepsilon$ , so we may use Proposition 3.5 to construct an infinite sequence  $\Psi_{x_n}^{\eta_n} \in \mathcal{A}$  such that

- (a)  $\Psi_{x_n}^{\eta_n}$   $\varepsilon$ -overlaps  $\Psi_{f^n(x)}^{\eta_n}$  and  $e^{-\varepsilon/3} \leq Q_\varepsilon(f^n(x))/Q_\varepsilon(x_n) \leq e^{\varepsilon/3}$ ;
- (b)  $\Psi_{f(x_n)}^{\eta_{n+1}}$   $\varepsilon$ -overlaps  $\Psi_{x_{n+1}}^{\eta_{n+1}}$ ;
- (c)  $\Psi_{f^{-1}(x_n)}^{\eta_{n-1}}$   $\varepsilon$ -overlaps  $\Psi_{x_{n-1}}^{\eta_{n-1}}$ ;
- (d)  $\Psi_{x_n}^{\eta'_n} \in \mathcal{A}$  for all  $\eta'_n \in I_\varepsilon$  s.t.  $\eta_n \leq \eta'_n \leq \min\{Q_\varepsilon(x_n), e^\varepsilon \eta_n\}$ .

Construct a sequence  $\{(p_n^u, p_n^s)\}_{n \in \mathbb{Z}}$  which is  $\varepsilon$ -subordinated to  $\{Q_\varepsilon(x_n)\}_{n \in \mathbb{Z}}$  and which satisfies  $p_n^u \wedge p_n^s \geq \eta_n$ .

*Claim 1.*  $\Psi_{x_n}^{p_n^u, p_n^s} \in \mathcal{V}$  for all  $n$ .

*Proof.* It is sufficient to show that  $1 \leq \frac{p_n^u \wedge p_n^s}{q_n^u \wedge q_n^s} \leq e^\varepsilon$  ( $n \in \mathbb{Z}$ ), because property (d) with  $\eta'_n := p_n^u \wedge p_n^s$  says that in this case  $\Psi_{x_n}^{p_n^u \wedge p_n^s} \in \mathcal{A}$ , whence  $\Psi_{x_n}^{p_n^u, p_n^s} \in \mathcal{V}$ .

We start by showing that there are infinitely many  $n < 0$  such that  $p_n^u \leq e^\varepsilon q_n^u$ . Since  $x \in \text{NUH}_\chi^\#(f)$ ,  $\limsup_{n \rightarrow \infty} q_n, \limsup_{n \rightarrow -\infty} q_n > 0$ . Therefore by Lemma 4.7, there are infinitely many  $n < 0$  for which  $q_n^u = e^{-\varepsilon/3} Q_\varepsilon(f^n(x))$ . Property (a) guarantees that for such  $n$ ,  $q_n^u > e^{-\varepsilon} Q_\varepsilon(x_n) \geq e^{-\varepsilon} p_n^u$ , whence  $p_n^u < e^\varepsilon q_n^u$ .

If  $p_n^u \leq e^\varepsilon q_n^u$ , then  $p_{n+1}^u \leq e^\varepsilon q_{n+1}^u$ , because

$$\begin{aligned} p_{n+1}^u &= \min\{e^\varepsilon p_n^u, Q_\varepsilon(x_{n+1})\} = e^\varepsilon \min\{p_n^u, e^{-\varepsilon} Q_\varepsilon(x_{n+1})\} \\ &\leq e^\varepsilon \min\{e^\varepsilon q_n^u, e^{-\varepsilon/3} Q_\varepsilon(f^{n+1}(x))\} \equiv e^\varepsilon q_{n+1}^u. \end{aligned}$$

It follows that  $p_n^u \leq e^\varepsilon q_n^u$  for all  $n \in \mathbb{Z}$ .

Working with positive  $n$ , one can show in the same manner that  $p_n^s \leq e^\varepsilon q_n^s$  for all  $n \in \mathbb{Z}$ . Combining the two results we see that  $p_n^u \wedge p_n^s \leq (e^\varepsilon q_n^u) \wedge (e^\varepsilon q_n^s) = e^\varepsilon (q_n^u \wedge q_n^s)$  for all  $n \in \mathbb{Z}$ . Since by construction  $p_n^u \wedge p_n^s \geq \eta_n = q_n^u \wedge q_n^s$ , we obtain  $1 \leq \frac{q_n^u \wedge q_n^s}{p_n^u \wedge p_n^s} \leq e^\varepsilon$  as needed.

*Claim 2.* For every  $n \in \mathbb{Z}$ ,  $\Psi_{x_n}^{p_n^u, p_n^s} \rightarrow \Psi_{x_{n+1}}^{p_{n+1}^u, p_{n+1}^s}$ , and  $\Psi_{x_n}^{p_n^u \wedge p_n^s}$   $\varepsilon$ -overlaps  $\Psi_{f^n(x)}^{p_n^u \wedge p_n^s}$ .

*Proof.* This follows from properties (a), (b), and (c) above, the inequality  $p_n^u \wedge p_n^s \geq \eta_n$ , and the monotonicity property of the overlap condition.  $\square$

**4.2. Admissible manifolds and the graph transform.** Suppose  $x \in \text{NUH}_\chi(f)$ . A  $u$ -manifold in  $\Psi_x$  is a manifold  $V^u \subset M$  of the form

$$V^u = \Psi_x\{(F^u(t), t) : |t| \leq q\},$$

where  $0 < q \leq Q_\varepsilon(x)$ , and  $F^u$  is a  $C^{1+\beta/3}$ -function s.t.  $\|F^u\|_\infty \leq Q_\varepsilon(x)$ .

An  $s$ -manifold in  $\Psi_x$  is a manifold  $V^s \subset M$  of the form

$$V^s = \Psi_x\{(t, F^s(t)) : |t| \leq q\},$$

where  $0 < q \leq Q_\varepsilon(x)$ , and  $F^s$  is a  $C^{1+\beta/3}$ -function s.t.  $\|F^s\|_\infty \leq Q_\varepsilon(x)$ .

We will use the superscript “ $u/s$ ” in statements which apply both to the  $s$  case and to the  $u$  case. The function  $F = F^{u/s}$  is called the *representing function* of  $V^{u/s}$  at  $\Psi_x$ . The *parameters* of a  $u/s$  manifold in  $\Psi_x$  are

- $\sigma$ -parameter:  $\sigma(V^{u/s}) := \|F'\|_{\beta/3} := \|F'\|_\infty + \text{Höl}_{\beta/3}(F')$ , where  $\text{Höl}_{\beta/3}(F') := \sup \left\{ \frac{|F'(t_1) - F'(t_2)|}{|t_1 - t_2|^{\beta/3}} \right\}$ ;
- $\gamma$ -parameter:  $\gamma(V^{u/s}) := |F'(0)|$ ;
- $\varphi$ -parameter:  $\varphi(V^{u/s}) := |F(0)|$ ;
- $q$ -parameter:  $q(V^{u/s}) := q$ .

A  $(u/s, \sigma, \gamma, \varphi, q)$ -manifold in  $\Psi_x$  is a  $u/s$ -manifold  $V^{u/s}$  in  $\Psi_x$  whose parameters satisfy  $\sigma(V^{u/s}) \leq \sigma$ ,  $\gamma(V^{u/s}) \leq \gamma$ ,  $\varphi(V^{u/s}) \leq \varphi$ , and  $q^{u/s}(V^u) = q$ .

**Definition 4.8.** Suppose  $\Psi_x^{p^u, p^s}$  is a double chart. A  $u/s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  is a  $(u/s, \sigma, \gamma, \varphi, q)$ -manifold in  $\Psi_x$  s.t.

$$\sigma \leq \frac{1}{2}, \quad \gamma \leq \frac{1}{2}(p^u \wedge p^s)^{\beta/3}, \quad \varphi \leq 10^{-3}(p^u \wedge p^s), \quad \text{and } q = \begin{cases} p^u & u\text{-manifolds} \\ p^s & s\text{-manifolds.} \end{cases}$$

This is similar to but stronger than the admissibility condition in Katok & Mendoza [KM, Definition S.3.4] or Katok [K1]. The bounds on  $\gamma$  and  $\varphi$  are distortion compensating bounds: the larger the distortion of the chart, the closer the  $u/s$ -admissible manifolds are to the  $u/s$ -axes. These bounds were designed to be sufficiently strong to imply Proposition 4.11 (4), but also sufficiently lax to remain invariant under the graph transform (Proposition 4.12 below).

Let  $F$  be the representing function of a  $u/s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$ . If  $\varepsilon < 1$  (as we always assume), then the conditions  $\sigma \leq \frac{1}{2}$ ,  $\varphi < 10^{-3}(p^u \wedge p^s)$  and  $p^u, p^s < Q_\varepsilon(x)$  force

$$\text{Lip}(F) < \varepsilon, \tag{4.1}$$

because for every  $t$  in the domain of  $F$ ,  $|t| \leq p^{u/s} \leq Q_\varepsilon(x) < \varepsilon^{3/\beta}$  and

$$|F'(t)| \leq |F'(0)| + \text{H\"ol}(F')|t|^{\frac{\beta}{3}} \leq \frac{1}{2}(p^u \wedge p^s)^{\frac{\beta}{3}} + \frac{1}{2}(p^{u/s})^{\frac{\beta}{3}} < (p^{u/s})^{\frac{\beta}{3}} < \varepsilon. \tag{4.2}$$

Another important fact is that if  $\varepsilon$  is small enough then  $\|F\|_\infty < 10^{-2}Q_\varepsilon(x)$ , because  $\|F\|_\infty \leq |F(0)| + \max |F'| \cdot p^{u/s} < \varphi + \varepsilon p^{u/s} \leq (10^{-3} + \varepsilon)p^{u/s} < 10^{-2}p^{u/s}$ .

**Definition 4.9.** Let  $V_1, V_2$  be two  $u$ -manifolds (resp.  $s$ -manifolds) in  $\Psi_x$  s.t.  $q(V_1) = q(V_2)$ , then  $\text{dist}(V_1, V_2) := \max |F_1 - F_2|$  where  $F_1$  and  $F_2$  are the representing functions of  $V_1$  and  $V_2$  in  $\Psi_x$ .

Occasionally we will also need the  $C^1$ -distance defined by

$$\text{dist}_{C^1}(V_1, V_2) := \max |F_1 - F_2| + \max |F'_1 - F'_2|.$$

Notice that  $\text{dist}$  and  $\text{dist}_{C^1}$  are defined using the Pesin charts, not the Riemannian metric. Riemannian distances are bounded by a constant times distances w.r.t. Pesin charts, because Pesin charts take the form  $\Psi_x = \exp_x \circ C_\chi(x)$  where  $C_\chi(x) : \mathbb{R}^2 \rightarrow M$  is a contraction.

**Definition 4.10.** Let  $V^s, V^u$  be a  $u$ -manifold and an  $s$ -manifold in  $\Psi_x$ , with representing functions  $F_s, F_u$ . Suppose  $V^s, V^u$  intersect at a unique point  $P = \Psi_x(u, v)$ , then  $\angle(V^s, V^u) := \angle((d\Psi_x)_{(u,v)}(F'_s(u)), (d\Psi_x)_{(u,v)}(F'_u(v)))$ .

*Remark:* Pesin charts preserve orientation, therefore there are only two possible choices to the pair of directions of  $V^s, V^u$  at  $P$ . Both lead to the same angle, and this angle is in  $(0, \pi)$ . Thus the angle of intersection is independent of the chart.

**Proposition 4.11.** The following holds for all  $\varepsilon$  small enough. Let  $V^u$  be a  $u$ -admissible manifold in  $\Psi_x^{p^u, p^s}$ , and  $V^s$  be an  $s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$ , then

- (1)  $V^u$  intersects  $V^s$  at a unique point  $P$ ;
- (2)  $P = \Psi_x(v, w)$  with  $|v|, |w| \leq 10^{-2}(p^u \wedge p^s)$ ;
- (3)  $P$  is a Lipschitz function of  $(V^u, V^s)$ , with Lipschitz constant less than 3;

(4) Suppose  $\eta := p^u \wedge p^s$ , then the angle of intersection at  $P$  satisfies

$$e^{-\eta^{\beta/4}} \leq \frac{\sin \angle(V^u, V^s)}{\sin \angle(E^u(x), E^s(x))} \leq e^{\eta^{\beta/4}}$$

$$|\cos \angle(V^u, V^s) - \cos \angle(E^u(x), E^s(x))| < 2\eta^{\beta/4}.$$

Parts (1),(2), and (3) follow from [KH, Corollary S.3.8]. Part (4) is a distortion compensating bound, which will be used in the proof of Proposition 6.5 below. It follows from the assumptions we made on  $\gamma$  and  $\sigma$ , and is the reason why we require more from admissible manifolds than Katok & Mendoza did in [KM]. See the appendix for proofs.

The following result describes the action of  $f$  on admissible manifolds. Results of this type (often called “graph transform” lemmas) are used to prove Pesin’s stable manifold theorem [BP, chapter 7], [P]. The version below says that the graph transform preserves admissibility as defined above. The proof is in the appendix.

**Proposition 4.12** (Graph Transform). *The following holds for all  $\varepsilon$  small enough. Suppose  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ , and  $V^u$  is a  $u$ -admissible manifold in  $\Psi_x^{p^u, p^s}$ , then*

(1)  $f(V^u)$  contains a  $u$ -manifold  $\widehat{V}^u$  in  $\Psi_y^{q^u, q^s}$  with parameters

$$\begin{aligned} \sigma(\widehat{V}^u) &\leq e^{\sqrt{\varepsilon}} e^{-2\chi} [\sigma(V^u) + \sqrt{\varepsilon}] \\ \gamma(\widehat{V}^u) &\leq e^{\sqrt{\varepsilon}} e^{-2\chi} [\gamma(V^u) + \varepsilon^{\beta/3} (q^u \wedge q^s)^{\beta/3}] \\ \varphi(\widehat{V}^u) &\leq e^{\sqrt{\varepsilon}} e^{-\chi} [\varphi + \sqrt{\varepsilon} (q^u \wedge q^s)] \\ q(\widehat{V}^u) &\geq \min\{e^{-\sqrt{\varepsilon}} e^{\chi} q(V^u), Q_\varepsilon(y)\} \end{aligned} \tag{4.3}$$

- (2)  $f(V^u)$  intersects any  $s$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  at a unique point.  
(3)  $\widehat{V}^u$  restricts to a  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ . This is the unique  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  inside  $f(V^u)$ . We call it  $\mathcal{F}_u[V^u]$ .  
(4) Suppose  $V^u$  is represented by the function  $F$ . If  $p := \Psi_x(F(0), 0)$ , then  $f(p) \in \mathcal{F}_u[V^u]$ .

Similar statements hold for the  $f^{-1}$ -image of an  $s$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ .

**Definition 4.13.** Suppose  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ . The graph transforms are the maps

- $\mathcal{F}_u$  which maps a  $u$ -admissible manifold  $V^u$  in  $\Psi_x^{p^u, p^s}$  to the unique  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  contained in  $f(V^u)$ ;
- $\mathcal{F}_s$  which maps an  $s$ -admissible manifold  $V^s$  in  $\Psi_y^{q^u, q^s}$  to the unique  $s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  contained in  $f^{-1}(V^s)$ .

The operators  $\mathcal{F}_s, \mathcal{F}_u$  depend on the edge  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ .

**Proposition 4.14.** *If  $\varepsilon$  is small enough then the following holds. Let  $t = s, u$ , then for any  $t$ -admissible manifolds  $V_1^t, V_2^t$  in  $\Psi_x^{p^u, p^s}$ ,*

$$\text{dist}(\mathcal{F}_t(V_1^t), \mathcal{F}_t(V_2^t)) \leq e^{-\chi/2} \text{dist}(V_1^t, V_2^t); \tag{4.4}$$

$$\text{dist}_{C^1}(\mathcal{F}_t(V_1^t), \mathcal{F}_t(V_2^t)) \leq e^{-\chi/2} [\text{dist}_{C^1}(V_1^t, V_2^t) + (\text{dist}(V_1^t, V_2^t))^{\beta/3}]. \tag{4.5}$$

See [BP, chapter 7], [KM], and the appendix.

**4.3. A Markov extension.** Let  $\Sigma := \Sigma(\mathcal{G})$  denote the topological Markov shift of two sided infinite paths on the graph  $G(\mathcal{V}, \mathcal{E})$ :

$$\Sigma := \{(v_i)_{i \in \mathbb{Z}} : v_i \in \mathcal{V}, v_i \rightarrow v_{i+1} \text{ for all } i\}.$$

We equip  $\Sigma$  with the metric  $d(\underline{v}, \underline{w}) = \exp[-\min\{k : v_k \neq w_k\}]$ , and the action of the left shift map  $\sigma : \Sigma \rightarrow \Sigma$ ,  $\sigma : (v_i)_{i \in \mathbb{Z}} \mapsto (v_{i+1})_{i \in \mathbb{Z}}$ .

Our aim is to construct a map  $\pi : \Sigma \rightarrow M$  with a  $\chi$ -large image s.t.  $\pi \circ \sigma = f \circ \pi$ . In fact, the map we construct will be well-defined for all chains.

We begin with some comments on general chains of double charts. Suppose  $(v_i)_{i \in \mathbb{Z}}$ ,  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$  is a chain, and let  $V_{-n}^u$  be a  $u$ -admissible manifold in  $v_{-n}$ . The graph transform relative to  $v_{-n} \rightarrow v_{-n+1}$  maps  $V_{-n}^u$  to a  $u$ -admissible manifold in  $v_{-n+1}$ ,  $\mathcal{F}_u[V_{-n}]$ . Another application of the graph transform, this time relative to  $v_{-n+1} \rightarrow v_{-n+2}$ , maps  $\mathcal{F}_u[V_{-n}]$  to a  $u$ -admissible manifold in  $v_{-n+2}$ , which we denote by  $\mathcal{F}_u^2[V_{-n}]$ . Continuing this way, we eventually reach a  $u$ -admissible manifold in  $v_0$  which we denote by  $\mathcal{F}_u^n[V_{-n}^u]$ . Similarly, any  $s$ -admissible manifold in  $v_n$  is mapped by  $n$  applications of  $\mathcal{F}_s$  to an  $s$ -admissible manifold in  $v_0$ . The manifolds  $\mathcal{F}_u^n[V_{-n}^u]$  and  $\mathcal{F}_s^n[V_n^s]$  depend on  $(v_{-n}, \dots, v_n)$ .

Let  $V_n$  denote a sequence of  $u/s$ -manifolds in a chart  $\Psi_x$ . We say that  $V_n$  converges to a  $u/s$ -manifold  $V$ , if the representing functions of  $V_n$  converge uniformly to the representing function of  $V$ .

**Proposition 4.15.** *Suppose  $(v_i)_{i \in \mathbb{Z}}$  is a chain of double charts, and choose arbitrary  $u$ -admissible manifolds  $V_{-n}^u$  in  $v_{-n}$ , and  $s$ -admissible manifolds  $V_n^s$  in  $v_n$ .*

- (1) *The limits  $V^u[(v_i)_{i \leq 0}] := \lim_{n \rightarrow \infty} \mathcal{F}_u^n[V_{-n}^u]$ , and  $V^s[(v_i)_{i \geq 0}] := \lim_{n \rightarrow \infty} \mathcal{F}_s^n[V_n^s]$  exist, and are independent of the choice of  $V_{-n}^u$  and  $V_n^s$ .*
- (2)  *$V^u[(v_i)_{i \leq 0}]$  is a  $u$ -admissible manifold in  $v_0$ , and  $V^s[(v_i)_{i \geq 0}]$  is an  $s$ -admissible manifold in  $v_0$ .*
- (3)  *$f(V^s[(v_i)_{i \geq 0}]) \subset V^s[(v_{i+1})_{i \geq 0}]$  and  $f^{-1}(V^u[(v_i)_{i \leq 0}]) \subset V^u[(v_{i-1})_{i \leq 0}]$ ;*
- (4) *Write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$ , then*

$$V^s[(v_i)_{i \geq 0}] = \{p \in \Psi_{x_0}[R_{p_0^s}(\underline{0})] : \forall k \geq 0, f^k(p) \in \Psi_{x_k}[R_{10Q_\varepsilon(x_k)}(\underline{0})]\},$$

$$V^u[(v_i)_{i \leq 0}] = \{p \in \Psi_{x_0}[R_{p_0^u}(\underline{0})] : \forall k \geq 0, f^{-k}(p) \in \Psi_{x_{-k}}[R_{10Q_\varepsilon(x_{-k})}(\underline{0})]\}.$$

- (5) *The maps  $(u_i)_{i \in \mathbb{Z}} \mapsto V^u[(u_i)_{i \leq 0}], V^s[(u_i)_{i \geq 0}]$  are Hölder continuous: there exist constants  $K > 0$  and  $0 < \theta < 1$  s.t. for every  $n \geq 0$  and any two chains  $\underline{u}, \underline{v}$ , if  $u_i = v_i$  for all  $|i| \leq n$ , then*

$$\text{dist}_{C^1}(V^u[(u_i)_{i \leq 0}], V^u[(v_i)_{i \leq 0}]) < K\theta^n;$$

$$\text{dist}_{C^1}(V^s[(u_i)_{i \geq 0}], V^s[(v_i)_{i \geq 0}]) < K\theta^n.$$

Parts (1)–(4) are a version of Pesin’s Stable Manifold Theorem [P]. The new twist is that Proposition 4.15 generates local stable manifolds with a definite choice of size, whereas Pesin’s theorem speaks of a germ of local stable manifolds at a point. In §8.1 we’ll see that for many chains, this size is “almost maximal” and therefore “almost canonical”. This will be instrumental to the proof of local finiteness.

Part (5) should be compared to Brin’s Theorem on the Hölder continuity of the Oseledets distribution [Bri]. Whereas Brin’s theorem only states Hölder continuity on Pesin sets, part (5) gives Hölder continuity everywhere. The secret behind this “improvement” is the difference between the metric in the symbolic space and the Riemannian metric of the manifold.

*Proof.* We give the proof in the case of  $u$ -manifolds. The case of  $s$ -manifolds is symmetric. Before we begin, we mention the following obvious fact: for any double chart  $\Psi_x^{p^u, p^s}$  and any two  $u$ -manifolds  $V_1^u, V_2^u$  in  $\Psi_x^{p^u, p^s}$ ,

$$\text{dist}(V_1^u, V_2^u) \leq 2Q_\varepsilon(x) < 1.$$

*Part 1.* Existence of the limit.

By Proposition 4.12,  $\mathcal{F}_u^n[V_{-n}^u]$  is a  $u$ -admissible manifold in  $v_0$ . By Proposition 4.14, for any other choice of  $u$ -admissible manifolds  $W_{-n}^u$  in  $v_{-n}$ ,

$$\text{dist}(\mathcal{F}_u^n[V_{-n}^u], \mathcal{F}_u^n[W_{-n}^u]) < \exp[-\frac{1}{2}\chi n] \text{dist}(V_{-n}^u, W_{-n}^u) < \exp[-\frac{1}{2}\chi n].$$

Thus, if the limit exists then it is independent of  $V_{-n}^u$ .

For every  $m > n$ ,  $W_{-n}^u := \mathcal{F}_u^{m-n}[V_{-m}^u]$  is a  $u$ -admissible manifold in  $v_{-n}$ . It follows that for every  $m > n$ ,  $\text{dist}(\mathcal{F}_u^n[V_{-n}^u], \mathcal{F}_u^m[V_{-m}^u]) < \exp[-\frac{1}{2}\chi n]$ . It follows that  $\lim \mathcal{F}_u^n[V_{-n}^u]$  exists.

*Part 2.* Admissibility of the limit.

Write  $v_0 = \Psi_x^{p^u, p^s}$ , and let  $F_n$  denote the functions which represent  $\mathcal{F}_u^n[V_{-n}^u]$  in  $v_0$ . Since  $\mathcal{F}_u^n[V_{-n}^u]$  are  $u$ -admissible in  $v_0$ , for every  $n$ ,

- $\|F'_n\|_{\beta/3} \leq \frac{1}{2}$ ;
- $\|F'_n(0)\| \leq \frac{1}{2}(p^u \wedge p^s)^{\beta/3}$ ;
- $|F_n(0)| \leq 10^{-3}(p^u \wedge p^s)$ .

Since  $\mathcal{F}_u^n[V_{-n}^u] \xrightarrow{n \rightarrow \infty} V^u[(v_i)_{i \leq 0}]$ ,  $F_n \xrightarrow{n \rightarrow \infty} F$  uniformly on  $[-p^u, p^u]$ , where  $F$  represents  $V^u[(v_i)_{i \leq 0}]$ .

By the Arzela–Ascoli Theorem,  $\exists n_k \uparrow \infty$  s.t.  $F'_{n_k} \xrightarrow{k \rightarrow \infty} G$  uniformly, where  $\|G\|_{\beta/3} \leq \frac{1}{2}$ . Thus  $F_{n_k}(t) = F_{n_k}(-p^u) + \int_{-p^u}^t F'_{n_k}(t) dt \xrightarrow{k \rightarrow \infty} F(-p^u) + \int_{-p^u}^t G(t) dt$ , whence  $F$  is differentiable, and  $F' = G$ . We also see that  $\{F'_n\}$  can only have one limit point. Consequently,  $F'_n \xrightarrow{n \rightarrow \infty} F'$  uniformly.

It follows that  $\|F'\|_{\beta/3} \leq \frac{1}{2}$ ,  $|F'(0)| \leq \frac{1}{2}(p^u \wedge p^s)^{\beta/3}$ , and  $|F(0)| \leq 10^{-3}(p^u \wedge p^s)$ , whence the  $u$ -admissibility of  $V^u[(v_i)_{i \leq 0}]$ .

*Part 3.* Invariance properties of the limit.

Let  $V^u := V^u[(v_i)_{i \leq 0}] = \lim \mathcal{F}_u^n[V_{-n}^u]$ , and  $W^u := V^u[(v_{i-1})_{i \leq 0}] = \lim \mathcal{F}_u^n[V_{-n-1}^u]$ .

$$\begin{aligned} \text{dist}(V^u, \mathcal{F}_u(W^u)) &\leq \text{dist}(V^u, \mathcal{F}_u^n(V_{-n}^u)) + \text{dist}(\mathcal{F}_u^n(V_{-n}^u), \mathcal{F}_u^{n+1}(V_{-n-1}^u)) \\ &\quad + \text{dist}(\mathcal{F}_u^{n+1}(V_{-n-1}^u), \mathcal{F}_u(W^u)) \\ &\leq \text{dist}(V^u, \mathcal{F}_u^n(V_{-n}^u)) + e^{-\frac{1}{2}\chi n} \text{dist}(V_{-n}^u, \mathcal{F}_u(V_{-n-1}^u)) + e^{-\frac{1}{2}\chi n} \text{dist}(\mathcal{F}_u^n(V_{-n-1}^u), W^u). \end{aligned}$$

The first and third summands tend to zero, by the definition of  $V^u$  and  $W^u$ . The second summand tends to zero, because  $\text{dist}(V_{-n}^u, \mathcal{F}_u(V_{-n-1}^u)) < 2Q_\varepsilon(x) < 1$ . It follows that  $V^u = \mathcal{F}_u(W^u) \subset f(W^u)$ .

*Part 4.* Suppose  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$ , then

$$V^u = \{p \in \Psi_{x_0}[R_{p_0^u}(\underline{0})] : \forall k \geq 0, f^{-k}(p) \in \Psi_{x_{-k}}[R_{10Q_\varepsilon(x_{-k})}(\underline{0})]\}.$$



The inclusion  $\subseteq$  is simple: Every  $u$ -admissible manifold  $W_i^u$  in  $\Psi_{x_i}^{p_i^u, p_i^s}$  is contained in  $\Psi_{x_i}[R_{p_i^u}(\mathbb{0})]$ , because if  $W_i^u$  is represented by the function  $F$  then any  $p = \Psi_{x_i}(v, w)$  in  $W_i^u$  satisfies  $|w| \leq p_i^u$ , and

$$|v| = |F(w)| \leq |F(0)| + \max |F'| \cdot |w| \leq \varphi + \varepsilon |w| \leq (10^{-3} + \varepsilon)p_i^u < p_i^u.$$

Applying this to  $V^u := V^u[(v_i)_{i \leq 0}]$ , we see that for every  $p \in V^u$ ,  $p \in \Psi_{x_0}[R_{p_0^u}(\mathbb{0})]$ , and by part 3 for every  $k \geq 0$

$$f^{-k}(p) \in f^{-k}(V^u) \subseteq V^u[(v_{i-k})_{i \leq 0}] \subset \Psi_{x_{-k}}[R_{p_{-k}^u}(\mathbb{0})] \subset \Psi_{x_{-k}}[R_{10Q_\varepsilon(x_{-k})}(\mathbb{0})].$$

We have  $\subseteq$ .

We prove  $\supseteq$ . Suppose  $z \in \Psi_{x_0}[R_{p_0^u}(\mathbb{0})]$  and  $f^{-k}(z) \in \Psi_{x_{-k}}[R_{10Q_\varepsilon(x_{-k})}(\mathbb{0})]$  for all  $k \geq 0$ . Write  $z = \Psi_{x_0}(v_0, w_0)$ . We show that  $z \in V^u$  by proving that  $v_0 = F(w_0)$ , where  $F$  is the function which represents  $V^u$ .

Introduce for this purpose the point  $\bar{z} = \Psi_{x_0}(\bar{v}_0, \bar{w}_0)$ , where  $\bar{w}_0 = w_0$  and  $\bar{v}_0 = F(\bar{w}_0)$ . For every  $k \geq 0$ ,  $f^{-k}(z), f^{-k}(\bar{z}) \in \Psi_{x_{-k}}[R_{10Q_\varepsilon(x_{-k})}(\mathbb{0})]$ , the first point by assumption, and the second point because  $f^{-k}(\bar{z}) \in f^{-k}(V^u) \subset V^u[(v_{i-k})_{i \leq 0}]$ . It is therefore possible to write

$$f^{-k}(z) = \Psi_{x_{-k}}(v_{-k}, w_{-k}) \text{ and } f^{-k}(\bar{z}) = \Psi_{x_{-k}}(\bar{v}_{-k}, \bar{w}_{-k}) \quad (k \geq 0),$$

where  $|v_{-k}|, |w_{-k}|, |\bar{v}_{-k}|, |\bar{w}_{-k}| \leq 10Q_\varepsilon(x_{-k})$  for all  $k \geq 0$ .

Proposition 3.4, in its version for  $f^{-1}$ , says that for every  $k \geq 0$ ,  $f_{x_{-k-1}x_{-k}}^{-1} = \Psi_{x_{-k-1}}^{-1} \circ f^{-1} \circ \Psi_{x_{-k}}$  can be put in the form

$$f_{x_{-k-1}x_{-k}}^{-1}(v, w) = (A_k^{-1}v + g_1^{(k)}(v, w), B_k^{-1}w + g_2^{(k)}(v, w)),$$

where  $|A_k| < e^{-\chi/2}$ ,  $|B_k| > e^{\chi/2}$ , and  $\max_{R_{10Q_\varepsilon(x_{-k})}} \|\nabla g_i^{(k)}\| < \varepsilon$  (provided  $\varepsilon$  is small enough).

Let  $\Delta v_{-k} := v_{-k} - \bar{v}_{-k}$  and  $\Delta w_{-k} := w_{-k} - \bar{w}_{-k}$ . Since for every  $k \leq 0$ ,  $(v_{-k-1}, w_{-k-1}) = f_{x_{-k-1}x_{-k}}^{-1}(v_{-k}, w_{-k})$  and  $(\bar{v}_{-k-1}, \bar{w}_{-k-1}) = f_{x_{-k-1}x_{-k}}^{-1}(\bar{v}_{-k}, \bar{w}_{-k})$ ,

$$\begin{aligned} |\Delta v_{-k-1}| &\geq |A_k^{-1}| \cdot |\Delta v_{-k}| - \max \|\nabla g_1^{(k)}\| \cdot (|\Delta v_{-k}| + |\Delta w_{-k}|) \\ &\geq (e^{\chi/2} - \varepsilon) |\Delta v_{-k}| - \varepsilon |\Delta w_{-k}|. \\ |\Delta w_{-k-1}| &\leq |B_k^{-1}| \cdot |\Delta w_{-k}| + \max \|\nabla g_2^{(k)}\| \cdot (|\Delta v_{-k}| + |\Delta w_{-k}|) \\ &\leq (e^{-\chi/2} + \varepsilon) |\Delta w_{-k}| + \varepsilon |\Delta v_{-k}|. \end{aligned}$$

Write for short  $a_k := |\Delta v_{-k}|$  and  $b_k := |\Delta w_{-k}|$ . If we assume, as we may, that  $\varepsilon$  is so small that  $e^{-\chi/2} + \varepsilon < e^{-\chi/3}$  and  $e^{\chi/2} - \varepsilon \geq e^{\chi/3}$ , then we obtain

$$\begin{aligned} a_{k+1} &\geq e^{\chi/3} a_k - \varepsilon b_k, \\ b_{k+1} &\leq e^{-\chi/3} b_k + \varepsilon a_k. \end{aligned}$$

By definition,  $b_0 = 0$ .

Suppose  $\varepsilon$  is so small that  $e^{-\chi/3} + \varepsilon < 1$  and  $e^{\chi/3} - \varepsilon > 1$ . We claim that  $a_k \leq a_{k+1}$  and  $b_k \leq a_k$  for all  $k$ . For  $k = 0$ , this is because  $b_0 = 0$ . Assume by induction that  $a_k \leq a_{k+1}$  and  $b_k \leq a_k$ , then

$$\begin{aligned} b_{k+1} &\leq e^{-\chi/3} b_k + \varepsilon a_k \leq (e^{-\chi/3} + \varepsilon) a_k < a_k \leq a_{k+1} \\ a_{k+2} &\geq e^{\chi/3} a_{k+1} - \varepsilon b_{k+1} \geq (e^{\chi/3} - \varepsilon) a_{k+1} > a_{k+1}. \end{aligned}$$

We see that  $a_{k+1} \geq (e^{\chi/3} - \varepsilon)a_k$  for all  $k$ , whence  $a_k \geq (e^{\chi/3} - \varepsilon)^k a_0$ . Either  $a_0 = 0$  or  $a_k \xrightarrow[k \rightarrow \infty]{} \infty$ . But  $a_k = |v_{-k} - \bar{v}_{-k}| \leq 20|Q_\varepsilon(x_{-k})| < 20\varepsilon$ , so  $a_0 = 0$ . Since  $a_0 = 0$ ,  $v_0 = \bar{v}_0$ , and therefore  $F(\bar{w}_0) = F(w_0)$ . Thus  $z = \Psi_x(F(w_0), w_0) \in V^u$ .

*Part 5.* Hölder continuity of  $\underline{u} \mapsto V^u[(u_i)_{i \in \mathbb{Z}}]$ : If two chains  $\underline{v} = (v_i)_{i \in \mathbb{Z}}$ ,  $\underline{w} = (w_i)_{i \in \mathbb{Z}}$  satisfy  $v_i = w_i$  for  $i = -N, \dots, N$ , then  $\text{dist}(V^u[(v_i)_{i \leq 0}], V^u[(w_i)_{i \leq 0}]) \leq e^{-\frac{1}{2}N\chi}$ .

Given  $n > N$ , let  $V_{-n}^u$  be a  $u$ -admissible manifold in  $v_{-n}$ , and let  $W_{-n}^u$  be a  $u$ -admissible manifold in  $w_{-n}$ .

Let  $\mathcal{F}_u^\ell(V_{-n}^u)$  (resp.  $\mathcal{F}_u^\ell(W_{-n}^u)$ ) denote the result of applying  $\mathcal{F}_u$   $\ell$  times to  $V_{-n}^u$  using the path  $u_{-n} \rightarrow \dots \rightarrow u_{-n+\ell}$  (resp. using  $w_{-n} \rightarrow \dots \rightarrow w_{-n+\ell}$ ).

$\mathcal{F}_u^{n-N}(V_{-n}^u)$  and  $\mathcal{F}_u^{n-N}(W_{-n}^u)$  are  $u$ -admissible manifolds in  $v_{-N}$  ( $= w_{-N}$ ). Let  $F_N, G_N$  be their representing functions. Admissibility implies that

$$\|F_N - G_N\|_\infty \leq \|F_N\|_\infty + \|G_N\|_\infty < 2Q_\varepsilon < 1$$

$$\|F'_N - G'_N\|_\infty \leq \|F'_N\|_\infty + \|G'_N\|_\infty < 2\varepsilon < 1.$$

Represent  $\mathcal{F}_u^{n-k}[V_{-n}^u]$  and  $\mathcal{F}_u^{n-k}[W_{-n}^u]$  by functions  $F_k$  and  $G_k$ . By(4.5),

$$\|F_{k-1} - G_{k-1}\|_\infty \leq e^{-\chi/2} \|F_k - G_k\|_\infty \quad (4.6)$$

$$\|F'_{k-1} - G'_{k-1}\|_\infty \leq e^{-\chi/2} (\|F'_k - G'_k\|_\infty + 2\|F_k - G_k\|_\infty^{\beta/3}). \quad (4.7)$$

Iterating (4.6) starting at  $k = N$  and going down, we get  $\|F_k - G_k\|_\infty \leq e^{-\frac{1}{2}\chi(N-k)}$ , whence  $\text{dist}(\mathcal{F}_u^n[V_{-n}^u], \mathcal{F}_u^n[W_{-n}^u]) \leq e^{-\frac{1}{2}\chi N}$ . Passing to the limit  $n \rightarrow \infty$ , we get

$$\text{dist}(V^u[(v_i)_{i \leq 0}], V^u[(w_i)_{i \leq 0}]) \leq e^{-\frac{1}{2}N\chi}.$$

Now substitute  $\|F_k - G_k\|_\infty \leq e^{-\frac{1}{2}\chi(N-k)}$  in (4.7), and set  $c_k := \|F'_k - G'_k\|_\infty$ ,  $\theta_1 := e^{-\chi/2}$ , and  $\theta_2 := e^{-\frac{1}{6}\beta\chi}$ , then  $c_{k-1} \leq \theta_1(c_k + 2\theta_2^{N-k})$ . It is easy to see by induction that for every  $0 \leq k \leq N$ ,

$$c_0 \leq \theta_1^k c_k + 2(\theta_1^k \theta_2^{N-k} + \theta_1^{k-1} \theta_2^{N-k+1} + \dots + \theta_1 \theta_2^{N-1}).$$

We now take  $k = N$ , paying attention to the inequalities  $\theta_1 < \theta_2$  and  $c_N \leq 1$ :  $c_0 \leq \theta_1^N + 2N\theta_2^N < (2N+1)\theta_2^N$ .

It follows that  $\text{dist}_{C^1}(\mathcal{F}_u^n[V_{-n}^u], \mathcal{F}_u^n[W_{-n}^u]) \leq 2(N+1)\theta_2^N$ . In part 2, we saw that  $\mathcal{F}_u^n[V_{-n}^u]$  and  $\mathcal{F}_u^n[W_{-n}^u]$  converge to  $V^u[(v_i)_{i \leq 0}]$  in  $C^1$ . Therefore if we pass to the limit as  $n \rightarrow \infty$ , we get  $\text{dist}_{C^1}(V^u[(v_i)_{i \leq 0}], V^u[(w_i)_{i \leq 0}]) \leq 2(N+1)\theta_2^N$ . Now pick two constants  $\theta \in (\theta_2, 1)$  and  $K > 0$  s.t.  $2(N+1)\theta_2^N \leq K\theta^N$  for all  $N \geq 0$ .  $\square$

**Theorem 4.16.** *Given a chain of double charts  $(v_i)_{i \in \mathbb{Z}}$ , let  $\pi(\underline{v}) :=$  unique intersection point of  $V^u[(v_i)_{i \leq 0}]$  and  $V^s[(v_i)_{i \geq 0}]$ .*

- (1)  $\pi$  is well-defined and  $\pi \circ \sigma = f \circ \pi$ ;
- (2)  $\pi : \Sigma \rightarrow M$  is Hölder continuous;
- (3)  $\pi(\Sigma) \supset \pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f)$ , therefore  $\pi(\Sigma)$  and  $\pi(\Sigma^\#)$  have full probability w.r.t. any ergodic invariant probability measure with entropy larger than  $\chi$ .

*Proof.* Proposition 4.11 guarantees that  $\pi$  is well defined for every chain.

*Part 1.*  $\pi \circ \sigma = f \circ \pi$ .

Suppose  $\underline{v}$  is a chain, and write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$  and  $z = \pi(\underline{v})$ . We claim that

$$f^k(z) \in \Psi_{x_k} [R_{Q_\varepsilon(x_k)}(\mathbb{Q})] \quad (k \in \mathbb{Z}). \quad (4.8)$$

For  $k = 0$ , this is because  $z \in V^s[(v_i)_{i \geq 0}]$  and  $V^s[(v_i)_{i \geq 0}]$  is  $s$ -admissible in  $\Psi_{x_0}^{p_0^u, p_0^s}$ . For  $k > 0$ , we use Proposition 4.15 part (3) to see that

$$f^k(z) \in f^k(V^s[(v_i)_{i \geq 0}]) \subset V^s[(v_{i+k})_{i \geq 0}].$$

Since  $V^s[(v_{i+k})_{i \geq 0}]$  is an  $s$ -admissible manifold in  $\Psi_{x_k}^{p_k^u, p_k^s}$ ,  $f^k(z) \in \Psi_{x_k}[R_{Q_\varepsilon(x_k)}(\underline{0})]$ . The case  $k < 0$  can be handled in the same way, using  $V^u[(v_i)_{i \leq 0}]$ . Thus  $z = \pi(\underline{v})$  satisfies (4.8).

Any point which satisfies (4.8) must equal  $z$ , because by Proposition 4.15 part (4), it must lie on  $V^u[(v_i)_{i \leq 0}] \cap V^s[(v_i)_{i \geq 0}]$ . So (4.8) characterizes  $\pi(\underline{v})$ .

It is now a simple matter to deduce that  $\pi(\sigma(\underline{v})) = f(\pi(\underline{v}))$ :  $f^k[f(\pi(\underline{v}))] = f^{k+1}[\pi(\underline{v})]$  belongs to  $\Psi_{x_{k+1}}[R_{Q_\varepsilon(x_{k+1})}(\underline{0})]$  for all  $k$ , and this is the condition which characterizes  $\pi(\sigma(\underline{v}))$ .

*Part 2.*  $\pi$  is Hölder continuous.

We saw that  $\underline{u} \mapsto V^u[(u_i)_{i \leq 0}]$  and  $\underline{u} \mapsto V^s[(u_i)_{i \geq 0}]$  are Hölder continuous (Proposition 4.15). Since the intersection point of an  $s$ -admissible manifold and a  $u$  admissible manifold is a Lipschitz function of these manifolds (Proposition 4.11 (3)),  $\pi$  is also Hölder continuous.

*Part 3.*  $\pi(\Sigma)$  has full probability with respect to any ergodic invariant probability measure with entropy larger than  $\chi$ .

We prove that  $\pi(\Sigma) \supset \text{NUH}_\chi^\#(f)$ . Suppose  $x \in \text{NUH}_\chi^\#(f)$ . By Proposition 4.5, there exist  $\Psi_{x_k}^{p_k^u, p_k^s} \in \mathcal{V}$  s.t.  $\Psi_{x_k}^{p_k^u, p_k^s} \rightarrow \Psi_{x_{k+1}}^{p_{k+1}^u, p_{k+1}^s}$  for all  $k$ , and s.t.  $\Psi_{x_k}^{p_k^u, p_k^s}$   $\varepsilon$ -overlaps  $\Psi_{f^k(x)}^{p_k^u \wedge p_k^s}$  for all  $k \in \mathbb{Z}$ . By Proposition 3.2(1), this implies that

$$f^k(x) = \Psi_{f^k(x)}(\underline{0}) \in \Psi_{x_k}[R_{p_k^u \wedge p_k^s}(\underline{0})] \subset \Psi_{x_k}[R_{Q_\varepsilon(x_k)}(\underline{0})] \text{ for all } k \in \mathbb{Z}.$$

Thus  $x$  satisfies (4.8) with  $\underline{v} = (\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$ . It follows that  $z = \pi(\underline{v})$ .

In fact this argument proves something stronger, that will be of use to us later. Looking closely into the proof of Proposition 4.5, we see that the chain we constructed above satisfies the property  $p_i^u \wedge p_i^s \geq e^{-\varepsilon/3} q_\varepsilon(f^i(x))$ . By the definition of  $\text{NUH}_\chi^\#(f)$ , there exist sequences  $i_k, j_k \uparrow \infty$  for which  $p_{i_k}^u \wedge p_{i_k}^s$  and  $p_{-j_k}^u \wedge p_{-j_k}^s$  are bounded away from zero. By the discreteness property of  $\mathcal{A}$  (Proposition 3.5),  $\Psi_{x_i}^{p_i^u, p_i^s}$  must repeat some symbol infinitely often in the past, and (possibly a different symbol) in the future. Thus the above actually proves that

$$\pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f), \tag{4.9}$$

where  $\Sigma^\# := \{\underline{v} \in \Sigma : \exists v, w \in \mathcal{V}, \exists n_k, m_k \uparrow \infty \text{ s.t. } v_{n_k} = v, \text{ and } v_{-m_k} = w\}$ .  $\square$

**4.4. The relevant part of the extension.** We cannot rule out the possibility that some of the vertices in  $\mathcal{V}$  do not appear in the coding of any point in  $\text{NUH}_\chi(f)$ . Such vertices are called *irrelevant*. More precisely,

**Definition 4.17.** A double chart  $v = \Psi_x^{p^u, p^s}$  is called relevant if there exists a chain  $(v_i)_{i \in \mathbb{Z}}$  s.t.  $v_0 = v$  and  $\pi(\underline{v}) \in \text{NUH}_\chi(f)$ . A double chart which is not relevant, is called irrelevant.

**Definition 4.18.** The relevant part of  $\Sigma$  is  $\Sigma_{rel} := \{\underline{v} \in \Sigma : v_i \text{ is relevant for all } i\}$ .

$\Sigma_{rel}$  is the topological Markov shift corresponding to the restriction of the graph  $G(\mathcal{V}, \mathcal{E})$  to the relevant vertices.

**Proposition 4.19.** *Theorem 4.16 holds with  $\Sigma_{rel}$  replacing  $\Sigma$ .*

*Proof.* All the properties of  $\pi : \Sigma_{rel} \rightarrow M$  are obvious, except for the statement that  $\pi(\Sigma_{rel}^\#) \supset \text{NUH}_\chi^\#(f)$ , where  $\Sigma_{rel}^\# := \Sigma^\# \cap \Sigma_{rel}$ .

Suppose  $p \in \text{NUH}_\chi^\#(f)$ , then the proof of Theorem 4.16 shows that  $\exists \underline{v} \in \Sigma^\#$  s.t.  $\pi(\underline{v}) = p$ . Since  $\text{NUH}_\chi^\#(f)$  is  $f$ -invariant and  $f \circ \pi = \pi \circ \sigma$ ,  $\pi(\sigma^i(\underline{v})) = f^i(p) \in \text{NUH}_\chi^\#(f)$ , so  $v_i$  is relevant for all  $i \in \mathbb{Z}$ . It follows that  $\underline{v} \in \Sigma_{rel}^\#$ .  $\square$

The proposition shows that we do not need the irrelevant vertices to code a  $\chi$ -large set of orbits. *Henceforth we assume w.l.o.g. that all irrelevant vertices have been removed from  $\mathcal{V}$ , and we set  $\Sigma := \Sigma_{rel}$ .* This is needed for the proof of Proposition 7.3 below.

## Part 2. Regular chains which shadow the same orbit are close

### 5. THE INVERSE PROBLEM FOR REGULAR CHAINS

In the previous section we constructed a map  $\pi$  from the space of chains to  $M$ , and showed that every  $x \in \text{NUH}_\chi^\#(f)$  takes the form  $x = \pi(\underline{v})$  for some chain  $\underline{v} \in \Sigma^\#$ . In principle, there could be infinitely many chains  $\underline{v}$  s.t.  $\pi(\underline{v}) = x$ . We ask what one can say about the solutions  $\underline{v}$  to the equation  $\pi(\underline{v}) = x$ .

Under the additional assumption that one of the pre-images of  $x$  is regular (see below), we shall see that the coordinates  $v_i$  of  $\underline{v}$  are determined “up to bounded error”. Here is the precise statement:

**Definition 5.1.** *A chain  $(v_i)_{i \in \mathbb{Z}}$  is called regular if every  $v_i$  is relevant (see §4.4), and if there are  $v, u$  s.t. for some  $n_k, m_k \uparrow \infty$   $v_{-m_k} = u$ ,  $v_{n_k} = v$  for all  $k$ .*

Every element of  $\Sigma^\#$  is regular, because of the convention stated in §4.4.

**Theorem 5.2.** *The following holds for all  $\varepsilon$  small enough. Suppose  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$ ,  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  are regular chains s.t.  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then for all  $i$ ,*

- (1)  $d(x_i, y_i) < \varepsilon$ ;
- (2)  $(\Psi_{y_i}^{-1} \circ \Psi_{x_i})(\underline{u}) = (-1)^{\sigma_i} \underline{u} + \underline{c}_i + \Delta_i(\underline{u})$  for all  $\underline{u} \in R_\varepsilon(0)$ , where  $\sigma_i \in \{0, 1\}$ ,  $\underline{c}_i$  is a constant vector s.t.  $\|\underline{c}_i\| < 10^{-1}(q_i^u \wedge q_i^s)$ , and  $\Delta_i$  is a vector field s.t.  $\Delta_i(0) = \underline{0}$  and  $\|(d\Delta_i)_{\underline{v}}\| < \sqrt[3]{\varepsilon}$  on  $R_\varepsilon(0)$ ;
- (3)  $p_i^u/q_i^u, p_i^s/q_i^s \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$ .

The proof of Theorem 5.2 is long, so we broke it into several sections (§6,7,8). Here is an overview. Suppose  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$ ,  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  are two chains in  $\Sigma^\#$  s.t.

$$\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}] = x \quad (5.1)$$

We want to show that  $\Psi_{x_i}$  is close to  $\Psi_{y_i}$  for all  $i$ .

Equation (5.1) implies that  $f^i(x)$  is the intersection of a  $u$ -admissible and an  $s$ -admissible manifold in  $\Psi_{x_i}^{p_i^u, p_i^s}$ , therefore (Proposition 4.11),  $f^i(x) = \Psi_{x_i}(v_i, w_i)$  where  $|v_i|, |w_i| \leq 10^{-2}(p_i^u \wedge p_i^s)$ . By construction, Pesin charts are 2-Lipschitz, therefore  $d(f^i(x), x_i) < 50^{-1}(p_i^u \wedge p_i^s)$ . Similarly  $d(f^i(x), y_i) < 50^{-1}(q_i^u \wedge q_i^s)$ . It follows that  $d(x_i, y_i) < 25^{-1} \max\{p_i^u \wedge p_i^s, q_i^u \wedge q_i^s\} < \varepsilon$  for all  $i \in \mathbb{Z}$ .

Assume without loss of generality that  $\varepsilon$  is smaller than the Lebesgue number of the cover  $\mathcal{D}$  which we had constructed in §3.1, then  $x_i, y_i$  belong to the same

element  $D_i$  of  $\mathcal{D}$ . This allows us to write

$$\begin{aligned}\Psi_{x_i} &= \exp_{x_i} \circ \vartheta_{x_i} \circ C_{x_i} \\ \Psi_{y_i} &= \exp_{y_i} \circ \vartheta_{y_i} \circ C_{y_i}\end{aligned}$$

where  $\vartheta_{z_i} : \mathbb{R}^2 \rightarrow T_{z_i}M$  ( $z_i = x_i, y_i$ ) are the isometries we constructed in §3.1, and  $C_{x_i}, C_{y_i} \in \text{GL}(2, \mathbb{R})$  are given by  $C_\chi(x_i) = \vartheta_{x_i} \circ C_{x_i}$  and  $C_\chi(y_i) = \vartheta_{y_i} \circ C_{y_i}$ .

Let  $z_i = x_i, y_i$ , then  $C_\chi(z_i)$  is the unique linear operator which maps  $\underline{e}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  to  $s_\chi(z_i)^{-1} \underline{e}^s(z_i)$ , and  $\underline{e}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  to  $u_\chi(z_i)^{-1} \underline{e}^u(z_i)$ . Writing as usual  $\alpha(z_i) := \angle(\underline{e}^s(z_i), \underline{e}^u(z_i))$ , we see that

$$C_{z_i} = R_{z_i} \begin{pmatrix} s_\chi(z_i)^{-1} & u_\chi(z_i)^{-1} \cos \alpha(z_i) \\ 0 & u_\chi(z_i)^{-1} \sin \alpha(z_i) \end{pmatrix}, \quad (5.2)$$

where  $R_{z_i}$  is the unique orientation preserving orthogonal matrix which rotates  $\underline{e}^1$  to the direction of  $\vartheta_{z_i}^{-1}(\underline{e}^s(z_i))$  ( $z_i = x_i, y_i$ ). Some terminology:

- $z_i$  are called *position parameters*,
- $R_{z_i}$  and  $\alpha(z_i)$  are called *axes parameters*,
- $s_\chi(z_i), u_\chi(z_i)$  are called *scaling parameters*,
- $(p_i^u, p_i^s)$  are called *window parameters*.

The proof is done by comparing the parameters of  $\Psi_{x_i}^{p_i^u, p_i^s}$  to those of  $\Psi_{y_i}^{q_i^u, q_i^s}$ .

The comparison of the position parameters had already been done above. We record the conclusion for future reference:

**Proposition 5.3.** *Let  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}, (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  be two chains s.t.  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then  $d(x_i, y_i) < 25^{-1} \max\{p_i^u \wedge p_i^s, q_i^u \wedge q_i^s\}$  ( $i \in \mathbb{Z}$ ).*

Regularity is not needed here. We shall make use of it when we analyze the scaling parameters and the window parameters.

## 6. AXES PARAMETERS

Let  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}, (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  be two chains s.t.  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ . We compare  $R_{x_i}$  to  $R_{y_i}$  and  $\alpha(x_i)$  to  $\alpha(y_i)$ . The analysis relies on a special property of  $V^u[(z_k)_{k \leq i}]$  and  $V^s[(z_k)_{k \geq i}]$  ( $z_k = x_k, y_k$ ), which we call “*staying in windows*”. We begin by discussing this property.

### 6.1. Staying in windows.

**Definition 6.1.** *Let  $V^u$  be a  $u$ -admissible manifold in the double chart  $\Psi_x^{p^u, p^s}$ .  $V^u$  stays in windows if there is a negative chain  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \leq 0}$  with  $\Psi_{x_0}^{p_0^u, p_0^s} = \Psi_x^{p^u, p^s}$  and  $u$ -admissible manifolds  $W_i^u$  in  $\Psi_{x_i}^{p_i^u, p_i^s}$  s.t.  $f^{-|i|}(V_i^u) \subseteq W_i^u$  for all  $i \leq 0$ .*

**Definition 6.2.** *Let  $V^s$  be an  $s$ -admissible manifold in the double chart  $\Psi_x^{p^u, p^s}$ .  $V^s$  stays in windows if there is a positive chain  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \geq 0}$  with  $\Psi_{x_0}^{p_0^u, p_0^s} = \Psi_x^{p^u, p^s}$  and  $s$ -admissible manifolds  $W_i^s$  in  $\Psi_{x_i}^{p_i^u, p_i^s}$  s.t.  $f^i(V_i^s) \subseteq W_i^s$  for all  $i \geq 0$ .*

If  $\underline{v}$  is a chain, then  $V_i^u := V^u[(v_k)_{k \leq i}]$  and  $V_i^s := V^s[(v_k)_{k \geq i}]$  stay in windows, because  $f^{-k}(V_i^u) \subset V_{i-k}^u$  and  $f^k(V_i^s) \subset V_{i+k}^s$  for all  $k \geq 0$  (Proposition 4.15).

The following proposition says that  $s/u$ -admissible manifolds which stay in windows are local stable/unstable manifolds in the sense of Pesin [P]:

**Proposition 6.3.** *The following holds for all  $\varepsilon$  small enough. Let  $V^s$  be an admissible  $s$ -manifold in  $\Psi_x^{p^u, p^s}$ , and suppose  $V^s$  stays in windows.*

- (1) For every  $y, z \in V^s$ ,  $d(f^k(y), f^k(z)) < e^{-\frac{1}{2}k\chi}$  for all  $k \geq 0$ .
- (2) For every  $y \in V^s$ , let  $\underline{e}^s(y)$  denote the positively oriented unit tangent vector to  $V^s$  at  $y$ , then  $\|df_y^k \underline{e}^s(y)\|_{f^k(y)} \leq 6\|C_\chi(x)^{-1}\|e^{-\frac{1}{2}k\chi}$  for all  $k \geq 0$ .
- (3)  $|\log \|df_y^k \underline{e}^s(y)\|_{f^k(y)} - \log \|df_z^k \underline{e}^s(z)\|_{f^k(z)}| < Q_\varepsilon(x)^{\beta/4}$  ( $y, z \in V^s, k \geq 0$ ).

The symmetric statement holds for  $u$ -admissible manifolds which stay in windows: replace the  $s$ -tags by  $u$ -tags, and  $f$  by  $f^{-1}$ .

The proof is modeled on the proof of Pesin's Stable Manifold Theorem [BP, chapter 7]:  $f^n : V^s \rightarrow f^n(V^s)$  is given in coordinates by

$$\Psi_{x_n}^{-1} \circ f^n \circ \Psi_{x_0} = f_{x_{n-1}x_n} \circ \cdots \circ f_{x_0x_1}.$$

Since  $V^s$  stays in windows, the orbits of points in  $V^s$  remain in the "windows" where  $f_{x_i x_{i+1}}$  is close to a linear hyperbolic map. One can then prove the proposition by direct calculations. See the appendix for details.

**Proposition 6.4.** *The following holds for all  $\varepsilon$  small enough. Let  $V^s$  (resp.  $U^s$ ) be an  $s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  (resp. in  $\Psi_y^{q^u, q^s}$ ). Suppose  $V^s, U^s$  stay in windows. If  $x = y$  then either  $V^s, U^s$  are disjoint, or one contains the other.*

The same statement holds for  $u$ -admissible manifolds.

See the appendix for a proof.

## 6.2. Comparison of $\alpha(x_i)$ to $\alpha(y_i)$ .

**Proposition 6.5.** *Let  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}, (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  be chains s.t.  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then for all  $i \in \mathbb{Z}$*

- (1)  $e^{-\sqrt{\varepsilon}} \leq \frac{\sin \alpha(x_i)}{\sin \alpha(y_i)} \leq e^{\sqrt{\varepsilon}}$
- (2)  $|\cos \alpha(x_i) - \cos \alpha(y_i)| < \sqrt{\varepsilon}$

*Proof.* Write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}, u_i = \Psi_{y_i}^{q_i^u, q_i^s}, x := \pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , and

$$\begin{aligned} V_{x_k}^s &:= V^s[(v_i)_{i \geq k}] & V_{x_k}^u &:= V^u[(v_i)_{i \leq k}] & E_{x_k}^{s/u} &:= T_{f^k(x)} V_{x_k}^{s/u} \\ V_{y_k}^s &:= V^s[(u_i)_{i \geq k}] & V_{y_k}^u &:= V^u[(u_i)_{i \leq k}] & E_{y_k}^{s/u} &:= T_{f^k(x)} V_{y_k}^{s/u}. \end{aligned}$$

We claim that

- (i)  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_{f^k(x)}^n \underline{w}\| < 0$  on  $E_{x_k}^s \setminus \{0\}$  and  $E_{y_k}^s \setminus \{0\}$ ,
- (ii)  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_{f^k(x)}^n \underline{w}\| > 0$  on  $E_{x_k}^u \setminus \{0\}$  and  $E_{y_k}^u \setminus \{0\}$ .

We give the details for  $E_{x_k}^{s/u}$ . The case of  $E_{y_k}^{s/u}$  is identical.

Part (i) follows from Proposition 6.3 (2), applied to  $V_{x_k}^s$  and  $V_{y_k}^s$ .

The proof of (ii) is slightly more complicated. Suppose  $\underline{w} \in E_{x_k}^u \setminus \{0\}$ , then  $\underline{w}$  is tangent to  $V_{x_k}^u$  at  $f^k(x)$ . For every  $n$ ,  $f^{k+n}(x) = \pi[(v_{i+k+n})_{i \in \mathbb{Z}}] \in V_{x_{k+n}}^u$ , so

$$f^k(x) = f^{-n}(f^{k+n}(x)) \in f^{-n}[V_{x_{k+n}}^u].$$

It follows that  $df_{f^k(x)}^n \underline{w} \in T_{f^{k+n}(x)}[V_{x_{k+n}}^u] \setminus \{0\}$ .

We apply Proposition 6.3 (2) in its version for  $u$ -admissible manifolds to the manifold  $V_{x_{k+n}}^u$  and the vector  $df_{f^k(x)}^n \underline{w}$ . This gives the estimate

$$\begin{aligned} \|\underline{w}\| &= \|df_{f^{k+n}(x)}^{-n}[df_{f^k(x)}^n \underline{w}]\| \leq 6e^{-\frac{1}{2}n\chi} \|C_\chi(x_{k+n})^{-1}\| \cdot \|df_{f^k(x)}^n \underline{w}\| \\ &\leq 6e^{-\frac{1}{2}n\chi} Q_\varepsilon(x_{k+n})^{-1} \|df_{f^k(x)}^n \underline{w}\| \quad (\text{definition of } Q_\varepsilon) \end{aligned}$$

$$\begin{aligned} &\leq 6e^{-\frac{1}{2}n\chi}(p_{k+n}^u \wedge p_{k+n}^s)^{-1} \|df_{f^k(x)}^n \underline{w}\| \\ &\leq 6e^{-\frac{1}{2}n\chi+n\varepsilon}(p_k^u \wedge p_k^s)^{-1} \|df_{f^k(x)}^n \underline{w}\| \quad (\text{Lemma 4.4}). \end{aligned}$$

Thus  $\|df_{f^k(x)}^n \underline{w}\| \geq \frac{1}{6}e^{\frac{1}{2}n\chi+n\varepsilon}(p_k^u \wedge p_k^s)\|\underline{w}\|$ . Part (ii) follows.

By (i) and (ii),  $E_{x_k}^s, E_{y_k}^s = \{\underline{w} \in T_{f^k(x)}M : \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_{f^k(x)}^n \underline{w}\| < 0\}$ . For reasons of symmetry,  $E_{x_k}^u, E_{y_k}^u = \{\underline{w} \in T_{f^k(x)}M : \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_{f^k(x)}^{-n} \underline{w}\| < 0\}$ . It follows that  $E_{x_k}^s = E_{y_k}^s$  and  $E_{x_k}^u = E_{y_k}^u$ .

As a result,  $\angle(V_{x_k}^s, V_{x_k}^u) = \angle(V_{y_k}^s, V_{y_k}^u)$ . By Proposition 4.11  $\sin \angle(V_{x_k}^s, V_{x_k}^u) = e^{\pm(p_i^u \wedge p_i^s)^{\beta/4}} \sin \alpha(x_k)$  and  $\sin \angle(V_{y_k}^s, V_{y_k}^u) = e^{\pm(q_i^u \wedge q_i^s)^{\beta/4}} \sin \alpha(y_k)$ . Since  $p_i^u \wedge p_i^s \leq Q_\varepsilon(x_i) < \varepsilon^{3/\beta}$  and  $q_i^u \wedge q_i^s \leq Q_\varepsilon(y_i) < \varepsilon^{3/\beta}$ ,  $e^{-2\varepsilon^{3/4}} < \sin \alpha(x_k) / \sin \alpha(y_k) < e^{2\varepsilon^{3/4}}$ . Similarly one sees that  $|\cos \alpha(x_k) - \cos \alpha(y_k)| < 4\varepsilon^{3/4}$ , and the proposition follows for all  $\varepsilon$  so small that  $4\varepsilon^{3/4} < \sqrt{\varepsilon}$ .  $\square$

The proof actually gives the following stronger estimates, which we will serve their purpose as distortion compensating bounds in §9 below.

**Lemma 6.6.** *Under the assumptions of the previous proposition,*

- (1)  $e^{-(p_i^u \wedge p_i^s)^{\beta/4} - (q_i^u \wedge q_i^s)^{\beta/4}} < \frac{\sin \alpha(x_i)}{\sin \alpha(y_i)} < e^{(p_i^u \wedge p_i^s)^{\beta/4} + (q_i^u \wedge q_i^s)^{\beta/4}}$ ;
- (2)  $|\cos \alpha(x_i) - \cos \alpha(y_i)| < 4[(p_i^u \wedge p_i^s)^{\beta/4} + (q_i^u \wedge q_i^s)^{\beta/4}]$ .

### 6.3. Comparison of $R_{x_i}$ to $R_{y_i}$ .

**Proposition 6.7.** *The following holds for all  $\varepsilon$  small enough. For any two chains  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$  and  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$ , if  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then for all  $i$*

$$R_{y_i}^{-1} R_{x_i} = (-1)^{\sigma_i} \text{Id} + \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{pmatrix},$$

where  $\sigma_i \in \{0, 1\}$  and  $|\varepsilon_{jk}| < [(p_i^u \wedge p_i^s)^{\beta/5} + (q_i^u \wedge q_i^s)^{\beta/5}] < \sqrt{\varepsilon}$ .

*Proof.* In order to keep the notation as light as possible, we only do the case  $i = 0$ , and write  $\Psi_{x_0}^{p_0^u, p_0^s} = \Psi_x^{p^u, p^s}$ ,  $\Psi_{y_0}^{q_0^u, q_0^s} = \Psi_y^{q^u, q^s}$ ,  $p := p^u \wedge p^s$ , and  $q := q^u \wedge q^s$ . We also set as usual  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$  and  $u_i = \Psi_{y_i}^{q_i^u, q_i^s}$ .

Let  $z = \pi[v] = \pi[u]$ . The manifold  $V^s[(v_i)_{i \geq 0}]$  inherits an orientation from the chart  $\Psi_x$ . Let  $\underline{e}_x^s(z)$  denote the positively oriented unit tangent vector to  $V^s[(v_i)_{i \geq 0}]$  at  $z$ . The manifold  $V^s[(u_i)_{i \geq 0}]$  inherits an orientation from the chart  $\Psi_y$ . Let  $\underline{e}_y^s(z)$  denote the positively oriented unit tangent vector to  $V^s[(u_i)_{i \geq 0}]$  at  $z$ . Since  $T_z V^s[(v_i)_{i \in \mathbb{Z}}] = T_z V^s[(u_i)_{i \in \mathbb{Z}}]$  (see the proof of Proposition 6.5),  $\underline{e}_x^s(z) = \pm \underline{e}_y^s(z)$ .

We write  $z$  and  $\underline{e}_x^s(z), \underline{e}_y^s(z)$  in coordinates in  $\Psi_x$  and  $\Psi_y$ :

- $z = \Psi_x(\underline{\zeta})$  and  $\underline{e}_x^s(z) = \frac{[(d\Psi_x)_{\underline{\zeta}}]\underline{a}}{\|[(d\Psi_x)_{\underline{\zeta}}]\underline{a}\|}$ , where  $\underline{\zeta} \in R_{10-2p}(0)$ ,  $\underline{a} = \begin{pmatrix} 1 \\ a \end{pmatrix}$ , and  $|a| \leq p^{\beta/3}$  (see Proposition 4.11 and (4.2)).
- $z = \Psi_y(\underline{\eta})$  and  $\underline{e}_y^s(z) = \frac{[(d\Psi_y)_{\underline{\eta}}]\underline{b}}{\|[(d\Psi_y)_{\underline{\eta}}]\underline{b}\|}$ , where  $\underline{\eta} \in R_{10-2q}(0)$ ,  $\underline{b} = \begin{pmatrix} 1 \\ b \end{pmatrix}$ , and  $|b| \leq q^{\beta/3}$  (see Proposition 4.11 and (4.2)).

Since  $\underline{e}_x^s(z) = \pm \underline{e}_y^s(z)$ , there is a non-zero (signed) scalar  $\lambda$  such that

$$C_x \underline{a} = \lambda [(d \exp_x \circ \vartheta_x)_{C_x \underline{\zeta}}]^{-1} [(d \exp_y \circ \vartheta_y)_{C_y \underline{\eta}}] C_y \underline{b}, \quad (6.1)$$

where  $C_x, C_y$  are given by (5.2).

*Claim 1.*  $C_x \underline{a} \propto R_x \begin{pmatrix} 1 \pm p^{\beta/4} \\ 0 \pm p^{\beta/4} \end{pmatrix}$  and  $C_y \underline{b} \propto R_y \begin{pmatrix} 1 \pm q^{\beta/4} \\ 0 \pm q^{\beta/4} \end{pmatrix}$ . Here  $\vec{a} \propto \vec{b}$  means that  $\vec{a} = t\vec{b}$  for some  $t \neq 0$ , and  $a \pm c$  means a quantity in  $[a - c, a + c]$ .

$$\begin{aligned} \text{Proof. } C_x \underline{a} &= R_x \begin{pmatrix} s_\chi(x)^{-1} + u_\chi(x)^{-1} \cos \alpha(x) a \\ u_\chi(x)^{-1} \sin \alpha(x) a \end{pmatrix} \\ &\propto R_x \begin{pmatrix} 1 \pm \|C_\chi(x)^{-1}\| \cdot |a| \\ 0 \pm \|C_\chi(x)^{-1}\| \cdot |a| \end{pmatrix}, \text{ because } u_\chi > 1 \text{ and } s_\chi = \|C_\chi(x)^{-1} \underline{e}^s(x)\| \\ &= R_x \begin{pmatrix} 1 \pm p^{\beta/4} \\ 0 \pm p^{\beta/4} \end{pmatrix}, \text{ because } |a| < p^{\beta/3} \leq Q_\chi(x)^{\beta/12} p^{\beta/4} < \frac{p^{\beta/4}}{\|C_\chi(x)^{-1}\|}. \end{aligned}$$

Similarly,  $C_y \underline{b} \propto R_y \begin{pmatrix} 1 \pm q^{\beta/4} \\ 0 \pm q^{\beta/4} \end{pmatrix}$ .

*Claim 2.* There exists a constant  $J > 1$  (which only depends on  $M$ ) s.t. for all  $D \in \mathcal{D}$ ,  $x, y \in D$ , and  $\|\underline{w}_1\|, \|\underline{w}_2\| < 2$ ,

$$\|[(d \exp_x \circ \vartheta_x)_{\underline{w}_1}]^{-1} [(d \exp_y \circ \vartheta_y)_{\underline{w}_2}] - \text{Id}\| < J(d(x, y) + \|\underline{w}_1 - \underline{w}_2\|).$$

*Proof.* Let  $J_1$  denote a common Lipschitz constant for the maps

$$(w, \underline{w}) \mapsto (d \exp_w \circ \vartheta_w)_{\underline{w}}$$

on  $D \times B_2(\underline{0})$  for all  $D \in \mathcal{D}$ . Let  $J_2$  denote the maximum over  $D \in \mathcal{D}$  of  $\sup\{\|(d \exp_w \circ \vartheta_w)_{\underline{w}}^{-1}\| : w \in D, \|\underline{w}\| < 2\}$ . The claim holds with  $J := J_1 J_2 + 1$ .

*Claim 3.*  $R_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underline{\varepsilon}_1 \propto R_y \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underline{\varepsilon}_2$  where  $\|\underline{\varepsilon}_1\|$  and  $\|\underline{\varepsilon}_2\|$  are less than  $3J(p^{\beta/4} + q^{\beta/4})$ .

*Proof.*  $C_\chi(\cdot)$  is a contraction, so  $\|C_x \underline{\zeta} - C_y \underline{\eta}\| < \|\underline{\zeta}\| + \|\underline{\eta}\| < 10^{-2}(p + q)$ . Also, by Proposition 5.3,  $d(x, y) < 25^{-1}(p + q)$ . Therefore, by Claim 2,

$$[(d \exp_x \circ \vartheta_x)_{C_\varepsilon(x) \underline{\zeta}}]^{-1} [(d \exp_y \circ \vartheta_y)_{C_\varepsilon(y) \underline{\eta}}] = \text{Id} + E$$

where  $E$  is a matrix s.t.  $\|E\| < J(p + q)$ . The claim follows from (6.1) by direct calculation.

We can now prove the proposition.  $R_x$  and  $R_y$  are rotation matrices, therefore  $R_y^{-1} R_x$  is a rotation matrix. The problem is to estimate the angle. Claim 3 allows us to write

$$R_y^{-1} R_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + R_y^{-1} \underline{\varepsilon}_2 - c^{-1} R_y^{-1} \underline{\varepsilon}_1 \right], \quad (6.2)$$

where  $c$  is a scalar s.t.  $|c| = \frac{1 \pm \|\underline{\varepsilon}_1\|}{1 \pm \|\underline{\varepsilon}_2\|}$ . Since  $\|\underline{\varepsilon}_i\| < 3J(p^{\beta/4} + q^{\beta/4}) < 6J\varepsilon^{3/4}$ ,  $|c| \in [e^{-10J\sqrt{\varepsilon}}, e^{10J\sqrt{\varepsilon}}]$ , at least provided  $\varepsilon$  is small enough.

Since  $R_x$  and  $R_y$  are orthogonal matrices, the vector on the right-hand side of (6.2) is a unit vector. Put it in the form  $(-1)^{\sigma_0} (\cos \theta, \sin \theta)$  where  $\sigma_0 \in \{0, 1\}$  and  $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ , then

$$\begin{aligned} |\theta| &\leq \tan^{-1} \left( \frac{\|\underline{\varepsilon}_2\| + |c|^{-1} \cdot \|\underline{\varepsilon}_1\|}{1 - \|\underline{\varepsilon}_2\| - |c|^{-1} \|\underline{\varepsilon}_1\|} \right) < \frac{\|\underline{\varepsilon}_2\| + |c|^{-1} \cdot \|\underline{\varepsilon}_1\|}{1 - \|\underline{\varepsilon}_2\| - |c|^{-1} \|\underline{\varepsilon}_1\|} \\ &< \frac{3J(1 + e^{10J\sqrt{\varepsilon}})}{1 - 6J(1 + e^{10J\sqrt{\varepsilon}})\varepsilon^{3/4}} (p^{\beta/4} + q^{\beta/4}). \end{aligned}$$

Since  $p, q < \varepsilon^{3/\beta}$ , if  $\varepsilon$  is small enough, then this is less than  $p^{\beta/5} + q^{\beta/5} < 2\varepsilon^{3/5} < \sqrt{\varepsilon}$ . It follows that  $(-1)^{\sigma_0} R_y^{-1} R_x$  is a rotation by angle less than  $p^{\beta/5} + q^{\beta/5} < \sqrt{\varepsilon}$ .  $\square$



## 7. SCALING PARAMETERS

**7.1. The  $s_\chi$  and  $u_\chi$  parameters of admissible manifolds.** In §2.1 we defined  $s_\chi(\cdot)$  on  $\text{NUH}_\chi(f)$ . We now extend this definition to all points lying on  $s$ -admissible manifolds  $V^s$  which stay in windows.

Suppose  $y \in V^s$ . If  $y \in \text{NUH}_\chi(f)$  define  $\underline{e}^s(y)$  as in §2.1, and note that by proposition 6.3(2),  $\underline{e}^s(y)$  is tangent to  $V^s$  at  $y$ . Motivated by this, we define  $\underline{e}^s(y)$  for  $y \notin \text{NUH}_\chi(f)$  to be one of the two unit tangent vectors to  $V^s$  at  $y$  (it doesn't matter which), and then we let

$$s_\chi(y) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \|df_y^k \underline{e}^s(y)\|_{f^k(y)}^2 \right)^{\frac{1}{2}} \in (\sqrt{2}, \infty].$$

Similarly, for any  $u$ -admissible manifold  $V^u$  which stays in windows, and any  $y \in V^u$  we define  $\underline{e}^u(y)$  as in §2.1 when  $y \in \text{NUH}_\chi(f)$ , and we let  $\underline{e}^u(y)$  be one of the two unit tangent vectors to  $V^u$  at  $y$  when  $y \notin \text{NUH}_\chi(f)$ . Then we let

$$u_\chi(y) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \|df_y^{-k} \underline{e}^u(y)\|_{f^{-k}(y)}^2 \right)^{\frac{1}{2}} \in (\sqrt{2}, \infty].$$

Although these numbers depend on  $y$ , they are not very sensitive to its value: by Proposition 6.3 part 3, for any pair of points  $y, z$  in the same  $s$ -admissible manifold, if  $s_\chi(y)$  is finite then  $s_\chi(z)$  is finite, and

$$e^{-\sqrt{\varepsilon}} < s_\chi(y)/s_\chi(z) < e^{\sqrt{\varepsilon}}.$$

A similar statement holds for  $u_\chi$ -parameters on  $u$ -admissible manifolds.

**Definition 7.1.** Let  $V^s$  be an  $s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  with representing function  $F^s$ . Let  $V^u$  be a  $u$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  with representing function  $F^u$ . If  $V^s$  and  $V^u$  stay in windows, then

- (1)  $s_\chi(V^s)$ , the  $s_\chi$ -parameter of  $V^s$ , is  $s_\chi(p)$  where  $p := \Psi_x(0, F^s(0))$ ,
- (2)  $u_\chi(V^u)$ , the  $u_\chi$ -parameter of  $V^u$ , is  $u_\chi(q)$  where  $q := \Psi_x(F^u(0), 0)$ .

**Lemma 7.2.** The following holds for all  $\varepsilon$  small enough. Suppose  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ , and let  $V^s$  be an  $s$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  which stays in windows. If  $s_\chi(V^s) < \infty$  then  $s_\chi(\mathcal{F}_s(V^s)) < \infty$ , and for every  $\rho \geq \exp(\sqrt{\varepsilon})$ ,

$$\frac{s_\chi(V^s)}{s_\chi(y)} \in [\rho^{-1}, \rho] \implies \frac{s_\chi(\mathcal{F}_s(V^s))}{s_\chi(x)} \in \left[ \rho^{-1} e^{Q_\varepsilon(x)^{\beta/4}}, \rho e^{-Q_\varepsilon(x)^{\beta/4}} \right]. \quad (7.1)$$

A similar statement holds for  $u$ -admissible manifolds in  $\Psi_x^{p^u, p^s}$  and  $\mathcal{F}_u$ .

Note that the ratio bound in (7.1) improves.

*Proof.* Suppose  $V^s$  is represented by the function  $G$ , and  $U^s := \mathcal{F}_s[V^s]$  is represented by the function  $F$ . Let  $p := \Psi_x(0, F(0))$  and  $q := \Psi_y(0, G(0))$ .

Suppose  $s_\chi(V^s) < \infty$ , then  $s_\chi(q) < \infty$ . By Proposition 4.12(4) (in its version for  $s$ -manifolds),  $f^{-1}(q) \in U^s$ . Since  $U^s$  is one-dimensional,  $df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q)) = \pm \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\|_q \cdot \underline{e}^s(q)$ , and so

$$\begin{aligned} s_\chi(f^{-1}(q))^2 &\equiv 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k\chi} \|df_q^{k-1} df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\|_{f^{k-1}(q)}^2 \right) \\ &= 2 + e^{2\chi} \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\|_q^2 \cdot s_\chi(q)^2 < \infty. \end{aligned}$$

Since  $f^{-1}(q) \in U^s$ ,  $s_\chi(U^s) \leq e^{\sqrt{\varepsilon}} s_\chi(f^{-1}(q)) < \infty$ .

Next assume that  $s_\chi(V^s)$  is finite, and

$$\frac{s_\chi(V^s)}{s_\chi(y)} \in [\rho^{-1}, \rho].$$

where  $\rho \geq \exp(\sqrt{\varepsilon})$ . Since  $s_\chi(U^s) = s_\chi(p)$ ,

$$\frac{s_\chi(U^s)}{s_\chi(x)} = \frac{s_\chi(p)}{s_\chi(f^{-1}(q))} \cdot \frac{s_\chi(f^{-1}(q))}{s_\chi(f^{-1}(y))} \cdot \frac{s_\chi(f^{-1}(y))}{s_\chi(x)}. \quad (7.2)$$

The three terms are well-defined and finite, because (proceeding from right to left):

- $s_\chi(x), s_\chi(f^{-1}(y))$  are well-defined and finite, because  $x, y \in \text{NUH}_\chi(f)$ ;
- $s_\chi(f^{-1}(q))$  is finite by the argument at the beginning of the proof;
- $s_\chi(p) < \infty$ , because  $s_\chi(p) = S_\chi(U_s) < \infty$  (see above).

The first factor in (7.2) belongs to  $[e^{-Q_\varepsilon(x)^{\beta/4}}, e^{Q_\varepsilon(x)^{\beta/4}}]$  by Proposition 6.3(3). The third factor in (7.2) takes values in  $[e^{-Q_\varepsilon(x)^{\beta/4}}, e^{Q_\varepsilon(x)^{\beta/4}}]$  because  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ , see Lemma 3.3. To prove the proposition, it is enough to show that

$$\frac{1}{\rho} \exp[3Q_\varepsilon(x)^{\beta/4}] < \frac{s_\chi(f^{-1}(q))}{s_\chi(f^{-1}(y))} < \rho \exp[-3Q_\varepsilon(x)^{\beta/4}]. \quad (7.3)$$

We begin with some identities. We omit the tags of the Riemannian norm, to avoid heavy notation. Since  $df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y)) = \pm \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\| \cdot \underline{e}^s(y)$ ,

$$\begin{aligned} s_\chi(f^{-1}(y))^2 &= 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k\chi} \|df_y^{k-1} df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2 \right) \\ &= 2 + e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2. \end{aligned} \quad (7.4)$$

Similarly,  $df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q)) = \pm \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\| \cdot \underline{e}^s(q)$ , so

$$\begin{aligned} s_\chi(f^{-1}(q))^2 &= 2 + e^{2\chi} s_\chi(q)^2 \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\|^2 \\ &\leq 2 + \rho^2 e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\|^2 \left( \because \frac{s_\chi(q)}{s_\chi(y)} = \frac{s_\chi(V^s)}{s_\chi(y)} \leq \rho \right) \\ &\leq \left( 2 + \rho^2 e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2 \right) \times \\ &\quad \times \exp \left( 2 \left| \log \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\| - \log \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\| \right| \right). \end{aligned}$$

We obtain the estimate

$$\begin{aligned} \frac{s_\chi(f^{-1}(q))^2}{s_\chi(f^{-1}(y))^2} &\leq \left( \frac{2 + \rho^2 e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2}{2 + e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2} \right) \times \\ &\quad \times \exp \left( 2 \left| \log \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\| - \log \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\| \right| \right). \end{aligned} \quad (7.5)$$

Call the first factor I and the second factor II.

*Analysis of I.*

$$\begin{aligned}
\text{I} &= \rho^2 - \frac{2(\rho^2 - 1)}{2 + e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2} \\
&= \rho^2 - \frac{2(\rho^2 - 1)}{s_\chi(f^{-1}(y))^2}, \text{ by (7.4)} \\
&\leq \rho^2 - \frac{e^{-2\varepsilon^{6/\beta}} \cdot 2(\rho^2 - 1)}{s_\chi(x)^2}, \text{ because } \frac{s_\chi(f^{-1}(y))}{s_\chi(x)} = \exp[\pm\varepsilon^{6/\beta}] \text{ by Lemma 3.3} \\
&\leq \rho^2 \left( 1 - \frac{2e^{-2\varepsilon^{6/\beta}} (1 - \rho^{-2})}{\|C_\chi(x)^{-1}\|^2} \right), \text{ since } s_\chi(x) = \|C_\chi(x)^{-1} \underline{e}^s(x)\| \leq \|C_\chi(x)^{-1}\| \\
&\leq \rho^2 \left( 1 - \frac{\varepsilon^{1/2}}{\|C_\chi(x)^{-1}\|^2} \right) \text{ for all } \varepsilon \text{ small enough, because } \rho \geq e^{\sqrt{\varepsilon}}.
\end{aligned}$$

By the definition of  $Q_\varepsilon(x)$ ,

$$\frac{\varepsilon^{1/2}}{\|C_\chi(x)^{-1}\|^2} > Q_\varepsilon(x)^{\beta/6} = Q_\varepsilon(x)^{-\beta/12} Q_\varepsilon(x)^{\beta/4} > \varepsilon^{-1/4} Q_\varepsilon(x)^{\beta/4}.$$

In particular, for all  $\varepsilon$  small enough,  $\frac{\varepsilon^{1/2}}{\|C_\chi(x)^{-1}\|^2} > 7Q_\varepsilon(x)^{\beta/4}$ , and by the inequality  $1 - x < e^{-x}$  for  $0 < x < 1$ ,  $\text{I} \leq \rho^2 \exp[-7Q_\varepsilon(x)^{\beta/4}]$ .

*Analysis of II.* Since  $f$  is a  $C^{1+\beta}$ -diffeomorphism,  $(p, \vec{v}) \mapsto df_p \vec{v}$  can be written in coordinates as a linear map of the coordinates of  $\vec{v}$ , with coefficients which are  $\beta$ -Hölder continuous functions of the coordinates of  $p$ . Since  $\|\underline{e}^s(\cdot)\| = 1$  and  $\|df\|$  is uniformly bounded, there exists a constant  $K_0 = K_0(f)$  so that

$$\text{II} \leq \exp \left[ K_0 d_M(f^{-1}(q), f^{-1}(y))^\beta + K_0 d_{TM}(\underline{e}^s(f^{-1}(q)), \underline{e}^s(f^{-1}(y))) \right],$$

where  $d_M$  and  $d_{TM}$  are the Riemannian distance functions on  $M$  and  $TM$ .

Since  $f$  is a  $C^{1+\beta}$  diffeomorphism and  $\underline{e}^s(\cdot)$  are unit vectors, there is another constant  $H_1$  (which only depends on  $f$ ), such that

$$\text{II} \leq \exp \left[ H_1 d_M(q, y)^\beta + H_1 d_{TM}(\underline{e}^s(q), \underline{e}^s(y))^\beta \right].$$

We estimate  $d(q, y)$ . By definition  $q = \Psi_y(0, G(0))$  and  $y = \Psi_y(0, 0)$ . Since Pesin charts have Lipschitz constant smaller than or equal to 2,

$$d(q, y) < 2|G(0)| \leq 2 \cdot 10^{-3} (q^u \wedge q^s) \leq 2 \cdot 10^{-3} \cdot e^\varepsilon (p^u \wedge p^s)$$

(see Lemma 4.4). In particular,  $d(q, y) < Q_\varepsilon(x)$ .

We estimate  $d_{TM}(\underline{e}^s(q), \underline{e}^s(y))$ . By the definition of  $\Psi_y$ ,  $\underline{e}^s(y)$  is the normalization of  $(d\Psi_y)_0 \binom{1}{0} = (d \exp_y)_0 [C_\chi(y) \binom{1}{0}]$ , and  $\underline{e}^s(q)$  is the normalization of

$$(d\Psi_y)_{(0, G(0))} \binom{1}{G'(0)} = (d \exp_y)_{C_\chi(y)(G(0))} \left[ C_\chi(y) \binom{1}{G'(0)} \right].$$

It is not difficult to see using the admissibility of  $V^s$  and Lemma 4.4 that  $|G(0)| < Q_\varepsilon(x)$  and  $|G'(0)| < Q_\varepsilon(x)^{\beta/3}$ . Since  $C_\chi(y)$  is a contraction,  $p \mapsto \exp_p$  is smooth, and  $d(q, y) < Q_\varepsilon(x)$ , there exists a constant  $G_0$  (which only depends on the smoothness of the exponential function) such that  $d_{TM}(\underline{e}^s(q), \underline{e}^s(y)) < G_0 Q_\varepsilon(x)^{\beta/3}$ .

We see that  $\text{II} \leq \exp[(H_1 + H_1 G_0) Q_\varepsilon(x)^{\beta/3}]$ . It follows that for all  $\varepsilon$  sufficiently small,  $\text{II} \leq \exp[Q_\varepsilon(x)^{\beta/4}]$ .

*Summary.* Combining the estimates of I and II, we find that

$$\frac{s_\chi(f^{-1}(q))}{s_\chi(f^{-1}(y))} \leq \rho \exp[-3Q_\varepsilon(x)^{\beta/4}].$$

The other half of (7.3) is proved in a similar way. First, one proves that

$$\begin{aligned} \frac{s_\chi(f^{-1}(q))^2}{s_\chi(f^{-1}(y))^2} &\geq \left( \frac{2 + \rho^{-2} e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2}{2 + e^{2\chi} s_\chi(y)^2 \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\|^2} \right) \times \\ &\times \exp\left(-2 \left| \log \|df_{f^{-1}(q)} \underline{e}^s(f^{-1}(q))\| - \log \|df_{f^{-1}(y)} \underline{e}^s(f^{-1}(y))\| \right| \right), \end{aligned}$$

and then one analyzes the two terms as before.  $\square$

## 7.2. Comparison of $s_\chi(x_i), u_\chi(x_i)$ to $s_\chi(y_i), u_\chi(y_i)$ .

**Proposition 7.3.** *The following holds for all  $\varepsilon$  small enough. For any two regular chains  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}, (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$ , if  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then*

$$e^{-4\sqrt{\varepsilon}} \leq \frac{s_\chi(x_i)}{s_\chi(y_i)} \leq e^{4\sqrt{\varepsilon}} \quad \text{and} \quad e^{-4\sqrt{\varepsilon}} \leq \frac{u_\chi(x_i)}{u_\chi(y_i)} \leq e^{4\sqrt{\varepsilon}} \quad \text{for all } i \in \mathbb{Z}.$$

*Proof.* Write  $\underline{v} := (\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}, \underline{u} := (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$ , and  $p := \pi(\underline{v}) = \pi(\underline{u})$ .

Let  $V_k^s := V^s[(v_i)_{i \geq k}]$ ,  $V_k^u := V^u[(v_i)_{i \leq k}]$ ,  $U_k^s := V^s[(u_i)_{i \geq k}]$ ,  $U_k^u := V^u[(u_i)_{i \leq k}]$ . We claim that it is enough to prove that

$$\frac{s_\chi(V_k^s)}{s_\chi(x_k)}, \frac{u_\chi(V_k^s)}{u_\chi(x_k)}, \frac{s_\chi(U_k^s)}{s_\chi(y_k)}, \frac{u_\chi(U_k^s)}{u_\chi(y_k)} \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]. \quad (7.6)$$

Here is the reason. The manifolds  $V_k^s$  stay in windows and contain  $f^k(p)$ , therefore by Proposition 6.3(3)  $s_\chi(V_k^s)/s_\chi(f^k(p)) \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$ . The same argument applies to  $U_k^s, V_k^u, U_k^u$ , so  $\frac{s_\chi(V_k^s)}{s_\chi(f^k(p))}, \frac{u_\chi(V_k^u)}{u_\chi(f^k(p))}, \frac{s_\chi(U_k^s)}{s_\chi(f^k(p))}, \frac{u_\chi(U_k^u)}{u_\chi(f^k(p))} \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$ . Decomposing  $\frac{s_\chi(x_k)}{s_\chi(y_k)} = \frac{s_\chi(x_k)}{s_\chi(V_k^s)} \cdot \frac{s_\chi(V_k^s)}{s_\chi(f^k(p))} \cdot \frac{s_\chi(f^k(p))}{s_\chi(U_k^s)} \cdot \frac{s_\chi(U_k^s)}{u_\chi(y_k)}$ , we see that (7.6) implies that  $s_\chi(x_k)/s_\chi(y_k) \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}]$ . Similarly,  $u_\chi(x_k)/u_\chi(y_k) \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}]$ .

We show that  $s_\chi(V_0^s)/s_\chi(x_0) \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$ . The other parts of (7.6) are proved in the same way, and are left to the reader.

We are assuming that  $\underline{v}$  is regular, therefore there exists a relevant double chart  $v$  and a sequence  $n_k \uparrow \infty$  s.t.  $v_{n_k} = v$  for all  $k$ . Write  $v = \Psi_x^{p^u, p^s}$ .

*Claim 1.* There exists some  $\rho \geq \exp(\sqrt{\varepsilon})$  which only depends on  $v$  such that  $s_\chi(V_{n_k}^s)/s_\chi(x_{n_k}) \in [\rho^{-1}, \rho]$  for all  $k$ .

*Proof.* By convention  $v$  is relevant (see §4.4). Choose a chain  $\underline{w}$  s.t.  $w_0 = v$  and  $w := \pi(\underline{w}) \in \text{NUH}_\chi(f)$ . Let  $W^s := V^s[(w_i)_{i \geq 0}]$ . This manifold has a finite  $s_\chi$ -parameter, because  $s_\chi(W^s) \leq e^{\sqrt{\varepsilon}} s_\chi(w)$  and  $w \in \text{NUH}_\chi(f)$  so  $s_\chi(w) < \infty$ . Let

$$\rho_0 := \max \left\{ \frac{s_\chi(W^s)}{s_\chi(x)}, \frac{s_\chi(x)}{s_\chi(W^s)}, \exp(\sqrt{\varepsilon}) \right\}.$$

$W^s$  is an admissible manifold in  $v_{n_k} = v$ . By Proposition 4.15, if we take  $W^s$  at  $v_{n_k+\ell}$  and apply to it the graph transform  $\mathcal{F}_s$   $n_k+\ell - n_k$  times using the path  $(v_{n_k}, \dots, v_{n_k+\ell})$ , then the resulting manifold

$$W_\ell^s := \mathcal{F}_s^{n_k+\ell - n_k} [W^s]$$

is an  $s$ -admissible manifold in  $v_{n_k}$ , which converges to  $V_{n_k}^s$ . By Lemma 7.2,

$$\frac{s_\chi(W_\ell^s)}{s_\chi(x)} \in [\rho_0^{-1}, \rho_0]. \quad (7.7)$$

The convergence of  $W_\ell^s$  to  $V_{n_k}^s$  means that if  $W_\ell^s$  is represented in  $v_{n_k} = \Psi_x^{p^u, p^s}$  by the function  $F_\ell$ , and  $V_{n_k}^s$  is represented in  $\Psi_x^{p^u, p^s}$  by  $F$ , then  $\|F_\ell - F\|_\infty \xrightarrow{\ell \rightarrow \infty} 0$ . In fact, since  $\sup \|F'_\ell\|_{\beta/3} < \infty$ , we have the stronger statement that

$$\|F_\ell - F\|_\infty + \|F'_\ell - F'\|_\infty \xrightarrow{\ell \rightarrow \infty} 0,$$

see part 2 of the proof of Proposition 4.15. Therefore, if  $\xi := \Psi_x(0, F(0))$  and  $\xi_\ell = \Psi_x(0, F_\ell(0))$ , then  $\xi_\ell \xrightarrow{\ell \rightarrow \infty} \xi$  and  $e^s(\xi_\ell) \xrightarrow{\ell \rightarrow \infty} e^s(\xi)$ .

Fix some  $N$  large and  $\delta > 0$  small. Since  $df$  is continuous, there exists  $\ell$  so large that

$$\sqrt{2} \left( \sum_{j=0}^N e^{2j\chi} \|df_\xi^j e^s(f^j(\xi))\|_{f^j(\xi)}^2 \right)^{\frac{1}{2}} \leq e^\delta \cdot \sqrt{2} \left( \sum_{j=0}^N e^{2j\chi} \|df_{\xi_\ell}^j e^s(f^j(\xi_\ell))\|_{f^j(\xi_\ell)}^2 \right)^{\frac{1}{2}}.$$

The expression on the right is smaller than  $e^\delta s_\chi(W_\ell^s)$ , and therefore by (7.7), smaller than  $e^\delta \rho_0 s_\chi(x)$ . Since this is true for all  $N$  and  $\delta$ ,  $s_\chi(V_{n_k}^s) \leq \rho_0 \cdot s_\chi(x)$ .

Recalling that  $x_{n_k} = x$  and that  $s_\chi(V_{n_k}^s) \geq \sqrt{2}$ , we see that  $s_\chi(V_{n_k}^s)/s_\chi(x_{n_k}) \in [\sqrt{2}/s_\chi(x), \rho_0]$ . The claim follows with  $\rho = \rho_0 \cdot s_\chi(x)$ .

*Claim 2.*  $s_\chi(V_0^s)/s_\chi(x_0) \in [\exp(-\sqrt{\varepsilon}), \exp(\sqrt{\varepsilon})]$ .

*Proof.* Fix  $k$  large. By claim 1,

$$\frac{s_\chi(V_{n_k}^s)}{s_\chi(x_{n_k})} \in [\rho^{-1}, \rho].$$

By Proposition 4.15 (3),  $\mathcal{F}_s(V_{n_k}^s) = V_{n_k-1}^s$ , and by Lemma 7.2, the bounds for  $\frac{s_\chi(V_{n_k}^s)}{s_\chi(x_{n_k})}$  improve. We ignore these improvements and write  $\frac{s_\chi(V_{n_k-1}^s)}{s_\chi(x_{n_k-1})} \in [\rho^{-1}, \rho]$ .

Another application of  $\mathcal{F}_s$  gives  $\frac{s_\chi(V_{n_k-2}^s)}{s_\chi(x_{n_k-2})} \in [\rho^{-1}, \rho]$ . Continuing this way, we eventually reach the index  $n_{k-1} + 1$  and the bound

$$\frac{s_\chi(V_{n_{k-1}+1}^s)}{s_\chi(x_{n_{k-1}+1})} \in [\rho^{-1}, \rho]$$

Since  $x_{n_k} = x$ , the next application of  $\mathcal{F}_s$  improves the ratio bound by at least  $\exp[Q_\varepsilon(x)^{\beta/4}]$ :

$$\frac{s_\chi(V_{n_{k-1}}^s)}{s_\chi(x_{n_{k-1}})} \in [\rho^{-1} e^{Q_\varepsilon(x)^{\beta/4}}, \rho e^{-Q_\varepsilon(x)^{\beta/4}}].$$

We repeat the procedure by applying  $\mathcal{F}_s$   $n_{k-1} - n_{k-2} + 1$  times, whilst ignoring the potential improvements of the error bounds, and then applying  $\mathcal{F}_s$  once more and arriving at

$$\frac{s_\chi(V_{n_{k-2}}^s)}{s_\chi(x_{n_{k-2}})} \in [\rho^{-1} e^{2Q_\varepsilon(x)^{\beta/4}}, \rho e^{-2Q_\varepsilon(x)^{\beta/4}}].$$

We are free to choose  $k$  as large as we want. If we make it so large that  $\exp[kQ_\varepsilon(x)^{\beta/4}] > \rho \exp(-\sqrt{\varepsilon})$ , then eventually we will reach a time  $n_{k_0}$  when the ratio bound is smaller than or equal to  $\exp(\sqrt{\varepsilon})$ :

$$\frac{s_\chi(V_{n_{k_0}}^s)}{s_\chi(x_{n_{k_0}})} \in [\exp(-\sqrt{\varepsilon}), \exp(\sqrt{\varepsilon})].$$

This is the threshold of applicability of Lemma 7.2. Henceforth we cannot claim that the ratio bound improves. On the other hand it is guaranteed that the ratio bound does not deteriorate. Therefore, after additional  $n_{k_0}$  iterations, we obtain  $\frac{s_\chi(V_0^s)}{s_\chi(x_0)} \in [\exp(-\sqrt{\varepsilon}), \exp(\sqrt{\varepsilon})]$  as desired.  $\square$

## 8. WINDOW PARAMETERS

**8.1.  $\varepsilon$ -maximality.** Let  $\underline{v} = (\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$ ,  $\underline{u} = (\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  be two regular chains such that  $\pi[\underline{v}] = \pi[\underline{u}]$ . We compare  $p_i^u$  to  $q_i^u$ , and  $p_i^s$  to  $q_i^s$ . The idea is to use regularity to see that the  $q$ -parameters of  $V^u[(v_i)_{i \leq 0}]$  and  $V^s[(v_i)_{i \geq 0}]$  are ‘‘almost maximal’’ in a certain sense that we describe below.

But first, some notation and terminology: (a) a positive or negative chain is called *regular*, if it can be completed to a regular chain (equiv. every coordinate is relevant, and some double chart appears infinitely many times); (b) if  $v$  is a double chart, then  $p^u(v)$  and  $p^s(v)$  means the  $p^u$  and  $p^s$  in  $v = \Psi_x^{p^u, p^s}$ .

**Definition 8.1.** A negative chain  $(v_i)_{i \leq 0}$  is called  $\varepsilon$ -maximal if it is regular, and

$$p^u(v_0) \geq e^{-\sqrt[3]{\varepsilon}} p^u(u_0)$$

for every regular chain  $(u_i)_{i \in \mathbb{Z}}$  for which there is a positive regular chain  $(v_i)_{i \geq 0}$  s.t.  $\pi[(v_i)_{i \in \mathbb{Z}}] = \pi[(u_i)_{i \in \mathbb{Z}}]$ .

**Definition 8.2.** A positive chain  $(v_i)_{i \geq 0}$  is called  $\varepsilon$ -maximal if it is regular, and

$$p^s(v_0) \geq e^{-\sqrt[3]{\varepsilon}} p^s(u_0)$$

for every regular chain  $(u_i)_{i \in \mathbb{Z}}$  for which there is a negative regular chain  $(v_i)_{i \leq 0}$  s.t.  $\pi[(v_i)_{i \in \mathbb{Z}}] = \pi[(u_i)_{i \in \mathbb{Z}}]$ .

**Proposition 8.3.** The following holds for all  $\varepsilon$  small enough: for every regular chain  $(v_i)_{i \in \mathbb{Z}}$ ,  $(v_i)_{i \leq 0}$  and  $(v_i)_{i \geq 0}$  are  $\varepsilon$ -maximal.

*Proof.* The proof is made of several steps.

*Step 1.* The following holds for all  $\varepsilon$  small enough: Let  $\underline{u}$  and  $\underline{v}$  be two regular chains s.t.  $\pi[\underline{u}] = \pi[\underline{v}]$ . If  $u_0 = \Psi_x^{p^u, p^s}$  and  $v_0 = \Psi_y^{q^u, q^s}$ , then  $Q_\varepsilon(x)/Q_\varepsilon(y) \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$ .

*Proof.* Propositions 6.5 and 7.3 say that  $\frac{\sin \alpha(x)}{\sin \alpha(y)} \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$ ,  $\frac{s_\chi(x)}{s_\chi(y)} \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}]$ , and  $\frac{u_\chi(x)}{u_\chi(y)} \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}]$ . By Lemma 2.4  $\frac{\|C_\chi(x)^{-1}\|_{Fr}}{\|C_\chi(y)^{-1}\|_{Fr}} \in [\exp(-5\sqrt{\varepsilon}), \exp(5\sqrt{\varepsilon})]$ , whence  $Q_\varepsilon(x)/Q_\varepsilon(y) \in [\exp(-\frac{60}{\beta}\sqrt{\varepsilon} - \frac{1}{3}\varepsilon), \exp(\frac{60}{\beta}\sqrt{\varepsilon} + \frac{1}{3}\varepsilon)]$ . If  $\varepsilon$  is small enough, then  $Q_\varepsilon(x)/Q_\varepsilon(y) \in [\exp(-\sqrt[3]{\varepsilon}), \exp(\sqrt[3]{\varepsilon})]$ .

*Step 2.* The following holds for all  $\varepsilon$  small enough: Every regular negative chain  $(v_i)_{i \leq 0}$  s.t.  $v_0 = \Psi_x^{p^u, p^s}$  where  $p^u = Q_\varepsilon(x)$  is  $\varepsilon$ -maximal, and every regular positive chain  $(v_i)_{i \geq 0}$  s.t.  $v_0 = \Psi_x^{p^u, p^s}$  where  $p^s = Q_\varepsilon(x)$  is  $\varepsilon$ -maximal.

*Proof.* Suppose  $(v_i)_{i \leq 0}$  is regular, and  $v_0 = \Psi_x^{p^u, p^s}$  where  $p^u = Q_\varepsilon(x)$ . We show that  $(v_i)_{i \leq 0}$  is  $\varepsilon$ -maximal.

Suppose  $(v_i)_{i \in \mathbb{Z}}$  is a regular extension of  $(v_i)_{i \leq 0}$  and let  $(u_i)_{i \in \mathbb{Z}}$  be some regular chain s.t.  $\pi[(u_i)_{i \in \mathbb{Z}}] = \pi[(v_i)_{i \in \mathbb{Z}}]$ . Write  $u_0 = \Psi_y^{q^u, q^s}$ . We have to show that  $p^u \geq e^{-\sqrt[3]{\varepsilon}} q^u$ . Indeed, by step 1,  $p^u = Q_\varepsilon(x) \geq e^{-\sqrt[3]{\varepsilon}} Q_\varepsilon(y) \geq e^{-\sqrt[3]{\varepsilon}} q^u$ .

The proof of the second half of step 2 is similar, and we therefore omit it.

*Step 3.* Let  $(v_i)_{i \leq 0}$  be a regular negative chain and suppose  $v_0 \rightarrow v_1$ . If  $(v_i)_{i \leq 0}$  is  $\varepsilon$ -maximal, then  $(v_i)_{i \leq 1}$  is  $\varepsilon$ -maximal. Let  $(v_i)_{i \geq 0}$  be a regular positive chain, and suppose  $v_{-1} \rightarrow v_0$ . If  $(v_i)_{i \geq 0}$  is  $\varepsilon$ -maximal, then  $(v_i)_{i \geq -1}$  is  $\varepsilon$ -maximal.

*Proof.* Let  $(v_i)_{i \leq 0}$  be an  $\varepsilon$ -maximal regular negative chain, and suppose  $v_0 \rightarrow v_1$ . We prove that  $(v_i)_{i \leq 1}$  is  $\varepsilon$ -maximal.

Suppose  $(u_i)_{i \in \mathbb{Z}}$ ,  $(v_i)_{i \leq 1}$  are regular and there is an extension of  $(v_i)_{i \leq 1}$  to a regular chain  $(v_i)_{i \in \mathbb{Z}}$  s.t.  $\pi[(v_{i+1})_{i \in \mathbb{Z}}] = \pi[(u_{i+1})_{i \in \mathbb{Z}}]$ . We write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$ ,  $u_i = \Psi_{y_i}^{q_i^u, q_i^s}$ , and show that  $p_1^u \geq e^{-\sqrt[3]{\varepsilon}} q_1^u$ .

Since  $\pi[(v_{i+1})_{i \in \mathbb{Z}}] = \pi[(u_{i+1})_{i \in \mathbb{Z}}]$  and  $\pi \circ \sigma = f \circ \pi$ ,  $\pi[(v_i)_{i \in \mathbb{Z}}] = \pi[(u_i)_{i \in \mathbb{Z}}]$ . Therefore, since  $(v_i)_{i \leq 0}$  is  $\varepsilon$ -maximal,  $p_0^u \geq e^{-\sqrt[3]{\varepsilon}} q_0^u$ . Also, by step 1,  $Q_\varepsilon(x_1) \geq e^{-\sqrt[3]{\varepsilon}} Q_\varepsilon(y_1)$ . It follows that

$$\begin{aligned} p_1^u &= \min\{e^\varepsilon p_0^u, Q_\varepsilon(x_1)\} \quad (\because v_0 \rightarrow v_1) \\ &\geq \min\{e^\varepsilon \cdot e^{-\sqrt[3]{\varepsilon}} q_0^u, e^{-\sqrt[3]{\varepsilon}} Q_\varepsilon(y_1)\} \\ &= e^{-\sqrt[3]{\varepsilon}} \min\{e^\varepsilon q_0^u, Q_\varepsilon(y_1)\} = e^{-\sqrt[3]{\varepsilon}} q_1^u \quad (\because u_0 \rightarrow u_1). \end{aligned}$$

This proves the part of step 3 dealing with negative chains. The case of positive chains is similar, and we leave it to the reader.

*Step 4.* Proof of the proposition.

Suppose  $(v_i)_{i \in \mathbb{Z}}$  is a regular chain, and write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$ . Since  $(v_i)_{i \in \mathbb{Z}}$  is a chain,  $\{(p_i^u, p_i^s)\}_{i \in \mathbb{Z}}$  is  $\varepsilon$ -subordinated to  $\{Q_\varepsilon(x_i)\}_{i \in \mathbb{Z}}$ . Since  $(v_i)_{i \in \mathbb{Z}}$  is regular,  $\limsup_{i \rightarrow \pm\infty} (p_i^u \wedge p_i^s) > 0$ , therefore by Lemma 4.7,  $p_n^u = Q_\varepsilon(x_n)$  for some  $n < 0$  and  $p_\ell^s = Q_\varepsilon(x_\ell)$  for some  $\ell > 0$ .

By step 2,  $(v_i)_{i \leq n}$  is an  $\varepsilon$ -maximal negative chain, and  $(v_i)_{i \geq \ell}$  is an  $\varepsilon$ -maximal positive chain.

By step 3,  $(v_i)_{i \leq 0}$  is an  $\varepsilon$ -maximal negative chain, and  $(v_i)_{i \geq 0}$  is an  $\varepsilon$ -maximal positive chain.  $\square$

**8.2. Comparison of  $p_i^{u/s}$  to  $q_i^{u/s}$ .** We can now easily compare the window parameters of all regular chains with the same  $\pi$  image.

**Proposition 8.4.** *Let  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$  and  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  be two regular chains such that  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$ , then  $p_i^u/q_i^u, p_i^s/q_i^s \in [\exp(-\sqrt[3]{\varepsilon}), \exp(\sqrt[3]{\varepsilon})]$  for all  $i \in \mathbb{Z}$ .*

*Proof.* By Proposition 8.3  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \leq 0}$  is  $\varepsilon$ -maximal, so  $p_0^u \geq e^{-\sqrt[3]{\varepsilon}} q_0^u$ .  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \leq 0}$  is also  $\varepsilon$ -maximal, so  $q_0^u \geq e^{-\sqrt[3]{\varepsilon}} p_0^u$ . It follows that  $p_0^u/q_0^u \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$ . Similarly,  $p_0^s/q_0^s \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$ .

Working with the shifted sequences  $(\Psi_{x_{i+k}}^{p_{i+k}^u, p_{i+k}^s})_{i \in \mathbb{Z}}$  and  $(\Psi_{y_{i+k}}^{q_{i+k}^u, q_{i+k}^s})_{i \in \mathbb{Z}}$ , we obtain  $p_k^s/q_k^s, p_k^u/q_k^u \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$ .  $\square$

## 9. PROOF OF THEOREM 5.2

Parts (1) and (3) of the theorem are handled by Propositions 5.3 and 8.4, so we focus on part (2).

Suppose  $\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$  where  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$  and  $(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$  are regular chains. We compare  $\Psi_{x_i}$  and  $\Psi_{y_i}$ . Write, as in §5,  $\Psi_{x_i} = \exp_{x_i} \circ \vartheta_{x_i} \circ C_{x_i}$  and  $\Psi_{y_i} = \exp_{y_i} \circ \vartheta_{y_i} \circ C_{y_i}$ . We also let  $p_i := p_i^u \wedge p_i^s$  and  $q_i := q_i^u \wedge q_i^s$ .

*Claim 1.*  $C_{y_i}^{-1} C_{x_i} = (-1)^{\sigma_i} \text{Id} + E$  where  $\sigma_i \in \{0, 1\}$  and  $E$  is a matrix all of whose entries have absolute value less than  $7\sqrt{\varepsilon}$ .

*Proof.* By (5.2) and Proposition 6.7,

$$\begin{aligned} C_{y_i}^{-1} C_{x_i} &= \begin{pmatrix} s_\chi(y_i) & -\frac{s_\chi(y_i)}{\tan \alpha(y_i)} \\ 0 & \frac{u_\chi(y_i)}{\sin \alpha(y_i)} \end{pmatrix} R_{y_i}^{-1} R_{x_i} \begin{pmatrix} s_\chi(x_i)^{-1} & u_\chi(x_i)^{-1} \cos \alpha(x_i) \\ 0 & u_\chi(x_i)^{-1} \sin \alpha(x_i) \end{pmatrix} \\ &= \begin{pmatrix} s_\chi(y_i) & -\frac{s_\chi(y_i)}{\tan \alpha(y_i)} \\ 0 & \frac{u_\chi(y_i)}{\sin \alpha(y_i)} \end{pmatrix} [(-1)^{\sigma_i} \text{Id} + E'] \begin{pmatrix} s_\chi(x_i)^{-1} & u_\chi(x_i)^{-1} \cos \alpha(x_i) \\ 0 & u_\chi(x_i)^{-1} \sin \alpha(x_i) \end{pmatrix}, \end{aligned}$$

where  $\sigma_i \in \{0, 1\}$  and  $E' = (\varepsilon_{ij})_{2 \times 2}$  and  $|\varepsilon_{ij}| < p_i^{\beta/5} + q_i^{\beta/5} < \sqrt{\varepsilon}$ .

We call the contribution of  $(-1)^{\sigma_i} \text{Id}$  the ‘‘main term’’, and the contribution of  $E'$ , the ‘‘error term’’.

*Main term:* This equals  $(-1)^{\sigma_i} \begin{pmatrix} \frac{s_\chi(y_i)}{s_\chi(x_i)} & \frac{s_\chi(y_i) \sin[\alpha(y_i) - \alpha(x_i)]}{u_\chi(x_i) \sin \alpha(y_i)} \\ 0 & \frac{u_\chi(y_i) \sin \alpha(x_i)}{u_\chi(x_i) \sin \alpha(y_i)} \end{pmatrix}$ .

Proposition 7.3 says that  $\frac{s_\chi(y_i)}{s_\chi(x_i)}$  and  $\frac{u_\chi(y_i)}{u_\chi(x_i)}$  belong to  $[\exp(-4\sqrt{\varepsilon}), \exp(4\sqrt{\varepsilon})]$ , and Proposition 6.5 says that  $\frac{\sin \alpha(x_i)}{\sin \alpha(y_i)} \in [\exp(-\sqrt{\varepsilon}), \exp \sqrt{\varepsilon}]$ . It follows that the (1,1) and (2,2) terms of the main term are, up to a sign  $(-1)^{\sigma_i}$ , in  $[\exp(-5\sqrt{\varepsilon}), \exp(5\sqrt{\varepsilon})]$ .

We bound the (1,2) term: Since  $u_\chi(y_i) \geq \sqrt{2} > 1$  and  $\frac{s_\chi(y_i)}{|\sin \alpha(y_i)|} < \|C_\chi(y_i)^{-1}\|_{Fr}$  (Lemma (2.4)),

$$\begin{aligned} \left| \frac{s_\chi(y_i) \sin[\alpha(y_i) - \alpha(x_i)]}{u_\chi(x_i) \sin \alpha(y_i)} \right| &\leq \|C_\chi(y_i)^{-1}\|_{Fr} \cdot |\sin(\alpha(y_i) - \alpha(x_i))| \\ &\leq \|C_\chi(y_i)^{-1}\|_{Fr} \cdot (|\sin \alpha(y_i) - \sin \alpha(x_i)| + |\cos \alpha(y_i) - \cos \alpha(x_i)|). \end{aligned}$$

By Lemma 6.6, if  $\varepsilon$  is small enough,

$$\left| \frac{s_\chi(y_i) \sin[\alpha(y_i) - \alpha(x_i)]}{u_\chi(x_i) \sin \alpha(y_i)} \right| \leq \|C_\chi(y_i)^{-1}\|_{Fr} \cdot 6(p_i^{\beta/4} + q_i^{\beta/4}).$$

By Proposition 8.4,  $p_i \leq e^{\sqrt[3]{\varepsilon}} q_i$ , therefore

$$p_i^{\beta/4} + q_i^{\beta/4} < (e^{\sqrt[3]{\varepsilon}} + 1) q_i^{\beta/4} < 2q_i^{\beta/4} < 2Q_\varepsilon(y_i)^{\beta/4} < 2\varepsilon^{3/4} \|C_\chi(y_i)^{-1}\|_{Fr}^{-3}.$$

Since  $\|C_\chi(\cdot)^{-1}\|_{Fr} > 1$ ,  $\left| \frac{s_\chi(y_i) \sin[\alpha(y_i) - \alpha(x_i)]}{u_\chi(x_i) \sin \alpha(y_i)} \right| < \sqrt{\varepsilon}$ , for all  $\varepsilon$  small enough. We see that the main term equals  $(-1)^{\sigma_i} \text{Id} + (m_{ij})_{2 \times 2}$  where  $|m_{ij}| < 6\sqrt{\varepsilon}$ .

*Error term:* This is

$$\begin{pmatrix} s_\chi(y_i) & -\frac{s_\chi(y_i)}{\tan \alpha(y_i)} \\ 0 & \frac{u_\chi(y_i)}{\sin \alpha(y_i)} \end{pmatrix} \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{pmatrix} \begin{pmatrix} s_\chi(x_i)^{-1} & u_\chi(x_i)^{-1} \cos \alpha(x_i) \\ 0 & u_\chi(x_i)^{-1} \sin \alpha(x_i) \end{pmatrix}.$$



Every entry of the product matrix is the sum of four products, each consisting of three terms, one for each matrix.

The term from the left matrix is bounded by  $\|C_\chi(y_i)^{-1}\|_{Fr}$  (Lemma 2.4). The term from the middle matrix is bounded by

$$p_i^{\beta/5} + q_i^{\beta/5} < q_i^{\beta/5}(1 + e^{\sqrt[3]{\varepsilon}\beta/5}) < 2Q_\varepsilon(y_i)^{\beta/5}.$$

The term from the right matrix is bounded by one. The product of these terms is bounded by  $4\|C_\chi(y_i)^{-1}\|_{Fr} \cdot 2Q_\varepsilon(y_i)^{\beta/5} \cdot 1$ . By the definition of  $Q_\varepsilon(y_i)$ , this is less than  $8\varepsilon^{3/5} < \sqrt{\varepsilon}$ .

Combining the two estimates we see that every entry of  $C_{y_i}^{-1}C_{x_i} - (-1)^{\sigma_i} \text{Id}$  is less than  $7\sqrt{\varepsilon}$  in absolute value.

*Claim 2.*  $\Psi_{y_i}^{-1} \circ \Psi_{x_i}$  is well defined on  $R_\varepsilon(\underline{0})$ .

*Proof.* We use the constants  $L_1, \dots, L_4$  introduced in the proof of Proposition 3.2, and the ball notation of §2.3. We assume that  $\varepsilon$  satisfies (3.2).

Suppose  $\underline{v} \in R_\varepsilon(\underline{0})$ . By Proposition 5.3,  $d(x_i, y_i) < 25^{-1}(p_i + q_i)$ , and by Proposition 8.4,  $p_i \leq e^{\sqrt[3]{\varepsilon}}q_i$ , so  $d(x_i, y_i) < q_i$ . By the definition of  $L_1$  (page 13),

$$d((\exp_{x_i} \circ \vartheta_{x_i})(C_{x_i}\underline{v}), (\exp_{y_i} \circ \vartheta_{y_i})(C_{x_i}\underline{v})) \leq L_1 d(x_i, y_i) < L_1 q_i.$$

Therefore,  $\Psi_{x_i}(\underline{v}) \in B := B_{L_1 q_i}(\exp_{y_i} \circ \vartheta_{y_i}(C_{x_i}\underline{v}))$ .

As in the proof of Proposition 3.2,  $\exp_{y_i}^{-1}$  is well defined on  $B$ , and has Lipschitz constant at most  $L_3$  there, so

$$\exp_{y_i}^{-1}(B) \subset B_{L_1 L_3 q_i}^{y_i}(\vartheta_{y_i}(C_{x_i}\underline{v})).$$

It follows that  $\Psi_{x_i}(\underline{v}) \in \exp_{y_i}[\exp_{y_i}^{-1}(B)] \subset \exp_{y_i}[B_{L_1 L_3 q_i}^{y_i}(\vartheta_{y_i}(C_{x_i}\underline{v}))] \equiv \Psi_{y_i}[E]$ , where  $E := C_\chi(y_i)^{-1}[B_{L_1 L_3 q_i}^{y_i}(\vartheta_{y_i}(C_{x_i}\underline{v}))] \subset B_{L_1 L_3 \|C_{y_i}^{-1}\|_{q_i}}(C_{y_i}^{-1}C_{x_i}\underline{v})$ .

We now use the inequalities  $q_i \leq Q_\varepsilon(y_i) < \varepsilon^{3/\beta} \|C_\chi(y_i)^{-1}\|^{-1}$  and (claim 1)

$$\|C_{y_i}^{-1}C_{x_i} - (-1)^{\sigma_i} \text{Id}\| \leq \|C_{y_i}^{-1}C_{x_i} - (-1)^{\sigma_i} \text{Id}\|_{Fr} < 14\sqrt{\varepsilon}.$$

These give  $E \subset B_{L_1 L_3 \varepsilon^{3/\beta} + 14\sqrt{\varepsilon}\|\underline{v}\|}((-1)^{\sigma_i}\underline{v}) \subset B_{L_1 L_3 \varepsilon^{3/\beta} + 14\sqrt{\varepsilon}\|\underline{v}\| + \|\underline{v}\|}(\underline{0})$ . Since  $\underline{v} \in R_\varepsilon(\underline{0})$ , for all  $\varepsilon$  small enough

$$L_1 L_3 \varepsilon^{3/\beta} + 14\sqrt{\varepsilon}\|\underline{v}\| + \|\underline{v}\| < (L_1 L_2 \varepsilon^2 + 14\sqrt{\varepsilon} + 1)\sqrt{2\varepsilon} < 2\varepsilon < r(M),$$

where  $r(M)$  is given in (2.1). It follows that  $E \subset B_{r(M)}(\underline{0})$ .

We just showed that for every  $\underline{v} \in R_\varepsilon(\underline{0})$ ,  $\Psi_{x_i}(\underline{v}) \in \Psi_{y_i}[B_{r(M)}(\underline{0})]$ . In other words,  $\Psi_{x_i}[R_\varepsilon(\underline{0})] \subset \Psi_{y_i}[B_{r(M)}(\underline{0})]$ . By the definition of  $r(M)$ ,  $\Psi_{y_i} : B_{r(M)}(\underline{0}) \rightarrow M$  is a diffeomorphism onto its image. It follows that  $\Psi_{y_i}^{-1} \circ \Psi_{x_i}$  is well defined and smooth on  $R_\varepsilon(\underline{0})$ .

*Claim 3.*  $\Psi_{y_i}^{-1} \circ \Psi_{x_i}(\underline{v}) = (-1)^{\sigma_i}\underline{v} + \underline{c}_i + \Delta_i(\underline{v})$  where  $\sigma_i \in \{0, 1\}$ ,  $\underline{c}_i$  is a constant vector s.t.  $\|\underline{c}_i\| < 10^{-1}q_i$ , and  $\Delta_i(\cdot)$  is a vector field s.t.  $\Delta_i(\underline{0}) = \underline{0}$  and  $\|(d\Delta_i)_\underline{v}\| < \sqrt[3]{\varepsilon}$  on  $R_\varepsilon(\underline{0})$ .

*Proof.* Choose  $\sigma_i$  as in claim 1. One can always put  $\Psi_{y_i}^{-1} \circ \Psi_{x_i}$  in the form

$$\Psi_{y_i}^{-1} \circ \Psi_{x_i}(\underline{v}) = (-1)^{\sigma_i}\underline{v} + \underline{c}_i + \Delta_i(\underline{v})$$

where  $\underline{c}_i := (\Psi_{y_i}^{-1} \circ \Psi_{x_i})(\underline{0})$  and  $\Delta_i(\underline{v}) := (\Psi_{y_i}^{-1} \circ \Psi_{x_i})(\underline{v}) - (\Psi_{y_i}^{-1} \circ \Psi_{x_i})(\underline{0}) - (-1)^{\sigma_i} \underline{v}$ .

$$\begin{aligned} \Delta_i(\underline{v}) &= [C_{y_i}^{-1} \vartheta_{y_i}^{-1} \exp_{y_i}^{-1} \exp_{x_i} \vartheta_{x_i} C_{x_i}](\underline{v}) - \underline{c}_i - (-1)^{\sigma_i} \underline{v} \\ &= C_{y_i}^{-1} (\vartheta_{y_i}^{-1} \exp_{y_i}^{-1} \exp_{x_i} \vartheta_{x_i} - \text{Id}) C_{x_i} \underline{v} + (C_{y_i}^{-1} C_{x_i} - (-1)^{\sigma_i} \text{Id}) \underline{v} - \underline{c}_i \\ &= C_{y_i}^{-1} (\vartheta_{y_i}^{-1} \exp_{y_i}^{-1} - \vartheta_{x_i}^{-1} \exp_{x_i}^{-1}) (\Psi_{x_i}(\underline{v})) + (C_{y_i}^{-1} C_{x_i} - (-1)^{\sigma_i} \text{Id}) \underline{v} - \underline{c}_i. \end{aligned}$$

It is clear that  $\Delta_i(\underline{0}) = \underline{0}$ , and that for all  $\underline{v} \in R_\varepsilon(\underline{0})$

$$\begin{aligned} \|(d\Delta_i)_{\underline{v}}\| &\leq \|C_{y_i}^{-1}\| \cdot \|d(\vartheta_{y_i}^{-1} \exp_{y_i}^{-1})_{\Psi_{x_i}(\underline{v})} - d(\vartheta_{x_i}^{-1} \exp_{x_i}^{-1})_{\Psi_{x_i}(\underline{v})}\| \| (d\Psi_{x_i})_{\underline{v}} \| \\ &\quad + \|C_{y_i}^{-1} C_{x_i} - (-1)^{\sigma_i} \text{Id}\| \\ &\leq 2 \|C_{y_i}^{-1}\| \cdot \|d(\vartheta_{y_i}^{-1} \exp_{y_i}^{-1})_{\Psi_{x_i}(\underline{v})} - d(\vartheta_{x_i}^{-1} \exp_{x_i}^{-1})_{\Psi_{x_i}(\underline{v})}\| + 14\sqrt{\varepsilon} \\ &\leq 2 \|C_{y_i}^{-1}\| \cdot L_2 d(x_i, y_i) + 14\sqrt{\varepsilon}, \end{aligned}$$

where  $L_2$  is a common Lipschitz constant for the maps  $x \mapsto \vartheta_x^{-1} \exp_x^{-1}$  from  $D$  to  $C^2(D, \mathbb{R}^2)$  ( $D \in \mathcal{D}$ ). As we saw above,  $d(x_i, y_i) < q_i < \varepsilon^{3/\beta} \|C_{y_i}^{-1}\|^{-1}$ , whence

$$\|(d\Delta_i)_{\underline{v}}\| \leq 2L_2 \varepsilon^{3/\beta} + 14\sqrt{\varepsilon}.$$

This is smaller than  $\sqrt[3]{\varepsilon}$  for all  $\varepsilon$  small enough.

Finally we estimate  $\underline{c}_i$ . Let  $z := f^i(\pi[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}]) = f^i(\pi[(\Psi_{y_i}^{q_i^u, q_i^s})_{i \in \mathbb{Z}}])$ . This is the intersection of a  $u$ -admissible manifold and an  $s$ -admissible manifold in  $\Psi_{x_i}^{p_i^u, p_i^s}$ , therefore by Proposition 4.11,  $f^i(z) = \Psi_{x_i}^{p_i^u, p_i^s}(\underline{\zeta})$ , for some  $\underline{\zeta} \in R_{10^{-2}p_i}(\underline{0})$ . Similarly,  $f^i(z) = \Psi_{y_i}^{q_i^u, q_i^s}(\underline{\eta})$ , for some  $\underline{\eta} \in R_{10^{-2}q_i}(\underline{0})$ . It follows that

$$\underline{\eta} = (\Psi_{y_i}^{-1} \circ \Psi_{x_i})(\underline{\zeta}) = (-1)^{\sigma_i} \underline{\zeta} + \underline{c}_i + \Delta_i(\underline{\zeta}),$$

and consequently  $\|\underline{c}_i\| \leq \|\underline{\eta}\| + \|\underline{\zeta}\| + \|\Delta_i(\underline{\zeta})\|$ .

Now  $\|\underline{\zeta}\| < 10^{-2} \sqrt{2} p_i < 10^{-2} \sqrt{2} e^{\sqrt[3]{\varepsilon}} q_i$ ,  $\|\underline{\eta}\| < 10^{-2} \sqrt{2} q_i$ , and by the bound on  $\|d\Delta_i\|$ ,  $\|\Delta_i(\underline{\zeta})\| \leq \sqrt[3]{\varepsilon} \|\underline{\zeta}\|$ . It follows that  $\|\underline{c}_i\| < 10^{-1} q_i$ .  $\square$

### Part 3. Markov partitions and symbolic dynamics

#### 10. A LOCALLY FINITE COUNTABLE MARKOV COVER

**10.1. The cover.** In §4 we constructed a countable Markov shift  $\Sigma$  with countable alphabet  $\mathcal{V}$ , and a Hölder continuous map  $\pi : \Sigma \rightarrow M$  which commutes with the left shift  $\sigma : \Sigma \rightarrow \Sigma$ , so that  $\pi(\Sigma)$  has full measure w.r.t. any ergodic invariant probability measure with entropy larger than  $\chi$ . Moreover, if<sup>6</sup>

$$\begin{aligned} \Sigma^\# &= \{\underline{v} \in \Sigma : \underline{v} \text{ is a regular chain}\} \\ &= \{\underline{v} \in \Sigma : \exists v, w \in \mathcal{V} \exists n_k, m_k \uparrow \infty \text{ s.t. } v_{n_k} = v, v_{-m_k} = w\}, \end{aligned}$$

then  $\pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f)$ , therefore  $\pi(\Sigma^\#)$  has full probability w.r.t. any ergodic invariant probability measure with entropy larger than  $\chi$ .

In this section we study the following countable cover of  $\text{NUH}_\chi^\#(f)$ :

**Definition 10.1.**  $\mathcal{Z} := \{Z(v) : v \in \mathcal{V}\}$ , where  $Z(v) := \{\pi(\underline{v}) : \underline{v} \in \Sigma^\#, v_0 = v\}$ .

This is a cover of  $\text{NUH}_\chi^\#(f)$ . The following property of  $\mathcal{Z}$  is the hinge on which our entire approach turns (see §1.5):

<sup>6</sup>This uses the convention from §4.4 that every element of  $\mathcal{V}$  is relevant.

**Theorem 10.2.** *For every  $Z \in \mathcal{Z}$ ,  $|\{Z' \in \mathcal{Z} : Z' \cap Z \neq \emptyset\}| < \infty$ .*

*Proof.* Fix some  $Z = Z(\Psi_x^{p^u, p^s})$ . If  $Z' = Z(\Psi_y^{q^u, q^s})$  intersects  $Z$ , then there must exist two chains  $\underline{v}, \underline{w} \in \Sigma^\#$  s.t.  $v_0 = \Psi_x^{p^u, p^s}$ ,  $w_0 = \Psi_y^{q^u, q^s}$ , and  $\pi(\underline{v}) = \pi(\underline{w})$ . Proposition 8.4 says that in this case

$$q^u \geq e^{-\sqrt[3]{\varepsilon}} p^u \text{ and } q^s \geq e^{-\sqrt[3]{\varepsilon}} p^s.$$

It follows that  $Z'$  belongs to  $\{Z(\Psi_y^{q^u, q^s}) : \Psi_y^{q^u, q^s} \in \mathcal{V}, q^u \wedge q^s \geq e^{-\sqrt[3]{\varepsilon}}(p^u \wedge p^s)\}$ . By the definition of  $\mathcal{V}$ , this set has cardinality less than or equal to

$$|\{\Psi_y^\eta \in \mathcal{A} : \eta \geq e^{-\sqrt[3]{\varepsilon}}(p^u \wedge p^s)\}| \times |\{(q^u, q^s) \in I_\varepsilon \times I_\varepsilon : q^u \wedge q^s \geq e^{-\sqrt[3]{\varepsilon}}(p^u \wedge p^s)\}|.$$

This is a finite number, because of the discreteness of  $\mathcal{A}$  (Proposition 3.5).  $\square$

**10.2. Product structure.** Suppose  $x \in Z(v) \in \mathcal{Z}$ , then  $\exists \underline{v} \in \Sigma^\#$  s.t.  $v_0 = v$  and  $\pi(\underline{v}) = x$ . Associated to  $\underline{v}$  are two admissible manifolds in  $v$ :  $V^s[(v_i)_{i \leq 0}]$  and  $V^u[(v_i)_{i \geq 0}]$  (Proposition 4.15). These manifolds do not depend on the choice of  $\underline{v}$ : if  $\underline{w} \in \Sigma^\#$  is another chain s.t.  $w_0 = v$  and  $\pi(\underline{w}) = x$ , then

$$V^u[(w_i)_{i \leq 0}] = V^u[(v_i)_{i \leq 0}] \text{ and } V^s[(w_i)_{i \geq 0}] = V^s[(v_i)_{i \geq 0}],$$

because of Proposition 6.4 and the equalities  $p^{u/s}(w_0) = p^{u/s}(v_0) = p^{u/s}(v)$ . We are therefore free to make the following definition:

**Definition 10.3.** *Suppose  $Z = Z(v) \in \mathcal{Z}$ . For any  $x \in Z$ :*

- (1)  $V^s(x, Z) := V^s[(v_i)_{i \geq 0}]$  for some (every)  $\underline{v} \in \Sigma^\#$  s.t.  $v_0 = v$  and  $\pi(\underline{v}) = x$ .  
 $W^s(x, Z) := V^s(x, Z) \cap Z$ .
- (2)  $V^u(x, Z) := V^u[(v_i)_{i \leq 0}]$  for some (every)  $\underline{v} \in \Sigma^\#$  s.t.  $v_0 = v$  and  $\pi(\underline{v}) = x$ .  
 $W^u(x, Z) := V^u(x, Z) \cap Z$ .

It is important to understand the difference between  $V^{s/u}(x, Z)$  and  $W^{s/u}(x, Z)$ . Whereas  $V^{u/s}(x, Z)$  are smooth manifolds,  $W^{u/s}(x, Z)$  could in principle be totally disconnected. Whereas  $V^{u/s}(x, Z)$  extend all the way across  $\Psi_x[R_{p^{u/s}}(\mathbb{Q})]$  (assuming  $v = \Psi_x^{p^u, p^s}$ ),  $W^{u/s}(x, Z)$  are subsets of the much smaller set  $\Psi_x[R_{10^{-2}(p^u \wedge p^s)}(\mathbb{Q})]$ , because every point in  $W^{u/s}(x, Z)$  is the intersection of an  $s$ -admissible manifold in  $v$  and a  $u$ -admissible manifold in  $v$  (Proposition 4.11).

**Proposition 10.4.** *Suppose  $Z \in \mathcal{Z}$ . For every  $x, y \in Z$ ,  $V^u(x, Z)$  and  $V^u(y, Z)$  are either equal or they are disjoint. Similarly for  $V^s(x, Z)$  and  $V^s(y, Z)$ , for  $W^u(x, Z)$  and  $W^u(y, Z)$ , and for  $W^s(x, Z)$  and  $W^s(y, Z)$ .*

*Proof.* The statement holds for  $V^{u/s}$  because of Proposition 6.4. The statement for  $W^{u/s}$  is an immediate corollary.  $\square$

**Proposition 10.5.** *Suppose  $Z \in \mathcal{Z}$  and  $x, y \in Z$ , then  $V^u(x, Z)$  and  $V^s(y, Z)$  intersect at a unique point  $z$ , and  $z \in Z$ . Thus  $W^u(x, Z) \cap W^s(y, Z) = \{z\}$ .*

*Proof.* Write  $Z = Z(v)$  where  $v \in \mathcal{V}$ .  $V^u(x, Z)$  is a  $u$ -admissible manifold in  $v$ , and  $V^s(y, Z)$  is an  $s$ -admissible manifold in  $v$ . Consequently,  $V^u(x, Z)$  and  $V^s(y, Z)$  intersect at a unique point  $z$  (Proposition 4.11).

We claim that  $z \in Z$ . There are chains  $\underline{v}, \underline{w} \in \Sigma^\#$  s.t.  $v_0 = w_0 = v$  and so that  $V^u(x, Z) = V^u[(v_i)_{i \leq 0}]$  and  $V^s(y, Z) = V^s[(w_i)_{i \geq 0}]$ . Define  $\underline{u} = (u_i)_{i \in \mathbb{Z}}$  by

$$u_i = \begin{cases} v_i & i \leq 0 \\ w_i & i \geq 0 \end{cases}.$$

It is easy to see that  $\underline{u} \in \Sigma^\#$  and  $u_0 = v$ , therefore  $\pi(\underline{u}) \in Z$ . By definition,

$$\{\pi(\underline{u})\} = V^u[(u_i)_{i \leq 0}] \cap V^s[(u_i)_{i \geq 0}] = V^u[(v_i)_{i \leq 0}] \cap V^s[(w_i)_{i \geq 0}] = V^u(x, Z) \cap V^s(y, Z).$$

It follows that  $z = \pi(\underline{u}) \in Z$ .  $\square$

**Definition 10.6.** *The Smale bracket of two points  $x, y \in Z \in \mathcal{Z}$  is the unique point  $[x, y]_Z \in W^u(x, Z) \cap W^s(y, Z)$ .*

This definition is motivated by [Sm] (see also [B4, chapter 3]).

**Lemma 10.7.** *Suppose  $x, y \in Z(v_0)$  and  $f(x), f(y) \in Z(v_1)$ . If  $v_0 \rightarrow v_1$ , then  $f([x, y]_{Z(v_0)}) = [f(x), f(y)]_{Z(v_1)}$ .*

*Proof.* Write  $Y = Z(v_0)$ ,  $Z = Z(v_1)$ , and  $w := [x, y]_Y$ . By definition

$$\{f(w)\} = f[W^u(x, Y) \cap W^s(y, Y)] \subset f[V^u(x, Y)] \cap f[V^s(y, Y)]. \quad (10.1)$$

*Claim:*  $f[V^s(y, Y)] \subset V^s(f(y), Z)$  and  $f[V^u(x, Y)] \supset V^u(f(x), Z)$ .

*Proof.* Since  $f(y) \in Z(v_1) = Z$ ,  $V^s := V^s(f(y), Z)$  is an  $s$ -admissible manifold in  $v_1$ , and this manifold stays in windows. Applying the graph transform (Proposition 4.12) we see that  $f^{-1}[V^s(f(y), Z)]$  contains an  $s$ -admissible manifold  $\mathcal{F}_s[V^s]$  in  $v_0$ . Since  $V^s$  stays in windows,  $\mathcal{F}_s[V^s]$  stays in windows.

Since  $\mathcal{F}_s[V^s]$  is  $s$ -admissible in  $v_0$ , it intersects every  $u$ -admissible manifold in  $v_0$ . The larger set  $f^{-1}(V^s)$  intersects  $V^u(y, Y)$  at a unique point (Proposition 4.12 (2)). This point must be  $y$ , so  $\mathcal{F}_s[V^s] \cap V^u(y, Y) = \{y\}$ , whence  $\mathcal{F}_s[V^s] \ni y$ .

This means that  $\mathcal{F}_s[V^s]$  intersects  $V^s(y, Y)$ . These manifolds are  $s$ -admissible in  $v_0$ , and they stay in windows. Since they intersect, they are equal. It follows that  $f^{-1}(V^s) \supset \mathcal{F}_s[V^s] = V^s(y, Y)$ , whence  $f[V^s(y, Y)] \subset V^s$ , which is the first half of the claim. The other half of the claim is proved in the same way.

Returning to (10.1) we see that  $f(w) \in f[V^u(x, Y)] \cap V^s(f(y), Z)$ . By the second half of the claim,

$$f[V^u(x, Y)] \cap V^s(f(y), Z) \supseteq V^u(f(x), Z) \cap V^s(f(y), Z) \ni \{[f(x), f(y)]_Z\},$$

thus  $f[V^u(x, Y)] \cap V^s(f(y), Z) \ni f(w), [f(x), f(y)]_Z$ . But Proposition 4.12 part (2) says that  $f[V^u(x, Y)]$  intersects  $V^s(f(y), Z)$  at a single point. It follows that  $f(w) = [f(x), f(y)]_Z$ .  $\square$

Occasionally we will need to form the Smale bracket of points belonging to different elements of  $\mathcal{Z}$ :

**Lemma 10.8.** *The following holds for all  $\varepsilon$  small enough: Suppose  $Z, Z' \in \mathcal{Z}$ . If  $Z \cap Z' \neq \emptyset$ , then for any  $x \in Z$  and  $y \in Z'$ ,  $V^u(x, Z)$  and  $V^s(y, Z')$  intersect at a unique point.*

We do not claim that this point is in  $Z$  or  $Z'$ . The proof is in the appendix.

### 10.3. The symbolic Markov property.

**Proposition 10.9.** *If  $x = \pi[(v_i)_{i \in \mathbb{Z}}]$  where  $\underline{v} \in \Sigma^\#$ , then  $f[W^s(x, Z(v_0))] \subset W^s(f(x), Z(v_1))$  and  $f^{-1}[W^u(f(x), Z(v_1))] \subset W^u(x, Z(v_0))$ .*

*Proof.* We prove the inclusion for the  $s$ -manifolds. The case of  $u$ -manifolds follows by symmetry.

*Step 1.*  $f[W^s(x, Z(v_0))] \subset V^s(f(x), Z(v_1))$ .

By definition,  $W^s(x, Z(v_0)) \subset V^s(x, Z(v_0)) \equiv V^s[(v_i)_{i \geq 0}]$ . By Proposition 4.15,  $f(V^s[(v_i)_{i \geq 0}]) \subseteq V^s[(v_{i+1})_{i \geq 0}]$ . Since  $f(x) = \pi[(v_{i+1})_{i \in \mathbb{Z}}]$ , the last manifold is equal to  $V^s(f(x), Z(v_1))$ . Thus  $f[W^s(x, Z(v_0))] \subset V^s(f(x), Z(v_1))$ .

*Step 2.*  $f[W^s(x, Z(v_0))] \subset Z(v_1)$ .

Suppose  $y \in W^s(x, Z(v_0))$ .

- Since  $y \in Z(v_0)$ ,  $y \in \Psi_{x_0}[R_{10^{-2}(p_0^u \wedge p_0^s)}(\mathbb{0})]$  (it is the intersection of a  $u$  and an  $s$ -admissible manifold in  $v_0$ ).
- Since  $y \in V^s[(v_i)_{i \geq 0}]$ ,  $f^k(y) \in V^s[(v_{i+k})_{i \geq 0}] \subset \Psi_{x_k}[R_{Q_\varepsilon(x_k)}(\mathbb{0})]$  for all  $k > 0$ , where  $v_k = \Psi_{x_k}^{p_k^u, p_k^s}$ .
- Since  $y \in Z(v_0)$ ,  $\exists \underline{w} \in \Sigma^\#$  s.t.  $w_0 = v_0$  and  $y = \pi(\underline{w}) \in V^u[(w_i)_{i \leq 0}]$ . It follows that  $f^{-k}(y) \in V^u[(w_{i-k})_{i \leq 0}] \subset \Psi_{y-k}[R_{Q_\varepsilon(y-k)}(\mathbb{0})]$  for all  $k \geq 0$ , where  $w_i = \Psi_{y_i}^{q_i^u, q_i^s}$ .

Writing  $u_i = \begin{cases} w_i & i \leq 0 \\ v_i & i > 0 \end{cases}$  and  $u_i = \Psi_{z_i}^{r_i^u, r_i^s}$ , we see that  $\underline{u} \in \Sigma^\#$ ,  $u_0 = v_0$ ,  $y \in$

$\Psi_{z_0}[R_{p_0^u \wedge p_0^s}(\mathbb{0})]$ , and  $f^k(y) \in \Psi_{z_k}[R_{Q_\varepsilon(z_k)}(\mathbb{0})]$  for all  $k \in \mathbb{Z}$ . By Proposition 4.15 part (4),  $y = \pi(\underline{u})$ . It follows that  $f(y) = \pi[\sigma(\underline{u})] \in Z(u_1) \equiv Z(v_1)$ .  $\square$

**Lemma 10.10.** *Suppose  $Z, Z' \in \mathcal{Z}$  and  $Z \cap Z' \neq \emptyset$ .*

- (1) *If  $Z = Z(\Psi_{x_0}^{p_0^u, p_0^s})$  and  $Z' = Z(\Psi_{y_0}^{q_0^u, q_0^s})$ , then  $Z \subset \Psi_{y_0}[R_{q_0^u \wedge q_0^s}(\mathbb{0})]$ .*
- (2) *For any  $x \in Z \cap Z'$ ,  $W^u(x, Z) \subset V^u(x, Z')$  and  $W^s(x, Z) \subset V^s(x, Z')$ .*

See the appendix for the proof.

## 11. A COUNTABLE MARKOV PARTITION

In the previous section we described a locally finite countable cover  $\mathcal{Z}$  of  $\text{NUH}_\chi^\#(f)$  by sets equipped with a Smale bracket and satisfying the symbolic Markov property. Here we produce a pairwise disjoint cover of  $\text{NUH}_\chi^\#(f)$  with similar properties.

Sinai and Bowen showed how to do this in the case of finite covers [Si1], [B4]. Thanks to the finiteness property of  $\mathcal{Z}$ , their ideas apply to our case almost without change. The only difference is that in our case, the sets  $Z \in \mathcal{Z}$  are not the closure of their interior, and therefore we cannot use “relative boundaries” and “relative interiors” of  $Z \in \mathcal{Z}$  as done in [Si1] and [B4]. The price is that we cannot claim that the coding we get is one-to-one almost everywhere.

**11.1. The Bowen–Sinai refinement.** Write  $\mathcal{Z} = \{Z_1, Z_2, Z_3, \dots\}$ . Following [B4], we define for every  $Z_i, Z_j \in \mathcal{Z}$  s.t.  $Z_i \cap Z_j \neq \emptyset$ ,

$$\begin{aligned} T_{ij}^{us} &:= \{x \in Z_i : W^u(x, Z_i) \cap Z_j \neq \emptyset, W^s(x, Z_i) \cap Z_j \neq \emptyset\}, \\ T_{ij}^{u\emptyset} &:= \{x \in Z_i : W^u(x, Z_i) \cap Z_j \neq \emptyset, W^s(x, Z_i) \cap Z_j = \emptyset\}, \\ T_{ij}^{\emptyset s} &:= \{x \in Z_i : W^u(x, Z_i) \cap Z_j = \emptyset, W^s(x, Z_i) \cap Z_j \neq \emptyset\}, \\ T_{ij}^{\emptyset\emptyset} &:= \{x \in Z_i : W^u(x, Z_i) \cap Z_j = \emptyset, W^s(x, Z_i) \cap Z_j = \emptyset\}. \end{aligned}$$

Let  $\mathcal{T} := \{T_{ij}^{\alpha\beta} : i, j \in \mathbb{N}, Z_i \cap Z_j \neq \emptyset, \alpha \in \{u, \emptyset\}, \beta \in \{s, \emptyset\}\}$ .

Notice that  $T_{ii}^{us} = Z_i$ , therefore  $\mathcal{T}$  covers the same set as  $\mathcal{Z}$ , namely  $\pi(\Sigma^\#)$ . Another useful identity is  $T_{ij}^{us} = Z_i \cap Z_j$ . The inclusion  $\supseteq$  is trivial. To see  $\subseteq$  suppose  $x \in T_{ij}^{us}$ . Choose some  $y \in W^u(x, Z_i) \cap Z_j$ , then  $y \in Z_i \cap Z_j$ , so  $W^u(x, Z_i) = W^u(y, Z_i) \subset V^u(y, Z_j)$  (Lemma 10.10). Similarly, for every  $z \in W^s(x, Z_i) \cap Z_j$ ,  $W^s(x, Z_i) \subset V^s(z, Z_j)$ . It follows that

$$\{x\} = W^u(x, Z_i) \cap W^s(x, Z_i) \subseteq V^u(y, Z_j) \cap V^s(z, Z_j) \subset Z_j,$$

whence  $x \in Z_i \cap Z_j$ .

**Definition 11.1.** For every  $x \in \pi(\Sigma^\#)$ , let  $R(x) := \bigcap \{T \in \mathcal{T} : T \ni x\}$ , and set  $\mathcal{R} := \{R(x) : x \in \pi(\Sigma^\#)\}$ .

**Proposition 11.2.**  $\mathcal{R}$  is a countable pairwise disjoint cover of  $\text{NUH}_\chi^\#(f)$ .

*Proof.* We prove that  $\mathcal{R}$  is countable by observing that thanks to Theorem 10.2,  $R(x)$  is a finite intersection of elements of  $\mathcal{T}$ . Since  $\mathcal{T}$  is countable, there are at most countably many finite subsets of  $\mathcal{T}$ , and therefore at most countably many different  $R(x)$ 's.

Next we claim that  $\mathcal{R}$  covers  $\text{NUH}_\chi^\#(f)$ . Every  $x \in T \in \mathcal{T}$  belongs to  $R(x) \in \mathcal{R}$ , so  $\bigcup \mathcal{R} = \bigcup \mathcal{T}$ . We saw above that for every  $Z_i \in \mathcal{Z}$ ,  $T_{ii}^{us} = Z_i$ . Consequently,  $\bigcup \mathcal{T} = \bigcup \mathcal{Z} = \pi(\Sigma^\#)$ . Since  $\pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f)$  (see the proof of Theorem 4.16),  $\mathcal{R}$  covers  $\text{NUH}_\chi^\#(f)$ .

It remains to prove that  $\mathcal{R}$  is pairwise disjoint. We do this by proving that  $R(x)$  is the equivalence class of  $x$  for the following equivalence relation on  $\bigcup \mathcal{R}$ :

$$x \sim y \text{ iff } \forall Z, Z' \in \mathcal{Z}, \begin{pmatrix} x \in Z \Leftrightarrow y \in Z \\ W^u(x, Z) \cap Z' \neq \emptyset \Leftrightarrow W^u(y, Z) \cap Z' \neq \emptyset \\ W^s(x, Z) \cap Z' \neq \emptyset \Leftrightarrow W^s(y, Z) \cap Z' \neq \emptyset \end{pmatrix} \quad (11.1)$$

So for every  $x, y \in \bigcup \mathcal{R}$ , either  $R(x) = R(y)$ , or  $R(x) \cap R(y) = \emptyset$ .

*Part 1.* If  $x \sim y$ , then  $x \in R(y)$ .

If  $x \sim y$ , then  $x$  and  $y$  belong to exactly the same elements of  $\mathcal{T}$ . So  $R(x) = R(y)$ .

*Part 2.* If  $x \in R(y)$ , then  $x \sim y$ .

Fix some  $Z_i \in \mathcal{Z}$ . We claim that  $x \in Z_i \Leftrightarrow y \in Z_i$ . Recall that  $Z_i = T_{ii}^{us}$ .

If  $y \in Z_i$ , then  $T_{ii}^{us}$  is one of the sets in the intersection which defines  $R(y)$ . Consequently,  $x \in R(y) \subseteq T_{ii}^{us} = Z_i$ , and  $x \in Z_i$ .

Next suppose  $x \in Z_i$ . Pick some  $Z_k \in \mathcal{Z}$  which contains both  $x$  and  $y$  (any  $k$  s.t.  $T_{k\ell}^{\alpha\beta} \ni y$  will do, because for such  $k$   $Z_k \supset R(y) \ni x, y$ ). Since  $y \in Z_k$  and  $Z_k \cap Z_i \neq \emptyset$ ,  $y \in T_{ki}^{\alpha\beta}$  for some  $\alpha, \beta$ . By the definition of  $R(y)$ ,  $R(y) \subset T_{ki}^{\alpha\beta}$ , whence  $x \in T_{ki}^{\alpha\beta}$ . But  $x \in Z_k \cap Z_i \equiv T_{ki}^{us}$ , so necessarily  $(\alpha, \beta) = (u, s)$ . Thus  $y \in T_{ki}^{us} = Z_k \cap Z_i \subset Z_i$ . This completes the proof that  $x \in Z_i \Leftrightarrow y \in Z_i$ .

Next we show that if  $x \in R(y)$ , then  $W^u(x, Z_i) \cap Z_j \neq \emptyset \Leftrightarrow W^u(y, Z_i) \cap Z_j \neq \emptyset$ . If  $W^u(x, Z_i) \cap Z_j \neq \emptyset$ , then  $x \in T_{ij}^{u*}$ , where  $*$  stands for  $s$  or  $\emptyset$ . In particular  $x \in Z_i$ . By the previous paragraph,  $y \in Z_i$ , and as a result  $y \in T_{ij}^{\alpha\beta}$  for some  $\alpha, \beta$ . Therefore  $x \in R(y) \subset T_{ij}^{\alpha\beta}$ , and since  $T_{ij}^{u*} \cap T_{ij}^{\alpha\beta} = \emptyset$ ,  $\alpha = u$ . It follows that  $y \in T_{ij}^{u*}$ , whence  $W^u(y, Z_i) \cap Z_j \neq \emptyset$  as required. The other implication is trivial: If  $W^u(y, Z_i) \cap Z_j \neq \emptyset$ , then  $y \in T_{ij}^{u*}$ , whence  $x \in R(y) \subseteq T_{ij}^{u*}$ , and so  $W^u(x, Z_i) \cap Z_j \neq \emptyset$ .

The proof that if  $x \in R(y)$ , then  $W^s(x, Z_i) \cap Z_j \neq \emptyset \Leftrightarrow W^s(y, Z_i) \cap Z_j \neq \emptyset$  is exactly the same.  $\square$

**Lemma 11.3.**  $\mathcal{R}$  is a locally finite refinement of  $\mathcal{Z}$ :

- (1) for every  $R \in \mathcal{R}$  and  $Z \in \mathcal{Z}$ , if  $R \cap Z \neq \emptyset$  then  $R \subset Z$ ;
- (2) for every  $Z \in \mathcal{Z}$ ,  $|\{R \in \mathcal{R} : Z \supset R\}| < \infty$ .

*Proof.* Suppose  $R \cap Z \neq \emptyset$  and let  $x \in R \cap Z$ . If  $Z = Z_i$ , then  $Z = T_{ii}^{us}$ . Since  $x \in Z$ ,  $R = R(x) \subseteq T_{ii}^{us} = Z_i = Z$ , whence  $R \subseteq Z$ .

We turn to the second part. If  $R \subset Z$ , then  $R$  is the intersection of a subset of  $\mathcal{T}(Z) := \{T_{ij}^{\alpha\beta} \in \mathcal{T} : T_{ij}^{\alpha\beta} \cap Z \neq \emptyset\}$ . If  $T_{ij}^{\alpha\beta} \cap Z \neq \emptyset$ , then  $Z_i \cap Z \neq \emptyset$ ,  $Z_j \cap Z_i \neq \emptyset$ , and  $\{\alpha, \beta\} \subset \{u, s, \emptyset\}$ . By Theorem 10.2, there are finitely many possibilities for  $Z_i$ , and therefore also finite many possibilities for  $Z_j$ . Thus  $\mathcal{T}(Z)$  is finite.

Since  $\mathcal{T}(Z)$  is finite, and any  $R \subset Z$  is the intersection of a subset of  $\mathcal{T}(Z)$ ,  $|\{R \in \mathcal{R} : R \subset Z\}| \leq 2^{|\mathcal{T}(Z)|} < \infty$ .  $\square$

## 11.2. Product structure and hyperbolicity.

**Definition 11.4.** For any  $R \in \mathcal{R}$  and  $x \in R$ , let

$$W^s(x, R) := \bigcap \{W^s(x, Z_i) \cap T_{ij}^{\alpha\beta} : T_{ij}^{\alpha\beta} \in \mathcal{T} \text{ contains } R\},$$

$$W^u(x, R) := \bigcap \{W^u(x, Z_i) \cap T_{ij}^{\alpha\beta} : T_{ij}^{\alpha\beta} \in \mathcal{T} \text{ contains } R\}.$$

**Proposition 11.5.** Suppose  $R \in \mathcal{R}$  and  $x, y \in R$ .

- (1)  $W^u(x, R), W^s(x, R) \subset R$  and  $W^u(x, R) \cap W^s(x, R) = \{x\}$ .
- (2) Either  $W^u(x, R), W^u(y, R)$  are equal, or they are disjoint. Similarly for  $W^s(x, R)$  and  $W^s(y, R)$ .
- (3)  $W^u(x, R)$  and  $W^s(y, R)$  intersect at a unique point  $z$ , and  $z \in R$ .
- (4) If  $\xi, \eta \in W^s(x, R)$ , then  $d(f^n(\xi), f^n(\eta)) \xrightarrow{n \rightarrow \infty} 0$ . If  $\xi, \eta \in W^u(x, R)$ , then  $d(f^{-n}(\xi), f^{-n}(\eta)) \xrightarrow{n \rightarrow \infty} 0$ .

*Proof.* Suppose  $R \in \mathcal{R}$  and  $x, y \in R$ .

*Part (1).* By definition,  $W^{u/s}(x, R) \subset \bigcap \{T_{ij}^{\alpha\beta} \in \mathcal{T} : T_{ij}^{\alpha\beta} \supset R\} \equiv R$ . It follows that  $W^{u/s}(x, R) \subset R$ .

If  $x \in R$ , then for every  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which contains  $R$ ,  $x \in W^{s/u}(x, Z_i) \cap R \subset W^{s/u}(x, Z_i) \cap T_{ij}^{\alpha\beta}$ . Passing to the intersection, we see that  $x \in W^{s/u}(x, R)$ . Thus  $x \in W^u(x, R) \cap W^s(x, R)$ . On the other hand for every  $Z_i \supseteq R$ ,  $W^s(x, R) \cap W^u(x, R) \subset W^u(x, Z_i) \cap W^s(x, Z_i) = \{x\}$ , so  $W^u(x, R) \cap W^s(x, R) = \{x\}$ .

*Part (2).* Suppose  $W^u(x, R) \cap W^u(y, R) \neq \emptyset$ , then  $W^u(x, Z_i) \cap W^u(y, Z_i) \neq \emptyset$  for every  $i$  s.t. there is some  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which contains  $R$ . By Proposition 10.4,  $W^u(x, Z_i) = W^u(y, Z_i)$ , whence  $W^u(x, Z_i) \cap T_{ij}^{\alpha\beta} = W^u(y, Z_i) \cap T_{ij}^{\alpha\beta}$ . Passing to the intersection, we see that  $W^u(x, R) = W^u(y, R)$ . Similarly, one shows that if  $W^s(x, R) \cap W^s(y, R) \neq \emptyset$ , then  $W^s(x, R) = W^s(y, R)$ .

*Part (3).* For every  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which covers  $R$  and for every  $z \in R$ , let

$$W^u(z, T_{ij}^{\alpha\beta}) := W^u(z, Z_i) \cap T_{ij}^{\alpha\beta} \text{ and } W^s(z, T_{ij}^{\alpha\beta}) := W^s(z, Z_i) \cap T_{ij}^{\alpha\beta}.$$

Fix  $x, y \in R$ . For every  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which contains  $R$ ,  $W^u(x, Z_i) \cap W^s(y, Z_i) = \{z_i\}$  where  $z_i := [x, y]_{Z_i}$ . By Proposition 10.4,  $W^u(z_i, Z_i) = W^u(x, Z_i)$  and  $W^s(z_i, Z_i) = W^s(y, Z_i)$ . It follows that  $z_i \in T_{ij}^{\alpha\beta}$ , whence

$$W^u(x, T_{ij}^{\alpha\beta}) \cap W^s(y, T_{ij}^{\alpha\beta}) = \{z_i\}.$$

Since  $z_i = [x, y]_{Z_i}$ ,  $z_i$  is independent of  $j, \alpha$ , and  $\beta$ . In fact  $z_i$  is also independent of  $i$ : If  $T_{k\ell}^{\gamma\delta} \in \mathcal{T}$  also covers  $R$ , then  $x, y \in Z_i \cap Z_k$  and so

$$\begin{aligned} \{z_i\} &= W^u(x, Z_i) \cap W^s(y, Z_i) \subset V^u(x, Z_i) \cap V^s(y, Z_i) \\ \{z_k\} &= W^u(x, Z_k) \cap W^s(y, Z_k) \subset V^u(x, Z_i) \cap V^s(y, Z_i) \quad (\text{Lemma 10.10}). \end{aligned}$$

Since  $V^u(x, Z_i) \cap V^s(y, Z_i)$  is a singleton,  $z_i = z_k$ .

Denote the common value of  $z_i$  by  $z$ , then  $W^u(x, T_{ij}^{\alpha\beta}) \cap W^s(y, T_{ij}^{\alpha\beta}) = \{z\}$  for all  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which cover  $R$ . Passing to the intersection, we obtain that  $W^u(x, R) \cap W^s(y, R) = \{z\}$ . By part (1) of the lemma,  $z \in R$ .

*Part (4).* Fix some  $Z \in \mathcal{Z}$  such that  $R \subseteq Z$ , then  $x = \pi(\underline{v})$  where  $\underline{v}$  is a regular chain such that  $Z := Z(v_0)$ . By construction,  $W^s(x, R) \subset V^s[(v_i)_{i \geq 0}]$  and  $W^u(x, R) \subset V^u[(v_i)_{i \leq 0}]$ . Part (4) follows from Proposition 6.3(1).  $\square$

Recall the definition of the Smale bracket (Definition 10.6). In the course of the proof we showed the following:

**Lemma 11.6.** *Suppose  $R \in \mathcal{R}$  and  $x, y \in R$ . Let  $[x, y]$  denote the unique element of  $W^u(x, R) \cap W^s(x, R)$ , then  $[x, y] = [x, y]_Z$  for any  $Z \in \mathcal{Z}$  which contains  $R$ .*

**11.3. The Markov property.**  $\mathcal{R}$  satisfies Sinai's Markov property [Si1]:

**Proposition 11.7.** *Let  $R_0, R_1 \in \mathcal{R}$ . If  $x \in R_0$  and  $f(x) \in R_1$ , then*

$$f[W^s(x, R_0)] \subset W^s(f(x), R_1) \text{ and } f^{-1}[W^u(f(x), R_1)] \subset W^u(x, R_0).$$

*Proof.* The proof is an easy adaptation of an argument in [B4, pages 54,55], except that our "rectangles"  $R \in \mathcal{R}$  are defined differently. We give all the details to convince the reader that everything works out as it should.

It is enough to show that  $f[W^s(x, R_0)] \subset W^s(f(x), R_1)$ : the statement for  $W^u$  follows by symmetry.

Suppose  $y \in W^s(x, R_0)$ . We prove that  $f(y) \in W^s(f(x), R_1)$  by checking that for every  $T_{ij}^{\alpha\beta} \in \mathcal{T}$  which covers  $R_1$ ,  $f(y) \in W^s(f(x), Z_i) \cap T_{ij}^{\alpha\beta}$ .

That  $f(y) \in W^s(f(x), Z_i)$  can be shown as follows. Since  $T_{ij}^{\alpha\beta}$  covers  $R_1$ ,  $T_{ij}^{\alpha\beta}$  contains  $f(x)$ . Thus  $f(x) \in T_{ij}^{\alpha\beta} \subset Z_i$ . Write  $Z_i = Z(v)$  and  $f(x) = \pi(\sigma \underline{v})$  where  $\underline{v} \in \Sigma^\#$  satisfies  $v_1 = v$ . Since  $f \circ \pi = \pi \circ \sigma$ ,  $x = \pi(\underline{v}) \in Z(v_0)$ . It follows that  $Z(v_0) \supseteq R(x) = R_0$ , whence  $y \in W^s(x, R_0) \subset W^s(x, Z(v_0))$ . By the symbolic Markov property (Proposition 10.9),

$$f[W^s(x, Z(v_0))] \subset W^s[f(x), Z(v_1)],$$

so  $f(y) \in f[W^s(x, R_0)] \subset f[W^s(x, Z(v_0))] \subset W^s(f(x), Z(v_1)) \equiv W^s(f(x), Z_i)$ .

It remains to prove that if  $y \in W^s(x, R_0)$ , then  $f(x) \in T_{ij}^{\alpha\beta} \Leftrightarrow f(y) \in T_{ij}^{\alpha\beta}$ . Since  $y \in W^s(x, R_0) \Leftrightarrow W^s(x, R_0) = W^s(y, R_0)$ , this is equivalent to showing that if  $W^s(x, R_0) = W^s(y, R_0)$ , then for every  $Z_i, Z_j \in \mathcal{Z}$  s.t.  $Z_i \cap Z_j \neq \emptyset$ ,

- $f(x) \in Z_i \Leftrightarrow f(y) \in Z_i$ ;
- $W^s(f(x), Z_i) \cap Z_j \neq \emptyset \Leftrightarrow W^s(f(y), Z_i) \cap Z_j \neq \emptyset$ ;



- $W^u(f(x), Z_i) \cap Z_j \neq \emptyset \Leftrightarrow W^u(f(y), Z_i) \cap Z_j \neq \emptyset$ .

We only prove  $\Rightarrow$ . The other implication follows by symmetry.

*Step 1.*  $f(x) \in Z_i \Rightarrow f(y) \in Z_i$ .

If  $f(x) \in Z_i$ , then  $f(x) \in T_{ii}^{us} \equiv Z_i$ . Thus  $T_{ii}^{us} \supseteq R(f(x)) = R_1$ . We saw above that if  $T_{ij}^{\alpha\beta}$  covers  $R_1$ , then  $f(y) \in W^s(f(x), Z_i)$ . Applying this to  $T_{ii}^{us}$ , we see that  $f(y) \in W^s(f(x), Z_i) \subset Z_i$ .

*Step 2.*  $W^s(f(x), Z_i) \cap Z_j \neq \emptyset \Rightarrow W^s(f(y), Z_i) \cap Z_j \neq \emptyset$ .

Write  $Z_i = Z(v)$ . Since  $f(x) \in Z_i$ ,  $f(x) = \pi[\sigma\underline{v}]$  where  $\underline{v} \in \Sigma^\#$  and  $v_1 = v$ . Since  $f \circ \pi = \pi \circ \sigma$ ,  $x = \pi(\underline{v})$ . By the symbolic Markov property,  $f[W^s(x, Z(v_0))] \subset W^s(f(x), Z(v_1)) = W^s(f(x), Z_i)$ . Since  $x = \pi(\underline{v})$ ,  $x \in Z(v_0)$ , whence  $R_0 \equiv R(x) \subset Z(v_0)$ . Consequently,

$$\begin{aligned} f(y) &\in f[W^s(y, R_0)] = f[W^s(x, R_0)] \quad (\text{by assumption}) \\ &\subset f[W^s(x, Z(v_0))] \subset W^s(f(x), Z(v_1)) \equiv W^s(f(x), Z_i). \end{aligned}$$

Since  $f(y) \in W^s(f(x), Z_i)$ ,  $W^s(f(y), Z_i) = W^s(f(x), Z_i)$ . It is now clear that  $W^s(f(x), Z_i) \cap Z_j \neq \emptyset \Rightarrow W^s(f(y), Z_i) \cap Z_j \neq \emptyset$ .

*Step 3.*  $W^u(f(x), Z_i) \cap Z_j \neq \emptyset \Rightarrow W^u(f(y), Z_i) \cap Z_j \neq \emptyset$ .

In order to reduce the number of indices, we write  $Z_i = Z$ ,  $Z_j = Z^*$ . We pick some  $f(z) \in W^u(f(x), Z) \cap Z^*$ , and show that  $W^u(f(y), Z) \cap Z^* \ni f(w)$  where  $w := [y, z]_Y$  for some suitable  $Y \in \mathcal{Z}$  that we proceed to construct.

Since  $f(x) \in Z$ , there exists  $\underline{v} \in \Sigma^\#$  such that  $\pi(\sigma\underline{v}) = f(x)$  and  $Z = Z(v_1)$ . Let  $Y := Z(v_0)$ , then  $x = \pi(\underline{v}) \in Y$ . By assumption,  $R(x) = R_0 = R(y)$ , therefore,  $x \sim y$  in the sense of (11.1). Since  $x \in Y$  and  $y \sim x$ ,  $y \in Y$ .

By construction,  $f(z) \in Z^*$  so there exists  $\underline{v}^* \in \Sigma^\#$  such that  $\pi(\sigma\underline{v}^*) = f(z)$  and  $Z^* = Z(v_1^*)$ . Let  $Y^* := Z(v_0^*)$ , then  $z = \pi(\underline{v}^*) \in Y^*$ . By the symbolic Markov property,  $z \in f^{-1}[W^u(f(x), Z)] \equiv f^{-1}[W^u(f(x), Z(v_1))] \subset W^u(x, Z(v_0)) \equiv W^u(x, Y)$ . Thus  $z \in W^u(x, Y) \cap Y^*$ . In particular,  $z \in Y \cap Y^*$ .

Since  $y, z \in Y$ , the Smale bracket  $w := [y, z]_Y$  is well defined. We show that  $f(w) \in W^u(f(y), Z) \cap Z^*$ .

By construction,  $w = [y, z]_Y$ . Since  $f(y) \in Z$  (by Step 1),  $f(z) \in Z$  (by choice), and  $Y = Z(v_0)$ ,  $Z = Z(v_1)$  and  $v_0 \rightarrow v_1$  (by construction), we have by Lemma 10.7 that  $f(w) = f([y, z]_Y) = [f(y), f(z)]_Z \in W^u(f(y), Z)$ .

Next recall that  $W^u(x, Y) \cap Y^*$  is non-empty (it contains  $z$ ). Since  $x \sim y$ ,  $W^u(y, Y) \cap Y^*$  is non-empty. Pick some  $y' \in W^u(y, Y) \cap Y^*$ . Since  $y', z \in Y \cap Y^*$ , we have by Lemma 10.10 that

$$\{w\} = W^u(y', Y) \cap W^s(z, Y) \subset V^u(y', Y^*) \cap V^s(z, Y^*) \equiv \{[y', z]_{Y^*}\}.$$

Thus  $w = [y', z]_{Y^*} \in W^s(z, Y^*)$ . Now  $Y^* = Z(v_0^*)$ ,  $Z^* = Z(v_1^*)$  and  $z = \pi(\underline{v}^*)$ , therefore by the symbolic Markov property,

$$f(w) \in f[W^s(z, Y^*)] \subset W^s(f(z), Z^*) \subset Z^*.$$

It follows that  $f(w) \in Z^*$ . This completes the proof of Step 3. The proposition follows from the discussion before Step 1.  $\square$

## 12. SYMBOLIC DYNAMICS

**12.1. A directed graph.** In the previous section we constructed a Markov partition  $\mathcal{R}$  for  $f$ . Here we use this partition to relate  $f$  to a topological Markov shift. The shift is  $\Sigma(\widehat{\mathcal{G}})$  where  $\widehat{\mathcal{G}}$  is the directed graph with vertices  $\widehat{\mathcal{V}} := \mathcal{R}$  and edges

$$\widehat{\mathcal{E}} := \{(R_1, R_2) \in \mathcal{R}^2 : R_1, R_2 \in \widehat{\mathcal{V}} \text{ s.t. } R_1 \cap f^{-1}(R_2) \neq \emptyset\}.$$

If  $(R_1, R_2) \in \widehat{\mathcal{E}}$ , then we write  $R_1 \rightarrow R_2$ .

For every finite path  $R_m \rightarrow R_{m+1} \rightarrow \cdots \rightarrow R_n$  in  $\widehat{\mathcal{G}}$ , let  ${}_\ell[R_m, \dots, R_n] := \bigcap_{k=\ell}^{\ell+n-m} f^{-k}(R_{k+m-\ell})$ . In particular,

$${}_m[R_m, \dots, R_n] = \bigcap_{k=m}^n f^{-k}(R_k).$$

**Lemma 12.1.** *Suppose  $m \leq n$  and  $R_m \rightarrow R_{m+1} \rightarrow \cdots \rightarrow R_n$  is a finite path on  $\widehat{\mathcal{G}}$ , then  ${}_m[R_m, \dots, R_n] \neq \emptyset$ .*

*Proof.* We use induction on  $n$ .

If  $n = m$ , then the statement is obvious.

Suppose by induction the statement is true for  $n-1$ , and let  $R_m \rightarrow \cdots \rightarrow R_{n-1}$  be a path on  $\widehat{\mathcal{G}}$ . By the induction hypothesis,  ${}_m[R_m, \dots, R_{n-1}] \neq \emptyset$ , therefore there exists a point  $y \in \bigcap_{k=m}^{n-1} f^{-k}(R_k)$ . Since  $R_{n-1} \rightarrow R_n$ , there exists a point  $z \in R_{n-1} \cap f^{-1}(R_n)$ . Let  $x$  be the point such that

$$\{f^{n-1}(x)\} = W^u(f^{n-1}(y), R_{n-1}) \cap W^s(z, R_{n-1}).$$

We claim that  $x \in {}_m[R_m, \dots, R_n]$ . This follows from the Markov property (Theorem 11.7):

- $f^n(x) \in R_n$ , because  $f^n(x) \in f[W^s(z, R_{n-1})] \subset W^s(f(z), R_n) \subset R_n$ ;
- $f^{n-1}(x) \in R_{n-1}$  by construction;
- $f^{n-2}(x) \in R_{n-2}$ , because  $f^{n-1}(x) \in W^u(f^{n-1}(y), R_{n-1}) \subset R_{n-1}$  so  $f^{n-2}(x) \in f^{-1}[W^u(f^{n-1}(y), R_{n-1})] \subset W^u(f^{n-2}(y), R_{n-2}) \subset R_{n-2}$ .
- $f^{n-3}(x) \in R_{n-3}$ , because  $f^{n-2}(x) \in W^u(f^{n-2}(y), R_{n-2})$  so  $f^{n-3}(x) \in f^{-1}[W^u(f^{n-2}(y), R_{n-2})] \subset W^u(f^{n-3}(y), R_{n-3}) \subset R_{n-3}$ .

Continuing this way, we see that  $f^{n-k}(x) \in R_{n-k}$  for all  $0 \leq k \leq n-m$ . □

We compare the paths on  $\widehat{\mathcal{G}}$  to the paths on  $\mathcal{G}$  (the graph we introduced in §4). Recall the map  $\pi : \Sigma \rightarrow M$  from Theorem 4.16, and define for any finite path  $v_m \rightarrow \cdots \rightarrow v_n$  on  $\mathcal{G}$ ,

$$Z_m(v_m, \dots, v_n) := \{\pi(\underline{w}) : \underline{w} \in \Sigma^\#, w_i = v_i \text{ for all } i = m, \dots, n\}.$$

**Lemma 12.2.** *For every infinite path  $\cdots \rightarrow R_i \rightarrow R_{i+1} \rightarrow \cdots$  in  $\widehat{\mathcal{G}}$  there exists a chain  $(v_i)_{i \in \mathbb{Z}} \in \Sigma$  such that for every  $i$ ,  $R_i \subset Z(v_i)$ , and for every  $n$ ,  ${}_{-n}[R_{-n}, \dots, R_n] \subset Z_{-n}(v_{-n}, \dots, v_n)$ .*

*Proof.* Fix, using Lemma 12.1, points  $y_n \in {}_{-n}[R_{-n}, \dots, R_n]$ .

Pick some  $v_0 \in \mathcal{V}$  s.t.  $R_0 \subset Z(v_0)$ . Since  $y_n \in R_0$ , there is a chain  $\underline{v}^{(n)} = (v_i^{(n)})_{i \in \mathbb{Z}} \in \Sigma^\#$  such that  $v_0^{(n)} = v_0$  and  $y_n = \pi[\underline{v}^{(n)}]$ .

For every  $|k| \leq n$ ,  $f^k(y_n) = \pi[\sigma^k(\underline{v}^{(n)})] \in Z(v_k^{(n)})$ , therefore  $Z(v_k^{(n)})$  covers  $R(f^k(y_n))$ . Since, by construction,  $f^k(y_n) \in R_k$ ,  $R(f^k(y_n)) = R_k$ . It follows that

$$R_k \subset Z(v_k^{(n)}) \text{ for every } k = -n, \dots, n.$$

Every vertex in the graph  $\mathcal{G}$  has finite degree (Lemma 4.4). Therefore, there are only finitely many paths of length  $k$  on  $\mathcal{G}$  which start at  $v_0$ . As a result, every set of the form  $\{v_k^{(n)} : n \in \mathbb{N}\}$  is finite. Using the diagonal argument, choose a subsequence  $n_i \uparrow \infty$  s.t. for every  $k$  the sequence  $\{v_k^{(n_i)}\}_{i \geq 1}$  is eventually constant. Call the constant  $v_k$ .

The sequence  $\underline{v} := (v_k)_{k \in \mathbb{Z}}$  is a chain, and  $R_k \subset Z(v_k)$  for all  $k \in \mathbb{Z}$ . We claim that  ${}_n[R_{-n}, \dots, R_n] \subset Z_{-n}(v_{-n}, \dots, v_n)$  for all  $n$ .

Suppose  $y \in {}_{-n}[R_{-n}, \dots, R_n]$ . Since  $f^n(y) \in R_n$  and  $R_n \subset Z(v_n)$ , there exists a chain  $\underline{w} \in \Sigma^\#$  s.t.  $f^n(y) = \pi[\sigma^n(\underline{w})]$  and  $w_n = v_n$ . Since  $f^{-n}(y) \in R_{-n}$  and  $R_{-n} \subset Z(v_{-n})$ , there exists a chain  $\underline{u} \in \Sigma^\#$  s.t.  $f^{-n}(y) = \pi[\sigma^{-n}(\underline{u})]$  and  $u_{-n} = v_{-n}$ . Let

$$\underline{a} = (a_i)_{i \in \mathbb{Z}} \text{ where } a_i = \begin{cases} u_i & i \leq -n \\ v_i & -n \leq i \leq n \\ w_i & i \geq n. \end{cases}$$

For every  $k$ ,  $f^k(y) \in Z(a_k)$ , because

- for all  $k \leq -n$ ,  $f^k(y) \in V^u[(u_i)_{i \leq k}] \subset Z(u_i) = Z(a_i)$ ,
- for all  $-n \leq k \leq n$ ,  $f^k(y) \in R_k \subset Z(v_k) = Z(a_k)$ ,
- for all  $k \geq n$ ,  $f^k(y) \in V^s[(w_i)_{i \geq k}] \subset Z(w_i) = Z(a_i)$ .

Writing  $a_i = \Psi_{x_i}^{p_i^u, p_i^s}$ , we see that  $y \in \Psi_{x_i}[R_{Q_\varepsilon(x_i)}(\mathbb{Q})]$  for all  $i \in \mathbb{Z}$ . By Proposition 4.15 part 4,  $y \in V^u[(a_i)_{i \leq 0}] \cap V^s[(a_i)_{i \geq 0}]$ , so  $y = \pi(\underline{a}) \in Z_{-n}(v_{-n}, \dots, v_n)$ .  $\square$

**Proposition 12.3.** *Every vertex of  $\widehat{\mathcal{G}}$  has finite degree.*

*Proof.* Fix  $R_0 \in \mathcal{R}$ . We bound the number of paths  $R_{-1} \rightarrow R_0 \rightarrow R_1$ .

Consider all the possible paths  $v_{-1} \rightarrow v_0 \rightarrow v_1$  on  $\mathcal{G}$  s.t.  ${}_{-1}[R_{-1}, R_0, R_1] \subset Z_{-1}(v_{-1}, v_0, v_1)$ . There are finitely many possibilities for  $v_0$ , because any two possible choices  $v_0, v'_0$  satisfy  $Z(v_0) \cap Z(v'_0) \supset R_0 \neq \emptyset$ , and  $\mathcal{Z}$  has the finiteness property (Theorem 10.2). Since every vertex of  $\mathcal{G}$  has finite degree, there are also only finitely many possibilities for  $v_{-1}$  and  $v_1$ . By Lemma 11.3(1),  $R_i \subset Z(v_i)$  ( $|i| \leq 1$ ). By Lemma 11.3(2) the number of possible  $R_{-1}$ ,  $R_0$  or  $R_1$  is finite.  $\square$

**12.2. The Markov extension.** Let

$$\widehat{\Sigma} := \Sigma(\widehat{\mathcal{G}}) = \{(R_i)_{i \in \mathbb{Z}} \in \mathcal{R}^{\mathbb{Z}} : R_i \rightarrow R_{i+1} \text{ for all } i \in \mathbb{Z}\}.$$

Abusing notation, we denote the left shift map on  $\widehat{\Sigma}$  by  $\sigma$ , and the natural metric on  $\widehat{\Sigma}$  by  $d(\cdot, \cdot)$ :  $d(\underline{x}, \underline{y}) = \exp[-\min\{|k| : x_k \neq y_k\}]$ . Since every vertex of  $\widehat{\mathcal{G}}$  has finite degree,  $\widehat{\Sigma}$  is locally compact. Define as before

$$\widehat{\Sigma}^\# := \{(R_i)_{i \in \mathbb{Z}} : \exists R, S \in \mathcal{R}, \exists n_k, m_k \uparrow \infty \text{ s.t. } R_{n_k} = R \text{ and } R_{-m_k} = S\}.$$

Clearly  $\widehat{\Sigma}^\#$  contains every periodic point for  $\sigma$ . By Poincaré's Recurrence Theorem, every  $\sigma$ -invariant probability measure on  $\widehat{\Sigma}$  is supported on  $\widehat{\Sigma}^\#$ .

The Markov extension  $\pi : \Sigma \rightarrow M$  is not finite-to-one. Our aim is to construct a finite-to-one Hölder continuous map  $\widehat{\pi} : \widehat{\Sigma} \rightarrow M$  which intertwines  $\sigma$  and  $f$ , and such that  $\widehat{\pi}(\widehat{\Sigma})$  (and even  $\widehat{\pi}(\widehat{\Sigma}^\#)$ ) has full probability w.r.t any ergodic invariant probability measure with entropy larger than  $\chi$ .

We start with the following simple observation:

**Lemma 12.4.** *There exist constants  $C$  and  $0 < \theta < 1$  s.t. for every  $(R_i)_{i \in \mathbb{Z}} \in \widehat{\Sigma}$ ,  $\text{diam}({}_{-n}[R_n, \dots, R_n]) < C\theta^n$ .*

*Proof.* Recall that  $\pi : \Sigma \rightarrow M$  is Hölder continuous, therefore there are  $C$  and  $0 < \theta < 1$  s.t. for every  $\underline{v}, \underline{u} \in \Sigma$ , if  $v_i = u_i$  for all  $|i| \leq n$  then  $d(\pi(\underline{u}), \pi(\underline{v})) < C\theta^n$ . By Lemma 12.2 there exists a chain  $(v_i)_{i \in \mathbb{Z}} \in \Sigma$  s.t.

$${}_{-n}[R_{-n}, \dots, R_n] \subset Z_{-n}(v_{-n}, \dots, v_n).$$

The diameter of  $Z_{-n}(v_{-n}, \dots, v_n)$  is less than or equal to  $C\theta^n$ . Therefore the diameter of  ${}_{-n}[R_{-n}, \dots, R_n]$  is less than or equal to  $C\theta^n$ .  $\square$

Suppose  $(R_i)_{i \in \mathbb{Z}} \in \widehat{\Sigma}$ , and let  $F_n := \overline{{}_{-n}[R_{-n}, \dots, R_n]}$  (closure in  $M$ ). Lemmas 12.1 and 12.4 say that  $\{F_n\}_{n \geq 1}$  is a decreasing sequence of non-empty compact subsets of  $M$ , whose diameters tend to zero. It follows that  $\bigcap_{n \geq 1} F_n$  consists of a single point. We call this point  $\widehat{\pi}[(R_i)_{i \in \mathbb{Z}}]$ :

$$\{\widehat{\pi}[(R_i)_{i \in \mathbb{Z}}]\} = \bigcap_{n=0}^{\infty} \overline{{}_{-n}[R_{-n}, \dots, R_n]}$$

**Theorem 12.5.**  $\widehat{\pi} : \widehat{\Sigma} \rightarrow M$  has the following properties:

- (1)  $\widehat{\pi} \circ \sigma = f \circ \widehat{\pi}$ ;
- (2)  $\widehat{\pi}$  is Hölder continuous;
- (3)  $\widehat{\pi}(\widehat{\Sigma}) \supset \widehat{\pi}(\widehat{\Sigma}^\#) \supset \text{NUH}_\chi^\#(f)$ , therefore the image of  $\widehat{\pi}$  has full measure w.r.t every ergodic invariant probability measure with entropy larger than  $\chi$ ;

*Proof.* The commutation relation is because for every  $\underline{R} = (R_i)_{i \in \mathbb{Z}}$  in  $\widehat{\Sigma}$ ,

$$\begin{aligned} \{\pi[\sigma(\underline{R})]\} &= \bigcap_{n=0}^{\infty} \overline{{}_{-n}[R_{-n+1}, \dots, R_{n+1}]} \supset \bigcap_{n=0}^{\infty} \overline{{}_{-n-2}[R_{-n-1}, \dots, R_{n+1}]} \\ &= \bigcap_{n=0}^{\infty} \overline{\bigcap_{k=-n-2}^n f^{-k}(R_{k+1})} = \bigcap_{N=0}^{\infty} \overline{f({}_{-N}[R_{-N}, \dots, R_N])} \\ &= \bigcap_{N=0}^{\infty} f(\overline{{}_{-N}[R_{-N}, \dots, R_N]}), \text{ because } f \text{ is a homeomorphism} \\ &= f\left(\bigcap_{N=0}^{\infty} \overline{{}_{-N}[R_{-N}, \dots, R_N]}\right), \text{ because } f \text{ is a bijection} \\ &\equiv f(\{\pi(\underline{R})\}) = \{f[\pi(\underline{R})]\}. \end{aligned}$$

The Hölder continuity of  $\pi$  is because if  $\underline{R}, \underline{S} \in \widehat{\Sigma}$  and  $R_i = S_i$  for all  $|i| \leq N$ , then  $\widehat{\pi}(\underline{R}), \widehat{\pi}(\underline{S}) \in {}_{-N}[R_{-N}, \dots, R_N]$ , whence by Lemma 12.4

$$d(\widehat{\pi}(\underline{R}), \widehat{\pi}(\underline{S})) \leq \text{diam}({}_{-N}[R_{-N}, \dots, R_N]) \leq C\theta^N.$$

Finally we claim that  $\widehat{\pi}(\widehat{\Sigma})$  and  $\widehat{\pi}(\widehat{\Sigma}^\#)$  contain  $\text{NUH}_\chi^\#(f)$ . Suppose  $x \in \text{NUH}_\chi^\#(f)$ . By Theorem 4.16,  $\pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f)$ , therefore there exists a chain  $\underline{v} \in \Sigma^\#$  s.t.

$\pi(\underline{v}) = x$ .  $\Sigma^\#$  is  $\sigma$ -invariant and  $f \circ \pi = \pi \circ \sigma$ , so  $f^i(x) \in \pi(\Sigma^\#)$  for all  $i \in \mathbb{Z}$ . The collection  $\mathcal{R}$  covers  $\pi(\Sigma^\#)$ , therefore for every  $i \in \mathbb{Z}$  there is some  $R_i \in \mathcal{R}$  s.t.  $f^i(x) \in R_i$ . Obviously  $R_i \rightarrow R_{i+1}$ , so  $\underline{R} := (R_i)_{i \in \mathbb{Z}}$  belongs to  $\widehat{\Sigma}$ . Also,

$$x \in \bigcap_{n=0}^{\infty} \overline{-n[R_{-n}, \dots, R_n]}$$

(even without the closure), so  $x = \pi(\underline{R})$ . It follows that  $\widehat{\pi}(\widehat{\Sigma}) \supset \text{NUH}_\chi^\#(f)$ .

We claim that the sequence  $\underline{R}$  which was constructed above belongs to  $\widehat{\Sigma}^\#$ , and deduce that  $\widehat{\pi}(\widehat{\Sigma}^\#) \supset \text{NUH}_\chi^\#(f)$ .

The sequence  $\underline{v}$  is in  $\Sigma^\#$  by construction, therefore there exists  $v$  and  $u$  s.t.  $v_i = u$  for infinitely many negative  $i$ , and  $v_i = v$  for infinitely many positive  $i$ .

The sets  $R_i$  and  $Z(v_i)$  intersect, because they both contain  $f^i(x)$ . By Lemma 11.3,  $R_i \subset Z(v_i)$  for all  $i \in \mathbb{Z}$ . It follows that there are infinitely many negative  $i$  s.t.  $R_i \subset Z(u)$ , and infinitely many positive  $i$  s.t.  $R_i \subset Z(v)$ .

The sets  $\mathcal{R}(w) := \{R \in \mathcal{R} : R \subset Z(w)\}$  ( $w = u, v$ ) are finite (Lemma 11.3). Therefore  $\exists n_k \uparrow \infty$  and  $\exists R \in \mathcal{R}(v)$  s.t.  $R_{n_k} = R$  for all  $k$ , and  $\exists m_k \uparrow \infty$  and  $\exists S \in \mathcal{R}(u)$  s.t.  $R_{-m_k} = S$  for all  $k$ . Thus  $\underline{R} \in \widehat{\Sigma}^\#$  as required.  $\square$

The following result is not needed for the purposes of this paper, but we anticipate some future applications.

**Proposition 12.6.** *For every  $x \in \widehat{\pi}(\widehat{\Sigma})$ ,  $T_x M = E^s(x) \oplus E^u(x)$  where*

- (a)  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_x^n \underline{v}\|_{f^n(x)} \leq -\frac{\chi}{2}$  on  $E^s(x) \setminus \{0\}$ ;
- (b)  $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df_x^{-n} \underline{v}\|_{f^{-n}(x)} \leq -\frac{\chi}{2}$  on  $E^u(x) \setminus \{0\}$ .

*The maps  $\underline{R} \mapsto E^{u/s}(\widehat{\pi}(\underline{R}))$  are Hölder continuous as maps from  $\widehat{\Sigma}$  to  $TM$ .*

*Proof.* Suppose  $x = \widehat{\pi}(\underline{R})$  where  $\underline{R} \in \widehat{\Sigma}$ . By Lemma 12.2, there is a chain  $(v_i)_{i \in \mathbb{Z}}$  s.t.  $R_i \subset Z(v_i)$  for all  $i$  and  $-n[R_{-n}, \dots, R_n] \subset Z_{-n}(v_{-n}, \dots, v_n)$  for every  $n$ . Then  $f^n(x) \in \overline{Z(v_n)}$  for all  $n$ . Every element of  $Z(v_n)$  is the intersection of  $s/u$ -admissible manifolds in  $v_n$ , so if  $v_n = \Psi_{x_n}^{p_n^s, p_n^u}$ , then  $\overline{Z(v_n)} \subset \Psi_{x_n}[R_{p_n^s} \wedge p_n^u(0)]$  (Proposition 4.11 (2)). By Proposition 4.15 (4),  $x \in V^u[(v_i)_{i \leq 0}] \cap V^s[(v_i)_{i \geq 0}]$ .

Let  $E^s(x) := T_x V^s[(v_i)_{i \geq 0}]$  and  $E^u(x) := T_x V^u[(v_i)_{i \leq 0}]$ . These spaces satisfy (a) and (b), because they are tangent to admissible manifolds which stay in windows (Proposition 6.3). This definition of  $E^s(x), E^u(x)$  is independent of the choice of  $(v_i)_{i \in \mathbb{Z}}$ , because there can be only one decomposition of  $T_x M$  into two spaces which satisfy (a) and (b).

Suppose  $x = \widehat{\pi}(\underline{R})$  and  $y = \widehat{\pi}(\underline{S})$  where  $R_i = S_i$  for  $i = -N, \dots, N$ , and let  $v = (v_i)_{i \in \mathbb{Z}}$  be as before. The argument in the first paragraph shows that  $x = \pi(\underline{v})$ . We claim that  $y = \pi(\underline{w})$  where  $\underline{w}$  is a chain s.t.  $w_i = v_i$  for all  $|i| \leq N$ .

By assumption,  $y \in -n[S_{-N}, \dots, S_N] = -n[R_{-N}, \dots, R_N] \subset \overline{Z_{-N}(v_{-N}, \dots, v_N)}$ , so  $y = \lim \pi(\underline{w}^{(n)})$  where  $\underline{w}^{(n)} \in \Sigma$  satisfy  $w_i^{(n)} = v_i$  for all  $|i| \leq N$ . Since every vertex of  $\mathcal{G}$  has finite degree, each of the sets  $\{w_i^{(n)} : n \in \mathbb{N}\}$  is finite. It follows that there is a convergent subsequence  $\underline{w}^{(n_k)} \xrightarrow[k \rightarrow \infty]{} \underline{w}$ . The limit is a chain  $\underline{w}$  s.t.  $y = \pi(\underline{w})$  and  $w_i = v_i$  for all  $|i| \leq N$ .

Write  $v_0 = \Psi_{x_0}^{p_0^u, p_0^s}$ , and let  $F_u, F_s$  be the representing functions in  $\Psi_{x_0}$  for  $V^u[(v_i)_{i \leq 0}]$ ,  $V^s[(v_i)_{i \geq 0}]$ . Let  $G_u, G_s$  be the representing functions for  $V^u[(w_i)_{i \leq 0}]$ ,  $V^s[(w_i)_{i \geq 0}]$ .

The intersection of the (vertical) graph of  $F_u$  and the (horizontal) graph of  $F_s$  is the point  $\underline{\xi} \in \mathbb{R}^2$  s.t.  $\Psi_{x_0}(\underline{\xi}) = x$ . The intersection of the vertical and horizontal graphs of  $G_u$  and  $G_s$  is the point  $\underline{\eta} \in \mathbb{R}^2$  s.t.  $\Psi_{x_0}(\underline{\eta}) = y$ . By Proposition 4.11 and the uniform hyperbolicity of  $f$  in coordinates,  $\|\underline{\xi} - \underline{\eta}\| < K\theta^N(p_0^u \wedge p_0^s)$  for some global constants  $K > 0, \theta \in (0, 1)$ .

By admissibility,  $F_u, F_s, G_u, G_s$  have  $\frac{\beta}{3}$ -Hölder exponent at most  $\frac{1}{2}$ . This implies  $|F'_s(\xi_1) - G'_s(\eta_1)|, |F'_u(\xi_2) - G'_u(\eta_2)| = O(\theta^{\frac{1}{3}\beta N} Q_\varepsilon(x_0)^{\frac{\beta}{3}})$ . It follows that  $\text{dist}_{T\mathbb{R}^2}(T_{\underline{\xi}}[\text{graph}(F_s)], T_{\underline{\eta}}[\text{graph}(G_s)]) = O(\theta^{\frac{1}{3}\beta N} Q_\varepsilon(x_0)^{\frac{\beta}{3}})$ .

$E^s(x), \bar{E}^s(y)$  are the images of  $T_{\underline{\xi}}[\text{graph}(F_s)]$  and  $T_{\underline{\eta}}[\text{graph}(G_s)]$  under  $d\Psi_{x_0}$ . By Lemma 2.9(2),  $\text{dist}_{TM}(E^s(x), E^s(y)) = O(\theta^{\frac{1}{3}\beta N})$ . Similarly,  $\text{dist}_{TM}(E^u(x), E^u(y)) = O(\theta^{\frac{1}{3}\beta N})$ . All implied constants are uniform, so  $\underline{R} \mapsto E^{s/u}(\widehat{\pi}(\underline{R}))$  are Hölder continuous.  $\square$

**12.3. The extension is finite-to-one.** Say that  $R, R' \in \mathcal{R}$  are *affiliated*, if there exist  $Z, Z' \in \mathcal{Z}$  s.t.  $R \subset Z, R' \subset Z'$ , and  $Z \cap Z' \neq \emptyset$ . For every  $R \in \mathcal{R}$ , let

$$N(R) := |\{(R', Z') \in \mathcal{R} \times \mathcal{Z} : R' \text{ is affiliated to } R \text{ and } Z' \text{ contains } R'\}|.$$

**Lemma 12.7.**  $N(R) < \infty$ .

*Proof.* Suppose  $R \in \mathcal{R}$ . The set  $A(R) := \{Z \in \mathcal{Z} : Z \supset R\}$  is finite, because if  $Y \in \mathcal{Z}$  contains  $R$  then every  $Z \in A(R)$  intersects  $Y$ , and the number of such  $Z$  is finite (Theorem 10.2).

Since  $A(R)$  is finite,  $B(R) := \{Z' \in \mathcal{Z} : \exists Z \in A(R) \text{ s.t. } Z' \cap Z \neq \emptyset\}$  is finite (Theorem 10.2). For every  $Z' \in B$  there are at most finitely many  $R' \in \mathcal{R}$  s.t.  $R' \subset Z'$  (Lemma 11.3). Therefore,  $C(R) := \{R' \in \mathcal{R} : R, R' \text{ are affiliated}\}$  is finite. It follows that  $N(R) = \sum_{R' \in C(R)} |A(R')| < \infty$ .  $\square$

**Theorem 12.8.** *Every  $x \in \widehat{\pi}(\widehat{\Sigma}^\#)$  has a finite number of  $\widehat{\pi}$ -pre-images. More precisely, if  $x = \widehat{\pi}(\underline{R})$  where  $R_i = R$  for infinitely many  $i < 0$  and  $R_i = S$  for infinitely many  $i > 0$ , then  $|\widehat{\pi}^{-1}(x)| < \varphi_\chi(R, S) := N(R)N(S)$ .*

*Proof.* The proof is based on an idea of Bowen's [B3, pp. 13–14] (see also [PP, page 229]), who used it in the context of Axiom A diffeomorphisms. We show that the product structure described above is sufficient to implement his argument in our setting.

Suppose  $x \in \widehat{\pi}(\widehat{\Sigma}^\#)$ , then  $x$  has a  $\widehat{\pi}$ -preimage  $\underline{R} \in \widehat{\Sigma}$  s.t.  $R_i = R$  for infinitely many negative  $i$ , and  $R_i = S$  for infinitely many positive  $i$ . Let  $N := N(R)N(S)$  and assume by way of contradiction that there are  $N+1$  different points in  $\widehat{\Sigma}$  whose image under  $\widehat{\pi}$  is equal to  $x$ . Call these points  $\underline{R}^{(j)} = (R_i^{(j)})_{i \in \mathbb{Z}}$  ( $j = 0, \dots, N$ ). Assume w.l.o.g. that  $\underline{R}^{(0)} = \underline{R}$ .

By Lemma 12.2 there are chains  $\underline{v}^{(j)} = (v_i^{(j)})_{i \in \mathbb{Z}} \in \Sigma$  s.t. for every  $n$

$$R_n^{(j)} \subset Z(v_n^{(j)}) \text{ and } {}_{-n}[R_{-n}^{(j)}, \dots, R_n^{(j)}] \subset Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)}). \quad (12.1)$$

*Claim 1.*  $\pi(\underline{v}^{(j)}) = x$  for every  $0 \leq j \leq N$ .

The following inclusions hold:

$$\pi(\underline{v}^{(j)}) \in \bigcap_{n=0}^{\infty} Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)}) \subset \bigcap_{n=0}^{\infty} \overline{Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)})}, \quad (12.2)$$

$$x = \widehat{\pi}(\underline{R}^{(j)}) \in \bigcap_{n=0}^{\infty} \overline{-n[R_{-n}^{(j)}, \dots, R_n^{(j)}]} \subset \bigcap_{n=0}^{\infty} \overline{Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)})}.$$

Since  $\pi$  is Hölder continuous,  $\text{diam} \left[ \overline{Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)})} \right] \xrightarrow{n \rightarrow \infty} 0$ , so  $\pi(\underline{v}^{(j)}) = x$ .

*Claim 2:* Suppose  $i \in \mathbb{Z}$ , then  $R_i^{(0)}, \dots, R_i^{(N)}$  are affiliated.

*Proof.* By (12.2)  $x = \pi(\underline{v}^{(j)}) \in \bigcap_{n=0}^{\infty} Z_{-n}(v_{-n}^{(j)}, \dots, v_n^{(j)})$ , so  $f^i(x) \in Z(v_i^{(j)})$ . Thus  $Z(v_i^{(0)}), \dots, Z(v_i^{(N)})$  have a common intersection. Since  $R_i^{(j)} \subset Z(v_i^{(j)})$ ,  $R_i^{(0)}, \dots, R_i^{(N)}$  are affiliated.

*Claim 3:* There exist  $k, \ell \geq 0$  and  $0 \leq j_1, j_2 \leq N$  such that

- $(R_{-k}^{(j_1)}, \dots, R_{\ell}^{(j_1)}) \neq (R_{-k}^{(j_2)}, \dots, R_{\ell}^{(j_2)})$ ;
- $R_{-k}^{(j_1)} = R_{-k}^{(j_2)}$  and  $R_{\ell}^{(j_1)} = R_{\ell}^{(j_2)}$ ;
- $v_{-k}^{(j_1)} = v_{-k}^{(j_2)}$  and  $v_{\ell}^{(j_1)} = v_{\ell}^{(j_2)}$ .

*Proof.* We are assuming that  $\underline{R}^{(j)}$  are different, therefore there exists some  $m$  such that the words  $(R_{-m}^{(j)}, \dots, R_m^{(j)})$  ( $0 \leq j \leq N$ ) are different.

We are assuming that  $R_i^{(0)}$  equals  $R$  for infinitely many negative  $i$ , and equals  $S$  for infinitely many positive  $i$ . Choose  $k, \ell \geq m$  s.t.  $R_{-k}^{(0)} = R$  and  $R_{\ell}^{(0)} = S$ . The words  $(R_{-k}^{(j)}, \dots, R_{\ell}^{(j)})$  ( $0 \leq j \leq N$ ) are different.

By claims 1 and 2,  $R_{-k}^{(j)}$  are all affiliated to  $R_{-k}^{(0)} = R$ , and by (12.1)  $R_{-k}^{(j)} \subset Z(v_{-k}^{(j)})$ , therefore  $|\{(R_{-k}^{(j)}, v_{-k}^{(j)}) : j = 0, \dots, N\}| \leq N(R)$ . In the same way, one can show that  $|\{(R_{\ell}^{(j)}, v_{\ell}^{(j)}) : j = 0, \dots, N\}| \leq N(S)$ . It follows that

$$|\{(R_{-k}^{(j)}, v_{-k}^{(j)}; R_{\ell}^{(j)}, v_{\ell}^{(j)}) : j = 0, \dots, N\}| \leq N(R)N(S) = N.$$

By the pigeonhole principle, at least two quadruples coincide, proving the claim.

To ease up the notation, we let  $\underline{A} := \underline{R}^{(j_1)}$ ,  $\underline{B} := \underline{R}^{(j_2)}$ ,  $\underline{a} := \underline{v}^{(j_1)}$  and  $\underline{b} := \underline{v}^{(j_2)}$ , and we write  $A_{-k} = B_{-k} =: B$ ,  $A_{\ell} = B_{\ell} =: A$ ,  $a_{-k} = b_{-k} =: b$ ,  $a_{\ell} = b_{\ell} =: a$ . By Lemma 12.1, there are two points

$$x_A \in -k[A_{-k}, \dots, A_{\ell}] \text{ and } x_B \in -k[B_{-k}, \dots, B_{\ell}].$$

By definition,  $f^{-k}(x_A), f^{-k}(x_B) \in B \subset Z(b)$  and  $f^{\ell}(x_A), f^{\ell}(x_B) \in A \subset Z(a)$ . Define two points  $z_A, z_B$  by the equations

$$\begin{aligned} f^{-k}(z_A) &\in W^u(f^{-k}(x_B), B) \cap W^s(f^{-k}(x_A), B); \\ f^{\ell}(z_B) &\in W^u(f^{\ell}(x_B), A) \cap W^s(f^{\ell}(x_A), A). \end{aligned}$$

*Claim 4.*  $z_A \neq z_B$ .

*Proof.* By construction,  $f^{-k}(z_A) \in W^s(f^{-k}(x_A), A_{-k})$ . By the Markov property (Theorem 11.7),

$$\begin{aligned} f^{-k+1}(z_A) &\in f[W^s(f^{-k}(x_A), A_{-k})] \subset W^s(f^{-k+1}(x_A), A_{-k+1}) \\ f^{-k+2}(z_A) &\in f[W^s(f^{-k+1}(x_A), A_{-k+1})] \subset W^s(f^{-k+2}(x_A), A_{-k+2}) \end{aligned}$$

and so on. It follows that  $f^{-k}(z_A) \in {}_{-k}[A_{-k}, \dots, A_\ell]$ . Similarly, if we start from  $f^\ell(z_B) \in W^u(f^\ell(x_B), B_\ell)$  and apply  $f^{-1}$  repeatedly, then the Markov property will give us that  $f^{-k}(z_B) \in {}_{-k}[B_{-k}, \dots, B_\ell]$ .

But  $(A_{-k}, \dots, A_\ell) \equiv (R_{-k}^{(j_1)}, \dots, R_\ell^{(j_1)}) \neq (R_{-k}^{(j_2)}, \dots, R_\ell^{(j_2)}) \equiv (B_{-k}, \dots, B_\ell)$ , and the elements of  $\mathcal{R}$  are pairwise disjoint, so  ${}_{-k}[A_{-k}, \dots, A_\ell] \cap {}_{-k}[B_{-k}, \dots, B_\ell] = \emptyset$  and  $z_A \neq z_B$ .

*Claim 5.*  $z_A = z_B$  (a contradiction).

*Proof.* We saw above that  $f^{-k}(z_A) \in {}_{-k}[A_{-k}, \dots, A_\ell]$ ,  $f^{-k}(z_B) \in {}_{-k}[B_{-k}, \dots, B_\ell]$ . In particular,  $f^{-k}(z_B) \in B_{-k} = B \subset Z(b)$  and  $f^\ell(z_A) \in A_\ell = A \subset Z(a)$ .

Construct chains  $\underline{\alpha}, \underline{\beta} \in \Sigma^\#$  such that  $z_A = \pi(\underline{\alpha})$ ,  $\alpha_\ell = a$  and  $z_B = \pi(\underline{\beta})$ ,  $\beta_{-k} = b$ . Define a sequence  $\underline{c}$  by

$$c_i = \begin{cases} \beta_i & i \leq -k \\ a_i & -k+1 \leq i \leq \ell-1 \\ \alpha_i & i \geq \ell. \end{cases}$$

This is a chain because  $\beta_{-k} = b = a_{-k}$  and  $\alpha_\ell = a = a_\ell$ . This chain belongs to  $\Sigma^\#$ , because  $\underline{\alpha}, \underline{\beta} \in \Sigma^\#$ . We write  $c_i := \Psi_{x_i}^{p_i^u, p_i^s}$ .

We claim that  $f^{-k}(z_A), f^{-k}(z_B) \in V^u[(c_i)_{i \leq -k}]$ . Note firstly that both points belong to  $W^u(f^{-k}(x_B), B)$ :  $f^{-k}(z_A)$  by definition, and  $f^{-k}(z_B)$  because  $f^\ell(z_B) \in W^u(f^\ell(x_B), B_\ell)$ . Since  $B \subset Z(b)$ ,

$$W^u(f^{-k}(x_B), B) \subset V^u(f^{-k}(x_B), Z(b)) = V^u[(\beta_i)_{i \leq -k}] \equiv V^u[(c_i)_{i \leq -k}].$$

It follows that  $f^{-k}(z_A), f^{-k}(z_B) \in V^u[(c_i)_{i \leq -k}]$ .

This together with the fact that  $f^{-k}(z_A), f^{-k}(z_B) \in Z(b) = Z(c_{-k})$  implies that

$$f^i(z_A), f^i(z_B) \in Z(c_i) \subset \Psi_{x_i}[R_{p_i^u \wedge p_i^s}(\underline{0})] \text{ for all } i \leq -k. \quad (12.3)$$

Similarly, one can show that  $f^\ell(z_A), f^\ell(z_B) \in V^s[(c_i)_{i \geq \ell}]$ , whence

$$f^i(z_A), f^i(z_B) \in Z(c_i) \subset \Psi_{x_i}[R_{p_i^u \wedge p_i^s}(\underline{0})] \text{ for all } i \geq \ell. \quad (12.4)$$

Using the inclusions  $f^{-k}(z_A) \in {}_{-k}[A_{-k}, \dots, A_\ell]$ ,  $f^{-k}(z_B) \in {}_{-k}[B_{-k}, \dots, B_\ell]$  (see the proof of claim 4), we see that if  $-k < i < \ell$  then  $f^i(z_A), f^i(z_B) \in A_i \cup B_i$ . Therefore  $f^i(z_A), f^i(z_B) \in Z(a_i) \cup Z(b_i)$ . The sets  $Z(a_i), Z(b_i)$  intersect, because by claim 1  $f^i(x) = \pi[\sigma^i(\underline{a})] = \pi[\sigma^i(\underline{b})] \in Z(a_i) \cap Z(b_i)$ . Thus by Lemma 10.10,

$$f^i(z_A), f^i(z_B) \in Z(a_i) \cup Z(b_i) \subset \Psi_{x_i}[R_{Q_\varepsilon(x_i)}(\underline{0})] \text{ for all } -k < i < \ell. \quad (12.5)$$

In summary,  $f^i(z_A), f^i(z_B) \in \Psi_{x_i}[R_{Q_\varepsilon(x_i)}(\underline{0})]$ , where  $c_i = \Psi_{x_i}^{p_i^u, p_i^s}$  is a chain. By Proposition 4.15(4),  $z_A, z_B \in V^u[(c_i)_{i \leq 0}] \cap V^s[(c_i)_{i \geq 0}]$ . So  $z_A = \pi(\underline{c}) = z_B$ , and the claim is proved.

The contradiction between claims 4 and 5 shows that  $x$  cannot have more than  $N$  pre-images.  $\square$



## 13. INVARIANT MEASURES

Let  $\sigma : \widehat{\Sigma} \rightarrow \widehat{\Sigma}$  denote the finite-to-one Markov extension of  $f$  which we constructed in part 3. We compare the invariant Borel measures of  $\sigma : \widehat{\Sigma} \rightarrow \widehat{\Sigma}$  to the invariant Borel measures of  $f : M \rightarrow M$ . We restrict our attention to measures whose entropy is larger than  $\chi$ .

**Proposition 13.1.** *Suppose  $\widehat{\mu}$  is an ergodic Borel probability measure on  $\widehat{\Sigma}$ , then  $\mu := \widehat{\mu} \circ \widehat{\pi}^{-1}$  is an ergodic Borel probability measure on  $M$ , and  $h_\mu(f) = h_{\widehat{\mu}}(\sigma)$*

*Proof.* It is clear that  $\mu$  is well-defined, ergodic and invariant.

By Poincaré's Recurrence Theorem (applied to  $\widehat{\mu}$ ) there is a vertex  $R \in \mathcal{R}$  s.t.

$$\Upsilon := \{\underline{R} \in \widehat{\Sigma} : \exists n_k, m_k \uparrow \infty \text{ s.t. } R_{n_k}, R_{-m_k} = R\}$$

has full measure with respect to  $\widehat{\mu}$ . The map  $\widehat{\pi} : \Upsilon \rightarrow M$  is bounded-to-one (the bound is  $\varphi_\chi(R, R)$ ). Finite extensions preserve entropy, so  $h_\mu(f) = h_{\widehat{\mu}}(\sigma)$ .  $\square$

The other direction, “every invariant measure  $\mu$  supported on  $\widehat{\pi}(\widehat{\Sigma})$  lifts to an invariant measure on  $\widehat{\Sigma}$ ”, is less clear.<sup>7</sup> Lifting measures to Markov extensions is a difficult issue in general, and it has received considerable attention (see e.g. [Hof1],[Ke1],[Bru],[BT],[PSZ],[Bu2],[Z]). But our case is very simple, because our Markov extension is finite-to-one.

Indeed, suppose  $\mu$  is an ergodic  $f$ -invariant probability measure on  $M$  s.t.  $h_\mu(f) > \chi$ . Define a measure  $\widetilde{\mu}$  on  $\widehat{\Sigma}$  by

$$\widetilde{\mu}(E) := \int_M \left( \frac{1}{|\widehat{\pi}^{-1}(x)|} \sum_{\widehat{\pi}(\underline{R})=x} 1_E(\underline{R}) \right) d\mu(x). \quad (13.1)$$

**Proposition 13.2.** *Suppose  $\mu$  is an ergodic  $f$ -invariant Borel probability measure on  $M$  s.t.  $h_\mu(f) > \chi$ .*

- (1)  $\widetilde{\mu}$  is a well-defined  $\sigma$ -invariant Borel probability measure on  $\widehat{\Sigma}$ .
- (2) Almost every ergodic component  $\widehat{\mu}$  of  $\widetilde{\mu}$  is an ergodic  $\sigma$ -invariant probability measure such that  $\widehat{\mu} \circ \widehat{\pi}^{-1} = \mu$  and  $h_{\widehat{\mu}}(\sigma) = h_\mu(f)$ .

*Proof.* The first thing to do is to verify that the integrand in (13.1) is measurable. We recall some basic facts from set theory (see e.g. [Sr, §4.5, §4.12]): Let  $X, Y$  be two complete separable metric spaces.

- (I)  $F : X \rightarrow Y$  is Borel iff  $\text{graph}(F)$  is a Borel subset of  $X \times Y$ .
- (II) Suppose  $F : X \rightarrow Y$  is Borel and countable-to-one (i.e.  $F^{-1}(y)$  is finite or countable for all  $y \in Y$ ). If  $E \subset X$  is Borel, then  $F(E) \subset Y$  is Borel.
- (III) Lusin's Theorem: Suppose  $B \subset X \times Y$  is Borel. If  $B_x := \{y : (x, y) \in B\}$  is finite or countable for every  $x \in X$ , then  $B$  is a countable disjoint union of Borel graphs of partially defined Borel functions.

Since  $h_\mu(f) > \chi$ ,  $\mu$  is carried by  $\widehat{\pi}(\Sigma^\#)$ . Since  $\widehat{\pi} : \Sigma^\# \rightarrow M$  is finite-to-one,  $\widehat{\pi}(\Sigma^\#)$  is Borel. Henceforth we work inside  $\widehat{\pi}(\Sigma^\#)$ .

*Step 1.*  $x \mapsto |\widehat{\pi}^{-1}(x)|$  is constant on a Borel set  $\Omega$  s.t.  $\mu(\Omega) = 1$ .

*Proof.* Since  $\widehat{\pi} \circ \sigma = f \circ \widehat{\pi}$  and  $f$  is a bijection,  $x \mapsto |\widehat{\pi}^{-1}(x)|$  is  $f$ -invariant.

We show that the restriction of  $x \mapsto |\widehat{\pi}^{-1}(x)|$  to  $\widehat{\pi}(\Sigma^\#)$  is Borel measurable. The claim will then follow from the ergodicity of  $\mu$ .

<sup>7</sup> $\mu \circ \widehat{\pi}$  does not work: it is not even  $\sigma$ -additive.

Graphs of Borel functions are Borel, therefore  $B := \{(\widehat{\pi}(\underline{R}), \underline{R}) : \underline{R} \in \widehat{\Sigma}^\#\}$  is a Borel subset of  $M \times \widehat{\Sigma}$ .

By Lusin's theorem, there exist partially defined Borel functions  $\varphi_n : M_n \rightarrow \widehat{\Sigma}^\#$  s.t.  $M_n$  are pairwise disjoint Borel subsets of  $M$  and  $B = \{(x, \varphi_n(x)) : x \in M_n, n \in \mathbb{N}\}$ . In particular,  $\widehat{\pi}^{-1}(x) = \{\varphi_i(x) : i \in \mathbb{N} \text{ s.t. } M_i \ni x\}$ . The graphs of  $\varphi_n$  are pairwise disjoint, so  $i \neq j \Rightarrow \varphi_i(x) \neq \varphi_j(x)$ . Consequently,

$$|\widehat{\pi}^{-1}(x)| = \sum_{i=1}^{\infty} 1_{M_i}(x) \text{ on } \widehat{\pi}(\widehat{\Sigma}^\#).$$

Since  $M_i$  are Borel,  $x \mapsto |\widehat{\pi}^{-1}(x)|$  is Borel on  $\widehat{\pi}(\widehat{\Sigma}^\#)$ .

*Step 2.* Let  $\Upsilon := \widehat{\pi}^{-1}(\Omega)$  and let  $N$  denote the number of pre-images of points  $x \in \Omega$ . There exists a Borel partition  $\Upsilon = \bigsqcup_{i=1}^N \Upsilon_i$  such that  $\widehat{\pi} : \Upsilon_i \rightarrow \Omega$  is one-to-one and onto for every  $i$ .

*Proof.* This is a consequence of Lusin's Theorem.

Let  $B_1 := \{(\widehat{\pi}(\underline{y}), \underline{y}) : \underline{y} \in \widehat{\pi}^{-1}(\Omega)\}$ . Each  $x$ -fibre of  $B_1$  has  $N$  elements. By Lusin's Theorem  $B_1 = \bigsqcup_{n \geq 1} \text{graph}(\varphi_n)$  where  $\varphi_n : M_n \rightarrow \widehat{\Sigma}$  are Borel.  $\Omega = \bigsqcup_{n \geq 1} M_n$ .

Define  $\psi_1 : \Omega \rightarrow \widehat{\Sigma}$  by  $\psi_1 = \varphi_i$  on  $M_i \setminus \bigcup_{j < i} M_j$  ( $i \in \mathbb{N}$ ), then  $\psi_1$  is Borel and  $\psi_1(x) \in \widehat{\pi}^{-1}(x)$  for all  $x$ . Since  $\widehat{\pi} \circ \psi_1 = \text{Id}$ ,  $\psi_1$  is one-to-one. It follows that  $\Upsilon_1 := \psi_1(\Omega)$  is Borel, and  $\widehat{\pi} : \Upsilon_1 \rightarrow \Omega$  is one-to-one and onto.

Now take  $B_2 := B_1 \setminus \text{graph}(\psi_1)$ . Each  $x$ -fibre of  $B_2$  has  $N - 1$  elements, and  $B_2$  is disjoint from  $\text{graph}(\psi_1)$ . Apply the previous process to  $B_2$  to obtain  $\Upsilon_2$ . After  $N$  steps, we are done.

*Step 3.* The restriction of the integrand in (13.1) to  $\Omega$  is Borel measurable.

*Proof.* Every  $x \in \Omega$  has exactly  $N$  pre-images, one in every  $\Upsilon_i$ . It follows that for every Borel set  $E \subset \widehat{\Sigma}$ ,

$$\frac{1}{|\widehat{\pi}^{-1}(x)|} \sum_{\widehat{\pi}(\underline{y})=x} 1_E(\underline{y}) = \frac{1}{N} \sum_{i=1}^N 1_{\widehat{\pi}(E \cap \Upsilon_i)}(x) \text{ on } \Omega.$$

Since  $\widehat{\pi}$  is one-to-one on  $\Upsilon_i$ ,  $\widehat{\pi}(E \cap \Upsilon_i)$  is a Borel set. It follows that the right-hand-side is Borel measurable.

*Step 4.*  $\tilde{\mu}$  is an invariant Borel probability measure such that  $\tilde{\mu} \circ \widehat{\pi}^{-1} = \mu$  and  $h_{\tilde{\mu}}(\sigma) = h_{\mu}(f)$ .

*Proof.* We saw that  $\tilde{\mu}(E)$  is well-defined for all Borel sets  $E \subset \widehat{\Sigma}$ . This set function is obviously  $\sigma$ -additive, and it is clear that  $\tilde{\mu}(\widehat{\Sigma}) = 1$ . Thus  $\tilde{\mu}$  is a Borel probability measure.

This measure is  $\sigma$ -invariant, because

$$\begin{aligned} \tilde{\mu}(\sigma^{-1}E) &= \int_M \left( \frac{1}{|\widehat{\pi}^{-1}(x)|} \sum_{\widehat{\pi}(\underline{R})=x} 1_E(\sigma(\underline{R})) \right) d\mu(x) \\ &= \int_M \left( \frac{1}{|\widehat{\pi}^{-1}(f(x))|} \sum_{\widehat{\pi}(\sigma \underline{R})=f(x)} 1_E(\sigma(\underline{R})) \right) d\mu(x) \quad (\because \widehat{\pi} \circ \sigma = f \circ \widehat{\pi}) \end{aligned}$$

$$\begin{aligned}
&= \int_M \left( \frac{1}{|\widehat{\pi}^{-1}(f(x))|} \sum_{\widehat{\pi}(\underline{S})=f(x)} 1_E(\underline{S}) \right) d\mu(x) \\
&= \tilde{\mu}(E) \quad (\because \mu \circ f^{-1} = \mu).
\end{aligned}$$

It is a lift of  $\mu$  because

$$\tilde{\mu}(\widehat{\pi}^{-1}E) = \int_M \left( \frac{1}{|\widehat{\pi}^{-1}(x)|} \sum_{\widehat{\pi}(\underline{R})=x} 1_E(\widehat{\pi}(\underline{R})) \right) d\mu(x) = \int_M 1_E(x) d\mu(x) = \mu(E).$$

Finally  $\tilde{\mu}$  and  $\mu$  have the same entropy, because  $\widehat{\pi}$  is  $N$ -to-one on a set of full measure, and finite extensions preserve entropy.

*Step 5.* Almost every ergodic component of  $\tilde{\mu}$  satisfies  $\widehat{\mu} \circ \widehat{\pi}^{-1} = \mu$  and  $h_{\widehat{\mu}}(\sigma) = h_{\mu}(\mu)$ .

Let  $\tilde{\mu} = \int \widehat{\mu}_y d\nu(y)$  be the ergodic decomposition of  $\tilde{\mu}$ , then  $\mu = \tilde{\mu} \circ \widehat{\pi}^{-1} = \int \widehat{\mu}_y \circ \widehat{\pi}^{-1} d\nu_y$ . Each of the measures  $\widehat{\mu}_y \circ \widehat{\pi}^{-1}$  is  $f$ -invariant. Since  $\mu$  is ergodic,  $\widehat{\mu}_y \circ \widehat{\pi}^{-1} = \mu$  for a.e.  $y$ .

The equality of the entropies follows as before from the fact that finite extensions preserve entropy.  $\square$

#### Part 4. Appendix: Proofs of standard results in Pesin Theory

**Proof of Theorem 2.3** This is an adaptation of the proof of Theorem 3.5.5 in [BP]. The idea is to evaluate  $A_{\chi}(x) := C_{\chi}(f(x))^{-1} \circ df_x \circ C_{\chi}(x)$  on the standard basis of  $\mathbb{R}^2$ .

We start from the identity  $df_x E^s(x) = E^s(f(x))$ . Both sides of the equation are one-dimensional, therefore  $df_x \underline{e}^s(x) = \pm \|df_x \underline{e}^s(x)\|_{f(x)} \underline{e}^s(f(x))$ . It follows that

$$\begin{aligned}
A_{\chi}(x) \underline{e}_1 &= s_{\chi}(x)^{-1} [C_{\chi}(f(x))^{-1} \circ df_x] \underline{e}^s(x) \\
&= \pm s_{\chi}(x)^{-1} \|df_x \underline{e}^s(x)\|_{f(x)} C_{\chi}(f(x))^{-1} \underline{e}^s(f(x)) \\
&= \pm \frac{s_{\chi}(f(x))}{s_{\chi}(x)} \|df_x \underline{e}^s(x)\|_{f(x)} \underline{e}_1.
\end{aligned}$$

We see that  $\underline{e}_1$  is an eigenvector of  $A_{\chi}(x)$  with eigenvalue

$$\lambda_{\chi}(x) := \pm \frac{s_{\chi}(f(x))}{s_{\chi}(x)} \|df_x \underline{e}^s(x)\|_{f(x)}. \quad (\text{A.1})$$

Similarly,  $\underline{e}_2$  is an eigenvector of  $A_{\chi}(x)$  with eigenvalue

$$\mu_{\chi}(x) := \pm \frac{u_{\chi}(f(x))}{u_{\chi}(x)} \|df_x \underline{e}^u(x)\|_{f(x)}. \quad (\text{A.2})$$

We estimate the eigenvalues:

$$\begin{aligned}
s_\chi(x)^2 &\equiv 2 \sum_{k=0}^{\infty} e^{2k\chi} \|(df^k)_x \underline{e}^s(x)\|_{f^k(x)}^2 > 2 \sum_{k=1}^{\infty} e^{2k\chi} \|(df^k)_x \underline{e}^s(x)\|_{f^k(x)}^2 \\
&= 2 \sum_{k=0}^{\infty} e^{2(k+1)\chi} \|(df^k)_{f(x)} df_x \underline{e}^s(x)\|_{f^{k+1}(x)}^2 \\
&= 2 \|df_x \underline{e}^s(x)\|_{f(x)}^2 \sum_{k=0}^{\infty} e^{2(k+1)\chi} \|(df^k)_{f(x)} \underline{e}^s(f(x))\|_{f^{k+1}(x)}^2 \\
&= e^{2\chi} \|df_x \underline{e}^s(x)\|_{f(x)}^2 s_\chi(f(x))^2.
\end{aligned}$$

Rearranging terms, we find that  $e^{-2\chi} > \frac{s_\chi(f(x))^2}{s_\chi(x)^2} \|df_x \underline{e}^s(x)\|_{f(x)}^2 = \lambda_\chi(x)^2$ . It follows that  $|\lambda_\chi(x)| < e^{-\chi}$ . Similarly, one shows that  $|\mu_\chi(x)| > e^\chi$ .

Since  $f$  is a diffeomorphism, the number  $M_f := \max\{\|df_x\|, \|df_x^{-1}\| : x \in M\}$  is well defined and finite. It is easy to see that  $M_f \geq 1$ . By [KH, Cor. 3.2.10],  $h_{top}(f) \leq 2 \log M_f$ .

By definition of  $s_\chi(x)$ , and the identity  $df_x \underline{e}^s(x) = \pm \|df_x \underline{e}^s(x)\| \underline{e}^s(f(x))$ ,

$$\begin{aligned}
s_\chi(x)^2 &= 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k\chi} \|df_{f(x)}^{k-1} \underline{e}^s(f(x))\|_{f^k(x)}^2 \|df_x \underline{e}^s(x)\|_x^2 \right) \\
&\leq 2 \left( 1 + e^{2\chi} M_f^2 \sum_{k=0}^{\infty} e^{2k\chi} \|df_{f(x)}^k \underline{e}^s(f(x))\|_{f^{k+1}(x)}^2 \right) \\
&\leq 2 + e^{2\chi} M_f^2 s_\chi(f(x))^2 \\
&\leq (M_f^6 + 1) s_\chi(f(x))^2 \quad (\because s_\chi > \sqrt{2} \text{ and } \chi < h_{top}(f) \leq 2 \log M_f).
\end{aligned}$$

Therefore by (A.1)

$$|\lambda_\chi(x)| > (1 + M_f^6)^{-1/2} \|df_x \underline{e}^s(x)\|_{f(x)} \geq M_f^{-1} (1 + M_f^6)^{-1/2}. \quad (\text{A.3})$$

Similarly, one can bound  $|\mu_\chi(x)|$  from above by a function of  $M_f$ .  $\square$

**Proof of Lemma 2.4** We put the standard basis  $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  on  $\mathbb{R}^2$ , and the basis  $\underline{e}^s(x), \underline{e}^s(x)^\perp$  on  $T_x M$ , where  $\underline{v}^\perp$  denotes the unique vector s.t. the signed angle from  $\underline{v}$  to  $\underline{v}^\perp$  is  $\pi/2$ . The linear map  $C_\chi(x) : \mathbb{R}^2 \rightarrow T_x$  is represented in these bases by the matrix

$$\begin{pmatrix} s_\chi(x)^{-1} & u_\chi(x)^{-1} \cos \alpha(x) \\ 0 & u_\chi(x)^{-1} \sin \alpha(x) \end{pmatrix}.$$

Inverting, we find that  $C_\chi(x)^{-1} : T_x M \rightarrow \mathbb{R}^2$  is represented by

$$\begin{pmatrix} s_\chi(x) & -s_\chi(x)/\tan \alpha(x) \\ 0 & u_\chi(x)/\sin \alpha(x) \end{pmatrix}.$$

The lemma follows by direct calculation, using the fact that the Frobenius norm of a linear map represented by a matrix  $(a_{ij})$  is equal to  $(\sum a_{ij}^2)^{1/2}$ .  $\square$

**Proof of Lemma 2.5** Define an inner product  $\langle \cdot, \cdot \rangle_x^*$  on  $T_x M$  by the conditions (a)  $\|\underline{e}^s(x)\|_x^* = s_\chi(x)$ , (b)  $\|\underline{e}^u(x)\|_x^* = u_\chi(x)$ , and (c)  $\langle \underline{e}^u(x), \underline{e}^s(x) \rangle_x^* = 0$  (compare

with [BP, §3.5.1]). The inner product  $\|\cdot\|_x^*$  satisfies  $\|\cdot\|_x^* \geq \|\cdot\|_x$ , because for every  $\xi, \eta \in \mathbb{R}$

$$\begin{aligned} \|\xi \underline{e}^s(x) + \eta \underline{e}^u(x)\|_x^* &= \sqrt{\xi^2 s_\chi(x)^2 + \eta^2 u_\chi(x)^2} > \sqrt{2(\xi^2 + \eta^2)} \quad (\because s_\chi, u_\chi > \sqrt{2}) \\ &\geq |\xi| + |\eta| = \|\xi \underline{e}^s(x)\|_x + \|\eta \underline{e}^u(x)\|_x \geq \|\xi \underline{e}^s(x) + \eta \underline{e}^u(x)\|_x. \end{aligned}$$

$\therefore \|C_\chi(x) \begin{pmatrix} \xi \\ \eta \end{pmatrix}\|_x \leq \|C_\chi(x) \begin{pmatrix} \xi \\ \eta \end{pmatrix}\|_x^* = \|\xi s_\chi(x)^{-1} \underline{e}^s(x) + \eta u_\chi(x)^{-1} \underline{e}^u(x)\|_x^* = \sqrt{\xi^2 + \eta^2}$ .  
The lemma follows.  $\square$

**Proof of Lemma 2.6** Let  $A_\chi(x) := C_\chi(f(x))^{-1} \circ df_x \circ C_\chi(x)$ . Extend  $A_\chi$  to a cocycle  $A_\chi^{(n)}$  using the identities  $A_\chi^{(1)} := A_\chi$  and  $A_\chi^{(m+n)}(x) = A_\chi^{(m)}(f^n(x)) A_\chi^{(n)}(x)$ . The extension is unique, and is given by  $A_\chi^{(n)}(x) = C_\chi(f^n(x))^{-1} df_x^n C_\chi(x)$ .

Theorem 2.3 says that  $A_\chi(x)$  is a diagonal matrix with entries in  $[C_f^{-1}, C_f]$  for every  $x \in \text{NUH}_\chi(f)$ . In particular,  $\log \|A_\chi^{(1)}\|$  and  $\log \|(A_\chi^{(0)})^{-1}\|$  are uniformly bounded on  $\text{NUH}_\chi(f)$ , whence absolutely integrable w.r.t any ergodic invariant probability measure with entropy larger than  $\chi$ . This allows us to apply the Multiplicative Ergodic Theorem to  $A_\chi^{(n)}$  w.r.t. every ergodic invariant probability measure with entropy larger than  $\chi$ .

Let  $\text{NUH}_\chi^\dagger(f)$  denote the set of points  $x \in \text{NUH}_\chi(f)$  for which for every  $y \in \{f^k(x) : k \in \mathbb{Z}\}$  there is a decomposition  $T_y \mathbb{R}^2 = E_\chi^s(y) \oplus E_\chi^u(y)$  so that

- (1)  $E_\chi^s(y) = \text{span}\{\underline{e}_\chi^s(y)\}$ ,  $\|\underline{e}_\chi^s(y)\| = 1$ ,  $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|A_\chi^{(n)}(y) \underline{e}_\chi^s(y)\| < 0$ ;
- (2)  $E_\chi^u(y) = \text{span}\{\underline{e}_\chi^u(y)\}$ ,  $\|\underline{e}_\chi^u(y)\| = 1$ ,  $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|A_\chi^{(n)}(y) \underline{e}_\chi^u(y)\| > 0$ ;
- (3)  $\lim_{n \rightarrow \infty} \frac{1}{n} \log |\sin \alpha_\chi(f^n(y))| = 0$ , where  $\alpha_\chi(y) := \angle(\underline{e}_\chi^s(y), \underline{e}_\chi^u(y))$ ;
- (4)  $A_\chi(x)[E_\chi^s(y)] = E_\chi^s(f(y))$  and  $A_\chi(y)[E_\chi^u(y)] = E_\chi^u(f(y))$ .

By the discussion above,  $\text{NUH}_\chi^\dagger(f)$  has full measure w.r.t. to any ergodic invariant probability measure with entropy larger than  $\chi$ .

Let  $\text{NUH}_\chi^*(f)$  denote the subset of  $\text{NUH}_\chi^\dagger(f)$  which consists of all points  $x$  for which there exist a sequence  $n_k \uparrow \infty$  s.t.  $C_\chi(f^{n_k}(x)) \xrightarrow[k \rightarrow \infty]{} C_\chi(x)$  and a sequence  $m_k \downarrow -\infty$  s.t.  $C_\chi(f^{m_k}(x)) \xrightarrow[k \rightarrow \infty]{} C_\chi(x)$ . By the Poincaré Recurrence Theorem, every invariant probability measure which is carried by  $\text{NUH}_\chi^\dagger(f)$  is carried by  $\text{NUH}_\chi^*(f)$ , so  $\text{NUH}_\chi^*(f)$  has full measure w.r.t. to every ergodic invariant measure with entropy greater than  $\chi$ .

Applying the Multiplicative Ergodic Theorem to the cocycles  $df_x$  and  $A_\chi^{(n)}(x)$  on  $\text{NUH}_\chi^*(f)$ , we obtain the existence of the following limits:

$$\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|df_x^n C_\chi(x) \underline{e}_i\|_{f^n(x)}, \quad \lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|C_\chi(f^n(x))^{-1} df_x^n C_\chi(x) \underline{e}_i\|. \quad (\text{A.4})$$

Let  $n_k \uparrow \infty$  be a subsequence for which  $C_\chi(f^{n_k}(x)) \xrightarrow[k \rightarrow \infty]{} C_\chi(x)$ . The norms of  $C_\chi(f^{n_k}(x))$  and  $C_\chi(f^{n_k}(x))^{-1}$  are bounded along this sequence, so

$$\|C_\chi(f^{n_k}(x))^{-1} df_x^{n_k} C_\chi(x) \underline{e}_i\| \asymp \|df_x^{n_k} C_\chi(x) \underline{e}_i\|.$$

We see that the limits in (A.4) agree. As a result  $E_\chi^s(x) = \mathbb{R} \times \{\underline{0}\}$ ,  $E_\chi^u(x) = \{\underline{0}\} \times \mathbb{R}$ , and  $x$  has Lyapunov exponents  $\log \lambda(x)$  and  $\log \mu(x)$  w.r.t.  $A_\chi^{(n)}$ .

Let  $\Lambda_\chi(x) := \begin{pmatrix} \lambda(x) & 0 \\ 0 & \mu(x) \end{pmatrix}$ , then the limits (A.4) mean that

$$\|(A_\chi^{(n)}(x)\Lambda_\chi(x)^{-n})^{\pm 1}\|^{1/n} \xrightarrow{n \rightarrow \pm\infty} 1.$$

Similarly, if  $\Lambda(x)$  is the linear operator s.t.  $\Lambda(x)\underline{e}^s(x) = \lambda(x)\underline{e}^s(x)$  and  $\Lambda(x)\underline{e}^u(x) = \mu(x)\underline{e}^u(x)$ , then

$$\|(df_x^n \Lambda(x)^{-n})^{\pm 1}\|^{1/n} \xrightarrow{n \rightarrow \pm\infty} 1.$$

Since  $\Lambda_\chi(x) = C_\chi(x)^{-1}\Lambda(x)C_\chi(x)$  and  $A_\chi^{(n)}(x) = C_\chi(f^n(x))^{-1} \circ df_x^n \circ C_\chi(x)$ ,

$$\begin{aligned} \|C_\chi(f^n(x))^{-1}\|^{1/n} &= \|A_\chi^{(n)}(x)C_\chi(x)^{-1}(df_x^n)^{-1}\|^{1/n} \\ &= \|A_\chi^{(n)}(x)C_\chi(x)^{-1}\Lambda(x)^{-n}C_\chi(x) \cdot C_\chi(x)^{-1} \cdot \Lambda(x)^n(df_x^n)^{-1}\|^{1/n} \\ &\leq \|A_\chi^{(n)}(x)\Lambda_\chi(x)^{-n}\|^{1/n} \|C_\chi(x)^{-1}\|^{1/n} \|(df_x^n \Lambda(x)^{-n})^{-1}\|^{1/n} \rightarrow 1. \end{aligned}$$

Thus  $\limsup \frac{1}{n} \log \|C_\chi(f^n(x))^{-1}\| \leq 0$ . On the other hand  $C_\chi$  is a contraction (Lemma 2.5), so  $\|C_\chi(f^n(x))^{-1}\|^{1/n} \geq 1$ , whence  $\liminf \frac{1}{n} \log \|C_\chi(f^n(x))^{-1}\| \geq 0$ . The first part of the Lemma is proved.

We prove the second part of the Lemma:  $\frac{1}{n} \log \|C_\chi(f^n(x))\underline{e}_i\|_{f^n(x)} \xrightarrow{n \rightarrow \pm\infty} 0$ .

We do this for  $i = 1$ , and leave the case  $i = 2$  to the reader. Since the  $A_\chi^{(n)}(\cdot)$  is diagonal,  $A_\chi^{(n)}(x)\underline{e}_1$  is proportional to  $\underline{e}_1$ . The multiplicative ergodic theorem for  $A_\chi^{(n)}(x)$  says that  $A_\chi^{(n)}(x)\underline{e}_1 = \pm\lambda(x)^n \exp[o(n)]\underline{e}_1$ , therefore

$$\begin{aligned} \lim_{n \rightarrow \pm\infty} \|C_\chi(f^n(x))\underline{e}_1\|_{f^n(x)}^{1/n} &= \lambda(x)^{-1} \lim_{n \rightarrow \pm\infty} \|C_\chi(f^n(x))A_\chi^{(n)}(x)\underline{e}_1\|_{f^n(x)}^{1/n} \\ &= \lambda(x)^{-1} \lim_{n \rightarrow \pm\infty} \|(df_x^n)C_\chi(x)\underline{e}_1\|_{f^n(x)}^{1/n} \\ &= \lambda(x)^{-1} \lim_{n \rightarrow \pm\infty} \|(df_x^n)\underline{e}^s(x)\|_{f^n(x)}^{1/n} = 1, \end{aligned}$$

proving that  $\frac{1}{n} \log \|C_\chi(f^n(x))\underline{e}_1\|_{f^n(x)} \xrightarrow{n \rightarrow \pm\infty} 0$ .

Finally, we prove that  $\frac{1}{n} \log |\det C_\chi(f^n(x))| \xrightarrow{n \rightarrow \pm\infty} 0$ . We begin with some general comments on determinants.

Suppose  $L : V \rightarrow W$  is a linear operator between two 2-dimensional vector spaces with inner product. The determinant of  $L$  can be defined as  $\det(L\Theta)$  for some (every) isometry  $\Theta : W \rightarrow V$ . The following fact holds:<sup>8</sup> If  $\underline{u}, \underline{v}$  span  $V$ , then

$$\frac{\sin \angle(L\underline{u}, L\underline{v})}{\sin \angle(\underline{u}, \underline{v})} = \frac{\|\underline{u}\|\|\underline{v}\| \det L}{\|L\underline{u}\|\|L\underline{v}\|}. \quad (\text{A.5})$$

It follows that  $|\det L| = \frac{\|L\underline{u}\|\|L\underline{v}\| \sin \angle(L\underline{u}, L\underline{v})}{\|\underline{u}\|\|\underline{v}\| \sin \angle(\underline{u}, \underline{v})}$  ( $\underline{u}, \underline{v}$  independent).

Applying this to  $L = A_\chi^{(n)}$  with  $\underline{u} = \underline{e}_1$ ,  $\underline{v} = \underline{e}_2$ , and to  $L = df_x^n$  with  $\underline{u} = \underline{e}^s(x)$ ,  $\underline{v} = \underline{e}^u(x)$ , we find that

$$\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log |\det A_\chi^{(n)}(x)| = \log \lambda(x) + \log \mu(x) = \lim_{n \rightarrow \pm\infty} \frac{1}{n} \log |\det df_x^n|.$$

<sup>8</sup>Proof: Let  $\omega_V, \omega_W$  denote the volume 2-forms on  $V, W$ , then  $\omega_V(\underline{u}, \underline{v}) = \|\underline{u}\|\|\underline{v}\| \sin \angle(\underline{u}, \underline{v})$  and  $\omega_W(\underline{u}, \underline{v}) = \|\underline{u}\|\|\underline{v}\| \sin \angle(\underline{u}, \underline{v})$ . Since  $\omega_W(L\underline{u}, L\underline{v})$  is also a 2-form on  $V$ , and any two 2-forms on  $V$  are proportional,  $\exists c$  s.t.  $\omega_W(L\underline{u}, L\underline{v}) = c\omega_V(\underline{u}, \underline{v})$ . Evaluating on an orthonormal basis of  $V$ , we find that  $c = \det L$ . Consequently,  $\|L\underline{u}\|\|L\underline{v}\| \sin \angle(L\underline{u}, L\underline{v}) = \det L \|\underline{u}\|\|\underline{v}\| \sin \angle(\underline{u}, \underline{v})$ .

Since  $|\det A_\chi^{(n)}(x)| = |\det C_\chi(f^n(x))|^{-1} |\det df_x^n| |\det C_\chi(x)|$ ,  
 $\frac{1}{n} \log |\det C_\chi(f^n(x))| \xrightarrow{n \rightarrow \infty} 0$  as required.  $\square$

**Proof of Lemma 2.9** Parts (1) and (3) are obvious, and part (4) is a consequence of Lemma 2.6 and the estimate  $Q_\varepsilon(f^n(x)) \asymp \|C_\chi(f^n(x))^{-1}\|^{-12/\beta}$ . For part (6), define  $q_\varepsilon(x)$  on  $\text{NUH}^*(f)$  by the formula

$$\frac{1}{q_\varepsilon(x)} = \frac{1}{\varepsilon} \sum_{k=-\infty}^{\infty} e^{-\frac{1}{3}|k|\varepsilon} \frac{1}{Q_\varepsilon(f^k(x))}.$$

The sum converges because  $\frac{1}{k} \log Q_\varepsilon(f^k(x)) \xrightarrow{k \rightarrow \pm\infty} 0$ , and it is easy to check that  $q_\varepsilon(x)$  behaves as required, see [BP, Lemma 3.5.7].

It remains to prove parts (2) and (5). First we prove the following claim.

*Claim.* There exists a constant  $C$ , which only depends on  $M, f$  and  $\chi$ , such that  $C^{-1} \leq \|C_\chi(f(x))^{-1}\| / \|C_\chi(x)^{-1}\| \leq C$  on  $\text{NUH}_\chi(f)$ .

*Proof.* By Lemma 2.4 it is enough to show that

$$\frac{s_\chi \circ f}{s_\chi}, \quad \frac{u_\chi \circ f}{u_\chi}, \quad \frac{|\sin \alpha \circ f|}{|\sin \alpha|}$$

are uniformly bounded away from zero and infinity on  $\text{NUH}_\chi(f)$ .

The following quantity is well defined and finite, because  $f$  is a diffeomorphism and  $M$  is compact:

$$F_0 := \max\{\|df_x\|, \|df_x^{-1}\|, |\det(df_x)|, |\det(df_x^{-1})| : x \in M\}.$$

Notice that  $F_0 > 1$ .

Equation (A.1) makes it clear that  $\frac{s_\chi(f(x))}{s_\chi(x)} = F_0^{\pm 1} |\lambda_\varepsilon(x)| \in [(C_f F_0)^{-1}, C_f F_0]$  on  $\text{NUH}_\chi(f)$ . Similarly,  $\frac{u_\chi(f(x))}{u_\chi(x)}$  takes values in  $[(C_f F_0)^{-1}, C_f F_0]$  on  $\text{NUH}_\chi(f)$ . Finally, by (A.5) and the fact that  $e^{s/u}(f(x))$  have the same direction as  $df_x e^{s/u}(x)$  up to a sign,

$$\frac{|\sin \alpha(f(x))|}{|\sin \alpha(x)|} = \frac{|\sin \angle(\underline{e}^s(f(x)), \underline{e}^u(f(x)))|}{|\sin \angle(\underline{e}^s(x), \underline{e}^u(x))|} = \frac{|\det df_x|}{\|df_x \underline{e}^s(x)\| \|df_x \underline{e}^u(x)\|}.$$

The last quantity takes values in  $[F_0^{-3}, F_0^3]$ . The claim follows.

Part (5) follows directly from the claim. For part (2), we start by noting that  $Q_\varepsilon(x) < \varepsilon^{3/\beta} \|C_\chi(x)^{-1}\|_{Fr}^{-12/\beta} < \varepsilon^{3/\beta} \|C_\chi(x)^{-1}\|^{-12}$ , therefore also  $Q_\varepsilon(x) < (\varepsilon^{3/\beta} C^{12/\beta}) \cdot \|C_\chi(f^{\pm 1}(x))^{-1}\|^{-12}$ . If  $\varepsilon$  is small enough then  $\varepsilon^{1/\beta} C^{12/\beta} < 1$ , and the proof of part (2) is complete.  $\square$

**Proof of Theorem 2.7** What follows is based on [BP, Theorem 5.6.1].

Recall the following basic fact from differential geometry [Sp, chapter 9]: Every  $p \in M$  has an open neighborhood  $W_p$  and a positive number  $r > 0$  s.t.

- (1) any  $q, q' \in W_p$  are connected by a unique geodesic of length less than  $r$ ;
- (2) for each  $q \in W_p$ ,  $\exp_q$  maps  $B_r^q(\underline{0}) \subset T_q M$  diffeomorphically onto an open set  $U_q \supseteq W_p$  in a 2-bi-Lipschitz way, and  $d(\exp_q)_\underline{0} = \text{Id}$ ;
- (3) for every  $q, q' \in W_p$ , there is a unique vector  $\underline{v}(q, q') \in T_q M$  s.t.  $\|\underline{v}(q, q')\|_q < r$  and  $\exp_q[\underline{v}(q, q')] = q'$ ;
- (4)  $(q, q') \mapsto \underline{v}(q, q')$  is a well-defined  $C^\infty$  map from  $W_p \times W_p$  to  $M$ .

Since  $M$  is compact, there exist positive constants  $r(M), \rho(M)$  s.t. for every  $p \in M$ ,  $\exp_p$  maps  $B_{r(M)}^p(\underline{0}) \subseteq T_p M$  diffeomorphically onto a neighborhood of  $B_{\rho(M)}(p) \subset M$ , in a 2-bi-Lipschitz way. Let

$$r_0 := \frac{\min\{1, r(M), \rho(M)\}}{10[\text{Lip}(f) + \text{Lip}(f^{-1})]}. \quad (\text{A.6})$$

Note that  $r_0 < 1$ .

Suppose  $\varepsilon < r_0/5$ . By the definition of  $Q_\varepsilon(x)$ ,  $Q_\varepsilon(x) < \varepsilon^3$ , so  $10Q_\varepsilon(x) < r_0/\sqrt{2}$ . By Lemma 2.5,  $C_\chi(x)$  maps  $R_{10Q_\varepsilon(x)}(\underline{0})$  contractively into  $B_{r_0}(\underline{0})$ . Therefore  $\Psi_x = \exp_x \circ C_\chi(x)$  maps  $R_{10Q_\varepsilon(x)}(\underline{0})$  diffeomorphically in a 2-Lipschitz way into  $M$ . The first part of the theorem is proved.

Next we show that  $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$  is well defined on  $R_{10Q_\varepsilon(x)}(\underline{0})$  and establish its properties.

Since  $\exp_x$  is 2-Lipschitz,  $C_\chi(x)$  is a contraction, and  $10Q_\varepsilon(x) < r_0/\sqrt{2}$ ,

$$\Psi_x \text{ maps } R_{10Q_\varepsilon(x)}(\underline{0}) \text{ diffeomorphically into } B_{2r_0}(x).$$

It follows that  $f \circ \Psi_x$  maps  $R_{10Q_\varepsilon(x)}(\underline{0})$  diffeomorphically into  $B_{2\text{Lip}(f)r_0}(f(x))$ , which by the definition of  $r_0$  is a subset of  $B_{\rho(M)}(f(x))$ , whence a subset of  $\exp_{f(x)}[B_{r(M)}^x(\underline{0})]$ . It follows that  $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$  is well defined, smooth and injective on  $R_{10Q_\varepsilon(x)}(\underline{0})$ .

For every  $p \in M$ ,  $\exp_p(\underline{0}) = p$  and  $d(\exp_p)_\underline{0} = \text{Id}$ . It easily follows that  $f_x(\underline{0}) = \underline{0}$ , and  $(df_x)_\underline{0} = C_\chi(f(x))^{-1} \circ (df)_{f(x)} \circ C_\chi(x)$ . By Theorem 2.3, this is a diagonal matrix with diagonal elements  $A(x) = \lambda_\varepsilon(x)$ ,  $B(x) = \mu_\varepsilon(x)$ , and  $C_f^{-1} < |A(x)| < e^{-x}$ ,  $e^x < |B(x)| < C_f$ .

We compare  $f_x$  to its linearization at  $\underline{0}$  by analyzing

$$r_x(\underline{u}) := f_x(\underline{u}) - (df_x)_\underline{0}(\underline{u}).$$

By assumption  $f$  is  $C^{1+\beta}$ , so there is a constant  $L$  s.t. for all  $\underline{u}, \underline{v} \in R_{r_0}(\underline{0})$ ,  $\|d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_\underline{u} - d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_\underline{v}\| \leq L\|\underline{u} - \underline{v}\|^\beta$ . For every  $\underline{u}, \underline{v} \in R_{r_0}(\underline{0})$ ,

$$\begin{aligned} \|(dr_x)_\underline{u} - (dr_x)_\underline{v}\| &= \|C_\chi(f(x))^{-1} d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_{C_\chi(x)\underline{u}} C_\chi(x) \\ &\quad - C_\chi(f(x))^{-1} d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_{C_\chi(x)\underline{v}} C_\chi(x)\| \\ &= \|C_\chi(f(x))^{-1} [d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_{C_\chi(x)\underline{u}} \\ &\quad - d(\exp_{f(x)}^{-1} \circ f \circ \exp_x)_{C_\chi(x)\underline{v}}] C_\chi(x)\| \\ &\leq \|C_\chi(f(x))^{-1}\| \cdot L \|C_\chi(x)\|^\beta \|\underline{u} - \underline{v}\|^\beta \cdot \|C_\chi(x)\| \\ &\leq (\|C_\chi(f(x))^{-1}\| \cdot L \|\underline{u} - \underline{v}\|^{\beta/2}) \cdot \|\underline{u} - \underline{v}\|^{\beta/2} \quad (\because \|C_\chi(x)\| < 1). \end{aligned}$$

If  $\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\underline{0})$ , then the term in the brackets is smaller than

$$\|C_\chi(f(x))^{-1}\| \cdot L(20\sqrt{2}Q_\varepsilon(x))^{\beta/2}.$$

Plugging in the definition of  $Q_\varepsilon(x)$  from (2.3), and recalling that  $\|C_\chi(\cdot)^{-1}\| > 1$  (because  $C_\chi(\cdot)$  is a contraction), we see that the term in the brackets is smaller than  $30^{\beta/2} L \varepsilon^{3/2}$ . Thus, if  $\varepsilon < \frac{1}{3} \cdot 30^{-\beta/2} L^{-1}$ , then

$$\|(dr_x)_\underline{u} - (dr_x)_\underline{v}\| \leq \frac{1}{3} \varepsilon \|\underline{u} - \underline{v}\|^{\beta/2} \quad (\underline{u}, \underline{v} \in R_{10Q_\varepsilon(x)}(\underline{0})).$$



Since  $(dr_x)_0 = 0$ , we have that  $\|(dr_x)_u\| \leq \frac{1}{3}\varepsilon\|u\|^{\beta/2}$  on  $R_{10Q_\varepsilon(x)}(\underline{0})$ . Now  $Q_\varepsilon(x) < \varepsilon^{3/\beta}$ , so  $\|u\| \leq (10\sqrt{2})Q_\varepsilon(x) < 15\varepsilon^{3/\beta}$ . If  $\varepsilon < 15^{-\beta/3}$ , then  $\|u\| < 1$ , so

$$\|(dr_x)_u\| \leq \frac{1}{3}\varepsilon \text{ on } R_{10Q_\varepsilon(x)}(\underline{0}).$$

Since  $r_x(\underline{0}) = \underline{0}$ , we have by the mean value theorem that

$$\|r_x(u)\| \leq \frac{1}{3}\varepsilon\|u\| < \frac{1}{3}\varepsilon \text{ on } R_{10Q_\varepsilon(x)}(\underline{0}).$$

In summary, if  $\varepsilon$  is small enough, then the  $C^{1+\beta/2}$ -distance between  $r_x$  and 0 on  $R_{10Q_\varepsilon(x)}(\underline{0})$  is less than  $\varepsilon$ . This shows that the  $C^{1+\beta/2}$ -distance between  $f_x$  and  $(df_x)_0$  on this set is less than  $\varepsilon$ .

The treatment of  $f_x^{-1}$  is similar, and is left to the reader.  $\square$

**Proof of Proposition 4.11** The proof of parts (1),(2) and (3) of the proposition is taken from [KM]. Part (4) is new, but routine.

Assume that  $0 < \varepsilon < \frac{1}{2}$ . Write  $V^u = \Psi_x\{(F(w), w) : |w| \leq p^u\}$  and  $V^s = \Psi_x\{(v, G(v)) : |v| \leq p^s\}$ , and let  $\eta := p^u \wedge p^s$ . Note that  $\eta < \varepsilon$ , and that  $|F(0)|, |G(0)| \leq 10^{-3}\eta$  and  $\text{Lip}(F), \text{Lip}(G) \leq \varepsilon$ , see (4.1).

The maps  $H = F, G$  are contractions (with Lipschitz constant less than  $\varepsilon$ ), and they map the interval  $[-10^{-2}\eta, 10^{-2}\eta]$  into itself, because for every  $|t| < 10^{-2}\eta$ ,

$$|H(t)| \leq |H(0)| + \text{Lip}(H)|t| < 10^{-3}\eta + \varepsilon \cdot 10^{-2}\eta = (10^{-1} + \varepsilon)10^{-2}\eta < 10^{-2}\eta.$$

It follows that  $G \circ F$  is a  $\varepsilon^2$ -contraction of  $[-10^{-2}\eta, 10^{-2}\eta]$  into itself. By the Banach Fixed Point Theorem,  $G \circ F$  has a unique fixed point:  $(G \circ F)(w) = w$ .

Let  $v := F(w)$ . We claim that  $V^u, V^s$  intersect at  $P := \Psi_x(v, w)$ .

- $P \in V^u$ , because  $v = F(w)$  and  $|w| \leq 10^{-2}\eta < p^u$ ;
- $P \in V^s$ , because  $w = (G \circ F)(w) = G(v)$ , and  $|v| < |F(0)| + \text{Lip}(F)|w| \leq 10^{-3}\eta + \varepsilon \cdot 10^{-2}\eta < 10^{-2}\eta < p^s$ .

We also see that  $|v|, |w| \leq 10^{-2}\eta$ .

We claim that  $P$  is the unique intersection point of  $V^u$  and  $V^s$ . Let  $\xi := p^u \vee p^s$  and extend  $F, G$  (arbitrarily) to  $\varepsilon$ -Lipschitz continuous functions  $\tilde{F}, \tilde{G} : [-\xi, \xi] \rightarrow [-Q_\varepsilon(x), Q_\varepsilon(x)]$ . Let  $\tilde{V}^u$  and  $\tilde{V}^s$  denote the  $u/s$ -sets represented by  $\tilde{F}, \tilde{G}$ . Any intersection point of  $V^u, V^s$  is an intersection point of  $\tilde{V}^u, \tilde{V}^s$ . Such points take the form  $\tilde{P} = \Psi_x(\tilde{v}, \tilde{w})$  where  $\tilde{v} = \tilde{F}(\tilde{w})$  and  $\tilde{w} = \tilde{G}(\tilde{v})$ . Notice that  $\tilde{w}$  is a fixed point of  $\tilde{G} \circ \tilde{F}$ . The same calculations as before show that  $\tilde{G} \circ \tilde{F}$  contracts  $[-\xi, \xi]$  into itself. Such a map has a unique fixed point, therefore  $\tilde{w} = w$ , whence  $\tilde{P} = P$ .

Next we show that  $P$  is a Lipschitz function of  $V^u, V^s$ . Suppose  $V_i^u, V_i^s$  ( $i = 1, 2$ ) are represented by  $F_i$  and  $G_i$  ( $i = 1, 2$ ) respectively. Let  $P_i$  denote the intersection points of  $V_i^u \cap V_i^s$ . We saw above that  $P_i = \Psi_x(v_i, w_i)$  where  $w_i$  is a fixed point of  $G_i \circ F_i : [-10^{-2}\eta, 10^{-2}\eta] \rightarrow [-10^{-2}\eta, 10^{-2}\eta]$ . The maps  $f_i := G_i \circ F_i$  are  $\varepsilon^2$ -contractions of  $[-10^{-2}\eta, 10^{-2}\eta]$  into itself, therefore

$$\begin{aligned} |w_1 - w_2| &= |f_1^n(w_1) - f_2^n(w_2)| \leq |f_1(f_1^{n-1}(w_1)) - f_2(f_1^{n-1}(w_1))| \\ &\quad + |f_2(f_1^{n-1}(w_1)) - f_2(f_2^{n-1}(w_2))| \\ &\leq \|f_1 - f_2\|_\infty + \varepsilon^2 |f_1^{n-1}(w_1) - f_2^{n-1}(w_2)| \\ &\leq \dots \leq \|f_1 - f_2\|_\infty (1 + \varepsilon^2 + \dots + \varepsilon^{2(n-1)}) + \varepsilon^{2n} |w_1 - w_2| \end{aligned}$$

Passing to the limit as  $n \rightarrow \infty$  we obtain  $|w_1 - w_2| \leq (1 - \varepsilon^2)^{-1} \|f_1 - f_2\|_\infty$ . Similarly,  $v_i$  is a fixed point of  $F_i \circ G_i : [-10^{-2}\eta, 10^{-2}\eta] \rightarrow [-10^{-2}\eta, 10^{-2}\eta]$ , and

the same argument gives that  $|v_1 - v_2| \leq (1 - \varepsilon^2)^{-1} \|g_1 - g_2\|_\infty$  where  $g_i = F_i \circ G_i$ . Since  $\Psi_x$  is 2-Lipschitz, this means that

$$d(P_1, P_2) < \frac{2}{1 - \varepsilon^2} (\|G_1 \circ F_1 - G_2 \circ F_2\|_\infty + \|F_1 \circ G_1 - F_2 \circ G_2\|_\infty).$$

Now

$$\begin{aligned} \|F_1 \circ G_1 - F_2 \circ G_2\|_\infty &\leq \|F_1 \circ G_1 - F_1 \circ G_2\|_\infty + \|F_1 \circ G_2 - F_2 \circ G_2\|_\infty \\ &\leq \text{Lip}(F_1) \|G_1 - G_2\|_\infty + \|F_1 - F_2\|_\infty \\ \|G_1 \circ F_1 - G_2 \circ F_2\|_\infty &\leq \text{Lip}(G_1) \|F_1 - F_2\|_\infty + \|G_1 - G_2\|_\infty \end{aligned}$$

Since  $\text{Lip}(F_i), \text{Lip}(G_i) \leq \varepsilon^2$ ,  $d(P_1, P_2) < \frac{2(1+\varepsilon)}{1-\varepsilon^2} [\text{dist}(V_1^u, V_2^u) + \text{dist}(V_1^s, V_2^s)]$ . The coefficient is less than 3 for all  $\varepsilon$  small enough. For such  $\varepsilon$ ,  $P$  is a 3-Lipschitz function of  $V^u, V^s$ .

Finally, we analyze the angle of intersection at  $P$ . We assume throughout that  $\varepsilon$  is so small that  $0 < t \leq \varepsilon \implies e^{-2t} < 1 - t < 1 + t < e^{2t}$ . In what follows we drop the subscript  $x$  in  $\|\cdot\|_x$ .

Let  $\underline{v} = (v, w)$  be the  $\Psi_x$ -coordinates of  $P$  (i.e.  $P = \Psi_x(\underline{v})$ ), and write  $E^s = E^s(x)$ ,  $E^u = E^u(x)$ . The following identities hold:

$$\begin{aligned} \angle(E^s, E^u) &= \angle((d\Psi_x)_0 \underline{e}^1, (d\Psi_x)_0 \underline{e}^2), \text{ where } \underline{e}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ and } \underline{e}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \angle(V^s, V^u) &= \angle((d\Psi_x)_v \underline{v}^s, (d\Psi_x)_v \underline{v}^u), \text{ where } \underline{v}^s = \begin{pmatrix} 1 \\ G'(v) \end{pmatrix} \text{ and } \underline{v}^u = \begin{pmatrix} F'(w) \\ 1 \end{pmatrix}. \end{aligned}$$

It is not difficult to see that the admissibility of  $V^s, V^u$  and the inequalities  $|v|, |w| < 10^{-2}\eta$  imply that  $|F'(w)|, |G'(v)| < \eta^{\beta/3}$ .

We begin with the estimate of  $\frac{\sin \angle(V^s, V^u)}{\sin \angle(E^s, E^u)} = \frac{\sin \angle((d\Psi_x)_v \underline{v}^s, (d\Psi_x)_v \underline{v}^u)}{\sin \angle((d\Psi_x)_0 \underline{e}^1, (d\Psi_x)_0 \underline{e}^2)}$ . By (A.5),

$$\frac{\sin \angle(V^s, V^u)}{\sin \angle(E^s, E^u)} = \frac{\sin \angle(\underline{v}^s, \underline{v}^u)}{\sin \angle(\underline{e}^1, \underline{e}^2)} \cdot \frac{\|\underline{v}^s\| \|\underline{v}^u\|}{\|\underline{e}^1\| \|\underline{e}^2\|} \cdot \frac{\det(d\Psi_x)_v}{\det(d\Psi_x)_0} \cdot \frac{\|(d\Psi_x)_0 \underline{e}^1\| \|(d\Psi_x)_0 \underline{e}^2\|}{\|(d\Psi_x)_v \underline{v}^s\| \|(d\Psi_x)_v \underline{v}^u\|}.$$

**First factor:** The first factor equals  $\sin \angle(\underline{v}^s, \underline{v}^u)$ . Using the formula for the sine of the difference of two angles it is not difficult to see that

$$\sin \angle(\underline{v}^s, \underline{v}^u) = \frac{1}{\|\underline{v}^s\| \|\underline{v}^u\|} \det \begin{pmatrix} 1 & F'(w) \\ G'(v) & 1 \end{pmatrix}.$$

Since  $|G'(v)|, |F'(w)| < \eta^{\beta/3}$ , the first factor is  $e^{\pm 2\eta^{2\beta/3}}$ .

**Second factor:** Since  $|G'(v)|, |F'(w)| < \eta^{\beta/3}$ , the numerator is  $e^{\pm \eta^{2\beta/3}}$ . Since the denominator is equal to one, the second factor is  $e^{\pm \eta^{2\beta/3}}$ .

**Third factor:**  $\det(d\Psi_x)_v = \det(d\exp_x)_{C_\chi(x)v} \cdot \det C_\chi(x)$ , and  $\det(d\Psi_x)_0 = \det C_\chi(x)$ , therefore the third factor is equal to  $\det(d\exp_x)_{C_\chi(x)v}$ .

The exponential map on  $M$  is smooth, and  $\det(d\exp_x)_0 = 1$ , therefore there exists a constant  $K_1$  which only depends on  $M$  s.t.

$$|\det[(d\exp_x)_v] - 1| < K_1 \|\underline{v}\| \text{ for all } x \in M \text{ and } \|\underline{v}\| < 1.$$

Since  $C_\chi(x)$  is a contraction (Lemma 2.5) and  $\|\underline{v}\| < 2\eta$ ,  $\det(d\exp_x)_{C_\chi(x)v} = 1 \pm 2K_1\eta$ . Since  $0 < \eta < \varepsilon$ ,  $2K_1\eta \ll \sqrt{\eta}$  for all  $\varepsilon$  small enough. For such  $\varepsilon$ , the third factor is  $e^{\pm \sqrt{\eta}}$  (provided  $\varepsilon$  is small enough).

**Fourth factor:** Find a global constant  $K_2$  s.t.  $\|(\Theta_D d \exp_x)_{\underline{u}} - \text{Id}\| < K_2 \|\underline{u}\|$  for all  $x \in D \in \mathcal{D}$  and  $\|\underline{u}\| < 1$  (cf. §3.1).

Write  $\underline{u} = C_\chi(x)\underline{v}$ , and choose some  $D \in \mathcal{D}$  which contains  $\Psi_x[R_{Q_\varepsilon(x)}(0)]$ , then

$$\begin{aligned} \|\Theta_D(d\Psi_x)_{\underline{v}\underline{v}^s} - \Theta_D(d\Psi_x)_{\underline{0}\underline{e}^1}\| &\leq \|\Theta_D(d\Psi_x)_{\underline{v}} - \Theta_D(d\Psi_x)_{\underline{0}}\| \|\underline{v}^s\| \\ &\quad + \|\Theta_D(d\Psi_x)_{\underline{0}}\| \|\underline{v}^s - \underline{e}^1\| \\ &\leq \|\Theta_D(d \exp_x)_{\underline{u}} - \text{Id}\| \|C_\chi(x)\| \|\underline{v}^s\| \\ &\quad + 2\|C_\chi(x)\| \|\underline{v}^s - \underline{e}^1\| \\ &< 3K_2\eta + 2\eta^{\beta/3}, \end{aligned} \tag{A.7}$$

because  $C_\chi(x)$  is a contraction,  $\|\underline{v}\| < 2\eta$ , and  $\underline{v}^s = (\frac{1}{0 \pm \eta^{\beta/3}})$ . Consequently,  $\|(\Theta_D(d\Psi_x)_{\underline{v}\underline{v}^s}) - (\Theta_D(d\Psi_x)_{\underline{0}\underline{e}^1})\| < (3K_2 + 2)\eta^{\beta/3}$ . Since also

$$\|(\Theta_D(d\Psi_x)_{\underline{0}\underline{e}^1})\| = \|C_\chi(x)\underline{e}^1\| \geq \|C_\chi(x)^{-1}\|^{-1}, \tag{A.8}$$

$$\left| \frac{\|(\Theta_D(d\Psi_x)_{\underline{v}\underline{v}^s})\|}{\|(\Theta_D(d\Psi_x)_{\underline{0}\underline{e}^1})\|} - 1 \right| < (3K_2 + 2)\|C_\chi(x)^{-1}\|\eta^{\beta/3}.$$

Since  $\eta \leq Q_\varepsilon(x)$  and  $Q_\varepsilon(x) < \varepsilon^{3/\beta}\|C_\chi(x)^{-1}\|^{-12/\beta}$ ,

$$\|C_\chi(x)^{-1}\|\eta^{\beta/3} \leq \|C_\chi(x)^{-1}\|\eta^{\beta/12} \cdot \eta^{\beta/4} < \varepsilon^{1/4}\eta^{\beta/4}. \tag{A.9}$$

It follows that for all  $\varepsilon$  small enough,  $\frac{\|(\Theta_D(d\Psi_x)_{\underline{v}\underline{v}^s})\|}{\|(\Theta_D(d\Psi_x)_{\underline{0}\underline{e}^1})\|} = \exp[\pm(\frac{1}{3}\eta^{\beta/4})]$ . How small depends only on  $K_2$ , and therefore only on the surface  $M$ .

Similarly, one can show that  $\frac{\|(\Theta_D(d\Psi_x)_{\underline{u}\underline{v}^u})\|}{\|(\Theta_D(d\Psi_x)_{\underline{0}\underline{e}^2})\|} = \exp[\pm\frac{1}{3}\eta^{\beta/4}]$ , with the result that the fourth factor is  $\exp[\pm\frac{2}{3}\eta^{\beta/4}]$ .

Putting all these estimates together, we see that

$$\frac{\sin \angle(V^u, V^s)}{\sin \angle(E^u, E^s)} = \exp \left[ \pm(2\eta^{2\beta/3} + \eta^{2\beta/3} + \sqrt{\eta} + \frac{2}{3}\eta^{\beta/4}) \right].$$

Since  $0 < \eta < \varepsilon$ , for all  $\varepsilon$  small enough, this is  $e^{\pm\eta^{\beta/4}}$ . How small just depends on  $K_1$ ,  $K_2$ , and  $\beta$ .

Next we estimate  $|\cos \angle(V^s, V^u) - \cos \angle(E^s, E^u)|$ . This is equal to

$$\begin{aligned} &\left| \frac{\langle (d\Psi_x)_{\underline{v}\underline{v}^s}, (d\Psi_x)_{\underline{v}\underline{v}^u} \rangle}{\|(d\Psi_x)_{\underline{v}\underline{v}^s}\| \|(d\Psi_x)_{\underline{v}\underline{v}^u}\|} - \frac{\langle (d\Psi_x)_{\underline{0}\underline{e}^1}, (d\Psi_x)_{\underline{0}\underline{e}^2} \rangle}{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|} \right| \leq \\ &\leq \frac{|\langle (d\Psi_x)_{\underline{v}\underline{v}^s}, (d\Psi_x)_{\underline{v}\underline{v}^u} \rangle|}{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|} \times \left| \frac{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|}{\|(d\Psi_x)_{\underline{v}\underline{v}^s}\| \|(d\Psi_x)_{\underline{v}\underline{v}^u}\|} - 1 \right| + \\ &\quad + \frac{1}{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|} \times |\langle (d\Psi_x)_{\underline{v}\underline{v}^s}, (d\Psi_x)_{\underline{v}\underline{v}^u} \rangle - \langle (d\Psi_x)_{\underline{0}\underline{e}^1}, (d\Psi_x)_{\underline{0}\underline{e}^2} \rangle| \\ &\leq \frac{\|(d\Psi_x)_{\underline{v}\underline{v}^s}\| \|(d\Psi_x)_{\underline{v}\underline{v}^u}\|}{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|} \times \left| \frac{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|}{\|(d\Psi_x)_{\underline{v}\underline{v}^s}\| \|(d\Psi_x)_{\underline{v}\underline{v}^u}\|} - 1 \right| + \\ &\quad + \frac{1}{\|(d\Psi_x)_{\underline{0}\underline{e}^1}\| \|(d\Psi_x)_{\underline{0}\underline{e}^2}\|} \times |\langle (d\Psi_x)_{\underline{v}\underline{v}^s}, (d\Psi_x)_{\underline{v}\underline{v}^u} \rangle - \langle (d\Psi_x)_{\underline{0}\underline{e}^1}, (d\Psi_x)_{\underline{0}\underline{e}^2} \rangle|. \end{aligned}$$

By (A.8) and the estimate of the ‘‘fourth factor’’ above, this is smaller than

$$e^{\frac{2}{3}\eta^{\beta/4}} \cdot \eta^{\beta/4} + \|C_\chi(x)^{-1}\|^2 |\langle (d\Psi_x)_{\underline{v}\underline{v}^s}, (d\Psi_x)_{\underline{v}\underline{v}^u} \rangle - \langle (d\Psi_x)_{\underline{0}\underline{e}^1}, (d\Psi_x)_{\underline{0}\underline{e}^2} \rangle|. \tag{A.10}$$

Since  $\Theta_D$  is an isometry, the difference of the inner products is equal to

$$\begin{aligned}
& \left| \langle \Theta_D(d\Psi_x)_v \underline{v}^s, \Theta_D(d\Psi_x)_v \underline{v}^u \rangle - \langle \Theta_D(d\Psi_x)_0 \underline{e}^1, \Theta_D(d\Psi_x)_0 \underline{e}^2 \rangle \right| \\
& \leq \|\Theta_D(d\Psi_x)_v \underline{v}^s - \Theta_D(d\Psi_x)_0 \underline{e}^1\| \cdot \|(d\Psi_x)_v \underline{v}^u\| \\
& \quad + \|\Theta_D(d\Psi_x)_0 \underline{e}^1\| \cdot \|\Theta_D(d\Psi_x)_v \underline{v}^u - \Theta_D(d\Psi_x)_0 \underline{e}^2\| \\
& \leq 3(\|\Theta_D(d\Psi_x)_v \underline{v}^s - \Theta_D(d\Psi_x)_0 \underline{e}^1\| + \|\Theta_D(d\Psi_x)_v \underline{v}^u - \Theta_D(d\Psi_x)_0 \underline{e}^2\|) \\
& \leq 3(\|\Theta_D(d\Psi_x)_v\| \|\underline{v}^s - \underline{e}^1\| + 2\|\Theta_D(d\Psi_x)_v - \Theta_D(d\Psi_x)_0\| \\
& \quad + \|\Theta_D(d\Psi_x)_v\| \|\underline{v}^u - \underline{e}^2\|) \\
& \leq 3[2\eta^{\beta/3} + 2 \cdot 2K_2\eta + 2\eta^{\beta/3}],
\end{aligned}$$

because  $\Theta_D$  is an isometry,  $\|d\Psi_x\| \leq 2$  on  $R_{Q_\varepsilon(x)}(\mathbb{0})$ , and  $\|\underline{v}^{s/u} - \underline{e}^{1/2}\| < \eta^{\beta/3}$ . Thus  $|\langle (d\Psi_x)_v \underline{v}^s, (d\Psi_x)_v \underline{v}^u \rangle - \langle (d\Psi_x)_0 \underline{e}^1, (d\Psi_x)_0 \underline{e}^2 \rangle| < K_3\eta^{\beta/3}$ , where  $K_3$  only depends on  $M$ . It now follows from (A.10) and the inequality  $\eta < \varepsilon$  that

$$|\cos \angle(V^s, V^u) - \cos \angle(E^s, E^u)| \leq e^{\frac{2}{3}\varepsilon^{3/4}} \eta^{\beta/4} + \|C_\chi(x)^{-1}\|^2 \cdot K_3\eta^{\beta/3}.$$

We now argue as in (A.9) and deduce that

$$|\cos \angle(V^s, V^u) - \cos \angle(E^s, E^u)| \leq (e^{\frac{2}{3}\varepsilon^{3/4}} + K_3\varepsilon^{1/4})\eta^{\beta/4}.$$

This is smaller than  $2\eta^{\beta/4}$ , for all  $\varepsilon$  small enough.  $\square$

**Proof of Proposition 4.12 (Graph Transform)** The proof is a straightforward adaptation of the arguments in [KM] and [BP, chapter 7] (see also [P]).

Let  $V^u = \Psi_x\{(F(t), t) : |t| \leq p^u\}$  be a  $u$ -admissible manifold in  $\Psi_x^{p^u, p^s}$ . We denote the parameters of  $V^u$  by  $\sigma, \gamma, \varphi$ , and  $q$ , and let  $\eta := p^u \wedge p^s$ .  $V^u$  is admissible, so

$$\sigma \leq \frac{1}{2}, \gamma \leq \frac{1}{2}\eta^{\beta/3}, \varphi \leq 10^{-3}\eta, q = p^u, \text{ and } \text{Lip}(F) < \varepsilon, \quad (\text{A.11})$$

see Definition 4.8 and Equation (4.1).

We analyze  $\Gamma_y^u := \Psi_y^{-1}[f(V^u)] \subset \mathbb{R}^2$ , looking for parameterizations of large  $u$ -sub-manifolds. Notice that

$$\Gamma_y^u = f_{xy}[\text{graph}(F)],$$

where  $f_{xy} = \Psi_y^{-1} \circ f \circ \Psi_x$  and  $\text{graph}(F) := \{(F(t), t) : |t| \leq q\}$ .

Since  $V^u$  is admissible,  $\text{graph}(F) \subset R_{Q_\varepsilon(x)}(\mathbb{0})$ . On this domain,  $f_{xy}$  can be expanded as follows (Proposition 3.4):

$$f_{xy}(u, v) = (Au + h_1(u, v), Bv + h_2(u, v)) \quad (\text{A.12})$$

where  $C_f^{-1} < |A| < e^{-\chi}$ ,  $e^\chi < |B| < C_f$ ; and  $h_i$  are  $C^{1+\frac{\beta}{3}}$ -functions s.t.  $|h_i(0)| < \varepsilon\eta$ ,  $\|\nabla h_i(\mathbb{0})\| < \varepsilon\eta^{\beta/3}$ , and  $\|\nabla h_i(\underline{u}) - \nabla h_i(\underline{v})\| \leq \varepsilon\|\underline{u} - \underline{v}\|^{\beta/3}$ . Necessarily,  $\|\nabla h_i\| < \varepsilon\eta^{\beta/3} + \varepsilon[\sqrt{2}Q_\varepsilon(x)]^{\beta/3} < 3\varepsilon Q_\varepsilon(x)^{\beta/3}$  and  $|h_i| < \varepsilon\eta + 3\varepsilon Q_\varepsilon(x)^{\beta/3} \cdot Q_\varepsilon(x)$ . Since  $\eta \leq Q_\varepsilon(x)$ , and  $Q_\varepsilon(x) < \varepsilon^{3/\beta}$ , the following holds for  $\varepsilon$  small enough:

$$\|\nabla h_i\| < 3\varepsilon^2 \text{ and } |h_i| < \varepsilon^2 \text{ on } \text{graph}(F). \quad (\text{A.13})$$

Using (A.12), we can put  $\Gamma_y^u$  in the following form:

$$\Gamma_y^u = \{(AF(t) + h_1(F(t), t), Bt + h_2(F(t), t)) : |t| \leq q\}. \quad (\text{A.14})$$

The idea is to call the second coordinate  $\tau$ , solve  $t = t(\tau)$ , and substitute the result in the first coordinate.

*Claim 1.* The following holds for all  $\varepsilon$  small enough:  $Bt + h_2(F(t), t) = \tau$  has a unique solution  $t = t(\tau)$  for all  $\tau \in [-e^{X-\sqrt{\varepsilon}}q, e^{X-\sqrt{\varepsilon}}q]$ , and

- (a)  $\text{Lip}(t) < e^{-X+\varepsilon}$ ;
- (b)  $|t(0)| < 2\varepsilon\eta$ ;
- (c) the  $C^{\beta/3}$ -norm of  $t'$  is smaller than  $|B|^{-1}e^{3\varepsilon}$ .

*Proof.* Let  $\tau(t) := Bt + h_2(F(t), t)$ . For every  $|t| \leq q$ ,

$$\begin{aligned} |\tau'(t)| &\geq |B| - \max \|\nabla h_2\| \cdot \|(F'(t), 1)\| > |B| - 3\varepsilon^2 \sqrt{1 + \varepsilon^2} \quad (\because \text{(A.13)}, \text{(A.11)}) \\ &> |B|(1 - 3\varepsilon^2 \sqrt{1 + \varepsilon^2}) \quad (\because |B| > e^X > 1) \\ &> e^{-\varepsilon}|B| > 1 \text{ provided } \varepsilon \text{ is small enough.} \end{aligned}$$

It follows that  $\tau$  is  $e^{-\varepsilon}|B|$ -expanding, whence one-to-one.

Since  $\tau$  is one-to-one,  $\tau^{-1}$  is well-defined on  $\tau[-q, q]$ . We estimate this set. Since  $\tau$  is continuous and  $e^{-\varepsilon}B$ -expanding,  $\tau[-q, q] \supseteq (\tau(0) - e^{-\varepsilon}|B|q, \tau(0) + e^{-\varepsilon}|B|q)$ . The center of the interval can be estimated as follows:

$$\begin{aligned} |\tau(0)| &= |h_2(F(0), 0)| \leq |h_2(\underline{0})| + \max \|\nabla h_2\| \cdot |F(0)| \\ &\leq \varepsilon\eta + 3\varepsilon^2 \cdot 10^{-3}\eta < 2\varepsilon\eta \quad (\text{admissibility and (A.13)}). \end{aligned}$$

Recall that  $\eta \equiv p^u \wedge p^s \leq p^u \equiv q$ , therefore  $|\tau(0)| < 2\varepsilon q$ . Since  $|\tau'| > e^{-\varepsilon}|B|$ ,

$$\begin{aligned} \tau[-q, q] &\supseteq [2\varepsilon q - e^{-\varepsilon}|B|q, -2\varepsilon q + e^{-\varepsilon}|B|q] \supseteq [-(|B|e^{-\varepsilon} - 2\varepsilon)q, (|B|e^{-\varepsilon} - 2\varepsilon)q] \\ &\supseteq [-|B|(e^{-\varepsilon} - 2\varepsilon)q, |B|(e^{-\varepsilon} - 2\varepsilon)q]. \end{aligned}$$

Since  $|B|(e^{-\varepsilon} - 2\varepsilon) > e^X(e^{-2\varepsilon} - 2\varepsilon) > e^{X-\sqrt{\varepsilon}}$  for all  $\varepsilon$  small enough,  $\tau^{-1}$  is well defined on  $[-e^{X-\sqrt{\varepsilon}}q, e^{X-\sqrt{\varepsilon}}q]$ .

Since  $t(\cdot)$  is the inverse of a  $|B|e^{-\varepsilon}$ -expanding map,  $\text{Lip}(t) \leq e^\varepsilon|B|^{-1} < e^{-X+\varepsilon}$ , proving (a).

We saw above that  $|\tau(0)| < 2\varepsilon\eta$ . For all  $\varepsilon$  small enough, this is (much) smaller than  $e^{X-\sqrt{\varepsilon}}q$ , therefore  $\tau(0)$  belongs to the domain of  $t$ . It follows that

$$|t(0)| = |t(0) - t(\tau(0))| < \text{Lip}(t)|\tau(0)| < e^{-X+\varepsilon} \cdot 2\varepsilon\eta.$$

For all  $\varepsilon$  small enough, this is less than  $2\varepsilon\eta$ , proving (b).

Next we calculate the  $C^{\beta/3}$ -norm of  $t'(\cdot)$ .

We remind the reader that the  $C^\alpha$ -norm of  $\varphi : [-q, q]^{d_1} \rightarrow \mathbb{R}^{d_2}$  ( $0 < \alpha < 1$ ) is defined by  $\|\varphi\|_\alpha := \|\varphi\|_\infty + \text{Höl}_\alpha(\varphi)$ , where

$$\text{Höl}_\alpha(\varphi) := \sup \left\{ \frac{\|\varphi(\underline{u}) - \varphi(\underline{v})\|}{\|\underline{u} - \underline{v}\|^\alpha} : \underline{u}, \underline{v} \in [-q, q]^{d_1} \text{ different} \right\}.$$

The following inequalities are easy to verify:

- (H1)  $\|\varphi \cdot \psi\|_\alpha \leq \|\varphi\|_\alpha \|\psi\|_\alpha$  for all  $\varphi, \psi \in C^\alpha[-q, q]$ ;
- (H2)  $\|\varphi \circ g\|_\alpha \leq \|\varphi\|_\alpha + \text{Höl}_\alpha(\varphi) \text{Lip}(g)^\alpha$  for all  $\varphi$   $\alpha$ -Hölder and  $g$  Lipschitz;
- (H3) In case  $d_2 = 1$  and  $\|\varphi\|_\alpha < 1$ ,  $\|1/(1 + \varphi)\|_\alpha \leq (1 - \|\varphi\|_\alpha)^{-1}$ .

Differentiating the identity  $s = \tau(t(s)) = Bt(s) + h_2(F(t(s)), t(s))$  w.r.t  $s$ , we obtain after some manipulations

$$t'(s) = B^{-1} \left( 1 + B^{-1} \frac{\partial h_2}{\partial x}(F(t(s)), t(s)) F'(t(s)) + B^{-1} \frac{\partial h_2}{\partial y}(F(t(s)), t(s)) \right)^{-1}.$$

We write this in the form  $t'(s) = B^{-1}(1 + T(s))^{-1}$ , where

$$T(s) := B^{-1} \frac{\partial h_2}{\partial x}(F(t(s)), t(s)) F'(t(s)) + B^{-1} \frac{\partial h_2}{\partial y}(F(t(s)), t(s)).$$

By (H3), it is enough to find  $\|T\|_{\beta/3}$ . Here is the estimation:

$$\begin{aligned} \left\| \frac{\partial h_2}{\partial x}(F(t(s)), t(s)) \right\|_{\beta/3} &\leq \left\| \frac{\partial h_2}{\partial x} \right\|_{\infty} + \text{Höl}_{\beta/3}(\nabla h_2)[\text{Lip}(F \circ t, t)]^{\beta/3} \quad \cdot (\text{H2}) \\ &< 3\varepsilon^2 + \varepsilon \cdot [\text{Lip}(F)^2(\text{Lip}(t))^2 + (\text{Lip}(t))^2]^{\beta/6} \\ &< 3\varepsilon^2 + \varepsilon[\sqrt{\varepsilon^2 + 1}(e^\varepsilon |B|^{-1})]^{\beta/3} \quad \cdot (\text{A.11}), (\text{A.13}) \\ &< \varepsilon, \text{ provided } \varepsilon \text{ is small enough.} \end{aligned}$$

$$\left\| \frac{\partial h_2}{\partial y}(F(t(s)), t(s)) \right\|_{\beta/3} < \varepsilon \text{ (same proof).}$$

$$\begin{aligned} \|F'(t(s))\|_{\beta/3} &\leq \|F'\|_{\infty} + \|F'\|_{\beta/3} \text{Lip}(t)^{\beta/3} \quad (\text{see (H2) above}) \\ &\leq \sigma + \sigma \cdot (e^{-\chi+\varepsilon})^{\beta/3} < 1 \text{ provided } \varepsilon \text{ is small enough.} \end{aligned}$$

Putting these estimates together, we see that  $\|T\|_{\beta/3} < 2\varepsilon$ . It now follows from (H3) that  $\|t'\|_{\beta/3} < |B|^{-1}(1 - 2\varepsilon)^{-1}$ . This is smaller than  $e^{3\varepsilon}|B|^{-1}$  for all  $\varepsilon$  small enough. This proves (c), and completes the proof of the claim.

We now return to (A.14). Substituting  $t = t(\tau)$ , we find that

$$\Gamma_y^u \supset \{(G(\tau), \tau) : |\tau| < e^{\chi - \sqrt{\varepsilon}} q\},$$

where  $G(\tau) := AF(t(\tau)) + h_1(F(t(\tau)), t(\tau))$ . Claim 1 guarantees that  $G(\tau)$  is well-defined and  $C^{1+\beta/3}$  on  $[-e^{\chi - \sqrt{\varepsilon}} q, e^{\chi - \sqrt{\varepsilon}} q]$ . We find the parameters of  $G$ .

*Claim 2.* For all  $\varepsilon$  small enough,  $|G(0)| < e^{-\chi + \sqrt{\varepsilon}}[\varphi + \sqrt{\varepsilon}(q^u \wedge q^s)]$ , and  $|G(0)| < 10^{-3}(q^u \wedge q^s)$ .

*Proof.* Claim 1 says that  $|t(0)| < 2\varepsilon\eta$ . Since  $\text{Lip}(F) < \varepsilon$ ,  $|F(0)| < \varphi$  and  $\varphi \leq 10^{-3}\eta$ ,  $|F(t(0))| < \varphi + 2\varepsilon^2\eta < \eta$  provided  $\varepsilon$  is small enough. Thus

$$\begin{aligned} |G(0)| &\leq |A| \cdot |F(t(0))| + |h_1(F(t(0)), t(0))| \\ &\leq |A|(\varphi + 2\varepsilon^2\eta) + [|h_1(0)| + \max \|\nabla h_1\| \cdot \|(F(t(0)), t(0))\|] \\ &\leq |A|(\varphi + 2\varepsilon^2\eta) + \left[ \varepsilon\eta + 3\varepsilon^2 \cdot \sqrt{\eta^2 + (2\varepsilon\eta)^2} \right] \quad (\because |F(t(0))| < \eta) \\ &\leq |A| \left[ \varphi + \eta(2\varepsilon^2 + \varepsilon + 3\varepsilon^2\sqrt{1 + 4\varepsilon^2}) \right]. \end{aligned}$$

Recalling that  $|A| < e^{-\chi}$  and  $\eta \equiv (p^u \wedge p^s) \leq e^\varepsilon(q^u \wedge q^s)$  (Lemma 4.4), we see that  $|G(0)| < e^{-\chi + \varepsilon}[\varphi + 2\varepsilon(q^u \wedge q^s)]$  for all  $\varepsilon$  small enough.

Since  $\varphi \leq 10^{-3}(p^u \wedge p^s) \leq 10^{-3}e^\varepsilon(q^u \wedge q^s)$ ,  $|G(0)| < e^{-\chi + \varepsilon}[10^{-3} + 2\varepsilon](q^u \wedge q^s)$ . This is less than  $10^{-3}(q^u \wedge q^s)$  for all  $\varepsilon$  sufficiently small. The claim follows.

*Claim 3.* For all  $\varepsilon$  small enough,  $|G'(0)| < e^{-2\chi + \sqrt{\varepsilon}}[\gamma + \varepsilon^{\beta/3}(q^u \wedge q^s)^{\beta/3}]$ , and  $|G'(0)| < \frac{1}{2}(q^u \wedge q^s)^{\beta/3}$ .

*Proof.*  $|G'(0)| \leq |t'(0)|[|A| \cdot |F'(t(0))| + \|\nabla h_1(F(t(0)), t(0))\| \cdot \|(F'(t(0)), 1)\|]$ , and

$$\bullet |t'(0)| \leq \text{Lip}(t) < e^{-\chi + \varepsilon} \text{ (Claim 1).}$$

- $|F'(t(0))| < \gamma + \frac{2}{3}\varepsilon^{\beta/3}\eta^{\beta/3}$ , because  $\text{Höl}_{\beta/3}(F') := \sup \frac{|F'(t_1) - F'(t_2)|}{|t_1 - t_2|^{\beta/3}} \leq \frac{1}{2}$  and therefore by Claim 1(b)

$$|F'(t(0))| < |F'(0)| + \text{Höl}_{\beta/3}(F')|t(0)|^{\beta/3} < \gamma + \sigma \cdot (2\varepsilon\eta)^{\beta/3} < \gamma + \frac{2}{3}\varepsilon^{\beta/3}\eta^{\beta/3}.$$

- $\|\nabla h_1(F(t(0)), t(0))\| \leq 3\varepsilon\eta^{\beta/3}$ , because  $|F(t(0))| < \eta$  (proof of Claim 2), and  $|t(0)| < 2\varepsilon\eta$  (Claim 1), so by the Hölder regularity of  $\nabla h_i$ ,

$$\begin{aligned} \|\nabla h_1(F(t(0)), t(0))\| &\leq \|\nabla h_1(0)\| + \varepsilon \left( \sqrt{|F(t(0))|^2 + |t(0)|^2} \right)^{\beta/3} \\ &\leq \varepsilon\eta^{\beta/3} + \varepsilon(\sqrt{\eta^2 + (2\varepsilon\eta)^2})^{\beta/3} < 3\varepsilon\eta^{\beta/3}. \end{aligned}$$

- $\|(F'(t(0)), 1)\| < \sqrt{1 + \varepsilon^2} < 2$ .

Putting these estimates together, we see that

$$\begin{aligned} |G'(0)| &< e^{-\chi + \varepsilon}|A| \left[ \gamma + \frac{2}{3}\varepsilon^{\beta/3}\eta^{\beta/3} + |A|^{-1} \cdot 3\varepsilon\eta^{\beta/3} \cdot 2 \right] \\ &< e^{-2\chi + \varepsilon} \left[ \gamma + \left( \frac{2}{3}\varepsilon^{\beta/3} + 6C_f\varepsilon \right) \eta^{\beta/3} \right], \quad \because C_f^{-1} < |A| < e^{-\chi} \\ &\leq e^{-2\chi + \varepsilon} \left[ \gamma + \left( \frac{2}{3}\varepsilon^{\beta/3} + 6C_f\varepsilon \right) e^{\varepsilon\beta/3} (q^u \wedge q^s)^{\beta/3} \right] \quad \because p^u \wedge p^s \leq e^\varepsilon (q^u \wedge q^s). \end{aligned}$$

This implies that for all  $\varepsilon$  small enough,  $|G'(0)| < e^{-2\chi + \varepsilon} [\gamma + \varepsilon^{\beta/3} (q^u \wedge q^s)^{\beta/3}]$ , which is stronger than the estimate in the claim.

Since  $\gamma \leq \frac{1}{2}(p^u \wedge p^s)^{\beta/3}$  and  $(p^u \wedge p^s) \leq e^\varepsilon (q^u \wedge q^s)$ , we also get that for all  $\varepsilon$  small enough,  $|G'(0)| < \frac{1}{2}(q^u \wedge q^s)^{\beta/3}$ , as required.

*Claim 4.* For all  $\varepsilon$  small enough,  $\|G'\|_{\beta/3} < e^{-2\chi + \sqrt{\varepsilon}}[\sigma + \sqrt{\varepsilon}]$ , and  $\|G'\|_{\beta/3} < \frac{1}{2}$ .

*Proof.* Differentiating, we see that  $G' = t' \cdot [AF' \circ t + \frac{\partial h_1}{\partial x}(F \circ t, t)F' \circ t + \frac{\partial h_1}{\partial y}(F \circ t, t)]$ . By Claim 1 and its proof

- $\|t'\|_{\beta/3} \leq |B|^{-1}e^{3\varepsilon}$ ,
- $\|F' \circ t\|_{\beta/3} \leq \sigma$ , because  $\|F'\|_{\beta/3} \leq \sigma$  and  $t$  is a contraction,
- $\|\frac{\partial h_1}{\partial x}(F \circ t, t)\|_{\beta/3} < \varepsilon$ , and  $\|\frac{\partial h_1}{\partial y}(F \circ t, t)\|_{\beta/3} < \varepsilon$ .

Thus by (H1),  $\|G'\|_{\beta/3} \leq |B|^{-1}e^{3\varepsilon} [|A|\sigma + \varepsilon\sigma + \varepsilon]$ . Since  $\sigma \leq \frac{1}{2}$ ,  $e^\chi < |B| < C_f$ , and  $C_f^{-1} < |A| < e^{-\chi}$ ,  $\|G'\|_{\beta/3} \leq e^{-2\chi + 3\varepsilon} [\sigma + \frac{3}{2}C_f\varepsilon]$ . If  $\varepsilon$  is small enough, then  $\|G'\|_{\beta/3} < e^{-2\chi + \sqrt{\varepsilon}}[\sigma + \sqrt{\varepsilon}]$ , and  $\|G'\|_{\beta/3} < \frac{1}{2}$ .

*Claim 5.* For all  $\varepsilon$  small enough,  $\widehat{V}^u := \Psi_y\{(G(\tau), \tau) : |\tau| \leq \min\{e^{\chi - \sqrt{\varepsilon}}q, Q_\varepsilon(y)\}\}$  is a  $u$ -manifold in  $\Psi_y$ , the parameters of  $\widehat{V}^u$  satisfy (4.3), and  $\widehat{V}^u$  contains a  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ .

*Proof.* To see that  $\widehat{V}^u$  is a  $u$ -manifold in  $\Psi_y$  we have to check that  $G$  is  $C^{1+\beta/3}$  and  $\|G\|_\infty \leq Q_\varepsilon(y)$ .

Claim 1 shows that  $G$  is  $C^{1+\beta/3}$ . To see that  $\|G\|_\infty \leq Q_\varepsilon(y)$ , we first observe that for all  $\varepsilon$  small enough,  $\text{Lip}(G) < \sqrt{\varepsilon}$ , because

$$|G'| \leq |G'(0)| + \text{Höl}_{\beta/3}(G')Q_\varepsilon(y)^{\beta/3} \leq \varepsilon + \frac{1}{2}\varepsilon < \sqrt{\varepsilon}, \quad \text{provided } \varepsilon \text{ is small enough.}$$

It follows that  $\|G\|_\infty \leq |G(0)| + \sqrt{\varepsilon}Q_\varepsilon(y) < (10^{-3} + \sqrt{\varepsilon})Q_\varepsilon(y) < Q_\varepsilon(y)$ .

Next we claim that  $\widehat{V}^u$  contains a  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ . Since  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ ,  $q^u = \min\{e^\varepsilon p^u, Q_\varepsilon(y)\}$ . Consequently, for every  $\varepsilon$  small enough,

$$e^{X-\sqrt{\varepsilon}}q \equiv e^{X-\sqrt{\varepsilon}}p^u > e^\varepsilon p^u \geq q^u, \quad (\text{A.15})$$

so  $\widehat{V}^u$  restricts to a  $u$ -manifold with  $q$ -parameter equal to  $q^u$ . Claims 2–4 guarantee that this manifold is  $u$ -admissible in  $\Psi_y^{q^u, q^s}$ , and that (4.3) holds.

*Claim 6.*  $f(V^u)$  contains exactly one  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ . This manifold contains  $f(p)$  where  $p = \Psi_x(F(0), 0)$ .

*Proof.* The previous claim shows existence. We prove uniqueness. By formula (A.14), any  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  which is contained in  $f(V^u)$  must be a subset of

$$\Psi_y\{(AF(t) + h_1(F(t), t), Bt + h_2(F(t), t)) : |t| \leq q, |Bt + h_2(F(t), t)| \leq q^u\}.$$

We saw in (A.15) that for all  $\varepsilon$  small enough,  $q^u < e^{X-\sqrt{\varepsilon}}q$ . By claim 1, the equation

$$\tau = Bt + h_2(F(t), t)$$

has a unique solution  $t = t(\tau) \in [-q, q]$  for all  $|\tau| \leq q^u$ . Our manifold must therefore equal  $\Psi_y\{(AF(t(\tau)) + h_1(F(t(\tau)), t(\tau)), \tau) : |\tau| \leq q^u\}$ . This is exactly the  $u$ -admissible manifold that we constructed above.

Let  $\mathcal{F}_u[V^u]$  denote the unique  $u$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  contained in  $f(V^u)$ . We claim that  $\mathcal{F}_u[V^u] \ni f(p)$  where  $p = \Psi_x(F(0), 0)$ . By the previous paragraph, it is enough to check that the second coordinate of  $\Psi_y^{-1}[f(p)]$  has absolute value less than  $q^u$ . Call this second coordinate  $\tau$ , then

$$\begin{aligned} |\tau| &= \text{second coordinate of } f_{xy}(F(0), 0) = |h_2(F(0), 0)| \\ &\leq |h_2(0)| + \max \|\nabla h_2\| \cdot |F(0)| < \varepsilon\eta + 3\varepsilon^2 \cdot 10^{-3}\eta < e^{-\varepsilon}\eta < (q^u \wedge q^s) \leq q^u. \end{aligned}$$

*Claim 7.*  $f(V^u)$  intersects any  $s$ -admissible manifold in  $\Psi_y^{q^u, q^s}$  at a unique point.

*Proof.* Let  $W^s$  be an  $s$ -admissible manifold in  $\Psi_y^{q^u, q^s}$ . We saw in the previous claim that  $f(V^u)$  contains a  $u$ -admissible manifold  $W^u$  in  $\Psi_y^{q^u, q^s}$ . By Proposition 4.11,  $W^u$  and  $W^s$  intersect. Therefore  $f(V^u)$  and  $W^s$  intersect at least at one point.

We claim that the intersection point is unique. Recall that one can put  $f(V^u)$  in the form

$$f(V^u) = \Psi_y\{(AF(t) + h_1(F(t), t), Bt + h_2(F(t), t)) : |t| \leq q\}.$$

We saw in the proof of claim 1 that the second coordinate,  $\tau(t) := Bt + h_2(F(t), t)$ , is a one-to-one continuous map whose image is an interval  $[\alpha, \beta]$  with endpoints  $\alpha < -e^{X-\sqrt{\alpha}}q < -q^u$ ,  $\beta > e^{X-\sqrt{\varepsilon}}q > q^u$ . We also saw that  $|\tau'| > e^{-\varepsilon}|B| \geq e^{X-\varepsilon}$ . Consequently, the inverse function  $t : [\alpha, \beta] \rightarrow [-q, q]$  satisfies  $|t'(\tau)| < 1$ , and so

$$f(V^u) = \Psi_y\{(G(\tau), \tau) : \tau \in [\alpha, \beta]\}, \text{ where } \text{Lip}(G) \leq \varepsilon.$$

Let  $H : [-q^u, q^u] \rightarrow \mathbb{R}$  denote the function which represents  $W^s$  in  $\Psi_y$ , then  $\text{Lip}(H) \leq \varepsilon$ . Extend it to an  $\varepsilon$ -Lipschitz function on  $[\alpha, \beta]$ . The extension represents a Lipschitz manifold  $\widetilde{W}^s \supset W^s$ . The same argument we used to prove Proposition 4.11 shows that  $f(V^u)$  and  $\widetilde{W}^u$  intersect at a unique point. We see that  $f(V^u)$  and  $W^s$  intersect at most at one point.



This completes the proof of the proposition, in the case of  $u$ -manifolds. The case of  $s$ -manifolds follows from the symmetry between  $s$  and  $u$ -manifolds:

- (1)  $V$  is a  $u$ -admissible manifold w.r.t.  $f$  iff  $V$  is a  $s$ -admissible manifold w.r.t.  $f^{-1}$ , and the parameters are the same.
- (2)  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$  w.r.t.  $f$  iff  $\Psi_y^{q^u, q^s} \rightarrow \Psi_x^{p^u, p^s}$  w.r.t.  $f^{-1}$ .  $\square$

**Proof of Proposition 4.14.** We prove the proposition for  $\mathcal{F}_u$ , and leave the case of  $\mathcal{F}_s$  to the reader.

Suppose  $\Psi_x^{p^u, p^s} \rightarrow \Psi_y^{q^u, q^s}$ , and let  $V_i^u$  be two  $u$ -admissible manifolds in  $\Psi_x^{p^u, p^s}$ . We take  $\varepsilon$  to be small enough for the arguments of the previous proof to work.

We saw in the proof of Proposition 4.12 that if  $V_i = \Psi_x\{(F_i(t), t) : |t| \leq p^u\}$ , then  $\mathcal{F}_u[V_i] = \Psi_y\{(G_i(\tau), \tau) : |\tau| \leq q^u\}$ , where

- $G_i(\tau) = AF_i(t_i(\tau)) + h_1(F_i(t_i(\tau)), t_i(\tau))$ ;
- $t_i(\tau)$  is defined implicitly by  $Bt_i(\tau) + h_2(F_i(t_i(\tau)), t_i(\tau)) = \tau$ , and  $|t'_i| < 1$ ;
- $C_f^{-1} < |A| < e^{-\chi}$ ,  $e^\chi < |B| < C_f$ ;
- $|h_i(0)| < \varepsilon(p^u \wedge p^s)$ ,  $\text{Höl}_{\beta/3}(\nabla h_u) \leq \varepsilon$ , and  $\max \|\nabla h_i\| < 3\varepsilon^2$ .

In order to prove the proposition, we need to estimate  $\|G_1 - G_2\|_\infty$  and  $\|G'_1 - G'_2\|_\infty$  in terms of  $\|F_1 - F_2\|_\infty$  and  $\|F'_1 - F'_2\|_\infty$ .

*Part 1.* For all  $\varepsilon$  small enough,  $\|t_1 - t_2\|_\infty \leq \varepsilon\|F_1 - F_2\|_\infty$ .

By definition,  $Bt_i(\tau) + h_2(F_i(t_i(\tau)), t_i(\tau)) = \tau$ . Taking differences, we see that

$$\begin{aligned} |B| \cdot |t_1 - t_2| &\leq |h_2(F_1(t_1), t_1) - h_2(F_2(t_2), t_2)| \\ &\leq \left\| \frac{\partial h_2}{\partial x} \right\|_\infty |F_1(t_1) - F_2(t_2)| + \left\| \frac{\partial h_2}{\partial x} \right\|_\infty |t_1 - t_2| \\ &\leq 3\varepsilon^2(|F_1(t_1) - F_2(t_1)| + |F_2(t_1) - F_2(t_2)| + |t_1 - t_2|) \\ &\leq 3\varepsilon^2(\|F_1 - F_2\|_\infty + (\text{Lip}(F_2) + 1)|t_1 - t_2|) \\ &\leq 3\varepsilon^2\|F_1 - F_2\|_\infty + 3\varepsilon^2(1 + \varepsilon)|t_1 - t_2|, \text{ see (4.1)}. \end{aligned}$$

Rearranging terms, and recalling that  $|B| > e^{\chi - \varepsilon}$ , we see that

$$\|t_1 - t_2\|_\infty < \frac{3\varepsilon^2\|F_1 - F_2\|_\infty}{e^{\chi - \varepsilon} - 3\varepsilon^2(1 + \varepsilon)}.$$

The claim follows.

*Part 2.* For all  $\varepsilon$  small enough,  $\|G_1 - G_2\|_\infty < e^{-\chi/2}\|F_1 - F_2\|_\infty$ , whence (4.4).

Subtracting the defining equations for  $G_i$ , we find that

$$\begin{aligned} |G_1 - G_2| &\leq |A| \cdot |F_1(t_1) - F_2(t_2)| + |h_1(F_1(t_1), t_1) - h_1(F_2(t_2), t_2)| \\ &\leq |A| \cdot |F_1(t_1) - F_2(t_2)| + \|\nabla h_1\| \sqrt{|F_1(t_1) - F_2(t_2)|^2 + |t_1 - t_2|^2} \\ &\leq (|A| + 3\varepsilon^2)|F_1(t_1) - F_2(t_2)| + 3\varepsilon^2|t_1 - t_2| \\ &\leq (|A| + 3\varepsilon^2)(|F_1(t_1) - F_2(t_1)| + |F_2(t_1) - F_2(t_2)|) + 3\varepsilon^2|t_1 - t_2| \\ &\leq (|A| + 3\varepsilon^2)(\|F_1 - F_2\|_\infty + \text{Lip}(F_2)|t_1 - t_2|) + 3\varepsilon^2|t_1 - t_2| \\ &\leq (|A| + 3\varepsilon^2)(1 + \varepsilon \cdot \varepsilon + 3\varepsilon^2 \cdot \varepsilon)\|F_1 - F_2\|_\infty, \text{ see part 1} \\ &\leq |A|(1 + 3C_f\varepsilon^2)(1 + \varepsilon^2 + 3\varepsilon^3)\|F_1 - F_2\|_\infty \\ &\leq e^{-\chi}(1 + 3C_f\varepsilon^2)(1 + \varepsilon^2 + 3\varepsilon^3)\|F_1 - F_2\|_\infty. \end{aligned}$$

It follows that for every  $\varepsilon$  small enough,  $\|G_1 - G_2\|_\infty < e^{-\chi/2}\|F_1 - F_2\|_\infty$ .

*Part 3.* For all  $\varepsilon$  small enough,  $\|t'_1 - t'_2\|_\infty < \sqrt{\varepsilon}(\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3})$ .

Differentiating both sides of the defining equation of  $t_i$  gives

$$t'_i \left[ B + \frac{\partial h_2}{\partial x}(F_i \circ t_i, t_i) F'_i \circ t_i + \frac{\partial h_2}{\partial y}(F_i \circ t_i, t_i) \right] = 1.$$

Taking differences, we obtain after some re-arrangement

$$\begin{aligned} (t'_1 - t'_2) \left[ B + \frac{\partial h_2}{\partial x}(F_1 \circ t_1, t_1) F'_1 \circ t_1 + \frac{\partial h_2}{\partial y}(F_1 \circ t_1, t_1) \right] &= \\ - t'_2 \left[ \frac{\partial h_2}{\partial x}(F_1 \circ t_1, t_1) - \frac{\partial h_2}{\partial x}(F_2 \circ t_2, t_2) \right] F'_1 \circ t_1 &=: \text{I} \\ - t'_2 \frac{\partial h_2}{\partial x}(F_2 \circ t_2, t_2) [(F'_1 \circ t_1 - F'_2 \circ t_1) + (F'_2 \circ t_1 - F'_2 \circ t_2)] &=: \text{II} \\ - t'_2 \left[ \frac{\partial h_2}{\partial y}(F_1 \circ t_1, t_1) - \frac{\partial h_2}{\partial y}(F_2 \circ t_2, t_2) \right] &=: \text{III} \end{aligned}$$

Since  $|B| > e^\chi$ ,  $|F'_1| < 1$  and  $\|\nabla h_2\| < 3\varepsilon^2$ ,

$$\|t'_1 - t'_2\|_\infty \leq \frac{1}{e^\chi - 6\varepsilon^2} \|\text{I} + \text{II} + \text{III}\|_\infty.$$

Since I, II and III involve partial derivatives of  $h_2$  evaluated at  $(F_i \circ t_i, t_i)$ , we begin by analyzing  $\nabla h_2(F_i \circ t_i, t_i)$ . Since  $\text{Höl}_{\beta/3}(\nabla h_i) \leq \varepsilon$ ,

- $\|\nabla h_2(F_1 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_1, t_1)\| \leq \varepsilon \|F_1 - F_2\|_\infty^{\beta/3}$ ;
- $\|\nabla h_2(F_2 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_2, t_1)\| \leq \varepsilon \|t_1 - t_2\|_\infty^{\beta/3}$  (because  $\text{Lip}(F_2) < 1$ );
- $\|\nabla h_2(F_2 \circ t_2, t_1) - \nabla h_2(F_2 \circ t_2, t_2)\| \leq \varepsilon \|t_1 - t_2\|_\infty^{\beta/3}$ .

By part 1,  $\|t_1 - t_2\|_\infty \leq \varepsilon \|F_1 - F_2\|_\infty$ . It follows that

$$\|\nabla h_2(F_1 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_2, t_2)\| < 3\varepsilon \|F_1 - F_2\|_\infty^{\beta/3}.$$

Using the facts that  $|t'_1| < 1$ ,  $|F'_1| < 1$ ,  $\text{Lip}(F_2) < 1$ , and  $\text{Höl}_{\beta/3}(F'_2) < 1$  (see the definition of admissible manifolds and the proof of Proposition 4.12), we get that

$$\begin{aligned} |\text{I}| &\leq 3\varepsilon \|F_1 - F_2\|_\infty^{\beta/3}; \\ |\text{II}| &\leq 3\varepsilon^2 (\|F'_1 - F'_2\|_\infty + \|t_1 - t_2\|_\infty^{\beta/3}) \leq 3\varepsilon^2 \|F'_1 - F'_2\|_\infty + 3\varepsilon^2 \|F_1 - F_2\|_\infty^{\beta/3}; \\ |\text{III}| &\leq 3\varepsilon \|F_1 - F_2\|_\infty^{\beta/3}. \end{aligned}$$

So for all  $\varepsilon$  sufficiently small,  $\|t'_1 - t'_2\|_\infty < \sqrt{\varepsilon}(\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3})$ .

*Part 4.*  $\|G'_1 - G'_2\|_\infty < e^{-\chi/2}(\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3})$ .

By the definition of  $G_i$ ,  $G'_i = t'_i[AF'_i \circ t_i + \frac{\partial h_1}{\partial x}(F_i \circ t_i, t_i)F'_i \circ t_i + \frac{\partial h_1}{\partial y}(F_i \circ t_i, t_i)]$ . Taking differences, we see that

$$\begin{aligned}
|G'_1 - G'_2| &\leq |t'_1 - t'_2| \cdot \left| AF'_1 \circ t_1 + \frac{\partial h_1}{\partial x}(F_1 \circ t_1, t_1)F'_1 \circ t_1 + \frac{\partial h_1}{\partial y}(F_1 \circ t_1, t_1) \right| &=: \text{I}' \\
&\quad + |t'_2| \cdot |A| \cdot (|F'_1 \circ t_1 - F'_2 \circ t_1| + |F'_2 \circ t_1 - F'_2 \circ t_2|) &=: \text{II}' \\
&\quad + |t'_2| \left| \frac{\partial h_1}{\partial x}(F_1 \circ t_1, t_1) - \frac{\partial h_1}{\partial x}(F_2 \circ t_2, t_2) \right| |F'_1 \circ t_1| &=: \text{III}' \\
&\quad + |t'_2| \left| \frac{\partial h_1}{\partial x}(F_2 \circ t_2, t_2) \right| |F'_1 \circ t_1 - F'_2 \circ t_2| &=: \text{IV}' \\
&\quad + |t'_2| \left| \frac{\partial h_1}{\partial y}(F_1 \circ t_1, t_1) - \frac{\partial h_1}{\partial y}(F_2 \circ t_2, t_2) \right| &=: \text{V}'
\end{aligned}$$

Using the same arguments that we used in part 3, one can show that

$$\begin{aligned}
\text{I}' &\leq \|t'_1 - t'_2\|_\infty (e^{-\chi} + 6\varepsilon^2) < \sqrt{\varepsilon} (\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3}) \\
\text{II}' &\leq e^{-\chi} (\|F'_1 - F'_2\|_\infty + \|t_1 - t_2\|_\infty^{\beta/3}) \leq e^{-\chi} (\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3}) \quad (\text{part 1}) \\
\text{III}' &\leq 3\varepsilon \|F_1 - F_2\|_\infty^{\beta/3} \quad (\text{see the estimate of I in part 3}) \\
\text{IV}' &\leq 3\varepsilon^2 \|F'_1 - F'_2\|_\infty + 3\varepsilon^3 \|F_1 - F_2\|_\infty^{\beta/3} \quad (\text{see the estimate of II in part 3}) \\
\text{V}' &\leq 3\varepsilon \|F_1 - F_2\|_\infty^{\beta/3} \quad (\text{see the estimate of III in part 3}).
\end{aligned}$$

It follows that  $\|G'_1 - G'_2\|_\infty < (e^{-\chi} + 10\varepsilon + \sqrt{\varepsilon})(\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3})$ . If  $\varepsilon$  is small enough, then  $\|G'_1 - G'_2\|_\infty < e^{-\chi/2}(\|F'_1 - F'_2\|_\infty + \|F_1 - F_2\|_\infty^{\beta/3})$ .  $\square$

**Proof of Proposition 6.3** The following proof is based on [BP, Chapter 7].

Suppose  $V^s$  is an  $s$ -admissible manifold in  $\Psi_x^{p^u, p^s}$  which stays in windows, then there is a positive chain  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \geq 0}$  s.t.  $\Psi_{x_0}^{p_0^u, p_0^s} = \Psi_x^{p^u, p^s}$ , and there are  $s$ -admissible manifolds  $W_i^s$  in  $\Psi_{x_i}^{p_i^u, p_i^s}$  s.t.  $f^i(V^s) \subset W_i^s$  for all  $i \geq 0$ . We write

- $V^s = \Psi_x \{(t, F_0(t)) : |t| \leq p^s\}$ ,
- $W_i^s = \Psi_{x_i} \{(t, F_i(t)) : |t| \leq p_i^s\}$ ,
- $\eta_i := p_i^u \wedge p_i^s$ .

Admissibility means that  $\|F'_i\|_{\beta/3} \leq \frac{1}{2}$ ,  $|F'_i(0)| \leq \frac{1}{2}\eta_i^{\beta/3}$  and  $|F_i(0)| \leq 10^{-3}\eta_i$ . By Lemma 4.4,  $e^{-\varepsilon} \leq \eta_i/\eta_{i+1} \leq e^\varepsilon$ . By (4.1),  $\text{Lip}(F_i) < \varepsilon$ .

*Part 1.* If  $\varepsilon$  is so small that  $e^{-\chi} + 4\varepsilon^2 < e^{-\chi/2}$ , then for every  $y, z \in V^s$ ,  $d(f^k(y), f^k(z)) \leq 6p_0^s e^{-\frac{1}{2}k\chi}$  for all  $k \geq 0$ .

*Proof.* Since  $V^s$  stays in windows,  $f^k(V^s) \subset \Psi_{x_k} [R_{Q_\varepsilon(x_k)}(\mathbb{0})]$  for all  $k \geq 0$ . Therefore, for any  $y, z \in V^s$ , one can write  $f^k(y) = \Psi_{x_k}(\underline{y}_k)$  and  $f^k(z) = \Psi_{x_k}(\underline{z}_k)$ , where  $\underline{y}_k = (y_k, F_k(y_k))$ ,  $\underline{z}_k = (z_k, F_k(z_k))$  belong to  $R_{Q_\varepsilon(x_k)}(\mathbb{0})$ .

For every  $k$ ,  $\underline{y}_{k+1} = f_{x_k x_{k+1}}(\underline{y}_k)$  and  $\underline{z}_{k+1} = f_{x_k x_{k+1}}(\underline{z}_k)$ , where  $f_{x_k x_{k+1}} := \Psi_{x_{k+1}}^{-1} \circ f \circ \Psi_{x_k}$ . By (3.3),

$$f_{x_k x_{k+1}}(v, w) = (A_k v + h_1(v, w), B_k w + h_2(v, w)) \text{ on } R_{Q_\varepsilon(x_k)}(\mathbb{0}),$$

where  $C_f^{-1} < |A_k| < e^{-\chi}$ ,  $e^\chi < |B_k| < C_f$ , and  $\max \|\nabla h_i\| < 3\varepsilon^2$ . Thus

$$\begin{aligned} |y_{k+1} - z_{k+1}| &\leq |A_k| \cdot |y_k - z_k| + 3\varepsilon^2(|y_k - z_k| + \text{Lip}(F_k)|y_k - z_k|) \\ &\leq (e^{-\chi} + 4\varepsilon^2)|y_k - z_k| < e^{-\frac{1}{2}\chi}|y_k - z_k| \leq \dots \leq e^{-\frac{1}{2}(k+1)\chi}|y_0 - z_0|. \end{aligned}$$

Since  $\underline{y}_0, \underline{z}_0$  are on the graph of an  $s$ -admissible manifold in  $\Psi_{x_0}^{p_0^s, p_0^s}$ , their  $x$ -coordinates are in  $[-p_0^s, p_0^s]$ , so  $|y_0 - z_0| \leq 2p_0^s$ . Thus  $|y_k - z_k| \leq 2e^{-\frac{1}{2}k\chi}p_0^s$ . Since  $\underline{y}_k = (y_k, F_k(y_k))$ ,  $\underline{z}_k = (z_k, F_k(z_k))$ , and  $\text{Lip}(F_k) < \varepsilon$ ,  $\|\underline{y}_k - \underline{z}_k\| < 3p_0^s e^{-\frac{1}{2}k\chi}$ .

Pesin charts have Lipschitz constant less than two, so  $d(f^k(y), f^k(z)) < 6p_0^s e^{-\frac{1}{2}k\chi}$ .

*Part 2.* Suppose  $\varepsilon$  is so small that  $e^{-\chi} + 3\varepsilon^2 + 3\varepsilon^3 < e^{-\frac{2}{3}\chi}$  and  $C_f\varepsilon + 3\varepsilon^2 < 1$ . For every  $y \in V^s$ , let  $\underline{e}^s(y)$  denote the positively oriented unit tangent vector to  $V^s$  at  $y$ . If  $y \in V^s$ , then  $\|df_y^k \underline{e}^s(y)\| \leq 6e^{-\frac{2}{3}k\chi} \|C_\chi(x_0)^{-1}\|$  for all  $k \geq 0$ .

*Proof.* If  $y \in V^s$ , then  $f^k(y) \in W_k^s \subset \Psi_{x_k}[R_{Q_\varepsilon(x_k)}(\underline{0})]$ . So  $df_y^k \underline{e}^s(y) = (d\Psi_{x_k})_{\underline{y}_k} \binom{a_k}{b_k}$  where  $\binom{a_k}{b_k}$  is tangent to the graph of  $F_k$ . Since  $\text{Lip}(F_k) < \varepsilon$ ,  $|b_k| \leq \varepsilon|a_k|$  for all  $k$ . The identity  $\binom{a_{k+1}}{b_{k+1}} = (df_{x_k x_{k+1}})_{\underline{y}_k} \binom{a_k}{b_k}$  holds. Since  $\|\nabla h_i\| \leq 3\varepsilon^2$ ,

$$\binom{a_{k+1}}{b_{k+1}} = \begin{pmatrix} A_k + \frac{\partial h_1}{\partial x}(\underline{y}_k) & \frac{\partial h_1}{\partial y}(\underline{y}_k) \\ \frac{\partial h_2}{\partial x}(\underline{y}_k) & B_k + \frac{\partial h_2}{\partial y}(\underline{y}_k) \end{pmatrix} \binom{a_k}{b_k} = \begin{pmatrix} (A_k \pm 3\varepsilon^2)a_k \pm 3\varepsilon^2|b_k| \\ (B_k \pm 3\varepsilon^2)b_k \pm 3\varepsilon^2|a_k| \end{pmatrix}.$$

It follows that  $|a_{k+1}| \leq (|A_k| + 3\varepsilon^2 + 3\varepsilon^3)|a_k|$ . By the bounds on  $A_k$  and  $B_k$  and the assumption on  $\varepsilon$ ,

$$|a_k| \leq e^{-\frac{2}{3}k\chi}|a_0| \text{ and } |b_k| \leq \varepsilon|a_k| \leq e^{-\frac{2}{3}k\chi}|a_0|.$$

Returning to the defining relation  $df_y^k \underline{e}^s(y) = (d\Psi_{x_k})_{\underline{y}_k} \binom{a_k}{b_k}$ , and recalling that  $\|d\Psi_{x_k}\| \leq 2$  (Theorem 2.7), we see that  $\|df_y^k \underline{e}^s(y)\| \leq 2\sqrt{2}e^{-\frac{2}{3}k\chi}|a_0|$ .

Since  $\binom{a_0}{b_0} = (d\Psi_{x_0})_{\underline{y}_0}^{-1} \underline{e}^s(y)$ ,  $|a_0| \leq \|d\Psi_{x_0}^{-1}\|$ , so  $\|df_y^k \underline{e}^s(y)\| \leq 2\sqrt{2}e^{-\frac{2}{3}k\chi} \|d\Psi_{x_0}^{-1}\|$ .

For every  $x$ ,  $\|d\Psi_x^{-1}\| \leq 2\|C_\chi(x)^{-1}\|$  because  $C_\chi(x)^{-1}$  maps  $B_{2Q_\varepsilon(x)}^x(\underline{0})$  into  $B_{2\varepsilon^{3/\beta}}(\underline{0}) \subset B_{2\varepsilon}(\underline{0}) \subset B_{\rho(M)}(\underline{0})$ , provided  $\varepsilon < \frac{1}{2}\rho(M)$ , and by the definition  $\rho(M)$  is so small that  $\|(d \exp_x^{-1})_y\| \leq 2$  for all  $x \in M$  and  $y \in B_{\rho(M)}(\underline{0})$ .

It follows that  $\|df_y^k \underline{e}^s(y)\| \leq 6\|C_\chi(x_0)^{-1}\| e^{-\frac{2}{3}k\chi}$ .

*Part 3.* The following holds for all  $\varepsilon$  small enough: for all  $y, z \in V^s$  and  $n \geq 0$ ,  $|\log \|df_y^n \underline{e}^s(y)\| - \log \|df_z^n \underline{e}^s(z)\|| \leq Q_\varepsilon(x_0)^{\beta/4}$ .

*Proof.* Call the quantity to be estimated  $A$ . For every  $p \in V^s$ ,

$$\begin{aligned} df_p^n[\underline{e}^s(p)] &= df_{f(p)}^{n-1}[df_p \underline{e}^s(p)] = \pm \|df_p \underline{e}^s(p)\| \cdot df_{f(p)}^{n-1}[\underline{e}^s(f(p))] \\ &= \dots = \pm \prod_{k=0}^{n-1} \|df_{f^k(p)} \underline{e}^s(f^k(p))\| \cdot \underline{e}^s(f^n(p)). \end{aligned}$$

Thus  $A := \left| \log \frac{\|df_y^n \underline{e}^s(y)\|}{\|df_z^n \underline{e}^s(z)\|} \right| \leq \sum_{k=0}^{n-1} \left| \log \|df_{f^k(y)} \underline{e}^s(f^k(y))\| - \log \|df_{f^k(z)} \underline{e}^s(f^k(z))\| \right|$ .

We shall estimate the sum term-by-term, using the Hölder continuity of  $df$ .

In section 3.1 we covered  $M$  by a finite collection  $\mathcal{D}$  of open sets  $D$ , equipped with a smooth map  $\Theta_D : TD \rightarrow \mathbb{R}^2$  s.t.  $\Theta_D|_{T_x M} : T_x M \rightarrow \mathbb{R}^2$  is an isometry, and  $\vartheta_x := \Theta_D^{-1}|_{\mathbb{R}^2} : \mathbb{R}^2 \rightarrow TD$  has the property that  $(x, \underline{v}) \mapsto \vartheta_x(\underline{v})$  is Lipschitz on  $D \times B_1(\underline{0})$ . Since  $f$  is a  $C^{1+\beta}$ -diffeomorphism and  $M$  is compact,  $df_p[\underline{v}]$  depends in a  $\beta$ -Hölder way on  $p$ , and in a Lipschitz way on  $\underline{v}$ . It follows that there exists a

constant  $H_0 > 1$  s.t. for every  $D \in \mathcal{D}$ , for every  $y, z \in D$ , and for every  $\underline{u}, \underline{v} \in \mathbb{R}^2$  of length one,  $\left| \log \|df_y(\vartheta_y(\underline{u}))\| - \log \|df_z(\vartheta_z(\underline{v}))\| \right| < H_0(d(y, z)^\beta + \|\underline{u} - \underline{v}\|)$ .

Choose  $D_k \in \mathcal{D}$  s.t.  $D_k \ni f^k(y), f^k(z)$ . Such sets exist provided  $\varepsilon$  is much smaller than the Lebesgue number of  $\mathcal{D}$ , because by part 1  $d(f^k(y), f^k(z)) < 6\varepsilon$ . Writing  $\text{Id} = \Theta_{D_k} \circ \vartheta_{f^k(y)}$  and  $\text{Id} = \Theta_{D_k} \circ \vartheta_{f^k(z)}$ , we see that

$$\begin{aligned} A &\leq \sum_{k=0}^{n-1} \left| \log \|df_{f^k(y)} \vartheta_{f^k(y)} \Theta_{D_k} \underline{e}^s(f^k(y))\| - \log \|df_{f^k(z)} \vartheta_{f^k(z)} \Theta_{D_k} \underline{e}^s(f^k(z))\| \right| \\ &\leq \sum_{k=0}^{n-1} H_0 (d(f^k(y), f^k(z))^\beta + \|\Theta_{D_k} \underline{e}^s(f^k(y)) - \Theta_{D_k} \underline{e}^s(f^k(z))\|) \\ &\leq \frac{H_0(6p_0^s)^\beta}{1 - e^{-\frac{1}{2}\beta\chi}} + H_0 \sum_{k=0}^{n-1} \|\Theta_{D_k} \underline{e}^s(f^k(y)) - \Theta_{D_k} \underline{e}^s(f^k(z))\|, \text{ by part 1.} \quad (\text{A.16}) \end{aligned}$$

We estimate  $N_k := \|\Theta_{D_k} \underline{e}^s(f^k(y)) - \Theta_{D_k} \underline{e}^s(f^k(z))\|$ . By definition,  $\underline{e}^s(f^k(y))$  and  $\underline{e}^s(f^k(z))$  are the positively oriented unit tangent vectors to  $f^k(V^s) \subset W_k^s$ , at  $f^k(y)$  and  $f^k(z)$ . Defining  $\underline{y}_k$  and  $\underline{z}_k$  as before, we obtain

$$\underline{e}^s(f^k(y)) = \frac{(d\Psi_{x_k})_{\underline{y}_k} \begin{pmatrix} 1 \\ F'_k(y_k) \end{pmatrix}}{\|(d\Psi_{x_k})_{\underline{y}_k} \begin{pmatrix} 1 \\ F'_k(y_k) \end{pmatrix}\|}, \quad \underline{e}^s(f^k(z)) = \frac{(d\Psi_{x_k})_{\underline{z}_k} \begin{pmatrix} 1 \\ F'_k(z_k) \end{pmatrix}}{\|(d\Psi_{x_k})_{\underline{z}_k} \begin{pmatrix} 1 \\ F'_k(z_k) \end{pmatrix}\|}.$$

We saw in part 1 that  $\|(d\Psi_{x_k})_{\underline{y}_k}^{-1}\|$  and  $\|(d\Psi_{x_k})_{\underline{z}_k}^{-1}\|$  are bounded by  $2\|C_\chi(x_k)^{-1}\|$ , so the denominators are bounded below by  $\frac{1}{2}\|C_\chi(x_k)^{-1}\|^{-1}$ . Since for any two non-zero vectors  $\underline{v}, \underline{u}$ ,  $\|\underline{v}/\|\underline{v}\| - \underline{u}/\|\underline{u}\|\| < 2\|\underline{v} - \underline{u}\|/\|\underline{v}\|$ ,

$$N_k \leq 2\|C_\chi(x_k)^{-1}\| \cdot \left\| \Theta_{D_k} (d\Psi_{x_k})_{\underline{y}_k} \begin{pmatrix} 1 \\ F'_k(y_k) \end{pmatrix} - \Theta_{D_k} (d\Psi_{x_k})_{\underline{z}_k} \begin{pmatrix} 1 \\ F'_k(z_k) \end{pmatrix} \right\|.$$

On  $D_k$  we can write  $\Psi_{x_k} = \exp_{x_k} \circ \vartheta_{x_k} \circ C_{x_k}$ , where  $\vartheta_{x_k} \circ C_{x_k} = C_\chi(x_k)$ . Let

$$\underline{u}_k := C_\chi(x_k) \underline{y}_k, \underline{u}'_k := C_\chi(x_k) \underline{z}_k, \text{ and } \underline{v}_k := C_{x_k} \begin{pmatrix} 1 \\ F'_k(y_k) \end{pmatrix}, \underline{v}'_k := C_{x_k} \begin{pmatrix} 1 \\ F'_k(z_k) \end{pmatrix},$$

then  $N_k \leq 2\|C_\chi(x_k)^{-1}\| \cdot \left\| \Theta_{D_k} (d \exp_{x_k})_{\underline{u}_k} [\vartheta_{x_k}(\underline{v}_k)] - \Theta_{D_k} (d \exp_{x_k})_{\underline{u}'_k} [\vartheta_{x_k}(\underline{v}'_k)] \right\|$ . Since  $\Theta_{D_k}, \vartheta_{x_k}$  are isometries,  $C_{x_k}$  are contractions,  $\|(d \exp_{x_k})_{\underline{u}_k}\| \leq 2$ , and  $|F'_k(y_k) - F'_k(z_k)| \leq \frac{1}{2}|y_k - z_k|^{\beta/3}$ ,

$$\begin{aligned} N_k &\leq 2\|C_\chi(x_k)^{-1}\| \cdot \left\| \Theta_{D_k} (d \exp_{x_k})_{\underline{u}_k} [\vartheta_{x_k}(\underline{v}_k)] - \Theta_{D_k} (d \exp_{x_k})_{\underline{u}_k} [\vartheta_{x_k}(\underline{v}'_k)] \right\| + \\ &\quad + 2\|C_\chi(x_k)^{-1}\| \cdot \left\| \Theta_{D_k} (d \exp_{x_k})_{\underline{u}_k} [\vartheta_{x_k}(\underline{v}'_k)] - \Theta_{D_k} (d \exp_{x_k})_{\underline{u}'_k} [\vartheta_{x_k}(\underline{v}'_k)] \right\| \\ &\leq 2\|C_\chi(x_k)^{-1}\| \cdot |y_k - z_k|^{\beta/3} + \\ &\quad + 2\|C_\chi(x_k)^{-1}\| \cdot \left\| \Theta_{D_k} (d \exp_{x_k})_{\underline{u}_k} [\vartheta_{x_k}(\underline{v}'_k)] - \Theta_{D_k} (d \exp_{x_k})_{\underline{u}'_k} [\vartheta_{x_k}(\underline{v}'_k)] \right\|. \end{aligned}$$

We study this expression. In what follows we identify the differential of a linear map with the map itself.

By construction, the map  $(x, \underline{u}, \underline{v}) \mapsto [\Theta_D \circ (d \exp_x)_{\underline{u}}] [\vartheta_x(\underline{v})]$  is smooth on  $D \times B_2(\underline{0}) \times B_2(\underline{0})$  for every  $D \in \mathcal{D}$ . Therefore there exists a constant  $E_0 > 1$  s.t. for

every  $(x, \underline{u}_i, \underline{v}_i) \in D \times B_2(\underline{0}) \times B_2(\underline{0})$  and every  $D \in \mathcal{D}$ ,

$$\|\Theta_D(d \exp_x)_{\underline{u}_1}[\vartheta_x(\underline{v}_1)] - \Theta_D(d \exp_x)_{\underline{u}_2}[\vartheta_x(\underline{v}_2)]\| \leq E_0(\|\underline{u}_1 - \underline{u}_2\| + \|\underline{v}_1 - \underline{v}_2\|).$$

It follows that

$$\begin{aligned} N_k &\leq 2\|C_\chi(x_k)^{-1}\| \cdot \left( |y_k - z_k|^{\beta/3} + E_0(\|\underline{u}_k - \underline{u}'_k\| + \|\underline{v}_k - \underline{v}'_k\|) \right) \\ &\leq 2\|C_\chi(x_k)^{-1}\| \cdot \left( |y_k - z_k|^{\beta/3} + E_0(\|\underline{y}_k - \underline{z}_k\| + |y_k - z_k|^{\beta/3}) \right) \\ &\leq 6E_0\|C_\chi(x_k)^{-1}\| \|\underline{y}_k - \underline{z}_k\|^{\beta/3} \quad (\because E_0 > 1) \\ &\leq 6E_0\|C_\chi(x_k)^{-1}\| (3p_0^s)^{\beta/3} e^{-\frac{1}{6}\beta\chi k} \quad \text{because } \|\underline{y}_k - \underline{z}_k\| < 3p_0^s e^{-\frac{1}{2}k\chi} \text{ (part 1)} \\ &\leq 9E_0\|C_\chi(x_k)^{-1}\| (p_0^s)^{\beta/3} e^{-\frac{1}{6}\beta\chi k}. \end{aligned}$$

By the definition of  $Q_\varepsilon(\cdot)$ ,  $\|C_\chi(x_k)^{-1}\| \leq \varepsilon^{1/4} Q_\varepsilon(x_k)^{-\beta/12} \leq \varepsilon^{1/4} (p_k^s)^{-\beta/12}$ , and therefore  $N_k \leq 9\varepsilon^{1/4} E_0 (p_k^s)^{-\beta/12} (p_0^s)^{\beta/3} e^{-\frac{1}{6}\beta\chi k}$ . Since  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$  is a chain,  $p_i^s = \min\{e^\varepsilon p_{i+1}^s, Q_\varepsilon(x_i)\} \leq e^\varepsilon p_{i+1}^s$  for all  $i$ , whence  $p_0^s \leq e^{k\varepsilon} p_k^s$ . It follows that for all  $\varepsilon$  small enough,

$$N_k \leq 9\varepsilon^{1/4} E_0 (p_0^s)^{\beta/4} \exp[-\frac{1}{7}\beta\chi k]. \quad (\text{A.17})$$

Plugging this in (A.16), we obtain

$$\begin{aligned} \left| \log \frac{\|df_y^n \underline{e}^s(y)\|}{\|df_z^n \underline{e}^s(z)\|} \right| &\leq \left( \frac{6^\beta H_0 (p_0^s)^{3\beta/4}}{1 - e^{-\frac{1}{2}\beta\chi}} + \frac{9\varepsilon^{1/4} E_0 H_0}{1 - e^{-\frac{1}{7}\beta\chi}} \right) (p_0^s)^{\beta/4} \\ &< \left( \frac{9\varepsilon^{3\beta/4} E_0 H_0}{1 - e^{-\frac{1}{7}\beta\chi}} \right) Q_\varepsilon(x_0)^{\beta/4}. \end{aligned}$$

The term in the brackets is less than one for every  $\varepsilon$  small enough. How small depends only on  $M$  (through  $E_0$ ),  $f$  (through  $H_0$  and  $\beta$ ), and  $\chi$ .  $\square$

**Proof of Proposition 6.4** We continue to use the notation of the previous proof.

Assume that  $V^s \cap U^s \neq \emptyset$ . We show that  $V^s \subseteq U^s$  or  $U^s \subseteq V^s$ .

Since  $V^s$  stays in windows, there is a positive chain  $(\Psi_{x_i}^{p_i^u, p_i^s})_{i \geq 0}$  such that  $\Psi_{x_i}^{p_i^u, p_i^s} = \Psi_x^{p_i^u, p_i^s}$  and such that for all  $i \geq 0$ ,  $f^i(V^s) \subset W_i^s$  where  $W_i^s$  is an  $s$ -admissible manifold in  $\Psi_{x_i}^{p_i^u, p_i^s}$ .

*Claim 1.* The following holds for all  $\varepsilon$  small enough.  $f^n(V^s) \subseteq \Psi_{x_n}[R_{\frac{1}{2}Q_\varepsilon(x_n)}(\underline{0})]$  for all  $n$  large enough.

*Proof.* Suppose  $y \in V^s$ , and write as in part 1 of the previous proof,  $f^n(y) = \Psi_{x_n}(\underline{y}_n)$  where  $\underline{y}_n = (y_n, F_n(y_n))$  and  $F_n$  is the function which represents  $W_n^s$  in  $\Psi_{x_n}$ . We have  $\underline{y}_{n+1} = f_{x_n x_{n+1}}(\underline{y}_n)$ , which implies in the notation of the previous proof that if  $\varepsilon$  is small enough, then

$$\begin{aligned} |y_{n+1}| &\leq |A_n| \cdot |y_n| + |h_1(\underline{y}_n)| \leq |A_n| \cdot |y_n| + |h_1(\underline{0})| + \|\nabla h_1\|(|y_n| + |F_n(y_n)|) \\ &< e^{-\chi}|y_n| + \varepsilon\eta_n + 3\varepsilon^2(|y_n| + p_n^s) < (e^{-\chi} + 3\varepsilon^2)|y_n| + 2\varepsilon p_n^s \\ &< (e^{-\chi} + 3\varepsilon^2)|y_n| + 2\varepsilon \min\{e^\varepsilon p_{n+1}^s, Q_\varepsilon(x_n)\} \\ &< (e^{-\chi} + 3\varepsilon^2)|y_n| + 2e^\varepsilon \varepsilon p_{n+1}^s < e^{-\chi/2}|y_n| + 4\varepsilon p_{n+1}^s. \end{aligned}$$

We see that  $|y_n| \leq a_n$  where  $a_n$  is defined by induction by

$$a_0 := Q_\varepsilon(x_0) \text{ and } a_{n+1} = e^{-\chi/2} a_n + 4\varepsilon p_{n+1}^s.$$

We claim that if  $\varepsilon$  is small enough, then  $a_n < \frac{1}{4}p_n^s$  for some  $n$ . Otherwise,  $p_n^s \leq 4a_n$  for all  $n$ , whence  $a_{n+1} \leq (e^{-\chi/2} + 16\varepsilon)a_n$  for all  $n$ , which implies that  $a_n < (e^{-\frac{1}{2}\chi} + 16\varepsilon)^n a_0$ . But by assumption,  $a_n \geq \frac{1}{4}p_n^s \geq \frac{1}{4}(p_n^u \wedge p_n^s) \geq \frac{1}{4}e^{-\varepsilon n}(p_0^u \wedge p_0^s)$  (Lemma 4.4), so necessarily  $e^{-\varepsilon} \leq e^{-\chi/2} + 16\varepsilon$ . If  $\varepsilon$  is small enough, this is false and we obtain a contradiction. It follows that  $\exists n$  s.t.  $a_n < \frac{1}{4}p_n^s$ .

It is clear from the definition of  $a_n$ , that if  $\varepsilon$  is small enough then  $a_n < \frac{1}{4}p_n^s \implies a_{n+1} < \frac{1}{4}p_{n+1}^s$ . Thus  $a_n < \frac{1}{4}p_n^s$  for all  $n$  large enough.

In particular,  $|y_n| < \frac{1}{4}Q_\varepsilon(x_n)$  for all  $n$  large enough. Since  $\underline{y}_n = (y_n, F_n(y_n))$  and  $|F_n(y_n)| \leq |F_n(0)| + \text{Lip}(F_n)|y_n| < (10^{-3} + \varepsilon)Q_\varepsilon(x_n)$ ,  $\|\underline{y}_n\| < \frac{1}{2}Q_\varepsilon(x_n)$  for all  $n$  large enough.

*Claim 2.* The following holds for all  $\varepsilon$  small enough:  $f^n(U^s) \subseteq \Psi_{x_n}[R_{Q_\varepsilon(x_n)}(\underline{0})]$  for all  $n$  large enough.

*Proof.*  $U^s$  stays in windows, so there exists a positive chain  $\{\Psi_{y_i}^{q_i^u, q_i^s}\}_{i \geq 0}$  such that  $\Psi_{y_0}^{q_0^u, q_0^s} = \Psi_y^{q^u, q^s}$  and such that for all  $i \geq 0$ ,  $f^i(U^s)$  is a subset of an  $s$ -admissible manifold in  $\Psi_{y_i}^{q_i^u, q_i^s}$ .

Let  $z$  be a point in  $U^s \cap V^s$ . By Part 1 of Theorem 6.3, for any  $w \in U^s$   $d(f^n(z), f^n(w)) \leq 6q_0^s e^{-\frac{1}{2}n\chi}$ . Therefore  $f^n(z), f^n(w) \in B_{Q_\varepsilon(x_n) + 6q_0^s}(x_n) \subset B_{7\varepsilon}(x_n)$ . If  $\varepsilon < \frac{1}{7}\rho(M)$  (cf. §2.3), then  $\|\exp_{x_n}^{-1}[f^n(z)] - \exp_{x_n}^{-1}[f^n(w)]\| < 12e^{-\frac{1}{2}n\chi}q_0^s$ , so

$$\|\Psi_{x_n}^{-1}[f^n(z)] - \Psi_{x_n}^{-1}[f^n(w)]\| < \|C_\chi(x_n)^{-1}\| \cdot 12e^{-\frac{1}{2}n\chi}q_0^s.$$

Since  $p_n^s \leq Q_\varepsilon(x_n) \ll \|C_\chi(x_n)^{-1}\|^{-1}$ ,  $\|\Psi_{x_n}^{-1}[f^n(z)] - \Psi_{x_n}^{-1}[f^n(w)]\| \leq 12(p_n^s)^{-1}q_0^s e^{-\frac{1}{2}n\chi}$ .

Since  $\{\Psi_{x_i}^{p_i^u, p_i^s}\}_{i \in \mathbb{Z}}$  is a chain,  $p_i^s = \max\{e^\varepsilon p_{i+1}^s, Q_\varepsilon(x_i)\} \leq e^\varepsilon p_{i+1}^s$  for all  $i$ . It follows that  $p_0^s \leq e^{n\varepsilon} p_n^s$ , whence

$$\|\Psi_{x_n}^{-1}[f^n(z)] - \Psi_{x_n}^{-1}[f^n(w)]\| < 12 \left( \frac{q_0^s}{p_0^s} \right) e^{-\frac{1}{2}n\chi + n\varepsilon} \xrightarrow{n \rightarrow \infty} 0 \text{ exponentially fast.}$$

Since  $Q_\varepsilon(x_n) \geq (p_n^u \wedge p_n^s) \geq e^{-\varepsilon n}(p_0^u \wedge p_0^s)$ , for all  $n$  large enough

$$\|\Psi_{x_n}^{-1}[f^n(z)] - \Psi_{x_n}^{-1}[f^n(w)]\| < \frac{1}{2}Q_\varepsilon(x_n).$$

How large depends only on  $(p_0^s, p_0^u)$  and  $q_0^s$ .

Since, by claim 1,  $\|\Psi_{x_n}^{-1}(f^n(z))\| < \frac{1}{2}Q_\varepsilon(x_n)$  for all  $n$  large enough, we have that  $\|\Psi_{x_n}^{-1}(f^n(w))\| < Q_\varepsilon(x_n)$  for all  $n$  large enough. All the estimates are uniform in  $w \in U^s$ , so the claim is proved.

*Claim 3.* Recall that  $V^s$  is  $s$ -admissible in  $\Psi_x^{p_x^u, p_x^s}$  and  $U^s$  is  $s$ -admissible in  $\Psi_y^{q_y^u, q_y^s}$ . If  $p^s \leq q^s$  then  $V^s \subseteq U^s$ , and if  $q^s \leq p^s$  then  $U^s \subseteq V^s$ .

*Proof.* W.l.o.g.  $p^s \leq q^s$ . Pick  $n_0$  s.t.  $f^n(U^s), f^n(V^s) \subset \Psi_{x_n}[R_{Q_\varepsilon(x_n)}(\underline{0})]$  for all  $n \geq n_0$ , then  $f^{n_0}(V^s), f^{n_0}(U^s) \subset W^s := V^s[(\Psi_{x_i}^{p_i^u, p_i^s})_{i \geq n_0}]$  (Proposition 4.15 (4)).

Let  $G$  denote the function which represents  $W^s$  in  $\Psi_{x_{n_0}}$ , then  $\Psi_{x_n}^{-1}[f^n(U^s)]$  and  $\Psi_{x_n}^{-1}[f^n(V^s)]$  are two connected subsets of  $\text{graph}(G)$ . Write

$$\begin{aligned} f^n(V^s) &= \Psi_{x_n} \{(t, G(t)) : t \in [\alpha, \beta]\}, \\ f^n(U^s) &= \Psi_{x_n} \{(t, G(t)) : t \in [\alpha', \beta']\}. \end{aligned}$$

The manifold  $f^n(V^s)$  has endpoints  $A := \Psi_{x_n}(\alpha, G(\alpha))$ ,  $B := \Psi_{x_n}(\beta, G(\beta))$ , and the manifold  $f^n(U^s)$  has endpoints  $A' := \Psi_{x_n}(\alpha', G(\alpha'))$ ,  $B' := \Psi_{x_n}(\beta', G(\beta'))$ .

Since  $V^s$  and  $U^s$  intersect,  $f^n(V^s)$  and  $f^n(U^s)$  intersect. Consequently,  $[\alpha, \beta]$  and  $[\alpha', \beta']$  overlap. We use the assumption that  $p^s \leq q^s$  to show that  $[\alpha, \beta] \subseteq [\alpha', \beta']$ .

Otherwise  $\alpha < \alpha'$  or  $\beta > \beta'$ . Assume by contradiction that  $\alpha < \alpha'$ . Then  $A'$  is in the relative interior of  $f^n(V^s)$ . Since  $f$  is a homeomorphism,  $f^{-n}(A')$  is in the relative interior of  $V^s$ . Since  $f^{-n}(A')$  is an endpoint of  $U^s$ , we obtain that  $U^s$  has an endpoint at the relative interior of  $V^s$ .

We now use the assumption that  $x = y$ , and view  $V^s$  and  $U^s$  as sub-manifolds of the chart  $\Psi_x$ . The endpoints of  $U^s$  have  $s$ -coordinates equal in absolute value to  $q^s$ , and the points on  $V^s$  have  $s$ -coordinates in  $[-p^s, p^s]$ . It follows that  $q^s < p^s$ , in contradiction to our assumption. The contradiction shows that  $\alpha \geq \alpha'$ . Similarly one shows that  $\beta \leq \beta'$ , with the conclusion that  $[\alpha, \beta] \subseteq [\alpha', \beta']$ . It follows that  $f^n(V^s) \subseteq f^n(U^s)$ , whence  $V^s \subseteq U^s$ .  $\square$

**Proof of Lemma 10.8** Suppose  $Z = Z(\Psi_{x_0}^{p_0^u, p_0^s})$ ,  $Z' = Z(\Psi_{y_0}^{q_0^u, q_0^s})$  intersect. We are asked to show that for every  $x \in Z$  and  $y \in Z'$ ,  $V^u(x, Z)$  and  $V^s(y, Z')$  intersect at a unique point. Loosely speaking:

- Since  $Z, Z'$  intersect, the parameters of  $\Psi_{x_0}^{p_0^u, p_0^s}, \Psi_{y_0}^{q_0^u, q_0^s}$  are close.
- This implies that  $u$ -admissible manifolds in  $\Psi_{x_0}^{p_0^u, p_0^s}$  are very close to being  $u$ -admissible manifolds in  $\Psi_{y_0}^{q_0^u, q_0^s}$ .
- Therefore they intersect  $s$ -admissible in  $\Psi_{y_0}^{q_0^u, q_0^s}$  at a unique point.

The details follow.

Fix some  $z \in Z \cap Z'$ , then there are  $\underline{v}, \underline{w} \in \Sigma^\#$  s.t.  $v_0 = \Psi_{x_0}^{p_0^u, p_0^s}$ ,  $w_0 = \Psi_{y_0}^{q_0^u, q_0^s}$ , and  $z = \pi(\underline{v}) = \pi(\underline{w})$ . Write  $p := p_0^u \wedge p_0^s$  and  $q := q_0^u \wedge q_0^s$ . By Theorem 5.2,  $p_0^u/q_0^u, p_0^s/q_0^s, p/q \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$  and

$$\Psi_{y_0}^{-1} \circ \Psi_{x_0} = (-1)^\sigma \text{Id} + \underline{c} + \Delta \text{ on } R_\varepsilon(\underline{0}),$$

where  $\sigma \in \{0, 1\}$ ,  $\underline{c}$  is a constant vector s.t.  $\|\underline{c}\| < 10^{-1}q$ , and  $\Delta : R_\varepsilon(\underline{0}) \rightarrow \mathbb{R}^2$  satisfies  $\Delta(\underline{0}) = \underline{0}$ , and  $\|(d\Delta)_{\underline{u}}\| < \sqrt[3]{\varepsilon}$  for all  $\underline{u} \in R_\varepsilon(\underline{0})$ . By the Mean Value Theorem,  $\|\Delta(\underline{u})\| \leq \sqrt[3]{\varepsilon}\|\underline{u}\|$  for all  $\underline{u} \in R_\varepsilon(\underline{0})$ .

Now suppose  $x \in Z$ .  $V^u := V^u(x, Z)$  is a  $u$ -admissible in  $\Psi_{x_0}^{p_0^u, p_0^s}$ , therefore it can be put in the form  $V^u(x, Z) = \Psi_{x_0}\{(F(t), t) : |t| \leq p_0^u\}$ , where  $F : [-p_0^u, p_0^u] \rightarrow \mathbb{R}$  satisfies  $|F(0)| \leq 10^{-3}p$ ,  $\|F\|_\infty \leq 10^{-2}p_0^u$  and  $\text{Lip}(F) < \varepsilon$ .

We write  $V^u(x, Z)$  in  $\Psi_{y_0}$ -coordinates. Let  $\underline{c} = (c_1, c_2)$ ,  $\Delta = (\Delta_1, \Delta_2)$ , then

$$\begin{aligned} V^u(x, Z) &= [\Psi_{y_0} \circ (\Psi_{y_0}^{-1} \circ \Psi_{x_0})]\{(F(t), t) : |t| \leq p_0^u\} \\ &= \Psi_{y_0}\{((-1)^\sigma F(t) + c_1 + \Delta_1(F(t), t), (-1)^\sigma t + c_2 + \Delta_2(F(t), t)) : |t| \leq p_0^u\} \\ &= \Psi_{y_0}\{(\tilde{F}(\theta) + c_1 + \tilde{\Delta}_1(\tilde{F}(\theta), \theta), \underbrace{\theta + c_2 + \tilde{\Delta}_2(\tilde{F}(\theta), \theta)}_{=:\tau(\theta)}) : |\theta| \leq p_0^u\}, \end{aligned}$$

where we have used the transformations  $\theta := (-1)^\sigma t$ ,  $\tilde{F}(s) := (-1)^\sigma F((-1)^\sigma s)$ , and  $\tilde{\Delta}_i(u, v) := \Delta_i((-1)^\sigma u, (-1)^\sigma v)$ . Notice that  $|\tilde{F}(0)| = |F(0)| \leq 10^{-3}p$ ,  $\|\tilde{F}\|_\infty = \|F\|_\infty \leq 10^{-2}p_0^u$  and  $\text{Lip}(\tilde{F}) = \text{Lip}(F) < \varepsilon$ . Also  $\tilde{\Delta}(\underline{0}) = \underline{0}$  and  $\|(d\tilde{\Delta})_{\underline{u}}\| = \|(d\Delta)_{\underline{u}}\| < \sqrt[3]{\varepsilon}$  on  $R_\varepsilon(\underline{0})$ .

Let  $\tau(\theta) := \theta + c_2 + \tilde{\Delta}_2(\tilde{F}(\theta), \theta)$ . Assuming  $\varepsilon$  is small enough, we have

- $\tau' \in [e^{-2\sqrt[3]{\varepsilon}}, e^{2\sqrt[3]{\varepsilon}}]$ ;
- $|\tau(0)| \leq |c_2| + |\tilde{\Delta}_2(\tilde{F}(0), 0)| < 10^{-1}q + \sqrt[3]{\varepsilon} \cdot 10^{-3}p < \frac{1}{6}p$  ( $\because p \leq e^{\sqrt[3]{\varepsilon}}q$ ).



It follows that  $\tau$  is one-to-one, and  $\tau[-p_0^u, p_0^u] = [\alpha, \beta]$  where  $\alpha := \tau(-p_0^u)$  and  $\beta := \tau(p_0^u)$ . It is easy to see that  $|\alpha + p_0^u| < \frac{1}{6}p_0^u$  and  $|\beta - p_0^u| < \frac{1}{6}p_0^u$ : both quantities are less than  $|c_2| + \sup_{R_{p_0^u}(\mathbb{Q})} |\tilde{\Delta}_2|$ , which is less than  $\frac{1}{6}p_0^u$  provided  $\varepsilon$  is small enough. It follows that  $\tau[-p_0^u, p_0^u] = [\alpha, \beta] \supset [-\frac{2}{3}q, \frac{2}{3}q]$ .

Since  $\tau : [-p_0^u, p_0^u] \rightarrow [\alpha, \beta]$  is one-to-one and onto, it has a well defined inverse function  $\theta : [\alpha, \beta] \rightarrow [-p_0^u, p_0^u]$ . Let  $G(s) := \tilde{F}(\theta(s)) + c_1 + \tilde{\Delta}_1(\tilde{F}(\theta(s)), \theta(s))$ , then

$$V^u(x, Z) = \Psi_{y_0}\{(G(s), s) : s \in [\alpha, \beta]\}.$$

Using the properties of  $\tau$ , it is not difficult to check that  $\theta' \in [e^{-2\sqrt[3]{\varepsilon}}, e^{2\sqrt[3]{\varepsilon}}]$  and  $|\theta(0)| = |\theta(0) - \theta(\tau(0))| \leq e^{2\sqrt[3]{\varepsilon}}|\tau(0)| < \frac{1}{6}e^{2\sqrt[3]{\varepsilon}}p$ . It follows that  $|\tilde{F}(\theta(0))| \leq |\tilde{F}(0)| + \varepsilon|\theta(0)| < (10^{-3} + \frac{1}{6}e^{2\sqrt[3]{\varepsilon}}\varepsilon)p < 10^{-2}p$ , whence

$$\begin{aligned} |G(0)| &\leq 10^{-2}p + 10^{-1}q + \sqrt[3]{\varepsilon}p < \min\{\frac{1}{6}p, \frac{1}{6}q\} \quad (\because q/p \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]) \\ |G'| &\leq \|\tilde{F}'\|_\infty|\theta'| + \sqrt[3]{\varepsilon}\sqrt{1 + |\tilde{F}'|^2} \cdot |\theta'| < 2\sqrt[3]{\varepsilon}. \end{aligned}$$

It follows that (for all  $\varepsilon$  small enough)  $G[-\frac{2}{3}p, \frac{2}{3}p] \subset [-\frac{2}{3}p, \frac{2}{3}p]$ .

We can now show that  $|V^u(x, Z) \cap V^s(y, Z')| \geq 1$  (compare with [KM, S.3.7]). Represent

$$V^s(y, Z') = \Psi_{y_0}\{(t, H(t)) : |t| \leq q_0^s\}.$$

By admissibility,  $|H(0)| < 10^{-3}q$  and  $\text{Lip}(H) < \varepsilon$ , so  $H[-\frac{2}{3}p, \frac{2}{3}p] \subset [-\frac{2}{3}p, \frac{2}{3}p]$ . It follows that  $H \circ G$  is a contraction of  $[-\frac{2}{3}p, \frac{2}{3}p]$  into itself. Such a map has a (unique) fixed point  $(H \circ G)(s_0) = s_0$ . It is easy to see that  $\Psi_{y_0}(G(s_0), s_0)$  belongs to  $V^u(x, Z) \cap V^s(y, Z')$ .

Next we claim that  $V^u(x, Z) \cap V^s(y, Z')$  contains at most one point. Extend  $G$  and  $H$  to  $\varepsilon$ -Lipschitz functions  $\tilde{G}, \tilde{H}$  on  $[-a, a]$  where  $a := \max\{|\alpha|, |\beta|, q_0^s\}$ . By construction,  $|\tilde{G}(0)| \leq \frac{1}{6}a$ , so  $\tilde{G}[-a, a] \subset [-a, a]$ . Also  $|\tilde{H}(0)| \leq 10^{-3}a$ , so  $\tilde{H}[-a, a] \subset [-a, a]$ . It follows that  $\tilde{H} \circ \tilde{G}$  is a contraction of  $[-a, a]$  into itself, and therefore it has a unique fixed point. Every point in  $V^u(x, Z) \cap V^s(y, Z')$  takes the form  $\Psi_{y_0}(G(s), s)$  where  $s \in [\alpha, \beta]$  and  $s = (H \circ G)(s) \equiv (\tilde{H} \circ \tilde{G})(s)$ . Since the equation  $s = (\tilde{H} \circ \tilde{G})(s)$  has at most one solution in  $[-a, a]$ , it has at most one solution in  $[\alpha, \beta]$ . It follows that  $|V^u(x, Z) \cap V^s(y, Z')| \leq 1$ .  $\square$

**Proof of Lemma 10.10** We have to show: If  $Z = Z(\Psi_{x_0}^{p_0^u, p_0^s})$  and  $Z' = Z(\Psi_{y_0}^{q_0^u, q_0^s})$  intersect, then (1)  $Z \subset \Psi_{y_0}[R_{q_0^u \wedge q_0^s}(\mathbb{Q})]$ , and (2) for any  $x \in Z \cap Z'$ ,  $W^u(x, Z) \subset V^u(x, Z')$  and  $W^s(x, Z) \subset V^s(x, Z')$ .

Fix some  $x \in Z \cap Z'$ . Write  $x = \pi(\underline{v})$ ,  $x = \pi(\underline{w})$  where  $\underline{v}, \underline{w} \in \Sigma^\#$  satisfy  $v_0 = \Psi_{x_0}^{p_0^u, p_0^s}$  and  $w_0 = \Psi_{y_0}^{q_0^u, q_0^s}$ . Write  $p := p_0^u \wedge p_0^s$  and  $q := q_0^u \wedge q_0^s$ . Since  $\pi(\underline{v}) = \pi(\underline{w})$ , we have by Theorem 5.2 that  $p/q \in [e^{-\sqrt[3]{\varepsilon}}, e^{\sqrt[3]{\varepsilon}}]$  and

$$\Psi_{y_0}^{-1} \circ \Psi_{x_0} = (-1)^\sigma \text{Id} + \underline{c} + \Delta \text{ on } R_\varepsilon(\mathbb{Q}),$$

where  $\sigma \in \{0, 1\}$ ,  $\underline{c}$  is a constant vector s.t.  $\|\underline{c}\| < 10^{-1}q$ , and  $\Delta : R_\varepsilon(\mathbb{Q}) \rightarrow \mathbb{R}^2$  satisfies  $\Delta(\mathbb{Q}) = \mathbb{Q}$ , and  $\|(d\Delta)_{\underline{u}}\| < \sqrt[3]{\varepsilon}$  for all  $\underline{u} \in R_\varepsilon(\mathbb{Q})$ . By the Mean Value Theorem,  $\|\Delta(\underline{u})\| \leq \sqrt[3]{\varepsilon}\|\underline{u}\|$  for all  $\underline{u} \in R_\varepsilon(\mathbb{Q})$ .

Every point in  $Z$  is the intersection of a  $u$ -admissible and an  $s$ -admissible manifold in  $\Psi_{x_0}^{p_0^u, p_0^s}$ , therefore  $Z$  is contained in  $\Psi_{x_0}[R_{10^{-2}p}(\underline{0})]$  (Proposition 4.11). Thus

$$\begin{aligned} Z &\subseteq \Psi_{y_0}[(\Psi_{y_0}^{-1} \circ \Psi_{x_0})[R_{10^{-2}p}(\underline{0})]] \subset \Psi_{y_0}[(\Psi_{y_0}^{-1} \circ \Psi_{x_0})[B_{\sqrt{2} \cdot 10^{-2}p}(\underline{0})]] \\ &\subseteq \Psi_{y_0}[B_{(1+\sqrt[3]{\varepsilon})\sqrt{2} \cdot 10^{-2}p}(\underline{c})] \subseteq \Psi_{y_0}[B_{(1+\sqrt[3]{\varepsilon})\sqrt{2} \cdot 10^{-2}e^{\sqrt[3]{\varepsilon}q+10^{-1}q}}(\underline{0})] \\ &\subseteq \Psi_{y_0}[R_{(1+\sqrt[3]{\varepsilon})\sqrt{2} \cdot 10^{-2}e^{\sqrt[3]{\varepsilon}q+10^{-1}q}}(\underline{0})] \subset \Psi_{y_0}[R_q(\underline{0})] \quad (\because 0 < \varepsilon < 1). \end{aligned}$$

This proves the first statement of the lemma.

Next we show that  $W^s(x, Z) \subset V^s(x, Z')$ . Write  $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$  and  $w_i = \Psi_{y_i}^{q_i^u, q_i^s}$ . Since  $x = \pi(\underline{v})$  and  $Z = Z(v_0)$ , we have by the symbolic Markov property that

$$f^k[W^s(x, Z)] \subset W^s(f^k(x), Z(v_k)) \quad (k \geq 0).$$

The sets  $Z(v_k)$  and  $Z(w_k)$  intersect, because they both contain  $f^k(x)$ . By the first part of the lemma,  $Z(v_k) \subset \Psi_{y_k}[R_{q_k^u \wedge q_k^s}(\underline{0})]$ . It follows that

$$f^k[W^s(x, Z)] \subset \Psi_{y_k}[R_{q_k^u \wedge q_k^s}(\underline{0})] \subset \Psi_{y_k}[R_{Q_\varepsilon(y_k)}(\underline{0})]$$

for all  $k \geq 0$ . By Proposition 4.15 part 4,  $W^s(x, Z) \subset V^s[(w_i)_{i \geq 0}] \equiv V^s(x, Z')$ .  $\square$

**Acknowledgements.** This work was partially supported by the NSF grant DMS-0400687 and by the ERC award ERC-2009-StG n $^\circ$  239885. The author would like to thank J. Buzzi, A. Katok, F. Ledrappier, and M. Pollicott for useful discussions.

**Note added in proof.** Recently Pierre Berger has come up with a construction of countable Markov partitions for certain Hénon like diffeomorphisms. For these maps he proved that the measure of maximal entropy is unique [**Brg**].

## REFERENCES

- [AW] R. Adler and B. Weiss: *Entropy, a complete metric invariant for automorphisms of the torus*, Proc. of the Nat. Acad. of Sciences **57** (1967), 1573–1576.
- [BP] L. Barreira and Y. Pesin: Nonuniform hyperbolicity: dynamics of systems with nonzero Lyapunov exponents. *Encycl. of Math. and Its Appl.* **115**, Cambridge UP (2007), xiv+513pp.
- [BY] M. Benedicks; L.-S. Young: *Markov extensions and decay of correlations for certain Hénon maps*. (English, French summary) *Géométrie complexe et systèmes dynamiques* (Orsay, 1995). *Astérisque* **261** (2000), xi, 13–56.
- [Be] K. Berg: *On the conjugacy problem for K-systems*, Ph.D. thesis., University of Minnesota (1967).
- [Brg] P. Berger: *Properties of the maximal entropy measure and geometry of Hénon attractors*. arXiv:1202.2822v1 (2012), 53 pages.
- [B1] R. Bowen: *Markov partitions for axiom A diffeomorphisms*, American J. Math. **92** (1970), 725–747.
- [B2] R. Bowen: *Periodic points and measures for Axiom A diffeomorphisms*, Transactions of the Amer. Math. Soc. **154** (1971), 377–397.
- [B3] R. Bowen: *On Axiom A diffeomorphisms*. Regional Conference Series in Mathematics, No. **35**. American Mathematical Society, Providence, R.I., 1978. vii+45 pp.
- [B4] R. Bowen: *Equilibrium states and the ergodic theory of Anosov diffeomorphisms. Second revised edition. With a preface by David Ruelle. Edited by Jean-René Chazottes. Lecture Notes in Mathematics, 470 Springer-Verlag, Berlin, 2008. viii+75 pp*
- [BD] M. Boyle, T. Downarowicz: *The entropy theory of symbolic extensions*, Invent. Math. **156** (2004), no. 1, 119–161.
- [BFF] M. Boyle, D. Fiebig, and U. Fiebig: *Residual entropy, conditional entropy and subshift covers*. Forum Math. **14** (2002), no. 5, 713–757.

- [Bri] M. Brin: *Hölder continuity of invariant distributions*, in Smooth Ergodic Theory and its Applications, edited by A. Katok, R. de la Llave, Ya. Pesin, and H. Weiss, Proc. Symp. Pure Math. **69**, AMS 2001, pp. 99–101.
- [Bru] H. Bruin: *Induced maps, Markov extensions and invariant measures in one-dimensional dynamics*. Comm. Math. Phys. **168** (1995), no. 3, 571–580.
- [BT] H. Bruin, M. Todd: *Markov extensions and lifting measures for complex polynomials*. Ergodic Theory Dynam. Systems **27** (2007), no. 3, 743–768.
- [BS] L.A. Bunimovich; Ya.G. Sinai: *Markov partitions for dispersed billiards*. Comm. Math. Phys. **78** (1980/81), 247–280. *Erratum*: Comm. Math. Phys. **107** (1986), 357–358.
- [BSC] L.A. Bunimovich; Ya.G. Sinai; N.I. Chernov: *Markov partitions for two-dimensional hyperbolic billiards*. (Russian) Uspekhi Mat. Nauk **45** (1990), no. 3(273), 97–134, 221; translation in Russian Math. Surveys **45** (1990), no. 3, 105–152.
- [Bur] D. Burguet:  *$C^2$  surface diffeomorphisms have symbolic extensions*, Invent. Math. **186** (2011), no. 1, 191–236.
- [Bu1] J. Buzzi: *Intrinsic ergodicity of smooth interval maps*, Israel J. Math. **100** (1997), 125–161.
- [Bu2] J. Buzzi: *Markov extensions for multi-dimensional dynamical systems*. Israel J. Math. **112** (1999), 357–380.
- [Bu3] J. Buzzi: *Subshifts of quasi-finite type*. Invent. Math. **159** (2005), no. 2, 369–406.
- [Bu4] J. Buzzi: *Maximal entropy measures for piecewise affine surface homeomorphisms*, Ergod. Th. & Dynam. Sys. (2009), **29**, 1723–1763.
- [Bu5] J. Buzzi: *Puzzles of quasi-finite type, zeta functions, and symbolic dynamics for multi-dimensional maps*, Ann. Inst. Fourier **60** (2010), no. 3, 801–852.
- [DN] T. Downarowicz, and S. Newhouse: *Symbolic extensions and smooth dynamical systems*, Invent. Math. **160** (2005), no. 3, 453–499.
- [FS] A. Fathi and M. Shub: *Some dynamics of pseudo-Anosov diffeomorphisms*, Astérisque **66-67** (1979), 181–207.
- [G] T.N.T. Goodman: *Relating topological entropy and measure entropy*. Bull. London Math. Soc. **3** (1971) 176–180.
- [Gu1] B.M. Gurevič: *Topological entropy of a countable Markov chain*. Dokl. Akad. Nauk SSSR **187** (1969) 715–718. *Engl. Trans.* Soviet Math. Dokl. **10** (1969), 911–915.
- [Gu2] B.M. Gurevič: *Shift entropy and Markov measures in the space of paths of a countable graph*, Dokl. Akad. Nauk SSSR **192** (1970) 963–965.
- [Hof1] F. Hofbauer: *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy*. Israel J. Math. **34** (1979), no. 3, 213–237 (1980).
- [Hof2] F. Hofbauer: *The structure of piecewise monotonic transformations with positive entropy*, Ergodic Th. & Dynam. Syst. **1** (1981), 159–178.
- [IT] G. Iommi and M. Todd: *Natural equilibrium states for multimodal maps*. Comm. Math. Phys. **300** (2010), no. 1, 65–94.
- [Kal] V. Kaloshin: *Generic diffeomorphisms with superexponential growth of number of periodic orbits*. Comm. Math. Phys. **211** (2000), no. 1, 253–271.
- [K1] A. Katok: *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Pub. Math. de l’Institut des Hautes Études Scientifiques, **51** (1980), 137–173.
- [K2] A. Katok: *Non-uniform hyperbolicity and periodic orbits for smooth dynamical systems*, Proc. Inter. Congress of Math., PWN-North Holland (1984), 1245–1253.
- [K3] A. Katok: *Fifty years of entropy in dynamics: 1958–2007*, J. Modern Dynamics **1** no. 4 (2007), 545–596.
- [KH] A. Katok and B. Hasselblatt: *Introduction to the modern theory of dynamical systems. Encycl. of Math. and Its Appl.* **54**, Cambridge UP (1995), xviii+802pp.
- [KM] A. Katok and L. Mendoza: *Dynamical systems with non-uniformly hyperbolic behavior*, Supplement to “Introduction to the modern theory of dynamical systems.” Cambridge UP (1995), 659–700.
- [Ke1] G. Keller: *Lifting measures to Markov extensions*. Monatsh. Math. **108** (1989), no. 2-3, 183–200.
- [Ke2] G. Keller: *Markov extensions, zeta functions, and Fredholm theory for piecewise invertible dynamical systems*. Trans. Amer. Math. Soc. **314** (1989), no. 2, 433–497.
- [Ki] B.P. Kitchens: *Symbolic Dynamics. One-sided, two-sided and countable state Markov shifts*. Universitext. Springer-Verlag, Berlin, 1998.  $x+252$  pp.

- [KT] T. Krüger; S. Troubetzkoy: *Markov partitions and shadowing for non-uniformly hyperbolic systems with singularities*. Ergodic Theory Dynam. Systems **12** (1992), no. 3, 487–508.
- [M] G. Margulis: *On some aspects of the theory of Anosov systems*, With a survey by Richard Sharp: *Periodic orbits of hyperbolic flows*, Springer 2003.
- [N] S. Newhouse: *Continuity properties of entropy*. Ann. of Math. (2) **129** (1989), no. 2, 215–235.
- [Os] V. Oseledets: *A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems*, Trans. Mosc. Math. Soc. **19** (1968), 197–221.
- [PP] W. Parry and M. Pollicott: Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* **187–188** (1990), 268 pp.
- [P] Ja. B. Pesin: *Families of invariant manifolds corresponding to nonzero characteristic exponents*, Math. USSR Izvestija **10** (1976) 1261–1305. (*English Transl.*)
- [PSZ] Ya. Pesin, S. Senti, K. Zhang: *Lifting measures to inducing schemes*. Ergodic Theory Dynam. Systems **28** (2008), no. 2, 553–574.
- [Ru] D. Ruelle: *An inequality for the entropy of differentiable dynamical systems*, Bol. Soc. Bras. Mat. **9** (1978), 83–87.
- [Si1] Ya.G. Sinai: *Construction of Markov partitions*, Func. Anal. and its Applic. **2** (1968), 70–80.
- [Si2] Ya.G. Sinai: *Gibbs measures in ergodic theory*, Uspehi Mat. Nauk. **27** (1972), 21–64.
- [Sm] S. Smale: *Differentiable dynamical systems*, Bull. AMS **73** (1967), 747–817.
- [Sp] M. Spivak: A comprehensive introduction to differential geometry, vol 1. *Publish or Perish Inc., Houston, Texas, Second Edition, 1979*.
- [Sr] S.M. Srivastava: A course on Borel sets, *Graduate Texts in Math.* **180**, Springer-Verlag, New York, 1998. *xvi+261 pp.*
- [Ta] Y. Takahashi: *Isomorphisms of  $\beta$ -transformations to Markov automorphisms*. Osaka J. Math. **10** (1973), 175–184.
- [Y] L.-S. Young: *Statistical properties of dynamical systems with some hyperbolicity*. Ann. of Math. (2) **147** (1998), no. 3, 585–650.
- [Z] R. Zweimüller: *Ergodic properties of infinite measure-preserving interval maps with indifferent fixed points*. Ergodic Theory Dynam. Systems **20** (2000), no. 5, 1519–1549.

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, THE WEIZMANN INSTITUTE OF SCIENCE,  
 234 HERZL STREET, POB 26, REHOVOT 76100, ISRAEL  
*E-mail address:* omsarig@gmail.com