SYMBOLIC DYNAMICS FOR SURFACE DIFFEOMORPHISMS WITH POSITIVE ENTROPY

OMRI M. SARIG

Abstract. Let $f$ be a $C^r$ diffeomorphism ($r > 1$) on a compact orientable smooth surface. Suppose the topological entropy $h_{\text{top}}(f)$ is positive. Given $0 < \chi < h_{\text{top}}(f)$, we construct a countable Markov partition for the restriction of $f$ to an invariant set which is “large” in the sense that it has full measure with respect to every ergodic invariant probability measure with entropy greater than $\chi$. The following results follow: (1) $f$ has at most countably many ergodic measures of maximal entropy (a conjecture of J. Buzzi), and (2) if $f$ is $C^\infty$, then $\limsup_{n \to \infty} e^{-n h_{\text{top}}(f)} \# \{ x : f^n(x) = x \} > 0$ (a conjecture of A. Katok).

Contents

Part 0. Introduction and statement of results
  1.1. Results
  1.2. Symbolic dynamics
  1.3. Markov partitions
  1.4. Comparison to other results in the literature
  1.5. Overview of the construction of a Markov partition
  1.6. Notational conventions

Part 1. Chains as pseudo–orbits
  2. Pesin charts
    2.1. Non-uniform hyperbolicity
    2.2. Lyapunov change of coordinates
    2.3. Pesin Charts
    2.4. NUH $\chi$ $(f)$
  3. Overlapping charts
    3.1. The overlap condition
    3.2. The form of $f$ in overlapping charts
    3.3. Coarse graining
  4. $\varepsilon$–chains and an infinite-to-one Markov extension of $f$
    4.1. Double charts and $\varepsilon$–chains
    4.2. Admissible manifolds and the graph transform
    4.3. A Markov extension
4.4. The relevant part of the extension

Part 2. Regular chains which shadow the same orbit are close
5. The inverse problem for regular chains
6. Axes parameters
6.1. Staying in windows
6.2. Comparison of \( \alpha(x_i) \) to \( \alpha(y_i) \)
6.3. Comparison of \( R_{x_i} \) to \( R_{y_i} \)
7. Scaling parameters
7.1. The \( s_\chi \) and \( u_\chi \) parameters of admissible manifolds
7.2. Comparison of \( s_\chi(x_i), u_\chi(x_i) \) to \( s_\chi(y_i), u_\chi(y_i) \).
8. Window parameters
8.1. \( \varepsilon \)-maximality
8.2. Comparison of \( p_i^{u/s} \) to \( q_i^{u/s} \)
9. Proof of Theorem 5.2

Part 3. Markov partitions and symbolic dynamics
10. A locally finite countable Markov cover
10.1. The cover
10.2. Product structure
10.3. The symbolic Markov property
11. A countable Markov partition
11.1. The Bowen–Sinai refinement
11.2. Product structure and hyperbolicity
11.3. The Markov property
12. Symbolic dynamics
12.1. A directed graph
12.2. The Markov extension
12.3. The extension is finite-to-one
13. Invariant measures

Part 4. Appendix: Proofs of standard results in Pesin Theory

Acknowledgements
References

Part 0. Introduction and statement of results

1.1. Results. Let \( M \) be a compact orientable \( C^\infty \) Riemannian manifold of dimension two, and let \( f: M \to M \) be a \( C^{1+\beta} \) diffeomorphism, where \( 0 < \beta < 1 \). We assume throughout that the topological entropy of \( f \) is positive.

Let \( P_n(f) := \{|x \in M : f^n(x) = x|\} \). Anatole Katok showed in [K1] and [K2] that \( \limsup_{n \to \infty} \frac{1}{n} \log P_n(f) \geq h_{\text{top}}(f) \), and conjectured in [K3] that if \( f \) is \( C^\infty \) then \( \limsup_{n \to \infty} e^{-nh_{\text{top}}(f)} P_n(f) > 0 \).
Theorem 1.1. Suppose \( f \) is a \( C^{1+\beta} \) diffeomorphism of a compact orientable smooth surface, and assume \( h_{\text{top}}(f) > 0 \). If \( f \) has a measure of maximal entropy, then \( \exists p \in \mathbb{N} \) s.t. \( \liminf_{n \to \infty} e^{-n h_{\text{top}}(f)} P_n(f) > 0 \).

This proves Katok’s conjecture, because \( C^\infty \) diffeomorphisms on compact manifolds have measures of maximal entropy (Newhouse [N]). Theorem 1.1 was conjectured to hold as stated above by Jérôme Buzzi [Bu4].

It was also conjectured in [Bu4] that \( f \) admits at most countably many different ergodic measures of maximal entropy. This turns out to be correct:

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

Buzzi 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], continuous piecewise affine homeomorphisms of affine surfaces [Bu4]. There are also results on non-invertible maps, see [Hof1, Hof2] and [Bu4, Bu5].

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 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

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

We equip \( \Sigma \) with the natural metric:

\[
d(u,v) := \exp[-\min\{|i| : u_i \neq v_i\}],
\]

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 \( \sigma : \Sigma \to \Sigma \) is defined by \( \sigma[(v_i)_{i \in \mathbb{Z}}] = (v_{i+1})_{i \in \mathbb{Z}} \).

Let \( \Sigma^\# := \{(v_i)_{i \in \mathbb{Z}} \in \Sigma : \exists u, v \in \mathcal{V} \forall n_k, m_k \uparrow \infty \text{ s.t. } v_{-m_k} = u, v_{n_k} = v\} \). \( \Sigma^\# \) 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 \( 0 < \chi < h_{\text{top}}(f) \) there exists a locally compact topological Markov shift \( \Sigma_\chi \) and a Hölder continuous map \( \pi_\chi : \Sigma_\chi \to M \) 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 \( \mathcal{Y}_\chi \). There exists a function \( \varphi_\chi : \mathcal{Y}_\chi \times \mathcal{Y}_\chi \to \mathbb{N} \) 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.

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 1.3, 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 \( \sigma : \Sigma(\mathcal{G}') \to \Sigma(\mathcal{G}) \), it is also a measure of maximal entropy for \( \sigma : \Sigma(\mathcal{G}') \to \Sigma(\mathcal{G}) \).

The irreducibility of \( \mathcal{G}' \) means that \( \sigma : \Sigma(\mathcal{G}') \to \Sigma(\mathcal{G}) \) 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 \( \exists \rho \in \mathbb{N} \) s.t. for every vertex \( v_0 \) in \( \mathcal{G}' \),

\[
|\{ \nu \in \Sigma(\mathcal{G}') : \sigma^n(\nu) = \nu, v_0 = v \}| \leq \exp[nh_{\max}(\Sigma(\mathcal{G}'))]
\]

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. \( \sigma \). Here and throughout, \( \sim \) means equality up to bounded multiplicative error.

Since \( \pi_\chi \circ \sigma = f \circ \pi_\chi \), the collection \( \{ \nu \in \Sigma(\mathcal{G}') : \sigma^n(\nu) = \nu, v_0 = v \} \) 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 \( \liminf_{n \to \infty} e^{-nh_{\max}(\Sigma(\mathcal{G}'))) P_\nu(f) > 0 \). By construction, \( h_{\max}(\Sigma(\mathcal{G}')) = h_\mu(\sigma) = h_\mu(f) = \max \{ h_\nu(f) : \nu \text{ } f \text{-inv.} \} \). The last quantity is equal to \( h_{\top}(f) \) by the variational principle [C]. 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) \( \Sigma(\mathcal{G}') \) has a measure of maximal entropy, and (3) \( h_{\max}(\Sigma(\mathcal{G}')) = h_{\max}(\Sigma(\mathcal{G}')) \).

By a theorem of Salama [Sal] (see also Ruette [Rut]), if \( \Sigma(\mathcal{G}') \) carries a measure of maximal entropy, then every addition of a vertex or an edge to \( \mathcal{G}' \) increases \( h_{\max}(\Sigma(\mathcal{G}')) \). This implies that the subgraphs \( \mathcal{G}' \subset \mathcal{G} \) which satisfy (1), (2), and (3) have disjoint sets of vertices. Since \( \mathcal{G} \) is countable, there can be at most countably many such subgraphs, and 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 \( \mathcal{B} \) of Borel sets with the following properties:

1. **Covering property:** The union of \( \mathcal{B} \) is \( \chi \)-large.
2. **Product structure:** There are \( W^+(x, R), W^+(x, R) \subset R \) (\( x \in R \in \mathcal{B} \)) s.t.
   - (a) \( W^+(x, R) \cap W^+(x, R) = \{ x \} \).
   - (b) \( \forall x, y \in R, \exists x \in R \text{ s.t. } W^+(x, R) \cap W^+(y, R) = \{ z \} \).
   - (c) \( \forall x, y \in R, W^+(x, R) \text{ and } W^+(y, R) \text{ are equal, or they are disjoint.} \)

   Similarly for \( W^-(x, R), W^-(y, R) \).
3. **Hyperbolicity:** If \( y, z \in W^+(x, R) \), then \( d(f^n(y), f^n(z)) \xrightarrow{n \to \infty} 0 \). If \( y, z \in W^-(x, R) \), then \( d(f^{-n}(y), f^{-n}(z)) \xrightarrow{n \to \infty} 0 \).
(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, KM], 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. $f : \Lambda_\varepsilon \to \Lambda_\varepsilon$ 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. But (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) seem to be unavoidable.

**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, 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 1.3; see also [BD, DK]). 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 $\sigma : \Lambda \to \Lambda$ where $\Lambda \subseteq \{1, \ldots, N\}^Z$ is closed and shift invariant and $\sigma$ is the left shift (“subshift”). Recently, Burguet has shown that every $C^2$ surface diffeomorphism has a symbolic extension [Bur].

Unlike symbolic extensions, our coding is by a non–compact shift space. On the positive side, our space has Markov structure. This gives us access to many results which are not true for general subshifts, e.g. Gurevich’s theory mentioned in the end of §1.2.

### 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 $x = \{x_i\}_{i \in Z}$ such that $d(x_{i+1}, f(x_i)) < \varepsilon$ for all $i$. A pseudo–orbit $x$ is said to $\delta$–shadow a real orbit $\{f^i(x)\}_{i \in Z}$ if $d(x_i, f^i(x)) < \delta$ for all $i \in Z$. Anosov showed that for every $\delta$ small enough, there exists an $\varepsilon > 0$ s.t.

(A1) Every $\varepsilon$–pseudo–orbit $x$ $\delta$–shadows the real orbit of some unique point $\pi(x)$.

(A2) “Finite alphabet suffices”: There exists a finite set of points $A$ such that $\{\pi(x) : x \in A^Z\}$ is an $\varepsilon$–pseudo-orbit is the entire manifold.

(A3) “Inverse problem”: If two pseudo–orbits $x, y$ $\delta$–shadow the same orbit, then their corresponding coordinates are close, $d(x_i, y_i) < 2\delta$ for all $i \in Z$. 

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

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

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

The sets $\overline{0}[a] := \{ \overline{x} \in \Sigma : x_0 = a \}$ form a natural Markov partition for the left shift on $\Sigma$. Their projections $Z(a) = \{ \pi(\overline{x}) : \overline{x} \in \Sigma, x_0 = a \}$ ($a \in A$) would have been natural candidates for a Markov partition, had they not overlapped. Sinai came up with a set–theoretic procedure for refining $Z := \{ Z(a) : a \in A \}$ 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 use a different device to generate orbits from symbolic sequences. 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 $\overline{u}$ corresponds to a unique real orbit $\pi(\overline{u})$;
(A2') There is a countable set $A$ of symbols s.t. $\{ \pi(\overline{u}) : \overline{u} \in A^Z \text{ is an } \varepsilon \text{–chain} \}$ 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 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 $\overline{u}, \overline{v}$ are “regular” and $\pi(\overline{u}) = \pi(\overline{v})$, 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.

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

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) $\{\pi(u) : u \in A^2, u$ is a regular $\varepsilon$–chain $\}$ has full measure w.r.t. any ergodic invariant probability measure with entropy more than $\chi$. As a result, the sets $Z(\varepsilon) := \{\pi(u) : u \in A^2$ is a regular $\varepsilon$–chain $\}$ form a locally finite cover $\mathcal{X}$ of a large set.

Sinai’s refinement procedure can now be safely applied to $\mathcal{X}$. In part 3, we check that the elements of $\mathcal{X}$ 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 modifications of well–known results in Pesin Theory. Part 4 collects their proofs.

1.6. Notational conventions. In what follows, $M$ is a compact orientable $C^\infty$ Riemannian manifold of dimension two, and $f : M \to M$ is a $C^{1+\beta}$ diffeomorphism where $0 < \beta < 1$. We assume that the topological entropy of $f$ is positive, and we fix once and for all $0 < \chi < h_{\text{top}}(f)$.

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”.

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

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

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 $\frac{a_n}{b_n} \to 1$, and $a_n \asymp b_n$ means that $\exists N, c \text{ where } \forall n > N \ (e^{-c}b_n \leq a_n \leq e^{c}b_n)$. Finally, $\wedge b := \min\{a, b\}$.

Some abbreviations: s.t. “such that”, w.r.t. “with respect to”, i.o. “infinitely often”, resp. “respectively”, and w.l.o.g “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 [O3] (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 (invariant) set $\text{NUH}_\chi(f)$ of points $x \in M$ for which there is a decomposition $T_xM = E^s(x) \oplus E^u(x)$ so that

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


The splitting $T_x M = E^s(x) \oplus E^u(x)$ is unique, but the vectors $e^s(x), e^u(x)$ are only determined up to a sign. Fix a measurable family of positively oriented bases $(e^s_1, e^s_2)$ of $T_x M \ (x \in M)$. Choose the signs of $e^s/u(x)$ in such a way that $\angle(e^s_1, e^s(x)) \in [0, \pi)$ and $(e^u_1(x), e^u(x))$ has positive orientation.

The set $\text{NUH}(f) := \bigcup_{\chi > 0} \text{NUH}_\chi(f)$ 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

$$\log \lambda(x) := \lim_{n \to \pm \infty} \frac{1}{n} \log \| (df^n)_x \chi^s(x) \|_{f^n(x)}\quad (x \in \text{NUH}(f))$$

$$\log \mu(x) := \lim_{n \to \pm \infty} \frac{1}{n} \log \| (df^n)_x \chi^u(x) \|_{f^n(x)}$$

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

$$s_\chi(x) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2\chi k} \| (df^k)_x \chi^s(x) \|^2_{f^k(x)} \right)^{1/2};$$

$$u_\chi(x) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2\chi k} \| (df^{-k})_x \chi^u(x) \|^2_{f^{-k}(x)} \right)^{1/2}.$$  

(The factor $\sqrt{2}$ is needed for Lemma 2.5 below.)

Definition 2.2. The Lyapunov change of coordinates (with parameter $\chi$) is the linear map $C_\chi(x) : \mathbb{R}^2 \to T_x M \ (x \in \text{NUH}_\chi(f))$ s.t. $C_\chi(x) e_1 = s_\chi(x)^{-1} \chi^s(x)$, and $C_\chi(x) e_2 = u_\chi(x)^{-1} \chi^u(x)$, where $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

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 $e^{-2\chi k} \lambda(x)^{-2k}$ or $e^{-2\chi k} \mu(x)^{2k}$ replacing $e^{2\chi k}$. 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 \( \|L\| := \max\{\|Lg\|_W : \|g\|_V = 1\} \), and the Frobenius norm of \( L \) is \( \|L\|_{Fr} := \sqrt{\text{tr}(L^*L)} \), where \( \Theta \) is some (any) isometry \( \Theta : W \rightarrow V \). \( \|L\|_{Fr} \) is well defined\(^2\) and \( \|L\| \leq \|L\|_{Fr} \leq \sqrt{2}\|L\| \). 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 \( \|L\|_{Fr} = \left( \sum_{ij} a_{ij}^2 \right)^{1/2} \). 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: \( \|C_\chi(x)(x)\|_x \leq \|x\|_x \) for all \( x \in \mathbb{R} \).

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

\[
\begin{align*}
(1) \lim_{k \rightarrow \pm \infty} \frac{1}{k} \log |C_\chi(f^k(x))|^{-1} &= 0; \\
(2) \lim_{k \rightarrow \pm \infty} \frac{1}{k} \log \|C_\chi(f^k(x))\|_{L^2} = 0, \text{ where } \xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \\
(3) \lim_{k \rightarrow \pm \infty} \frac{1}{k} \log |\det C_\chi(f^k(x))| &= 0.
\end{align*}
\]

### 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:

\[
\begin{align*}
B_\eta(x) &:= \{ y \in M : d(x, y) < \eta \} & B_\eta(\underline{0}) &:= \{ \underline{u} \in \mathbb{R}^2 : \underline{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \sqrt{u_1^2 + u_2^2} < \eta \} \\
B_\eta^* &:= \{ \underline{u} \in T_x M : \|\underline{u}\|_x < \eta \} & R_\eta(\underline{0}) &:= \{ \underline{u} \in \mathbb{R}^2 : \underline{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, |u_1|, |u_2| < \eta \}
\end{align*}
\]

Since \( M \) is compact, there exist \( r(M), \rho(M) > 0 \) s.t. for every \( x \in M \), \( \exp_x \) maps \( B_{2r(M)}(\underline{0}) \) diffeomorphically onto a neighborhood of \( B_{\rho(M)}(x) \). (2.1)

We take \( \rho(M) \) so small that \( (x, y) \mapsto \exp_x^{-1}(y) \) is well defined and \( 2 \)-Lipschitz on \( B_{\rho(M)}(x) \times B_{\rho(M)}(z) \) 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 \( C_\chi(x) \) is a contraction,

\[
\Psi_x := \exp_x \circ C_\chi(x)
\]

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

Let \( f_x := (\Psi_x^{-1})_f \circ \Psi_x \), 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.

\(^2\text{Proof: } \text{tr}(\Theta^2 L'^*L \Theta_1) = \text{tr}(\Theta^2 \Theta_1 \Theta_1^* L' L \Theta_1) = \text{tr}(\Theta^2 L'^*L \Theta_1).\)

\(^3\text{Proof: Let } s_1(L) \geq s_2(L) \text{ denote the singular values of } L \text{ (equal by definition to the eigenvalues of } \sqrt{L'^*L}, \text{ then } \|L\| = s_1(L), \text{ and } \|L\|_{Fr} = \sqrt{s_1(L)^2 + s_2(L)^2}.\)

\(^4\text{Proof: Let } \Theta : W \rightarrow V \text{ be the isometry which maps the base we chose for } W \text{ to the base we chose for } V, \text{ then } L \Theta : W \rightarrow W \text{ is represented w.r.t. the base we chose for } W \text{ by the matrix } (a_{ij}). \text{ A calculation shows that } \text{tr}(\Theta^* L' \Theta) = \sum a_{ij}^2.\)
For reasons that will become clear later, we prefer to define it as a quantity taking values in \( I_\varepsilon := \{ e^{-\frac{t}{\ell}} : \ell \in \mathbb{N} \} \), where \( \varepsilon \) will be determined later. Set

\[
Q_\varepsilon(x) := \max\{ q \in I_\varepsilon : q \leq \bar{Q}_\chi(x) \}
\]

\[
\bar{Q}_\chi(x) := \varepsilon^{3/\beta} (\|C_\chi(x)^{-1}\|_{F_r})^{-12/\beta}
\]

(2.3)

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

1. \( \Psi_x : R_{10Q_\varepsilon(x)}(U) \to M \) is a diffeomorphism, \( \Psi_x(U) = x \), and \( \|d\Psi_x\| \leq 2 \) on \( R_{10Q_\varepsilon(x)}(U) \);
2. \( f_x := \Psi_x^{-1} \circ f \circ \Psi_x \) is well defined and injective on \( R_{10Q_\varepsilon(x)}(U) \) and
   (a) \( f_x(U) = U \) and \( (d\xi_x, 0) \)
   \[
   \left( \begin{array}{cc} A(x) & 0 \\ 0 & B(x) \end{array} \right)
   \]
   where \( C_f^{-1} < |A(x)| < \varepsilon^{-\chi} \)
   and \( \varepsilon^\chi < |B(x)| < C_f \) (cf. Theorem 2.8);
   (b) The \( C^{1+\frac{3}{2}} \) distance between \( f_x \) and \( (d\xi_x, 0) \) on \( R_{10Q_\varepsilon(x)}(U) \) is less than \( \varepsilon \).
3. The symmetric statement holds for \( f_x = \Psi_x^{-1} \circ f^{-1} \circ \Psi_f(x) \).

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 \( \Psi^\chi_x \) is the map \( \Psi^\chi_x : R_{\eta}(U) \to M \).

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. \( \|C_\chi(F^i(x))^{-1}\|_{F_r} < \varepsilon^{2/\beta}/Q_\varepsilon(x) \) 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. \( \frac{1}{n} \log Q_x(f^n(x)) \to 0 \) 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 \( q_\varepsilon : \text{NUH}_\chi^*(f) \to (0, 1) \) so that \( q_\varepsilon(x) < \varepsilon Q_\varepsilon(x) \) and
   \[
   e^{-\varepsilon/3} \leq q_\varepsilon \circ f/q_\varepsilon \leq e^{\varepsilon/3} \text{ on } \text{NUH}_\chi^*(f).
   \]

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

\[
\text{NUH}_\chi^*(f) := \{ x \in \text{NUH}_\chi^*(f) : \limsup_{n \to \infty} q_\varepsilon(f^n(x)) = \limsup_{n \to \infty} q_\varepsilon(f^{-n}(x)) \neq 0 \}
\]

(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 \( \mathcal{C} := \{ \Psi^\chi_x : x \in \text{NUH}_\chi^*(f), 0 < \eta \leq Q_\varepsilon(x) \} \) 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 the dynamics of \( f \).

3.1. The overlap condition. We need to compare the maps \( C_\chi(x) : \mathbb{R}^2 \to T_xM \) 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 \( \Theta_D : TD \to \mathbb{R}^2 \) s.t.

1. \( \Theta_D : T_xM \to \mathbb{R}^2 \) is a linear isometry for every \( x \in D \);
Lemma 2.5 and the general inequality \( \| \cdot \| \).

Two Pesin charts \( \Psi_{x_1}, \Psi_{x_2} \) \( \varepsilon \)-overlap if \( e^{-\varepsilon} < \frac{\eta_i}{\eta_j} < e^{\varepsilon} \), and for some \( D \in \mathcal{D} \), \( x_1, x_2 \in D \) and \( d(x_1, x_2) + \| \Theta_D \circ C_{\chi}(x_1) - \Theta_D \circ C_{\chi}(x_2) \| < \eta_1^2 \eta_2^2 \).

The overlap condition is symmetric. It is also monotone: if \( \Psi_{x_1}^\gamma \) \( \varepsilon \)-overlap, then \( \Psi_{x_2}^\xi \) \( \varepsilon \)-overlap for all \( \eta_i \le \xi_i \le Q_\varepsilon(x_i) \) s.t. \( e^{-\varepsilon} < \xi_1 / \xi_2 < e^\varepsilon \). Notice that the overlap requirement is stronger at areas of NUH \( \chi(x) \) where \( s_i(x) \) or \( u_i(x) \) are large or where \( e^{u}(x) \) and \( e^{s}(x) \) are nearly parallel. This is because construction

\[
\eta_i \le Q_\varepsilon(x_i) \ll \| C_{\chi}(x_i)^{-1} \|_{F_\varepsilon} = \frac{\sqrt{s_i(x)^2 + u_i(x)^2}}{|\sin \alpha(x)|}.
\]

The following proposition explains what the overlap condition means.

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

1. \( \Psi_{x_1}[R_{e^{-2\varepsilon}\eta_1}(\mathcal{U})] \subset \Psi_{x_2}[R_{\eta_2}(\mathcal{U})] \) and \( \Psi_{x_2}[R_{e^{-2\varepsilon}\eta_2}(\mathcal{U})] \subset \Psi_{x_1}[R_{\eta_1}(\mathcal{U})] \);
2. \( \text{dist}_{C^{1+\frac{\varepsilon}{2}}(\Psi_{x_i}^{-1} \circ \Psi_{x_j}^{-1}) \circ \text{Id}} < \varepsilon \eta_i^2 \eta_j^2 \{ i, j \} = \{ 1, 2 \} \}, \) where the \( C^{1+\frac{\varepsilon}{2}} \) distance is calculated on \( R_{e^{-r(M)}}(\mathcal{U}) \) and \( r(M) \) is defined in (2.4).

**Proof.** Suppose \( \Psi_{x_1}^\gamma \varepsilon \)-overlap, and fix some \( D \in \mathcal{D} \) which contains \( x_1 \) and \( x_2 \) such that \( d(x_1, x_2) + \| \Theta_D \circ C_{\chi}(x_1) - \Theta_D \circ C_{\chi}(x_2) \| < \eta_1^2 \eta_2^2 \). Write \( C_i := \Theta_D \circ C_{\chi}(x_i) \), then \( \Psi_{x_i} = \exp_{x_i} \circ \partial_x \circ C_i \).

By the definition of Pesin charts, \( \eta_i < Q_\varepsilon(x_i) \), where \( Q_\varepsilon(x_i) \) is given by (2.3). Lemma 3.5 and the general inequality \( \| \cdot \|_{F_\varepsilon} \ge \| \cdot \| \) (see page 9) guarantee that

\[
\eta_i < e^{3/\beta} \| C_{\chi}(x_i)^{-1} \|_{12/\beta}.
\]

In particular, \( \eta_i < e^{3/\beta} \).

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

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

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

1. \( L_1 \) is a common Lipschitz constant for the maps \( (x, u) \mapsto (\exp_x \circ \partial_x)(u) \) on \( D \times B_{r(M)}(\mathcal{U}) \) (\( D \in \mathcal{D} \));
2. \( L_2 \) is a common Lipschitz constant for the maps \( x \mapsto \partial_x^{-1} \circ \exp_x \) from \( D \) into \( C^2(D, \mathbb{R}^2) \) (\( D \in \mathcal{D} \));
3. \( L_3 \) is a common Lipschitz constant for \( \exp_x^{-1} : B_{r(M)}(M) \to T_{x,M} \) (\( x \in M \));
4. \( L_4 \) is a common Lipschitz constant for \( \exp_x : B_{r(M)}^2(\mathcal{U}) \to M \) (\( x \in M \)).

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

**Part 1.** \( \Psi_{x_1}[R_{e^{-2\varepsilon}\eta_1}(\mathcal{U})] \subset \Psi_{x_2}[R_{\eta_2}(\mathcal{U})] \).
Proof. Suppose \( v \in R_{e^{-2\eta_1}}(0) \). Lemma 2.5 says that \( C_\chi(x_1) \) is a contraction, therefore \( \|C_\chi(v)\| = \|C_\chi(x_1)v\| \leq \|v\| \), and \( (x_1, C_\chi(x_1), x_2, C_\chi(x_2)) \in D \times B_{r(M)}(0) \).

Since \( d(x_1, x_2) < \eta_1^2 \),

\[
d_{\left(\exp_{x_2} \circ \vartheta_{x_2}(C_\chi v), \exp_{x_1} \circ \vartheta_{x_1}(C_\chi v)\right)} < L_1 \eta_1^2 \eta_2^2.
\]

It follows that \( \Psi_{\chi_1}(v) \in B_{L_2 \eta_1^2 \eta_2^2}(\exp_{x_2} \circ \vartheta_{x_2}(C_\chi v)) \). Call this ball \( B \).

The radius of \( B \) is less than \( \rho(M) \) because of our assumptions on \( \varepsilon \). Therefore \( \exp_{x_2} \) is well defined and Lipschitz on \( B \), and its Lipschitz constant is at most \( L_3 \).

Writing \( B = \exp_{x_2}[\exp_{x_1}(B)] \), we deduce that

\[
\Psi_{\chi_1}(v) \in B \subset \exp_{x_2}[B_{L_2 \eta_1^2 \eta_2^2}(\vartheta_{x_2}(C_\chi v))] =: \Psi_{\chi_2}[E],
\]

where \( E := C_\chi(x_2)^{-1}[B_{L_2 \eta_1^2 \eta_2^2}(\vartheta_{x_2}(C_\chi v))] \).

We claim that \( E \subset R_{\eta_2}(0) \). First note that \( E \subset B_{L_2 \eta_1^2 \eta_2^2}(C_\chi^{-2}C_\chi v) \), therefore if \( w \in E \), then

\[
\|w\|_\infty \leq \|C_\chi^{-1}v\|_\infty + \|C_\chi(x_2)^{-1}L_3 \eta_1 \eta_2^2\|_\infty
\]

\[
\leq \|C_\chi^{-1}C_\chi - \text{Id}\|_\infty + \|w\|_\infty + \|C_\chi(x_2)^{-1}L_3 \eta_1 \eta_2^2\|_\infty
\]

\[
\leq \|w\|_\infty + \sqrt{2}\|C_\chi^{-1}\|_\infty \|C_\chi - C_2\|_\infty + \|C_\chi(x_2)^{-1}L_3 \eta_1 \eta_2^2\|_\infty
\]

\[
\leq e^{-2\eta_1} + \|C_\chi(x_2)^{-1}\|_\infty \|\eta_1 \eta_2^2\|_\infty \sqrt{2}e^{-2\eta_1} + L_3 \eta_1 \eta_2^2\|_\infty
\]

\[
\leq e^{-2\eta_1} + \|C_\chi(x_2)^{-1}\|_\infty \|\eta_1 \eta_2^2\|_\infty \sqrt{2}e^{-2\eta_1} + L_3 L_1 \eta_2^3\|_\infty \eta_1
\]

\[
< e^{-2\eta_1} + e^2 \eta_1,
\]

because of (3.1) and (3.2).

It follows that \( E \subset R_{\eta_2}(0) \). Thus \( \Psi_{\chi_1}(v) = \Psi_{\chi_2}[E] \in R_{\eta_2}(0) \). Part 1 follows.

Part 2. The \( C^{1+\beta/2} \)-distance between \( \Psi_{\chi_1}^{-1} \circ \Psi_{\chi_2} \) on \( R_{e^{-\varepsilon r(M)}}(0) \) is less than \( \varepsilon \eta_1 \).

Proof. One can show exactly as in the proof of part 1 that \( \Psi_{\chi_1}[R_{e^{-\varepsilon r(M)}}(0)] \subset \Psi_{\chi_2}[R_{r(M)}(0)] \), therefore \( \Psi_{\chi_1}^{-1} \circ \Psi_{\chi_2} \) is well defined on \( R_{e^{-\varepsilon r(M)}}(0) \). We calculate the distance of this map from the identity:

\[
\Psi_{\chi_1}^{-1} \circ \Psi_{\chi_2} = C_1^{-1} \circ \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} \circ \exp_{x_2} \circ \vartheta_{x_2} \circ C_2
\]

\[
= C_1^{-1} \circ \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} + \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} \circ \vartheta_{x_2} \circ \exp_{x_2} \circ \vartheta_{x_2} \circ C_2
\]

\[
= C_1^{-1} C_2 + C_1^{-1} \circ \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} \circ \vartheta_{x_2} \circ \exp_{x_2} \circ \vartheta_{x_2} \circ \Psi_{x_2}
\]

\[
= \text{Id} + C_1^{-1} \circ \vartheta_{x_2}^{-1} \circ \exp_{x_2}^{-1} \circ \vartheta_{x_2} \circ \exp_{x_2} \circ \vartheta_{x_2} \circ \Psi_{x_2}.
\]

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

\[
\|C_1^{-1}\| \cdot L_2 d(x_1, x_2) \cdot L_4^{1+\beta/2}.
\]

This is less than \( \|C_1^{-1}\| L_2 L_4^2 \eta_1^4 \eta_2^2 \).

It follows that \( \text{dist}_{C^{1+\beta/2}}(\Psi_{\chi_1}^{-1} \circ \Psi_{\chi_2}, \text{Id}) < \|C_1^{-1}\| \|1 + L_2 L_4^2\| \eta_1 \eta_2^2 \). This is (much) smaller than \( \varepsilon \eta_1 \eta_2^2 \), because of (3.1) and (3.2).

We record the following fact for future reference:

**Lemma 3.3.** Suppose \( \Psi_{\chi_1}^u, \Psi_{\chi_2}^u \) \( \varepsilon \)-overlap, then

\[
\frac{s_\chi(x_1)}{s_\chi(x_2)} \frac{u_\chi(x_1)}{u_\chi(x_2)} \in \left[e^{-Q_\varepsilon(x_1)Q_\varepsilon(x_2)}, e^{Q_\varepsilon(x_1)Q_\varepsilon(x_2)}\right].
\]
Proof. We use the notation of the previous proof. \( \Psi^{-1}_{x_2} \circ \Psi_{x_1} \) maps \( R_{c-\eta_1}(0) \) into \( \mathbb{R}^2 \). Its derivative at the origin is

\[
A := C_2^{-1} \left( d(\exp_{x_2}) C_\chi \right) (x_1) = C_2^{-1} \left( d(\varphi^{-1}_{x_2}) \exp_{x_2} \right) (x_1) \partial x_1 C_1 \\
= C_2^{-1} C_1 + C_2^{-1} \left( d(\varphi^{-1}_{x_2}) \exp_{x_2} \right) (x_1) \partial x_1 C_1 \\
\equiv C_2^{-1} C_1 + C_2^{-1} \left( d(\varphi^{-1}_{x_2}) \exp_{x_2} \right) (x_1) \partial x_1 C_1.
\]

Since \( \| d(\varphi^{-1}_{x_2}) \exp_{x_2} \| \| x_1 - d(\varphi^{-1}_{x_2}) \exp_{x_2} \| \| x_1 \| < L_2 \| x_1 \| < L_2 \eta_1^4 \eta_2^4 < \varepsilon \eta_1^2 \eta_2^2 \),

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

It follows that \( \| C_2 - C_1 \| < 2 \varepsilon \| C_2^{-1} \| \eta_1^2 \eta_2^2 \).

Recall that \( s_\chi(x_i)^{-1} = \| C_\chi(x_i) \chi_1 \| \) and \( s_\chi(x_i) = \| C_\chi(x_i)^{-1} \chi_1(x_i) \| \), so

\[
\left| \frac{s_\chi(x_1)^{-1}}{s_\chi(x_2)^{-1}} - 1 \right| = \left| \frac{s_\chi(x_2)^{-1} - s_\chi(x_1)^{-1}}{s_\chi(x_1)^{-1}} \right| \\
\leq \| C_\chi(x_1)^{-1} \| \cdot \| C_\chi(x_1) \chi_1 \| - \| C_\chi(x_2) \chi_1 \| \\
= \| C_\chi^{-1} \| \cdot \| C_\chi \chi_1 \| - \| C_\chi \chi_1 \| \\
\leq \| C_\chi^{-1} \| \cdot \| C_\chi - C_\chi \| < 2 \varepsilon \| C_\chi^{-1} \| \| C_\chi \| \eta_1^2 \eta_2^2 < \varepsilon \eta_1 \eta_2.
\]

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

\[ \square \]

3.2. The form of \( f \) in overlapping charts. Theorem 2.7 says that \( \Psi^{-1}_{f(x)} \circ f \circ \Psi_x \) 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 \( \Psi_{f(x)} \varepsilon \)-overlaps \( \Psi_y \), then \( f_{xy} := \Psi^{-1}_{f(x)} \circ f \circ \Psi_x \) is a well defined injective map from \( R_{10Q_\varepsilon(x)}(0) \) 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 \( C_f^{-1} < |A| < e^{-x}, e^x < |B| < C_f \) (cf. Theorem 2.3), \( |h_i(0)| < \varepsilon \eta_i \), \( \| \nabla h_i(0) \| < \varepsilon \eta_i^{\beta/3} \), and \( \| \nabla h_i(0) - \nabla h_i(\varphi) \| < \varepsilon \| \varphi \|^{\beta/3} \) on \( R_{10Q_\varepsilon(x)}(0) \).

A similar statement holds for \( f_{x^{-1}} \), assuming that \( \Psi^{-1}_{f^{-1}(y)} \varepsilon \)-overlaps \( \Psi_y^{-1} \).

**Proof.** We write \( f_{xy} = (\Psi^{-1}_{f(x)} \circ f \circ \Psi_x) \circ f_x \) where \( f_x = \Psi^{-1}_{f(x)} \circ f \circ \Psi_x \), 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 \( R_{10Q_\varepsilon(x)}(0) \).
2. \( f_x(0) = 0 \) and \( (df_x)_0 = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \) where \( C_f^{-1} < |A| < e^{-x}, e^x < |B| < C_f \).
3. For all \( u, v \in R_{10Q_\varepsilon(x)}(0) \), \( \| (df_x)_0 \| - \| (df_x)_0 \| \leq \varepsilon \| u - v \|^{\beta/2} \) (because the \( C^{1+\beta} \) distance between \( f_x \) and \( (df_x)_0 \) on \( R_{10Q_\varepsilon(x)}(0) \) is less than \( \varepsilon \)).
4. For every \( 0 < \eta < 10Q_\varepsilon(x) \) and \( \varphi \in R_{\varepsilon}(0) \), \( \| (df_x)_0 \| < 3C_f \), provided \( \varepsilon \) is small enough (because \( \| (df_x)_0 \| \leq \| (df_x)_0 \| + \varepsilon \eta^{\beta/2} < 2C_f + \varepsilon \)).
(2) and (4) imply that $f_{x}[R_{10Q}(x)(\mathbb{U})] \subset B_{30Q_{x}}(e_{C_{1}}(\mathbb{U})].$ Since $Q_{\varepsilon}(x) < \varepsilon^{3/\beta},$ $f_{x}[R_{10Q}(x)(\mathbb{U})] \subset B_{30C_{x}}(e_{C_{1}}(\mathbb{U})].$ If $\varepsilon$ is small enough that $30C_{x} \varepsilon^{3/\beta} < e^{-\varepsilon r}(M),$ then $f_{x}[R_{10Q}(x)(\mathbb{U})] \subset R_{e^{-\varepsilon r}(M)}(\mathbb{U}).$ $R_{e^{-\varepsilon r}(M)}(\mathbb{U})$ 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 $R_{10Q}(x)(\mathbb{U}).$

Equation (3.3) can be used to define the functions $h_{x}(u,v).$ We check that the resulting functions satisfy the properties proclaimed by the proposition.

We have $(h_{1}(\mathbb{U}),h_{2}(\mathbb{U})) = f_{xy}(\mathbb{U}) = \Psi_{y}^{-1}(f(x)) = (\Psi_{y}^{-1} \circ \Psi_{f(x)}(\mathbb{U}),$ therefore $||(h_{1}(\mathbb{U}),h_{2}(\mathbb{U}))|| \leq \text{dist}_{C_{0}}(\Psi_{y}^{-1} \circ \Psi_{f(x)},\text{Id}) < e\eta^{2}(\eta)^{2} < e\eta.$

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

\begin{equation}
(df_{xy})_{\mathbb{U}} = [d(\Psi_{y}^{-1} \circ \Psi_{f(x)} \circ f_{x})_{\mathbb{U}} - \text{Id}] \cdot (df_{x})_{\mathbb{U}} + [(df_{x})_{\mathbb{U}} - (df_{x})_{\mathbb{U}}] + (df_{x})_{\mathbb{U}}.
\end{equation}

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 $\varepsilon \cdot (e^{2/\beta})^{2} < 2\varepsilon\eta^{2/\beta}.$ The third term is $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$ Thus

\begin{align*}
\left\| \frac{\partial(h_{1},h_{2})}{\partial(u,v)} \right\| &= \left\| (df_{xy})_{\mathbb{U}} - \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \right\| < e[3C_{f} + 2]\eta^{2/\beta} \\
&< e\eta^{2/\beta} \cdot [3C_{f} + 2]\eta^{2/\beta} < e\eta^{2/\beta} \cdot [3C_{f} + 2]\eta^{2/\beta} \text{ by (3.3).}
\end{align*}

If $\varepsilon$ is so small that $[3C_{f} + 2]\eta^{2/\beta} < 1,$ then $||\nabla h_{i}|| < e\eta^{2/\beta}$ on $R_{\eta}(\mathbb{U}).$ In particular, $||\nabla h_{i}(\mathbb{U})|| < e\eta^{2/\beta}.$

Equation (3.2) also shows that for every $\mathbb{U},\mathbb{V} \in R_{10Q}(x)(\mathbb{U}),$

\begin{align*}
||(df_{xy})_{\mathbb{U}} - (df_{xy})_{\mathbb{V}}|| &= ||d(\Psi_{y}^{-1} \circ \Psi_{f(x)} \circ f_{x})_{\mathbb{U}} - d(\Psi_{y}^{-1} \circ \Psi_{f(x)} \circ f_{x})_{\mathbb{V}}|| \\
&+ \||d(\Psi_{y}^{-1} \circ \Psi_{f(x)} \circ f_{x})_{\mathbb{U}} - (df_{xy})_{\mathbb{U}} + (df_{xy})_{\mathbb{U}}|| + \||d(\Psi_{y}^{-1} \circ \Psi_{f(x)} \circ f_{x})_{\mathbb{V}} - (df_{xy})_{\mathbb{V}} + (df_{xy})_{\mathbb{V}}||
\end{align*}

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

\begin{align*}
||(df_{xy})_{\mathbb{U}} - (df_{xy})_{\mathbb{V}}|| &\leq e\eta^{2}(\eta)^{2} \cdot ||f_{x}(\mathbb{U}) - f_{x}(\mathbb{V})||^{2} \cdot 3C_{f} + e||\mathbb{U} - \mathbb{V}||^{2} (e\eta^{2}(\eta)^{2} + 2) \\
&\leq e\eta^{2} \cdot \sup_{\mathbb{W} \in R_{10Q}(x)(\mathbb{U})} ||(df_{xy})_{\mathbb{W}}||^{2} \cdot ||\mathbb{U} - \mathbb{W}||^{2} \cdot 3C_{f} + 3e||\mathbb{U} - \mathbb{V}||^{2} \\
&\leq e((3C_{f})^{1+4\beta/3} + 3)||\mathbb{U} - \mathbb{V}||^{2} \leq \varepsilon((3C_{f})^{1+4\beta/3} + 3)||\mathbb{U} - \mathbb{V}||^{2} \\
&\leq 4\varepsilon||\mathbb{U} - \mathbb{V}||^{2}, \text{ provided } \varepsilon \text{ is small enough}
\end{align*}

\begin{align*}
&\leq 3e(30Q_{\varepsilon}(x))^{1/6}||\mathbb{U} - \mathbb{V}||^{2} < 6e^{3/2}||\mathbb{U} - \mathbb{V}||^{2/3} (\because Q_{\varepsilon} < e^{3/2}) \\
&\leq \frac{1}{3}e||\mathbb{U} - \mathbb{V}||^{2/3}, \text{ provided } \varepsilon \text{ is small enough}
\end{align*}

It follows that $\frac{\partial_{x}h_{i}}{\partial_{x}u}(\mathbb{U}) - \frac{\partial_{x}h_{i}}{\partial_{x}u}(\mathbb{V}) < \frac{1}{3}e||\mathbb{U} - \mathbb{V}||^{2/3}$ for all $\mathbb{U},\mathbb{V} \in R_{10Q}(x)(\mathbb{U}),$ whence $||\nabla h_{i}(\mathbb{U}) - \nabla h_{i}(\mathbb{V})|| \leq \frac{1}{3}e||\mathbb{U} - \mathbb{V}||^{2/3}$ for all $\mathbb{U},\mathbb{V} \in R_{10Q}(x)(\mathbb{U}).$

3.3. Coarse graining. We replace $\mathcal{C} := \{\Psi_{x}^{2} : x \in \text{NUH}_{x}(f), 0 < \eta \leq Q_{x}(x)\}$ by a “sufficient” countable subset $\mathcal{A}.$ We remind the reader that $\text{NUH}_{x}$ is defined in Lemma 2.6 and that $I_{\varepsilon} = \left\{ e^{-\frac{1}{4}ke} : k \in \mathbb{N} \right\}.$
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. Sufficiency: For every $x \in \text{NUH}^\ast(f)$ and for every sequence of positive numbers $0 < \eta_n \leq e^{-\varepsilon/3}Q_\varepsilon(f^n(x))$ in $I_x$ s.t. $e^{-\varepsilon} \leq \eta_n/\eta_{n+1} \leq e^{\varepsilon}$, there exists a sequence $\{\Psi^{0_n}_{x_n}\}_{n \in \mathbb{Z}}$ of elements of $\mathcal{A}$ s.t. for every $n$,
   (a) $\Psi^{0_n}_{x_n} \varepsilon$-overlaps $\Psi^{0_n}_{f^n(x)}$ and $e^{-\varepsilon/3} \leq Q_\varepsilon(f^n(x))/Q_\varepsilon(x_n) \leq e^{\varepsilon/3}$;
   (b) $\Psi^{0_{n+1}}_{x_n} \varepsilon$-overlaps $\Psi^{0_{n+1}}_{f^{n+1}(x)}$;
   (c) $\Psi^{0_{n-1}}_{x_n} \varepsilon$-overlaps $\Psi^{0_{n-1}}_{f^{-1}(x)}$;
   (d) $\Psi^{0_n}_{x_n} \in \mathcal{A}$ for all $\eta_n \in I_x$ s.t. $\eta_n \leq \eta'_n \leq \min\{Q_\varepsilon(x_n), e^\varepsilon\eta_n\}$.

2. Discreteness: $\{\Psi^1_n \in \mathcal{A} : \eta > t\}$ is finite for every $t > 0$.

Proof. The proposition would have been easy had $C_\varepsilon(x)$ been a continuous function of $x$. In general it is not, and as a result there is no clear connection between conditions (a), (b), and (c). We must treat the three conditions separately, and simultaneously.

The following construction will help us to do this. Let

$$X := \bigcup_{D_0, D_1, D_{-1} \in \mathcal{D}} D_0 \times D_1 \times D_{-1} \times \mathbb{R}^3 \times \text{GL}(2, \mathbb{R})^3.$$ 

Here $\mathcal{D}$ is the finite open cover of $M$ which we constructed in §3.1. $X$ is a subset of $M^3 \times \mathbb{R}^3 \times \text{GL}(2, \mathbb{R})^3$. We equip it with the relative product topology.

Let $Y \subset X$ denote the collection of all $(x, Q, C) \in X$ where

- $x = (x, f(x), f^{-1}(x))$, $x \in \text{NUH}^\ast(f)$;
- $Q = (Q_0(x), Q_1(f(x)), Q_1(f^{-1}(x)))$ (cf. (2.3));
- $C = (\Theta_{D_0} \circ C_\varepsilon(x), \Theta_{D_1} \circ C_\varepsilon(f(x)), \Theta_{D_{-1}} \circ C_\varepsilon(f^{-1}(x)))$, 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 $Y_k := \{(x, Q, C) \in Y : x \in \text{NUH}^\ast(f), e^{-(k+1)} \leq Q_\varepsilon(x) \leq e^{-(k-1)}\}$ ($k \in \mathbb{N}$). $Y_k$ is a pre-compact subset of $X$. To see this, pick some $(x, Q, C) \in Y_k$. The vector $x$ belongs to the compact set $M^3$. $Q$ belongs to a compact subset of $\mathbb{R}^3$ because by Lemma 2.9 for each $i = -1, 0, 1$,

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

$C$ belongs to a compact subset of $\text{GL}(2, \mathbb{R})$, because (a) $\Theta_{D_i}$ are isometries; (b) $\|C_\varepsilon(f^i(x))\| < 1$ (Lemma 2.9); and (c) $\|C_\varepsilon(f^i(x))^{-1}\| \leq (e^{3/3}Fe^{k+1})^{3/12}$ by (2.3). It follows that $Y_k$ is a subset of a compact subset of $M^3 \times \mathbb{R}^3 \times \text{GL}(2, \mathbb{R})^3$.

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

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

5Here 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 $C > 0$. 
Define $\mathcal{A}$ to be the collection of all Pesin charts $\Psi^g$ such that for some $k, m \in \mathbb{N}$, $x$ is the first coordinate of some element $(x, Q, C) \in Y_{k,m}$, and

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

Part 1. Discreteness.

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

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

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

Part 2. Sufficiency.

Proof. Suppose $x \in \text{NUH}^\ast(f)$, and $\eta_n \in I_\epsilon$ satisfy $0 < \eta_n \leq e^{-\epsilon/3}Q_\epsilon(f^n(x))$ and $e^{-\epsilon} \leq \eta_n / \eta_n+1 \leq e^\epsilon$ 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 $Q_\epsilon(f^n(x)) \in [e^{-k_n-1}, e^{-k_n+1}]$. 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 $i = -1, 0, 1$,

\begin{align*}
(A_n) & \ d(f^i(f^n(x)), f^i(x_n)) < \frac{1}{2} \epsilon(\mathcal{D}); \\
(B_n) & \ d(f^i(f^n(x)), f^i(x_n)) + \| \Theta_D \circ C_X(f^i(f^n(x))) - \Theta_D \circ C_X(f^i(x_n)) \| < e^{-8(m_n+2)} \\
(C_n) & \ e^{-\epsilon/3} < Q_\epsilon(f^i(f^n(x)))/Q_\epsilon(f^i(x_n)) < e^{\epsilon/3}.
\end{align*}

Claim 1. $\Psi^g_{x_n} \in \mathcal{A}$ and $\Psi^g_{f^n(x)} \in \mathcal{A}$ for all $\eta' \in I_\epsilon$ s.t. $\eta_n \leq \eta_n' \leq \min\{e^\epsilon \eta_n, Q_\epsilon(x_n)\}$.

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 $\eta_n \leq \eta_n' \leq e^\epsilon \eta_n$, $\eta_n' \in [e^{-m_n-2}, e^{m_n+2}]$. It remains to check that $\eta_n, \eta_n' \leq Q_\epsilon(x_n)$. In case of $\eta_n$, (C_n) with $i = 0$ says that $Q_\epsilon(x_n) > e^{-\epsilon/3}Q_\epsilon(f^n(x)) \geq \eta_n$.

Claim 2. $\Psi^g_{x_n}$ and $\Psi^g_{f^n(x)}$ $\epsilon$-overlap.

Proof. (A_n) with $i = 0$ says that $d(f^n(x), x_n)$ is smaller than the Lebesgue number of $\mathcal{D}$, so there exists some $D \in \mathcal{D}$ s.t. $f^n(x), x_n \in D$. (B_n) with $i = 0$ says that

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

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

Claim 3. $\Psi^g_{f^n(x)}$ $\epsilon$-overlaps $\Psi^g_{f^{n+1}(x)}$ 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 $d(f(x_n), f(f^n(x))) < \frac{1}{2} \epsilon(\mathcal{D})$. Setting $i = 0$ in (A_{n+1}), we see that $d(f^{n+1}(x), x_{n+1}) < \frac{1}{2} \epsilon(\mathcal{D})$. 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 $i = 0$,

$$d(f(x_n), x_{n+1}) + \| \Theta_D \circ C_X(f(x_n)) - \Theta_D \circ C_X(x_{n+1}) \| \leq$$
\[
\left( d(f(x_n), f(f^n(x))) + \left\| \Theta_D \circ C_x(f(x_n)) - \Theta_D \circ C_x(f(f^n(x))) \right\| \right) + \\
+ \left( d(f^{n+1}(x), x_{n+1}) + \left\| \Theta_D \circ C_x(f^{n+1}(x)) - \Theta_D \circ C_x(x_{n+1}) \right\| \right) \\
\leq e^{-8(m_n+2)} + e^{-8(m_n+2)} \\
< e^{-8}(\eta^8_n + \eta^8_{n+1}) < 2e^{-8}(1 + e^{8\varepsilon})\eta^4_{n+1}\eta^4_{n+1} \leq \eta^4_{n+1}\eta^4_{n+1}.
\]

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

4. \(\varepsilon\)-chains and an infinite-to-one Markov extension of \(f\)

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

**Definition 4.1.** $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$ means

- $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$ \(\varepsilon\)-overlaps (recall that $a \land b := \min\{a, b\}$);
- $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$ \(\varepsilon\)-overlaps;
- $q^u = \min\{e^\varepsilon p^u, Q_x(y)\}$ and $p^\sigma = \min\{e^\varepsilon q^s, Q_x(x)\}$.

**Definition 4.2.** $\{\Psi^\eta_x \Psi^\sigma_y \}_{i \in \mathbb{Z}}$ (resp. $\{\Psi^\eta_x \Psi^\sigma_y \}_{i \geq 0}$, $\{\Psi^\eta_x \Psi^\sigma_y \}_{i \leq 0}$) is called an \(\varepsilon\)-chain (resp. positive \(\varepsilon\)-chain, negative \(\varepsilon\)-chain), if $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_{x+i} \Psi^\sigma_{y+i}$ 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 [33] and recall that $L_\varepsilon = \{e^{-k\varepsilon/3} : k \in \mathbb{N}\}$.

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

- $\mathcal{Y} := \{\Psi^\eta_x \Psi^\sigma_y : \Psi^\eta_x \Psi^\sigma_y \in \mathcal{A}, p^u, p^\sigma \in L_\varepsilon, p^u, p^\sigma \leq Q_x(x)\}$;
- $\mathcal{E} := \{(\Psi^\eta_x \Psi^\sigma_y, \Psi^\eta_y \Psi^\sigma_x) : \Psi^\eta_y \Psi^\sigma_x \in \mathcal{Y}, \Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x \}$.

This is a countable directed graph. Every vertex has finite degree, because of the following lemma, and Proposition 3.3.2:

**Lemma 4.4.** If $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$, then $e^{-\varepsilon} \leq (q^u \land q^s)/(p^u \land p^\sigma) \leq e^\varepsilon$. Therefore for every $\Psi^\eta_x \Psi^\sigma_y \in \mathcal{Y}$ there are only finitely many $\Psi^\eta_y \Psi^\sigma_x \in \mathcal{Y}$ s.t. $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$ or $\Psi^\eta_x \Psi^\sigma_y \to \Psi^\eta_y \Psi^\sigma_x$.

**Proof.** The proof is a manipulation of the following relations:

\[
q^u = \min\{e^\varepsilon p^u, Q_x(y)\}, \quad p^\sigma \leq Q_x(x); \\
p^\sigma = \min\{e^\varepsilon q^s, Q_x(x)\}, \quad q^s \leq Q_x(y). \tag{4.1}
\]

Let $p := p^u \land p^s$ and $q := q^u \land q^s$. We show that $e^{-\varepsilon} \leq p/q \leq e^\varepsilon$ by considering each of the following cases separately:

1. $p = p^u$, $q = q^u$.
2. $p = p^s$, $q = q^s$.
3. $p = p^u$, $q = q^s$.
4. $p = p^s$, $q = q^u$.

**Case 1.** If $e^\varepsilon p^u \leq Q_x(y)$, then $q^u = \min\{e^\varepsilon p^u, Q_x(y)\} = e^\varepsilon p^u$, and $\frac{p}{q} = \frac{p^u}{q^u} = e^\varepsilon$. If $e^\varepsilon p^u > Q_x(y)$, then $p \leq p^u \leq e^\varepsilon q^u \leq e^\varepsilon Q_x(y) = e^\varepsilon \min\{e^\varepsilon p^u, Q_x(y)\} = e^\varepsilon q^u = e^\varepsilon q$, so $\frac{p}{q} \leq e^\varepsilon$. Also, $q = q^u = \min\{Q_x(y), e^\varepsilon p^u\} \leq e^\varepsilon p^u = e^\varepsilon p$, so $\frac{q}{p} \leq e^\varepsilon$.  

SYMBOLIC DYNAMICS FOR SURFACE Diffeomorphisms 17
Case 2. This is the same as case 1.

Case 3. In this case $p^u \leq p^s$, so $p = p^u \leq p^s \leq e^x q^s = e^x q$, whence $p/q \leq e^x$. Also, $q^s \leq q^u$, so $q = q^u \leq q^s \leq e^x p^s = e^x p$, whence $q/p \leq e^x$.

Case 4. In this case $p^s \leq p^u$ and $q^u \leq q^s$. Since $p^u = \min\{e^x q^u, Q_e(x)\}$, either $p^s = e^x q^u$, or $p^s = Q_e(x)$.

Suppose $p^s = e^x q^u$, then $q^u \leq q^s = e^{-x} p^u \leq e^{-x} e^x p^s = e^x p^u$. The inequality $q^u < e^x p^u$ and the identity $q^u = \min\{e^x p^u, Q_e(y)\}$ force $q^u = Q_e(y)$. But $q^u \leq q^s = Q_e(y)$, so $q^u = q^s = Q_e(y)$. It follows that $p = p^s = e^x q^u = e^x q^s = e^x q$, and we are done.

Next suppose that $p^s = Q_e(x)$. Since $p^s \leq p^u \leq Q_e(x)$, we must have

$$p = p^s = p^u = Q_e(x).$$

At the same time, $q^u = \min\{e^x p^u, Q_e(y)\} \leq e^x p^u = e^x p$. If there is an equality, then we are done. Otherwise $q^u < e^x p$, and since $q^u \leq q^s \leq Q_e(y)$,

$$q = q^u = q^s = Q_e(y).$$

Since $\min\{e^x p^u, Q_e(y)\} = q^u = Q_e(y)$, $e^x p^u \geq Q_e(y)$. Thus $e^x Q_e(x) \geq Q_e(y)$. Similarly, $\min\{e^x q^u, Q_e(x)\} = p^s = Q_e(x)$ implies that $e^x q^s \geq Q_e(x)$, whence $e^x Q_e(x) \geq Q_e(y)$. It follows that $p/q = Q_e(x)/Q_e(y) \in [e^{-x}, e^x]$. □

We claim that the collection of infinite admissible paths on $\mathcal{G}$ is as rich as the set of orbits of $f$ in NUH$^\#(f)$. Recall that NUH$^\#(f)$ has full measure w.r.t. every $f$–ergodic invariant probability measure with entropy greater than $\chi$.

**Proposition 4.5.** For every $x \in \text{NUH}^\#(f)$, 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 $I_e = \{e^{-\ell\varepsilon}/3 : \ell \in \mathbb{N}\}$. 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 $k \in \mathbb{Z}$, $0 < p_k^u, p_k^s \leq Q_k$, $p_k^u, p_k^s \in I_e$, and

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

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 $\{Q_e(x_k)\}_{k \in \mathbb{Z}}$.

**Lemma 4.6.** Let $(Q_k)_{k \in \mathbb{Z}}$ be a sequence in $I_e$, and suppose $q_k \in I_e$ satisfy $0 < q_k \leq Q_k$ and $e^{-x} \leq q_k/q_{k+1} \leq e^x$ 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 $q_k, Q_e(x_{k-n}), Q_e(x_{k+n}) \geq e^{-\varepsilon} q_k$ for all $n \geq 0$, therefore the following definitions make sense:

$$p_k^u := \max\{t \in I_e : e^{-\varepsilon n} t \leq Q_e(x_{k-n}) \text{ for all } n \geq 0\};$$

$$p_k^s := \max\{t \in I_e : e^{-\varepsilon n} t \leq Q_e(x_{k+n}) \text{ for all } n \geq 0\}.$$ 

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

**Lemma 4.7.** Suppose $\{(p_n^u, p_n^s)\}_{n \in \mathbb{Z}}$ is $\varepsilon$–subordinated to a sequence $\{Q_n\}_{n \in \mathbb{Z}} \subset I_e$. If $\limsup\{p_n^u \wedge p_n^s\} > 0$, 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. Let $M := \sup_n Q_n$ be finite, because $Q_n \in I_s$ for all $n$. Let $p_n := p_n^u \land p_n^s$, and define $m := \frac{1}{2} \min \{ \limsup_{n \to \infty} p_n, \limsup_{n \to \infty} p_n \}$. We prove the statement for $p_n^u$. Otherwise, by $\varepsilon$-subordination,

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

which is false.

We can now prove Proposition 4.3. Suppose $x \in \text{NUH}^\#(f)$, and recall the definition of $q_n(x)$ from Lemma 2.9. Choose $q_n \in I_s \cap [e^{-\varepsilon/3} q_n(f^n(x)), e^{\varepsilon/3} q_n(f^n(x))]$. 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 $\{ e^{-\varepsilon/3} Q_n(f^n(x)) \}_{n \in \mathbb{Z}}$ and that satisfies $q_n^u \land q_n^s \geq q_n$.

Let $\eta_n := q_n^u \land q_n^s$. By Lemma 4.1, $e^{-\varepsilon} \leq \eta_{n+1}/\eta_n \leq e^{\varepsilon}$, so we are free to use Proposition 4.5 to construct an infinite sequence $\Psi_{x_n}^{p_n^u}$ such that

(a) $\Psi_{x_n}^{p_n^u}$ $\varepsilon$-overlaps $\Psi_{x_n}^{p_n^u(f^n(x))}$ and $e^{-\varepsilon/3} \leq Q_n(f^n(x))/Q_n(x_n) \leq e^{\varepsilon/3}$;
(b) $\Psi_{x_n}^{p_{n+1}^u}$ $\varepsilon$-overlaps $\Psi_{x_n}^{p_{n+1}^u(f^n(x))}$;
(c) $\Psi_{x_n}^{p_{n-1}^u}$ $\varepsilon$-overlaps $\Psi_{x_n}^{p_{n-1}^u(f^n(x))}$;
(d) $\Psi_{x_n}^{p_n^u} \in \mathcal{A}$ for all $\eta_n \in I_s$ s.t. $\eta_n \leq \eta_n \leq \min \{ Q_n(x_n), e^\varepsilon \eta_n \}$.

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

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

Proof. It is sufficient to show that $1 \leq \frac{p_n^u \land p_n^s}{p_n^u \land p_n^s} \leq e^\varepsilon$ (n \in \mathbb{Z})$, because property (d) with $\eta_n := p_n^u \land p_n^s$ says that in this case $\Psi_{x_n}^{p_n^u} \land p_n^s \in \mathcal{A}$, whence $\Psi_{x_n}^{p_n^u} \land p_n^s \in \mathcal{V}$. We start by showing that there are infinitely many $n < 0$ such that $p_n^u \leq e^\varepsilon q_n^u$. Since $x \in \text{NUH}^\#(f)$, $\limsup_{n \to \infty} q_n = \limsup_{n \to \infty} q_n > 0$. Therefore by Lemma 4.7 there are infinitely many $n < 0$ for which $q_n^u = e^{-\varepsilon/3} Q_n(f^n(x))$. Property (a) guarantees that for such $n$, $q_n^u \geq e^{-\varepsilon} Q_n(x_n)$, and that $p_n^u \leq e^\varepsilon q_n^u$.

If $p_n^u \leq e^\varepsilon q_n^u$, then $p_{n+1}^u = \min \{ e^\varepsilon p_n^u, Q_n(x_{n+1}) \} = \varepsilon^\varepsilon \min \{ p_n^u, e^{-\varepsilon} Q_n(x_{n+1}) \}$

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

Claim 2. For every $n \in \mathbb{Z}$, $\Psi_{x_n}^{p_n^u} \land p_n^s \to \Psi_{x_{n+1}}^{p_{n+1}^u \land p_{n+1}^s}$, and $\Psi_{x_n}^{p_n^u} \land p_n^s \varepsilon$-overlaps $\Psi_{x_n}^{p_n^u \land p_n^s}$.

Proof. This follows from properties (a), (b), and (c) above, the inequality $p_n^u \land p_n^s \geq \eta_n$, and the monotonicity property of the overlap condition.
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
\[
\begin{align*}
\sigma &- \text{parameter: } \sigma(V^{u/s}) := \| F' \|_{\beta/3} := \| F' \|_\infty + \sup \left\{ \frac{|F'(t_1) - F'(t_2)|}{|t_1 - t_2|^{\beta/3}} \right\}; \\
\gamma &- \text{parameter: } \gamma(V^{u/s}) := |F'(0)|; \\
\varphi &- \text{parameter: } \varphi(V^{u/s}) := |F(0)|; \\
q &- \text{parameter: } q(V^{u/s}) := q.
\end{align*}
\]

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 \( \Psi_x^{u/p} \) is a double chart. A \((u/s, \sigma, \gamma, \varphi, q)\)-manifold in \( \Psi_x \) s.t.
\[
\sigma \leq 1/2, \quad \gamma \leq 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 & \text{\( u \)-manifolds} \\ p^s & \text{\( s \)-manifolds}. \end{cases}
\]

This is similar, but stronger, than the admissibility condition in Katok & Mendoza [KM, Definition S.3.4] or Katok [K1]. We needed to strengthen the condition to get Proposition [LI] (4) below.

Let \( F \) be the representing function of a \( u/s \)-admissible manifold in \( \Psi_x^{u/p} \). If \( \varepsilon < 1 \) (as we always assume), then the conditions \( \sigma \leq 1/2 \), \( \varphi < 10^{-3}(p^u \wedge p^s) \) and \( p^u, p^s < Q_\varepsilon(x) \) force
\[
\text{Lip}(F) < \varepsilon, \quad (4.2)
\]
because for every \( t \) in the domain of \( F \), \(|t| \leq p^{u/s} \leq Q_\varepsilon(x) < \varepsilon^{3/\beta} \) and
\[
|F'(t)| \leq |F'(0)| + \text{H"{o}l}(F')|t|^{\frac{\beta}{3}} \leq \frac{1}{2}(p^u \wedge p^s)^{\frac{\beta}{3}} + \frac{1}{2}(p^{u/s})^{\frac{\beta}{3}} < (p^{u/s})^{\frac{\beta}{3}} < \varepsilon. \quad (4.3)
\]

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 dist and \( \text{dist}_{C^1} \) are defined using the Pesin charts, not its “natural” charts. Distances using natural charts 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 \( C_\chi(x) : \mathbb{R}^2 \to M \) is a contraction.
**Definition 4.10.** Let $V^s, V^u$ be a $u$–manifold and an $s$–manifold in $Ψ_x$, with representing functions $F_x, F_u$. Suppose $V^s, V^u$ intersect at a unique point $P = Ψ_x(u, v)$, then $\angle(V^s, V^u) := \angle((dΨ_x)_x(v, w)(F'_x(u)), (dΨ_x)_x(v, w)(F'_x(v)))$.

**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 $Ψ_x^{p^u, p^s}$, and $V^s$ be an $s$–admissible manifold in $Ψ_x^{p^u, p^s}$, then

1. $V^u$ intersects $V^s$ at a unique point $P$;
2. $P = Ψ_x(v, w)$ 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
   $$e^{-\eta} \leq \frac{\sin \angle(V^u, V^s)}{\sin \angle(E^u(x), E^u(x))} \leq e^{\eta}$$
   $$|\cos \angle(V^u, V^s) - \cos \angle(E^u(x), E^u(x))| < 2\eta^{3/4}.$$

Parts (1), (2), and (3) follow from [KH] Corollary S.3.8. Part (3) is because of the assumptions on $\gamma$ and $\sigma$, and is the reason why we require more 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 proof is in the appendix.

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

1. $f(V^u)$ contains a $u$–manifold $V^u$ in $Ψ_y^{q^u, q^s}$ with parameters
   $$\sigma(V^u) \leq e^{\sqrt{\varepsilon}} e^{-2x|\sigma(V^u) + \varepsilon|}$$
   $$\gamma(V^u) \leq e^{\sqrt{\varepsilon}} e^{-2x[\gamma(V^u) + \varepsilon^{3/2}(p^u \wedge p^s)^{3/2}]}$$
   $$\varphi(V^u) \leq e^{\sqrt{\varepsilon}} e^{-x[\varphi + \varepsilon(q^u \wedge q^s)]}$$
   $$q(V^u) \geq \min\{e^{-\sqrt{\varepsilon}} e^{\varepsilon q(V^u), Q_e(y)}\}\}
   $$\varphi(V^u) \leq e^{\sqrt{\varepsilon}} e^{-x[\gamma(V^u) + \varepsilon^{3/2}(p^u \wedge p^s)^{3/2}]}$$

2. $f(V^u)$ intersects any $s$–admissible manifold in $Ψ_y^{q^u, q^s}$ at a unique point.
3. $V^u$ restricts to a $u$–admissible manifold in $Ψ_y^{q^u, q^s}$. This is the unique $u$–admissible manifold in $Ψ_y^{q^u, q^s}$ inside $f(V^u)$. We call it $F_u[V^u]$.
4. Suppose $V^u$ is represented by the function $F$. If $p := Ψ(x(F(0), 0)$, then $f(p) ∈ F_u[V^u]$.

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

**Definition 4.13.** Suppose $Ψ_x^{p^u, p^s} → Ψ_y^{q^u, q^s}$. The graph transforms are the maps

- $F_u$ which maps a $u$–admissible manifold $V^u$ in $Ψ_x^{p^u, p^s}$ to the unique $u$–admissible manifold in $Ψ_y^{q^u, q^s}$ contained in $f(V^u)$;
- $F_s$ which maps an $s$–admissible manifold $V^s$ in $Ψ_x^{p^u, p^s}$ to the unique $s$–admissible manifold in $Ψ_x^{p^u, p^s}$ contained in $f^{-1}(V^s)$.

(The operators $F_s, F_u$ depend on the edge $Ψ_x^{p^u, p^s} → Ψ_y^{q^u, q^s}$.)
We equip $\Sigma$ with the metric $d$ distribution on Pesin sets $(5)$ should be compared to Brin’s Theorem on the Hölder continuity of the Oseledets u/s Parts (1)–(4) should be compared with Pesin’s Stable Manifold Theorem $v$ in $\mathbb{R}$.

Suppose $G$ of two sided infinite paths on the graph $\delta$:

In fact, the map we construct will be well-defined for all chains.

Our aim is to construct a map $\pi$ defined on all chains. Suppose $(v_i)_{i \in \mathbb{Z}}$, $v_i = \Psi_{x_i}^s$, $\Psi_{x_i}^s$ is a chain, and let $V^\mu_n$ be a $u$–admissible manifold in $v_n$. The graph transform relative to $v_n \to v_{n+1}$ maps $V^\mu_n$ to a $u$–admissible manifold in $v_{n+1}$, $F_{v_n}$. Another application of the graph transform, this time relative to $v_{n+1} \to v_{n+2}$, maps $F_{v_n}$ to a $u$–admissible manifold in $v_{n+2}$, which we denote by $F_{v_n}^u$. Continuing this way, we eventually reach a $u$–admissible manifold in $v_0$ which we denote by $F_{v_0}$.

Similarly, any $s$–admissible manifold in $v_n$ is mapped by $n$ applications of $F_{v_n}$ to an $s$–admissible manifold in $v_0$. The manifolds $F_{v_n}^u$ and $F_{v_n}$ depend on $(v_n, \ldots, v_0)$.

Let $V_n$ denote a sequence of $u$–s–manifolds in a chart $\Psi_{x_i}$. 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 $V$. Compare with definition $\llbracket 4 \rrbracket$.

**Proposition 4.14.** Suppose $(v_i)_{i \in \mathbb{Z}}$ is a chain of double charts, and choose arbitrary $u$–admissible manifolds $V^\mu_n$ in $v_n$, and $s$–admissible manifolds $V^s_n$ in $v_n$.

1. The limits $V^\mu_n((v_i)_{i \leq 0}) := \lim_{n \to \infty} F_{v_n}^u(V^\mu_n)$, and $V^s_n((v_i)_{i \geq 0}) := \lim_{n \to \infty} F_{v_n}^s(V^s_n)$ exist, and are independent of the choice of $V^\mu_n$ and $V^s_n$.
2. $V^\mu_n((v_i)_{i \leq 0})$ is a $u$–admissible manifold in $v_0$, and $V^s_n((v_i)_{i \geq 0})$ is an $s$–admissible manifold in $v_0$.
3. $f(V^s_n((v_i)_{i \geq 0}) \subset V^s_n((v_i)_{i \geq 0} = V^s_n((v_i)_{i < 0}) \subset V^s_n((v_i)_{i < 0})$.
4. Write $v_i = \Psi_{x_i}^s$, then

   \begin{align*}
   V^s_n((v_i)_{i \geq 0}) &= \{ p \in \Psi_{x_0}^s \cap R_{00q}^s : \forall k \geq 0, f^k(p) \in \Psi_{x_{k+1}}^s \cap R_{00q}^s \}, \\
   V^u_n((v_i)_{i \leq 0}) &= \{ p \in \Psi_{x_0}^s \cap R_{00q}^s : \forall k \geq 0, f^{-k}(p) \in \Psi_{x_{k-1}}^s \cap R_{00q}^s \}.
   \end{align*}

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

   \begin{align*}
   \text{dist} C_1(V^\mu_n((u_i)_{i \leq 0}), V^\mu_n((v_i)_{i \leq 0})) &< K \theta^n, \\
   \text{dist} C_1(V^s_n((u_i)_{i \geq 0}), V^s_n((v_i)_{i \geq 0})) &< K \theta^n.
   \end{align*}

Parts (1)–(4) should be compared with Pesin’s Stable Manifold Theorem [P]. Part (5) should be compared to Brin’s Theorem on the Hölder continuity of the Oseledets distribution on Pesin sets [BT].
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 \( \Psi_x^{\alpha;\beta} \) and any two \( u \)-manifolds \( V_1^u, V_2^u \) in \( \Psi_x^{\alpha;\beta} \),
\[
\text{dist}(V_1^u, V_2^u) \leq 2Q_\varepsilon(x) < 1.
\]

Part 1. Existence of the limit.

By Proposition 4.12, \( F_0^n[V_{-n}^u] \) is a \( u \)-admissible manifold in \( v_0 \). By Proposition 4.13, for any other choice \( u \)-admissible manifolds \( W_{-n}^u \) in \( v_{-n} \),
\[
\text{dist}(F_0^n[V_{-n}^u], F_0^n[W_{-n}^u]) < \exp[-\frac{1}{2} \chi n] \text{dist}(V_{-n}^u, W_{-n}^u) < \exp[-\frac{1}{2} \chi n].
\]
Thus, if the limit exists then it is independent of \( V_{-n}^u \).

For every \( m > n \), \( W_{-n}^u := F_0^n[V_{-n}^u] \) is a \( u \)-admissible manifold in \( v_{-n} \). It follows that for every \( m > n \), \( \text{dist}(F_0^n[V_{-n}^u], F_0^n[V_{-m}^u]) < \exp[-\frac{1}{2} \chi n] \). It follows that \( \lim_{n \to \infty} F_0^n[V_{-n}^u] \) exists.

Part 2. Admissibility of the limit.

Write \( v_0 = \Psi_x^{\alpha;\beta} \), and let \( F_n \) denote the functions which represent \( F_0^n[V_{-n}^u] \) in \( v_0 \). Since \( F_0^n[V_{-n}^u] \) are \( u \)-admissible in \( v_0 \), for every \( n \),
\begin{itemize}
  \item \( \| F_n \|_{\beta/3} \leq \frac{1}{2} \);
  \item \( \| F_n'(0) \| \leq \frac{1}{2} (p^u \wedge p^s)^{\beta/3} \);
  \item \( |F_n(0)| \leq 10^{-3} (p^u \wedge p^s) \).
\end{itemize}

Since \( F_0^n[V_{-n}^u] \to V^u[(v_i)_{i \leq 0}], F_n \to F \) uniformly on \([p^u, p^s]\), where \( F \) represents \( V^u[(v_i)_{i \leq 0}] \).

By the Arzela–Ascoli Theorem, \( \exists n_k \uparrow \infty \) s.t. \( F'_{n_k} \to G \) uniformly, where \( \| G \|_{\beta/3} \leq \frac{1}{2} \). Thus \( F_{n_k}(t) = F_{n_k}(p^u) + \int_{p^u}^t F_{n_k}'(s)ds \to F(p^u) + \int_{p^u}^t G(s)ds \), whence \( F \) is differentiable, and \( F' = G \). We also see that \( \{ F_{n_k} \} \) can only have one limit point. Consequently, \( F' \to F' \) 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 \( V^u := V^u[(v_i)_{i \leq 0}] = \lim F_0^n[V_{-n}^u] \), and \( W^u := V^u[(v_{i-1})_{i \leq 0}] = \lim F_0^n[V_{-n-1}^u] \).
\[
\text{dist}(V^u, F_u(W^u)) \leq \text{dist}(V^u, F_0^n(V_{-n}^u)) + \text{dist}(F_0^n(V_{-n}^u), F_0^{n+1}(V_{-n-1}^u)) + \text{dist}(F_0^{n+1}(V_{-n-1}^u), F_u(W^u)) \leq \text{dist}(V^u, F_0^n(V_{-n}^u)) + e^{-\frac{1}{2} \chi n} \text{dist}(V_{-n}^u, F_u(V_{-n-1}^u)) + e^{-\frac{1}{2} \chi n} \text{dist}(F_0^n(V_{-n-1}^u), W^u).
\]

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

Part 4. Suppose \( v_i = \Psi_x^{\alpha;\beta} \), then
\[
V^u = \{ p \in [R_0^\beta(x)] : \forall k \geq 0, f^{-k}(p) \in [R_{10Q_\varepsilon(x-\delta)}(x)] \}.
\]
The inclusion $\subseteq$ is simple: Every $u$–admissible manifold $W^u_i$ in $\Psi^u_{x_i}[R_{p^u_i}(\Omega)]$, because if $W^u_i$ is represented by the function $F$ then any $p = \Psi_{x_i}(v, w)$ in $W^u_i$ satisfies $|w| \leq p^u_i$, and

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

Applying this to $V^u := V^u[\{v_i\}_{i \leq 0}]$, we see that $\forall p \in V^u, p \in V^u \subseteq \Psi_{x_0}[R_{p^u_0}(\Omega)]$, and by part 3 for every $k \geq 0$

$$f^{-k}(p) \in f^{-k}(V^u) \subseteq V^u[\{v_i\}_{i \leq 0}] \subseteq \Psi_{x_k}[R_{p^u_k}(\Omega)] \subseteq \Psi_{x_k}[R_{10Q_\varepsilon(x-k)}(\Omega)].$$

We have $\subseteq$.

We prove $\supseteq$. Suppose $z \in \Psi_{x_0}[R_{p^u_0}(\Omega)]$ and $f^{-k}(z) \in \Psi_{x_k}[R_{10Q_\varepsilon(x-k)}(\Omega)]$ 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 $\Xi = \Psi_{x_0}(v_0, w_0)$, where $v_0 = w_0$ and $w_0 = F(\Xi)$, for every $k \geq 0$. This point is so small that $\Xi, k \leq 10Q_\varepsilon(x-k)$ for all $k \geq 0$. Proposition 3.4, in its version for $f^{-1}$, says that for every $k \geq 0, f_{x_k-1}^{-1}x_k = \Psi_{x_k}(\Xi, w_k) \subseteq \Psi_{x_k}[R_{10Q_\varepsilon(x-k)}(\Omega)]$ can be put in the form

$$f_{x_k-1}^{-1}(v, w) = (A^{-1}_k v + g_1(k) v, B^{-1}_k w + g_2(k) v, w),$$

where $|A_k| < e^{-\chi/2}, |B_k| > e\chi/2$, and $\max R_{10Q_\varepsilon(x-k)} \| \nabla g_1(k) \| < \varepsilon$ (provided $\varepsilon$ is small enough).

Let $\Delta v_k := v_k - v_{k-1}$ and $\Delta w_k := w_k - w_{k-1}$. Since for every $k \leq 0$,

$$\Delta v_{k-1} := \Delta v_k - \Delta v_{k-1} = f_{x_k-1}^{-1}x_k (v_k, w_k) - (\Xi, w_k),$$

$$\Delta w_{k-1} := \Delta w_k - \Delta w_{k-1} = f_{x_k-1}^{-1}x_k (\Xi, w_k) - (\Xi, w_{k-1}),$$

$$\Delta v_{k-1} \geq A^{-1}_k \cdot |\Delta v_k| - \max \| \nabla g_1(k) \| \cdot (|\Delta v_k| + |\Delta w_k|) \geq (e^{\chi/2} - \varepsilon)|\Delta v_k| - \varepsilon|\Delta w_k|,$$

$$\Delta w_{k-1} \leq B^{-1}_k \cdot |\Delta w_k| + \max \| \nabla g_1(k) \| \cdot (|\Delta v_k| + |\Delta w_k|) \leq (e^{\chi/2} + \varepsilon)|\Delta w_k| + \varepsilon|\Delta v_k|.$$
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 \( a_k \to \infty \) as \( k \to \infty \). But \( a_k = |v_{\theta-k} - \theta_k| \leq 20|Q_\varepsilon(x_{\theta-k})| < 20\varepsilon \), so \( a_0 = 0 \). Since \( a_0 = 0, v_0 = \theta_0 \), and therefore \( F(\theta_0) = F(\theta_0) \). Thus \( z = \Psi_x(F(\theta_0), \theta_0) \in V^u \).

Part 5. Hölder continuity of \( \underline{u} \mapsto V^u((u_i)_{i \in \mathbb{Z}}) \).

Suppose two chains \( \underline{u} = (v_i)_{i \in \mathbb{Z}}, \underline{w} = (w_i)_{i \in \mathbb{Z}} \) satisfy \( v_i = w_i \) for \( i = -N, \ldots, N \). Given \( n > N \), let \( V^u_{-n} \) be a \( u \)-admissible manifold in \( v_{-n} \), and let \( W^u_{n} \) be a \( w \)-admissible manifold in \( v_{w-n} \).

Let \( F^u_{n}(V^u_{-n}) \) (resp. \( F^u_{n}(W^u_{-n}) \)) denote the result of applying \( F \) \( n \) times to \( V^u_{-n} \) using the path \( u_{-n} \to \cdots \to u_{n-1} \to u_n \) (resp. \( w_{-n} \to \cdots \to w_{n-1} \to w_n \)).

\( F^u_{n-N}(V^u_{-n}) \) and \( F^u_{n-N}(W^u_{-n}) \) are \( u \)-admissible manifolds in \( v_{-n}(=w_{-n}) \). Let \( F^n, G^n \) be their representing functions. Admissibility implies that

\[
\|F^n - G^n\| \leq \|F^n\| + \|G^n\| < 2Q_\varepsilon < 1\]

Represent \( F_{u}^{n-k}[V_u] \) and \( F_{u}^{n-k}[W_u] \) by functions \( F_k \) and \( G_k \). By \( 4.11 \),

\[
\|F_k - G_{k-1}\| \leq e^{-\chi/2}\|F_k - G_k\| \quad (4.7)
\]

\[
\|F_k - G_{k-1}\| \leq e^{-\chi/2}(\|F_k - G_k\| + 2\|F_k - G_k\|^{3/2}) \quad (4.8)
\]

Iterating \( (4.7) \) starting at \( k = N \) and going down, we get \( \|F_k - G_k\| \leq e^{-\chi(N-k)} \), whence \( \text{dist}(F^n_u[V^n_u], F^{n-N}_u[W^n_u]) \leq e^{-\chi N}. \]

Passing to the limit \( n \to \infty \), we get

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

Now let \( k = N \), paying attention to the inequalities \( \theta_1 < \theta_2 \) and \( c_N \leq 1: c_0 \leq \theta_1^N + 2\theta_1^{2N} < (2N + 1)\theta_1^{N+1} \).

It follows that \( \text{dist}((F^n_u[V^n_u]), F^{n-N}_u[W^n_u]) \leq 2(N + 1)\theta_1^{N+1} \) in \( 4.11 \). In part 2, we saw that \( F^n_u[V^n_u] \) and \( F^n_u[W^n_u] \) converge to \( V^u((v_i)_{i \leq 0}) \) in \( C^1 \). Therefore if we pass to the limit as \( n \to \infty \), we get \( \text{dist}(V^n_u((v_i)_{i \leq 0}), V^n_u((w_i)_{i \leq 0})) \leq 2(N + 1)\theta_1^{N+1} \). Now pick two constants \( \theta \in (\theta_2, 1) \) and \( K > 0 \) such that \( 2(N + 1)\theta_1^{N+1} \leq K\theta^N \) for all \( N \geq 0 \).

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

1. \( \pi \) is a well-defined and \( \pi \circ \sigma = f \circ \pi \);
2. \( \sigma : \Sigma \to M \) is Hölder continuous map;
3. \( \pi(\Sigma) \subset \pi(\Sigma^\#) \supset \text{NUH}_\chi^\#(f) \), 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 \( v_i = \Psi^{v_i}_{\theta_i} \psi_i \) and \( z = \pi(\underline{v}) \). We claim that

\[
f^k(z) \in \Psi_{\theta_{z_k}}[R_{Q_{\varepsilon}(x_k)}(\underline{v})] \quad (k \in \mathbb{Z}).
\]
For $k = 0$, this is because $z \in V^s[(v_1)_{i \geq 0}]$ and $V^s[(v_1)_{i \geq 0}]$ is $s$--admissible in $\Psi_{x_i}^{p_i^u, p_i^v}$.

For $k > 0$, we use Proposition 4.15 part (3) to see that

$$f^k(z) \in f^k(V^s[(v_1)_{i \geq 0}]) \subset V^s[(v_{i+k})_{i \geq 0}].$$

Since $V^s[(v_{i+k})_{i \geq 0}]$ is an $s$--admissible manifold in $\Psi_{x_i}^{p_i^u, p_i^v}$, $f^k(z) \in \Psi_{x_k}[R_{Q_k(x_k)}(0)]$.

The case $k < 0$ can be handled in the same way, using $V^u[(v_1)_{i \leq 0}]$. Thus $z = \pi(\underline{w})$ satisfies (4.9).

Any point which satisfies (4.9) must equal $z$, because by Proposition 4.15 part (4), it must lie on $V^u[(v_1)_{i \leq 0}] \cap V^s[(v_1)_{i \geq 0}]$. So (4.9) characterizes $\pi(\underline{w})$.

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

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

We saw that $\underline{u} \mapsto V^u[(u_1)_{i \leq 0}]$ and $\underline{w} \mapsto V^s[(v_1)_{i \geq 0}]$ are Hölder continuous (Proposition 4.15). Since the 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 $\pi(\Sigma) \supset \text{NUH}^\#(f)$. Suppose $x \in \text{NUH}^\#(f)$. By Proposition 4.16 there exist $\Psi_{x_k}^{p_k, p_k} \in \mathcal{Y}$ s.t. $\Psi_{x_k}^{p_k, p_k} \rightarrow \Psi_{x_{k+1}}^{p_{k+1}, p_{k+1}}$ for all $k$, and s.t. $\Psi_{x_k}^{p_k, p_k} \varepsilon$--overlaps $\Psi_{f^k(x)}^{p_k, p_k}$ for all $k \in \mathbb{Z}$. By Proposition 3.2(1), this implies that

$$f^k(x) = \Psi_{f^k(x)}^{p_k(x)}(\underline{w}) \in \Psi_{x_k}[R_{p_k(x)}^{p_k(x)}(\underline{w})] \subset \Psi_{x_k}[R_{Q_k(x_k)}(0)]$$

for all $k \in \mathbb{Z}$.

Thus $x$ satisfies (4.9) with $\underline{w} = (\Psi_{x_i}^{p_i^u, p_i^v})_{i \in \mathbb{Z}}$. It follows that $z = \pi(\underline{w})$.

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 $p_i^u \land p_i^v \geq q_\varepsilon(f^i(x))$. By the definition of $\text{NUH}^\#(f)$, there exist sequences $i_k, j_k \uparrow \infty$ for which $p_{i_k}^u \land p_{i_k}^v$ and $p_{j_k}^u \land p_{j_k}^v$ are bounded away from zero. By the discreteness property of $\mathcal{Y}$ (Proposition 3.7), $\Psi_{x_i}^{p_i, p_i}$ must repeat some symbol infinitely often in the past, and (possibly a different symbol) in the future. Thus the above actually proves that

$$\pi(\Sigma^\#) \supset \text{NUH}^\#(f),$$

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

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

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

$\Sigma_{rel}$ 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 [1.16] holds with $\Sigma_{rel}$ replacing $\Sigma$.

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

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

Henceforth we assume w.l.o.g. that all irrelevant vertices have been removed from $\mathcal{V}$, and we set $\Sigma := \Sigma_{rel}$.

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 $x \in NUH_{\chi}^\#(f)$ takes the form $x = \pi(\varpi)$ for some chain $\varpi \in \Sigma^\#$. In principle, there could be infinitely many chains $\varpi$ s.t. $\pi(\varpi) = x$. We ask what one can say about the solutions $\varpi$ to the equation $\pi(\varpi) = x$.

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 $\varpi$ 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 $n_k, m_k \uparrow \infty v_{-m_k} = u, v_n = v$ 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 $(\Psi_{y_i}^{p_i, q_i})_{i \in \mathbb{Z}}$, $(\Psi_{y_i}^{p_i, q_i})_{i \in \mathbb{Z}}$ are regular chains s.t. $\pi(\Psi_{y_i}^{p_i, q_i}) = \pi(\Psi_{y_i}^{p_i, q_i})$, then for all $i$,

1. $d(x, y_i) < \varepsilon$;
2. $(\Psi_{y_i}^{1, \sigma})_{\{y \in R} = (-1)^{\sigma_i} + \Delta_{\{y \in R}$ for all $y \in R_i(\varpi)$, where $\sigma_i \in \{0, 1\}$,
3. $\psi_i$ is a constant vector s.t. $\|\psi_i\| \leq 10^{-1}(q_i \cap q_i)$, and $\Delta_{\{y \in R}$ is a vector field s.t. $\Delta_{\{y \in R} = 0$ and $\|\Delta_{\{y \in R} \| < \sqrt{\varepsilon}$ on $R_i(\varpi)$;
4. $p_i / q_i \cap q_i \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$. 

The proof of Theorem 5.2 is long, so we broke it into several sections (5.6–8). Here is an overview. Suppose $(\Psi_{y_i}^{p_i, q_i})_{i \in \mathbb{Z}}$, $(\Psi_{y_i}^{p_i, q_i})_{i \in \mathbb{Z}}$ are two chains in $\Sigma^\#$ s.t.

\[ \pi(\Psi_{y_i}^{p_i, q_i}) = \pi(\Psi_{y_i}^{p_i, q_i}) = x \] (5.1)

We want to show that $\Psi_{y_i}$ 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 $\Psi_{y_i}^{p_i, q_i}$, therefore (Proposition 4.11), $f^i(x) = \Psi_{y_i}(v_i, u_i)$ where $|v_i|, |u_i| \leq 10^{-2}(p_i \cap q_i)$ and $\|f^i(x, v_i) \| < 50^{-1}(p_i \cap q_i)$). Similarly $\|f^i(x, u_i) \| < 50^{-1}(q_i \cap q_i)$. It follows that $d(x_i, y_i) < 25^{-1} \max\{|p_i \cap q_i|, |q_i \cap q_i|\} \leq \varepsilon$ for all $i \in \mathbb{Z}$.

Assume without loss of generality that $\varepsilon$ is smaller than the Lebesgue number of the cover $\mathcal{V}$ which we had constructed in [3.3] then $x_i, y_i$ belong to the same element $D_i$ of $\mathcal{Y}$. This allows us to write

\[ \Psi_{x_i} = \exp_{x_i} \circ \vartheta_{x_i} \circ C_{x_i} \]
\[ \Psi_{y_i} = \exp_{y_i} \circ \vartheta_{y_i} \circ C_{y_i} \]
where $\partial_z: \mathbb{R}^2 \to T_z M$ ($z = x_i, y_i$) are the isometries we constructed in [3, 1] and $C_{x_i}, C_{y_i} \in \text{GL}(2, \mathbb{R})$ are given by $C_{x_i}(x_i) = \iota_{x_i} \circ C_x$, and $C_{y_i}(y_i) = \iota_{y_i} \circ C_y$.

Let $z_i = x_i, y_i$, then $C_{x_i}(z_i)$ is the unique linear operator which maps $\xi^1 = (\psi^1_0)$ to $s_{x_i}(z_i)^{-1} \xi^u(z_i)$, and $\xi^2 = (\psi^2_0)$ to $u_{x_i}(z_i)^{-1} \xi^u(z_i)$. Writing as usual $\alpha(z_i) := \xi(\xi^u(z_i), \xi^s(z_i))$, we see that

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

(5.2)

where $R_{z_i}$ is the unique orientation preserving orthogonal matrix which rotates $\xi^1$ to the direction of $\partial_z^{+1}(\xi^u(z_i))$ ($z_i = x_i, y_i$). Some terminology:

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

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

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

**Proposition 5.3.** Let $(\Psi^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$, $(\Psi^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$ be two chains s.t. $\pi[(\Psi^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$, 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 $(\Psi^{p_i^u, p_i^s})_{i \in \mathbb{Z}}$, $(\Psi^{q_i^u, q_i^s})_{i \in \mathbb{Z}}$ be two chains s.t. $\pi[(\Psi^{p_i^u, p_i^s})_{i \in \mathbb{Z}}] = \pi[(\Psi^{q_i^u, q_i^s})_{i \in \mathbb{Z}}]$. We compare $R_{z_i}$ 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.** Suppose $V^u$ is a $u$–admissible manifold in $\Psi^{p_i^u, p_i^s}$. We say that $V^u$ stays in windows if there is a negative chain $(\Psi^{p_i^u, p_i^s})_{i \leq 0}$ with $\Psi^{p_0^u, p_0^s} = \Psi^{p_i^u, p_i^s}$ and $u$–admissible manifolds $W_i^u$ in $\Psi^{p_i^u, p_i^s}$ s.t. $f^{-i}(V^u_i) \subseteq W_i^u$ for all $i \leq 0$.

**Definition 6.2.** Suppose $V^s$ is an $s$–admissible manifold in $\Psi^{p_i^u, p_i^s}$. We say that $V^s$ stays in windows if there is a positive chain $(\Psi^{p_i^u, p_i^s})_{i \geq 0}$ with $\Psi^{p_0^u, p_0^s} = \Psi^{p_i^u, p_i^s}$ and $s$–admissible manifolds $W_i^s$ in $\Psi^{p_i^u, p_i^s}$ s.t. $f^i(V^s_i) \subseteq W_i^s$ for all $i \geq 0$.

If $\xi$ is a chain, then $V^u_i := V^u[(v_k)_{k \leq i}]$ and $V^s_i := V^s[(v_k)_{k \geq i}]$ stay in windows, because $f^{-k}(V^u_i) \subseteq V^u_{i-k}$ and $f^k(V^s_i) \subseteq V^s_{i+k}$ for all $k \geq 0$ (Proposition [4, 13]).

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 $\Psi^{p_i^u, p_i^s}$, and suppose $V^s$ stays in windows.

1. For every $y, z \in V^s$, $d(f^k(y), f^k(z)) < e^{-\frac{\varepsilon}{4} k}$ for all $k \geq 0$.
2. For every $y \in V^s$, let $\xi^u(y)$ denote the positively oriented unit tangent vector to $V^s$ at $y$, then $\|df^k_x \xi^u(y)\|_{f^k(y)} \leq 6\|C_x(x)^{-1}\|e^{-\frac{\varepsilon}{4} k}$ for all $k \geq 0$. 
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]: \( f^n : V^s \to f^n(V^s) \) is given in coordinates by

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

Since \( V^s \) stays in windows, the orbits of points in \( V^s \) remain in the “windows” where \( f_{x_{i+1}} \) 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 \( \Psi^{\pm}_{x_0} \) (resp. in \( \Psi^{-}_0 \)). 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 \( (\Psi^{\pm}_{x_i}, \pi^i_{x_i})_{i \in \mathbb{Z}}, (\Psi^{\pm}_{y_i}, \pi^i_{y_i})_{i \in \mathbb{Z}} \) be chains s.t. \( \pi((\Psi^{\pm}_{x_i}, \pi^i_{x_i})_{i \in \mathbb{Z}}) = \pi((\Psi^{\pm}_{y_i}, \pi^i_{y_i})_{i \in \mathbb{Z}}) \), 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 \( v_i = \Psi^{\pm}_{x_i}, u_i = \Psi^{\pm}_{y_i}, x := (\Psi^{\pm}_{x_i}, \pi^i_{x_i})_{i \in \mathbb{Z}} = \pi((\Psi^{\pm}_{y_i}, \pi^i_{y_i})_{i \in \mathbb{Z}}), \) and

\[
V^s_{x_k} := V^s[(v_i)_{i \geq k}],
V^u_{x_k} := V^u[(v_i)_{i \leq k}],
E^s_{x_k} := T_{f^k(x)} V^s_{x_k},
E^u_{x_k} := T_{f^k(x)} V^u_{x_k}.
\]

We claim that

1. \( \limsup_{n \to \infty} \frac{1}{n} \log \|df^n_{f^k(x)}(w)\| < 0 \) on \( E^s_{x_k} \setminus \{0\} \) and \( E^u_{x_k} \setminus \{0\} \).
2. \( \limsup_{n \to \infty} \frac{1}{n} \log \|df^n_{f^k(x)}(w)\| > 0 \) on \( E^s_{x_k} \setminus \{0\} \) and \( E^u_{x_k} \setminus \{0\} \).

We give the details for \( E^s_{x_k} \). The case of \( E^u_{x_k} \) is identical.

Part (i) follows from Proposition 6.3 (2), applied to \( V^s_{x_k} \) and \( V^u_{x_k} \).

The proof of (ii) is slightly more complicated. Suppose \( w \in E^s_{x_k} \setminus \{0\} \), then \( w \) is tangent to \( V^u_{x_k} \) at \( f^k(x) \). For every \( n, f^{k+n}(x) = \pi((v_i+k+n)_{i \in \mathbb{Z}}) \in V^u_{x_k+n+k} \), so

\[ f^{k+n}(x) = f^{-n}(f^{k+n}(x)) \in f^{-n}[V^u_{x_k+n+k}] \]

It follows that \( df^n_{f^k(x)}(w) \in T_{f^{k+n}(x)}[V^u_{x_k+n+k}] \setminus \{0\} \).

We apply Proposition 6.3 (2) in its version for \( u \)-admissible manifolds to the manifold \( V^u_{x_k+n+k} \) and the vector \( df^n_{f^k(x)}(w) \). This gives the estimate

\[
\|w\| \leq \|df^n_{f^{k+n}(x)}(w)\| \leq 6e^{-\frac{1}{2}n\varepsilon} \|C_x(x_{k+n})^{-1}\| \cdot \|df^n_{f^k(x)}(w)\| \leq 6e^{-\frac{1}{2}n\varepsilon} Q_x(x_{k+n})^{-1} \|df^n_{f^k(x)}(w)\| \quad (\text{definition of } Q_x)
\]

\[
\leq 6e^{-\frac{1}{2}n\varepsilon} (p_k^n \wedge p_k^{n+n})^{-1} \|df^n_{f^k(x)}(w)\| \leq 6e^{-\frac{1}{2}n\varepsilon} (p_k^n \wedge p_k^n)^{-1} \|df^n_{f^k(x)}(w)\| \quad (\text{Lemma 4.4}).
\]
Similarly one sees that
\[ \epsilon \in \mathbb{Q} \] where
\[ v \text{ also set as usual} \]

At \[ v \] also set as usual

Since \[ p_i \leq Q_i(x_i) < \epsilon^{3/\beta} \] and \[ q_i \leq Q_i(y_i) < \epsilon^{3/\beta} \], \[ e^{-2\epsilon^{3/4}} < \sin \alpha(x_i) / \sin \alpha(y_k) < e^{2\epsilon^{3/4}} \]. Similarly one sees that \[ |\cos \alpha(x_i) - \cos \alpha(y_k)| < 4e^{3/4}, \] and the proposition follows for all \( \epsilon \) so small that \( 4e^{3/4} < \sqrt{x} \).

The proof actually gives the following stronger estimates, which we now record for future reference:

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

1. \[ e^{-(p_i^q + p_i^q)^{\beta/4} - (q_i^q + q_i^q)^{\beta/4}} < \sin \alpha(x_i) / \sin \alpha(y_k) < e^{(p_i^q + p_i^q)^{\beta/4} + (q_i^q + q_i^q)^{\beta/4}} \]

2. \[ |\cos \alpha(x_i) - \cos \alpha(y_k)| < 4(|p_i^q \wedge p_i^q|^{\beta/4} + |q_i^q \wedge q_i^q|^{\beta/4}) \]

### 6.3. Comparison of \( R_{x_i} \) to \( R_{y_i} \)

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

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

where \( \sigma_i \in \{0, 1\} \) and \( |\varepsilon_{jk}| < |(p_i^q \wedge p_i^q)^{\beta/5} + (q_i^q \wedge q_i^q)^{\beta/5}| < \sqrt{x} \).

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

Let \( z = \pi^{1/2} = \pi^{1/2} \). The manifold \( V^*[(v_i)_{i \geq 0}] \) inherits an orientation from the chart \( \Psi_x \). Let \( e_x(z) \) denote the positively oriented unit tangent vector to \( V^*[(v_i)_{i \geq 0}] \) at \( z \). The manifold \( V^*[(u_i)_{i \geq 0}] \) inherits an orientation from the chart \( \Psi_y \). Let \( e_y(z) \) denote the positively oriented unit tangent vector to \( V^*[(u_i)_{i \geq 0}] \) at \( z \). Since \( T_z V^*[(v_i)_{i \in \mathbb{Z}}] = T_z V^*[(u_i)_{i \in \mathbb{Z}}] \) (see the proof of Proposition 6.5), \( e_x(z) = \pm e_y(z) \).

We write \( z = \Psi_x(\zeta) \) and \( c_x(z) = \frac{(d\Psi_x)_{\zeta}}{|(d\Psi_x)_{\zeta}|} \), where \( \zeta \in R_{10^{-2p}}(0) \), \( a = (1, a) \), and \( |a| < p^{\beta/3} \) (see Proposition 4.11 and 4.3).

Let \( z = \Psi_y(\eta) \) and \( c_y(z) = \frac{(d\Psi_y)_{\eta}}{|(d\Psi_y)_{\eta}|} \), where \( \eta \in R_{10^{-2q}}(0) \), \( b = (1, b) \), and \( |b| < q^{\beta/3} \) (see Proposition 4.11 and 4.3).

Since \( e_x(z) = \pm e_y(z) \), there is a non-zero (signed) scalar \( \lambda \) such that
\[
C_x a = \lambda ([d\exp_x \circ \partial_x] c_x \zeta)^{-1} ([d\exp_y \circ \partial_y] c_x \zeta) C_y b.
\]

(6.1)

where \( C_x, C_y \) are given by (4.12).

**Claim 1.** \( C_x a \propto R_x (1 \pm p^{\beta/3}) \) and \( C_y b \propto R_y (1 \pm q^{\beta/3}) \). Here \( a \propto b \) means that \( a = t b \) for some \( t \neq 0 \), and \( a \pm c \) means a quantity in \( [a - c, a + c] \).
Proof. \( C_x \mathbf{u} = R_x \left( \frac{s_x(x)^{-1} + u_x(x)^{-1} \cos \alpha(x) a}{u_x(x)^{-1} \sin \alpha(x) a} \right) \)
\( \propto R_x \left( \frac{1 \pm \|C_x(x)^{-1}\| \cdot |a|}{0 \pm \|C_x(x)^{-1}\| \cdot |a|} \right) \), because \( u_x > 1 \) and \( s_x = \|C_x(x)^{-1} e^*(x)\| \)
\( = R_x \left( \frac{1 \pm p^{\beta/4}}{0 \pm p^{\beta/4}} \right), \) because \( |a| < p^{\beta/3} \leq Q_x(x)^{\beta/12} p^{\beta/4} < \frac{p^{\beta/4}}{\|C_x(x)^{-1}\|} \).

Similarly, \( C_y \mathbf{u} \propto R_y (1 \pm q^{\beta/4}) \).

Claim 2. There exists a constant \( J > 1 \) (which only depends on \( M \)) s.t. for all \( D \in \mathcal{D}, x, y \in D, \|w_x\|, \|w_y\| < 2, \) \( \|((d \exp_x \circ \partial_x)_w)^{-1}[(d \exp_y \circ \partial_y)_w] - \text{Id}\| < J (d(x, y) + \|w_x - w_y\|). \)

Proof. Let \( J_1 \) denote a common Lipschitz constant for the maps
\( (w, w) \mapsto (d \exp_w \circ \partial_w)_w \)
on \( D \times B_2(0) \) for all \( D \in \mathcal{D}. \) Let \( J_2 \) denote the maximum over \( D \in \mathcal{D} \) of \( \sup \|((d \exp_w \circ \partial_w)_w)^{-1} : w \in D, \|w\| < 2 \| \). The claim holds with \( J := J_1 J_2 + 1. \)

Claim 3. \( R_x (1 \mid 0) + \xi_1 \propto R_y (1 \mid 0) + \xi_2 \) where \( \|\xi_1\| \) and \( \|\xi_2\| \) are less than \( 3 J (p^{\beta/4} + q^{\beta/4}). \)

Proof. \( C_x(\cdot) \) is a contraction, so \( \|C_x \xi - C_y \eta\| < \|\xi\| + \|\eta\| < 10^{-2} (p + q) \). Also, by Proposition \( 5.3 \) \( d(x, y) < 25^{-1} (p + q) \). Therefore, by Claim 2,
\( \|((d \exp_x \circ \partial_x)_{C_x(x) \xi})^{-1}[(d \exp_y \circ \partial_y)_{C_y(y) \eta}] - \text{Id}\| = E \)
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
\( R_y^{-1} R_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c \begin{pmatrix} 1 \\ 0 \end{pmatrix} + R_y^{-1} \xi_2 - c^{-1} R_y^{-1} \xi_1 \), \( (6.2) \)
where \( c \) is a scalar s.t. \( |c| = \frac{1 + \|\xi_2\|}{1 - \|\xi_2\|} \). Since \( \|\xi_2\| < 3 J (p^{\beta/4} + q^{\beta/4}) < 6 J \epsilon^{3/4}, \) \( |c| \in \left[ e^{-10 J \epsilon^{\sqrt{2}}}, e^{10 J \epsilon^{\sqrt{2}}} \right], \) at least provided \( \epsilon \) 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 \( \theta \in (-\frac{\pi}{2}, \frac{\pi}{2}). \)

\( |\theta| \leq \tan^{-1} \left( \frac{\|\xi_2\| + |c|^{-1} \cdot \|\xi_1\|}{1 - \|\xi_2\| - |c|^{-1} \cdot \|\xi_1\|} \right) \leq \frac{3 J (1 + e^{10 J \epsilon^{\sqrt{2}}})}{1 - 6 J (1 + e^{10 J \epsilon^{\sqrt{2}}}) \epsilon^{3/4} (p^{\beta/4} + q^{\beta/4})}. \) 

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

7.1. The $s_\chi$ and $u_\chi$ parameters of admissible manifolds. In [2.1] we defined $s_\chi(\cdot)$ on $NU\chi_1(f)$. 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 $y \in NU\chi_1(f)$ define $\xi^s(y)$ as in [2.1] and note that by proposition 6.3(2), $\xi^s(y)$ is tangent to $V^s$ at $y$. Motivated by this, we define $\xi^s(y)$ for $y \notin NU\chi_1(f)$ to be one of the two unit tangent vectors to $V^s$ at $y$ (it doesn’t matter which), and then we let

$$s_\chi(y) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \| df^k_y \xi^s(y) \|_F^2 \right)^{\frac{1}{2}} \in (\sqrt{2}, \infty].$$

Similarly, for any $u$-admissible manifold $V^u$ which stays in windows, and any $y \in V^u$ we define $\xi^u(y)$ as in [2.1] when $y \in NU\chi(f)$, and we let $\xi^u(y)$ be one of the two unit tangent vectors to $V^u$ at $y$ when $y \notin NU\chi(f)$. Then we let

$$u_\chi(y) := \sqrt{2} \left( \sum_{k=0}^{\infty} e^{2k\chi} \| df^k_y \xi^u(y) \|_F^2 \right)^{\frac{1}{2}} \in (\sqrt{2}, \infty].$$

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

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

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

**Definition 7.1.** Let $V^s$ be an $s$-admissible manifold in $\Psi_{x^s,p^s}$ with representing function $F^s$. Let $V^u$ be a $u$-admissible manifold in $\Psi_{x^u,p^u}$ 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 $\epsilon$ small enough. Suppose $\Psi_{x^s,p^s} \to \Psi_{y^u,p^u}$, and let $V^s \to V^u$ be an $s$-admissible manifold in $\Psi_{x^s,p^s}$ which stays in windows. If $s_\chi(V^s) < \epsilon$ then $s_\chi(F_x(V^s)) < \epsilon$, and for every $\rho \geq \exp(\sqrt{\epsilon})$,

$$\frac{s_\chi(V^s)}{s_\chi(x)} \in [\rho^{-1}, \rho] \implies s_\chi(F_x(V^s)) \in \left[\rho^{-1}e^{Q_\tau(x)^{\beta/4}}, \rho e^{-Q_\tau(x)^{\beta/4}}\right]. (7.1)$$

A similar statement holds for $u$-admissible manifolds in $\Psi_{x^u,p^u}$ and $F_u$.

Note that the ratio bound in (7.1) improves.

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

Suppose $s_\chi(V^s) < \epsilon$, then $s_\chi(q) < \epsilon$. By Proposition 1.12.4 (in its version for $s$-manifolds), $f^{-1}(q) \in U^s$. Since $U^s$ is one-dimensional, $df_{f^{-1}(q)}\xi^s(f^{-1}(q)) = \pm ||df_{f^{-1}(q)}\xi^s(f^{-1}(q))||_F \cdot \xi^s(q)$, and so

$$s_\chi(f^{-1}(q))^2 = 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k\chi} ||df^k_{q} df_{f^{-1}(q)}\xi^s(f^{-1}(q))||_F^2 \right) = 2 + e^{2\chi} ||df_{f^{-1}(q)}\xi^s(f^{-1}(q))||_F^2 \cdot s_\chi(q)^2 < \infty.$$
Since \( f^{-1}(q) \in U^s \), \( s_\chi(U^s) \leq e^{\epsilon_\chi} s_\chi(f^{-1}(q)) < \infty \).

Next assume that \( s_\chi(V^s) \) is finite, and
\[
\frac{s_\chi(U^s)}{s_\chi(x)} = \frac{s_\chi(p)}{s_\chi(f^{-1}(q))} \cdot \frac{s_\chi(f^{-1}(q))}{s_\chi(f^{-1}(y))} \cdot \frac{s_\chi(f^{-1}(y))}{s_\chi(x)}.
\] (7.2)
The three terms are well-defined and finite, because (proceeding from right to left):
- \( s_\chi(x), s_\chi(f^{-1}(q)) \) 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 \( s_\chi(p) = S_\chi(U) < \infty \) (see above).

The first factor in (7.2) belongs to \([e^{-Q_\chi(x)^{3/4}}, e^{Q_\chi(x)^{3/4}}]\) by Proposition 6.3(3). The second factor in (7.2) takes values in \([e^{-Q_\chi(x)^{3/4}}, e^{Q_\chi(x)^{3/4}}]\) because \( \Psi_{\chi}^{\rho_{\chi}, \rho_{\chi}} \rightarrow \Psi_{\chi}^{\rho_{\chi}, \rho_{\chi}}\), see Lemma 3.3. To prove the proposition, it is enough to show that
\[
\frac{1}{\rho} \exp[3Q_\chi(x)^{3/4}] < \frac{s_\chi(f^{-1}(q))}{s_\chi(f^{-1}(y))} < \rho \exp[-3Q_\chi(x)^{3/4}].
\] (7.3)

We begin with some identities. We omit the tags of the Riemannian norm, to avoid heavy notation. Since \( df_{f^{-1}(y)} \xi^s(f^{-1}(y)) = \pm \|df_{f^{-1}(y)} \xi^s(f^{-1}(y))\| \cdot \xi^s(y) \),
\[
s_\chi(f^{-1}(y))^2 = 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k} \|df_{f^{-1}(y)} \xi^s(f^{-1}(y))\|^2 \right)
= 2 + e^{2} s_\chi(y)^2 \|df_{f^{-1}(y)} \xi^s(f^{-1}(y))\|^2.
\] (7.4)
Similarly, \( df_{f^{-1}(q)} \xi^s(f^{-1}(q)) = \pm \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\| \cdot \xi^s(q) \), so
\[
s_\chi(f^{-1}(q))^2 = 2 + e^{2} s_\chi(q)^2 \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\|^2
\leq 2 + \rho^2 e^{2} s_\chi(y)^2 \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\|^2 \leq 2 + \rho^2 e^{2} s_\chi(y)^2 \|df_{f^{-1}(y)} \xi^s(f^{-1}(y))\|^2 \times
\times \exp \left( 2 \log \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\| - \log \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\| \right).
\]
We obtain the estimate
\[
\frac{s_\chi(f^{-1}(q))^2}{s_\chi(f^{-1}(y))^2} \leq \left( 2 + \rho^2 e^{2} s_\chi(y)^2 \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\|^2 \right) \times
\times \exp \left( 2 \log \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\| - \log \|df_{f^{-1}(q)} \xi^s(f^{-1}(q))\| \right).
\] (7.5)
Call the first factor I and the second factor II.
Analysis of I.

\[ I = \rho^2 - \frac{2(\rho^2 - 1)}{2 + e^{2x}s_x(y)\|df_{f^{-1}(y)}e^x(f^{-1}(y))\|^2} \]

\[ = \rho^2 - \frac{2(\rho^2 - 1)}{s_x(f^{-1}(y))^2}, \text{ by (7.4)} \]

\[ \leq \rho^2 - \frac{e^{-2\varepsilon^{2/3}} \cdot 2(\rho^2 - 1)}{s_x(x)^2}, \text{ because } s_x(f^{-1}(y)) = \exp[\pm \varepsilon^{2/3}] \text{ by Lemma 3.3} \]

\[ \leq \rho^2 \left( 1 - \frac{2e^{-2\varepsilon^{2/3}}(1 - \rho^{-2})}{\|C_x(x)^{-1}\|^2} \right), \text{ since } s_x(x) = \|C_x(x)^{-1}\| \leq \|C_x(x)^{-1}\| \]

\[ \leq \rho^2 \left( 1 - \frac{e^{1/2}}{\|C_x(x)^{-1}\|^2} \right) \text{ for all } \varepsilon \text{ small enough, because } \rho \geq e^{\sqrt{\varepsilon}}. \]

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

\[ \frac{\varepsilon^{1/2}}{\|C_x(x)^{-1}\|^2} > Q_\varepsilon(x)^{\beta/6} = Q_\varepsilon(x)^{-\beta/12}Q_\varepsilon(x)^{\beta/4} > \varepsilon^{-1/4}Q_\varepsilon(x)^{\beta/4}. \]

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

Analysis of II. Since \( f \) is a \( C^{1+\beta} \)-diffeomorphism and \( \|e^x(\cdot)\| = 1, \) there exists a constant \( K_0, \) which only depends on \( f, \) so that

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

where \( d_M \) and \( d_{TM} \) are the Riemannian distance functions on \( M \) and its tangent bundle. Since \( f \) is a \( C^{1+\beta} \) diffeomorphism and \( e^x(\cdot) \) are unit vectors, there is another constant \( H_1 \) (which only depends on \( f, \) such that

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

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,

\[ d(q, y) < 2|G(0)| \leq 2 \cdot 10^{-3}(p^* \& q^*) \leq 2 \cdot 10^{-3} \cdot \varepsilon^{p + p^*} \]

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

We estimate \( d_{TM}(e^x(q), e^x(y)). \) By the definition of \( \Psi_y, e^x(\cdot) \) is the normalization of \( (d\Psi_y)_{\Psi_y}([1]) = (d\exp_y)_{\Psi_y}([C_x(y)]([1])) \), and \( e^x(q) \) is the normalization of

\[ (d\Psi_y)_{\Psi_y}([G'(0)]) = (d\exp_y)_{\Psi_y}([C_x(y)]([0])) \cdot \frac{1}{G'(0)(0)}. \]

It is not difficult to see using the admissibility of \( V^* \) and Lemma 4.4 that \( |G(0)| < Q_\varepsilon(x) \) and \( |G'(0)| < Q_\varepsilon(x)^{\beta/3}. \) Since \( C_x(y) \) is a contraction, \( p \mapsto \exp_p p \) 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 \( d_{TM}(e^x(q), e^x(y)) < G_0 Q_\varepsilon(x)^{\beta/3}. \)

We see that \( II \leq \exp[(H_1 + H_1 G_0)Q_\varepsilon(x)^{\beta/3}] \). It follows that for all \( \varepsilon \) sufficiently small, \( II \leq \exp[Q_\varepsilon(x)^{\beta/4}]. \)
Summary. Combining the estimates of I and II, we find that
\[
\frac{s_x(f^{-1}(q))}{s_x(f^{-1}(y))} \leq \rho \exp[-3Q_x(x)^{\beta/4}].
\]

The other half of (7.3) is proved in a similar way. First, one proves that
\[
\frac{s_x(f^{-1}(q))^2}{s_x(f^{-1}(y))^2} \geq \left( \frac{2 + \rho^{-2}e^{2x} s_x(y)^2 \|df^{-1}(y)\| e^{x/(f^{-1}(y))} \| \right) \times \exp\left( -2 \log \|df^{-1}(y)\| e^{x/(f^{-1}(y))} \| - \log \|df^{-1}(y)\| e^{x/(f^{-1}(y))} \| \right),
\]
and then one analyzes the two terms as before.

\[\square\]

7.2. Comparison of \(s_x(x_i), u_x(x_i)\) to \(s_x(y_i), u_x(y_i)\).

**Proposition 7.3.** The following holds for all \(\varepsilon\) small enough. For any two regular chains \((\Psi_{v_i}^{p_i}, \psi_i)\) in \(\mathcal{Z}\), if \(\pi[(\Psi_{v_i}^{p_i}, \psi_i)] \in \mathcal{Z}\), then
\[
e^{-\varepsilon} \leq \frac{s_x(x_i)}{s_x(y_i)} \leq e^{-\varepsilon/2} \leq \frac{u_x(x_i)}{u_x(y_i)} \leq e^{\varepsilon/2} \quad \text{for all} \quad i \in \mathbb{Z}.
\]

**Proof.** Write \(\psi := (\Psi_{v_i}^{p_i}, \psi_i)\) in \(\mathcal{Z}\), \(\psi := (\Psi_{v_i}^{p_i}, \psi_i)\) in \(\mathcal{Z}\), and \(p := \pi(\psi) = \pi(\psi)\).

Let \(V_k^s := V_s[(v_i), k] , V_k^u := V_u[(v_i), k] , U_k^s := V_s[(v_i), k] , U_k^u := V_u[(v_i), k] \). We claim that it is enough to prove that
\[
\frac{s_x(V_k^s)}{s_x(x_k)} \frac{u_x(V_k^s)}{u_x(x_k)} \frac{s_x(U_k^u)}{s_x(y_k)} \frac{u_x(U_k^u)}{u_x(y_k)} \in \left[ e^{-\varepsilon}, e^{\varepsilon} \right]. \quad (7.6)
\]

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

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

We are assuming that \(v\) 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 \(v := \Psi_{v_i}^{p_i}\).

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

**Proof.** By convention \(v\) is relevant (see [4.4]). Choose a chain \(w\) s.t. \(w_0 = v\) and \(w := \pi(\psi) \in \mathcal{Y}_v(f)\). Let \(W^s := V^s[(v_i)] \geq 0\). This manifold has a finite \(s_x-\)parameter, because \(s_x(W^s) \leq e^{\varepsilon} s_x(w)\) and \(w \in \mathcal{Y}_v(f)\) so \(s_x(w) < \infty\). Let \(\rho_0 := \max \left\{ s_x(W^s), \frac{s_x(x)}{s_x(W^s)} \right\} \cdot \exp(\sqrt{\varepsilon}) \).

\(W^s\) is an admissible manifold in \(v_{n_k} = v\). By Proposition 4.4.1 if we take \(W^s\) at \(v_{n_k+\ell}\) and apply to it the graph transform \(F_s n_{k+\ell} - n_k\) times using the path \((v_{n_k}, \ldots, v_{n_k+\ell})\), then the resulting manifold
\[
W_{\ell}^s := F_s^{n_{k+\ell} - n_k}[W^s]
\]
is an $s$–admissible manifold in $v_{n_k}$, which converges to $V^s_{n_k}$. By Lemma 7.2

$$\frac{s_\chi\left(W^s_{\ell}\right)}{s_\chi(x)} \in [\rho_0^{-1}, \rho_0].$$

(7.7)

The convergence of $W^s_{\ell}$ to $V^s_{n_k}$ means that if $W^s_{\ell}$ is represented in $v_{n_k} = \Psi^\beta_{x}: p^\beta$ by the function $F_\ell$, and $V^s_{n_k}$ is represented in $\Psi^\beta_{x}: p^\beta$ by $F$, then $\|F_\ell - F\|_\infty \to 0$. In fact, since $\sup \|F_\ell\|_{\beta/3} < \infty$, we have the stronger statement that

$$\|F_\ell - F\|_\infty + \|F_\ell - F\|_\infty \to 0, \quad \ell \to \infty,$$

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 $\xi_\ell \to \xi$ and $\xi_\ell \to \xi, \ell \to \infty$.

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

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

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

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

Claim 2. $s_\chi(V^s_{0})/s_\chi(x_0) \in [\exp(-\sqrt{2}), \exp(\sqrt{2})]$.

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

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

By Proposition 4.15 (3), $F_s(V^s_{n_k}) = V^s_{n_k-1}$, and by Lemma 7.2, the bounds for $s_\chi(V^s_{n_k})/s_\chi(x_{n_k})$ improve. We ignore these improvements and write $s_\chi(V^s_{n_k-1})/s_\chi(x_{n_k-1}) \in [\rho^{-1}, \rho]$. Another application of $F_s$ gives $s_\chi(V^s_{n_k-2})/s_\chi(x_{n_k-2}) \in [\rho^{-1}, \rho]$. Continuing this way, we eventually reach the index $n_k-1 + 1$ and the bound

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

Since $x_{n_k} = x$, the next application of $F_s$ improves the ratio bound by at least $\exp(Q(x)\beta/4)$:

$$\frac{s_\chi(V^s_{n_k})}{s_\chi(x_{n_k})} \in [\rho^{-1} e^{Q(x)\beta/4}, \rho e^{-Q(x)\beta/4}].$$

We repeat the procedure by applying $F_s$ $n_k-1 - n_k-2 + 1$ times, whilst ignoring the potential improvements of the error bounds, and then applying $F_s$ once more and arriving at

$$\frac{s_\chi(V^s_{n_k-2})}{s_\chi(x_{n_k-2})} \in [\rho^{-1} e^{2Q(x)\beta/4}, \rho e^{-2Q(x)\beta/4}].$$
We are free to choose \( k \) as large as we want. If we make it so large that 
\[
\exp(k Q_v(x) \frac{\beta}{4}) \geq \rho \exp(-\sqrt{\varepsilon})
\]
then eventually we will reach a time \( n_{k_0} \) when the ratio bound is smaller than or equal to \( \exp(\varepsilon) \):
\[
\frac{s_k(V_{n_{k_0}}^x)}{s_k(x_{n_{k_0}})} \in [\exp(-\sqrt{\varepsilon}), \exp(\varepsilon)].
\]

This is the threshold the 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 
\[
\frac{s_k(V_{n_{k_0}}^x)}{s_k(x_{n_{k_0}})} \in [\exp(-\sqrt{\varepsilon}), \exp(\varepsilon)]
\]
as desired. \(\square\)

8. Window Parameters

8.1. \(\varepsilon\)-maximality. Let \( \underline{u} = (\Psi(v^u_1, v^u_i)_{i \in \mathbb{Z}}, \underline{u} = (\Psi(v^u_i, v^u_i')_{i \in \mathbb{Z}} \) be two regular chains such that \( \pi(\underline{u}) = \pi(\underline{u}). \) We compare \( p_i^u \) to \( q_i^u \) and \( p_i^u \) to \( q_i^u \). The idea is to use regularity to see that the \( q \)-parameters of \( V^u[v_i, n_{i=0}] \) and \( V^u[v_i, n_{i=0}] \) 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 \( p^v(v) \) and \( p^v(v) \) means the \( p^v \) and \( p^v \) in \( v = \Psi^v \).

**Definition 8.1.** A negative chain \( (v_i)_{i \leq 0} \) is called \( \varepsilon \)-maximal if it is regular, and
\[
p^v(v_0) \geq e^{-\sqrt{\varepsilon}} p^v(u_0)
\]
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. \( \pi(u_i) = \pi(v_i) \).

**Definition 8.2.** A positive chain \( (v_i)_{i \geq 0} \) is called \( \varepsilon \)-maximal if it is regular, and
\[
p^v(v_0) \geq e^{-\sqrt{\varepsilon}} p^v(u_0)
\]
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. \( \pi(u_i) = \pi(v_i) \).

**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 \( u_0 = \Psi^u x_0^u \) and \( v_0 = \Psi^v x_0^v \), then \( Q_v(x) / Q_v(y) \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}] \).

**Proof.** Propositions and say that
\[
\frac{\sin \alpha(x)}{\alpha(y)} \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}], \quad \frac{\sin \alpha(x)}{\alpha(y)} \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}],
\]
and \( \frac{\sin \alpha(x)}{u_0^v(x)} \in [e^{-4\sqrt{\varepsilon}}, e^{4\sqrt{\varepsilon}}] \). By Lemma, \( \frac{\alpha(x)}{\alpha(y)} \in [\exp(-5\sqrt{\varepsilon}), \exp(5\sqrt{\varepsilon})] \), whence \( Q_v(y) / Q_v(y) \in [\exp(-5\sqrt{\varepsilon}), \exp(5\sqrt{\varepsilon})] \). If \( \varepsilon \) is small enough, then \( Q_v(x) / Q_v(y) \in [\exp(-\sqrt{\varepsilon}), \exp(\sqrt{\varepsilon})] \).

**Step 2.** The following holds for all \( \varepsilon \) small enough: Every regular negative chain \( (v_i)_{i \leq 0} \) s.t. \( v_0 = \Psi^v x_0^v \) where \( p^v = Q_v(x) \) is \( \varepsilon \)-maximal, and every regular positive chain \( (v_i)_{i \geq 0} \) s.t. \( v_0 = \Psi^v x_0^v \) where \( p^v = Q_v(x) \) is \( \varepsilon \)-maximal.

**Proof.** Suppose \( (v_i)_{i \leq 0} \) is regular, and \( v_0 = \Psi^v x_0^v \) where \( p^v = Q_v(x) \). 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 \(u_0 = \Psi_{u}^{u'}.\) We have to show that \(p_u^v \geq e^{-\sqrt{Q_u}} q_u^w\). Indeed, by step 1, \(p_u^v = Q_u(x) \geq e^{-\sqrt{Q_u}} Q_u(y) \geq e^{-\sqrt{Q_u}} q_u^w\).

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 \(v_0 \to v_1\). If \((v_i)_{i \leq 0}\) is \(\varepsilon\)-maximal, then \((v_i)_{i \leq 1}\) is \(\varepsilon\)-maximal. Let \((v_i)_{i \geq 1}\) be a regular positive chain, and suppose \(v_{-1} \to v_0\). 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 positive chain, and suppose \(v_0 \to v_1\). We prove that \((v_i)_{i \leq 1}\) is \(\varepsilon\)-maximal.

Suppose \((u_i)_{i \in \mathbb{Z}}, (v_i)_{i \geq 1}\) are regular and there is an extension of \((v_i)_{i \geq 1}\) to a regular chain \((u_i)_{i \in \mathbb{Z}}\) s.t. \(\pi([u_i]_{i \in \mathbb{Z}}) = \pi([v_i]_{i \in \mathbb{Z}})\). We write \(v_i = \Psi_{x_i}^{x_i}, u_i = \Psi_{y_i}^{y_i}\), and show that \(p_{v_i}^u \geq e^{-\sqrt{Q_u}} q_{y_i}^u\).

Since \(\pi([v_i]_{i \in \mathbb{Z}}) = \pi([v_i]_{i \in \mathbb{Z}})\) 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, \(p_0^v \geq e^{-\sqrt{Q_u}} q_0^v\). Also, by step 1, \(Q_u(x_1) \geq e^{-\sqrt{Q_u}} Q_u(y_1)\). It follows that

\[
\begin{align*}
\pi_1^u &= \min\{e^{v_0^u}, Q_u(x_1)\} \quad \vdots v_0 \to v_1 \\
&\geq \min\{e^{v_0^u}, e^{-\sqrt{Q_u}} Q_u(y_1)\} \\
&= e^{-\sqrt{Q_u}} \min\{e^{v_0^u}, Q_u(y_1)\} = e^{-\sqrt{Q_u}} q_{y_1}^u \quad \vdots u_0 \to u_1.
\end{align*}
\]

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 \(v_i = \Psi_{x_i}^{v_i}, u_i = \Psi_{y_i}^{u_i}\). Since \((v_i)_{i \in \mathbb{Z}}\) is a chain, \((p_i^v, p_i^u)_{i \in \mathbb{Z}}\) is \(\varepsilon\)-subordinated to \(\{Q_u(x_i)\}_{i \in \mathbb{Z}}\). Since \((v_i)_{i \in \mathbb{Z}}\) is regular, \(\limsup_{i \to \pm \infty} (p_i^v, p_i^u) > 0\), therefore by Lemma 3.7, \(p_n^v = Q_u(x_n)\) for some \(n < 0\) and \(p_\ell^v = Q_u(x_\ell)\) for some \(\ell > 0\).

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

\[\Box\]

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 \((\Psi_{x_i}^{v_i}, \Psi_{y_i}^{u_i})_{i \in \mathbb{Z}}\) and \((\Psi_{y_i}^{v_i}, \Psi_{x_i}^{u_i})_{i \in \mathbb{Z}}\) be two regular chains such that \(\pi([\Psi_{x_i}^{v_i}, \Psi_{y_i}^{u_i}]_{i \in \mathbb{Z}}) = \pi([\Psi_{y_i}^{v_i}, \Psi_{x_i}^{u_i}]_{i \in \mathbb{Z}})\), then \(p_i^v/q_i^v, p_i^u/q_i^u \in [\exp(-\sqrt{Q}), \exp(\sqrt{Q})]\) for all \(i \in \mathbb{Z}\).

**Proof.** By Proposition 6.3, \((\Psi_{x_i}^{v_i}, \Psi_{y_i}^{u_i})_{i \leq 0}\) is \(\varepsilon\)-maximal, so \(p_0^v \geq e^{-\sqrt{Q}} q_0^v\). \((\Psi_{y_i}^{v_i}, \Psi_{x_i}^{u_i})_{i \leq 0}\) is also \(\varepsilon\)-maximal, so \(q_0^v \geq e^{-\sqrt{Q}} p_0^v\). It follows that \(p_0^v/q_0^v \in [e^{-\sqrt{Q}}, e^{\sqrt{Q}}]\). Similarly, \(p_\ell^v/q_\ell^v \in [e^{-\sqrt{Q}}, e^{\sqrt{Q}}]\).

Working with the shifted sequences \((\Psi_{x_{i+k}}^{v_i}, \Psi_{y_{i+k}}^{u_i})_{i \in \mathbb{Z}}\) and \((\Psi_{y_{i+k}}^{v_i}, \Psi_{x_{i+k}}^{u_i})_{i \in \mathbb{Z}}\), we obtain \(p_k^v/q_k^v, p_k^u/q_k^u \in [e^{-\sqrt{Q}}, e^{\sqrt{Q}}]\). 

\[\Box\]
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 $\pi([\Psi^u_{x_i}, \Psi^s_{x_i}]_{i \in \mathbb{Z}}) = \pi([\Psi^u_{y_i}, \Psi^s_{y_i}]_{i \in \mathbb{Z}})$ where $(\Psi^u_{x_i}, \Psi^s_{x_i})_{i \in \mathbb{Z}}$ and $(\Psi^u_{y_i}, \Psi^s_{y_i})_{i \in \mathbb{Z}}$ are regular chains. We compare $\Psi_{x_i}$ and $\Psi_{y_i}$. Write, as in §5.3, $\Psi_{x_i} = \exp x_i \circ \theta_{x_i} \circ C_{x_i}$ and $\Psi_{y_i} = \exp y_i \circ \theta_{y_i} \circ C_{y_i}$. We also let $p_i := p^u_i \wedge p^s_i$ and $q_i := q^u_i \wedge q^s_i$.

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

$$C_{y_i}^{-1} C_{x_i} = \begin{pmatrix} s_x(y_i) & -s_x(y_i) / \tan \alpha(y_i) \\ 0 & s_x(y_i) / \tan \alpha(y_i) \end{pmatrix} R_{y_i}^{-1} R_{x_i} \begin{pmatrix} s_x(x_i)^{-1} & 0 \\ u_x(x_i)^{-1} \cos \alpha(x_i) & u_x(x_i)^{-1} \sin \alpha(x_i) \end{pmatrix}$$

$$= \begin{pmatrix} s_x(y_i) & -s_x(y_i) / \tan \alpha(y_i) \\ 0 & s_x(y_i) / \tan \alpha(y_i) \end{pmatrix} (-1)^{\sigma_i} \text{Id} + E' \begin{pmatrix} s_x(x_i)^{-1} & 0 \\ u_x(x_i)^{-1} \cos \alpha(x_i) & u_x(x_i)^{-1} \sin \alpha(x_i) \end{pmatrix},$$

where $\sigma_i \in \{0, 1\}$ and $E' = (\varepsilon_{ij})_{2 \times 2}$ and $|\varepsilon_{ij}| < 2^{\beta/5}$. By Proposition 5.3 says that $s_x(y_i)$ and $u_x(y_i)$ belong to $[\exp(-4\sqrt{\varepsilon}), \exp(4\sqrt{\varepsilon})$, and Proposition 6.5 says that $\sin \alpha(y_i) / \sin \alpha(y_i) \in [\exp(-\sqrt{\varepsilon}), \exp(\sqrt{\varepsilon})$. 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 $u_x(y_i) \geq \sqrt{2} > 1$ and $s_x(y_i) / \sin \alpha(y_i) < \|C_{x_i}(y_i)^{-1}\|_{F^r}$ (Lemma 2.4),

$$\left| \frac{s_x(y_i) \sin \alpha(y_i) - \alpha(x_i)}{u_x(x_i) \sin \alpha(y_i)} \right| \leq \|C_{x_i}(y_i)^{-1}\|_{F^r} \cdot \left| \sin \alpha(y_i) - \alpha(x_i) \right|$$

$$\leq \|C_{x_i}(y_i)^{-1}\|_{F^r} \cdot \left( \left| \sin \alpha(y_i) - \sin \alpha(x_i) \right| + \left| \cos \alpha(y_i) - \cos \alpha(x_i) \right| \right).$$

By Lemma 6.3 if $\varepsilon$ is small enough,

$$\left| \frac{s_x(y_i) \sin \alpha(y_i) - \alpha(x_i)}{u_x(x_i) \sin \alpha(y_i)} \right| \leq \|C_{x_i}(y_i)^{-1}\|_{F^r} \cdot 6(p_i^{\beta/4} + q_i^{\beta/4}).$$

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

$$p_i^{\beta/4} + q_i^{\beta/4} < (e^{\sqrt{\varepsilon}} p_i^{\beta/4} + q_i^{\beta/4}) < 2^{\beta/4} q_i^{\beta/4} < 2Q_i(\varepsilon_i)^{\beta/4} < 2\varepsilon^{\beta/4} \|C_{x_i}(y_i)^{-1}\|_{F^r}.$$
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 \( \|C_x(y_i)\|_{F^r} \) (Lemma 2.3). The term from the middle matrix is bounded by

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

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

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

**Claim 2.** \( \Psi_{yi}^{-1} \circ \Psi_{xi} \) is well defined on \( R_\varepsilon(\mathcal{U}) \).

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

Suppose \( \psi \in R_\varepsilon(\mathcal{U}) \). By Proposition 3.2, \( d(x_i, y_i) < 25^{-1}(p_i + q_i) \), and by Proposition 3.3, \( p_i \leq \varepsilon^{\beta/5} q_i \), so \( d(x_i, y_i) < q_i \). By the definition of \( L_1 \) (page 11),

\[
d((\exp_{x_i} \circ \vartheta_{x_i})(C_{x_i} \psi), (\exp_{y_i} \circ \vartheta_{y_i})(C_{y_i} \psi)) \leq L_1 d(x_i, y_i) < L_1 q_i.
\]

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

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

\[
\exp_{y_i}^{-1}(B) \subset B_{L_1 L_3 q_i}((\vartheta_{y_i}(C_{y_i} \psi)).
\]

It follows that \( \Psi_{x_i}(\psi) \in \exp_{y_i}([\exp_{y_i}^{-1}(B)]) \subset \exp_{y_i}([B_{L_1 L_3 q_i}((\vartheta_{y_i}(C_{y_i} \psi)))] = \Psi_{y_i}([E]), \)

where \( E := C_x(y_i)_{x_i}^{-1}[B_{L_1 L_3 q_i}((\vartheta_{y_i}(C_{x_i} \psi)))] \subset B_{L_1 L_3 \|C_{y_i}^{-1}C_{x_i} \|_{F^r}}([E]). \)

We now use the inequalities \( q_i \leq Q_\varepsilon(y_i) < \varepsilon^{3/5} \|C_x(y_i)^{-1}\|_{F^r} - 1 \) and (claim 1)

\[
\|C_{y_i}^{-1}C_{x_i} - (-1)^{\sigma_i} \text{Id} \|_{F^r} \leq \|C_{y_i}^{-1}C_{x_i} - (-1)^{\sigma_i} \text{Id} \|_{F^r} < 14\sqrt{\varepsilon}.
\]

These give \( E \subset B_{L_1 L_3 \varepsilon^{3/5} + 14\sqrt{\varepsilon} \|\psi\| + \|\psi\|}(((-1)^{\sigma_i} \psi) \subset B_{L_1 L_3 \varepsilon^{3/5} + 14\sqrt{\varepsilon} \|\psi\| + \|\psi\|}(0). \) Since \( \psi \in R_\varepsilon(\mathcal{U}) \), for all \( \varepsilon \) small enough

\[
L_1 L_3 \varepsilon^{3/5} + 14\sqrt{\varepsilon} \|\psi\| + \|\psi\| < (L_1 L_2 \varepsilon^{2} + 14\sqrt{\varepsilon} + 1)\sqrt{\varepsilon} < 2\varepsilon < r(M),
\]

where \( r(M) \) is given in (2.3). It follows that \( E \subset B_{r(M)}(0) \).

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

**Claim 3.** \( \Psi_{y_i}^{-1} \circ \Psi_{x_i}(\psi) = (-1)^{\sigma_i} \psi + c_i + \Delta_i(\psi) \) where \( \sigma_i \in \{0, 1\}, c_i \) is a constant vector s.t. \( \|c_i\| < 10^{-1} q_i \), and \( \Delta_i(\cdot) \) is a vector field s.t. \( \Delta_i(0) = 0 \) and \( \|d\Delta_i(\psi)\| < \sqrt{\varepsilon} \) on \( R_\varepsilon(\mathcal{U}) \).

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

\[
\Psi_{y_i}^{-1} \circ \Psi_{x_i}(\psi) = (-1)^{\sigma_i} \psi + c_i + \Delta_i(\psi)
\]
where \( C_i := (\Psi_i^{-1} \circ \Psi_x)(0) \) and \( \Delta_i(v) := (\Psi_i^{-1} \circ \Psi_x)(v) - (\Psi_i^{-1} \circ \Psi_x)(0) - (-1)^{\sigma_i} \xi_i \).

\[
\Delta_i(v) = [C_i^{-1} \Psi_i^{-1} \exp_{\Psi_i} \Psi_x, \theta_{\xi_i}, C_i] v - \xi_i - (-1)^{\sigma_i} \xi_i
\]

\[
= C_i^{-1} \Psi_i^{-1} \exp_{\Psi_i} \Psi_x, \theta_{\xi_i} - 1 \] \( C_i v + (C_i^{-1} \Psi_i - (-1)^{\sigma_i} \xi_i) v - \xi_i \)

\[
= C_i^{-1} \Psi_i^{-1} \exp_{\Psi_i} \Psi_x, \theta_{\xi_i} (\Psi_i(v)) + (C_i^{-1} \Psi_i - (-1)^{\sigma_i} \xi_i) v - \xi_i.
\]

It is clear that \( \Delta_i(0) = 0 \), and that for all \( v \in R_\epsilon(0) \)

\[
\| (d\Delta_i)_v \| \leq \| C_i^{-1} \| \cdot \| (d\Psi_i^{-1})_{\Psi_x, \theta_{\xi_i}} - (d\Psi_i^{-1})_{\Psi_x, \theta_{\xi_i}}(\Psi_i(v)) \| \| (d\Psi_i)_{\Psi_x, \theta_{\xi_i}} \|
\]

\[
+ \| C_i^{-1} \Psi_i - (-1)^{\sigma_i} \xi_i \| \| (d\Psi_i)_{\Psi_x, \theta_{\xi_i}} \|
\]

\[
\leq 2\| C_i^{-1} \| \cdot \| (d\Psi_i^{-1})_{\Psi_x, \theta_{\xi_i}} - (d\Psi_i^{-1})_{\Psi_x, \theta_{\xi_i}}(\Psi_i(v)) \| + 14\epsilon
\]

\[
\leq 2\| C_i^{-1} \| \cdot \| (d\Psi_i)_{\Psi_x, \theta_{\xi_i}} \| \cdot L_2 d(x, y) + 14\epsilon,
\]

where \( L_2 \) is a common Lipschitz constant for the maps \( x \mapsto \theta_{\xi_i}^{-1} \exp_x \) from \( D \) to \( C^2(D, \mathbb{R}^3) \) \( (D) \). As we saw above, \( d(x_i, y_i) < q_i < \epsilon/\sqrt{\beta} \), whence

\[
\| (d\Delta_i)_v \| \leq 2L_2 \epsilon^{3/\beta} + 14\epsilon.
\]

This is smaller than \( \sqrt{\epsilon} \) for all \( \epsilon \) small enough.

Finally we estimate \( C_i \). Let \( z := f^i(\pi([\Psi_{x_1}^i, \Psi_{x_i}^i]_{i \in \mathbb{Z}})) = f^i(\pi([\Psi_{y_i}^i, \Psi_{y_i}^i]_{i \in \mathbb{Z}})) \). This is the intersection of a \( u \)-admissible manifold and an \( s \)-admissible manifold in \( \Psi_{x_i}^i, \Psi_{x_i}^i \), therefore by Proposition 11.1 \( f^i(z) = \Psi_{x_i}^s \Psi_{y_i}^i(\zeta_i) \), for some \( \zeta_i \in R_{10^{-2}q_i} (0) \).

Similarly, \( z = \Psi_{y_i}^i, \Psi_{y_i}^i(\eta_i) \), for some \( \eta_i \in R_{10^{-2}q_i} (0) \). It follows that

\[
\eta_i = (\Psi_{y_i}^{-1} \circ \Psi_{x_i}) (\zeta_i) = (\Psi_{y_i}^{-1} - (-1)^{\sigma_i}) (\zeta_i) + \Delta_i(\zeta_i),
\]

and consequently \( \| C_i \| \leq \| \eta_i \| + \| \xi_i \| + \| \Delta_i(\zeta_i) \|.
\]

Now \( \| \zeta_i \| < 10^{-2} \sqrt{2} p_i < 10^{-2} \sqrt{2} \epsilon^{3/\beta} q_i, \eta_i < 10^{-2} \sqrt{2} q_i, \) and by the bound on \( \| d\Delta_i \|, \| \Delta_i(\zeta_i) \| \leq \sqrt{\epsilon} \| \zeta_i \| \). It follows that \( \| C_i \| < 10^{-1} q_i. \)

\[ \square \]

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 \( \mathcal{F} \), and a Hölder continuous map \( \pi : \Sigma \to M \) which commutes with the left shift \( \sigma : \Sigma \to \Sigma \), so that \( \pi(\Sigma) \) has full measure w.r.t. any ergodic invariant probability measure with entropy larger than \( \chi \). Moreover, if

\[
\Sigma^\# = \{ u \in \Sigma : u \text{ is a regular chain} \}
\]

\[
= \{ u \in \Sigma : \exists v, w \in \mathcal{F} \exists n_k, m_k \uparrow \infty \text{ s.t. } v_{n_k} = v, v_{-m_k} = w \},
\]

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.** \( \mathcal{Z} := \{ Z(v) : v \in \mathcal{F} \} \), where \( Z(v) := \{ \pi(u) : u \in \Sigma^\#, v_0 = v \}. \)

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 [15]):

\[ \text{this uses the convention from [4] that every element of } \mathcal{F} \text{ is relevant.} \]
Theorem 10.2. For every $Z \in \mathcal{Z}$, $|\{Z' \in \mathcal{Z} : Z' \cap Z \neq \emptyset\}| < \infty$.

Proof. Fix some $Z = Z(\Psi^u_x, \Psi^u_x)$. If $Z' = Z(\Psi^v_y, \Psi^v_y)$ intersects $Z$, then there must exist two chains $\nu, \nu' \in \Sigma^\#$ s.t. $v_0 = \Psi^u_x, w_0 = \Psi^v_y$, and $\pi(\nu) = \pi(\nu')$. Proposition 5.3 says that in this case

$$q^u \geq e^{-\sqrt[p]{p}}\pi^u$$ and $q^v \geq e^{-\sqrt[p]{p}}\pi^v$.

It follows that $Z'$ belongs to $\{Z(\Psi^u_x, \Psi^u_x) : \Psi^v_y, \Psi^v_y \in \mathcal{Y}', q^u \land q^v \geq e^{-\sqrt[p]{p}}(p^u \land p^v)\}$. By the definition of $\mathcal{Y}'$, this set has cardinality less than or equal to

$$(\{\Psi^u_x \in \mathcal{A} : \eta \geq e^{-\sqrt[p]{p}}(p^u \land p^v)\} \times \{(q^u, q^v) \in I_x \times I_z : q^u \land q^v \geq e^{-\sqrt[p]{p}}(p^u \land p^v)\})$$.

This is a finite number, because of the discreteness of $\mathcal{A}$ (Proposition 5.5).

10.2. Product structure. Suppose $x \in Z(v) \in \mathcal{Z}$, then $\exists \nu \in \Sigma^\#$ s.t. $v_0 = v$ and $\pi(\nu) = x$. Associated to $\nu$ are two admissible manifolds in $v$: $V^v[(v_i)_{i \leq 0}]$ and $V^u[(v_i)_{i \geq 0}]$ (Proposition 4.11). These manifolds do not depend on the choice of $\nu$: if $\nu' \in \Sigma^\#$ is another chain s.t. $v_0 = v$ and $\pi(\nu') = x$, then

$$V^v[(v_i)_{i \leq 0}] = V^u[(v_i)_{i \leq 0}] \land V^u[(v_i)_{i \geq 0}] = V^u[(v_i)_{i \geq 0}]$$.

because of Proposition 6.4. 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. $V^x(x, Z) := V^y[(v_i)_{i \geq 0}]$ for some (every) $\nu \in \Sigma^\# \subset v_0 = v$ and $\pi(\nu) = x$.
2. $W^x(x, Z) := V^x(x, Z) \land Z$.

It is important to understand the difference between $V^s/u(x, Z)$ and $W^s/u(x, Z)$. Whereas $V^s/u(x, Z)$ are smooth manifolds, $W^s/u(x, Z)$ could in principle be totally disconnected. Whereas $V^s/u(x, Z)$ extend all the way across $\Psi_x[R_{p/v}\nu]\{0\}$ (assuming $v = \Psi^v_x, \Psi^v_x)$, $W^s/u(x, Z)$ are subsets of the much smaller set $\Psi_x[R_{p/v}\nu\land\{0\}]$, because every point in $W^s/u(x, Z)$ 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 $V^y(x, Z)$ and $V^y(y, Z)$, for $W^u(x, Z)$ and $W^u(y, Z)$, and for $V^y(x, Z)$ and $W^u(y, Z)$.

Proof. The statement holds for $V^u/y$ because of Proposition 6.4. The statement for $W^u/y$ is an immediate corollary.

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

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

We claim that $z \in Z$. There are chains $\nu, \nu' \in \Sigma^\#$ s.t. $v_0 = v_0 = v$ and so that $V^x(x, Z) = V^x[(v_i)_{i \leq 0}]$ and $V^x(x, Z) = V^x[(v_i)_{i \geq 0}]$. Define $u = (u_i)_{i \in \mathbb{Z}}$ by

$$u_i \begin{cases} v_i & i \leq 0 \\ v_i & i \geq 0 \end{cases}.$$
It is easy to see that \( u \in \Sigma^\# \) and \( u_0 = v \), therefore \( \pi(u) \in Z \). By definition, 
\[
\{ \pi(u) \} = V^u[(u_i)_{i \leq 0}] \cap V^s[(w_i)_{i \geq 0}] = V^u[(v_i)_{i \leq 0}] \cap V^s[(w_i)_{i \geq 0}] = V^u(x, Z) \cap V^s(y, Z).
\]
It follows that \( z = \pi(u) \in Z \).

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

Compare with [Sm] or [13A] 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
\[
\{ f(w) \} = f[W^u(x, Y) \cap W^s(y, Y)] \subset f[W^u(x, Y)] \cap f[V^s(y, Y)].
\]

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

**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 \( F_s[V^s] \) in \( v_0 \). Since \( V^s \) stays in windows, \( F_s[V^s] \) stays in windows.

Since \( F_s[V^s] \) 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 \( F_s[V^s] \cap V^u(y, Y) = \{ y \} \), whence \( F_s[V^s] \ni y \).

This means that \( F_s[V^s] \) 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 \( f^{-1}(V^s) \supset F_s[V^s] = V^s(y, Y) \), 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 \( f(w) \in f[V^u(x, Y)] \cap f[V^s(y, Z)] \). By the second half of the claim,
\[
f[V^u(x, Y)] \cap V^s(f(y), Z) \supset V^u(f(x), Z) \cap V^s(f(y), Z) \ni \{ f(x), f(y) \}_{Z(v_1)},
\]
thus \( f[V^u(x, Y)] \cap V^s(f(y), Z) \ni f(w), [f(x), f(y)]_{Z(v_1)} \). 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 \( f(w) = [f(x), f(y)]_{Z(v_1)} \).

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 \( Z' \).

**Proof.** Suppose \( Z = Z(\Psi^{\varepsilon}_{x_0}p^\varepsilon_0), Z' = Z(\Psi^{\varepsilon}_{y_0}p^\varepsilon_0) \) and \( z \in Z \cap Z' \), then there are \( \sigma, \omega \in \Sigma^\# \) s.t. \( v_0 = \Psi^{\varepsilon}_{x_0}p^\varepsilon_0, w_0 = \Psi^{\varepsilon}_{y_0}p^\varepsilon_0, \) and \( z = \pi(\sigma) = \pi(\omega) \). Write \( p := p_0 \wedge p^\varepsilon_0 \) and \( q := q^\varepsilon_0 \wedge q_0 \). By Theorem 8.2, \( p^\varepsilon_0/q_0, p_0/q^\varepsilon_0, p/q \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}] \) and 
\[
\Psi^{-1} \circ \Psi_{x_0} = (-1)^z \Id + \xi + \Delta \text{ on } R_{x_0} \cup \Omega.
\]
where $\sigma \in \{0, 1\}$, $c$ is a constant vector s.t. $\|c\| < 10^{-1}q$, and $\Delta : R_c(0) \to \mathbb{R}^2$ satisfies $\Delta(0) = \mathbf{0}$ and $\|\Delta(u)\| < \sqrt{\varepsilon}$ for all $u \in R_c(0)$. By the Mean Value Theorem, $\|\Delta(u)\| \leq \frac{\sqrt{\varepsilon}}{2}\|u\|$ for all $u \in R_c(0)$.

Now suppose $x \in Z$. $V^u := V^u(x, Z)$ is a $u$-admissible in $\Psi_{p_0}^{n_0}$, therefore it can be put in the form $V^u(x, Z) = \Psi_{\tau, \varepsilon}(F(t), t) : |t| \leq p_0^u$, where $F : [-p_0^u, p_0^u] \to \mathbb{R}$ satisfies $|F(0)| \leq 10^{-3}p_0^u$, $\|F\|_{\infty} \leq 10^{-2}p_0^u$, and $\text{Lip}(F) < \varepsilon$. We write $V^u(x, Z)$ in $\Psi_{\tau, \varepsilon}$-coordinates. Let $\zeta = (c_1, c_2), \Delta = (\Delta_1, \Delta_2)$, then

$V^u(x, Z) = [\Psi_{\tau, \varepsilon} \circ (\Psi_{\tau, \varepsilon} \circ \Psi_{\tau, \varepsilon})\{F(t), t) : |t| \leq p_0^u]\$

$= \Psi_{\tau, \varepsilon}\{(|-1|^n t + c_1 + \Delta_1(F(t), t), (-1)^n t + c_2 + \Delta(t, \theta), \theta) : |\theta| \leq p_0^u\}$

where we have used the transformations $F := (-1)^n t, F := (-1)^n F((-1)^n s)$, and $\Delta_i(t, \theta) := \Delta_i((-1)^n t, (-1)^n \theta)$. Notice that $|F(0)| = |F(0)| \leq 10^{-3}p_0^u$, $\|F\|_{\infty} = |F(0)| \leq 10^{-2}p_0^u$, and $\text{Lip}(F) = \text{Lip}(F) < \varepsilon$. Also $\Delta(0) = 0$ and $\|\Delta(u)\| \leq \sqrt{\varepsilon}$ on $R_c(0)$.

Let $\tau(t) : = \theta + c_2 + \Delta_2(F(t), \theta)$. Assuming $\varepsilon$ is small enough, we have

- $\tau(0) \in [e^{-2\varepsilon}, \varepsilon^2 e^{-2\varepsilon}]$;
- $|\tau(0)| \leq |c_2| + |\Delta_2(\bar{F}(0), 0)| < 10^{-1}q + \sqrt{\varepsilon} \cdot 10^{-3}p < \frac{1}{4}p$ ($\because p \leq e\sqrt{\varepsilon} q$).

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 $|\alpha + p_0^u| < \frac{1}{4}p_0^u$ and $|\beta - p_0^u| < \frac{1}{4}p_0^u$: both quantities are less than $|c_2| + \sup_{R_{p_0^u}(0)} |\Delta_2|$, which is less than $\frac{1}{6}p_0^u$ provided $\varepsilon$ is small enough. It follows that $\tau([-p_0^u, p_0^u]) = [\alpha, \beta] \subset [-\frac{2}{3}q, \frac{2}{3}q]$.

Since $\tau : [-p_0^u, p_0^u] \to [\alpha, \beta]$ is one-to-one and onto, it has a well defined inverse function $\theta : [\alpha, \beta] \to [-p_0^u, p_0^u]$. Let $G(s) = \bar{F}(\theta(s)) + c_1 + \Delta_1(\bar{F}(\theta(s)), \theta(s))$, then

$V^u(x, Z) = \Psi_{\tau, \varepsilon}(G(s), s) : s \in [\alpha, \beta]$.

Using the properties of $\tau$, it is not difficult to check that $\theta' \in [e^{-2\varepsilon}, e^{2\varepsilon}]$ and $|\theta(0)| = |\theta(0) - \theta(\tau(0))| \leq e^{2\sqrt{\varepsilon}} |\theta(0)| < \frac{1}{4}e^{2\sqrt{\varepsilon}}$. It follows that $|\bar{F}(\theta(0))| \leq |\bar{F}(0)| + \varepsilon |\theta(0)| < (10^{-3} + \frac{1}{4}e^{2\sqrt{\varepsilon}}) p < 10^{-2}p_0^u$. Hence

$|G(0)| \leq 10^{-2}p + 10^{-1}q + \sqrt{\varepsilon} p < \min\{\frac{1}{4}p, \frac{1}{4}q\}$ ($\because q/p \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$)

$|G'| \leq \bar{F}'|_{\infty} |\theta'| + \sqrt{\varepsilon} p < \min\{\frac{1}{4}p, \frac{1}{4}q\}$ ($\because q/p \in [e^{-\sqrt{\varepsilon}}, e^{\sqrt{\varepsilon}}]$)

It follows that (for all $\varepsilon$ small enough) $G[-\frac{2}{3}q, \frac{2}{3}q] \subset [-\frac{2}{3}p, \frac{2}{3}p]$.

We can now show that $|V^u(x, Z) \cap V^u(y, Z')| \geq 1$ (compare with [KM S.3.7]). Represent

$V^u(x, Z') = \Psi_{\tau, \varepsilon}(\{t, H(t)\} : |t| \leq p_0^u)$. By admissibility, $|H(0)| < 10^{-3}q$ and $\text{Lip}(H) < \varepsilon$, so $H[-\frac{2}{3}q, \frac{2}{3}p] \subset [-\frac{2}{3}p, \frac{2}{3}q]$. It follows that $H \circ G$ is a contraction of $[-\frac{2}{3}p, \frac{2}{3}q]$ into itself. Such a map has a (unique) fixed point $(H \circ G)(s_0) = s_0$. It is easy to see that $\Psi_{\tau, \varepsilon}(G(s_0), s_0)$ belongs to $V^u(x, Z) \cap V^u(y, Z')$.

Next we claim that $V^u(x, Z) \cap V^u(y, Z')$ contains at most one point. Extend $G$ and $H$ to $\varepsilon$-Lipschitz functions $\tilde{G}$, $\tilde{H}$ on $[-a, a]$ where $a := \max\{|\alpha|, |\beta|, q_0^u\}$. By construction, $|\tilde{G}(0)| \leq \frac{1}{a}$, 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 \( \Psi_{v_0}(G(s), s) \) where \( s \in [\alpha, \beta] \) and \( s = (H \circ G)(s) = (\tilde{H} \circ \tilde{G})(s) \). 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 \( |V^u(x, Z) \cap V^s(y, Z')| \leq 1 \). \( \Box \)

10.3. The symbolic Markov property.

**Proposition 10.9.** If \( x = \pi((v_i)_{i \in \mathbb{Z}}) \) where \( v \in \Sigma^# \), 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, \( f(V^s([v_i]_{i \geq 0})) \subset V^s([v_{i+1}]_{i \geq 0}) \). 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.** \( f[W^u(x, Z(v_0))] \subset Z(v_1) \).

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

- Since \( y \in Z(v_0) \), \( y \in \Psi_{z_0}[R_{10-\epsilon}^{\beta} \wedge \tilde{a}_0^{\beta}](\tilde{1}) \) (it is the intersection of a \( u \) and an \( s \)-admissible manifolds in \( v_0 \)).
- Since \( y \in V^u([v_i]_{i \geq 0}) \), \( f^k(y) \in V^u([v_{i+k}]_{i \geq 0}) \) \( \subset \Psi_{z_k}[R_{Q_{\epsilon^t}(z_k)}(\tilde{1})] \) for all \( k \geq 0 \), where \( v_k = \Psi_{z_k}^{\beta} \cdot v_k \).
- Since \( y \in Z(v_0) \), \( \exists w \in \Sigma^# \) s.t. \( w_0 = v_0 \) and \( y = \pi(w) \in V^u([w_i]_{i \leq 0}) \).

It follows that \( f^{-k}(y) \in V^u([w_{i-k}]_{i \leq 0}) \) \( \subset \Psi_{y_k}[R_{Q_{\epsilon^t}(y_k)}(\tilde{1})] \) for all \( k \geq 0 \), where \( w_i = \Psi_{y_i}^{\beta} \cdot q_i \).

Writing \( u_i = \begin{cases} u_i & i \leq 0 \\ u_i & i > 0 \end{cases} \) and \( u_i = \Psi_{z_i}^{\beta} \cdot q_i \), we see that \( u \in \Sigma^# \), \( u_0 = v_0 \), \( y \in \Psi_{z_0}[R_{p_0^{\beta} \wedge q_0^{\beta}}(\tilde{1})] \), and \( f^k(y) \in \Psi_{z_k}[R_{Q_{\epsilon^t}(z_k)}(\tilde{1})] \) for all \( k \in \mathbb{Z} \). By Proposition 4.15 part (4), \( y = \pi(u) \). It follows that \( f(y) = \pi(\sigma(u)) \in Z(z_1) \equiv Z(v_1) \). \( \Box \)

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

1. If \( Z = Z(\Psi_{z_0}^{\beta} \wedge \tilde{a}_0^{\beta}) \) and \( Z' = Z(\Psi_{y_0}^{\beta} \wedge \tilde{a}_0^{\beta}) \), then \( Z \subset \Psi_{y_0}[R_{Q_{\epsilon^t}(y_0)}(\tilde{1})] \).
2. For any \( x \in Z \cap Z' \), \( W^u(x, Z) \subset V^s(x, Z') \) and \( W^s(x, Z) \subset V^u(x, Z') \).

**Proof.** Fix some \( x \in Z \cap Z' \). Write \( x = \pi(u) \), \( x = \pi(w) \) where \( u, w \in \Sigma^# \) satisfy \( v_0 = \Psi_{x_0}^{\beta} \wedge \tilde{a}_0^{\beta} \) and \( w_0 = \Psi_{y_0}^{\beta} \wedge \tilde{a}_0^{\beta} \). Write \( p := p_0 \wedge q_0^{\beta} \) and \( q := q_0^{\beta} \wedge \tilde{a}_0^{\beta} \). Since \( \pi(u) = \pi(w) \), we have by Theorem 5.2 that \( p/q \in [e^{-\sqrt{c}}, e^{\sqrt{c}}] \) and

\[
\Psi_{y_0}^{-1} \circ \Psi_{x_0} = (-1)^{\sigma} \text{Id} + \mathcal{C} + \Delta \text{ on } R_{\epsilon}(\tilde{1}),
\]

where \( \sigma \in \{0, 1\} \), \( \mathcal{C} \) is a constant vector s.t. \( \|\mathcal{C}\| < 10^{-1}q \), and \( \Delta : R_{\epsilon}(\tilde{1}) \to \mathbb{R}^2 \) satisfies \( \Delta(\tilde{1}) = \mathcal{0} \) and \( \|\Delta(\mathcal{u})\| < \sqrt{\mathcal{C}} \) for all \( \mathcal{u} \in R_{\epsilon}(\tilde{1}) \). By the Mean Value Theorem, \( \|\Delta(\mathcal{u})\| \leq \sqrt{\mathcal{C}} \|\mathcal{u}\| \) for all \( \mathcal{u} \in R_{\epsilon}(\tilde{1}) \).
Every point in $Z$ is the intersection of a $u$–admissible and an $s$–admissible manifold in $\Psi_{x_0}^{y_0}$, therefore $Z$ is contained in $\Psi_{x_0}[R_{10−2p}(0)]$ (Proposition 11.1). Thus
\[
Z \subseteq \Psi_{y_0}[(\Psi_{y_0}^{-1} \circ \Psi_{x_0})[R_{10−2p}(0)]] \subset \Psi_{y_0}[(\Psi_{y_0}^{-1} \circ \Psi_{x_0})[B_{\sqrt{10−2p}(0)}]] \\
\subseteq \Psi_{y_0}[B_{(1+\sqrt{3})\sqrt{10−2p}(0)}] \subseteq \Psi_{y_0}[B_{(1+\sqrt{3})\sqrt{10−2p}\epsilon\tau q+10−1q}(0)] \\
\subseteq \Psi_{y_0}[R_q(0)] \quad (: 0 < \epsilon < 1).
\]
This proves the first statement of the lemma.

Next we show that $W^s(x, Z) \subset V^s(x, Z')$. Write $v_i = \Psi_{y_i}^{x_i}$ and $w_i = \Psi_{y_i}^{x_i}$. Since $x = \pi(y)$ 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, $Z(v_k) \subset \Psi_{y_k}[R_{q_k}(0) \wedge q_k(0)]$. It follows that
\[
f^k[W^s(x, Z)] \subset \Psi_{y_k}[R_{q_k}(0) \wedge q_k(0)] \subset \Psi_{y_k}[R_{q_k}(0)],
\]
for all $k \geq 0$. By Proposition 11.5 part 4, $W^s(x, Z) \subset V^s[(w_i)_{i \geq 0}] \equiv V^s(x, Z').$ □

11. A countable Markov partition

In the previous section we described a locally finite countable cover $\mathcal{Z}$ of NUH$_{\chi}^f$ by sets equipped with a Smale bracket and satisfying the symbolic Markov property (Proposition 10.11). Here we produce a pairwise disjoint cover of NUH$_{\chi}^f$ with similar properties.

Sinai and Bowen showed how to do this in the case of finite covers [SI, 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 [SI] and [B4].

11.1. The Bowen–Sinai refinement. Write $\mathcal{Z} = \{Z_1, Z_2, Z_3, \ldots \}$. Following [B4], we define for every $Z_i, Z_j \in \mathcal{Z}$ s.t. $Z_i \cap Z_j \neq \emptyset$,
\[
T_{ij}^u := \{x \in Z_i : W^u(x, Z_i) \cap Z_j \neq \emptyset, W^s(x, Z_i) \cap Z_j \neq \emptyset\},
\]
\[
T_{ij}^s := \{x \in Z_i : W^s(x, Z_i) \cap Z_j \neq \emptyset, W^s(x, Z_i) \cap Z_j = \emptyset\}.
\]
Let $\mathcal{I} := \{T_{ij}^\beta : i, j \in \mathbb{N}, Z_i \cap Z_j \neq \emptyset, \alpha \in \{u, s\}, \beta \in \{s, \emptyset\}\}$. Notice that $T_{ii}^u = Z_i$, therefore $\mathcal{I}$ covers the same set as $\mathcal{Z}$, namely $\pi(\Sigma^#)$. Another useful identity is $T_{ij}^u = Z_i \cap Z_j$. The inclusion $\subseteq$ is trivial. To see $\supseteq$ suppose $x \in T_{ij}^u$. 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 $R(x) := \bigcap \{T \in \mathcal{I} : T \ni x\}$, and set $\mathcal{R} := \{R(x) : x \in \pi(\Sigma^#)\}$.
Proposition 11.2. $R$ is a countable pairwise disjoint cover of $\text{NUH}^\#(f)$.

Proof. We claim that each $R(x)$ is a finite intersection. By the finiteness property of $\mathcal{Z}$ (Theorem 10.2), there are at most finitely many $Z_i \in \mathcal{Z}$ which contain $x$. Again by Theorem 10.2, for every $Z_i \in \mathcal{Z}$ which contains $x$, there are at most finitely many $Z_j \in \mathcal{Z}$ which intersect $Z_i$. As a result, there are at most finitely many $T \in \mathcal{F}$ which contain $x$. Thus $R(x)$ is a finite intersection.

Since there are countably many finite subsets of $\mathcal{F}$, there are countably many elements in $\mathcal{R}$.

Since every $x \in T \in \mathcal{F}$ belongs to $R(x) \in \mathcal{R}$, $\bigcup \mathcal{R} = \bigcup \mathcal{F}$. We saw above that for every $Z_i \in \mathcal{Z}$, $T_i^{us} = Z_i$. Consequently, $\bigcup \mathcal{F} = \bigcup \mathcal{F} = \pi(\Sigma^\#)$. Since $\pi(\Sigma^\#) \supset \text{NUH}^\#(f)$ (see the proof of Theorem 11.10), $\mathcal{R}$ covers $\text{NUH}^\#(f)$.

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 $\bigcup \mathcal{R}$:

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

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 $\mathcal{F}$. 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 $Z_i = T_i^{us}$.

If $y \in Z_i$, then $T_i^{us}$ is one of the sets in the intersection which defines $R(y)$. Consequently, $x \in R(y) \subseteq T_i^{us} = Z_i$, 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. $T_{ki}^{\alpha\beta} \ni y$ 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 $x \in Z_k \cap Z_i \equiv T_i^{us}$, so necessarily $(\alpha, \beta) = (u, s)$. Thus $y \in T_i^{us} = Z_k \cap Z_i \subset Z_i$. 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 $x \in T_i^{us}$, 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 $T_{ij}^{\alpha\beta} \cap T_{ij}^{\alpha\beta} = \emptyset$, $\alpha = u$. It follows that $y \in T_{ij}^{us}$, 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 $y \in T_{ij}^{us}$, whence $x \in R(y) \subseteq T_{ij}^{us}$, 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. □

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 $Z = T_i^{us}$. Since $x \in Z$, $T_i^{us}$ appears in the intersection which defines $R(x)$, therefore $R(x) \subset T_i^{us}$. Since
$x \in R$, $R$ intersects $R(x)$, and therefore by the previous proposition $R = R(x)$. It follows that $R = R(x) \subset T^u_{ii} = Z$, which proves the first part of the proposition.

We turn to the second part. If $R \subset Z$, then $R$ is the intersection of a subset of $\mathcal{F}(Z) := \{ T^{\alpha\beta}_{ij} \in \mathcal{F} : T^{\alpha\beta}_{ij} \cap Z \neq \emptyset \}$. If $T^{\alpha\beta}_{ij} \cap Z \neq \emptyset$, 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 $\mathcal{F}(Z)$ is finite.

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

11.2. Product structure and hyperbolicity.

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

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

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

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

1. Either $W^s(x, R)$ and $W^u(y, R)$ are disjoint. Similarly for $W^u(x, R)$ and $W^s(y, R)$.
2. If $x, y \in W^s(x, R)$, then $d(f^n(x), f^n(y)) \rightarrow 0$. If $x, y \in W^u(x, R)$, then $d(f^{-n}(x), f^{-n}(y)) \rightarrow 0$.

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

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

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

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

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

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

Fix $x, y \in R$. For every $T^{\alpha\beta}_{ij} \in \mathcal{F}$ which contains $R$, $W^u(x, Z_i) \cap W^u(y, Z_i) = \{ z_i \}$ where $z_i := [x, y, Z_i]$. By Proposition 10.3, $W^u(z_i, Z_i) = W^u(x, Z_i)$ and $W^s(z_i, Z_i) = W^s(y, Z_i)$. It follows that $z_i \in T^{\alpha\beta}_{ij}$, whence

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

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 $T_{ij}^{\alpha\beta} \in \mathcal{T}$ also covers $R$, then $x, y \in Z_i \cap Z_k$ and so
\[
\{ z_i \} = W^u(x, Z_i) \cap W^s(y, Z_i) \subset V^u(x, Z_i) \cap V^s(y, Z_i)
\]
\[
\{ z_k \} = W^u(x, Z_k) \cap W^s(y, Z_k) \subset V^u(x, Z_i) \cap V^s(y, Z_i) \quad \text{(Lemma 11.10)}.
\]
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 covers $R$. 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 $x = \pi(y)$ where $y$ 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 11.11.1.

Given $x, y \in R$, we let $[x, y]$ denote the unique element of $W^u(x, R) \cap W^s(x, R)$. As the proof of the previous proposition shows, $[x, y]$ is equal to the Smale bracket of $x$ and $y$ in any of the $Z \in \mathcal{Z}$ which contain $R$.

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

**Proposition 11.6.** Let $R_0, R_1 \in \mathcal{R}$. If $x \in R_0$ and $f(x) \in R_1$, then $f[W^s(x, R_0)] \subset W^s(f(x), R_1)$ and $f^{-1}[W^u(f(x), R_1)] \subset W^u(x, R_0)$.

**Proof.** The proof is the same as Bowen’s [Bo4] pages 54, 55, except that our “rectangles” $R \in \mathcal{R}$ are different. 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 $f(x) = \pi(\sigma v)$ where $v \in \Sigma#$ satisfies $v_0 = v$. Since $f \circ \pi = \pi \circ \sigma$, $x = \pi(y) \in Z(v_0)$. 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 11.8),
\[
f[W^s(x, Z(v_0))] \subset W^s[f(x), Z(v_1)],
\]
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 $f(x) \in T_{ij}^{\alpha\beta} \iff f(y) \in T_{ij}^{\alpha\beta}$.

Since $y \in W^s(x, R_0) \iff W^s(x, R_0) = W^s(y, R_0)$, 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. $Z_i \cap Z_j \neq \emptyset$,
- $f(x) \in Z_i \iff f(y) \in Z_i$;
- $W^s(f(x), Z_i) \cap Z_j \neq \emptyset \iff W^s(f(y), Z_i) \cap Z_j \neq \emptyset$;
- $W^u(f(x), Z_i) \cap Z_j \neq \emptyset \iff W^u(f(y), Z_i) \cap Z_j \neq \emptyset$.

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

\[
f(y) \in f[W^s(y, R_0)] = f[W^s(x, R_0)] \quad \text{(by assumption)}
\]

\[
\subseteq f[W^s(x, Z(v_0))] \subseteq W^s(f(x), Z(v_1)) \equiv W^s(f(x), Z_i).
\]

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 \( Z_i = Z, \) \( Z_j = Z^*, \) and prove that \( W^u(f(x), Z) \cap Z^* \neq \emptyset \Rightarrow W^u(f(y), Z) \cap Z^* \neq \emptyset. \) We do this by picking some \( f(z) \in W^u(f(x), Z) \cap Z^* \) and showing that \( W^u(f(y), Z) \cap Z^* \ni f(w) \) where \( w := [y, z]_Y \) for some suitable \( Y \in \mathcal{Y} \) that we proceed to construct.

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

Since \( f(z) \in W^u(f(x), Z) \cap Z^* \) and \( f(z) \in Z^*. \) This means that there exists \( u^* \in \Sigma^\# \) such that \( \pi(\sigma u^*) = f(z) \) and \( Z^* = Z(v_0^*). \) Let \( Y^* := Z(v_0^*), \) then \( z = \pi(\sigma u^*) \in Y^*. \)

By the symbolic Markov property,

\[
z \in f^{-1}[W^u(f(x), Z)] \equiv f^{-1}[W^u(f(x), Z(v_1))] \subseteq W^u(x, Z(v_0)) \equiv W^u(x, Y).
\]

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

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

Thus \( w = [y', z]_Y \in W^u(z, Y^*). \) Now \( Y^* = Z(v_0^*), Z^* = Z(v_1^*) \) and \( z = \pi(z^*), \) therefore by the symbolic Markov property,

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

It follows that \( f(w) \in Z^*. \) This completes the proof of Step 3. The theorem 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 \( \Sigma(\mathcal{F}) \) where \( \mathcal{F} \) is the directed graph with vertices \( \mathcal{V} := \mathcal{R} \) and edges

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

If \( (R_1, R_2) \in \hat{\mathcal{E}}, \) then we write \( R_1 \to R_2. \)
For every finite path $R_m \to R_{m+1} \to \cdots \to R_n$ in $\mathcal{G}$, let $\ell[R_m,\ldots,R_n] := \bigcap_{k=\ell}^{\ell+n-m} f^{-k}(R_{k+m-\ell})$. In particular,

$$m[R_m,\ldots,R_n] = \bigcap_{k=m}^{n} f^{-k}(R_k).$$

**Lemma 12.1.** Suppose $m \leq n$ and $R_m \to R_{m+1} \to \cdots \to R_n$ is a finite path on $\mathcal{G}$, then $m[R_m,\ldots,R_n] \neq \emptyset$.

**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 $R_m \to \cdots \to R_{n-1}$ be a path on $\mathcal{G}$. By the induction hypothesis, $m[R_m,\ldots,R_{n-1}] \neq \emptyset$, therefore there exists a point $y \in \bigcap_{k=m}^{n-1} f^{-k}(R_k)$. Since $R_{n-1} \to R_n$, 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 $x \in m[R_m,\ldots,R_n]$. This follows from the Markov property (Theorem 11.18).

- $f^n(x) \in R_n$, 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 $\mathcal{G}$ to the paths on $\mathcal{G}$ (the graph we introduced in [4]). Recall the map $\pi : \Sigma \to M$ from Theorem 11.10 and define for any finite path $v_m \to v_n$ on $\mathcal{G}$,

$$Z_m(v_m,\ldots,v_n) := \{\pi(w) : w \in \Sigma^\#, w_i = v_i \text{ for all } i = m,\ldots,n\}.$$

**Lemma 12.2.** For every infinite path $\cdots \to R_i \to R_{i+1} \to \cdots$ in $\mathcal{G}$ there exists a chain $(v_i)_{i \in \mathbb{Z}} \subset \Sigma$ such that for every $i$, $R_i \subset Z(v_i)$, and for every $n$, $-n[R_{-n},\ldots,R_n] \subset Z_{-n}(v_{-n},\ldots,v_n)$.

**Proof.** Fix, using Lemma 12.1 points $y_n \in [-n[R_{-n},\ldots,R_n]$.

Pick some $v_0 \in \mathcal{G}$ s.t. $R_0 \subset Z(v_0)$. Since $y_n \in R_0$, there is a chain $v^{(n)} = (v^{(n)}_i)_{i \in \mathbb{Z}} \subset \Sigma$ such that $v^{(n)}_i = v_i$ and $y_n = \pi [v^{(n)}]$.

For every $|k| \leq n$, $f^k(y_n) = \pi [\sigma^k(v^{(n)}_i)] \in Z(v^{(n)}_k)$, therefore $Z(v^{(n)}_k)$ 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 $R_k \subset Z(v^{(n)}_k)$ for every $k = -n,\ldots,n$.

Every vertex in the graph $\mathcal{G}$ has finite degree (Lemma 11.14). 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 $\{v^{(n)}_k : n \in \mathbb{N}\}$ 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 \( \nu := (v_k)_{k \in \mathbb{Z}} \) is a chain, and \( R_k \subset Z(v_k) \) for all \( k \in \mathbb{Z} \). We claim that \( n[R_{-n}, \ldots, R_n] \subset Z_{-n}(v_{-n}, \ldots, v_n) \) for all \( n \).

Suppose \( y \in _{-n}[R_{-n}, \ldots, R_n] \). Since \( f^n(y) \in R_n \) and \( R_n \subset Z(v_n) \), there exists a chain \( w \in \Sigma^\# \) s.t. \( f^n(y) = \pi[\sigma^n(w)] \) and \( w_n = v_n \). Since \( f^{-n}(y) \in R_{-n} \) and \( R_{-n} \subset Z(v_{-n}) \), there exists a chain \( w \in \Sigma^\# \) s.t. \( f^{-n}(y) = \pi[\sigma^{-n}(w)] \) and \( u_{-n} = v_n \). Let

\[
\mathcal{A} = (a_i)_{i \in \mathbb{Z}} \text{ where } a_i = \begin{cases} u_i & i \leq -n \\ v_i & -n \leq i \leq n \\ w_i & i \geq n. \end{cases}
\]

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 \( a_i = \Psi_{x_l}[R_{Q_l(x_l)}(1)] \), we see that \( y \in \Sigma_{x_l}[R_{Q_l(x_l)}(1)] \) for all \( l \in \mathbb{Z} \). By Proposition 12.3, \( y \in V^s[[a_i]_{i \leq 0}] \cap V^s[[a_i]_{i \geq 0}] \), so \( y = \pi(a) \in \Omega_{n}(v_{-n}, \ldots, v_n) \). \( \square \)

**Proposition 12.3.** Every vertex of \( \hat{G} \) has finite degree.

**Proof.** Fix \( R_0 \in \mathcal{A} \). We bound the number of paths \( R_{-1} \to R_0 \to R_1 \).

Consider all the possible paths \( v_{-1} \to v_0 \to v_1 \) on \( G \) 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 \( \mathcal{A} \) has the finiteness property (Theorem 10.2). Since every vertex of \( G \) 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

\[
\hat{\Sigma} := \Sigma(\hat{G}) = \{(R_i)_{i \in \mathbb{Z}} \in \mathcal{A}^\mathbb{Z} : R_i \to R_{i+1} \text{ for all } i \in \mathbb{Z} \}.
\]

Abusing notation, we denote the left shift map on \( \hat{\Sigma} \) by \( \sigma \), and the natural metric on \( \hat{\Sigma} \) by \( d(\cdot, \cdot) : d(\hat{\Sigma}, \overline{y}) = \exp[-\min\{|k| : x_k \neq y_k\}] \). Since every vertex of \( G \) has finite degree, \( \hat{\Sigma} \) is locally compact. Define as before

\[
\hat{\Sigma}^\# := \{(R_i)_{i \in \mathbb{Z}} : \exists R, S \in \mathcal{A}, \exists n, m \uparrow \infty \text{ s.t. } R_{m_k} = R \text{ and } R_{-m_k} = S\}.
\]

Clearly \( \hat{\Sigma}^\# \) contains every periodic point for \( \sigma \). By Poincaré’s Recurrence Theorem, every \( \sigma \)-invariant probability measure on \( \hat{\Sigma} \) is supported on \( \hat{\Sigma}^\# \).

Our aim is to construct a finite-to-one Hölder continuous map \( \hat{\pi} : \hat{\Sigma} \to M \) which intertwines \( \sigma \) and \( f \), and such that \( \hat{\pi}(\hat{\Sigma}) \) (and even \( \hat{\pi}(\hat{\Sigma}^\#) \)) 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 \( (R_i)_{i \in \mathbb{Z}} \in \hat{\Sigma} \),

\[
\text{diam}(-n[R_{-n}, \ldots, R_n]) < C\theta^n.
\]

**Proof.** Recall that \( \pi : \Sigma \to M \) is Hölder continuous, therefore there are \( C \) and \( 0 < \theta < 1 \) s.t. for every \( v, v \in \Sigma, \) if \( v_i = u_i \) for all \( |i| \leq n \) then \( d(\pi(v), \pi(v)) < C\theta^n \).

By Lemma 12.2 there exists a chain \( (v_i)_{i \in \mathbb{Z}} \in \Sigma \) s.t.

\[
-n[R_{-n}, \ldots, R_n] \subset Z_{-n}(v_{-n}, \ldots, v_n).
\]
The diameter of \( Z_{-n}(v_{-n}, \ldots, v_n) \) is less than or equal to \( C\theta^n \). Therefore the diameter of \([-n[R_{-n}, \ldots, R_n]\) is less than or equal to \( C\theta^n \). \( \square \)

Suppose \((R_i)_{i \in \mathbb{Z}} \in \hat{\Sigma}\), and let \( F_n := -n[R_{-n}, \ldots, R_n] \) (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 \( \hat{\pi}[(R_i)_{i \in \mathbb{Z}}] \):

\[
\{\hat{\pi}[(R_i)_{i \in \mathbb{Z}}]\} = \bigcap_{n=0}^{\infty} -n[R_{-n}, \ldots, R_n]
\]

**Theorem 12.5.** \( \hat{\pi} : \hat{\Sigma} \to M \) has the following properties:

1. \( \hat{\pi} \circ \sigma = f \circ \hat{\pi} \);
2. \( \hat{\pi} \) is Hölder continuous;
3. \( \hat{\pi}(\hat{\Sigma}) \supset \hat{\pi}(\hat{\Sigma}^\#) \supset \text{NUH}_\chi^\#(f) \), therefore the image of \( \hat{\pi} \) has full measure w.r.t every ergodic invariant probability measure with entropy larger than \( \chi \).

**Proof.** The commutation relation is because for every \( R = (R_i)_{i \in \mathbb{Z}} \) in \( \hat{\Sigma} \),

\[
\{\pi(\sigma(R))\} = \bigcap_{n=0}^{\infty} -n[R_{-n+1}, \ldots, R_{n+1}] \supset \bigcap_{n=0}^{\infty} -n[R_{-n}, \ldots, R_n]
\]

\[
= \bigcap_{n=0}^{\infty} \bigcap_{k=-n-2}^{n} f^{-k}(R_{k+1}) = \bigcap_{N=0}^{\infty} f(-N[R_{-N}, \ldots, R_N])
\]

\[
= \bigcap_{N=0}^{\infty} f\left(-N[R_{-N}, \ldots, R_N]\right), \text{ because } f \text{ is a homeomorphism}
\]

\[
= f\left(\bigcap_{N=0}^{\infty} -N[R_{-N}, \ldots, R_N]\right), \text{ because } f \text{ is a bijection}
\]

\[
\equiv f(\{\pi(R)\}) = \{f(\pi(R))\}.
\]

The Hölder continuity of \( \pi \) is because if \( R, S \in \hat{\Sigma} \) and \( R_i = S_i \) for all \( |i| \leq N \), then \( \hat{\pi}(R), \hat{\pi}(S) \in -N[R_{-N}, \ldots, R_N] \), whence by Lemma \[12.4\]

\[
d(\hat{\pi}(R), \hat{\pi}(S)) \leq \text{diam}(-N[R_{-N}, \ldots, R_N]) \leq C\theta^N.
\]

Finally we claim that \( \hat{\pi}(\hat{\Sigma}) \) and \( \hat{\pi}(\hat{\Sigma}^\#) \) 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 \( \underline{\pi} \in \Sigma^\# \) s.t. \( \pi(\underline{\pi}) = x \). \( \Sigma^\# \) 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 \( \mathcal{F} \) covers \( \pi(\Sigma^\#) \), therefore for every \( i \in \mathbb{Z} \) there is some \( R_i \in \mathcal{F} \) s.t. \( f^i(x) \in R_i \). Obviously \( R_i \to R_{i+1} \), so \( R := (R_i)_{i \in \mathbb{Z}} \) belongs to \( \hat{\Sigma} \). Also,

\[
x \in \bigcap_{n=0}^{\infty} -n[R_{-n}, \ldots, R_n]
\]

(even without the closure), so \( x = \pi(R) \). It follows that \( \hat{\pi}(\hat{\Sigma}) \supset \text{NUH}_\chi^\#(f) \).

We claim that the sequence \( R \) which was constructed above belongs to \( \hat{\Sigma}^\# \), and deduce that \( \hat{\pi}(\hat{\Sigma}^\#) \supset \text{NUH}_\chi^\#(f) \).

The sequence \( v \) 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 4.11.3 \( R_i \subset Z(v) \) 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 4.11.3). Therefore \( \exists m_k \uparrow \infty \) 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 \( \overline{R} \in \tilde{\Sigma}^\pi \) as required.

The following result is not needed for the purposes of this paper, but we anticipate some future applications.

**Proposition 12.2.1.** For every \( x \in \hat{\pi} (\hat{\Sigma}) \), \( T_x M = E^s(x) \oplus E^u(x) \) where

(a) \( \limsup_{n \to \infty} \frac{1}{n} \log \| df^n_x \| f^n(x) \leq -\frac{\lambda}{2} \) on \( E^s(x) \setminus \{ 0 \} \);

(b) \( \limsup_{n \to \infty} \frac{1}{n} \log \| df^{-n}_x \| f^{-n}(x) \leq -\frac{\lambda}{2} \) on \( E^u(x) \setminus \{ 0 \} \).

The maps \( \mathcal{R} \mapsto E^u/\mathcal{R}(\mathcal{R}) \) are Hölder continuous as maps from \( \tilde{\Sigma} \) to \( TM \).

**Proof.** Suppose \( x = \tilde{\pi} (R) \) where \( R \in \tilde{\mathcal{G}} \). By Lemma 4.2.2 there is a chain \( (v_i)_{i \in \mathbb{Z}} \) s.t. \( R_i \subset Z(v_i) \) for all \( i \in \mathbb{Z} \), and \( R_{n} \subset \mathbb{Z} (v_{n}, \ldots, v_n) \) for every \( n \). Then \( f^n(x) \in \tilde{\mathcal{F}}(v_n) \) for all \( n \). Every element of \( Z(v_n) \) is the intersection of \( s/u \)-admissible manifolds in \( v_n \), so if \( v_n = \Psi_{x_n}^{\pi^n} \), then \( Z(v_n) \subset \Psi_{x_n}^{\pi^n} [R_{\pi^n} \setminus \tilde{\mathcal{G}} (\bar{u})] \) (Proposition 4.11.1 (2)). By Proposition 4.11.1 (4), \( x \in V^s([v_i]_{i \leq 0}) \cap V^u ([v_i]_{i \geq 0}) \).

Let \( E^s(x) : = T_x V^s ([v_i]_{i \leq 0}) \) 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 4.1.3). 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 \( x = \tilde{\pi} (\bar{R}) \) and \( y = \tilde{\pi} (\bar{S}) \) where \( R_i = S_i \) for \( i = -N, \ldots, N \), and let \( \bar{v} = (v_i)_{i \in \mathbb{Z}} \) be as before. The argument in the first paragraph shows that \( x = \pi (\bar{v}) \). We claim that \( y = \pi (\bar{w}) \) where \( \bar{w} \) is a chain s.t. \( w_i = v_i \) for all \( |i| \leq N \).

By assumption, \( y \in \mathbb{Z}(S_{-N}, \ldots, S_N) = \mathbb{Z}(R_{-N}, \ldots, R_N) \subset \mathbb{Z}(v_{-N}, \ldots, v_N) \), so \( y = \lim \pi (w^{(n)}) \) where \( w^{(n)} \in \Sigma \) satisfies \( w^{(n)}_i = v_i \) 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 \( w^{(n_k)} \to \bar{w} \). The limit is a chain \( w \) s.t. \( y = \pi (\bar{w}) \) and \( w_i = v_i \) for all \( |i| \leq N \).

Write \( v_0 = \Psi_{x_0}^{\pi_0} \), 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 \( V^u ([w_i]_{i \leq 0}), V^s ([w_i]_{i \geq 0}) \). By Proposition 4.1.5 (5),

\[
\| F_u - G_u \| + \| F'_u - G'_u \| < K \theta^N \\
\| F_s - G_s \| + \| F'_s - G'_s \| < K \theta^N
\]

(12.1)

for some global constants \( K > 0, \theta \in (0, 1) \).

The intersection of the (vertical) graph of \( F_u \) and the (horizontal) graph of \( F_s \) is the point \( \xi \in \mathbb{R}^2 \) s.t. \( \Psi_{x_0} (\xi) = x \). The intersection of the vertical and horizontal graphs of \( G_u \) and \( G_s \) is the point \( \eta \in \mathbb{R}^2 \) s.t. \( \Psi_{x_0} (\eta) = y \). (12.1) and Proposition 4.1.1 imply that \( \| \xi - \eta \| < 3K \theta^N \).
By admissibility, $F_u, F_s, G_u, G_s$ have $\frac{\alpha}{3}$ Hölder exponent at most $\frac{1}{3}$. Together with (12.1), this implies $|F_u'(z_1) - G_s'(q)|, |F_s'(z_2) - G_u'(p)| < \frac{1}{3}(3K\theta N)^{\beta/3} + K\theta N$. It follows that $\text{dist}_{T^2}(T_x[\text{graph}(F_s)], T_y[\text{graph}(G_s)]) = O(\theta N^{\alpha/3})$.

$E^\alpha(x), E^\alpha(y)$ are the images of $T_x[\text{graph}(F_s)]$ and $T_y[\text{graph}(G_s)]$ under $d\Psi_{x_0}$. Since the norm of the differential of a Pesin chart is bounded above by two, $\text{dist}_{TM}(E^\alpha(x), E^\alpha(y)) = O(\theta N^{\alpha/3})$. Similarly, $\text{dist}_{TM}(E^\alpha_i(x), E^\alpha_i(y)) = O(\theta N^{\alpha/3})$. All implied constants are uniform, so $R \mapsto E^{i/u}(\hat{\pi}(R))$ are Hölder continuous.

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 12.6. $N(R) < \infty$.

Proof. Suppose $R \in \mathcal{R}$. The set $A(R) := \{Z \in \mathcal{Z} : Z \supseteq R\}$ 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 $N(R) = \sum_{R' \in C(R)}|A(R')| < \infty$.

Theorem 12.7. Every $x \in \hat{\pi}(\hat{\Sigma})$ has a finite number of $\hat{\pi}$–pre-images. If $x = \hat{\pi}(R)$ where $R_i = R$ for infinitely many $i < 0$ and $R_i = S$ for infinitely many $i > 0$, then $|\hat{\pi}^{-1}(x)| < \varphi(R, S) := N(R)N(S)$.

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 $x \in \hat{\pi}(\hat{\Sigma})$, then $x$ has a $\hat{\pi}$–preimage $R \in \hat{\Sigma}$ s.t. $R_i = R$ for infinitely many negative $i$, and $R_i =$ $S$ for infinitely many positive $i$. We show that the number of $\hat{\pi}$–pre-images of $x$ is less than or equal to $N := N(R)N(S)$.

Suppose by way of contradiction that there are $N + 1$ different points in $\hat{\Sigma}$ whose image under $\hat{\pi}$ is equal to $x$. Call these points $R^{(j)} = (R_i^{(j)})_{i \in \mathbb{Z}}$ ($j = 0, \ldots, N$).

Assume w.l.o.g. that $R^{(0)} = R$.

By Lemma 12.2 there are chains $u^{(j)} = (v_i^{(j)})_{i \in \mathbb{Z}} \in \Sigma$ s.t. for every $n$

$$R_n^{(j)} \subset Z(v_n^{(j)}) \text{ and } -n[R_n^{(j)}, \ldots, R_0^{(j)}] \subset Z_n(v_n^{(j)}, \ldots, v_0^{(j)}).$$  \hspace{1cm} (12.2)

Claim 1. $\pi(u^{(j)}) = x$ for every $0 \leq j \leq N$.

The following inclusions hold:

$$\pi(u^{(j)}) \in \bigcap_{n=0}^{\infty} Z_n(v_n^{(j)}, \ldots, v_0^{(j)}) \subset \bigcap_{n=0}^{\infty} Z_n(v_n^{(j)}, \ldots, v_0^{(j)}),$$ \hspace{1cm} (12.3)

$$x = \hat{\pi}(R^{(j)}) \in \bigcap_{n=0}^{\infty} Z_n(-R_0^{(j)}, \ldots, R_n^{(j)}) \subset \bigcap_{n=0}^{\infty} Z_n(v_n^{(j)}, \ldots, v_0^{(j)}).$$
Since $\pi$ is H"{o}lder continuous, $\text{diam}\left[ Z_{-n}(v_{-n}^{(j)}, \ldots, v_{n}^{(j)}) \right] \xrightarrow{n \to \infty} 0$, so $\pi(v^{(j)}) = x$.

Claim 2: Suppose $i \in \mathbb{Z}$, then $R_{i}^{(0)}, \ldots, R_{i}^{(N)}$ are affiliated.

Proof. By (12.23) $x = \pi(v^{(j)}) \in \bigcap_{n=0}^{\infty} Z_{-n}(v_{-n}^{(j)}, \ldots, v_{n}^{(j)})$, so $f^{i}(x) \in Z(v_{i}^{(j)})$. Thus $Z(v_{i}^{(0)}), \ldots, Z(v_{i}^{(N)})$ have a common intersection. Since $R_{i}^{(j)} \subset Z(v_{i}^{(j)})$, $R_{i}^{(0)}, \ldots, R_{i}^{(N)}$ are affiliated.

Claim 3: There exist $k, \ell \geq 0$ and $0 \leq j_{1}, j_{2} \leq N$ such that
- $(R_{-k}^{(j)}), \ldots, R_{\ell}^{(j)} \neq (R_{-k}^{(j)}), \ldots, R_{\ell}^{(j)}$;
- $R_{-k}^{(j)} = R_{-k}^{(j)}$ and $R_{\ell}^{(j)} = R_{\ell}^{(j)}$;
- $v_{-k}^{(j)} = v_{-k}^{(j)}$ and $v_{\ell}^{(j)} = v_{\ell}^{(j)}$.

Proof. We are assuming that $R_{i}^{(j)}$ are different, therefore there exists some $m$ such that the words $(R_{-m}^{(j)}, \ldots, R_{m}^{(j)}) (0 \leq j \leq N)$ 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 $(R_{-k}^{(j)}, \ldots, R_{\ell}^{(j)}) (0 \leq j \leq N)$ are different.

By claims 1 and 2, $R_{-k}^{(j)}$ are all affiliated to $R_{-k}^{(0)} = R$, and by (12.22) $R_{-k}^{(j)} \subset Z(v_{-k}^{(j)})$, therefore $|\{(R_{-k}^{(j)}, v_{-k}^{(j)}): j = 0, \ldots, N\}| \leq N(R)$. In the same way, one can show that $|\{(R_{\ell}^{(j)}, v_{\ell}^{(j)}): j = 0, \ldots, N\}| \leq N(S)$. It follows that

$|\{(R_{-k}^{(j)}, v_{-k}^{(j)}; R_{\ell}^{(j)}, v_{\ell}^{(j)}): j = 0, \ldots, N\}| \leq N(R)N(S) = N$.

By the pigeonhole principle, at least two quadruples coincide, proving the claim.

To ease up the notation, we let $A := R^{(j)}_{\ell} - R^{(j)}_{k}, B := R^{(j)}_{\ell}, a := v^{(j)}_{\ell}$ and $b := v^{(j)}_{k}$, 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

$x_{A} \in -k[A_{-k}, \ldots, A_{\ell}]$ and $x_{B} \in -k[B_{-k}, \ldots, B_{\ell}]$.

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

$f^{-k}(z_{A}) \in W^{u}(f^{-k}(x_{B}), B) \cap W^{s}(f^{-k}(x_{A}), B)$;

$f^{\ell}(z_{B}) \in W^{u}(f^{\ell}(x_{B}), A) \cap W^{s}(f^{\ell}(x_{A}), A)$.

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 [11.4]),

$f^{-k+1}(z_{A}) \in f \left[ W^{s}(f^{-k}(x_{A}), A_{-k}) \right] \subset W^{s}(f^{-k+1}(x_{A}), A_{-k+1})$

and so on. It follows that $f^{-k}(z_{A}) \in -k[A_{-k}, \ldots, A_{\ell}]$. 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 $f^{-k}(z_{B}) \in -k[B_{-k}, \ldots, B_{\ell}]$. 

Suppose whose entropy is larger than \( f \), using the inclusions \( f \) the elements of \( \mathcal{R} \) are pairwise disjoint, so \( -k[A_{-k}, \ldots, A_{\ell}] \cap -k[B_{-k}, \ldots, B_{\ell}] = \emptyset \) and \( z_A \neq z_B \).

**Claim 5.** \( z_A = z_B \) (a contradiction).

**Proof.** We saw above that \( f^{-k}(z_A) \in -k[A_{-k}, \ldots, A_{\ell}], f^{-k}(z_B) \in -k[B_{-k}, \ldots, B_{\ell}] \).
In particular, \( f^{-k}(z_B) \in B_{-k} = B \subset Z(b) \) and \( f^i(z_A) \in A_{i} \subset Z(a) \).

Construct chains \( \alpha, \beta \in \Sigma^\# \) such that \( z_A = \pi(\alpha), \alpha_{\ell} = a \) and \( z_B = \pi(\beta), \beta_{-\ell} = b \).
Define a sequence \( \ell \) by

\[
c_i = \begin{cases} 
\beta_i & i \leq -k \\
\alpha_i & -k + 1 \leq i \leq \ell - 1 \\
\alpha_i & i \geq \ell. 
\end{cases}
\]

This is a chain because \( \beta_{-\ell} = b = a_{-k} \) and \( \alpha_{\ell} = a = a_{\ell} \). This chain belongs to \( \Sigma^\# \), because \( \alpha, \beta \in \Sigma^\# \).
We write \( c_i := \Psi_{x_i}^{p_i, p_i'} \).

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 \( V^u(f^{-k}(x_B), B) \), \( f^{-k}(z_A) \) by definition, and \( f^{-k}(z_B) \) because of the inclusion \( f^i(z_B) \in W^u(f^i(x_B), B) \) and the Markov property. 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

\[
f^i(z_A), f^i(z_B) \in Z(c_i) \subset \Psi_{x_i}[R_{\mathcal{P}_i \& \mathcal{P}_i'}(\mathcal{Q})] \quad \text{for all } i \leq -k.
\]  
(12.4)

Similarly, one can show that \( f^i(z_A), f^i(z_B) \in V^s[(c_i)_{i \geq \ell}] \), whence

\[
f^i(z_A), f^i(z_B) \in Z(c_i) \subset \Psi_{x_i}[R_{\mathcal{P}_i \& \mathcal{P}_i'}(\mathcal{Q})] \quad \text{for all } i \geq \ell.
\]  
(12.5)

Using the inclusions \( f^{-k}(z_A) \in -k[A_{-k}, \ldots, A_{\ell}], f^{-k}(z_B) \in -k[B_{-k}, \ldots, B_{\ell}] \) (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 \( f^i(x) = \pi[\sigma^i(\mathcal{Q})] = \pi[\sigma^i(\mathcal{H})] \in Z(a_i) \cap Z(b_i) \).
Thus by Lemma 10.10

\[
f^i(z_A), f^i(z_B) \in Z(a_i) \cup Z(b_i) \subset \Psi_{x_i}[R_{\mathcal{Q}_i}(x_i)](\mathcal{Q}) \quad \text{for all } -k < i < \ell.
\]  
(12.6)

In summary, \( f^i(z_A), f^i(z_B) \in \Psi_{x_i}[R_{\mathcal{Q}_i}(x_i)](\mathcal{Q}) \), where \( c_i = \Psi_{x_i}^{p_i, p_i'} \) is a chain. By Proposition 4.4, \( z_A, z_B \in V^u[(c_i)_{i \leq 0}] \cap V^s[(c_i)_{i \geq 0}] \). So \( z_A = \pi(\alpha) = z_B \), 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 \( \sigma : \hat{\Sigma} \to \hat{\Sigma} \) denote the finite-to-one Markov extension of \( f \) which we constructed in part 3. We compare the invariant Borel measures of \( \sigma : \hat{\Sigma} \to \hat{\Sigma} \) to the invariant Borel measures of \( f : M \to M \). We restrict our attention to measures whose entropy is larger than \( \chi \).

**Proposition 13.1.** Suppose \( \hat{\mu} \) is an ergodic Borel probability measure on \( \hat{\Sigma} \), then \( \mu := \hat{\mu} \circ \hat{\pi}^{-1} \) is an ergodic Borel probability measure on \( M \), and \( h_\mu(f) = h_\hat{\mu}(\sigma) \).
Proof. It is clear that $\mu$ is well-defined, ergodic and invariant.

By Poincaré’s Recurrence Theorem there exists a vertex $R \in \mathcal{R}$ s.t.

$$\Upsilon := \{R \in \widehat{\Sigma} : \exists m_k, m_k \uparrow \infty \text{ s.t. } R_{m_k}, R_{m_k} = R\}$$

has full measure with respect to $\hat{\mu}$. The map $\hat{\pi} : \Upsilon \to M$ is bounded-to-one (the bound is $\varphi_\sigma(R, R)$). Finite extensions preserve entropy, so $h_\mu(f) = h_{\hat{\mu}}(\sigma)$. □

The other direction, “every invariant measure $\mu$ supported on $\hat{\pi}(\widehat{\Sigma})$ lifts to an invariant measure on $\widehat{\Sigma}$”, is less clear. 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 $\bar{\mu}$ by

$$\bar{\mu}(E) := \int_M \left(\frac{1}{|\hat{\pi}^{-1}(x)|} \sum_{\hat{\pi}(R) = x} 1_E(R)\right) d\mu(x). \quad (13.1)$$

**Proposition 13.2.** Suppose $\mu$ is an ergodic $f$–invariant Borel probability measure on $M$ s.t. $h_\mu(f) > \chi$.

1. $\bar{\mu}$ is a well-defined $\sigma$–invariant Borel probability measure on $\widehat{\Sigma}$.
2. Almost every ergodic component $\hat{\mu}$ of $\bar{\mu}$ is an ergodic $\sigma$–invariant probability measure such that $\hat{\mu} \circ \hat{\pi}^{-1} = \mu$ and $h_{\hat{\mu}}(\sigma) = h_\mu(f)$.

**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) $F : X \to Y$ is Borel iff graph($F$) is a Borel subset of $X \times Y$.

(II) Suppose $F : X \to Y$ 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 $\hat{\pi}(\Sigma^\#)$. Since $\hat{\pi} : \Sigma^\# \to M$ is finite-to-one, $\hat{\pi}(\Sigma^\#)$ is Borel. Henceforth we work inside $\hat{\pi}(\Sigma^\#)$.

**Step 1.** $x \mapsto |\hat{\pi}^{-1}(x)|$ is constant on a Borel set $\Omega$ s.t. $\mu(\Omega) = 1$.

**Proof.** Since $\hat{\pi} \circ \sigma = f \circ \hat{\pi}$ and $f$ is a bijection, $x \mapsto |\hat{\pi}^{-1}(x)|$ is $f$–invariant.

We show that the restriction of $x \mapsto |\hat{\pi}^{-1}(x)|$ to $\hat{\pi}(\Sigma^\#)$ is Borel measurable. The claim will then follow from the ergodicity of $\mu$.

Graphs of Borel functions are Borel, therefore $B := \{(\hat{\pi}(R), R) : R \in \hat{\Sigma}^\#\}$ is a Borel subset of $M \times \hat{\Sigma}$.

By Lusin’s theorem, $\exists$ partially defined Borel functions $\varphi_n : M_n \to \hat{\Sigma}^\#$ 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, $\hat{\pi}^{-1}(x) = \{\varphi_i(x) : i \in \mathbb{N} \text{ s.t. } M_i \ni x\}$. The graphs of $\varphi_i$ are pairwise disjoint, so $i \neq j \Rightarrow \varphi_i(x) \neq \varphi_j(x)$. Consequently,

$$|\hat{\pi}^{-1}(x)| = \sum_{i=1}^\infty 1_{M_i}(x)$$
on $\hat{\pi}^{-1}(x)$ does not work: it is not even $\sigma$–additive.
Since $M_i$ are Borel, $x \mapsto |\tilde{\pi}^{-1}(x)|$ is Borel on $\tilde{\pi}(\hat{\Sigma})$.

**Step 2.** Let $\Upsilon := \tilde{\pi}^{-1}(\Omega)$ and let $N$ denote the number of pre-images of points $x \in \Omega$. There exists a Borel partition $\Upsilon = \bigcup_{i=1}^{N} \Upsilon_i$ such that $\tilde{\pi} : \Upsilon_i \to \Omega$ is one-to-one and onto for every $i$.

**Proof.** This is a consequence of Lusin’s Theorem.

Let $B_1 := \{(\tilde{\pi}(y), y) : y \in \tilde{\pi}^{-1}(\Omega)\}$. Each $x$-fibre of $B_1$ has $N$ elements. By Lusin’s Theorem $B_1 = \bigcup_{n \geq 1} \text{graph}(\varphi_n)$ where $\varphi_n : M_n \to \hat{\Sigma}$ are Borel. $\Omega = \bigcup_{n \geq 1} M_n$.

Define $\psi_1 : \Omega \to \hat{\Sigma}$ by $\psi_1 = \varphi_1$ on $M_i \setminus \bigcup_{j < i} M_j$ ($i \in \mathbb{N}$), then $\psi_1$ is Borel and $\psi_1(x) \in \tilde{\pi}^{-1}(x)$ for all $x$. Since $\tilde{\pi} \circ \psi_1 = \text{Id}$, $\psi_1$ is one-to-one. It follows that $\Upsilon_1 := \psi_1(\Omega)$ is Borel, and $\tilde{\pi} : \Upsilon_1 \to \Omega$ 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 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 $E \subset \hat{\Sigma}$,

$$\frac{1}{|\tilde{\pi}^{-1}(x)|} \sum_{\tilde{\pi}(y) = x} 1_E(y) = \frac{1}{N} \sum_{i=1}^{N} 1_{\tilde{\pi}(E \cap \Upsilon_i)}(x)$$

on $\Omega$.

Since $\tilde{\pi}$ is one-to-one on $\Upsilon_i$, $\tilde{\pi}(E \cap \Upsilon_i)$ 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 $\tilde{\mu} \circ \tilde{\pi}^{-1} = \mu$ and $h_{\tilde{\pi}}(\sigma) = h_\mu(f)$.

**Proof.** We saw that $\tilde{\mu}(E)$ is well-defined for all Borel sets $E \subset \hat{\Sigma}$. This set function is obviously $\sigma$–additive, and it is clear that $\tilde{\mu}(\hat{\Sigma}) = 1$. Thus $\tilde{\mu}$ is a Borel probability measure.

This measure is $\sigma$–invariant, because

$$\tilde{\mu}(\sigma^{-1}E) = \int_M \left( \frac{1}{|\tilde{\pi}^{-1}(x)|} \sum_{\tilde{\pi}(y) = x} 1_E(\sigma(y)) \right) d\mu(x)$$

$$= \int_M \left( \frac{1}{|\tilde{\pi}^{-1}(f(x))|} \sum_{\tilde{\pi}(\sigma(y)) = f(x)} 1_E(\sigma(y)) \right) d\mu(x) \quad (\because \tilde{\pi} \circ \sigma = f \circ \tilde{\pi})$$

$$= \int_M \left( \frac{1}{|\tilde{\pi}^{-1}(f(x))|} \sum_{\tilde{\pi}(\Sigma) = f(x)} 1_E(\Sigma) \right) d\mu(x)$$

$$= \tilde{\mu}(E) \quad (\because \mu \circ f^{-1} = \mu).$$

It is a lift of $\mu$ because

$$\tilde{\mu}(\tilde{\pi}^{-1}E) = \int_M \left( \frac{1}{|\tilde{\pi}^{-1}(x)|} \sum_{\tilde{\pi}(y) = x} 1_E(\tilde{\pi}(y)) \right) d\mu(x) = \int_M 1_E(x) d\mu(x) = \mu(E).$$

Finally $\tilde{\mu}$ and $\mu$ have the same entropy, because $\tilde{\pi}$ is $N$–to–one on a set of full measure, and finite extensions preserve entropy.
Step 5. Almost every ergodic component of \( \hat{\mu} \) satisfies \( \hat{\mu} \circ \hat{\pi}^{-1} = \mu \) and \( h_{\hat{\mu}}(\sigma) = h_{\mu}(\mu) \).

Let \( \hat{\mu} = \int \hat{\mu}_y dv(y) \) be the ergodic decomposition of \( \hat{\mu} \), then \( \mu = \hat{\mu} \circ \hat{\pi}^{-1} = \int \hat{\mu}_y \circ \hat{\pi}^{-1} dv_y \). Each of the measures \( \hat{\mu}_y \circ \hat{\pi}^{-1} \) is \( f \)-invariant. Since \( \mu \) is ergodic, \( \hat{\mu}_y \circ \hat{\pi}^{-1} = \mu \) 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 (Compare with 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 \( df_x E^\sigma(x) = E^\sigma(f(x)) \). Both sides of the equation are one-dimensional, therefore \( df_x L^\sigma(x) = \pm \| df_x \xi(x) \| f(x) \xi^\sigma(f(x)) \). It follows that

\[
A_\chi(x) \xi_1 = s_\chi(x)^{-1} \{ [C_\chi(f(x))^{-1} \circ df_x] \xi^\sigma(x) } \]
\[
= \pm s_\chi(x)^{-1} \| df_x e^\sigma(x) \| f(x) C_\chi(f(x))^{-1} \xi^\sigma(f(x)) \]
\[
= \pm \frac{s_\chi(f(x))}{s_\chi(x)} \| df_x \xi^\sigma(x) \| f(x) \xi_1.
\]

We see that \( \xi_1 \) is an eigenvector of \( A_\chi(x) \) with eigenvalue

\[
\lambda_\chi(x) := \pm \frac{s_\chi(f(x))}{s_\chi(x)} \| df_x \xi^\sigma(x) \| f(x).
\]  \( \text{(A.1)} \)

Similarly, \( \xi_2 \) is an eigenvector of \( A_\chi(x) \) with eigenvalue

\[
u_\chi(x) := \pm \frac{u_\chi(f(x))}{u_\chi(x)} \| df_x \xi^u(x) \| f(x). \]
\[
\text{(A.2)} \]

We estimate the eigenvalues:

\[
 s_\chi(x)^2 = 2 \sum_{k=0}^\infty e^{2k\chi} \| (df^k_x) \xi^\sigma(x) \| f(x)^2 > 2 \sum_{k=1}^\infty e^{2k\chi} \| (df^k_x) \xi^\sigma(x) \| f(x)^2,
\]
\[
= 2 \sum_{k=0}^\infty e^{2(k+1)\chi} \| (df^k_x) \xi^\sigma(x) \| f(x)^2
\]
\[
= 2 \| df_x \xi^\sigma(x) \| f(x)^2 \sum_{k=0}^\infty e^{2(k+1)\chi} \| (df^k_x) \xi^\sigma(f(x)) \| f(x)^2
\]
\[
= e^{2\chi} \| df_x \xi^\sigma(x) \| f(x)^2 s_\chi(f(x))^2.
\]

Rearranging terms, we find that \( e^{-2\chi} > \frac{s_\chi(f(x))^2}{s_\chi(x)^2} \| df_x \xi^\sigma(x) \| f(x)^2 = \lambda_\chi(x)^2 \). It follows that \( |\lambda_\chi(x)| < e^{-\chi} \). Similarly, one shows that \( |\nu_\chi(x)| > e^\chi \).

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, \( h_{\text{top}}(f) \leq 2 \log M_f \).
By definition of $s_\chi(x)$, and the identity $df_x\xi^\chi(x) = \pm\|df_x\xi^\chi(x)\|\xi^\chi(f(x))$, 

$$s_\chi(x)^2 = 2 \left( 1 + \sum_{k=1}^{\infty} e^{2k\chi} \|df_{f(x)}^{k-1}\xi^\chi(f(x))\|^2 \|df_x\xi^\chi(x)\|_x \right) \leq 2 \left( 1 + e^{2\chi} M_f^2 \sum_{k=0}^{\infty} e^{2k\chi} \|df_{f(x)}^k\xi^\chi(f(x))\|^2 \|df_x\xi^\chi(x)\|_x \right) \leq 2 + e^{2\chi} M_f^2 s_\chi(f(x))^2 \leq (M_f^6 + 1) s_\chi(f(x))^2 \quad (\because s_\chi > \sqrt{2} \text{ and } \chi < h_{top}(f) \leq 2 \log M_f).$$

Therefore by (A.1)

$$|\lambda_\chi(x)| > (1 + M_f^6)^{-1/2} \|df_x\xi^\chi(x)\|_f(x) \geq M^{-1}_f(1 + M_f^6)^{-1/2}. \quad (A.3)$$

Similarly, one can bound $|\mu_\chi(x)|$ from above by a function of $M_f$. \hfill \Box

Proof of Lemma 2.4 We put the standard basis $\xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ on $\mathbb{R}^2$, and the basis $\xi^\chi(x), \xi^{\chi(x)+}$ on $T_x M$, where $\xi^\chi$ denotes the unique vector s.t. the signed angle from $\xi$ to $\xi^\chi$ is $\pi/2$. The linear map $C_\chi(x) : \mathbb{R}^2 \rightarrow T_x$ is represented in these bases by the matrix

$$\begin{pmatrix} s_\chi(x)^{-1} u_x(x)^{-1} \cos \alpha(x) \\ 0 & u_x(x)^{-1} \sin \alpha(x) \end{pmatrix}.$$ 

Inverting, we find that $C_\chi(x)^{-1} : T_x M \rightarrow \mathbb{R}^2$ is represented by

$$\begin{pmatrix} s_\chi(x) & -s_\chi(x)/\tan \alpha(x) \\ 0 & u_x(x)/\sin \alpha(x) \end{pmatrix}.$$ 

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 $(\sum a_{ij}^2)^{1/2}$. \hfill \Box

Proof of Lemma 2.5 Define an inner product $\langle \cdot, \cdot \rangle^\chi_x$ on $T_x M$ by the conditions (a) $\|\xi^\chi(x)\|^\chi_x = s_\chi(x)$, (b) $\langle \xi^\chi(x), \xi^{(\chi(x)+)\chi} \rangle^\chi_x = u_x(x)$, and (c) $\langle \xi^{(\chi(x)+)\chi}, \xi^\chi(x) \rangle^\chi_x = 0$ (compare with [BP, §3.5.1]). The inner product $\| \cdot \|_\chi^\chi_x$ satisfies $\| \cdot \|_\chi^\chi_x \geq \| \cdot \|_x$ because for every $\xi, \eta \in \mathbb{R}$

$$\|\xi^\chi(x) + \eta \xi^{\tau}(x)\|_\chi^\chi_x = \sqrt{\xi^2 s_\chi(x)^2 + \eta^2 u_x(x)^2} > \sqrt{2(\xi^2 + \eta^2)} \quad (\because s_\chi(x), u_x(x) > \sqrt{2})$$

$$\geq \|\xi\| + |\eta| = \|\xi \xi^\chi(x)\|_x + \|\eta \xi^{\tau}(x)\|_x \geq \|\xi \xi^\chi(x) + \eta \xi^{\tau}(x)\|_x.$$ 

$$\therefore \|C_\chi(x)(\xi)\|_x \leq \|C_\chi(x)(\xi)\|_\chi^\chi_x = \|\xi s_\chi(x)^{-1} \xi^\chi(x) + \eta u(x)^{-1} \xi^{\tau}(x)\|_x^\chi = \sqrt{\xi^2 + \eta^2}.$$ 

The lemma follows. \hfill \Box

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 $A_\chi^{(0)} := A_\chi$ and $A_\chi^{(m+n)}(x) = A_\chi^{(m)}(f^n(x))A_\chi^{(n)}(x)$. The extension is unique, and is given by $A_\chi^{(n)}(x) = C_\chi(f^n(x))^{-1} df^n_x C_\chi(x)$.

Theorem 2.3 says that $A_\chi(x)$ is a diagonal matrix with entries in $[C^{-1}_f, C_f]$ for every $x \in \text{NUH}_\chi(f)$. In particular, $\log \|A_\chi^{(0)}\|$ 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 $\text{NUH}_\chi^{(1)}(f)$ denote the set points $x \in \text{NUH}_\chi(f)$ for which there is a decomposition $T_x \mathbb{R}^2 = E_x^s(x) \oplus E_x^u(x)$ so that
(1) $E^\ast_\chi(x) = \text{span}\{e^s_\chi(x)\}$, $\|e^s_\chi(x)\| = 1$, $\lim_{n \to \pm \infty} \frac{1}{n} \log \|A^{(n)}_\chi(x)e^s_\chi(x)\| < 0$;

(2) $E^\ast_n(x) = \text{span}\{e^s_n(x)\}$, $\|e^s_n(x)\| = 1$, $\lim_{n \to \pm \infty} \frac{1}{n} \log \|A^{(n)}_\chi(x)e^s_n(x)\| > 0$;

(3) $\lim_{n \to \infty} \frac{1}{n} \log |\sin \alpha(x)f^n(x)| = 0$, where $\alpha(x) := \zeta(e^s_n(x), e^s_n(x))$;

(4) $A_\chi(x)|E^\ast_n(x)| = E^\ast_n(f(x))$ and $A_\chi(x)|E^\ast_n(x)| = E^\ast_n(f(x))$.

By the discussion above, $\text{NUH}_\chi^1(f)$ has full measure w.r.t. to any ergodic invariant probability measure with entropy larger than $\chi$.

Let $\text{NUH}_\chi^1(f)$ denote the subset of $\text{NUH}_\chi^1(f)$ which consists of all points $x$ for which there exist a sequence $n_k \uparrow \infty$ s.t. $C_\chi(f^{n_k}(x)) \xrightarrow{k \to \infty} C_\chi(x)$ and a sequence $m_k \downarrow -\infty$ s.t. $C_\chi(f^{m_k}(x)) \xrightarrow{k \to \infty} C_\chi(x)$. By the Poincaré Recurrence Theorem, every invariant probability measure which is carried by $\text{NUH}_\chi^1(f)$ is carried by $\text{NUH}_\chi^1(f)$, so $\text{NUH}_\chi^1(f)$ has full measure w.r.t. to every ergodic invariant measure with entropy greater than $\chi$.

On the set $\text{NUH}_\chi^1(f)$, the Multiplicative Ergodic Theorem holds for both $df_x$ and $A^{(n)}_\chi(x)$, so the following two limits exist:

$$\lim_{n \to \pm \infty} \frac{1}{n} \log \|df_x^n C_\chi(x)e^{s_n(x)}\| f^n(x), \quad \lim_{n \to \pm \infty} \frac{1}{n} \log \|C_\chi(f^n(x))^{-1} df_x^n C_\chi(x)e^{s_n(x)}\|. \quad (A.4)$$

Let $n_k \uparrow \infty$ be a subsequence for which $C_\chi(f^{n_k}(x)) \xrightarrow{k \to \infty} C_\chi(x)$. The norms of $C_\chi(f^{n_k}(x))$ and $C_\chi(f^{n_k}(x))^{-1}$ are bounded along this sequence, so

$$\|C_\chi(f^{n_k}(x))^{-1} df_x^{n_k} C_\chi(x)e^{s_n(x)}\| \leq \|df_x^{n_k} C_\chi(x)e^{s_n(x)}\|.$$

We see that the limits in (A.4) agree. As a result $E^\ast_\chi(x) = \mathbb{R} \times \{0\}$, $E'_\chi(x) = \{0\} \times \mathbb{R}$, and $x$ has Lyapunov exponents $\lambda(x)$ and $\mu(x)$ w.r.t. $A^{(n)}_\chi$.

Let $A_\chi(x) := \begin{pmatrix} \lambda(x) & 0 \\ 0 & \mu(x) \end{pmatrix}$, then the limits (A.4) mean that

$$\|(A^{(n)}_\chi(x))^{(x)} - 1\|^{1/n} \xrightarrow{n \to \pm \infty} 1.$$ 

Similarly, if $A(x)$ is the linear operator s.t. $A(x)e^{s}(x) = \lambda(x)e^{s}(x)$ and $A(x)e^{u}(x) = \mu(x)e^{u}(x)$, then

$$\|(df_x^n A(x))^{-1}\|^{1/n} \xrightarrow{n \to \pm \infty} 1.$$ 

Since $A_\chi(x) = C_\chi(x)^{-1} A(x) C_\chi(x)$ and $A^{(n)}_\chi(x) = C_\chi(f^n(x))^{-1} \circ df_x \circ C_\chi(x)$,

$$\|(C_\chi \circ f^n)^{-1}\|^{1/n} = \|A^{(n)}_\chi C_\chi^{-1} (df_x^n)^{-1}\|^{1/n}$$

$$= \|A^{(n)}_\chi C_\chi^{-1} A^{-1} \cdot C_\chi^{-1} \cdot \Lambda_n(df_x^n)^{-1}\|^{1/n}$$

$$\leq \|A^{(n)}_\chi A^{-1}\|^{1/n} \|C_\chi^{-1}\|^{1/n} \|(df_x^n A^{-1})^{-1}\|^{1/n} \xrightarrow{n \to \pm \infty} 1.$$ 

Thus $\limsup \frac{1}{n} \log \|(C_\chi \circ f^n)^{-1}\| \leq 0$. On the other hand $C_\chi$ is a contraction (Lemma 2.3), so $\|(C_\chi \circ f^n)^{-1}\|^{1/n} \geq 1$, whence $\lim \inf \frac{1}{n} \log \|(C_\chi(f^n(x)))^{-1}\| \geq 0$.

The first part of the Lemma is proved.

We prove the second part of the Lemma: $\frac{1}{n} \log \|C_\chi(f^n(x))e^{s_n(x)}\| f^n(x) \xrightarrow{n \to \pm \infty} 0$.

We do this for $i = 1$, and leave the case $i = 2$ to the reader. Since the $A^{(n)}_\chi(\cdot)$ is
diagonal, \( A^{(n)}_\chi(x) \mathbf{e}_1 \) is proportional to \( \mathbf{e}_1 \). The multiplicative ergodic theorem for \( A^{(n)}_\chi(x) \) says that \( A^{(n)}_\chi(x) \mathbf{e}_1 = \pm \lambda(x)^n \exp[\alpha(n)]] \mathbf{e}_1 \), therefore
\[
\lim_{n \to \pm \infty} \| C_\chi(f^n(x)) \mathbf{e}_1 \|^{1/n}_f = \lambda(x)^{-1} \lim_{n \to \pm \infty} \| C_\chi(f^n(x)) A^{(n)}_\chi(x) \mathbf{e}_1 \|^{1/n}_f
\]
\[
= \lambda(x)^{-1} \lim_{n \to \pm \infty} \| (df^n) C_\chi(x) \mathbf{e}_1 \|^{1/n}_f
\]
\[
= \lambda(x)^{-1} \lim_{n \to \pm \infty} \| (df^n) \mathbf{e}_1 \|^{1/n}_f = 1,
\]
proving that \( \frac{1}{n} \log \| C_\chi(f^n(x)) \mathbf{e}_1 \|^{1/n}_f \xrightarrow{n \to \pm \infty} 0 \).

Finally, we prove that \( \frac{1}{n} \log \| \det C_\chi(f^n(x)) \| \xrightarrow{n \to \pm \infty} 0 \). We begin with some general comments on determinants.

Suppose \( L : V \to W \) is a linear operator between two two dimensional vector spaces with inner product. The determinant of \( \det L \) can be defined as \( \det(L \Theta) \) for some (every) isometry \( \Theta : W \to V \). The following fact holds: If \( \mathbf{u}, \mathbf{v} \) span \( V \), then
\[
\sin \angle(\mathbf{u}, \mathbf{v}) = \frac{\| \mathbf{u} \| \| \mathbf{v} \| \det L}{\| \mathbf{u} \| \| \mathbf{v} \|} \qquad (A.5)
\]
It follows that
\[
| \det L | = \frac{\| \mathbf{u} \| \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v})}{\| \mathbf{u} \| \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v})} \qquad (\mathbf{u}, \mathbf{v} \text{ independent}).
\]

Applying this to \( L = A^{(n)}_\chi \) with \( \mathbf{u} = \mathbf{e}_1, \mathbf{v} = \mathbf{e}_2 \), and to \( L = df^n \) with \( \mathbf{u} = \mathbf{e}_\beta(x), \mathbf{v} = \mathbf{e}_\beta(x) \), we find that
\[
\lim_{n \to \pm \infty} \frac{1}{n} \log | \det A^{(n)}_\chi(x) | = \log \lambda(x) + \log \mu(x) = \lim_{n \to \pm \infty} \frac{1}{n} \log | \det df^n |.
\]
But \( | \det A^{(n)}_\chi(x) | = | \det C_\chi(f^n(x)) | | \det df^n | | \det C_\chi(x) | \). It follows that
\[
\frac{1}{n} \log | \det C_\chi(f^n(x)) | \xrightarrow{n \to \pm \infty} 0
\]
as required. \( \Box \)

**Proof of Lemma 2.9** Parts (1) and (3) are obvious, and part (4) is a consequence of Lemma 2.8 and the estimate \( Q_\varepsilon(f^n(x)) \approx \| C_\chi(f^n(x)) \|^{-12/\beta} \). For part (6), define \( q_\varepsilon(x) \) on NUH \( f \) by the formula
\[
\frac{1}{q_\varepsilon(x)} = \frac{1}{\varepsilon} \sum_{k=\infty}^{\infty} e^{-\frac{1}{2} \|k\varepsilon\} \frac{1}{Q_\varepsilon(f^n(x))}.
\]
The sum converges because \( \frac{1}{\varepsilon} \log Q_\varepsilon(f^n(x)) \xrightarrow{k \to \pm \infty} 0 \), and it is easy to check that \( q_\varepsilon(x) \) behaves as required. (Compare with [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 NUH \( f \).

---

8Proof: Let \( \omega_V, \omega_W \) denote the volume 2–forms on \( V, W \), then \( \omega_V(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} \| \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v}) \) and \( \omega_V(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} \| \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v}) \). Since \( \omega_W(\mathbf{u}, \mathbf{v}) \) is also a 2–form on \( V \), and any two 2–forms on \( V \) are proportional, s.t. \( \omega_W(\mathbf{u}, \mathbf{v}) = e \omega_V(\mathbf{u}, \mathbf{v}) \). Evaluating on an orthonormal basis of \( V \), we find that \( e = \det L \). Consequently, \( \| \mathbf{u} \| \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v}) = \det L \mathbf{u} \| \mathbf{v} \| \sin \angle(\mathbf{u}, \mathbf{v}) \).
Proof. By Lemma 2.4 it is enough to show that

\[
\frac{s_x \circ f}{s_x}, \frac{u_x \circ f}{u_x}, \frac{|\sin \alpha \circ f|}{|\sin \alpha|}
\]

are uniformly bounded away from zero and infinity on 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.4) makes it clear that \(\frac{u_x(f(x))}{s_x(x)} = F_0^{\frac{1}{2}} \lambda_\epsilon(x) \in [(C_f F_0)^{-1}, C_f F_0]\) on NUH\(\chi(f)\). Similarly, \(\frac{u_x(f(x))}{u_x(x)}\) takes values in \([(C_f F_0)^{-1}, C_f F_0]\) on NUH\(\chi(f)\).

Finally, by (A.5) and the fact that \(\epsilon^{1/u(f(x))}\) have the same direction as \(df_x\epsilon^{1/u(x)}\) up to a sign,

\[
\frac{|\sin \alpha(f(x))|}{|\sin \alpha(x)|} = \frac{|\sin \angle (\epsilon^u(f(x)), \epsilon^{u}(f(x)))|}{|\det \epsilon^u(x)||\det \epsilon^{u}(x)|}.
\]

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 \(Q_\epsilon(x) < \epsilon^{3/\beta} C_\chi(x)^{-1} F_0^{-12/\beta} < \epsilon^{3/\beta} |C_\chi(x)|^{-1} - 12\), therefore also \(Q_\epsilon(x) < (C_\chi(f^{+1}(x))^{-1} - 12\). If \(\epsilon\) is small enough then \(\epsilon^{1/\beta} C_\chi^{12/\beta} < 1\), 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 \(B_r^\beta(q) \subset T_q M\) diffeomorphically onto an open set \(U_q \supseteq W_p\) in a \(2\)-bi-Lipschitz way, and \(d(\exp_q) = \text{Id}\);
3. for every \(q, q' \in W_p\), there is a unique vector \(v(q, q') \in T_q M\) s.t. \(\|v(q, q')\| < r\) and \(\exp_q(v(q, q')) = q'\);
4. \(v(q, q') \mapsto v(q, q')\) 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 \(B_{\rho(M)}(p) \subset T_p M\) diffeomorphically onto a neighborhood of \(B_{\rho(M)}(p) \subset M\), in a \(2\)-bi-Lipschitz way. Let

\[
r_0 := \min\{1, r(M), \rho(M)\} / 10[\text{Lip}(f) + \text{Lip}(f^{-1})].
\]

Note that \(r_0 < 1\).

Suppose \(\epsilon < r_0/5\). By the definition of \(Q_\epsilon(x)\), \(Q_\epsilon(x) < \epsilon^3\), so \(10Q_\epsilon(x) < r_0/\sqrt{2}\). By Lemma 2.5 \(C_\chi(x)\) maps \(R_{10Q_\epsilon(x)}(x)\) contractively into \(B_{r_0}(x)\). Therefore \(\Psi_x = \exp_x o C_\chi(x)\) maps \(R_{10Q_\epsilon(x)}(0)\) diffeomorphically in a \(2\)-Lipschitz way into \(M\). The first part of the theorem is proved.

Next we show that \(f_x := \Psi_x^{-1} o f \circ \Psi_x\) is well defined on \(R_{10Q_\epsilon(x)}(0)\) and establish its properties.

Since \(\exp_x\) is \(2\)-Lipschitz, \(C_\chi(x)\) is a contraction, and \(10Q_\epsilon(x) < r_0/\sqrt{2}\),

\(\Psi_x\) maps \(R_{10Q_\epsilon(x)}(0)\) diffeomorphically into \(B_{2r_0}(x)\).
It follows that \( f \circ \Psi \) maps \( R_{10Q_x}(0) \) diffeomorphically into \( B_{2\operatorname{Lip}(f)r_0}(f(x)) \), which by the definition of \( r_0 \) is a subset of \( B_{\rho(M)}(f(x)) \), whence a subset of \( \exp_{f(x)}[B^e_{r(M)}(p)] \). It follows that \( f_x := \Psi^{-1}_{f(x)} \circ f \circ \Psi \) is well defined, smooth and injective on \( R_{10Q_x}(0) \).

For every \( p \in M \), \( \exp_p(0) = p \) and \( d(\exp_p)_{0} = \operatorname{Id} \). It easily follows that \( f_x(0) = 0 \) and \( (df_x)_{0} = C_x(f(x))^{-1} \circ (df)_{0} \circ C_x(x) \). By Theorem 2.3 this is a diagonal matrix with diagonal elements \( A(x) = \lambda_x(x) \), \( B(x) = \mu_x(x) \), and \( C^{-1}_x < |A(x)| < e^{-x} \), \( e^x < |B(x)| < C_f \).

We compare \( f_x \) to its linearization at \( 0 \) by analyzing
\[
r_x(u) := f_x(u) - (df_x)_{0}(u).
\]
By assumption \( f \) is \( C^{1+\beta} \), so there is a constant \( L \) s.t. for all \( u, v \in R_{r_{0}}(0) \),
\[
|d(\exp_{f(x)}^f \circ f \circ \exp_x)_{u} - d(\exp_{f(x)}^f \circ f \circ \exp_x)_{v}| \leq L\|u - v\|^{\beta}.
\]
For every \( u, v \in R_{r_{0}}(0) \),
\[
|\|df_x(u)\| - \|df_x(v)\|| = |C_x(f(x))^{-1}d(\exp_{f(x)}^f \circ f \circ \exp_x)_{u} C_x(x)|
\]
\[
- C_x(f(x))^{-1}d(\exp_{f(x)}^f \circ f \circ \exp_x)_{v} C_x(x)|
\]
\[
= |C_x(f(x))^{-1}[d(\exp_{f(x)}^f \circ f \circ \exp_x)_{u} - d(\exp_{f(x)}^f \circ f \circ \exp_x)_{v}]| C_x(x)|
\]
\[
\leq |C_x(f(x))^{-1}| \cdot L|C_x(x)||\|u - v\|^{\beta} \cdot |C_x(x)|
\]
\[
\leq \left| (C_x(f(x))^{-1}) \cdot L\|u - v\|^{3/2} \cdot |u - v|^{\beta/2} \right| < 1.
\]
If \( u, v \in R_{10Q_x}(0) \), then the term in the brackets is smaller than
\[
|C_x(f(x))^{-1}| \cdot L(20\sqrt{2}Q_x(x))^{3/2}.
\]
Plugging in the definition of \( Q_x(x) \) from 2.3, and recalling that \( |C_x(\cdot)|^{3} > 1 \) (because \( C_x(\cdot) \) is a contraction), we see that the term in the brackets is smaller than \( 30^{3/2}Lx^{3/2} \). Thus, if \( \epsilon < \frac{3}{2} \cdot 30^{-3/2}L^{-1} \), then
\[
|\|df_x(u)\| - \|df_x(v)\|| \leq \frac{1}{3}\|u - v\|^{3/2} \quad (u, v \in R_{10Q_x}(0)).
\]
Since \( (df_x)_{0} = 0 \), we have that \( |\|df_x(u)\| \leq \frac{1}{3}\|u\|^{3/2} \) on \( R_{10Q_x}(0) \). Now \( Q_x(x) < \epsilon^{3/2} \), so \( \|u\| \leq (10\sqrt{2})Q_x(x) < 15\epsilon^{3/2} \). If \( \epsilon < 15^{-3/2} \), then \( \|u\| < 1 \), so
\[
|\|df_x(u)\| | \leq \frac{1}{3}\|u\| < \frac{1}{3}\|u\| \quad \text{on} \quad R_{10Q_x}(0).
\]
Since \( r_x(0) = 0 \), we have by the mean value theorem that
\[
|\|r_x(u)\| | \leq \frac{1}{3}\|u\| < \frac{1}{3}\|u\| \quad \text{on} \quad R_{10Q_x}(0).
\]
In summary, if \( \epsilon \) is small enough, then the \( C^{1+3/2} \)-distance between \( r_x \) and \( 0 \) on \( R_{10Q_x}(0) \) is less than \( \epsilon \). This shows that the \( C^{1+3/2} \)-distance between \( f_x \) and \( (df_x)_{0} \) on this set is less than \( \epsilon \).

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 \( \text{KM} \). Part (4) is new, but routine. Assume that \( 0 < \epsilon < \frac{1}{2} \).

Write \( V^u = \Psi_x \{ (F(u),v) : |v| \leq p^u \} \) and \( V^s = \Psi_x \{ (v,G(v)) : |v| \leq p^s \} \), and let \( \eta := p^u \wedge p^s \). Note that \( \eta < \epsilon \), and that \( |F(0)|, |G(0)| \leq 10^{-3}\eta \) and \( \operatorname{Lip}(F), \operatorname{Lip}(G) \leq \epsilon \), see 4.2.
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 $P := \Psi_s(v, w)$.

- $P \in V^u$, 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 $\xi := p^u \lor p^s$ and extend $F, G$ (arbitrarily) to $\varepsilon$–Lipschitz continuous functions $\bar{F}, \bar{G} : [-\xi, \xi] \to [-Q_e(\xi), Q_e(\xi)]$. Let $V^u$ and $V^s$ denote the $u/s$–sets represented by $\bar{F}, \bar{G}$. Any intersection point of $V^u, V^s$ is an intersection point of $\bar{V}^u, \bar{V}^s$. Such points take the form $P = \bar{\Psi}_s(\bar{v}, \bar{w})$ where $\bar{v} = \bar{F}(\bar{w})$ and $\bar{w} = \bar{G}(\bar{v})$. Notice that $\bar{w}$ is a fixed point of $\bar{G} \circ \bar{F}$. The same calculations as before show that $\bar{G} \circ \bar{F}$ contracts $[-\xi, \xi]$ into itself.

Such a point $P$ is a fixed point of $\Psi_s(v, w)$, whence $\bar{P} = P$.

Next we show that $P_i := G_i \circ F_i$ are $\varepsilon^2$–contractions of $[-10^{-2}\eta, 10^{-2}\eta]$ into itself, therefore

\[ |w_1 - w_2| = |f^n_1(w_1) - f^n_2(w_2)| \leq |f_1(f_1^{-1}(w_1)) - f_2(f_2^{-1}(w_2))| \]
\[ + |f_2(f_1^{-1}(w_1)) - f_2(f_2^{-1}(w_2))| \]
\[ \leq \|f_1 - f_2\|_{\infty} + \varepsilon^2 |f_1^{-1}(w_1) - f_2^{-1}(w_2)| \]
\[ \leq \cdots \leq \|f_1 - f_2\|_{\infty} (1 + \varepsilon^2 + \cdots + \varepsilon^{2n}) \]
\[ \leq \frac{1}{1 - \varepsilon^2} \|f_1 - f_2\|_{\infty}. \]

Similarly, $v_i$ is a fixed point of $F_i \circ G_i : [-10^{-2}\eta, 10^{-2}\eta] \to [-10^{-2}\eta, 10^{-2}\eta]$, 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 $\Psi_s$ 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
\[ \|F_1 \circ G_1 - F_2 \circ G_2\|_{\infty} \leq \|F_1 \circ G_1 - F_1 \circ G_2\|_{\infty} + \|F_1 \circ G_2 - F_2 \circ G_2\|_{\infty} \]
\[ \leq \text{Lip}(F_1) \|G_1 - G_2\|_{\infty} + \|F_1 - F_2\|_{\infty} \]
\[ \|G_1 \circ F_1 - G_2 \circ F_2\|_{\infty} \leq \text{Lip}(G_1) \|F_1 - F_2\|_{\infty} + \|G_1 - G_2\|_{\infty} \]

Since $\text{Lip}(F_1), \text{Lip}(G_1) \leq \varepsilon^2$, $d(P_1, P_2) < \frac{2(1 + \varepsilon^2)}{1 - \varepsilon^2} [\text{dist}(V^u_1, V^u_2) + \text{dist}(V^s_1, V^s_2)]$. 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 \( 0 < t < \varepsilon \implies e^{-2it} < 1 - t < t < e^{2it} \). In what follows we drop the subscript \( x \) in \( \| \cdot \|_x \).

Let \( \mathbf{u} = (v, w) \) be the \( \Psi_x \)-coordinates of \( P \) (i.e. \( P = \Psi_x(\mathbf{u}) \)), and write \( E^s = E^s(x) \), \( E^u = E^u(x) \). The following identities hold:

\[
\angle(E^s, E^u) = \angle\left((d\Psi_x)|_{E^s}, (d\Psi_x)|_{E^u}\right), \quad \text{where} \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]

\[
\angle(V^s, V^u) = \angle\left((d\Psi_x)|_{V^s}, (d\Psi_x)|_{V^u}\right), \quad \text{where} \quad \mathbf{u}^s = \begin{pmatrix} F'(v) \\ 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{u}^u = \begin{pmatrix} F'(w) \\ 1 \end{pmatrix}.
\]

It is not difficult to see that the admissibility of \( V^s, V^u \) and the inequalities \(|v|, |w| < 10^{-2}\eta \) imply that \( |F'(v)|, |G'(v)| < \eta^{\beta/3} \).

We begin with the estimate of \( \frac{\sin \angle(V^s, V^u)}{\sin \angle(E^s, E^u)} = \frac{\sin \angle((d\Psi_x)|_{V^s}, (d\Psi_x)|_{V^u})}{\sin \angle((d\Psi_x)|_{E^s}, (d\Psi_x)|_{E^u})} \). By (A.3),

\[
\frac{\sin \angle(V^s, V^u)}{\sin \angle(E^s, E^u)} = \frac{\sin \angle(u^s, u^u)}{\sin \angle(e^s, e^u)} = \frac{\|u^s\| \|u^u\|}{\|e^s\| \|e^u\|} \cdot \frac{\det(d\Psi_x)_{e^s}}{\det(d\Psi_x)_{e^u}} \cdot \frac{\|d\Psi_x|_{e^s}\| \|d\Psi_x|_{e^u}\|}{\|d\Psi_x|_{u^s}\| \|d\Psi_x|_{u^u}\|}.
\]

**First factor:** The first factor equals \( \sin \angle(u^s, u^u) \). Using the formula for the sine of the difference of two angles it is not difficult to see that

\[
\sin \angle(u^s, u^u) = \frac{1}{\|u^s\| \|u^u\|} \cdot \det\left( \begin{pmatrix} 1 \\ F'(v) \end{pmatrix} F'(w) \right).
\]

Since \( |F'(v)|, |F'(w)| < \eta^{\beta/3} \), the first factor is \( e^{\pm 2\eta^{2\beta/3}} \).

**Second factor:** Since \( |F'(v)|, |F'(w)| < \eta^{\beta/3} \), 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:** \( \det(d\Psi_x)_{e^s} = \det(d\exp_x)_{C_\lambda(x)} \det C_\lambda(x) \), and \( \det(d\Psi_x)_{e^u} = \det C_\lambda(x) \), therefore the third factor is equal to \( \det(d\exp_x)_{C_\lambda(x)} \).

The exponential map on \( M \) is smooth, and \( \det(d\exp_x)_{\mathbf{u}} = 1 \), therefore there exists a constant \( K_1 \) which only depends on \( M \) s.t.

\[
|\det[(d\exp_x)_{\mathbf{u}}] - 1| < K_1 \|\mathbf{u}\| \quad \text{for all} \quad x \in M \quad \text{and} \quad \|\mathbf{u}\| < 1.
\]

Since \( C_\lambda(x) \) is a contraction (Lemma 2.5) and \( \|\mathbf{u}\| < 2\eta \), \( \det(d\exp_x)_{C_\lambda(x)} = 1 \pm 2K_1\eta \). Since \( 0 < \eta < \varepsilon \), \( 2K_2\eta \ll \sqrt{\eta} \) 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. \( \| (\Theta_D d\exp_x)_{\mathbf{u}} - \text{Id} \| < K_2 \|\mathbf{u}\| \) for all \( x \in D \in S \) and \( \|\mathbf{u}\| < 1 \) (cf. 3.6).

Write \( \mathbf{u} = C_\lambda(x)\mathbf{u} \), and choose some \( D \in S \) which contains \( \Psi_x[RQ_\lambda(x)(\mathbf{u})] \), then

\[
\|\Theta_D(d\Psi_x)|_{u^s} - \Theta_D(d\Psi_x)_{e^s}\| \leq \|\Theta_D(d\Psi_x)_{u^s} - \Theta_D(d\Psi_x)_{e^s}\| \|u^s\| + \|\Theta_D(d\Psi_x)_{e^s}\| \|u^s - e^s\| \leq \|\Theta_D(d\exp_x)_{u^s} - \text{Id} \| \|C_\lambda(x)\| \|u^s\| + 2\|C_\lambda(x)\| \|u^s - e^s\| \leq 3K_2\eta + 2\eta^{3/3}.
\]
because \( C_\lambda(x) \) is a contraction, \( ||u|| < 2\eta \), and \( x^* = (0_{\pm \beta} / \eta) \). Consequently, 
\[
||((d\Psi_x)_{2}^u) - ||(d\Psi_x)_{2}^e|| < (3K_2 + 2)\eta^{\beta/3}.
\]
Since also 
\[
||d\Psi_x||_{L^2} \geq ||C_\lambda(x)||^{-1},
\]
\[
||((d\Psi_x)_{2}^u) - ||(d\Psi_x)_{2}^e|| < (3K_2 + 2)||C_\lambda(x)||^{-1}\eta^{\beta/3}.
\]
It follows that for all \( \varepsilon \) small enough, 
\[
||((d\Psi_x)_{2}^u, (d\Psi_x)_{2}^e) - ((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)|| < \varepsilon^{3/\beta}||C_\lambda(x)||^{-1}\eta^{\beta/3} \cdot \eta^{2/3} < \varepsilon^{1/4} \eta^{\beta/3}.
\]
Putting all these estimates together, we see that 
\[
||d\Psi_x||_{L^2} = \exp \left[ \pm \left( \frac{1}{4} \eta^{3/4} \right) \right].
\]
How small depends only on \( K_2 \), and therefore only on the surface \( M \).

Similarly, one can show that 
\[
||((d\Psi_x)_{2}^u, (d\Psi_x)_{2}^e) - ((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)|| < \varepsilon^{3/\beta}||C_\lambda(x)||^{-1}\eta^{\beta/3} \cdot \eta^{2/3} < \varepsilon^{1/4} \eta^{\beta/3}.
\]

Putting all these estimates together, we see that 
\[
\sin \angle (V^*, V^*) \sin \angle (E^*, E^*) = \exp \left[ \pm (2\eta^{2\beta/3} + \eta^{2\beta/3} + \sqrt{\eta} + \frac{2}{3} \eta^{\beta/4}) \right].
\]
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 
\[
|\cos \angle (V^*, V^*) - \cos \angle (E^*, E^*)|.
\]
This is equal to 
\[
\frac{||((d\Psi_x)_{2}^u, (d\Psi_x)_{2}^e) - ((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} \leq \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1 + \frac{1}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} \cdot \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1 + \frac{1}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} \cdot \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1 + \frac{1}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} \cdot \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1.
\]
By \ref{eq:A8} and the estimate of the “fourth factor” above, this is smaller than 
\[
e^{\frac{3}{4} \eta^{3/4} \cdot \eta^{\beta/4}} \cdot ||C_\lambda(x)||^{-1} \cdot \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1.
\]
Since \( \Theta_D \) is an isometry, the difference of the inner products is equal to 
\[
\frac{||\Theta_D((d\Psi_x)_{2}^u, (d\Psi_x)_{2}^e) - \Theta_D((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} \cdot \frac{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||}{||((d\Psi_x)_{0}^e, (d\Psi_x)_{0}^e)||} - 1.
\]
This is smaller than 
\[
2K_2 \eta + 2\eta^{\beta/3}.
\]
The idea is to call the second coordinate $\tau$ in the first coordinate.

Claim 1. K3 is unique solution expanded as follows (Proposition 3.4): $\varepsilon \eta < \varepsilon \eta \leq Q \beta / h$.

Proof. We now argue as in (A.9) and deduce that

$$|\cos \angle(V^u, V^u) - \cos \angle(E^u, E^u)| \leq \varepsilon \varepsilon \eta \beta / h = \varepsilon \eta \beta / h.$$  

This is smaller than $2\eta \beta / h$, for all $\varepsilon$ small enough.

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 $\Psi^u, \nu^u$. We denote the parameters of $V^u$ by $\eta, \gamma, \varphi$, and $q$, and let $\eta := p^u \wedge p^u$. $V^u$ is admissible, so

$$\sigma \leq \frac{1}{2}, \eta \leq \frac{1}{2} \eta \beta / h, \varphi \leq 10^{-3} \eta, q = p^u, \text{ and } \text{Lip}(F) < \varepsilon, \quad (A.11)$$

see Definition 4.8 and Equation (4.2).

We analyze $\Gamma^u := \Psi_x^{-1}[f(V^u)] \subset \mathbb{R}^2$, looking for parameterizations of large $u$-sub-manifolds. Notice that

$$\Gamma^u = f_{xy}[\text{graph}(F)],$$

where $f_{xy} = \Psi_x^{-1} \circ f \circ \Psi_x$ and $\text{graph}(F) := \{(F(t), t) : |t| \leq q \}$.

Since $V^u$ is admissible, $\text{graph}(F) \subset \mathcal{R}(x, y)$. On this domain, $f_{xy}$ can be expanded as follows (Proposition 3.4):

$$f_{xy}(u, v) = (Au + h_1(u, v), Bv + h_2(u, v)) \quad (A.12)$$

where $C_f^{-1} < |A| < e^{-\chi}$, $e^\chi < |B| < C_f$; and $h_i$ are $C^{1+\beta} \varepsilon$-functions s.t. $|h_i(0)| < \varepsilon \eta$, $\|\nabla h_i(0)\| < \varepsilon \eta \beta / h$, and $\|\nabla h_i(u) - \nabla h_i(w)\| \leq \varepsilon \|u - w\| \beta / h$. Necessarily, $\|\nabla h_i\| < \varepsilon \eta \beta / h + \varepsilon \sqrt{2Q} \beta / h < 3e \sqrt{Q} \beta / h$ and $|h_i| < \varepsilon \eta + 3e \sqrt{Q} \beta / h$. $Q(x)$. Since $\eta \leq Q(x)$, and $Q(x) < e^{3\beta / h}$, the following holds for provided $\varepsilon$ is small enough:

$$\|\nabla h_i\| < 3e^2 \varepsilon \text{ and } |h_i| < \varepsilon^2 \text{ on graph}(F). \quad (A.13)$$

Using (A.12), we can put $\Gamma^u_y$ in the following form:

$$\Gamma^u_y = \{(AF(t) + h_1(F(t), t), Bt + h_2(F(t), t)) : |t| \leq q \}. \quad (A.14)$$

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 $\tau \in [-e^\chi - \sqrt{Q}, e^\chi - \sqrt{Q}]$, and

(a) $\text{Lip}(t) < e^{-\chi + \varepsilon}$;
(b) $|t(0)| < 2\pi$;
(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 $|t| \leq q$,

$$|\tau'(t)| \geq |B| - \max \|\nabla h_2\| \cdot \|F(t', 1)\| > |B| - 3e^2 \sqrt{1 + \varepsilon^2} \quad (\because (A.13), (A.11))$$

$$> |B|(1 - 3e^2 \sqrt{1 + \varepsilon^2}) \quad (\because |B| > e^\chi > 1)$$

$$> e^{-\varepsilon}|B| > 1 \text{ provided } \varepsilon \text{ is small enough}.$$
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, \( \tau[-q, q] \supset (\tau(0) - e^{-\varepsilon}|B|q, \tau(0) + e^{\varepsilon}|B|q) \).

The center of the interval can be estimated as follows:

\[
|\tau(0)| = |h_2(F(0), 0)| \leq |h_2(0)| + \max \|\nabla h_2\| \cdot |F(0)| \\
\leq \varepsilon \eta + 3\varepsilon^2 \cdot 10^{-3} \eta < 2\varepsilon \eta \quad \text{(admissibility and } (A.13))
\]

Recall that \( \eta \equiv p^u \land p^s \leq p^u \equiv q \), therefore \( |\tau(0)| < 2\varepsilon q \). Since \( |\tau'| > e^{-\varepsilon}|B| \),

\[
\tau[-q, q] \supseteq [2\varepsilon q - e^{-\varepsilon}|B|q, -2\varepsilon q + e^{-\varepsilon}|B|q] \supseteq [-(|B|e^{-\varepsilon} - 2\varepsilon)q, (|B|e^{-\varepsilon} - 2\varepsilon)q] \\
\supseteq [-|B|(e^{-\varepsilon} - 2\varepsilon)q, |B|(e^{-\varepsilon} - 2\varepsilon)q].
\]

Since \( |B|(e^{-\varepsilon} - 2\varepsilon) > e^\chi(e^{-2\varepsilon} - 2\varepsilon) > e^{\chi - \sqrt{\varepsilon}} \) for all \( \varepsilon \) small enough, \( \tau^{-1} \) is well-defined on \([e^{-\chi - \sqrt{\varepsilon}}q, e^{\chi - \sqrt{\varepsilon}}q] \).

Since \( t(\cdot) \) is the inverse of a \( |B|e^{-\varepsilon} \)–expanding map, \( \text{Lip}(t) \leq e\varepsilon |B|^{-1} < e^{-\chi + \varepsilon} \), proving (a).

We saw above that \( |\tau(0)| < 2\varepsilon q \). For all \( \varepsilon \) small enough, this is (much) smaller than \( e^{\chi - \sqrt{\varepsilon}}q \), therefore \( \tau(0) \) belongs to the domain of \( t \). It follows that

\[
|t(0)| = |t(0) - t(\tau(0))| < \text{Lip}(t)|\tau(0)| < e^{-\chi + \varepsilon} \cdot 2\varepsilon \eta.
\]

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 \( \varphi : [-q, q]^{d_1} \to \mathbb{R}^{d_2} \) \((0 < \alpha < 1)\) is defined by \( \|\varphi\|_\alpha := \|\varphi\|_\infty + \text{Hö} \alpha(\varphi) \), where

\[
\text{Hö} \alpha(\varphi) := \sup \left\{ \frac{\|\varphi(u) - \varphi(v)\|}{\|u - v\|^\alpha} : u, v \in [-q, q]^{d_1} \text{ different} \right\}.
\]

The following inequalities are easy to verify:

(1) \( \|\varphi \cdot \psi\|_\alpha \leq \|\varphi\|_\alpha \|\psi\|_\alpha \) for all \( \varphi, \psi \in C^\alpha[-q, q] \);
(2) \( \|\varphi \circ g\|_\alpha \leq \|\varphi\|_\infty + \text{Hö} \alpha(\varphi) \text{Lip}(g)^\alpha \) for all \( \varphi \)–Hölder and \( g \) Lipschitz;
(3) \( \text{Hö} \alpha(\varphi) \leq \|\varphi\|_\alpha \leq 1, \|1/(1 + \varphi)\|_\alpha \leq 1 - \|\varphi\|_\alpha \).

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:

\[
\left\| \frac{\partial h_2}{\partial x} (F(t(s)), t(s)) \right\|_{\beta/3} \leq \left\| \frac{\partial h_2}{\partial x} \right\|_\infty + \text{Höl}_{\beta/3}(\nabla h_2)[\text{Lip}(F \circ t, t)]^{\beta/3} \quad \therefore \text{(H2)}
\]

\[
< 3\varepsilon^2 + \varepsilon \cdot [\text{Lip}(F)^2(\text{Lip}(t))^2 + (\text{Lip}(t))^2]^{\beta/6}
\]

\[
< 3\varepsilon^2 + \varepsilon[\sqrt{\varepsilon^2 + 1(e^\varepsilon|B|^{-1})}]^{\beta/3} \quad \therefore \text{(A.11), (A.13)}
\]

\[
< \varepsilon, \text{ provided } \varepsilon \text{ is small enough.}
\]

\[
\left\| \frac{\partial h_2}{\partial y} (F(t(s)), t(s)) \right\|_{\beta/3} < \varepsilon \text{ (same proof)}.
\]

\[
\|F'(t(s))\|_{\beta/3} \leq \|F'\|_{\infty} + \|F'\|_{\beta/3} \text{Lip}(t)^{\beta/3} \text{ (see (H2) above)}
\]

\[
\leq \sigma + \sigma \cdot (e^{-\chi + \varepsilon})^{\beta/3} < 1 \text{ provided } \varepsilon \text{ is small enough.}
\]

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

\[
\Gamma_u^\tau \supset \{ (G(\tau), \tau) : |\tau| < e^{\chi - \sqrt{\varepsilon}} q \},
\]

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 \eta \eta < \eta \text{ provided } \varepsilon \text{ is small enough. Thus}

\[
|G(0)| \leq |A| \cdot |F(t(0))| + |h_1(F(t(0)), t(0))|
\]

\[
\leq |A|(|\varphi + 2\varepsilon \eta|) + |h_1(0)| + \max \|\nabla h_1\| \cdot \|(F(t(0)), t(0))\|
\]

\[
\leq |A|(\varphi + 2\varepsilon \eta) + \left[ \varepsilon \eta + 3\varepsilon^2 \sqrt{1 + 4\varepsilon^2} \right] \quad \therefore \text{(}F(t(0))\text{)} < \eta
\]

\[
\leq |A| \left[ \varphi + \eta(2\varepsilon^2 + 2\varepsilon + 3\varepsilon^2 \sqrt{1 + 4\varepsilon^2}) \right].
\]

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. \( |G'(0)| \leq |t'(0)| \left[ |A| \cdot |F'(t(0))| + ||\nabla h_1(F(t(0)), t(0))|| \cdot \|(F'(t(0)), 1)\| \right], \)

- \( |t'(0)| \leq \text{Lip}(t) < e^{-\chi + \varepsilon} \) (Claim 1).
- \( |F'(t(0))| < \gamma + \frac{2}{3}\varepsilon^{\beta/3}\eta^{\beta/3}, \) because by Claim 1(b)

\[
|F'(t(0))| < |F'(0)| + \text{Höl}_{\beta/3}(F)|t(0)|^{\beta/3} < \gamma + \sigma(2\varepsilon \eta)^{\beta/3} < \gamma + \frac{2}{3}\varepsilon^{\beta/3}\eta^{\beta/3}.
\]
\[ \| \nabla h_1(F(t(0)), t(0)) \| \leq 3 \varepsilon \eta^{2/3}, \text{ because } |F(t(0))| < \eta \text{ (proof of Claim 2)}, \]
\[ \text{and } |t(0)| < 2 \varepsilon \eta \text{ (Claim 1)}, \text{ so by the H"older regularity of } \nabla h_i, \]
\[ \| \nabla h_1(F(t(0)), t(0)) \| \leq \| \nabla h_1(0) \| + \varepsilon \left( \sqrt{|F(t(0))|^2 + |t(0)|^2} \right)^{\beta/3} \]
\[ \leq \varepsilon \eta^{2/3} + \varepsilon (\sqrt{\eta^2 + (2 \varepsilon \eta)^2})^{\beta/3} < 3 \varepsilon \eta^{2/3}. \]

- Putting these estimates together, we see that
\[ |G'(0)| < e^{-x + \varepsilon} \left[ \gamma + \frac{2}{3} \varepsilon^{\beta/3} \eta^{\beta/3} + |A|^{-1} \cdot 3 \varepsilon \eta^{2/3} \cdot 2 \right] \]
\[ < e^{-2x + \varepsilon} \left[ \gamma + \left( \frac{2}{3} \varepsilon^{\beta/3} + 6 C_f \varepsilon \right) \eta^{\beta/3} \right], \quad \therefore C_f^{-1} < |A| < e^{-x} \]
\[ \leq e^{-2x + \varepsilon} \left[ \gamma + \left( \frac{2}{3} \varepsilon^{\beta/3} + 6 C_f \varepsilon \right) \varepsilon^{\beta/3} (q^u \wedge q^s)^{\beta/3} \right], \quad \therefore p^u \wedge p^s \leq \varepsilon (q^u \wedge q^s). \]

This implies that for all \( \varepsilon \) small enough, \( |G'(0)| < e^{-2x + \varepsilon} \left[ \gamma + \varepsilon^{\beta/3} (q^u \wedge q^s)^{\beta/3} \right] \), which is stronger than the estimate in the claim.

Since \( \gamma \leq \frac{1}{2} (p^u \wedge p^s)^{\beta/3} \) and \( (p^u \wedge p^s) \leq \varepsilon (q^u \wedge q^s) \), 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, \( \|G'\|_{\beta/3} < e^{-2x + \sqrt{\varepsilon} [\sigma + \sqrt{\varepsilon}]} \), and \( \|G'\|_{\beta/3} < \frac{1}{2} \).

**Proof.** Differentiating, we see that \( G' = t' \cdot [AF \circ t + \frac{\partial h}{\partial x}(F \circ t, t)F' \circ t + \frac{\partial h}{\partial y}(F \circ t, t)] \).

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,
\- \( \|\frac{\partial h}{\partial x}(F \circ t, t)\|_{\beta/3} \leq \varepsilon \), and \( \|\frac{\partial h}{\partial y}(F \circ t, t)\|_{\beta/3} \leq \varepsilon \).

Thus by (H1), \( \|G'\|_{\beta/3} \leq |B|^{-1} e^{3\varepsilon} \left[ |A| \sigma + \varepsilon \sigma + \varepsilon \right]. \) Since \( \frac{1}{2} \leq \frac{3}{2} \leq C_f \), and \( C_f^{-1} < |A| < e^{-x} \), \( \|G'\|_{\beta/3} \leq e^{-2x + 3\varepsilon} \left[ \sigma + \frac{3}{2} C_f \varepsilon \right] \). If \( \varepsilon \) is small enough, then \( \|G'\|_{\beta/3} < e^{-2x + \sqrt{\varepsilon} [\sigma + \sqrt{\varepsilon}]} \), and \( \|G\|_{\beta/3} < \frac{1}{2} \).

**Claim 5.** For all \( \varepsilon \) small enough, \( \hat{V}^u := \Psi_y \{ (G(\tau), \tau) : |\tau| \leq \min \{ e^{x - \sqrt{\varepsilon} q}, Q_\varepsilon(y) \} \} \) is a \( u \)-manifold in \( \Psi_y \), the parameters of \( V^u \) satisfy (44), and \( \hat{V}^u \) contains a \( u \)-admissible manifold in \( \Psi_y^{q^u} \).

**Proof.** To see that \( \hat{V}^u \) 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
\[ |G'| \leq |G'(0)| + \text{H"older}(G)Q_\varepsilon(y)^{\beta/3} \leq \varepsilon + \frac{1}{2} \varepsilon < \sqrt{\varepsilon}, \text{ provided } \varepsilon \text{ is small enough.} \]

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 \( \hat{V}^u \) contains a \( u \)-admissible manifold in \( \Psi_y^{q^u} \). Since \( \Psi_y^{p^u} \to \Psi_y^{q^u} \), \( q^u = \min \{ e^{\varepsilon p^u}, Q_\varepsilon(y) \} \). Consequently, for every \( \varepsilon \) small enough,
\[ e^{x - \sqrt{\varepsilon} q} \equiv e^{x - \sqrt{\varepsilon} p^u} > e^{\varepsilon p^u} \geq q^u, \quad (A.15) \]
so \( \hat{V}^u \) restricts to a \( u \)-manifold with \( q \)-parameter equal to \( q^u \). Claims 2–4 guarantee that this manifold is \( u \)-admissible in \( \Psi_y^{q^\alpha - q^\beta} \), and that (4.14) holds.

Claim 6. \( f(V^u) \) contains exactly one \( u \)-admissible manifold in \( \Psi_y^{q^\alpha - q^\beta} \). 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 \( \Psi_y^{q^\alpha - q^\beta} \) 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, \( q^u < e^{\chi - \sqrt{\varepsilon}}q \). 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 \( F_u[V^u] \) denote the unique \( u \)-admissible manifold in \( \Psi_y^{q^\alpha - q^\beta} \) contained in \( f(V^u) \). We claim that \( F_u[V^u] \ni f(p) \) 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

\[
|\tau| = \text{second coordinate of } f_{xy}(F(0), 0) = |h_2(F(0), 0)|
\]

\[
\leq |h_2(0)| + \max \|\nabla h_2\| \cdot |F(0)| < \varepsilon \eta + 3\varepsilon^2 \cdot 10^{-3} \eta < e^{-\varepsilon} \eta < (q^u \wedge q^s) \leq q^u.
\]

Claim 7. \( f(V^u) \) intersects any \( s \)-admissible manifold in \( \Psi_y^{q^\alpha - q^\beta} \) at a unique point.

Proof. Let \( W^s \) be an \( s \)-admissible manifold in \( \Psi_y^{q^\alpha - q^\beta} \). We saw in the previous claim that \( f(V^u) \) contains a \( u \)-admissible manifold \( W^u \) in \( \Psi_y^{q^\alpha - q^\beta} \). 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 it 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

\[
\alpha < -e^{\chi - \sqrt{\varepsilon}}q < -q^u, \quad \beta > e^{\chi - \sqrt{\varepsilon}}q > q^u.
\]

We also saw that \( |\tau'| > e^{-\varepsilon}|B| \geq e^{\chi - \varepsilon} \). Consequently, the inverse function \( t : [\alpha, \beta] \to [-q, q] \) satisfies \( |t'(\tau)| < 1 \), and so

\[
f(V^u) = \Psi_y\{(G(\tau), \tau) : \tau \in [\alpha, \beta]\}, \text{ where Lip}(G) \leq \varepsilon.
\]

Let \( H : [-q^u, q^u] \to \mathbb{R} \) 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 \( \widehat{W}^s \supset W^s \). The same argument we used to prove Proposition 4.11 shows that \( f(V^u) \) and \( \widehat{W}^u \) 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 an \( s \)-admissible manifold w.r.t. \( f^{-1} \), and the parameters are the same.
2. \( \Psi_x^{p^\alpha - p^\beta} \to \Psi_y^{q^\alpha - q^\beta} \) w.r.t. \( f \) iff \( \Psi_y^{q^\alpha - q^\beta} \to \Psi_x^{p^\alpha - p^\beta} \) w.r.t. \( f^{-1} \). \( \square \)
Proof of Proposition\[4.14\] We prove the proposition for $F_u$, and leave the case of $F_s$ to the reader.

Suppose $Ψ_p^u \to Ψ_q^v$, and let $V^u$ be two $u$-admissible manifolds in $Ψ_p^u \to Ψ_q^v$.

We take $ε$ to be small enough for the arguments of the previous section to work.

We saw in the previous section that if $V_i = Ψ_x\{(F_i(t), t) : |t| \leq p^u\}$, then $F_u[V_i] = Ψ_y\{(G_i(τ), τ) : |τ| \leq q^u\}$, where

- $G_i(τ) = A_{F_i}(t_i(τ)) + k_1(F_i(t_i(τ)), t_i(τ))$;
- $t_i(τ)$ is defined implicitly by $B_{t_i}(τ) + h_2(F_i(t_i(τ)), t_i(τ)) = τ$, and $|t_i'| < 1$;
- $C_{f_i}^{-1} < |A| < e^{-λ} < |B| < C_f$;
- $|h_i(0)| < ε(p^u ∧ p^v)$. Hölβ/3(∇$h_u$) ≤ $ε$, and $max\|∇$h$\| < 3ε^2$.

In order to prove the proposition, we need to estimate $∥G_1 - G_2∥$. Taking differences, we see that

$$\|B\| \cdot |t_1 - t_2| \leq |h_2(F_1(t_1), t_1) - h_2(F_2(t_2), t_2)|$$

$$\leq \frac{∂h_2}{∂x} \|F_1(t_1) - F_2(t_2)\| + \frac{∂h_2}{∂x} \|t_1 - t_2\|$$

$$\leq 3ε^2(|F_1(t_1) - F_2(t_1)| + |F_2(t_1) - F_2(t_2)| + |t_1 - t_2|)$$

$$\leq 3ε^2(\|F_1 - F_2\| + (\operatorname{Lip}(F_2) + 1)|t_1 - t_2|)$$

$$\leq 3ε^2(\|F_1 - F_2\| + 3ε^2(1 + ε)|t_1 - t_2|), \text{ see } (4.2).$$

Rearranging terms, and recalling that $|B| > e^{x - ε}$, we see that

$$\|t_1 - t_2\| < \frac{3ε^2\|F_1 - F_2\|}{e^ε - ε^2(1 + ε)}.$$ 

The claim follows.

Part 1. For all $ε$ small enough, $∥G_1 - G_2∥_\infty < e^{-ε/2}∥F_1 - F_2∥_\infty$, whence \[15\].

Subtracting the defining equations for $G_i$, we find that

$$∥G_1 - G_2∥ \leq |A| \cdot |F_1(t_1) - F_2(t_2)| + |h_1(F_1(t_1), t_1) - h_1(F_2(t_2), t_2)|$$

$$\leq |A| \cdot |F_1(t_1) - F_2(t_2)| + \|∇h_1\|\sqrt{∥F_1(t_1) - F_2(t_2)∥^2 + |t_1 - t_2|^2}$$

$$\leq |A| + 3ε^2(∥F_1(t_1) - F_2(t_2)∥ + 3ε^2|t_1 - t_2|)$$

$$\leq |A| + 3ε^2(∥F_1 - F_2\|_\infty + \operatorname{Lip}(F_2)|t_1 - t_2|) + 3ε^2|t_1 - t_2|$$

$$\leq |A| + 3ε^2(1 + ε \cdot ε + 3ε^2 \cdot ε)\|F_1 - F_2\|_\infty, \text{ see part } 1.$$ 

$$\leq |A|(1 + 3C_fε^2)(1 + ε^2 + 3ε^3)\|F_1 - F_2\|_\infty$$

$$\leq e^{-ε}(1 + 3C_fε^2)(1 + ε^2 + 3ε^3)\|F_1 - F_2\|_\infty.$$ 

It follows that for every $ε$ small enough, $∥G_1 - G_2∥_\infty < e^{-ε/2}∥F_1 - F_2∥_\infty$.

Part 2. For all $ε$ small enough, $∥t_i' - t_i''∥_\infty < \sqrt{ε}(∥F_i' - F_i''∥_\infty + ∥F_i - F_i''∥^{β/3})$.

Differentiating both sides of the defining equation of $t_i$ gives

$$t_i' \left[ B + \frac{∂h_2}{∂x}(F_i \circ t_i, t_i)F_i' \circ t_i + \frac{∂h_2}{∂y}(F_i \circ t_i, t_i) \right] = 1.$$
Taking differences, we obtain after some re-arrangement

\[(t'_1 - t'_2) \left[ B + \frac{\partial h_2}{\partial x}(F_1 \circ t_1, t_2)F'_1 \circ t_1 + \frac{\partial h_2}{\partial y}(F_1 \circ t_1, t_1) \right] =
\]

\[ - t'_2 \left[ \frac{\partial h_2}{\partial x}(F_1 \circ t_1, t_1) \right] F'_1 \circ t_1 =: I
\]

\[ - t'_2 \left[ \frac{\partial h_2}{\partial y}(F_2 \circ t_2, t_2) \right] \left[ (F'_1 \circ t_1 - F'_2 \circ t_1) + (F'_2 \circ t_1 - F'_2 \circ t_2) \right] =: II
\]

\[ - t'_2 \left[ \frac{\partial h_2}{\partial y}(F_1 \circ t_1, t_1) \right] \left[ \frac{\partial h_2}{\partial y}(F_2 \circ t_2, t_2) \right] =: III
\]

Since \(|B| > e^x, |F'_1| < 1 \) and \( \|\nabla h_2\| < 3e^2 \),

\[ \|t'_1 - t'_2\|_{\infty} \leq \frac{1}{e^x - 6e^2} \|I + II + III\|_{\infty}. \]

Since I, II and III involve partial derivatives of \( h_2 \) evaluated at \((F_i \circ t_1, t_1)\), we begin by analyzing \( \nabla h_2(F_1 \circ t_1, t_1) \). Since \( \text{Hö}l_{1/3}(\nabla h_1) \leq \varepsilon \),

- \( \|\nabla h_2(F_1 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_1, t_1)\| \leq \varepsilon \|F_1 - F_2\|_{\infty}^{1/3} \);
- \( \|\nabla h_2(F_2 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_2, t_1)\| \leq \varepsilon \|t_1 - t_2\|_{\infty}^{1/3} \) (because \( \text{Lip}(F_2) < 1 \));
- \( \|\nabla h_2(F_2 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_2, t_2)\| \leq \varepsilon \|t_1 - t_2\|_{\infty}^{1/3} \).

By part 1, \( \|t_1 - t_2\|_{\infty} \leq \varepsilon \|F_1 - F_2\|_{\infty} \). It follows that

\[ \|\nabla h_2(F_1 \circ t_1, t_1) - \nabla h_2(F_2 \circ t_2, t_2)\| < 3\varepsilon \|F_1 - F_2\|_{\infty}^{1/3}. \]

Using the facts that \(|t'_1| < 1, |F'_1| < 1, \text{Lip}(F_2) < 1, \) and \( \text{Hö}l_{1/3}(F'_2) < 1 \) (see the definition of admissible manifolds and the proof of Proposition 1.12), we get that

\[ |I| \leq 3\varepsilon \|F_1 - F_2\|_{\infty}^{1/3}; \]

\[ |II| \leq 3\varepsilon^2 (\|F'_1 - F'_2\|_{\infty} + \|t_1 - t_2\|_{\infty}^{1/3}) \leq 3\varepsilon^2 \|F'_1 - F'_2\|_{\infty} + 3\varepsilon^2 \|F_1 - F_2\|_{\infty}^{1/3}; \]

\[ |III| \leq 3\varepsilon \|F_1 - F_2\|_{\infty}^{1/3}. \]

So for all \( \varepsilon \) sufficiently small, \( \|t'_1 - t'_2\|_{\infty} < \sqrt{\varepsilon} (\|F'_1 - F'_2\|_{\infty} + \|F_1 - F_2\|_{\infty}^{1/3}) \).

**Part 4.** \( \|G'_1 - G'_2\|_{\infty} < e^{-x/2} (\|F'_1 - F'_2\|_{\infty} + \|F_1 - F_2\|_{\infty}^{1/3}) \).

By the definition of \( G_i \), \( G'_i = t'_i [AF'_i \circ t_i + \frac{\partial h_1}{\partial x}(F_1 \circ t_i, t_i)F'_i \circ t_i + \frac{\partial h_1}{\partial y}(F_1 \circ t_i, t_i)] \).

Taking differences, we see that

\[ |G'_1 - G'_2| \leq |t'_1 - t'_2| \left[ AF'_1 \circ t_1 + \frac{\partial h_1}{\partial x}(F_1 \circ t_1, t_1)F'_1 \circ t_1 + \frac{\partial h_1}{\partial y}(F_1 \circ t_1, t_1) \right] =: I'
\]

\[ + |t'_2| \cdot |A| \cdot \left( |F'_1 \circ t_1 - F'_2 \circ t_1| + |F'_2 \circ t_1 - F'_2 \circ t_2| \right) \] =: II'

\[ + |t'_2| \left( \frac{\partial h_1}{\partial x}(F_1 \circ t_1, t_1) - \frac{\partial h_1}{\partial x}(F_2 \circ t_2, t_2) \right) \left| F'_1 \circ t_1 \right| \] =: III'

\[ + |t'_2| \left( \frac{\partial h_1}{\partial y}(F_2 \circ t_2, t_2) \right) \left| F'_1 \circ t_1 - F'_2 \circ t_2 \right| \] =: IV'

\[ + |t'_2| \left( \frac{\partial h_1}{\partial y}(F_1 \circ t_1, t_1) - \frac{\partial h_1}{\partial y}(F_2 \circ t_2, t_2) \right) \] =: V'
Using the same arguments that we used in part 1, one can show that
\[ I' \leq \| t_1' - t_2' \|_\infty (e^{-\chi} + 6e^2) < \sqrt{2} (\| F_1' - F_2' \|_\infty + \| F_1 - F_2 \|_\infty^{\beta/3}) \]
\[ II' \leq e^{-\chi} (\| F_1' - F_2' \|_\infty + \| t_1 - t_2 \|_\infty^{\beta/3}) \leq e^{-\chi} (\| F_1' - F_2' \|_\infty + \| F_1 - F_2 \|_\infty^{\beta/3}) \] (part 1)
\[ III' \leq 3\varepsilon \| F_1 - F_2 \|_\infty^{\beta/3} \] (see the estimate of I in part 3)
\[ IV' \leq 3\varepsilon^2 \| F_1' - F_2' \|_\infty + 3\varepsilon^3 \| F_1 - F_2 \|_\infty^{\beta/3} \] (see the estimate of II in part 3)
\[ V' \leq 3\varepsilon \| F_1 - F_2 \|_\infty^{\beta/3} \] (see the estimate of III in part 3).

It follows that \( \| G_1' - G_2' \|_\infty < (e^{-\chi} + 10 \varepsilon + \sqrt{2}) (\| F_1' - F_2' \|_\infty + \| F_1 - F_2 \|_\infty^{\beta/3}) \). 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 \( V^* \) is an \( s \)-admissible manifold in \( \Psi_{p^s-p^s} \) which stays in windows, then there is a positive chain \( (\Psi_{p^s-p^s})_{i \geq 0} \) s.t. \( \Psi_{p_i^s-p_0^s} = \Psi_{p^s-p^s} \), and there are \( s \)-admissible manifolds \( W_i^s \) in \( \Psi_{p_i^s-p_i^s} \) s.t. \( f^i(V^*) \subset W_i^s \) for all \( i \geq 0 \). We write
\[ V^s = \Psi_x \{ (t, F_0(t)) : |t| \leq p_s \}, \]
\[ W_i^s = \Psi_x \{ (t, F_i(t)) : |t| \leq p_i^s \}, \]
\[ \eta_i := p_i^s \land p_0^s. \]

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 (4.12), Lip\( (F_i) < \varepsilon \).

**Part 1.** If \( \varepsilon \) is so small that \( e^{-\chi} + 4e^2 < e^{-\chi/2} \), then for every \( y, z \in V^* \),
\[ d(f^k(y), f^k(z)) \leq 6\varepsilon e^{-\frac{k}{2}} \] for all \( k \geq 0 \).

**Proof.** Since \( V^* \) stays in windows, \( f^k(V^*) \subset \Psi_x [R_{Q_s(x_s)}(\Omega)] \) for all \( k \geq 0 \). Therefore, for any \( y, z \in V^* \), one can write \( f^k(y) = \Psi_x(y_k) \) and \( f^k(z) = \Psi_x(z_k) \), where \( y_k = (y_k, F_k(y_k)), z_k = (z_k, F_k(z_k)) \) belong to \( R_{Q_s(x_s)}(\Omega) \).

For every \( k \), \( y_{k+1} = f_{x_{s+1}}(y_k) \) and \( z_{k+1} = f_{x_{s+1}}(z_k) \), where \( f_{x_{s+1}} := \Psi_x^{-1} \circ f \circ \Psi_x. \) By (3.3),
\[ f_{x_{s+1}}(v, w) = (A_{kv} + h_1(v, w), B_{kw} + h_2(v, w)) \] on \( R_{Q_s(x_s)}(\Omega) \),
where \( C^1_f < |A_k| < e^{-\chi}, e^\chi < |B_k| < C_f, \) and max \( |\nabla h_k| < 3e^2 \). Thus
\[ |y_{k+1} - z_{k+1}| \leq |A_k| \cdot |y_k - z_k| + 3e^2 (|y_k - z_k| + \text{Lip}(F_k)|y_k - z_k|) \]
\[ \leq (e^{-\chi} + 4e^2)|y_k - z_k| < e^{-\frac{1}{2}} |y_k - z_k| \leq \cdots \leq e^{-\frac{1}{2}(k+1)\chi} |y_0 - z_0|. \]
Since \( y_k, z_k \) are on the graph of an \( s \)-admissible manifold in \( \Psi_{p_i^s-p_i^s} \), 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 \( y_k = (y_k, F_k(y_k)), z_k = (z_k, F_k(z_k)) \), and Lip\( (F_k) < \varepsilon, \| y_k - z_k \| < 3p_0^s e^{-\frac{1}{2}k\chi} \).

Pesin charts have Lipschitz constant less than two, so \( d(f^k(y), f^k(z)) < 6\varepsilon p_0 e^{-\frac{1}{2}k\chi} \).

**Part 2.** Suppose \( \varepsilon \) is so small that \( e^{-\chi} + 3e^2 + 3e^3 < e^{-\frac{1}{2}\chi} \) and \( C_{\varepsilon} + 3e^2 < 1 \). For every \( y \in V^* \), let \( e^i(y) \) denote the positively oriented unit tangent vector to \( V^* \) at \( y \). If \( y \in V^* \), then \( \| d_k e^i(y) \| \leq 6e^{-\frac{1}{2}k\chi} |C_{\chi}(x_s) |^{-1} \) for all \( k \geq 0 \).

**Proof.** If \( y \in V^* \), then \( f^k(y) \in W_k^s \subset \Psi_x [R_{Q_s(x_s)}(\Omega)] \). So \( d_k f^i e^i(y) = (d\Psi_x)(y_k)(d_k h_k) \) where \( (d_k h_k) \) is tangent to the graph of \( F_k \). Since Lip\( (F_k) < \varepsilon, |h_k| < \varepsilon |a_k| \) for all \( k \).
The identity \((a_{k+1}) = (df_{x_kx_{k+1}})\) holds. Since \(\|\nabla h_i\| \leq 3\varepsilon^2\),
\[
\begin{pmatrix}
(a_{k+1}) \\
(b_{k+1})
\end{pmatrix}
= \begin{pmatrix}
A_k + \frac{\partial h_i}{\partial x}(y_k) \\
B_k + \frac{\partial h_i}{\partial y}(y_k)
\end{pmatrix}
\begin{pmatrix}
a_k \\
b_k
\end{pmatrix}
= \begin{pmatrix}
(A_k + 3\varepsilon^2) a_k + 3\varepsilon^2 |b_k| \\
(B_k + 3\varepsilon^2) b_k + 3\varepsilon^2 |a_k|
\end{pmatrix}.
\]
It follows that \(|a_{k+1}| \leq (|A_k| + 3\varepsilon^2 + 3\varepsilon^2)|a_k|\). By the bounds on \(A_k\) and \(B_k\) and the assumption on \(\varepsilon\),
\[
|a_k| \leq e^{-\frac{3}{2} \varepsilon^2} |a_0| \text{ and } |b_k| \leq e^{\frac{3}{2} \varepsilon^2} |a_0|.
\]

Returning to the defining relation \(df_{x_k}^n e^*(y) = (d\Psi_{x_k}) (a_k)\), and recalling that \(\|d\Psi_{x_k}\| \leq 2\) (Theorem 2.7), we see that \(\|df_{x_k}^n e^*(y)\| \leq 2\sqrt{2} e^{-\frac{3}{2} \varepsilon^2} |a_0|\).

Since \((a_k) = (d\Psi_{x_k})^{-1} e^*(y),|a_0| \leq \|d\Psi_{x_k}\|\), so \(\|df_{x_k}^n e^*(y)\| \leq 2\sqrt{2} e^{-\frac{3}{2} \varepsilon^2} \|d\Psi_{x_k}\|\).

For every \(x,\|d\Psi_{x_k}\| \leq 2\|C_{\chi}(x)^{-1}\|\) because \(C_{\chi}(x)^{-1}\) maps \(B_{2\rho(M)}(\emptyset)\) into \(B_{2\varepsilon}(\emptyset) \subset B_{\rho(M)}(\emptyset)\), provided \(\varepsilon < \frac{1}{4} \rho(M),\) and by the definition \(\rho(M)\) is so small that \(\|d\exp_{y_{x_k}}^{1}\| \leq 2\) for all \(x \in M\) and \(y \in B_{\rho(M)}(\emptyset)\).

It follows that \(\|df_{x_k}^n e^*(y)\| \leq 6\|C_{\chi}(x_0)^{-1}\| e^{-\frac{3}{2} \varepsilon^2}\).

**Part 3.** The following holds for all \(\varepsilon\) small enough: for all \(y, z \in V^*\) and \(n \geq 0,\)
\[
\|\log|df_{y}^n e^*(y)| - \log|df_{z}^n e^*(z)|\| \leq Q_v(x_0)\varepsilon^{3/2}.
\]

**Proof.** Call the quantity to be estimated \(A.\) For every \(p \in V^*,\)
\[
df_{y}^n e^*(p) = df_{f(p)}^{n-1}(df_{x_k} e^*(p)) = \pm \|df_{x_k} e^*(p)\| \cdot df_{f(p)}^{n-1}(e^*(f(p))
\]
\[
= \cdots = \sum_{k=0}^{n-1} \|df_{x_k} e^*(f(k(p)))\| e^*(f(p)).
\]

Thus \(A := \log|df_{x_k} e^*(y)| \leq \sum_{k=0}^{n-1} \|df_{x_k} e^*(f(k(y)))\| - \log|df_{x_k} e^*(f(k(z)))|\). We shall estimate the sum term-by-term, using the Hölder continuity of \(df\).

In section 3.3, we covered \(M\) by a finite collection \(\mathcal{D}\) of open sets \(D_i\), equipped with a smooth map \(\Theta_D : TD \to \mathbb{R}^2\) s.t. \(\Theta_D|_{T_x M} : T_x M \to \mathbb{R}^2\) is an isometry, and \(\partial x := \Theta_D^{-1}|_{\partial x} : \mathbb{R}^2 \to TD\) has the property that \((x, y) \mapsto \partial x(y)\) is Lipschitz on \(D \times B_1(0)\). Since \(f\) is \(C^{1+\beta}\) diffeomorphism and \(M\) is compact, \(df_{y}^n|_{\mathbb{R}^2}\) depends in a \(\beta\)-Hölder way on \(p,\) and in a Lipschitz way on \(y, z\). 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 \(u, v \in \mathbb{R}^2\) of length one, \(\|df_{y}(\partial y(u))\| = \|df_{z}(\partial z(v))\| < H_0 (d(y, z)^\beta + \|u - v\|)\).

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 \(\text{Id} = \Theta_{D_k} \circ \partial f^k(y)\) and \(\text{Id} = \Theta_{D_k} \circ \partial f^k(z),\) we see that
\[
A \leq \sum_{k=0}^{n-1} \log|df_{x_k} e^*(f(k(y)))\| - \log|df_{x_k} e^*(f(k(z)))|\|
\]
\[
\leq \sum_{k=0}^{n-1} \|df_{y}^n e^*(f(k(y)))\| - \|df_{x_k} e^*(f(k(z)))|\|
\]
\[
\leq H_0 (6\rho_0)^\beta + \sum_{k=0}^{n-1} \|\Theta_{D_k} e^*(f^k(y)) - \Theta_{D_k} e^*(f^k(z))\|,\] by part 1. (A.16)
We estimate $N_k := \|\Theta_{D_k}(\exp_k^* f^k(y)) - \Theta_{D_k}(\exp_k^* f^k(z))\|$. By definition, $\exp_k^* (f^k(y))$ and $\exp_k^* (f^k(z))$ are the positively oriented unit tangent vectors to $f^k(V^s) \subset W^s_k$, at $f^k(y)$ and $f^k(z)$. Defining $y_k$ and $z_k$ as before, we obtain

$$
\exp_k^* (f^k(y)) = \frac{(d\Psi_{y_k})_{y_k} \left( \frac{1}{\|F_k'(y_k)\|} \right)}{(d\Psi_{z_k})_{z_k} \left( \frac{1}{\|F_k'(z_k)\|} \right)}, \quad \exp_k^* (f^k(z)) = \frac{(d\Psi_{z_k})_{z_k} \left( \frac{1}{\|F_k'(z_k)\|} \right)}{(d\Psi_{y_k})_{y_k} \left( \frac{1}{\|F_k'(y_k)\|} \right)}.
$$

We saw in part 1 that $\|\exp_k^* F_k(z)\|^{-1}$ and $\|\exp_k^* F_k(y)\|^{-1}$ are bounded by $2\|C(x_k)^{-1}\|$, so the denominators are bounded below by $\frac{1}{2}\|C(x_k)^{-1}\|^{-1}$. Since for any two non-zero vectors $u, v$, $\|u - v\| < 2\|u - v\|$, we have

$$
N_k \leq 2\|C(x_k)^{-1}\| \cdot \|\Theta_{D_k}(\exp_k^* y_k) - \Theta_{D_k}(\exp_k^* z_k)\| = 2\|C(x_k)^{-1}\| \cdot \|\Theta_{D_k}(\exp_k^* y_k) - \Theta_{D_k}(\exp_k^* z_k)\|.
$$

On $D_k$ we can write $\Psi_{x_k} = \exp_{x_k} \circ \vartheta_{x_k} \circ C_{x_k}$, where $\vartheta_{x_k} \circ C_{x_k} = C(x_k)$. Let $y_k := C(x_k) y_k$, $z_k := C(x_k) z_k$, and $v_k := C_{x_k}(F_k'(y_k))$. Then $N_k \leq 2\|C(x_k)^{-1}\| \cdot \|\Theta_{D_k}(\exp_k^* y_k) - \Theta_{D_k}(\exp_k^* z_k)\| + 2\|C(x_k)^{-1}\| \cdot \|\Theta_{D_k}(\exp_k^* v_k) - \Theta_{D_k}(\exp_k^* v_k)\| + 2\|\exp_k^* v_k - \exp_k^* v_k\|.

We study this expression. In what follows we identify the differential of a linear map with the map itself.

By construction, the map $(x, y, v) \mapsto \Theta_D \circ (\exp_x)_y \circ \vartheta_x(v)$ is smooth on $D \times B_2(0) \times B_2(0)$. Therefore there exists a constant $E_0 > 1$ s.t. for every $(x, y, v) \in D \times B_2(0) \times B_2(0)$ and every $D \in D$,

$$
\|\Theta_{D}(\exp_x)_y \circ \vartheta_x(v)\| = \|\Theta_{D}(\exp_x)_y \circ \vartheta_x(v)\| \leq E_0(\|u_0 - v_0\| + \|v_0 - v_0\|).
$$

It follows that

$$
N_k \leq 2\|C(x_k)^{-1}\| \cdot \left( \|y_k - z_k\|^{\beta/3} + E_0(\|y_k - z_k\| + \|z_k - v_k\|) \right) \leq 2\|C(x_k)^{-1}\| \cdot \left( \|y_k - z_k\|^{\beta/3} + E_0(\|y_k - z_k\| + \|z_k - v_k\|) \right) \leq 6E_0\|C(x_k)^{-1}\| \cdot \|y_k - z_k\|^{\beta/3} \quad (\because E_0 > 1)
$$

$$
\leq 6E_0\|C(x_k)^{-1}\| \cdot \left(3\|F_k\|^{\beta/3} e^{-\frac{1}{2} \beta x_k} \right) \text{ because } \|y_k - z_k\| < 3\|F_k\|^{\beta/3} e^{-\frac{1}{2} \beta x_k} \text{ (part 1)}
$$

$$
\leq 9E_0\|C(x_k)^{-1}\| \cdot (p_k)^{\beta/3} e^{-\frac{1}{2} \beta x_k}.
$$

By the definition of $Q_k(\cdot)$, $\|C_k(x_k)^{-1}\| \leq \varepsilon_1/4 \|\exp_k^* F_k(y_k) - \exp_k^* F_k(z_k)\| \leq \varepsilon_1/4 (p_k)^{-\beta/12}$, and therefore $N_k \leq \varepsilon_1/4 E_0(p_k)^{-\beta/12} (p_k)^{\beta/3} e^{-\frac{1}{2} \beta x_k}$. Since $(\Psi_{x_k}^*, \vartheta_{x_k})_{x_k \in X}$ is a chain,
$p^*_n = \min\{\varepsilon p^*_{n+1}, Q_\varepsilon(x_i)\} \leq e^\varepsilon p^*_{n+1}$ for all $i$, whence $p^*_0 \leq e^\varepsilon p^*_{n}$. It follows that for all $\varepsilon$ small enough,

$$N_k \leq 9e^{1/4}E_0(p^*_0)^{3/4}\exp[-\frac{1}{4}\beta \chi]k.$$  \hfill (A.17)

Plugging this in (A.16), we obtain

$$\log \left| \frac{df^n_{\nu}e^*(y)}{df^0_{\nu}e^*(z)} \right| \leq \left( \frac{6^3 H_0(p^*_0)^{3/4} + 9e^{1/4}E_0H_0}{1 - e^{-\frac{1}{2}\beta \chi}} \right)(p^*_0)^{3/4}$$

$$\leq \left( \frac{9e^{3/4}E_0H_0}{1 - e^{-\frac{1}{2}\beta \chi}} \right) Q_\varepsilon(x_0)^{3/4}.$$  

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 $\chi$.

**Proof of Proposition 6.4** We continue to use the notation of the previous proof.

Assume that $V^* \cap U^* \neq \emptyset$. We show that $V^* \subseteq U^*$ or $U^* \subseteq V^*$.

Since $V^*$ stays in windows, there is a positive chain $(\Psi^*_{x_i}^{-1}p^*_i)_{i \geq 0}$ such that $\Psi^*_{x_i}^{-1}p^*_i = \Psi_{x_i}p^*_i$ and such that for all $i \geq 0$, $f^i(V^*) \subseteq W^*_{i+1}$ where $W^*_{i+1}$ is an $s$-admissible manifold in $\Psi^{i+1}_{x_i}p^*_i$.

**Claim 1.** The following holds for all $\varepsilon$ small enough. $f^n(V^*) \subseteq \Psi_{x_i}[R_{\frac{1}{2n}Q_\varepsilon(x_i)}(0)]$ for all $n$ large enough.

**Proof.** Suppose $y \in V^*$, and write as in part 1 of the previous proof, $f^n(y) = \Psi_{x_n}(y_n)$ where $y_n = (y_n,F_n(y_n))$ and $F_n$ is the function which represents $W^*_{i+1}$ in $\Psi_{x_n}$. We have $y_{n+1} = f_{x_n,x_{n+1}}(y_n)$, which implies in the notation of the previous proof that if $\varepsilon$ is small enough, then

$$|y_{n+1}| \leq |A_n| |y_n| + |h_\lambda(y_n)| \leq |A_n| |y_n| + |h_\lambda(0)| + \|\nabla h_\lambda\|(|y_n| + |F_n(y_n)|)$$

$$< e^{-\chi}|y_n| + \varepsilon_\eta_n + 3\varepsilon^2(|y_n| + p^*_n) < (e^{-\chi} + 3\varepsilon^2)|y_n| + 2\varepsilon p^*_n$$

$$< (e^{-\chi} + 3\varepsilon^2)|y_n| + 2\varepsilon \min\{e^\varepsilon p^*_{n+1}, Q_\varepsilon(x_n)\}$$

$$< (e^{-\chi} + 3\varepsilon^2)|y_n| + 2e\varepsilon p^*_{n+1} < e^{-\chi/2}|y_n| + 4\varepsilon p^*_{n+1}.$$  

We see that $|y_n| \leq a_n$ where $a_n$ is defined by induction by

$$a_0 := Q_\varepsilon(x_0)$$

and $a_{n+1} = e^{-\chi/2}a_n + 4\varepsilon p^*_{n+1}$. We claim that if $\varepsilon$ is small enough, then $a_n < \frac{1}{4}p^*_n$ for some $n$. Otherwise, $p^*_n \leq 4a_n$ for all $n$, whence $a_{n+1} \leq (e^{-\chi/2} + 16\varepsilon)a_n$ for all $n$, which implies that $a_n < (e^{-\chi/2} + 16\varepsilon)a_0$. But by assumption, $a_n \geq \frac{1}{4}p^*_n \geq \frac{1}{4}(p^*_n \wedge p^*_n) \geq \frac{1}{4}e^{-\varepsilon n}(p^*_n \wedge p^*_n)$ (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. $a_n < \frac{1}{4}p^*_n$.

It is clear from the definition of $a_n$, that if $\varepsilon$ is small enough then $a_n < \frac{1}{4}p^*_n \implies a_{n+1} < \frac{1}{4}p^*_{n+1}$. Thus $a_n < \frac{1}{4}p^*_n$ for all $n$ large enough.

In particular, $|y_n| < \frac{1}{4}Q_\varepsilon(x_n)$ for all $n$ large enough. Since $y_n = (y_n,F_n(y_n))$ and $|F_n(y_n)| \leq |F_n(0)| + \text{Lip}(F_n)|y_n| < (10^{-3} + \varepsilon)Q_\varepsilon(x_n), |y_n| < \frac{1}{2}Q_\varepsilon(x_n)$ for all $n$ large enough.

**Claim 2.** The following holds for all $\varepsilon$ small enough: $f^n(U^*) \subseteq \Psi_{x_n}[R_{Q_\varepsilon(x_n)}(0)]$ for all $n$ large enough.
Proof. $U^s$ stays in windows, so there exists a positive chain $\{\Psi_{y_i}^{q_i^{n-\alpha}}, \Omega_i\}_{i \geq 0}$ such that $\Psi_{y_0}^{q_0^{n-\alpha}} = \Psi_y^{q_i^{n-\alpha}}$ and such that for all $i \geq 0$, $f^i(U^s)$ is a subset of an $s$-admissible manifold in $\Psi_{y_i}^{q_i^{n-\alpha}}$.

Let $z$ be a point in $U^s \cap V^s$. By Part 1 of Theorem 6.3, for any $w \in U^s$ $d(f^n(z), f^n(w)) \leq 6 \Omega_0 e^{-\frac{1}{2}nx}$. Therefore $f^n(z), f^n(w) \in B_{\Omega_0(x_n) + 6\Omega_0}(x_n) \subset B_{\eta}(x_n)$.

If $\varepsilon < \frac{1}{\rho(M)}$ (cf. 6.2.3), then $\|\exp_{x_\varepsilon}^{-1}[f^n(z)] - \exp_{x_\varepsilon}^{-1}[f^n(w)]\| < 12e^{-\frac{1}{2}nx}q_0$, so

$$\|\Psi_{x_\varepsilon}^{-1}[f^n(z)] - \Psi_{x_\varepsilon}^{-1}[f^n(w)]\| < \|\Psi_{x_\varepsilon}^{-1}[f^n(z)] - \Psi_{x_\varepsilon}^{-1}[f^n(w)]\| < 12\left(\frac{q_0}{\rho_0}\right) e^{-\frac{1}{2}nx} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Since $p_i^n \leq \Omega_0(x_n) \ll \|\Psi_{x_\varepsilon}^{-1}[f^n(z)] - \Psi_{x_\varepsilon}^{-1}[f^n(w)]\| < \Omega_0(x_n)$ for all $n$ large enough. The estimates are uniform in $w \in U^s$, so the claim is proved.

Claim 3. Recall that $V^s$ is $s$-admissible in $\Psi_{x_i}^{q_i^{n-\alpha}}$ and $U^s$ is $s$-admissible in $\Psi_y^{q_i^{n-\alpha}}$.

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. $f^n(U^s), f^n(V^s) \subseteq \Psi_{x_n}[R_{\Omega_0(x_n)}(0)]$ for all $n \geq n_0$, then $f^{n_0}(V^s), f^{n_0}(U^s) \subseteq W^s := \Psi_{x_0}[\{\Psi_{x_i}^{P_i^{n_i}}, \Omega_i\}_{i \geq n_0}]$ (Proposition 1.12 (4)).

Let $G$ denote the function which represents $W^s$ in $\Psi_{x_0}$, then $\Psi_{x_0}^{-1}[f^n(U^s)]$ and $\Psi_{x_0}^{-1}[f^n(V^s)]$ are two connected subsets of graph($G$). Write

$$f^n(U^s) = \Psi_{x_n}\{(t, G(t)) : t \in [\alpha, \beta]\},$$
$$f^n(V^s) = \Psi_{x_n}\{(t, G(t)) : t \in [\alpha', \beta']\}.$$

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 $V^s \subseteq U^s$. \qed
Acknowledgements

The author would like to thank J. Buzzi, A. Katok, F. Ledrappier, and M. Pollicott for useful discussions.

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. dissertation, University of Minnesota (1967).

[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, 119161.

[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.

[Br1] 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), no. 2, 247–280. Erratum: Comm. Math. Phys. 107 (1986), no. 2, 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. DOI 10.1007/s00222-011-0317-8 (2011)

[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 (to appear)

[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. Gurevich: 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. Gurevich: 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.

[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] Ya. 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.

[Rut] S. Ruette: On the Vere–Jones classification and existence of maximal measures for countable topological Markov shifts, Pacific J. Math. 209 (2003), 365–380.

[Sal] I.A. Salama: Topological entropy and recurrence of countable chains, Pacific J. Math. 134 (1988) 325–341. Errata: 140 (1989), 397.

[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 β–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, POB 26, Rehovot, Israel.

E-mail address: omsarig@gmail.com