# SYMBOLIC DYNAMICS FOR THREE DIMENSIONAL FLOWS WITH POSITIVE TOPOLOGICAL ENTROPY

YURI LIMA AND OMRI M. SARIG

Abstract. We construct symbolic dynamics on sets of full measure (with respect to an ergodic measure of positive entropy) for $C^{1+\varepsilon}$ flows on closed smooth three dimensional manifolds. One consequence is that the geodesic flow on the unit tangent bundle of a closed $C^\infty$ surface has at least const $\times(e^{hT}/T)$ simple closed orbits of period less than $T$, whenever the topological entropy $h$ is positive – and without further assumptions on the curvature.

## 1. Introduction

The aim of this paper is to develop symbolic dynamics for smooth flows with topological entropy $h>0$ on three dimensional closed (compact and boundaryless) Riemannian manifolds.

Earlier works treated geodesic flows on hyperbolic surfaces [Ser81, Ser87, KU07], geodesic flows on surfaces with variable negative curvature [Rat69], and uniformly hyperbolic flows in any dimension [Rat73, Bow73]. This work only assumes that $h>0$ and that the flow has *positive speed* (i.e. the vector that generates the flow has no zeroes). This generality allows us to cover several cases of interest that could not be treated before, for example:

(1) *Geodesic flows with positive entropy in positive curvature:* There are many Riemannian metrics with positive curvature somewhere (even everywhere) whose geodesic flow has positive topological entropy [Don88, BG89, KW02, CBP02].
(2) *Reeb flows with positive entropy:* These arise from Hamiltonian flows on surfaces of constant energy, see [Hut10]. Examples with positive topological entropy are given in [MS11]. (This application was suggested to us by G. Forni.)
(3) *Abstract non-uniformly hyperbolic flows in three dimensions*, see [BP07, Pes76].

The statement of our main result is somewhat technical, therefore we begin with a down-to-earth corollary. Let $\varphi$ be a flow. A *simple closed orbit* of *length* $\ell$ is a parameterized curve $\gamma(t)=\varphi^t(p)$, $0\leq t\leq \ell$ s.t. $\gamma(0)=\gamma(\ell)$ and $\gamma(0)\neq\gamma(t)$ when $0<t<\ell$. The *trace* of $\gamma$ is defined to be the set $\{\gamma(t):0\leq t\leq \ell\}$. Let $[\gamma]$ denote the equivalence class of the relation $\gamma_1\sim\gamma_2 \Leftrightarrow \gamma_1,\gamma_2$ have equal lengths and traces. Let $\pi(T):=\#\{[\gamma]:\ell(\gamma)\leq T,\gamma \text{ is simple}\}$.

**Theorem 1.1.** *Suppose $\varphi$ is a $C^\infty$ flow with positive speed on a $C^\infty$ closed three dimensional manifold. If $\varphi$ has positive topological entropy $h$, then there is a positive constant $C$ s.t. $\pi(T)\geq Ce^{hT}/T$ for all $T$ large enough.*

*Date*: October 29, 2016.
2010 *Mathematics Subject Classification.* 37B10, 37C10 (primary), 37C35 (secondary).
*Key words and phrases.* Markov partitions, symbolic dynamics, geodesic flows.
O.S. was partially supported by ERC award ERC-2009-StG n° 239885.

The theorem strengthens Katok's bound $\liminf_{T\to\infty}\frac{1}{T}\log\pi(T)\geq h$, see [Kat80, Kat82]. It extends to flows of lesser regularity, under the additional assumption that they possess a measure of maximal entropy (Theorem 8.1). The lower bound $Ce^{ht}/T$ is sharp in many special cases [Hub59, Mar69, PP83, Kni97], but not in the general setup of this paper. For more on this, see §8.

We obtain Theorem 1.1 by constructing a symbolic model that is a finite-to-one extension of $\varphi$. The orbits of this model are easier to understand than those of the original flow. This technique, called *"symbolic dynamics"*, can be traced back to the work of Hadamard, Morse, Artin, and Hedlund.

We proceed to describe the symbolic models used in this work. Let $\mathscr{G}$ be a directed graph with a countable set of vertices $V$. We write $v\to w$ if there is an edge from $v$ to $w$, and we assume throughout that for every $v$ there are $u,w$ s.t. $u\to v, v\to w$.

Topological Markov shifts: The *topological Markov shift* associated to $\mathscr{G}$ is the discrete-time topological dynamical system $\sigma:\Sigma\to\Sigma$ where

$$\Sigma=\Sigma(\mathscr{G}):=\{\text{paths on }\mathscr{G}\}=\{\{v_i\}_{i\in\mathbb{Z}}:v_i\to v_{i+1}\text{ for all }i\in\mathbb{Z}\},$$

equipped with the metric $d(\underline{v},\underline{w}):=\exp[-\min\{|n|:v_n\neq w_n\}]$, and $\sigma:\Sigma\to\Sigma$ is the *left shift map*, $\sigma:\{v_i\}_{i\in\mathbb{Z}}\mapsto\{v_{i+1}\}_{i\in\mathbb{Z}}$.

Birkhoff cocycle: Suppose $r:\Sigma\to\mathbb{R}$ is a function. The *Birkhoff sums* of $r$ are $r_n:=r+r\circ\sigma+\cdots+r\circ\sigma^{n-1}$ $(n\geq 1)$. There is a unique way to extend the definition to $n\leq 0$ in such a way that the *cocycle identity* $r_{m+n}=r_n+r_m\circ\sigma^n$ holds for all $m,n\in\mathbb{Z}$: $r_0:=0$ and $r_n:=-r_{|n|}\circ\sigma^{-|n|}$ $(n<0)$.

Topological Markov flow: Suppose $r:\Sigma\to\mathbb{R}^+$ is Hölder continuous and bounded away from zero and infinity. The *topological Markov flow* with *roof function* $r$ and *base map* $\sigma:\Sigma\to\Sigma$ is the flow $\sigma_r:\Sigma_r\to\Sigma_r$ where

$$\Sigma_r:=\{(\underline{v},t):\underline{v}\in\Sigma, 0\leq t<r(\underline{v})\},\ \ \sigma_r^\tau(\underline{v},t)=(\sigma^n(\underline{v}),t+\tau-r_n(\underline{v}))$$

for the unique $n\in\mathbb{Z}$ s.t. $0\leq t+\tau-r_n(\underline{v})<r(\sigma^n(\underline{v}))$.

Informally, $\sigma_r$ increases the $t$ coordinate at unit speed subject to the identifications $(\underline{v},r(\underline{v}))\sim(\sigma(\underline{v}),0)$. The cocycle identity guarantees that $\sigma_r^{\tau_1+\tau_2}=\sigma_r^{\tau_1}\circ\sigma_r^{\tau_2}$. There is a natural metric $d_r(\cdot,\cdot)$ on $\Sigma_r$, called the *Bowen-Walters metric*, s.t. $\sigma_r$ is a continuous flow [BW72]. Moreover, $\exists C>0, 0<\kappa<1$ s.t. $d_r(\sigma_r^\tau(\omega_1),\sigma_r^\tau(\omega_2))\leq Cd_r(\omega_1,\omega_2)^\kappa$ for all $|\tau|<1$ and every $\omega_1,\omega_2\in\Sigma_r$ (Lemma 5.8).

Regular parts: The *regular part* of $\Sigma$ is the set

$$\Sigma^\#:=\left\{\underline{v}\in\Sigma:\exists v,w\in V\text{ s.t. }\begin{array}{l}v_n=v\text{ for infinitely many }n>0\\ v_n=w\text{ for infinitely many }n<0\end{array}\right\},$$

and the *regular part* of $\Sigma_r$ is $\Sigma_r^\#:=\{(\underline{v},t)\in\Sigma_r:\underline{v}\in\Sigma^\#\}$.

By the Poincaré recurrence theorem, $\Sigma_r^\#$ has full measure with respect to any $\sigma_r$–invariant probability measure, and it contains all the closed orbits of $\sigma_r$. We now state our main result. Let $M$ be a three dimensional closed $C^\infty$ Riemannian manifold, let $X$ be a $C^{1+\beta}$ $(0<\beta<1)$ vector field on $M$ s.t. $X_p\neq 0$ for all $p$, let $\varphi:M\to M$ be the flow determined by $X$, and let $\mu$ be a $\varphi$–invariant Borel probability measure.

**Theorem 1.2.** *If $\mu$ is ergodic and its Kolmogorov-Sinaĭ entropy is positive, then there is a topological Markov flow $\sigma_r:\Sigma_r\to\Sigma_r$ and a map $\pi_r:\Sigma_r\to M$ s.t.:*

(1) $r : \Sigma \to \mathbb{R}^+$ *is Hölder continuous and bounded away from zero and infinity.*
(2) $\pi_r$ *is Hölder continuous with respect to the Bowen-Walters metric (see §5).*
(3) $\pi_r \circ \sigma_r^t = \varphi^t \circ \pi_r$ *for all* $t \in \mathbb{R}$.
(4) $\pi_r[\Sigma_r^\#]$ *has full measure with respect to* $\mu$.
(5) *If* $p = \pi_r(\underline{x}, t)$ *where* $x_i = v$ *for infinitely many* $i < 0$ *and* $x_i = w$ *for infinitely many* $i > 0$*, then* $\#\{(\underline{y}, s) \in \Sigma_r^\# : \pi_r(\underline{y}, s) = p\} \leq N(v, w) < \infty$.
(6) $\exists N = N(\mu) < \infty$ *s.t.* $\mu$*–a.e.* $p \in M$ *has exactly* $N$ *pre-images in* $\Sigma_r^\#$.

Some of the applications we have in mind require a version of this result for non-ergodic measures. To state it, we need to recall some facts from smooth ergodic theory [BP07]. Let $T_pM$ be the tangent space at $p$ and let $(d\varphi^t)_p : T_pM \to T_{\varphi^t(p)}M$ be the differential of $\varphi^t$ at $p$. Suppose $\mu$ is a $\varphi$–invariant Borel probability measure on $M$ (not necessarily ergodic). By the Oseledets Theorem, for $\mu$–a.e. $p \in M$, for every $0 \neq \vec{v} \in T_pM$, the limit $\chi(p, \vec{v}) := \lim_{t\to\infty} \frac{1}{t} \log \|(d\varphi^t)_p\vec{v}\|_{\varphi^t(p)}$ exists. The values of $\chi(p, \cdot)$ are called the *Lyapunov exponents* at $p$. If $\dim(M) = 3$, then there are at most three distinct such values. At least one of them, $\chi(p, X_p)$, equals zero.

HYPERBOLIC MEASURES: Suppose $\chi_0 > 0$. A $\chi_0$*–hyperbolic* measure is an invariant measure $\mu$ s.t. $\mu$–a.e. $p \in M$ has one Lyapunov exponent in $(-\infty, -\chi_0)$, one Lyapunov exponent in $(\chi_0, \infty)$ and one Lyapunov exponent equal to zero.

In dimension three, every *ergodic* invariant measure with positive metric entropy is $\chi_0$–hyperbolic for any $0 < \chi_0 < h_\mu(\varphi)$, by the Ruelle inequality [Rue78]. But some hyperbolic measures, e.g. those carried by hyperbolic closed orbits, have zero entropy.

**Theorem 1.3.** *Suppose* $\mu$ *is a* $\chi_0$*–hyperbolic invariant probability measure for some* $\chi_0 > 0$*. Then there is a topological Markov flow* $\sigma_r : \Sigma_r \to \Sigma_r$ *and a map* $\pi_r : \Sigma_r \to M$ *satisfying* (1)–(5) *in Theorem* 1.2*. If* $\mu$ *is ergodic, then* (6) *holds as well.*

Results in this spirit were first proved by Ratner and Bowen for Anosov flows and Axiom A flows in any dimension [Rat69, Rat73, Bow73], using the technique of *Markov partitions* introduced by Adler & Weiss and Sinaĭ for discrete-time dynamical systems [AW67, AW70, Sin68a, Sin68b].[1]

In 1975 Bowen gave a new construction of Markov partitions for Axiom A diffeomorphisms, using shadowing techniques [Bow75, Bow78]. The second author extended these techniques to general $C^{1+\beta}$ surface diffeomorphisms with positive topological entropy [Sar13]. Our strategy is to apply these methods to a suitable Poincaré section for the flow. The main difficulty is that [Sar13] deals with diffeomorphisms, while Poincaré sections are discontinuous.

In part 1 of the paper, we construct a Poincaré section $\Lambda$ with the following property: If $f : \Lambda \to \Lambda$ is the Poincaré return map and $\mathfrak{S} \subset \Lambda$ is the set of discontinuities of $f$, then $\liminf_{|n|\to\infty} \frac{1}{n} \log \mathrm{dist}_\Lambda(f^n(p), \mathfrak{S}) = 0$ a.e. in $\Lambda$. This places us in the context of "non-uniformly hyperbolic maps with singularities" studied in [KSLP86].

In part 2 we explain why the methods of [Sar13] apply to $f : \Lambda \to \Lambda$ despite its discontinuities. The result is a countable Markov partition for $f : \Lambda \to \Lambda$, which

[1]For geodesic flows on hyperbolic surfaces, alternative geometric and number theoretic methods are possible, see [Ser81, Ser91, KU07] and references therein. These methods are more restrictive than those of Ratner and Bowen, but they make the coding procedure more transparent.

leads to a coding of $f$ as a topological Markov shift, and a coding of $\varphi : M \to M$ as a topological Markov flow.

In part 3, we provide two applications: Theorem 1.1 on the growth of the number of closed orbits, and a result saying that the set of measures of maximal entropy is finite or countable. The proof of Theorem 1.1 uses a mixing/constant suspension dichotomy for topological Markov flows, in the spirit of [Pla72].

**Standing assumptions.** Let $M$ be a three dimensional closed $C^\infty$ Riemannian manifold, with tangent bundle $TM = \bigcup_{p\in M} T_pM$, Riemannian metric $\langle\cdot,\cdot\rangle_p$, norm $\|\cdot\|_p$, and exponential map $\exp_p$ (this is different from the $\mathrm{Exp}_p$ in §3).

Given $Y \subset M$, $\mathrm{dist}_Y(y_1, y_2) := \inf\{$lengths of rectifiable curves in $Y$ from $y_1$ to $y_2\}$, where $\inf\varnothing := \infty$. Given two metric spaces $(A, d_A), (B, d_B)$ and a map $F : A \to B$, $\mathrm{Höl}_\alpha(F) := \sup_{x\neq y} \frac{d_B(F(x),F(y))}{d_A(x,y)^\alpha}$ for $0 < \alpha \leq 1$, and $\mathrm{Lip}(F) := \mathrm{Höl}_1(F)$.

We let $X : M \to TM$ be a $C^{1+\beta}$ vector field on $M$ $(0 < \beta < 1)$, and $\varphi : M \to M$ be the flow generated by $X$. This means that $\varphi$ is a one-parameter family of maps $\varphi^t : M \to M$ s.t. $\varphi^{t+s} = \varphi^t \circ \varphi^s$ for all $t, s \in \mathbb{R}$, and s.t. $X_p(f) = \frac{d}{dt}\big|_{t=0} f[\varphi^t(p)]$ for all $f \in C^\infty(M)$. In this case $(t, p) \mapsto \varphi^t(p)$ is a $C^{1+\beta}$ map $[-1, 1] \times M \to M$ [EM70, page 112]. We assume throughout that $X_p \neq 0$ for all $p$.

## Part 1. The Poincaré section

### 2. Poincaré sections

**Basic definitions.** Suppose $\varphi : M \to M$ is a flow.

Poincaré section: $\Lambda \subset M$ Borel set s.t. for every $p \in M$, $\{t > 0 : \varphi^t(p) \in \Lambda\}$ is a sequence tending to $+\infty$, and $\{t < 0 : \varphi^t(p) \in \Lambda\}$ is a sequence tending to $-\infty$.

Roof function: $R_\Lambda : \Lambda \to (0, \infty)$, $R_\Lambda(p) := \min\{t > 0 : \varphi^t(p) \in \Lambda\}$.

Poincaré map: $f_\Lambda : \Lambda \to \Lambda$, $f_\Lambda(p) := \varphi^{R_\Lambda(p)}(p)$.

Induced measure: Every $\varphi$–invariant probability measure $\mu$ on $M$ induces an $f_\Lambda$–invariant measure $\mu_\Lambda$ on $\Lambda$ defined by the equality

$$\int_M g d\mu = \frac{1}{\int_\Lambda R_\Lambda d\mu_\Lambda} \int_\Lambda \left( \int_0^{R_\Lambda(p)} g[\varphi^t(p)]dt \right) d\mu_\Lambda(p), \text{ for all } g \in L^1(\mu).$$

Uniform Poincaré section: The Poincaré section $\Lambda$ is called *uniform* if its roof function is bounded away from zero and infinity. If $\Lambda$ is uniform, then $\mu_\Lambda$ is finite and it can be normalized. With this normalization, for every Borel subset $E \subset \Lambda$ and $0 < \varepsilon < \inf R_\Lambda$ it holds $\mu_\Lambda(E) = \mu[\bigcup_{0<t<\varepsilon} \varphi^t(E)]/\mu[\bigcup_{0<t<\varepsilon} \varphi^t(\Lambda)]$.

All the Poincaré sections considered in this paper will be uniform, and each of them will be the disjoint union of finitely many embedded smooth two dimensional discs. Let $\partial\Lambda$ denote the union of the boundaries of these discs. The set $\partial\Lambda$ will introduce discontinuities to the Poincaré map of $\Lambda$.

Singular set: The *singular set* of a Poincaré section $\Lambda$ is

$$\mathfrak{S}(\Lambda) := \left\{ p \in \Lambda : \begin{array}{l} p \text{ does } \textit{not} \text{ have a relative neighborhood } V \subset \Lambda \setminus \partial\Lambda \text{ s.t.} \\ V \text{ is diffeomorphic to an open disc, and } f_\Lambda : V \to f_\Lambda(V) \\ \text{and } f_\Lambda^{-1} : V \to f_\Lambda^{-1}(V) \text{ are diffeomorphisms} \end{array} \right\}.$$

Regular set: $\Lambda' := \Lambda \setminus \mathfrak{S}(\Lambda)$.

**Basic constructions.** Let $\varphi$ be a flow satisfying our standing assumptions.

CANONICAL TRANSVERSE DISC: $S_r(p) := \{\exp_p(\vec{v}) : \vec{v} \in T_pM, \vec{v} \perp X_p, \|\vec{v}\|_p \le r\}$.

CANONICAL FLOW BOX: $\mathrm{FB}_r(p) := \{\varphi^t(q) : q \in S_r(p), |t| \le r\}$.

The following lemmas are standard, see the appendix for proofs.

**Lemma 2.1.** *There is a constant $\mathfrak{r}_s > 0$ which only depends on $M$ and $\varphi$ s.t. for every $p \in M$ and $0 < r < \mathfrak{r}_s$, $S := S_r(p)$ is a $C^\infty$ embedded closed disc, $|\measuredangle(X_q, T_qS)| \ge \frac{1}{2}$ radians for all $q \in S$, and $\mathrm{dist}_M(\cdot,\cdot) \le \mathrm{dist}_S(\cdot,\cdot) \le 2\,\mathrm{dist}_M(\cdot,\cdot)$.*

**Lemma 2.2.** *There are constants $\mathfrak{r}_f, \mathfrak{d} \in (0,1)$ which only depend on $M$ and $\varphi$ s.t. for every $p \in M$, $\mathrm{FB}_{\mathfrak{r}_f}(p)$ contains an open ball with center $p$ and radius $\mathfrak{d}$, and $(q,t) \mapsto \varphi^t(q)$ is a diffeomorphism from $S_{\mathfrak{r}_f}(p) \times [-\mathfrak{r}_f, \mathfrak{r}_f]$ onto $\mathrm{FB}_{\mathfrak{r}_f}(p)$.*

**Lemma 2.3.** *There are constants $\mathfrak{L}, \mathfrak{H} > 1$ which only depend on $M$ and $\varphi$ s.t. $\mathfrak{t}_p : \mathrm{FB}_{\mathfrak{r}_f}(p) \to [-\mathfrak{r}_f, \mathfrak{r}_f]$ and $\mathfrak{q}_p : \mathrm{FB}_{\mathfrak{r}_f}(p) \to S_{\mathfrak{r}_f}(p)$ defined by $z = \varphi^{\mathfrak{t}_p(z)}[\mathfrak{q}_p(z)]$ are well-defined maps with $\mathrm{Lip}(\mathfrak{t}_p), \mathrm{Lip}(\mathfrak{q}_p) \le \mathfrak{L}$ and $\|\mathfrak{t}_p\|_{C^{1+\beta}}, \|\mathfrak{q}_p\|_{C^{1+\beta}} \le \mathfrak{H}$.*

We call $\mathfrak{t}_p, \mathfrak{q}_p$ the *flow box coordinates.* Set $\mathfrak{r} := 10^{-1}\min\{1, \mathfrak{r}_s, \mathfrak{r}_f, \mathfrak{d}\}/(1+\max\|X_p\|)$.

STANDARD POINCARÉ SECTION: A Poincaré section $\Lambda$ is *standard* if it has the form

$$\Lambda = \Lambda(p_1, \ldots, p_N; r) := \biguplus_{i=1}^{N} S_r(p_i)$$

where $r < \mathfrak{r}$, $\sup R_\Lambda < \mathfrak{r}$, and $S_r(p_i)$ are pairwise disjoint. The points $p_1, \ldots, p_N$ are called the *centers* of $\Lambda$, and $r$ is called the *radius* of $\Lambda$. (Here and throughout, $\biguplus$ means the union of pairwise disjoint sets.)

Standard Poincaré sections are special cases of the "proper families" Bowen used in [Bow73, §2] to build Markov partitions for Axiom A flows. Their existence is discussed below (Lemma 2.7). For the moment, let us assume Standard Poincaré sections exist, and discuss some of their properties.

Fix a standard Poincaré section $\Lambda = \Lambda(p_1, \ldots, p_N; r)$ and write $f = f_\Lambda$, $R = R_\Lambda$, $\mathfrak{S} := \mathfrak{S}(\Lambda)$, and $\Lambda' := \Lambda \setminus \mathfrak{S}$.

**Lemma 2.4.** *Every standard Poincaré section is a uniform Poincaré section.*

*Proof.* We have $\sup R < \infty$ by the definition of standard sections, so it remains to see that $\inf R > 0$. Let $x \in S_r(p_i)$, $f(x) \in S_r(p_j)$. If $i = j$ then $R(x) > \mathfrak{r}_f$, otherwise there would exist $0 < t \le \mathfrak{r}_f$ s.t. $\varphi^0(f(x)) = f(x) = \varphi^t(x)$, which contradicts the last part of Lemma 2.2. If $i \ne j$ then $\{\varphi^t(x)\}_{0 \le t \le R(x)}$ is a curve from $S_r(p_i)$ to $S_r(p_j)$, thus $R(x) \ge \mathrm{dist}_M(S_r(p_i), S_r(p_j))/\max\|X_p\|$. Hence $\inf R > 0$. $\square$

**Lemma 2.5.** *$R$, $f$ and $f^{-1}$ are differentiable on $\Lambda'$, and $\exists \mathfrak{C} > 0$ only depending on $M$ and $\varphi$ s.t. $\sup_{x \in \Lambda'} \|dR_x\| < \mathfrak{C}$, $\sup_{x \in \Lambda'} \|df_x\| < \mathfrak{C}$, $\sup_{x \in \Lambda'} \|(df_x)^{-1}\| < \mathfrak{C}$, $\|f \restriction_U \|_{C^{1+\beta}} < \mathfrak{C}$ and $\|f^{-1} \restriction_U \|_{C^{1+\beta}} < \mathfrak{C}$ for all open and connected $U \subset \Lambda'$.*

*Proof.* Suppose $x \in \Lambda'$, then $\exists i, j, k$ s.t. $f^{-1}(x) \in S_r(p_i)$, $x \in S_r(p_j)$, and $f(x) \in S_r(p_k)$. Since $f$ is continuous at $x$ and the canonical discs composing $\Lambda$ are closed and disjoint, $x$ has an open neighborhood $V$ in $S_r(p_j)$ s.t. for all $y \in V$ it holds $f(y) \in S_r(p_k)$ and $f^{-1}(y) \in S_r(p_i)$. Since $\sup R < \mathfrak{r} < 10^{-1}\mathfrak{d}/\max\|X_p\|$, if $y \in V$ then $\mathrm{dist}_M(y, p_k) \le \mathrm{dist}_M(y, f(y)) + \mathrm{dist}_M(f(y), p_k) \le \max\|X_p\| \sup R + \mathfrak{r} < \mathfrak{d}$. Similarly, $\mathrm{dist}_M(y, p_i) < \mathfrak{d}$. Thus $V \subset B_{\mathfrak{d}}(p_i) \cap B_{\mathfrak{d}}(p_k) \subset \mathrm{FB}_{\mathfrak{r}_f}(p_i) \cap \mathrm{FB}_{\mathfrak{r}_f}(p_k)$, whence $R \restriction_V = -\mathfrak{t}_{p_k}$, $f \restriction_V = \mathfrak{q}_{p_k}$, $f^{-1} \restriction_V = \mathfrak{q}_{p_i}$. Now use Lemma 2.3. $\square$

Let $\mu$ be a $\varphi$–invariant probability measure, and let $\mu_\Lambda$ be the induced measure on $\Lambda$. If $\mu_\Lambda(\mathfrak{S})=0$, then $\mu_\Lambda[\bigcup_{n\in\mathbb{Z}} f^n(\mathfrak{S})]=0$, and the derivative cocycle $df^n_x : T_x\Lambda\to T_{f^n(x)}\Lambda$ is well-defined $\mu_\Lambda$–a.e. By Lemma 2.5, $\log\|df_x\|, \log\|df^{-1}_x\|$ are integrable (even bounded), so the Oseledets Multiplicative Ergodic Theorem applies, and $f$ has well-defined Lyapunov exponents $\mu_\Lambda$–a.e. Fix $\chi>0$.

**Lemma 2.6.** *Suppose $\mu_\Lambda(\mathfrak{S})=0$. If $\mu$ is $\chi$–hyperbolic then $f$ has one Lyapunov exponent in $(-|\ln\mathfrak{C}|,-\chi\inf R)$ and another in $(\chi\inf R,|\ln\mathfrak{C}|)$ for $\mu_\Lambda$–a.e. $x\in\Lambda$.*

*Proof.* Let $\Omega_\chi$ denote the set of points where the flow has one zero Lyapunov exponent, one Lyapunov exponent in $(-\infty,-\chi)$ and another in $(\chi,\infty)$. By assumption $\mu[\Omega_\chi^c]=0$, thus $\Lambda_\chi:=\{x\in\Lambda\setminus\bigcup_{n\in\mathbb{Z}}f^{-n}(\mathfrak{S}) : \exists t>0 \text{ s.t. } \varphi^t(x)\in\Omega_\chi\}$ has full measure with respect to $\mu_\Lambda$.

Let $\Lambda_\chi^*:=\{x\in\Lambda_\chi : \chi(x,\vec v):=\lim\limits_{n\to\pm\infty}\frac{1}{n}\log\|df^n_x\vec v\| \text{ exists for all } 0\neq\vec v\in T_x\Lambda\}$. By the Oseledets theorem, $\Lambda_\chi^*$ has full $\mu_\Lambda$–measure. By Lemma 2.5, $|\chi(x,\vec v)|\leq|\ln\mathfrak{C}|$. The Lyapunov exponents of $\varphi$ are constant along flow lines, therefore for every $x\in\Lambda^*_\chi$ there are vectors $\underline{e}^s_x,\underline{e}^u_x\in T_xM$ s.t. $\lim\limits_{t\to\infty}\frac{1}{t}\log\|d\varphi^t_x\underline{e}^s_x\|_{\varphi^t(x)}<-\chi$ and $\lim\limits_{t\to\infty}\frac{1}{t}\log\|d\varphi^t_x\underline{e}^u_x\|_{\varphi^t(x)}>\chi$. Let $\vec n(x):=\frac{X_x}{\|X_x\|}$. Since $\lim\limits_{t\to\infty}\frac{1}{t}\log\|d\varphi^t_x\vec n(x)\|_{\varphi^t(x)}=0$, $\{\underline{e}^s_x,\underline{e}^u_x,\vec n(x)\}$ span $T_xM$. Note that $\underline{e}^s_x,\underline{e}^u_x$ are *not* necessarily in $T_x\Lambda$.

Pick two independent vectors $\vec v_1,\vec v_2\in T_x\Lambda$ and write $\vec v_i=\alpha_i\underline{e}^s_x+\beta_i\underline{e}^u_x+\gamma_i\vec n(x)$, $i=1,2$. The vectors $\binom{\alpha_1}{\beta_1},\binom{\alpha_2}{\beta_2}$ must be linearly independent, otherwise some non-trivial linear combination of $\vec v_1,\vec v_2$ equals $\vec n(x)$, which is impossible since $\mathrm{span}\{\vec v_1,\vec v_2\}=T_x\Lambda$ and $\Lambda$ is tranverse to the flow. It follows that $T_x\Lambda$ contains two vectors of the form

$$\vec v^s_x=\underline{e}^s_x+\gamma_s\vec n(x)\,,\ \vec v^u_x=\underline{e}^u_x+\gamma_u\vec n(x).$$

These vectors are the projections of $\underline{e}^s_x,\underline{e}^u_x$ to $T_x\Lambda$ along $\vec n(x)$. We will estimate their Lyapunov exponents.

Write $\Lambda=\Lambda(p_1,\ldots,p_N;r)$. As in the proof of Lemma 2.5, for every $x\in\Lambda\setminus\mathfrak{S}$, if $f(x)\in S_r(p_i)$, then $x$ has a neighborhood $V$ in $\Lambda$ s.t. $V\subset \mathrm{FB}_{\mathfrak{r}_f}(p_i)$, $R\restriction_V=-\mathfrak{t}_{p_i}$, and $f\restriction_V=\mathfrak{q}_{p_i}$. More generally, suppose $f^n(x)\in S_r(q_n)$ for $q_n\in\{p_1,\ldots,p_N\}$. If $x\notin\bigcup_{k\in\mathbb{Z}}f^k(\mathfrak{S})$ then there are open neighborhoods $V_n$ of $x$ in $\Lambda$ s.t.

$$f^{n-1}(V_n)\in\mathrm{FB}_{\mathfrak{r}_f}(q_n),\ \text{and}\ f^n\restriction_{V_n}=(\mathfrak{q}_{q_n}\circ\cdots\circ\mathfrak{q}_{q_1})\restriction_{V_n}. \tag{2.1}$$

By the definition of the flow box coordinates, $\mathfrak{q}_{q_i}(\cdot)=\varphi^{-\mathfrak{t}_{q_i}(\cdot)}(\cdot)$ for every $i$. Since $x\notin\bigcup_{k\in\mathbb{Z}}f^k(\mathfrak{S})$, $\mathfrak{t}_{q_i}$ is continuous on a neighborhood of $f^{i-1}(x)$, hence the smaller $V_n$ the closer $-\mathfrak{t}_{q_i}\restriction_{f^{i-1}(V_n)}$ is to $R(f^{i-1}(x))$. If $V_n$ is small enough and

$$R_n:=R(x)+R(f(x))+\cdots+R(f^{n-1}(x)),$$

then $\varphi^{R_n}(y)\in\mathrm{FB}_{\mathfrak{r}_f}(q_n)$ for all $y\in V_n$, and we can decompose

$$(\mathfrak{q}_{q_n}\circ\cdots\circ\mathfrak{q}_{q_1})(y)=(\mathfrak{q}_{q_n}\circ\varphi^{R_n})(y),\ \text{for}\ y\in V_n. \tag{2.2}$$

We emphasize that the power $R_n$ is the same for all $y\in V_n$.

We use (2.1)–(2.2) to calculate $df^n_x\vec v^s_x$. First note that $(d\mathfrak{q}_{q_1})_x\vec n(x)=\vec 0$: let $\gamma(t)=\varphi^t(x)$, then $\mathfrak{q}_{q_1}[\gamma(t)]=\mathfrak{q}_{q_1}(x)$ for all $|t|$ small, so $\frac{d}{dt}\big|_{t=0}\mathfrak{q}_{q_1}[\gamma(t)]=\vec 0$. By (2.1), $df^n_x\vec v^s_x=d(\mathfrak{q}_{q_n}\circ\cdots\circ\mathfrak{q}_{q_1})\underline{e}^s_x$. By (2.2), $\|df^n_x\vec v^s_x\|\leq\max_i\|d\mathfrak{q}_{q_i}\|\cdot\|d\varphi^{R_n}_x\underline{e}^s_x\|_{\varphi^{R_n}(x)}$, whence $\limsup_{n\to\infty}\frac{1}{n}\log\|df^n_x\vec v^s_x\|_{f^n(x)}\leq-\chi\limsup_{n\to\infty}\frac{R_n}{n}\leq-\chi\inf R$.

Applying this argument to the reverse flow $\psi^t:=\varphi^{-t}$, we find that the other Lyapunov exponent belongs to $(\chi\inf R,\infty)$. $\square$

*Remark.* If $\mu$ is ergodic then $\limsup_{n\to\infty}\frac{R_n}{n}=\int Rd\mu_\Lambda=1$, and we get the stronger estimate that the Lyapunov exponents of $f$ are outside $(-\chi,\chi)$ almost surely.

**Adapted Poincaré sections.** Let $\Lambda$ be a standard Poincaré section, and let $\mathrm{dist}_\Lambda$ denote its intrinsic Riemannian distance (with the convention that the distance between different connected components of $\Lambda$ is infinite). Let $\mu$ be a $\varphi$–invariant probability measure, and let $\mu_\Lambda$ be the induced probability measure on $\Lambda$. Recall that $f_\Lambda:\Lambda\to\Lambda$ may have singularities. The following definition is motivated by the treatment of Pesin theory for maps with singularities in [KSLP86].

ADAPTED POINCARÉ SECTION: A standard Poincaré section $\Lambda$ is *adapted to* $\mu$ if:
(1) $\mu_\Lambda(\mathfrak{S})=0$, where $\mathfrak{S}=\mathfrak{S}(\Lambda)$ is the singular set of $\Lambda$.
(2) $\lim\limits_{n\to\infty}\frac{1}{n}\log\mathrm{dist}_\Lambda(f_\Lambda^n(p),\mathfrak{S})=0$ for $\mu_\Lambda$–a.e. $p\in\Lambda$.
(3) $\lim\limits_{n\to\infty}\frac{1}{n}\log\mathrm{dist}_\Lambda(f_\Lambda^{-n}(p),\mathfrak{S})=0$ for $\mu_\Lambda$–a.e. $p\in\Lambda$.

Notice that (2) implies (1), by the Poincaré recurrence theorem.

We wish to show that any $\varphi$–invariant Borel probability measure has adapted Poincaré sections. The idea is to construct a one-parameter family of standard Poincaré sections $\Lambda_r$, and show that $\Lambda_r$ is adapted to $\mu$ for a.e. $r$. The family is constructed in the next lemma.

**Lemma 2.7.** *For every $h_0>0,K_0>1$ there are $p_1,\ldots,p_N\in M$, $0<\rho_0<h_0/K_0$ s.t. for every $r\in[\rho_0,K_0\rho_0]$ the set $\Lambda(p_1,\ldots,p_N;r)$ is a standard Poincaré section with roof function and radius bounded above by $h_0$.*

The existence of standard Poincaré sections is treated in [Bow73, §2] as a self-evident fact, but we do not think it is completely obvious. We provide a detailed proof of Lemma 2.7 in the appendix. The next result shows the existence of adapted sections.

**Theorem 2.8.** *Every $\varphi$–invariant probability measure $\mu$ has adapted Poincaré sections with arbitrarily small roof functions.*

*Proof.* We use parameter selection, as in [LS82]. Let $\Lambda_r:=\Lambda(p_1,\ldots,p_N;r)$, $a\le r\le b$, be a one-parameter family of standard Poincaré sections as in Lemma 2.7. We will show that $\Lambda_r$ is adapted to $\mu$ for Lebesgue a.e. $r\in[a,b]$.

Without loss of generality $a,b,r,\sup R_{\Lambda_r}\le h_0<\mathfrak{r}=\frac{1}{10}\big[\frac{\min\{1,\mathfrak{r}_s,\mathfrak{r}_f,\mathfrak{d}\}}{S_0}\big]$ for all $r\in[a,b]$, where $\mathfrak{r}_s,\mathfrak{r}_f,\mathfrak{d}$ are given by Lemmas 2.1–2.3, and $S_0:=1+\max\|X_p\|$. We define the *boundary* of a canonical transverse disc $S_r(p)$ by the formula $\partial S_r(p):=\{\exp_p(\vec{v}):\vec{v}\in T_pM,\vec{v}\perp X_p,\|\vec{v}\|_p=r\}$. Let

$$\mathfrak{S}_r:=\bigcup\big\{\mathfrak{q}_{p_i}\big[\partial S_r(p_j)\big]:1\le i,j\le N\,,\mathrm{dist}_M(S_r(p_i),S_r(p_j))\le h_0S_0\big\},$$

where $\mathfrak{q}_{p_i}:\mathrm{FB}_{\mathfrak{r}_f}(p_i)\to S_{\mathfrak{r}_f}(p_i)$ is given by Lemma 2.2. The assumption that $\mathrm{dist}_M(S_r(p_i),S_r(p_j))\le h_0S_0$ ensures the inclusion $\partial S_r(p_j)\subset\mathrm{FB}_{\mathfrak{r}_f}(p_i)$, since for all $q\in\partial S_r(p_j)$, $\mathrm{dist}_M(q,p_i)\le\mathrm{diam}[S_r(p_j)]+\mathrm{dist}_M(S_r(p_j),S_r(p_i))+\mathrm{diam}[S_r(p_i)]<h_0S_0+4r<5\mathfrak{r}S_0<\mathfrak{d}$, whence $q\in B_{\mathfrak{d}}(p_i)\subset\mathrm{FB}_{\mathfrak{r}_f}(p_i)$.

CLAIM. *$\mathfrak{S}_r$ contains the singular set of $\Lambda_r$.*

*Proof.* Fix $r$ and write $R=R_{\Lambda_r}$, $f=f_{\Lambda_r}$. We show that if $p\in\Lambda_r\setminus\mathfrak{S}_r$ then $f,f^{-1}$ are local diffeomorphisms on a neighborhood of $p$. Let $i,j$ be the unique indices s.t. $p\in S_r(p_i)$ and $f(p)\in S_r(p_j)$. The speed of the flow is less than

$S_0$, so $\mathrm{dist}_M(p,p_j) \leq \mathrm{dist}_M(p,f(p)) + \mathrm{dist}_M(f(p),p_j) < h_0 S_0 + r < \mathfrak{d}$. Thus $p \in \mathrm{FB}_{\mathfrak{r}_f}(p_j)$. Similarly, $\mathrm{dist}_M(f(p),p_i) < \mathfrak{d}$, so $f(p) \in \mathrm{FB}_{\mathfrak{r}_f}(p_i)$. It follows that $R(p) = -\mathfrak{t}_{p_j}(p) = |\mathfrak{t}_{p_j}(p)|$ and $f(p) = \mathfrak{q}_{p_j}(p)$. Similarly, $p = \mathfrak{q}_{p_i}[f(p)]$. Since $\mathrm{dist}_M(S_r(p_i), S_r(p_j)) \leq \mathrm{dist}_M(p, \varphi^{R(p)}(p)) < h_0 S_0$ and $p \notin \mathfrak{S}_r$,

$$p \notin \partial S_r(p_i) \text{ and } f(p) \notin \partial S_r(p_j).$$

So $\exists V \subset \Lambda_r \setminus \partial\Lambda_r$ relatively open s.t. $V \ni p$ and $\mathfrak{q}_{p_j}(V) \subset \Lambda_r \setminus \partial\Lambda_r$. The map $\mathfrak{q}_{p_j} : V \to \mathfrak{q}_{p_j}(V)$ is a diffeomorphism, because $\mathfrak{q}_{p_j}$ is differentiable and $\mathfrak{q}_{p_i} \circ \mathfrak{q}_{p_j} = \mathrm{Id}$ on $V$. We will show that $f \restriction_W = \mathfrak{q}_{p_j} \restriction_W$ on some open $W \subset \Lambda_r \setminus \partial\Lambda_r$ containing $p$.

Since $R(p) = |\mathfrak{t}_{p_j}(p)|$, the curve $\{\varphi^t(p) : 0 < t < |\mathfrak{t}_{p_j}(p)|\}$ does not intersect $\Lambda_r$. The set $\Lambda_r$ is compact and $\varphi, \mathfrak{t}_{p_j}$ are continuous, so $p$ has a relatively open neighborhood $W \subset V$ s.t. $\{\varphi^t(q) : 0 < t < |\mathfrak{t}_{p_j}(q)|\}$ does not intersect $\Lambda_r$ for all $q \in W$. So $f \restriction_W = \mathfrak{q}_{p_j} \restriction_W$, and we see that $f$ is a local diffeomorphism at $p$. Similarly, $f^{-1}$ is a local diffeomorphism at $p$, which proves that $p \notin \mathfrak{S}(\Lambda_r)$, and hence the claim.

We now proceed to the proof of the theorem. We begin with some reductions. Let $f_r := f_{\Lambda_r}$. By the claim it is enough to show that

$$\mu_{\Lambda_r}\left\{p \in \Lambda_r : \liminf_{|n|\to\infty} \frac{1}{|n|} \log \mathrm{dist}_{\Lambda_r}(f_r^n(p), \mathfrak{S}_r) < 0\right\} = 0 \text{ for a.e. } r \in [a,b]. \quad (2.3)$$

Indeed, this implies $\exists r$ s.t. $\liminf\limits_{|n|\to\infty} \frac{1}{|n|} \log \mathrm{dist}_{\Lambda_r}(f_r^n(p), \mathfrak{S}(\Lambda_r)) \geq 0$ for $\mu_{\Lambda_r}$–a.e. $p \in \Lambda_r$, and the limit is non-positive, because $\mathrm{dist}_{\Lambda_r}(q, \mathfrak{S}(\Lambda_r)) \leq \mathrm{dist}_{\Lambda_r}(q, \partial\Lambda) \leq r$ for all $q \in \Lambda$. Let

$$A_\alpha(r) := \{p \in \Lambda_b : \exists \text{ infinitely many } n \in \mathbb{Z} \text{ s.t. } \tfrac{1}{|n|} \log \mathrm{dist}_{\Lambda_b}(f_b^n(p), \mathfrak{S}_r) < -\alpha\}.$$

We have $\Lambda_r \subset \Lambda_b$, so $\mu_{\Lambda_r} \ll \mu_{\Lambda_b}$, $\mathrm{dist}_{\Lambda_r} \geq \mathrm{dist}_{\Lambda_b}$, and $f_r(x) = f_b^{n(x)}(x)$ with $1 \leq n(x) \leq \frac{\sup R_r}{\inf R_b}$. Therefore (2.3) follows from the statement

$$\forall\ \alpha > 0 \text{ rational } \big(\mu_{\Lambda_b}[A_\alpha(r)] = 0 \text{ for a.e. } r \in [a,b]\big). \quad (2.4)$$

Let $I_\alpha(p) := \{a \leq r \leq b : p \in A_\alpha(r)\}$, then $1_{A_\alpha(r)}(p) = 1_{I_\alpha(p)}(r)$, whence by Fubini's Theorem $\int_a^b \mu_{\Lambda_b}[A_\alpha(r)]dr = \int_{\Lambda_b} \mathrm{Leb}[I_\alpha(p)]d\mu_{\Lambda_b}(p)$. So (2.4) follows from

$$\mathrm{Leb}[I_\alpha(p)] = 0 \text{ for all } p \in \Lambda_b. \quad (2.5)$$

In summary, (2.5) $\Rightarrow$ (2.4) $\Rightarrow$ (2.3) $\Rightarrow$ the theorem.

Proof of (2.5): Fix $p \in \Lambda_b$. If $r \in I_\alpha(p)$ then $\mathrm{dist}_{\Lambda_b}(f_b^n(p), \mathfrak{S}_r) < e^{-\alpha|n|}$ for infinitely many $n \in \mathbb{Z}$. The section $\Lambda_b$ is a finite union of canonical transverse discs $S_b(p_i)$, and $S_b(p_i) \cap \mathfrak{S}_r$ is a finite union of projections $S_b(p_i) \cap \mathfrak{q}_{p_i}[\partial S_r(p_j)]$, each satisfying $\mathrm{dist}_M(S_r(p_i), S_r(p_j)) \leq h_0 S_0$. It follows that there are infinitely many $n \in \mathbb{Z}$ such that for some *fixed* $1 \leq i,j \leq N$ s.t. $\mathrm{dist}_M(S_r(p_i), S_r(p_j)) \leq h_0 S_0$, it holds that

$$f_b^n(p) \in S_b(p_i) \text{ and } \mathrm{dist}_{\Lambda_b}(f_b^n(p), \mathfrak{q}_{p_i}[\partial S_r(p_j)]) < e^{-\alpha|n|}. \quad (2.6)$$

Since $\mathrm{dist}_M(S_r(p_i), S_r(p_j)) \leq h_0 S_0$, $\mathrm{dist}_M(p_i, p_j) \leq h_0 S_0 + 2r < \mathfrak{d}$. This, and our assumptions on $b$ and $h_0$, guarantee that $S_b(p_j), f_b^n(p), \mathfrak{q}_{p_i}(\partial S_r(p_j))$ are inside $\mathrm{FB}_{\mathfrak{r}_f}(p_i) \cap \mathrm{FB}_{\mathfrak{r}_f}(p_j)$, which are in the domains of definition of $\mathfrak{q}_{p_i}$ and $\mathfrak{q}_{p_j}$.

Suppose $n$ satisfies (2.6). Let $q$ be the point that minimizes $\mathrm{dist}_{\Lambda_b}(f_b^n(p), \mathfrak{q}_{p_i}(q))$ over all $q \in \partial S_r(p_j)$, then

$$\begin{aligned}
&e^{-\alpha|n|} > \mathrm{dist}_{\Lambda_b}(f_b^n(p), \mathfrak{q}_{p_i}(q)) \geq \mathrm{dist}_M(f_b^n(p), \mathfrak{q}_{p_i}(q)) \\
&\geq \mathfrak{L}^{-1}\,\mathrm{dist}_M(\mathfrak{q}_{p_j}[f_b^n(p)], \mathfrak{q}_{p_j}[\mathfrak{q}_{p_i}(q)]) && \because \mathrm{Lip}(\mathfrak{q}_j) \leq \mathfrak{L} \\
&= \mathfrak{L}^{-1}\,\mathrm{dist}_M(\mathfrak{q}_{p_j}[f_b^n(p)], q) && \because q \in S_r(p_j) \\
&\geq (2\mathfrak{L})^{-1}\,\mathrm{dist}_{\Lambda_b}(\mathfrak{q}_{p_j}[f_b^n(p)], q). && \because \mathrm{dist}_{S_b(p_j)} \leq 2\,\mathrm{dist}_M
\end{aligned}$$

Thus $\mathrm{dist}_{\Lambda_b}(\mathfrak{q}_{p_j}[f_b^n(p)], q) \leq 2\mathfrak{L}e^{-\alpha|n|}$, which implies that

$$|\,\mathrm{dist}_{\Lambda_b}(p_j, \mathfrak{q}_{p_j}[f_b^n(p)]) - \mathrm{dist}_{\Lambda_b}(p_j, q)| \leq 2\mathfrak{L}e^{-\alpha|n|}.$$

We now use the special geometry of canonical transverse discs: $q \in \partial S_r(p_j)$, so $\mathrm{dist}_{\Lambda_b}(p_j, q) = r$. Writing $D_{jn}(p) := \mathrm{dist}_{\Lambda_b}(p_j, \mathfrak{q}_{p_j}[f_b^n(p)])$, we see that for every $n$ which satisfies (2.6), $|D_{jn}(p) - r| \leq 2\mathfrak{L}e^{-\alpha|n|}$. Thus every $r \in I_\alpha(p)$ belongs to

$$\bigcup_{j=1}^{N} \{r \in [a,b] : \exists \text{ infinitely many } n \in \mathbb{Z} \text{ s.t. } |r - D_{jn}(p)| \leq 2\mathfrak{L}e^{-\alpha|n|}\}.$$

By the Borel-Cantelli Lemma, this set has zero Lebesgue measure. $\square$

Two standard Poincaré sections with the same set of centers are called *concentric.* Since $b/a$ can be chosen arbitrarily large, the last proof shows the following.

**Corollary 2.9.** *Let $\mu$ be a $\varphi$–invariant probability measure. For every $h_0 > 0$ there are two concentric standard Poincaré sections $\Lambda_i = \Lambda(p_1, \ldots, p_N; r_i)$ with height functions bounded above by $h_0$, s.t. $\Lambda_1$ is adapted to $\mu$ and $r_2 > 2r_1$.*

To see this take $r_1$ close to $a$ s.t. $\Lambda_{r_1}$ is adapted, and $r_2 = b$.

*Remark.* The adapted Poincaré section given in Theorem 2.8 and Corollary 2.9 depends on the measure $\mu$. It would be interesting to construct, for a given $\chi > 0$, a Poincaré section which is adapted to all ergodic hyperbolic measures with one Lyapunov exponent bigger than $\chi$ and one Lyapunov exponent smaller than $-\chi$.

## 3. Pesin charts for adapted Poincaré sections

One of the central tools in Pesin theory is a system of local coordinates which present a non-uniformly hyperbolic map as a perturbation of a uniformly hyperbolic linear map [Pes76, KH95, BP07]. We will construct such coordinates for the Poincaré map of an adapted Poincaré section. Adaptability is used, as in [KSLP86], to control the size of the coordinate patches along typical orbits (Lemma 3.3).

Suppose $\mu$ is a $\varphi$–invariant probability measure on $M$, and assume that $\mu$ is $\chi_0$–hyperbolic for some $\chi_0 > 0$. We do not assume ergodicity. Fix once and for all a standard Poincaré section $\Lambda = \Lambda(p_1, \ldots, p_N; r)$ for $\varphi$, which is adapted to $\mu$. Set $f := f_\Lambda, R := R_\Lambda, \mathfrak{S} := \mathfrak{S}(\Lambda)$, and let $\mu_\Lambda$ be the induced measure on $\Lambda$.

Without loss of generality, there is a larger concentric standard Poincaré section $\widetilde{\Lambda} := \Lambda(p_1, \ldots, p_N; \widetilde{r})$ s.t. $\widetilde{r} > 2r$. Thus $\widetilde{\Lambda} \supset \Lambda$, and $\mathrm{dist}_{\widetilde{\Lambda}}(\Lambda, \partial\widetilde{\Lambda}) > r$. We will use $\widetilde{\Lambda}$ as a safety margin in the following definition of the *exponential map* of $\Lambda$:

$$\mathrm{Exp}_x : \{\vec{v} \in T_x\Lambda : \|\vec{v}\|_x < r\} \to \widetilde{\Lambda},\ \mathrm{Exp}_x(\vec{v}) := \gamma_x(\|\vec{v}\|_x),$$

where $\gamma_x(\cdot)$ is the geodesic in $\widetilde{\Lambda}$ s.t. $\gamma(0) = x$ and $\dot{\gamma}(0) = \vec{v}$. This makes sense even near $\partial\Lambda$, because every geodesic of $\Lambda$ can be prolonged $r$ units of distance into $\widetilde{\Lambda}$ without falling off the edge. Notice that geodesics of $\widetilde{\Lambda}$ are usually not geodesics of $M$, therefore $\mathrm{Exp}_x$ is usually different from $\exp_x$. As in [Spi79, chapter 9], there are $\rho_{\mathrm{dom}}, \rho_{\mathrm{im}} \in (0, r)$ s.t. for every $x \in \Lambda$, $\mathrm{Exp}_x$ is a 2–bi-Lipschitz diffeomorphism from $\{\vec{v} \in T_x\Lambda : \|\vec{v}\|_x < \sqrt{2}\rho_{\mathrm{dom}}\}$ onto a relative neighborhood of $\{y \in \widetilde{\Lambda} : \mathrm{dist}_{\widetilde{\Lambda}}(y, x) < \rho_{\mathrm{im}}\}$.

**Non-uniform hyperbolicity.** Since $\Lambda$ is adapted to $\mu$, $\mu_\Lambda(\mathfrak{S}) = 0$. By Lemma 2.6, for $\mu_\Lambda$–a.e. $x \in \Lambda$, $f$ has one Lyapunov exponent in $(-\infty, -\chi_0 \inf R)$ and one Lyapunov exponent in $(\chi_0 \inf R, \infty)$. Let $\chi := \chi_0 \inf R$.

Non-uniformly hyperbolic set: Let $\mathrm{NUH}_\chi(f)$ be the set of $x \in \Lambda \setminus \bigcup_{n\in\mathbb{Z}} f^{-n}(\mathfrak{S})$ s.t. $T_{f^n(x)}\Lambda = E^u(f^n(x)) \oplus E^s(f^n(x))$, $n \in \mathbb{Z}$, where $E^u, E^s$ are one-dimensional linear subspaces, and:

(i) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log \|df_x^n \vec{v}\| < -\chi$ for all non-zero $\vec{v} \in E^s(x)$.
(ii) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log \|df_x^{-n} \vec{v}\| < -\chi$ for all non-zero $\vec{v} \in E^u(x)$.
(iii) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log |\sin \measuredangle(E^s(f^n(x)), E^u(f^n(x)))| = 0$.
(iv) $df_x E^s(x) = E^s(f(x))$ and $df_x E^u(x) = E^u(f(x))$.

By the Oseledets Theorem and Lemma 2.6, $\mu_\Lambda[\mathrm{NUH}_\chi(f)] = 1$.

**Pesin charts.** These are a system of coordinates on $\mathrm{NUH}_\chi(f)$ which simplifies the form of $f$. The following definition is slightly different than in Pesin's original work [Pes76], but the proofs are essentially the same.[2] Fix a measurable family of unit vectors $\underline{e}^u(x) \in E^u(x), \underline{e}^s(x) \in E^s(x)$ on $\mathrm{NUH}_\chi(f)$. Since $\dim E^{u/s}(x) = 1$, $\underline{e}^{u/s}(x)$ are determined up to a sign. To make the choice, let $(\underline{e}_x^1, \underline{e}_x^2)$ be a continuous choice of basis for $T_x\Lambda$ so that $\langle \underline{e}_x^1, \underline{e}_x^2, X_x\rangle$ has positive orientation. Pick $\underline{e}^u(x), \underline{e}^s(x)$ s.t. $\measuredangle(\underline{e}_x^1, \underline{e}^s(x)) \in [0, \pi)$, and $\measuredangle(\underline{e}^s(x), \underline{e}^u(x)) > 0$.

Pesin parameters: Given $x \in \mathrm{NUH}_\chi(f)$, let

- $\alpha(x) := \measuredangle(\underline{e}^s(x), \underline{e}^u(x))$,
- $s(x) := \sqrt{2}\left(\sum_{k=0}^{\infty} e^{2k\chi} \|df_x^k \underline{e}^s(x)\|_{f^k(x)}^2\right)^{\frac{1}{2}}$,
- $u(x) := \sqrt{2}\left(\sum_{k=0}^{\infty} e^{2k\chi} \|df_x^{-k} \underline{e}^u(x)\|_{f^{-k}(x)}^2\right)^{\frac{1}{2}}$.

The infinite series converge, because $x \in \mathrm{NUH}_\chi(f)$.

Oseledets-Pesin reduction: Define a linear transformation $C_\chi(x) : \mathbb{R}^2 \to T_x\Lambda$ by mapping $\binom{1}{0} \mapsto s(x)^{-1}\underline{e}^s(x)$ and $\binom{0}{1} \mapsto u(x)^{-1}\underline{e}^u(x)$.

This diagonalizes the derivative cocycle $df_x : T_xM \to T_{f(x)}M$ as follows.

**Theorem 3.1.** *$\exists C_\varphi$ s.t. $\forall x \in \mathrm{NUH}_\chi(f)$, $C_\chi(f(x))^{-1} \circ df_x \circ C_\chi(x) = \begin{pmatrix} A_x & 0 \\ 0 & B_x \end{pmatrix}$, where $C_\varphi^{-1} \leq |A_x| \leq e^{-\chi}$, and $e^\chi \leq |B_x| \leq C_\varphi$.*

The proof is a routine modification of the proofs in [BP07, theorem 3.5.5] or [KH95, theorem S.2.10], using the uniform bounds on $df \restriction_{\Lambda\setminus\mathfrak{S}}$ (Lemma 2.5).

[2] The difference is in the choice of $\chi$ in the exponential terms $e^{2k\chi}$ in the definitions of the Pesin parameters $s(x), u(x)$. In Pesin's work, $\chi =$ Lyapunov exponent of $x$ minus $\varepsilon$, while here it is constant.

Our conventions for $\underline{e}^s(x), \underline{e}^u(x)$ guarantee that $C_\chi(x)$ is orientation-preserving, and one can show exactly as in [Sar13, Lemmas 2.4–2.5] that

$$\|C_\chi(x)\| \leq 1 \text{ and } \frac{1}{\sqrt{2}}\frac{\sqrt{s(x)^2+u(x)^2}}{|\sin\alpha(x)|} \leq \|C_\chi(x)^{-1}\| \leq \frac{\sqrt{s(x)^2+u(x)^2}}{|\sin\alpha(x)|}. \tag{3.1}$$

We see that $\|C_\chi(x)^{-1}\|$ is large exactly when $E^s(x) \approx E^u(x)$ (small $|\sin\alpha(x)|$), or when it takes a long time to notice the exponential decay of $\frac{1}{n}\log\|df_x^n \underline{e}^s(x)\|$ or of $\frac{1}{n}\log\|df_x^{-n}\underline{e}^u(x)\|$ (large $s(x)$ or large $u(x)$). In summary, large $\|C_\chi(x)^{-1}\|$ means bad hyperbolicity.

Pesin Maps: The *Pesin map* at $x \in \mathrm{NUH}_\chi(f)$ (not to be confused with the Pesin *chart* defined below) is $\Psi_x : [-\rho_{\mathrm{dom}}, \rho_{\mathrm{dom}}]^2 \to \widetilde{\Lambda}$, given by

$$\Psi_x(u,v) = \mathrm{Exp}_x\left[C_\chi(x)\begin{pmatrix}u\\v\end{pmatrix}\right].$$

The map $\Psi_x$ is orientation-preserving, and it maps $[-\rho_{\mathrm{dom}}, \rho_{\mathrm{dom}}]^2$ diffeomorphically onto a neighborhood of $x$ in $\widetilde{\Lambda}\backslash\partial\widetilde{\Lambda}$. We have $\mathrm{Lip}(\Psi_x) \leq 2$, because $\|C_\chi(x)\| \leq 1$, but $\mathrm{Lip}(\Psi_x^{-1})$ is not uniformly bounded, because $\|C_\chi(x)^{-1}\|$ can be arbitrarily large.

Maximal size: Fix some parameter $0 < \varepsilon < 2^{-\frac{3}{2}}$ (which will be calibrated later). Although $\Psi_x$ is well-defined on all of $[-\rho_{\mathrm{dom}}, \rho_{\mathrm{dom}}]^2$, it will only be useful for us on the smaller set $[-Q_\varepsilon(x), Q_\varepsilon(x)]^2$, where

$$Q_\varepsilon(x) := \left\lfloor \varepsilon^{3/\beta}\left(\frac{\sqrt{s(x)^2+u(x)^2}}{|\sin\alpha(x)|}\right)^{-12/\beta} \wedge \big(\varepsilon\,\mathrm{dist}_\Lambda(x,\mathfrak{S})\big) \wedge \rho_{\mathrm{dom}}\right\rfloor_\varepsilon . \tag{3.2}$$

Here $\mathfrak{S}$ is the singular set of $\Lambda$, $a \wedge b := \min\{a,b\}$, $\beta$ is the constant in the $C^{1+\beta}$ assumption on $\varphi$, and $\lfloor t\rfloor_\varepsilon := \max\{\theta \in I_\varepsilon : \theta \leq t\}$ where $I_\varepsilon := \{e^{-\frac{1}{3}\ell\varepsilon} : \ell \in \mathbb{N}\}$.

The value $Q_\varepsilon(x)$ is called the *maximal size* (of the Pesin charts defined below).[3] Notice that $Q_\varepsilon \leq \varepsilon^{3/\beta}\|C_\chi^{-1}\|^{-12/\beta}$, so $Q_\varepsilon(x)$ is small when $x$ is close to $\mathfrak{S}$ or when the hyperbolicity at $x$ is bad. Another important property of $Q_\varepsilon$ is that, thanks to the inequalities $\|C_\chi\| \leq 1$ and $Q_\varepsilon < 2^{-\frac{3}{2}}\,\mathrm{dist}_\Lambda(x,\mathfrak{S})$,

$$\Psi_x\big([-Q_\varepsilon(x), Q_\varepsilon(x)]^2\big) \subset \Lambda \setminus \mathfrak{S}. \tag{3.3}$$

This is in contrast to $\Psi_x\big([-\rho_{\mathrm{dom}}, \rho_{\mathrm{dom}}]^2\big)$, which may intersect $\mathfrak{S}$ or $\widetilde{\Lambda}\setminus\Lambda$.

Pesin Charts: The *maximal Pesin chart* at $x \in \mathrm{NUH}_\chi(f)$ (with parameter $\varepsilon$) is $\Psi_x : [-Q_\varepsilon(x), Q_\varepsilon(x)]^2 \to \Lambda\setminus\mathfrak{S}$, $\Psi_x(u,v) = \mathrm{Exp}_x[C_\chi(x)\binom{u}{v}]$. The *Pesin chart of size* $\eta$ is $\Psi_x^\eta := \Psi_x \restriction_{[-\eta,\eta]^2}$ for $0 < \eta \leq Q_\varepsilon(x)$.

The Pesin charts provide a system of local coordinates on a neighborhood of $\mathrm{NUH}_\chi(f)$. The following theorem says that the Poincaré map "in coordinates"

$$f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x : [-Q_\varepsilon(x), Q_\varepsilon(x)]^2 \to \mathbb{R}^2$$

is close to a uniformly hyperbolic *linear* map. In what follows, $\underline{0} = \binom{0}{0}$ and the statement "*for all $\varepsilon$ small enough $P$ holds*" means "$\exists\varepsilon_0 > 0$ which depends only on $M, \varphi, \Lambda, \beta, \chi_0$ s.t. for all $0 < \varepsilon < \varepsilon_0$, $P$ holds".

[3] We do not claim that Theorem 3.2 below does not hold on larger boxes $[-Q', Q']^2$.

**Theorem 3.2** (Pesin)**.** *For all $\varepsilon$ small enough the following holds. For every $x \in \mathrm{NUH}_\chi(f)$, $f_x$ is well-defined and injective on $[-Q_\varepsilon(x), Q_\varepsilon(x)]^2$, and can be put there in the form $f_x(u,v) = \big(A_x u + h_x^1(u,v), B_x v + h_x^2(u,v)\big)$, where:*

(1) *$C_\varphi^{-1} \le |A_x| \le e^{-\chi}$ and $e^\chi \le |B_x| \le C_\varphi$, with $C_\varphi$ as in Theorem 3.1.*
(2) *$h_x^i$ are $C^{1+\frac{\beta}{2}}$ functions s.t. $h_x^i(\underline{0}) = 0$, $(\nabla h_x^i)(\underline{0}) = \underline{0}$.*
(3) *$\|h_x^i\|_{C^{1+\frac{\beta}{2}}} < \varepsilon$ on $[-Q_\varepsilon(x), Q_\varepsilon(x)]^2$.*

*A similar statement holds for $f_x^{-1} := \Psi_{f^{-1}(x)}^{-1} \circ f^{-1} \circ \Psi_x : [-Q_\varepsilon(x), Q_\varepsilon(x)]^2 \to \mathbb{R}^2$.*

*Proof.* Let $U := \Psi_x([-Q_\varepsilon(x), Q_\varepsilon(x)]^2)$. By (3.3), $f$ and $f^{-1}$ are $C^{1+\beta}$ on $U$, with uniform bounds on their $C^{1+\beta}$ norms (Lemma 2.5). Now continue as in [Sar13, Theorem 2.7] or [BP07, Theorem 5.6.1], replacing $M$ by $\Lambda$ and $\exp_p$ by $\mathrm{Exp}_p$. $\square$

**Adaptability and temperedness.** The maximal size of Pesin charts may not be bounded below on $\mathrm{NUH}_\chi(f)$. A central idea in Pesin theory is that it is nevertheless possible to control how fast $Q_\varepsilon$ decays along typical orbits.

Define for this purpose the set $\mathrm{NUH}_\chi^*(f)$ of all $x \in \mathrm{NUH}_\chi(f)$ which on top of the defining properties (i)–(iv) of $\mathrm{NUH}_\chi(f)$ also satisfy:

(v) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log \mathrm{dist}_\Lambda(f^n(x), \mathfrak{S}) = 0$.
(vi) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log \|C_\chi(f^n(x))^{-1}\| = 0$.
(vii) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log \|C_\chi(f^n(x))\underline{v}\| = 0$ for $\underline{v} = \binom{1}{0}, \binom{0}{1}$.
(viii) $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log |\det C_\chi(f^n(x))| = 0$.

**Lemma 3.3.** *$\mathrm{NUH}_\chi^*(f)$ is an $f$–invariant Borel set of full $\mu_\Lambda$–measure, and for every $x \in \mathrm{NUH}_\chi^*(f)$, $\lim\limits_{n\to\pm\infty} \frac{1}{n} \log Q_\varepsilon(f^n(x)) = 0$.*

*Proof.* Condition (v) holds $\mu_\Lambda$–a.e. because $\Lambda$ is adapted to $\mu$. Conditions (vi)–(viii) hold $\mu_\Lambda$–a.e. because of the Oseledets Theorem (apply the proof of [Sar13, Lemma 2.6] to the ergodic components of $\mu_\Lambda$). By (3.1), conditions (v)–(vi) imply that $\lim_{n\to\pm\infty} \frac{1}{n} \log Q_\varepsilon(f^n(x)) = 0$. $\square$

**Lemma 3.4** (Pesin's Temperedness Lemma)**.** *There exists a positive Borel function $q_\varepsilon : \mathrm{NUH}_\chi^*(f) \to (0,1)$ s.t. for every $x \in \mathrm{NUH}_\chi^*(f)$, $0 < q_\varepsilon(x) \le \varepsilon Q_\varepsilon(x)$ and $e^{-\varepsilon/3} \le \frac{q_\varepsilon \circ f}{q_\varepsilon} \le e^{\varepsilon/3}$.*

This lemma follows from Lemma 3.3 as in [BP07, Lemma 3.5.7]. It implies that

$$Q_\varepsilon(f^n(x)) > e^{-\frac{1}{3}|n|\varepsilon} q_\varepsilon(x) \text{ for all } n \in \mathbb{Z}, \tag{3.4}$$

which gives a control on the decay of $Q_\varepsilon$ along typical orbits.

**Overlapping Pesin charts.** Theorem 3.2 says that $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x$ is close to a linear hyperbolic map. This property is stable under perturbations, therefore we expect $f_{xy} := \Psi_y^{-1} \circ f \circ \Psi_x$ to be close to a linear hyperbolic map, whenever $\Psi_y$ is "sufficiently close" to $\Psi_{f(x)}$. We now specify the meaning of "sufficiently close".

Recall that $\Lambda$ is the disjoint union of a finite number of canonical transverse discs $S_r(p_i)$. Let $D_r(p_i) := S_r(p_i) \setminus \partial S_r(p_i) = \{\exp_{p_i}(\vec{v}) : \vec{v} \perp X_{p_i}, \|\vec{v}\| < r\}$. Choose for every $D = D_r(p_i)$ a map $\Theta : TD \to \mathbb{R}^2$ s.t.:

(1) $\Theta : T_x D \to \mathbb{R}^2$ is a linear isometry for all $x \in D$.

(2) Let $\vartheta_x = (\Theta \restriction_{T_xD})^{-1} : \mathbb{R}^2 \to T_xD$, then $(x,\underline{u}) \mapsto (\mathrm{Exp}_x \circ \vartheta_x)(\underline{u})$ is smooth and Lipschitz on $\Lambda \times \{\underline{u} \in \mathbb{R}^2 : \|\underline{u}\| < \rho_{\mathrm{dom}}\}$ with respect to the metric $d\big((x,\underline{u}),(x',\underline{u}')\big) := \mathrm{dist}_\Lambda(x,x') + \|\underline{u}-\underline{u}'\|$.

(3) $x \mapsto \vartheta_x^{-1} \circ \mathrm{Exp}_x^{-1}$ is a Lipschitz map from $D$ to $C^2(D,\mathbb{R}^2)$, the space of $C^2$ maps from $D$ to $\mathbb{R}^2$.

Recall that the Pesin map is $\Psi_x(u,v) = \mathrm{Exp}_x[C_\chi(x)\binom{u}{v}]$, and the Pesin chart of size $0<\eta<Q_\varepsilon(x)$ is $\Psi_x^\eta := \Psi_x \restriction_{[-\eta,\eta]^2}$.

OVERLAPPING CHARTS: Let $x_1, x_2 \in \mathrm{NUH}_\chi(f)$. We say that $\Psi_{x_1}^{\eta_1}, \Psi_{x_2}^{\eta_2}$ *$\varepsilon$–overlap*, and write $\Psi_{x_1}^{\eta_1} \overset{\varepsilon}{\approx} \Psi_{x_2}^{\eta_2}$, if $x_1, x_2$ lie in the same transversal disc of $\Lambda$, $e^{-\varepsilon} < \frac{\eta_1}{\eta_2} < e^{\varepsilon}$, and $\mathrm{dist}_\Lambda(x_1,x_2) + \|\Theta \circ C_\chi(x_1) - \Theta \circ C_\chi(x_2)\| < \eta_1^4\eta_2^4$.

**Proposition 3.5.** *The following holds for all $\varepsilon$ small enough. If $\Psi_{x_1}^{\eta_1} \overset{\varepsilon}{\approx} \Psi_{x_2}^{\eta_2}$ then:*

(1) $\Psi_{x_i}^{\eta_i}$ chart nearly the same patch: $\Psi_{x_i}\big([-e^{-2\varepsilon}\eta_i, e^{-2\varepsilon}\eta_i]^2\big) \subset \Psi_{x_j}\big([-\eta_j,\eta_j]^2\big)$.

(2) $\Psi_{x_i}$ define nearly the same coordinates: $\mathrm{dist}_{C^{1+\frac{\beta}{2}}}(\Psi_{x_i}^{-1} \circ \Psi_{x_j}, \mathrm{Id}) < \varepsilon\eta_i^2\eta_j^2$, *where the $C^{1+\frac{\beta}{2}}$ distance is calculated on $[-e^{-\varepsilon}\rho_{\mathrm{dom}}, e^{-\varepsilon}\rho_{\mathrm{dom}}]^2$.*

**Corollary 3.6.** *The following holds for all $\varepsilon$ small enough. If $x, y \in \mathrm{NUH}_\chi(f)$ and $\Psi_{f(x)}^{\eta'} \overset{\varepsilon}{\approx} \Psi_y^\eta$ then $f_{xy} := \Psi_y^{-1} \circ f \circ \Psi_x : [-Q_\varepsilon(x), Q_\varepsilon(x)]^2 \to \mathbb{R}^2$ is well-defined, injective, and can be put in the form $f_{xy}(u,v) = \big(A_{xy}u + h_{xy}^1(u,v), B_{xy}v + h_{xy}^2(u,v)\big)$, where:*

(1) $C_\varphi^{-1} \le |A_{xy}| \le e^{-\chi}$ *and* $e^\chi \le |B_{xy}| \le C_\varphi$, *with $C_\varphi$ as in Theorem* 3.1.

(2) $|h_{xy}^i(\underline{0})| < \varepsilon\eta$, $\|\nabla h_{xy}^i(\underline{0})\| < \varepsilon\eta^{\beta/3}$.

(3) $\|h_{xy}^i\|_{C^{1+\frac{\beta}{3}}} < \varepsilon$ *for $i = 1,2$, where the norm is taken on $[-Q_\varepsilon(x), Q_\varepsilon(x)]^2$.*

*A similar statement holds for $f_{xy}^{-1} := \Psi_x^{-1} \circ f^{-1} \circ \Psi_y$ whenever $\Psi_{f^{-1}(y)}^{\eta'} \overset{\varepsilon}{\approx} \Psi_x^\eta$.*

The proofs are routine modifications of [Sar13, Props. 3.2 and 3.4] once we replace $M$ by one of the canonical transverse discs in $\Lambda$ and $\exp_x$ by $\mathrm{Exp}_x$. For Proposition 3.5, we use the definition of overlap, and for Corollary 3.6 we treat $f_{xy} = (\Psi_y^{-1} \circ \Psi_{f(x)}) \circ f_x$ as a small perturbation of $f_x$ and then use Theorem 3.2.

## Part 2. Symbolic dynamics

Throughout this part we assume that $M, X$ and $\varphi$ satisfy our standing assumptions, and that $\mu$ is a $\chi_0$–hyperbolic $\varphi$–invariant probability measure on $M$. We fix a standard Poincaré section $\Lambda = \Lambda(p_1,\ldots,p_N; r)$ adapted to $\mu$, and a larger concentric standard section $\widetilde{\Lambda} := \Lambda(p_1,\ldots,p_N;\widetilde{r})$ s.t. $\widetilde{r} > 2r$. Let $f, R$ and $\mathfrak{S}$ denote the Poincaré map, roof function, and singular set of $\Lambda$, and let $\chi := \chi_0 \inf R$ (a bound for the Lyapunov exponents of $f$ at $\mu_\Lambda$–a.e. point, see Lemma 2.6).

In this part of the paper we construct a countable Markov partition for $f$ on a set of full measure with respect to $\mu_\Lambda$, and then use it to develop symbolic dynamics for $\varphi$. This was done in [Sar13] for surface diffeomorphisms, and the proof would have applied to our setup verbatim had $\mathfrak{S}$ been empty. We will indicate the changes needed to treat the case $\mathfrak{S} \neq \varnothing$.

Not many changes are needed, because most of the work is done inside Pesin charts, where $f$ and $f^{-1}$ are smooth with uniformly bounded $C^{1+\beta}$ norm. One point is worth mentioning, though: [Sar13] uses a uniform bound on $|\ln Q_\varepsilon(f(x))/Q_\varepsilon(x)|$, where $Q_\varepsilon(x)$ is the maximal size of a Pesin chart. This quantity is no longer bounded when $\mathfrak{S} \neq \varnothing$. When this or other effects of $\mathfrak{S}$ matter, we will give complete details. Otherwise, we will just sketch the general idea and refer to [Sar13] for details.

## 4. Generalized pseudo-orbits and shadowing

**Generalized pseudo-orbits (gpo).** Fix some small $\varepsilon > 0$. Recall that a pseudo-orbit with parameter $\varepsilon$ is a sequence of points $\{x_i\}_{i\in\mathbb{Z}}$ satisfying the nearest neighbor conditions $\mathrm{dist}(f(x_i), x_{i+1}) < \varepsilon$ for all $i \in \mathbb{Z}$. A gpo is also a sequence of objects satisfying nearest neighbor conditions, but the objects and the conditions are more complicated, because of the need to record the hyperbolic features of each point.

$\varepsilon$–Double charts: Ordered pairs $\Psi_x^{p^u,p^s} := (\Psi_x \restriction_{[-p^u,p^u]^2}, \Psi_x \restriction_{[-p^s,p^s]^2})$ where $x \in \mathrm{NUH}_\chi(f)$ and $0 < p^u, p^s \leq Q_\varepsilon(x)$ (same Pesin chart, different domains).

$\varepsilon$–Generalized pseudo-orbit (gpo): A sequence $\{\Psi_{x_i}^{p_i^u,p_i^s}\}_{i\in\mathbb{Z}}$ of $\varepsilon$–double charts which satisfies the following nearest neighbor conditions for all $i \in \mathbb{Z}$:

(GPO1) $\Psi_{f(x_i)}^{p_{i+1}^u \wedge p_{i+1}^s} \overset{\varepsilon}{\approx} \Psi_{x_{i+1}}^{p_{i+1}^u \wedge p_{i+1}^s}$ and $\Psi_{f^{-1}(x_{i+1})}^{p_i^u \wedge p_i^s} \overset{\varepsilon}{\approx} \Psi_{x_i}^{p_i^u \wedge p_i^s}$, cf. Prop. 3.5.
(GPO2) $p_{i+1}^u = \min\{e^\varepsilon p_i^u, Q_\varepsilon(x_{i+1})\}$ and $p_i^s = \min\{e^\varepsilon p_{i+1}^s, Q_\varepsilon(x_i)\}$.

A *positive gpo* is a one-sided sequence $\{\Psi_{x_i}^{p_i^u,p_i^s}\}_{i\geq 0}$ with (GPO1), (GPO2). A *negative gpo* is a one-sided sequence $\{\Psi_{x_i}^{p_i^u,p_i^s}\}_{i\leq 0}$ with (GPO1), (GPO2). Gpos were called "chains" in [Sar13].

Shadowing: A gpo $\{\Psi_{x_i}^{p_i^u,p_i^s}\}_{i\in\mathbb{Z}}$ *shadows the orbit of* $x$, if $f^i(x) \in \Psi_{x_i}\big([-\eta_i,\eta_i]^2\big)$ for all $i \in \mathbb{Z}$, where $\eta_i := p_i^u \wedge p_i^s$.

This notation is heavy, so we will sometimes abbreviate it by writing $v_i$ instead of $\Psi_{x_i}^{p_i^u,p_i^s}$, and letting $p^u(v_i) := p_i^u, p^s(v_i) := p_i^s, x(v_i) := x_i$. The nearest neighbor conditions [(GPO1) + (GPO2)] will be expressed by the notation $v_i \xrightarrow{\varepsilon} v_{i+1}$.

**Lemma 4.1.** *Suppose* $0 < p_i^u, p_i^s \leq Q_i$ *satisfy* $p_{i+1}^u = \min\{e^\varepsilon p_i^u, Q_{i+1}\}$ *and* $p_i^s = \min\{e^\varepsilon p_{i+1}^s, Q_i\}$ *for* $i = 0, 1$. *If* $\eta_i := p_i^u \wedge p_i^s$, *then* $\eta_{i+1}/\eta_i \in [e^{-\varepsilon}, e^\varepsilon]$.

See [Sar13, Lemma 4.4] for the proof. Let $v = \Psi_x^{p^u,p^s}$ be an $\varepsilon$–double chart.

Admissible manifolds: An $s$*–admissible manifold* in $v$ is a set of the form $\Psi_x\{(t, F(t)) : |t| \leq p^s\}$, where $F : [-p^s, p^s] \to \mathbb{R}$ satisfies:
(Ad1) $|F(0)| \leq 10^{-3}(p^u \wedge p^s)$.
(Ad2) $|F'(0)| \leq \frac{1}{2}(p^u \wedge p^s)^{\beta/3}$.
(Ad3) $F$ is $C^{1+\frac{\beta}{3}}$ and $\sup |F'| + \mathrm{H\ddot{o}l}_{\beta/3}(F) \leq \frac{1}{2}$.
Similarly, a $u$*–admissible manifold* in $v$ is a set of the form $\Psi_x\{(F(t), t) : |t| \leq p^u\}$, where $F : [-p^u, p^u] \to \mathbb{R}$ satisfies (Ad1–3).

The constant $10^{-3}$ in (Ad1) implies that $s$–admissible manifolds in $v$ intersect $u$–admissible manifolds in $v$ inside the smaller set $\Psi_x([-10^{-2}\eta, 10^{-2}\eta]^2)$, where $\eta = p^u \wedge p^s$ (see Theorem 4.2 below). We call $F$ the *representing function*, and we denote the collections of all $s/u$–admissible manifolds in $v$ by $\mathscr{M}^s(v)$ and $\mathscr{M}^u(v)$. The representing function satisfies $\|F\|_\infty \leq Q_\varepsilon(x)$, because $p^u, p^s \leq Q_\varepsilon(x)$, $|F(0)| \leq 10^{-3}(p^u \wedge p^s)$ and $|F'| \leq \frac{1}{2}$.[4] As a result, $s/u$–admissible manifolds are subsets of $\Psi_x([-Q_\varepsilon(x), Q_\varepsilon(x)]^2)$, a set where $f$ is smooth, and where if $\varepsilon$ is small enough

[4] In fact $|F'(t)| \leq |F'(0)| + \frac{1}{2}|t|^{\beta/3} \leq \varepsilon$ for $t \in \mathrm{dom}(F)$, since $|t| \leq p^{u/s} \leq Q_\varepsilon(x) \leq \varepsilon^{3/\beta}$.

then $f$ is a perturbation of a uniformly hyperbolic linear map in Pesin coordinates (Theorem 3.2). This implies the following fact.

GRAPH TRANSFORM LEMMA: *For all $\varepsilon$ small enough, if $v_i \xrightarrow{\varepsilon} v_{i+1}$, then the forward image of a $u$–admissible manifold $V^u \in \mathscr{M}^u(v_i)$ contains a unique $u$–admissible manifold $\mathscr{F}^u_{v_i v_{i+1}}[V^u]$, called the* (forward) graph transform *of $V^u$.*

*Sketch of proof* (see [Sar13, Prop. 4.12], [KH95, Supplement], or [BP07] for details). Let $f_{x_i x_{i+1}} := \Psi^{-1}_{x_{i+1}} \circ f \circ \Psi_{x_i} : [-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2 \to \mathbb{R}^2$. By (GPO1) and Corollary 3.6, $f_{x_i x_{i+1}}$ is $\varepsilon$ close in the $C^{1+\frac{\beta}{3}}$ norm on $[-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2$ to a linear map which contracts the $x$–coordinate by at least $e^{-\chi}$ and expands the $y$–coordinate by at least $e^{\chi}$. Direct calculations show that if $\varepsilon$ is much smaller than $\chi$ and $V^u \in \mathscr{M}^u(v_i)$, then $f(V^u) \supset \Psi_{x_{i+1}}\{(G(t), t) : t \in [a, b]\}$ where $G$ satisfies (Ad1–3), and $[a, b] \supset [-e^{\chi/2} p^u_i, e^{\chi/2} p^u_i]$. By (GPO2), if $\varepsilon < \chi/2$ then $[-e^{\chi/2} p^u_i, e^{\chi/2} p^u_i] \supset [-p^u_{i+1}, p^u_{i+1}]$, so $f(V^u)$ restricts to a $u$–admissible manifold in $v_{i+1}$.

There is also a *(backward) graph transform* $\mathscr{F}^s_{v_{i+1} v_i} : \mathscr{M}^s(v_{i+1}) \to \mathscr{M}^s(v_i)$, obtained by applying $f^{-1}$ to $s$–admissible manifolds in $v_{i+1}$ and restricting the result to an $s$–admissible manifold in $v_i$. Put a metric on $\mathscr{M}^u(v_i)$ and $\mathscr{M}^s(v_{i+1})$ by measuring the sup-norm distance between the representing functions. Using the form of $f_{x_i x_{i+1}}$ in coordinates, one can show by direct calculations that $\mathscr{F}^u_{v_i v_{i+1}} : \mathscr{M}^u(v_i) \to \mathscr{M}^u(v_{i+1})$ and $\mathscr{F}^s_{v_{i+1} v_i} : \mathscr{M}^s(v_{i+1}) \to \mathscr{M}^s(v_i)$ contract distances by at least $e^{-\chi/2}$, see [Sar13, Prop. 4.14].

Suppose $\underline{v}^- = \{v_i\}_{i \leq 0}$ is a *negative* gpo, and pick arbitrary $V^u_{-n} \in \mathscr{M}^u(v_{-n})$ ($n \geq 0$), then $V^u_{0,n} := (\mathscr{F}^u_{v_{-1} v_0} \circ \cdots \circ \mathscr{F}^u_{v_{-n+1} v_{-n+2}} \circ \mathscr{F}^u_{v_{-n} v_{-n+1}})(V^u_{-n}) \in \mathscr{M}^u(v_0)$. Using the uniform contraction of $\mathscr{F}_{v_{-i-1} v_{-i}}$, it is easy to see that $\{V^u_{0,n}\}_{n \geq 1}$ is a Cauchy sequence, and that its limit is independent of the choice of $V^u_{-n}$ [Sar13, Prop. 4.15]. Thus we can make the following definition for all $\varepsilon$ small enough.

THE UNSTABLE MANIFOLD OF A NEGATIVE GPO $\underline{v}^-$:

$$V^u[\underline{v}^-] := \lim_{n \to \infty} (\mathscr{F}^u_{v_{-1} v_0} \circ \cdots \circ \mathscr{F}^u_{v_{-n+1} v_{-n+2}} \circ \mathscr{F}^u_{v_{-n} v_{-n+1}})(V^u_{-n})$$

for some (any) choice of $V^u_{-n} \in \mathscr{M}^u(v_{-n})$.

Working with *positive* gpos and *backward* graph transforms, we can also make the following definition.

THE STABLE MANIFOLD OF A POSITIVE GPO $\underline{v}^+$:

$$V^s[\underline{v}^+] := \lim_{n \to \infty} (\mathscr{F}^s_{v_1 v_0} \circ \cdots \circ \mathscr{F}^s_{v_{n-1} v_{n-2}} \circ \mathscr{F}^s_{v_n v_{n-1}})(V^s_n)$$

for some (any) choice of $V^s_n \in \mathscr{M}^s(v_n)$.

The following properties hold:

(1) ADMISSIBILITY: $V^u[\underline{v}^-] \in \mathscr{M}^u(v_0)$ and $V^s[\underline{v}^+] \in \mathscr{M}^s(v_0)$. This is because $\mathscr{M}^u(v_0), \mathscr{M}^s(v_0)$ are closed in the supremum norm.
(2) INVARIANCE: $f^{-1}(V^u[\{v_i\}_{i \leq 0}]) \subset V^u[\{v_i\}_{i \leq -1}]$, $f(V^s[\{v_i\}_{i \geq 0}]) \subset V^s[\{v_i\}_{i \geq 1}]$. This is immediate from the definition.
(3) HYPERBOLICITY: if $x, y \in V^u[\underline{v}^-]$ then $\mathrm{dist}_\Lambda(f^{-n}(x), f^{-n}(y)) \xrightarrow[n \to \infty]{} 0$, and if $x, y \in V^s[\underline{v}^+]$ then $\mathrm{dist}_\Lambda(f^n(x), f^n(y)) \xrightarrow[n \to \infty]{} 0$. The rates are exponential.

(4) HÖLDER PROPERTY: The maps $\underline{v} \mapsto V^u[\{v_i\}_{i\leq 0}], V^s[\{v_i\}_{i\geq 0}]$ are Hölder continuos, i.e. there exist constants $K > 0$ and $0 < \theta < 1$ s.t. for all $n \geq 0$, if $\{u_i\}_{i\in\mathbb{Z}}, \{v_i\}_{i\in\mathbb{Z}}$ are gpo's with $u_i = v_i$ for all $|i| \leq n$ then

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

Above, $\mathrm{dist}_{C^1}$ is the $C^1$ distance between two admissible manifolds: if $V_1 = \Psi_x\{(t, F_1(t)) : |t| \leq p^s\}$, $V_2 = \Psi_x\{(t, F_2(t)) : |t| \leq p^s\}$ are $s$–admissible manifolds then

$$\mathrm{dist}_{C^1}(V_1, V_2) := \max |F_1 - F_2| + \max |F_1' - F_2'|,$$

and a similar definition holds for $u$–admissible manifolds.

To prove part (3) notice first that by the invariance property, $f^n(V^s[\underline{v}^+])$ and $f^{-n}(V^u[\underline{v}^-])$ remain inside Pesin charts. Therefore $f^n\restriction_{V^s[\underline{v}^+]}$ and $f^{-n}\restriction_{V^u[\underline{v}^-]}$ can be written in Pesin coordinates as compositions of $n$ uniformly hyperbolic maps on $\mathbb{R}^2$. One can then use direct calculations as in the proof of Pesin's Stable Manifold Theorem to prove (3). See e.g. [Sar13, Prop. 6.3].

Part (4) is proved almost verbatim as in the case of diffeomorphisms [Sar13, Prop. 4.15(5)]. Here is a crude explanation: the Pesin charts avoid the singular set hence their $C^{1+\beta}$ norms are uniformly bounded, and since $V^u[\{v_i\}_{i\leq 0}], V^s[\{v_i\}_{i\geq 0}]$ are limits of admissible manifolds via the graph transform method, they depend Hölder continuously on $\underline{v}$.

**The Shadowing Lemma.**

**Theorem 4.2.** *The following holds for all $\varepsilon$ small enough: Every gpo with parameter $\varepsilon$ shadows a unique orbit.*

*Sketch of proof.* Let $\underline{v} = \{v_i\}_{i\in\mathbb{Z}}$ be a gpo, $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$. We have to show that there exists a unique $x$ s.t. $f^i(x) \in \Psi_{x_i}([-\eta_i, \eta_i]^2)$ for all $i \in \mathbb{Z}$, where $\eta_i = p_i^u \wedge p_i^s$. The sets $V^u := V^u[\{v_i\}_{i\leq 0}]$ and $V^s := V^s[\{v_i\}_{i\geq 0}]$ are admissible manifolds in $v_0$. Because of properties (Ad1–3), $V^u$ and $V^s$ intersect at a unique point $x$, and $x$ belongs to $\Psi_{x_0}([-10^{-2}\eta_0, 10^{-2}\eta_0]^2)$ [Sar13, Prop. 4.11], see also [KH95, Cor. S.3.8]. By the invariance property,

$$f^n(x) \in \Psi_{x_n}([-Q_\varepsilon(x_n), Q_\varepsilon(x_n)]^2) \text{ for all } n \in \mathbb{Z}.$$

We will show that $\underline{v}$ shadows $x$, and that $x$ is the only such point.

Any $y$ s.t. $f^n(y) \in \Psi_{x_n}([-Q_\varepsilon(x_n), Q_\varepsilon(x_n)]^2)$ for all $n \in \mathbb{Z}$ equals $x$. The map $f_{x_n x_{n+1}} := \Psi_{x_{n+1}}^{-1} \circ f \circ \Psi_{x_n}$ is uniformly hyperbolic on $[-Q_\varepsilon(x_n), Q_\varepsilon(x_n)]^2$. If $\Psi_{x_0}^{-1}(x)$ and $\Psi_{x_0}^{-1}(y)$ have different $y$–coordinates, then successive application of $f_{x_n x_{n+1}}$ will expand the difference between the $y$–coordinates of $\Psi_{x_n}^{-1}(f^n(x))$, $\Psi_{x_n}^{-1}(f^n(y))$ exponentially as $n \to \infty$. If $\Psi_{x_0}^{-1}(x), \Psi_{x_0}^{-1}(y)$ have different $x$–coordinates, then successive application of $f_{x_{-n-1},x_{-n}}^{-1}$ will expand the difference between the $x$–coordinates of $\Psi_{x_n}^{-1}(f^n(x))$,$\Psi_{x_n}^{-1}(f^n(y))$ exponentially as $n \to -\infty$. But these differences are bounded by $2Q_\varepsilon(x_n)$ whence by a constant, so $\Psi_{x_0}^{-1}(x) = \Psi_{x_0}^{-1}(y)$, whence $x = y$.

Let $y_k$ denote the unique intersection point of $V^u[\{v_i\}_{i\leq k}]$ and $V^s[\{v_i\}_{i\geq k}]$, then $f^n(y_k), f^{n+k}(x) \in \Psi_{x_{n+k}}([-Q_\varepsilon(x_{n+k}), Q_\varepsilon(x_{n+k})]^2)$ for all $n \in \mathbb{Z}$. By the previous paragraph, $y_k = f^k(x)$. Since $y_k$ is the intersection of a $u$–admissible manifold and an $s$–admissible manifold in $v_k$, $y_k \in \Psi_{x_k}([-\eta_k, \eta_k]^2)$ where $\eta_k := p_k^u \wedge p_k^s$. It follows that $f^k(x) \in \Psi_{x_k}([-\eta_k, \eta_k]^2)$ for all $k \in \mathbb{Z}$. Thus $\underline{v}$ shadows the orbit of $x$, and $x$ is unique with this property. $\square$

**Which points are shadowed by gpos?** To appreciate where the difficulty lies, let us try the naïve approach: given $x \in \mathrm{NUH}_\chi(f)$, set $x_i := f^i(x)$, and look for $p_i^u, p_i^s$ s.t. $\{\Psi_{x_i}^{p_i^u, p_i^s}\}_{i\in\mathbb{Z}}$ is a gpo. (GPO1) is automatic, but without additional information on $Q_\varepsilon(f^i(x))$, it is not clear that there exist $p_i^u, p_i^s$ satisfying (GPO2). This is where we will use the adaptedness of $\Lambda$: $\lim\limits_{|n|\to\infty} \frac{1}{n}\log Q_\varepsilon(f^n(x)) = 0$ for a.e. $x$, whence by (3.4) there exist $q_\varepsilon(x) > 0$ s.t. $Q_\varepsilon(f^n(x)) > e^{-\frac{1}{3}\varepsilon|n|} q_\varepsilon(x) > e^{-\varepsilon|n|} q_\varepsilon(x)$ for all $n \in \mathbb{Z}$. So the following suprema range over non-empty sets:

$$
\begin{aligned}
p_i^u &:= \sup\{t > 0 : Q_\varepsilon(f^{i-n}(x)) \geq e^{-\varepsilon n} t \text{ for all } n \geq 0\},\\
p_i^s &:= \sup\{t > 0 : Q_\varepsilon(f^{i+n}(x)) \geq e^{-\varepsilon n} t \text{ for all } n \geq 0\}.
\end{aligned}
$$

It is easy to see that $p_i^u, p_i^s$ satisfy (GPO2), so $\{\Psi_{f^i(x)}^{p_i^u, p_i^s}\}_{i\in\mathbb{Z}}$ is a gpo shadowing $x$.

If we want to use the previous construction to shadow a set of full measure of orbits, then we need uncountably many "letters" $\Psi_x^{p^u, p^s}$. The following proposition achieves this with a countable discrete collection. Recall the definition of $\mathrm{NUH}_\chi^*(f)$ from Lemma 3.3, and let

$$
\mathrm{NUH}_\chi^\#(f) := \{x \in \mathrm{NUH}_\chi^*(f) : \limsup_{n\to\infty} q_\varepsilon(f^n(x)), \limsup_{n\to-\infty} q_\varepsilon(f^n(x)) \neq 0\}. \tag{4.1}
$$

Thus $\mathrm{NUH}_\chi^\#(f)$ is $f$–invariant of full measure (by the Poincaré recurrence theorem).

**Proposition 4.3.** *The following holds for all $\varepsilon$ small enough. There exists a countable collection of $\varepsilon$–double charts $\mathscr{A}$ with the following properties:*

(1) Discreteness: *Let $D(x) := \mathrm{dist}_\Lambda\big(\{x, f(x), f^{-1}(x)\}, \mathfrak{S}\big)$, then for every $t > 0$ the set $\{\Psi_x^{p^u,p^s} \in \mathscr{A} : D(x), p^u, p^s > t\}$ is finite.*

(2) Sufficiency: *For every $x \in \mathrm{NUH}_\chi^\#(f)$ there is a gpo $\{v_n\}_{n\in\mathbb{Z}} \in \mathscr{A}^{\mathbb{Z}}$ which shadows $x$, and which satisfies $p^u(v_n) \wedge p^s(v_n) \geq e^{-\varepsilon/3} q_\varepsilon(f^n(x))$ for all $n \in \mathbb{Z}$.*

(3) Relevance: *For every $v \in \mathscr{A}$ there is a gpo $\underline{v} \in \mathscr{A}^{\mathbb{Z}}$ s.t. $v_0 = v$ and $\underline{v}$ shadows a point in $\mathrm{NUH}_\chi(f)$.*

*Proof.* The proof for diffeomorphisms in [Sar13, Props. 3.5 and 4.5] does not extend to our case, because it uses a uniform bound $F^{-1} \leq Q_\varepsilon \circ f / Q_\varepsilon \leq F$ which does not hold in the presence of singularities. We bypass this difficulty as follows.

Let $X := [\Lambda \setminus \bigcup_{i=-1,0,1} f^i(\mathfrak{S})]^3 \times (0,\infty)^3 \times \mathrm{GL}(2,\mathbb{R})$, together with the product topology, and let $Y \subset X$ denote the subset of $(\underline{x}, \underline{Q}, \underline{C}) \in X$ of the form

$$
\begin{aligned}
\underline{x} &= (x, f(x), f^{-1}(x)), \text{ where } x \in \mathrm{NUH}_\chi^*(f),\\
\underline{Q} &= (Q_\varepsilon(x), Q_\varepsilon(f(x)), Q_\varepsilon(f^{-1}(x))),\\
\underline{C} &= (C_\chi(x), C_\chi(f(x)), C_\chi(f^{-1}(x))).
\end{aligned}
$$

Cut $Y$ into the countable disjoint union $Y = \biguplus_{(k,\underline{\ell})\in\mathbb{N}_0^4} Y_{k,\underline{\ell}}$ where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and

$$
Y_{k,\underline{\ell}} := \left\{ (\underline{x}, \underline{Q}, \underline{C}) \in Y : \begin{array}{c} x \in \mathrm{NUH}_\chi^*(f),\ e^{-(k+1)} < Q_\varepsilon(x) \leq e^{-k}, \text{ and} \\ e^{-(\ell_i+1)} < \mathrm{dist}_\Lambda(f^i(x), \mathfrak{S}) \leq e^{-\ell_i} \quad (i = 0, 1, -1) \end{array} \right\}.
$$

Precompactness Lemma. *$Y_{k,\underline{\ell}}$ are precompact in $X$.*

*Proof.* Suppose $(\underline{x},\underline{Q},\underline{C})\in Y_{k,\underline{\ell}}$. By (3.1) and (3.2), $\|C_\chi(x)^{-1}\|\leq \varepsilon^{\frac14}e^{\frac{\beta(k+1)}{12}}$. One can show as in [Sar13, page 403], that $C^{-1}\leq \|C_\chi(y)^{-1}\|/\|C_\chi(f(y))^{-1}\|\leq C$ for all $y\in \mathrm{NUH}_\chi(f)$ for some global constant $C$. It follows that

$$\|C_\chi(f^i(x))^{-1}\|\leq C\varepsilon^{\frac14}e^{\frac{\beta(k+1)}{12}} \text{ for } i=0,1,-1.$$

Since $C_\chi(\cdot)$ is a contraction, $\underline{C}\in G_k\times G_k\times G_k$, where $G_k$ is the compact set $\{A\in \mathrm{GL}(2,\mathbb{R}):\|A\|\leq 1, \|A^{-1}\|\leq C\varepsilon^{\frac14}e^{\frac{\beta(k+1)}{12}}\}$.

Next we bound $\underline{Q}$ in a compact set. By (3.1), if $(\underline{x},\underline{Q},\underline{C})\in Y_{k,\underline{\ell}}$, then

$$\frac{Q_\varepsilon(f^i(x))}{Q_\varepsilon(x)}\geq \left(\frac{\sqrt{2}\|C_\chi(f^i(x))^{-1}\|}{\|C_\chi(x)^{-1}\|}\right)^{-\frac{12}{\beta}}\wedge\left(\frac{\mathrm{dist}_\Lambda(f^i(x),\mathfrak{S})}{\mathrm{dist}_\Lambda(x,\mathfrak{S})}\right)\geq 2^{-\frac{6}{\beta}}C^{-\frac{12}{\beta}}\wedge e^{-(\ell_i+1)},$$

whence $Q_\varepsilon(f^i(x))\geq 2^{-\frac{6}{\beta}}C^{-\frac{12}{\beta}}e^{-(\ell_i+1)}e^{-(k+1)}$. By definition, $Q_\varepsilon(f^i(x))\leq \rho_{\mathrm{dom}}$, so $\underline{Q}\in (F_{k,\underline{\ell}})^3$ with $F_{k,\underline{\ell}}\subset\mathbb{R}$ compact.

Finally, $\underline{x}\in E_{\ell_0}\times f(E_{\ell_0})\times f^{-1}(E_{\ell_0})$, with $E_{\ell_0}:=\{y\in\Lambda:\mathrm{dist}_\Lambda(y,\mathfrak{S})\geq e^{-\ell_0-1}\}$. The set $E_{\ell_0}$ is compact because $\Lambda$ is compact, and $f(E_{\ell_0}), f^{-1}(E_{\ell_0})$ are compact because $f\restriction_{E_{\ell_0}}, f^{-1}\restriction_{E_{\ell_0}}$ are continuous.

In summary, $Y_{k,\underline{\ell}}\subset\prod_{i=0,1,-1}f^i(E_{\ell_0})\times(F_{k,\underline{\ell}})^3\times(G_k)^3$, a compact subset of $X$. So $Y_{k,\underline{\ell}}$ is precompact in $X$, proving the lemma.

Since $Y_{k,\underline{\ell}}$ is precompact in $X$, $Y_{k,\underline{\ell}}$ contains a finite set $Y_{k,\underline{\ell}}(m)$ s.t. for every $(\underline{x},\underline{Q},\underline{C})\in Y_{k,\underline{\ell}}$ there exists some $(\underline{y},\underline{Q}',\underline{C}')\in Y_{k,\underline{\ell}}(m)$ s.t. for every $|i|\leq 1$:

(a) $\mathrm{dist}_\Lambda(f^i(x),f^i(y))+\|\Theta\circ C_\chi(f^i(x))-\Theta\circ C_\chi(f^i(y))\|<e^{-8(m+3)}$.
(b) $e^{-\varepsilon/3}<Q_\varepsilon(f^i(x))/Q_\varepsilon(f^i(y))<e^{\varepsilon/3}$.

($\Theta$ is defined at the end of §3.)

Definition of $\mathscr{A}$: The set of $\varepsilon$–double charts $\Psi_x^{p^u,p^s}$ s.t. for some $k,\ell_0,\ell_1,\ell_{-1},m$:

(A1) $x$ is the first coordinate of some $(\underline{x},\underline{Q},\underline{C})\in Y_{k,\underline{\ell}}(m)$.
(A2) $0<p^u,p^s\leq Q_\varepsilon(x)$ and $p^u,p^s\in I_\varepsilon=\{e^{-\frac13\ell\varepsilon}:\ell\in\mathbb{N}\}$.
(A3) $p^u\wedge p^s\in[e^{-m-2},e^{-m+2}]$.

Proof that $\mathscr{A}$ is discrete: Fix $t>0$. Suppose $\Psi_x^{p^u,p^s}\in\mathscr{A}$ and let $k,\underline{\ell},m$ be as above. If $D(x),p^u,p^s>t$, then:

- $k\leq|\log t|$ because $t<p^u\leq Q_\varepsilon(x)\leq e^{-k}$.
- $\ell_i\leq|\log t|$ because $t<D(x)\leq \mathrm{dist}_\Lambda(f^i(x),\mathfrak{S})\leq e^{-\ell_i}$.
- $m\leq|\log t|+2$ because $t<p^u\wedge p^s\leq e^{-m+2}$.

So $\#\{x:\Psi_x^{p^u,p^s}\in\mathscr{A}, D(x),p^u,p^s>t\}\leq \displaystyle\sum_{k,\ell_0,\ell_1,\ell_{-1}=0}^{\lceil|\log t|\rceil}\sum_{m=0}^{\lceil|\log t|\rceil+2}\#Y_{k,\underline{\ell}}(m)<\infty$. Also, $\#\{(p^u,p^s):\Psi_x^{p^u,p^s}\in\mathscr{A},D(x),p^u,p^s>t\}\leq(\#(I_\varepsilon\cap[t,1]))^2<\infty$. Thus $\#\{\Psi_x^{p^u,p^s}\in\mathscr{A}:D(x),p^u,p^s>t\}<\infty$, proving that $\mathscr{A}$ is discrete.

The proof of sufficiency requires some preparation. A sequence $\{(p_n^u,p_n^s)\}_{n\in\mathbb{Z}}$ is called *$\varepsilon$–subordinated* to a sequence $\{Q_n\}_{n\in\mathbb{Z}}\subset I_\varepsilon$, if $0<p_n^u,p_n^s\leq Q_n$; $p_n^u,p_n^s\in I_\varepsilon$; $p_{n+1}^u=\min\{e^\varepsilon p_n^u,Q_{n+1}\}$ and $p_n^s=\min\{e^\varepsilon p_{n+1}^s,Q_n\}$ for all $n$.

First Subordination Lemma. *Let $\{q_n\}_{n\in\mathbb{Z}},\{Q_n\}_{n\in\mathbb{Z}}\subset I_\varepsilon$. If for every $n\in\mathbb{Z}$ $0<q_n\leq Q_n$ and $e^{-\varepsilon}\leq q_n/q_{n+1}\leq e^\varepsilon$, then there exists $\{(p_n^u,p_n^s)\}_{n\in\mathbb{Z}}$ which is $\varepsilon$–subordinated to $\{Q_n\}_{n\in\mathbb{Z}}$, and such that $p_n^u\wedge p_n^s\geq q_n$ for all $n$.*

Second Subordination Lemma. *Suppose $\{(p^u_n,p^s_n)\}_{n\in\mathbb{Z}}$ is $\varepsilon$–subordinated to $\{Q_n\}_{n\in\mathbb{Z}}$. If $\limsup\limits_{n\to\infty}(p^u_n\wedge p^s_n)>0$ and $\limsup\limits_{n\to-\infty}(p^u_n\wedge p^s_n)>0$, then $p^u_n$ (resp. $p^s_n$) is equal to $Q_n$ for infinitely many $n>0$, and for infinitely many $n<0$.*

These are Lemmas 4.6 and 4.7 in [Sar13].

Proof of sufficiency: Fix $x\in\mathrm{NUH}^{\#}_\chi(f)$. Recall the definition of $q_\varepsilon(\cdot)$ from Pesin's Temperedness Lemma (Lemma 3.4), and choose $q_n\in I_\varepsilon$ s.t. $q_n/q_\varepsilon(f^n(x))\in[e^{-\varepsilon/3},e^{\varepsilon/3}]$. Necessarily $e^{-\varepsilon}\le q_n/q_{n+1}\le e^{\varepsilon}$.

By the first subordination lemma there exists $\{(q^u_n,q^s_n)\}_{n\in\mathbb{Z}}$ s.t. $\{(q^u_n,q^s_n)\}_{n\in\mathbb{Z}}$ is $\varepsilon$–subordinated to $\{e^{-\varepsilon/3}Q_\varepsilon(f^n(x))\}_{n\in\mathbb{Z}}$, and $q^u_n\wedge q^s_n\ge q_n$ for all $n\in\mathbb{Z}$. Let $\eta_n:=q^u_n\wedge q^s_n$. By Lemma 4.1, $e^{-\varepsilon}\le\eta_{n+1}/\eta_n\le e^{\varepsilon}$. Since $\eta_n\ge q_n\ge e^{-\varepsilon/3}q_\varepsilon(f^n(x))$ and $x\in\mathrm{NUH}^{\#}_\chi(f)$, $\limsup_{n\to\pm\infty}\eta_n>0$.

Choose non-negative integers $m_n,k_n,\underline{\ell}_n=(\ell^n_0,\ell^n_1,\ell^n_{-1})$ s.t. for all $n\in\mathbb{Z}$:

- ○ $\eta_n\in[e^{-m_n-1},e^{-m_n+1}]$.
- ○ $Q_\varepsilon(f^n(x))\in(e^{-k_n-1},e^{-k_n}]$.
- ○ $\mathrm{dist}_\Lambda(f^{n+i}(x),\mathfrak{S})\in(e^{-\ell^n_i-1},e^{-\ell^n_i}]$ for $i=0,1,-1$.

Choose an element of $Y_{k_n,\underline{\ell}_n}$ with first coordinate $f^n(x)$, and approximate it by some element of $Y_{k_n,\underline{\ell}_n}(m_n)$ with first coordinate $x_n$ s.t. for $i=0,1,-1$:

(a$_n$) $\mathrm{dist}_\Lambda(f^i(f^n(x)),f^i(x_n))+\|\Theta\circ C_\chi(f^i(f^n(x)))-\Theta\circ C_\chi(f^i(x_n))\|<e^{-8(m_n+3)}$.
(b$_n$) $e^{-\varepsilon/3}<Q_\varepsilon(f^i(f^n(x)))/Q_\varepsilon(f^i(x_n))<e^{\varepsilon/3}$.

By (b$_n$) with $i=0$, $Q_\varepsilon(x_n)\ge e^{-\varepsilon/3}Q_\varepsilon(f^n(x))\ge\eta_n$. By the first subordination lemma, there exists $\{(p^u_n,p^s_n)\}_{n\in\mathbb{Z}}$ $\varepsilon$–subordinated to $\{Q_\varepsilon(x_n)\}_{n\in\mathbb{Z}}$ such that $p^u_n\wedge p^s_n\ge\eta_n$ for all $n\in\mathbb{Z}$. Necessarily, $p^u_n\wedge p^s_n\ge e^{-\varepsilon/3}q_\varepsilon(f^n(x))$. Let

$$\underline{v}:=\{\Psi^{p^u_n,p^s_n}_{x_n}\}_{n\in\mathbb{Z}}.$$

We will show that $\underline{v}\in\mathscr{A}^{\mathbb{Z}}$, $\underline{v}$ is a gpo, and $\underline{v}$ shadows the orbit of $x$.

*Proof that $\Psi^{p^u_n,p^s_n}_{x_n}\in\mathscr{A}$:* (A1), (A2) are clear, so we focus on (A3). It is enough to show that $1\le(p^u_n\wedge p^s_n)/\eta_n\le e$. The lower bound is by construction. For the upper bound, recall that $\limsup_{n\to\pm\infty}\eta_n>0$, so by the second subordination lemma $q^u_n=e^{-\varepsilon/3}Q_\varepsilon(f^n(x))$ for infinitely many $n<0$. By (b$_n$) with $i=0$, $q^u_n\ge e^{-\varepsilon}Q_\varepsilon(x_n)\ge e^{-\varepsilon}p^u_n$ for infinitely many $n<0$. If $q^u_n\ge e^{-\varepsilon}p^u_n$ then $q^u_{n+1}\ge e^{-\varepsilon}p^u_{n+1}$:

$$\begin{aligned}
q^u_{n+1}&=\min\{e^{\varepsilon}q^u_n,e^{-\varepsilon/3}Q_\varepsilon(f^{n+1}(x))\}\\
&\ge\min\{e^{\varepsilon}e^{-\varepsilon}p^u_n,e^{-2\varepsilon/3}Q_\varepsilon(x_{n+1})\},\ \text{by (b}_{n+1}\text{) with } i=0\\
&\ge e^{-\varepsilon}\min\{e^{\varepsilon}p^u_n,Q_\varepsilon(x_{n+1})\}=e^{-\varepsilon}p^u_{n+1}.
\end{aligned}$$

It follows that $q^u_n\ge e^{-\varepsilon}p^u_n$ for all $n\in\mathbb{Z}$. Similarly $q^s_n\ge e^{-\varepsilon}p^s_n$ for all $n\in\mathbb{Z}$, whence $\eta_n\ge e^{-\varepsilon}(p^u_n\wedge p^s_n)$ for all $n\in\mathbb{Z}$, giving us (A3).

*Proof that $\{\Psi^{p^u_n,p^s_n}_{x_n}\}_{n\in\mathbb{Z}}$ is a gpo.* (GPO2) is true by construction, so we just need to check (GPO1). We write (a$_n$) with $i=1$, and (a$_{n+1}$) with $i=0$:

- ○ $\mathrm{dist}_\Lambda(f^{n+1}(x),f(x_n))+\|\Theta\circ C_\chi(f^{n+1}(x))-\Theta\circ C_\chi(f(x_n))\|<e^{-8(m_n+3)}$.
- ○ $\mathrm{dist}_\Lambda(f^{n+1}(x),x_{n+1})+\|\Theta\circ C_\chi(f^{n+1}(x))-\Theta\circ C_\chi(x_{n+1})\|<e^{-8(m_{n+1}+3)}$.

So $x_{n+1},f(x_n),f^{n+1}(x)$ are all in the same canonical transverse disc, and

$$\mathrm{dist}_\Lambda(f(x_n),x_{n+1})+\|\Theta\circ C_\chi(f(x_n))-\Theta\circ C_\chi(x_{n+1})\|<e^{-8(m_n+3)}+e^{-8(m_{n+1}+3)}.\tag{4.2}$$

The proof of (A3) shows that $\xi_n := p^u_n \wedge p^s_n \in [e^{-m_n-2}, e^{-m_n+2}]$. Also $\xi_n/\xi_{n+1} \in [e^{-\varepsilon}, e^{\varepsilon}]$, because $\{(p^u_n, p^s_n)\}_{n\in\mathbb{Z}}$ is $\varepsilon$–subordinated (see Lemma 4.1). So the right hand side of (4.2) is less than $e^{-8}(1+e^{8\varepsilon})\xi^8_{n+1} < (p^u_{n+1}\wedge p^s_{n+1})^8$. Thus $\Psi^{p^u_{n+1}\wedge p^s_{n+1}}_{f(x_n)} \overset{\varepsilon}{\approx} \Psi^{p^u_{n+1}\wedge p^s_{n+1}}_{x_{n+1}}$. A similar argument with ($\mathrm{a}_n$) and $i=-1$, and with ($\mathrm{a}_{n-1}$) and $i=0$ shows that $\Psi^{p^u_{n-1}\wedge p^s_{n-1}}_{f^{-1}(x_n)} \overset{\varepsilon}{\approx} \Psi^{p^u_{n-1}\wedge p^s_{n-1}}_{x_{n-1}}$. So (GPO1) holds, and $\underline{v}$ is a gpo.

*Proof that $\underline{v}$ shadows $x$:* By ($\mathrm{a}_n$) with $i=0$, $\Psi^{p^u_n\wedge p^s_n}_{x_n} \overset{\varepsilon}{\approx} \Psi^{p^u_n\wedge p^s_n}_{f^n(x)}$ for all $n\in\mathbb{Z}$. By Proposition 3.5, $f^n(x) = \Psi_{f^n(x)}(\underline{0}) \in \Psi_{x_n}([-p^u_n\wedge p^s_n, p^u_n\wedge p^s_n]^2)$. So $\underline{v}$ shadows $x$.

Arranging relevance: Call an element $v\in\mathscr{A}$ *relevant*, if there is a gpo $\underline{v}\in\mathscr{A}^{\mathbb{Z}}$ s.t. $v_0 = v$ and $\underline{v}$ shadows a point in $\mathrm{NUH}_\chi(f)$. In this case every $v_i$ is relevant, because $\mathrm{NUH}_\chi(f)$ is $f$–invariant. So $\mathscr{A}' := \{v\in\mathscr{A} : v \text{ is relevant}\}$ is sufficient. It is discrete, because $\mathscr{A}'\subseteq\mathscr{A}$ and $\mathscr{A}$ is discrete. The theorem follows with $\mathscr{A}'$. $\square$

**The inverse shadowing problem.** The same orbit can be shadowed by many different gpos. The "inverse shadowing problem" is to control the set of gpos $\{\Psi^{p^u_i,p^s_i}_{x_i}\}_{i\in\mathbb{Z}}$ which shadow the orbit of a given point $x$. (GPO1) and (GPO2) were designed to make this possible. We need the following condition.

Regularity: Let $\mathscr{A}$ be as in Proposition 4.3. A gpo $\underline{v}\in\mathscr{A}^{\mathbb{Z}}$ is called *regular*, if $\{v_i\}_{i\geq 0}$, $\{v_i\}_{i\leq 0}$ have constant subsequences.

**Proposition 4.4.** *Almost every $x\in\Lambda$ is shadowed by a regular gpo in $\mathscr{A}^{\mathbb{Z}}$.*

*Proof.* We will show that this is the case for all $x\in\mathrm{NUH}^\#_\chi(f)$. Since $\mathscr{A}$ is sufficient (Proposition 4.3(2)), for every $x\in\mathrm{NUH}^\#_\chi(f)$ there is a gpo $\underline{v} = \{\Psi^{p^u_k,p^s_k}_{x_k}\}_{k\in\mathbb{Z}}\in\mathscr{A}^{\mathbb{Z}}$ which shadows $x$ s.t. for all $k\in\mathbb{Z}$, $\eta_k := p^u_k\wedge p^s_k \geq e^{-\varepsilon/3}q_\varepsilon(f^k(x))$.

Since $p^{u/s}\leq Q_\varepsilon(\cdot)\leq\varepsilon\,\mathrm{dist}_\Lambda(\cdot,\mathfrak{S})$,

$$\mathrm{dist}_\Lambda(x_k,\mathfrak{S}) \geq \varepsilon^{-1}e^{-\varepsilon/3}q_\varepsilon(f^k(x)) \text{ for all } k\in\mathbb{Z}. \tag{4.3}$$

Since $v_k \xrightarrow{\varepsilon} v_{k+1}$, $\Psi^{\eta_{k+1}}_{f(x_k)} \overset{\varepsilon}{\approx} \Psi^{\eta_{k+1}}_{x_{k+1}}$, whence $f(x_k)\in\Psi_{x_{k+1}}([-Q(x_{k+1}), Q(x_{k+1})]^2)$. Since $\mathrm{Lip}(\Psi_{x_{k+1}})\leq 2$, $\mathrm{dist}_\Lambda(f(x_k), x_{k+1}) \leq 2\sqrt{2}Q(x_{k+1}) \leq 3\varepsilon\,\mathrm{dist}_\Lambda(x_{k+1},\mathfrak{S})$. By the triangle inequality, relation (4.3), and the inequality $e^{-\varepsilon/3}\leq q_\varepsilon\circ f/q_\varepsilon\leq e^{\varepsilon/3}$,

$$\begin{aligned}\mathrm{dist}_\Lambda(f(x_k),\mathfrak{S}) &\geq \mathrm{dist}_\Lambda(x_{k+1},\mathfrak{S}) - \mathrm{dist}_\Lambda(f(x_k),x_{k+1}) \geq (1-3\varepsilon)\,\mathrm{dist}_\Lambda(x_{k+1},\mathfrak{S})\\ &\geq (1-3\varepsilon)\varepsilon^{-1}e^{-\varepsilon/3}q_\varepsilon(f^{k+1}(x)) > (1-3\varepsilon)\varepsilon^{-1}e^{-\varepsilon}q_\varepsilon(f^k(x)) > q_\varepsilon(f^k(x)),\end{aligned}$$

provided $\varepsilon$ is small enough. Similarly, $\mathrm{dist}_\Lambda(f^{-1}(x_k),\mathfrak{S}) > q_\varepsilon(f^k(x))$, and we obtain that $\min\{D(x_k), p^u_k, p^s_k\}\geq e^{-\varepsilon/3}q_\varepsilon(f^k(x))$ for all $k\in\mathbb{Z}$.

Since $x\in\mathrm{NUH}^\#_\chi(f)$, $\exists k_i,\ell_i\uparrow\infty$ and $c>0$ s.t. $q_\varepsilon(f^{-k_i}(x))\geq c$, $q_\varepsilon(f^{\ell_i}(x))\geq c$. Since $\mathscr{A}$ is discrete, there must be some constant subsequences $v_{-k_{i_j}}, v_{\ell_{i_j}}$. $\square$

The next theorem says, in a precise way, that if $\underline{u}$ is a regular gpo which shadows $x$, then $u_i$ is determined "up to bounded error". Together with the discreteness of $\mathscr{A}$, this implies that for every $i$ there are only finitely many choices for $u_i$.[5]

**Theorem 4.5.** *The following holds for all $\varepsilon$ small enough. Let $\underline{u},\underline{v}$ be regular gpos which shadow the orbit of the same point $x$. If $u_i = \Psi^{p^u_i,p^s_i}_{x_i}$ and $v_i = \Psi^{q^u_i,q^s_i}_{y_i}$, then:*
(1) $\mathrm{dist}_\Lambda(x_i,y_i) < 10^{-1}\max\{p^u_i\wedge p^s_i, q^u_i\wedge q^s_i\}$.

[5]But the set of all possible full sequences $\underline{u}$ can be uncountable.

(2) $\mathrm{dist}_\Lambda(f^k(x_i),\mathfrak{S})/\mathrm{dist}_\Lambda(f^k(y_i),\mathfrak{S})\in[e^{-\sqrt{\varepsilon}},e^{\sqrt{\varepsilon}}]$ *for* $k=0,1,-1$.
(3) $Q_\varepsilon(x_i)/Q_\varepsilon(y_i)\in[e^{-\sqrt[3]{\varepsilon}},e^{\sqrt[3]{\varepsilon}}]$.
(4) $(\Psi_{y_i}^{-1}\circ\Psi_{x_i})=(-1)^{\sigma_i}\mathrm{Id}+\underline{c}_i+\Delta_i$ *on* $[-\varepsilon,\varepsilon]^2$, *where* $\|\underline{c}_i\|<10^{-1}(q_i^u\wedge q_i^s)$, $\sigma_i\in\{0,1\}$ *are constants, and* $\Delta_i:[-\varepsilon,\varepsilon]^2\to\mathbb{R}^2$ *is a vector field s.t.* $\Delta_i(\underline{0})=\underline{0}$ *and* $\|(d\Delta_i)_{\underline{v}}\|<\sqrt[3]{\varepsilon}$ *on* $[-\varepsilon,\varepsilon]^2$.
(5) $p_i^u/q_i^u, p_i^s/q_i^s\in[e^{-\sqrt[3]{\varepsilon}},e^{\sqrt[3]{\varepsilon}}]$.

*Proof.* Denote the unstable and stable manifolds of $\underline{u}$ and $\underline{v}$ by $U^u,U^s$ and $V^u,V^s$. By the proof of the shadowing lemma, $U^u\cap U^s=V^u\cap V^s=\{x\}$.

PART (1). $U^{u/s}$ are admissible manifolds. By (Ad1–3), their intersection point must satisfy $x=\Psi_{x_0}(\underline{\xi})$ where $\|\underline{\xi}\|_\infty\le 10^{-2}(p_0^u\wedge p_0^s)$, see [Sar13, Prop. 4.11]. Since $\mathrm{Lip}(\Psi_{x_0})\le 2$, $\mathrm{dist}_\Lambda(x_0,x)\le 50^{-1}(p_0^u\wedge p_0^s)$. Similarly $\mathrm{dist}_\Lambda(y_0,x)\le 50^{-1}(q_0^u\wedge q_0^s)$, whence $\mathrm{dist}_\Lambda(x_0,y_0)\le 25^{-1}\max\{p_0^u\wedge p_0^s,q_0^u\wedge q_0^s\}$.

PART (2). In what follows $a=b\pm c$ means $b-c\le a\le b+c$.

$$\begin{aligned}\mathrm{dist}_\Lambda(x_0,\mathfrak{S})&=\mathrm{dist}_\Lambda(x,\mathfrak{S})\pm\mathrm{dist}_\Lambda(x,x_0)=\mathrm{dist}_\Lambda(x,\mathfrak{S})\pm 50^{-1}(p_0^u\wedge p_0^s),\ \text{by part 1}\\&=\mathrm{dist}_\Lambda(x,\mathfrak{S})\pm 50^{-1}Q_\varepsilon(x_0),\quad\text{because } p_0^u,p_0^s\le Q_\varepsilon(x_0)\\&=\mathrm{dist}_\Lambda(x,\mathfrak{S})\pm 50^{-1}\varepsilon\,\mathrm{dist}_\Lambda(x_0,\mathfrak{S}),\ \text{by the definition of } Q_\varepsilon(x_0).\end{aligned}$$

Therefore $\frac{\mathrm{dist}_\Lambda(x,\mathfrak{S})}{\mathrm{dist}_\Lambda(x_0,\mathfrak{S})}=1\pm 50^{-1}\varepsilon$. Similarly $\frac{\mathrm{dist}_\Lambda(x,\mathfrak{S})}{\mathrm{dist}_\Lambda(y_0,\mathfrak{S})}=1\pm 50^{-1}\varepsilon$. It follows that if $\varepsilon$ is small enough then $\frac{\mathrm{dist}_\Lambda(x_0,\mathfrak{S})}{\mathrm{dist}_\Lambda(y_0,\mathfrak{S})}\in[e^{-\varepsilon},e^{\varepsilon}]$. Applying this argument to suitable shifts of $\underline{u}$ and $\underline{v}$, we obtain $\frac{\mathrm{dist}_\Lambda(x_1,\mathfrak{S})}{\mathrm{dist}_\Lambda(y_1,\mathfrak{S})}\in[e^{-\varepsilon},e^{\varepsilon}]$ and $\frac{\mathrm{dist}_\Lambda(x_{-1},\mathfrak{S})}{\mathrm{dist}_\Lambda(y_{-1},\mathfrak{S})}\in[e^{-\varepsilon},e^{\varepsilon}]$.

Since $u_0\overset{\varepsilon}{\to}u_1$, $\Psi_{f(x_0)}^{p_1^u\wedge p_1^s}\overset{\varepsilon}{\approx}\Psi_{x_1}^{p_1^u\wedge p_1^s}$. So $x_1,f(x_0)\in\Psi_{x_1}\big([-Q_\varepsilon(x_1),Q_\varepsilon(x_1)]^2\big)$. Since $\mathrm{Lip}(\Psi_{x_1})\le 2$, $\mathrm{dist}_\Lambda\big(x_1,f(x_0)\big)\le 2\sqrt{2}Q_\varepsilon(x_1)<6\varepsilon\,\mathrm{dist}_\Lambda(x_1,\mathfrak{S})$. Thus

$$\begin{aligned}\mathrm{dist}_\Lambda(f(x_0),\mathfrak{S})&=\mathrm{dist}_\Lambda(x_1,\mathfrak{S})\pm\mathrm{dist}_\Lambda(x_1,f(x_0))=\mathrm{dist}_\Lambda(x_1,\mathfrak{S})\pm 6\varepsilon\,\mathrm{dist}_\Lambda(x_1,\mathfrak{S})\\&=e^{\pm 7\varepsilon}\,\mathrm{dist}_\Lambda(x_1,\mathfrak{S}),\ \text{provided } \varepsilon \text{ is small enough.}\end{aligned}$$

Similarly $\mathrm{dist}_\Lambda(f(y_0),\mathfrak{S})=e^{\pm 7\varepsilon}\,\mathrm{dist}_\Lambda(y_1,\mathfrak{S})$. Since $\frac{\mathrm{dist}_\Lambda(x_1,\mathfrak{S})}{\mathrm{dist}_\Lambda(y_1,\mathfrak{S})}\in[e^{-\varepsilon},e^{\varepsilon}]$,

$$\frac{\mathrm{dist}_\Lambda(f(x_0),\mathfrak{S})}{\mathrm{dist}_\Lambda(f(y_0),\mathfrak{S})}\in[e^{-15\varepsilon},e^{15\varepsilon}]\subset[e^{-\sqrt{\varepsilon}},e^{\sqrt{\varepsilon}}],\ \text{provided } \varepsilon \text{ is small enough.}$$

Similarly $\frac{\mathrm{dist}_\Lambda(f^{-1}(x_0),\mathfrak{S})}{\mathrm{dist}_\Lambda(f^{-1}(y_0),\mathfrak{S})}\in[e^{-\sqrt{\varepsilon}},e^{\sqrt{\varepsilon}}]$.

PART (3). One shows as in [Sar13, §6 and §7] that for all $\varepsilon$ small enough,

$$\frac{\sin\alpha(x_i)}{\sin\alpha(y_i)}\in[e^{-\sqrt{\varepsilon}},e^{\sqrt{\varepsilon}}]\quad\text{and}\quad\frac{s(x_i)}{s(y_i)},\frac{u(x_i)}{u(y_i)}\in[e^{-4\sqrt{\varepsilon}},e^{4\sqrt{\varepsilon}}].\tag{4.4}$$

The proof carries over without change, because all the calculations are done on $f^n(U^s),f^n(V^s)$ $(n\ge 0)$ and $f^n(U^u),f^n(V^u)$ $(n\le 0)$, and these sets stay inside Pesin charts, away from $\mathfrak{S}$. By (4.4), for all $\varepsilon$ small enough we have

$$\left(\frac{\sqrt{s(x_0)^2+u(x_0)^2}}{|\sin\alpha(x_0)|}\right)^{-\frac{12}{\beta}}\left(\frac{\sqrt{s(y_0)^2+u(y_0)^2}}{|\sin\alpha(y_0)|}\right)^{\frac{12}{\beta}}\in[e^{-\sqrt[3]{\varepsilon}},e^{-\sqrt[3]{\varepsilon}}].$$

By part (2) and the definition of $Q_\varepsilon$, $\frac{Q_\varepsilon(x_0)}{Q_\varepsilon(y_0)}\in[e^{-\sqrt[3]{\varepsilon}},e^{\sqrt[3]{\varepsilon}}]$.

PART (4). This is done exactly as in [Sar13, §9], except that one needs to add the constraint $\varepsilon<\rho_{\mathrm{dom}}$ to be able to use the smoothness of $p\mapsto\mathrm{Exp}_p$ on $\Lambda$.

PART (5). This is done exactly as in the proof of [Sar13, Prop. 8.3], except that step 1 there should be replaced by part (3) here. $\square$

*Remark.* Regularity is needed in parts (3), (4), (5). Parts (3), (4) also use the full force of (Ad1–3), and Part (5) is based on (GPO2). See [Sar13].

## 5. Countable Markov partitions and symbolic dynamics

Sinaĭ and Bowen gave several methods for constructing Markov partitions for uniformly hyperbolic diffeomorphisms [Sin68a, Sin68b, Bow70, Bow75]. One of these constructions, due to Bowen, uses pseudo-orbits and shadowing [Bow75]. The theory of gpos we developed in the previous section allows us to apply this method to adapted Poincaré sections. The result is a Markov partition for $f:\Lambda\to\Lambda$. It is then a standard procedure to code $f:\Lambda\to\Lambda$ by a topological Markov shift, and $\varphi: M\to M$ by a topological Markov flow.

**Step 1: A Markov extension.** Let $\mathscr{A}$ be the countable set of double charts we constructed in Proposition 4.3, and let $\mathscr{G}$ denote the countable directed graph with set of vertices $\mathscr{A}$ and set of edges $\{(v,w)\in\mathscr{A}\times\mathscr{A} : v\overset{\varepsilon}{\to} w\}$.

**Lemma 5.1.** *Every vertex of $\mathscr{G}$ has finite ingoing degree, and finite outgoing degree.*

*Proof.* We fix $v\in\mathscr{A}$, and bound the number of $w$ s.t. $v\overset{\varepsilon}{\to} w$, using the discreteness and relevance of $\mathscr{A}$ (cf. Prop. 4.3).

By the relevance property, $v\overset{\varepsilon}{\to} w$ extends to a path $v\overset{\varepsilon}{\to} w\overset{\varepsilon}{\to} u$. Write $v=\Psi_x^{p^u,p^s}$, $w=\Psi_y^{q^u,q^s}$, $u=\Psi_z^{r^u,r^s}$, then:

- $\circ$ $\frac{r^u\wedge r^s}{q^u\wedge q^s},\frac{q^u\wedge q^s}{p^u\wedge p^s}\in[e^{-\varepsilon},e^{\varepsilon}]$, by Lemma 4.1.
- $\circ$ $\mathrm{dist}_\Lambda(y,\mathfrak{S})\geq\varepsilon^{-1}e^{-\varepsilon}(p^u\wedge p^s)$, because $\mathrm{dist}_\Lambda(y,\mathfrak{S})\geq\varepsilon^{-1}Q_\varepsilon(y)$, $Q_\varepsilon(y)\geq q^u\wedge q^s$.
- $\circ$ $\mathrm{dist}_\Lambda(f(y),\mathfrak{S})\geq e^{-2\varepsilon}(\varepsilon^{-1}-3)(p^u\wedge p^s)$, because

$$
\begin{aligned}
&\mathrm{dist}_\Lambda(f(y),\mathfrak{S})\geq\mathrm{dist}_\Lambda(z,\mathfrak{S})-\mathrm{dist}_\Lambda(z,f(y))\\
&\geq\varepsilon^{-1}Q_\varepsilon(z)-2\sqrt{2}Q_\varepsilon(z)\quad\because f(y)\in\Psi_z([-Q_\varepsilon(z),Q_\varepsilon(z)]^2)\text{ and }\mathrm{Lip}(\Psi_z)\leq 2\\
&\geq(\varepsilon^{-1}-3)(r^u\wedge r^s)\geq e^{-2\varepsilon}(\varepsilon^{-1}-3)(p^u\wedge p^s).
\end{aligned}
$$

- $\circ$ $\mathrm{dist}_\Lambda(f^{-1}(y),\mathfrak{S})\geq e^{-2\varepsilon}(\varepsilon^{-1}-3)(p^u\wedge p^s)$, for similar reasons.

So $D(y):=\mathrm{dist}_\Lambda(\{y,f(y),f^{-1}(y)\},\mathfrak{S})\geq t:=e^{-2\varepsilon}(\varepsilon^{-1}-3)(p^u\wedge p^s)$.

By the discreteness of $\mathscr{A}$ (and assuming $\varepsilon<\frac{1}{3}$), $\#\{w\in\mathscr{A}:v\overset{\varepsilon}{\to} w\}<\infty$. The finiteness of the ingoing degree is proved in the same way. $\square$

THE MARKOV EXTENSION: Let $\Sigma(\mathscr{G})$ denote the set of two-sided paths on $\mathscr{G}$:

$$\Sigma(\mathscr{G}):=\{\underline{v}\in\mathscr{A}^{\mathbb{Z}}:v_i\overset{\varepsilon}{\to} v_{i+1}\text{ for all }i\in\mathbb{Z}\}.$$

We equip $\Sigma(\mathscr{G})$ with the metric $d(\underline{u},\underline{v})=\exp[-\min\{|n|:u_n\neq v_n\}]$, and with the action of the *left shift map* $\sigma:\Sigma(\mathscr{G})\to\Sigma(\mathscr{G})$, $\sigma:\{v_i\}_{i\in\mathbb{Z}}\mapsto\{v_{i+1}\}_{i\in\mathbb{Z}}$. The set $\Sigma(\mathscr{G})$ is exactly the collection of gpos in $\mathscr{A}^{\mathbb{Z}}$, hence $\pi:\Sigma(\mathscr{G})\to\Lambda$ given by

$$\pi(\underline{v}):=\text{unique point whose }f\text{–orbit is shadowed by }\underline{v}$$

is well-defined. Necessarily $f\circ\pi=\pi\circ\sigma$, so $\sigma:\Sigma(\mathscr{G})\to\Sigma(\mathscr{G})$ is an extension of $f:\Lambda\to\Lambda$ (at least on a subset of full measure, by Prop. 4.4).

It is easy to see, using the finite degree of the vertices of $\mathscr{G}$, that $(\Sigma(\mathscr{G}), d)$ is a locally compact, complete and separable metric space. The left shift map is a bi-Lipschitz homeomorphism. The subset of *regular gpos*

$$\Sigma^{\#}(\mathscr{G}) := \{\underline{v} \in \Sigma(\mathscr{G}) : \{v_i\}_{i\leq 0}, \{v_i\}_{i\geq 0} \text{ contain constant subsequences}\}$$

has full measure with respect to every $\sigma$–invariant Borel probability measure.

As we saw in the proof of the shadowing lemma (Theorem 4.2), $\pi(\underline{v})$ is the unique intersection of $V^u(\underline{v}^-)$ and $V^s(\underline{v}^+)$ where $\underline{v}^{\pm}$ are the half gpos determined by $\underline{v}$. The proof shows that the following holds for all $\varepsilon$ small enough:

(1) Hölder continuity: $\pi$ is Hölder continuous (because $\mathscr{F}^u, \mathscr{F}^s$ are contractions, see [Sar13, Thm 4.16(2)]).
(2) Almost surjectivity: $\mu_\Lambda(\Lambda \setminus \pi[\Sigma^{\#}(\mathscr{G})]) = 0$ (Proposition 4.4).
(3) Inverse Property: for all $x \in \Lambda, i \in \mathbb{Z}$, $\#\{v_i : \underline{v} \in \Sigma^{\#}(\mathscr{G}), \pi(\underline{v}) = x\} < \infty$. (Theorem 4.5 and the discreteness of $\mathscr{A}$.)

The inverse property does *not* imply that $\pi$ is finite-to-one or even countable-to-one. The following steps will lead us to an a.e. *finite-to-one* Markov extension.

**Step 2: A Markov cover.** Given $v \in \mathscr{A}$, let ${}_0[v] := \{\underline{v} \in \Sigma(\mathscr{G}) : v_0 = v\}$. This is a partition of $\Sigma(\mathscr{G})$. The projection to $\Lambda$,

$$\mathscr{Z} := \{Z(v) : v \in \mathscr{A}\}, \text{ where } Z(v) := \{\pi(\underline{v}) : \underline{v} \in \Sigma^{\#}(\mathscr{G}),\ v_0 = v\},$$

is not a partition. It could even be the case that $Z(v) = Z(w)$ for $v \neq w$ (in this case, we agree to think of $Z(v), Z(w)$ as different elements of $\mathscr{Z}$). Here are some important properties of $\mathscr{Z}$.

Covering property: *$\mathscr{Z}$ covers a set of full $\mu_\Lambda$–measure.*

*Proof.* $\mathscr{Z}$ covers $\mathrm{NUH}^{\#}_\chi(f)$.

Local finiteness: *For all $Z \in \mathscr{Z}$, $\#\{Z' \in \mathscr{Z} : Z' \cap Z \neq \varnothing\} < \infty$. Even better: $\#\{v \in \mathscr{A} : Z(v) \cap Z \neq \varnothing\} < \infty$ for all $Z \in \mathscr{Z}$.*

*Proof.* Write $Z = Z(\Psi_x^{p^u,p^s})$. If $Z(\Psi_y^{q^u,q^s}) \cap Z \neq \varnothing$, then $q^u \wedge q^s \geq e^{-\sqrt[3]{\varepsilon}}(p^u \wedge p^s)$ and $\mathrm{dist}_\Lambda(\{y, f(y), f^{-1}(y)\}, \mathfrak{S}) \geq e^{-\sqrt{\varepsilon}}\,\mathrm{dist}_\Lambda(\{x, f(x), f^{-1}(x)\}, \mathfrak{S})$ (Theorem 4.5). Since $\mathscr{A}$ is discrete, there are only finitely many such $\Psi_y^{q^u,q^s}$ in $\mathscr{A}$.

Product structure: *Suppose $v \in \mathscr{A}$ and $Z = Z(v)$. For every $x \in Z$ there are sets $W^u(x,Z)$ and $W^s(x,Z)$ called the* $s$–fibre *and* $u$–fibre *of* $x$ in $Z$ *s.t.:*

(1) *$Z = \bigcup_{x\in Z} W^u(x,Z)$, $Z = \bigcup_{x\in Z} W^s(x,Z)$.*
(2) *Any two $s$–fibres in $Z$ are either equal or disjoint, and the same for $u$–fibres.*
(3) *For every $x, y \in Z$, $W^u(x,Z) \cap W^s(y,Z)$ consists of a single point.*

*Notation: $[x,y]_Z :=$ unique point in $W^u(x,Z) \cap W^s(x,Z)$.*

*Proof.* Recall from §4 the notation for the stable and unstable manifolds of positive and negative gpos. Fix $Z = Z(v)$ in $\mathscr{Z}$, $x \in Z$, and let:

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

To see that the definition is proper, suppose $\underline{u}, \underline{v}$ are two regular gpos such that $u_0 = v_0 = v$. If $V^t[\underline{u}], V^t[\underline{v}]$ intersect at some point $z$ for $t = s$ or $u$, then $V^t[\underline{u}] = V^t[\underline{v}]$, because both are equal to the piece of the local stable/unstable manifold of $z$ at $\Psi_{x(v)}([-p^t(v), p^t(v)]^2)$. See [Sar13, Prop. 6.4] for details. In particular, $\pi(\underline{u}) = \pi(\underline{v}) \Rightarrow V^t[\underline{u}] = V^t[\underline{v}]$ for $t = u, s$.

This argument also shows that any two $t$–fibres ($t = s$ or $u$) are equal or disjoint, hence (2) holds. (1) is because $W^u(x, Z)$, $W^s(x, Z)$ both contain $x$. For (3), write $W^u(x, Z) = V^u[\underline{u}] \cap Z$ and $W^s(y, Z) = V^s[\underline{v}] \cap Z$ where $\underline{u}, \underline{v} \in \Sigma^\#(\mathscr{G})$ satisfy $u_0 = v_0 = v$. Let $\underline{w} := (\ldots, u_{-2}, u_{-1}, \dot{v}, v_1, v_2, \ldots)$ with the dot indicating the zeroth coordinate. Clearly $\pi(\underline{w}) \in W^u(x, Z) \cap W^s(y, Z)$. Since $W^u(x, Z) \cap W^s(y, Z) \subset V^u(x, Z) \cap V^s(y, Z)$ and a $u$–admissible manifold intersects an $s$–admissible at most once [KH95, Cor. S.3.8], [Sar13, Prop. 4.11], $W^u(x, Z) \cap W^s(y, Z) = \{\pi(\underline{w})\}$.

Symbolic Markov property: *If $x = \pi(\underline{v})$ with $\underline{v} \in \Sigma^\#(\mathscr{G})$, then*

$$f\big[W^s(x, Z(v_0))\big] \subset W^s(f(x), Z(v_1)) \text{ and } f^{-1}\big[W^u(f(x), Z(v_1))\big] \subset W^u(x, Z(v_0)).$$

*Proof.* Fix $y \in W^s(x, Z(v_0))$. Choose $\underline{u} \in \Sigma^\#(\mathscr{G})$ s.t. $u_0 = v_0$ and $y = \pi(\underline{u})$. Write $u_i = \Psi_{y_i}^{q_i^u, q_i^s}$ and $\eta_i := q_i^u \wedge q_i^s$, then $f^{-n}(y) \in \Psi_{y_{-n}}([-\eta_{-n}, \eta_{-n}]^2)$ for all $n \geq 0$. Write $v_i = \Psi_{x_i}^{p_i^u, p_i^s}$ and $\xi_i := p_i^u \wedge p_i^s$. Since $y \in W^s(x, Z(v_0)) \subset V^s[\underline{v}^+]$, $f^n(y) \in f^n(V^s[\underline{v}^+]) \subset V^s[\sigma^n \underline{v}^+] \subset \Psi_{x_n}([-\xi_n, \xi_n]^2)$ for all $n \geq 0$. It follows that the gpo $\underline{w} = (\ldots, u_{-2}, u_{-1}, \dot{v}, v_1, v_2, \ldots)$ shadows $y$ (the dot indicates the position of the zeroth coordinate). Necessarily, $f(y) = f[\pi(\underline{w})] = \pi[\sigma(\underline{w})] \in V^s[\{v_i\}_{i \geq 1}] \cap Z(v_1) = W^s(f(x), Z(v_1))$. Thus $f(y) \in W^s(f(x), Z(v_1))$.

Since $y \in W^s(x, Z(v_0))$ was arbitrary, $f[W^s(x, Z(v_0))] \subset W^s(f(x), Z(v_1))$. The other inequality is symmetric.

Overlapping charts property: *The following holds for all $\varepsilon$ small enough. Suppose $Z, Z' \in \mathscr{Z}$ and $Z \cap Z' \neq \varnothing$.*

(1) *If $Z = Z(\Psi_{x_0}^{p_0^u, p_0^s}), Z' = Z(\Psi_{y_0}^{q_0^u, q_0^s})$, then $Z \subset \Psi_{y_0}([-(q_0^u \wedge q_0^s), (q_0^u \wedge q_0^s)]^2)$.*
(2) *For all $x \in Z, y \in Z'$, $V^u(x, Z)$ intersects $V^s(y, Z')$ at a unique point.*
(3) *For any $x \in Z \cap Z'$, $W^u(x, Z) \subset V^u(x, Z')$ and $W^s(x, Z) \subset V^s(x, Z')$.*

*Sketch of proof.* If $Z \cap Z' \neq \varnothing$, then there are $\underline{u}, \underline{v} \in \Sigma^\#(\mathscr{G})$ s.t. $u_0 = \Psi_{x_0}^{p_0^u, p_0^s}$, $v_0 = \Psi_{y_0}^{q_0^u, q_0^s}$, and $\pi(\underline{u}) = \pi(\underline{v})$. By Theorem 4.5(4), $\Psi_{y_0}^{-1} \circ \Psi_{x_0}$ is close to $\pm\mathrm{Id}$. This is enough to prove (1)–(3), see Lemmas 10.8 and 10.10 in [Sar13] for details.

**Step 3 (Bowen, Sinaĭ): A countable Markov partition.** We refine $\mathscr{Z}$ into a partition without destroying the Markov property or the product structure. The refinement procedure we use below is due to Bowen [Bow75], building on earlier work of Sinaĭ [Sin68a, Sin68b]. It was designed for finite Markov covers, but works equally well for locally finite infinite covers. Local finiteness is essential: a general non-locally finite cover may not have a countable refining partition as can be seen in the example of the cover $\{(\alpha, \beta) : \alpha, \beta \in \mathbb{Q}\}$ of $\mathbb{R}$.

Enumerate $\mathscr{Z} = \{Z_i : i \in \mathbb{N}\}$. For every $Z_i, Z_j \in \mathscr{Z}$ s.t. $Z_i \cap Z_j \neq \varnothing$, let

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

This is a partition of $Z_i$. Let $\mathscr{T} := \left\{T_{ij}^{\alpha\beta} : i, j \in \mathbb{N}, \alpha \in \{u, \varnothing\}, \beta \in \{s, \varnothing\}\right\}$. This is a countable set, and $\mathscr{T} \supset \mathscr{Z}$ (since $T_{ii}^{us} = Z_i, \forall i$). Necessarily, $\mathscr{T}$ covers $\mathrm{NUH}_\chi^\#(f)$.

THE MARKOV PARTITION: $\mathscr{R}$:= the collection of sets which can be put in the form $R(x) := \bigcap\{T \in \mathscr{T} : T \ni x\}$ for some $x \in \bigcup_{i\geq 1} Z_i$.

**Proposition 5.2.** *$\mathscr{R}$ is a countable pairwise disjoint cover of $\mathrm{NUH}_\chi^\#(f)$. It refines $\mathscr{Z}$, and every element of $\mathscr{Z}$ contains only finitely many elements of $\mathscr{R}$.*

*Proof.* See [Bow75] or [Sar13, Prop. 11.2]. The local finiteness of $\mathscr{Z}$ is needed to show that $\mathscr{R}$ is countable: it implies that $\#\{T \in \mathscr{T} : T \ni x\} < \infty$ for all $x$. □

The following proposition says that $\mathscr{R}$ is a *Markov partition* in the sense of Sinaĭ [Sin68b]. First, some definitions. The *u–fibre* and *s–fibre* of $x \in R \in \mathscr{R}$ are

$$\begin{aligned}
W^u(x, R) &:= \bigcap\{W^u(x, Z_i) \cap T_{ij}^{\alpha\beta} : T_{ij}^{\alpha\beta} \in \mathscr{T} \text{ contains } R\}, \\
W^s(x, R) &:= \bigcap\{W^s(x, Z_i) \cap T_{ij}^{\alpha\beta} : T_{ij}^{\alpha\beta} \in \mathscr{T} \text{ contains } R\}.
\end{aligned}$$

**Proposition 5.3.** *The following properties hold.*

(1) PRODUCT STRUCTURE: *Suppose $R \in \mathscr{R}$.*
   (a) *If $x \in R$, then the $s$ and $u$ fibres of $x$ contain $x$, and are contained in $R$, therefore $R = \bigcup_{x\in R} W^u(x, R)$ and $R = \bigcup_{x\in R} W^s(x, R)$.*
   (b) *For all $x, y \in R$, either the $u$–fibres of $x, y$ in $R$ are equal, or they are disjoint. Similarly for $s$–fibres.*
   (c) *For all $x, y \in R$, $W^u(x, R)$ and $W^s(y, R)$ intersect at a unique point, denoted by $[x, y]$ and called the* Smale bracket *of $x, y$.*

(2) HYPERBOLICITY: *For all $z_1, z_2 \in W^s(x, R)$, $\mathrm{dist}_\Lambda(f^n(z_1), f^n(z_2)) \xrightarrow[n\to\infty]{} 0$, and for all $z_1, z_2 \in W^u(x, R)$, $\mathrm{dist}_\Lambda(f^{-n}(z_1), f^{-n}(z_2)) \xrightarrow[n\to\infty]{} 0$. The rates are exponential.*

(3) MARKOV PROPERTY: *Let $R_0, R_1 \in \mathscr{R}$. 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.* This follows from the Markov properties of $\mathscr{Z}$ as in [Bow75]. See [Sar13, Prop. 11.5–11.7] for a proof using the notation of this paper. □

**Step 4: Symbolic coding for $f : \Lambda \to \Lambda$ [AW67, Sin68b].** Let $\mathscr{R}$ denote the partition we constructed in the previous section. Suppose $R, S \in \mathscr{R}$. We say that $R$ *connects to* $S$, and write $R \to S$, whenever $\exists x \in R$ s.t. $f(x) \in S$. Equivalently, $R \to S$ iff $R \cap f^{-1}(S) \neq \varnothing$.

THE DYNAMICAL GRAPH OF $\mathscr{R}$: This is the directed graph $\widehat{\mathscr{G}}$ with set of vertices $\mathscr{R}$ and set of edges $\{(R, S) \in \mathscr{R} \times \mathscr{R} : R \to S\}$.

FUNDAMENTAL OBSERVATION [AW67, Sin68b]: Suppose $m \leq n$ are integers, and $R_m \to \cdots \to R_n$ is a finite path on $\widehat{\mathscr{G}}$, then

$$_\ell[R_m, \ldots, R_n] := f^{-\ell}(R_m) \cap f^{-\ell-1}(R_{m+1}) \cap \cdots \cap f^{-\ell-(n-m)}(R_n) \neq \varnothing.$$

*Proof.* This can be seen by induction on $n-m$ as follows: If $n-m=0$ or 1 there is nothing to prove. Assume by induction that the statement holds for $m-n$, then $\exists x \in {}_\ell[R_m, \ldots, R_n]$ and $\exists y \in R_n$ s.t. $f(y) \in R_{n+1}$. Let $z := [f^n(x), y]$, then $f^{-n}(z) \in {}_\ell[R_m, \ldots, R_{n+1}]$ by the Markov property. □

The sets ${}_\ell[R_m, \ldots, R_n]$ can be related to cylinders in $\Sigma^\#(\mathscr{G})$ as follows. Define for every path $v_m \to \cdots \to v_n$ on $\mathscr{G}$ (not $\widehat{\mathscr{G}}$) the set

$$Z_\ell(v_m, \ldots, v_n) := \{\pi(\underline{u}) : \underline{u} \in \Sigma^\#(\mathscr{G}), u_i = v_i \text{ for } i = \ell, \ldots, \ell+n-m\}.$$

**Lemma 5.4.** *For all doubly infinite path $\cdots \to R_0 \to R_1 \to \cdots$ on $\widehat{\mathscr{G}}$ there is a gpo $\underline{v} \in \Sigma(\mathscr{G})$ s.t. for every $n$, $R_n \subset Z(v_n)$ and ${}_{-n}[R_{-n}, \ldots, R_n] \subset Z_{-n}(R_{-n}, \ldots, R_n)$.*

The proof proceeds as follows: for each $n \geq 0$ take $x_n \in_{-n} [R_{-n}, \ldots, R_n]$, write $x_n = \pi(\underline{v}^{(n)})$ for $\underline{v}^{(n)} \in \Sigma^\#(\mathscr{G})$, and then apply a diagonal argument to construct $\underline{v}$. See [Sar13, Lemma 12.2] for the details.

**Proposition 5.5.** *Every vertex of $\widehat{\mathscr{G}}$ has finite outgoing and ingoing degrees.*

*Proof.* Fix $R_0 \in \mathscr{R}$. For every path $R_{-1} \to R_0 \to R_1$ in $\widehat{\mathscr{G}}$, find a path $v_{-1} \to v_0 \to v_1$ in $\mathscr{G}$ s.t. $Z(v_i) \supset R_i$ for $|i| \leq 1$. Since $\mathscr{Z}$ is locally finite, there are finitely many possibilities for $v_0$. Since every vertex of $\mathscr{G}$ has finite degree, there are also only finitely many possibilities for $v_{-1}, v_1$. Since every element in $\mathscr{Z}$ contains at most finitely many elements in $\mathscr{R}$, there is a finite number of possibilities for $R_{-1}, R_1$. □

Let $\Sigma(\widehat{\mathscr{G}}) := \{\text{doubly infinite paths on } \widehat{\mathscr{G}}\} = \{\underline{R} \in \mathscr{R}^{\mathbb{Z}} : R_i \to R_{i+1}, \forall i\}$, equipped with the metric $d(\underline{R}, \underline{S}) := \exp[-\min\{|i| : R_i \neq S_i\}]$, and the action of the left shift map $\sigma : \Sigma(\widehat{\mathscr{G}}) \to \Sigma(\widehat{\mathscr{G}})$, $\sigma(\underline{R})_i = R_{i+1}$. Let

$$\Sigma^\#(\widehat{\mathscr{G}}) := \left\{\underline{R} \in \Sigma(\widehat{\mathscr{G}}) : \{R_i\}_{i \leq 0}, \{R_i\}_{i \geq 0} \text{ contain constant subsequences}\right\}.$$

Since $\pi : \Sigma(\mathscr{G}) \to \Lambda$ is Hölder continuous, there are constants $C > 0, \theta \in (0,1)$ s.t. for every finite path $v_{-n} \to \cdots \to v_n$ on $\mathscr{G}$, $\operatorname{diam}(Z_{-n}(v_{-n}, \ldots, v_n)) \leq C\theta^n$. By Lemma 5.4, $\operatorname{diam}({}_{-n}[R_{-n}, \ldots, R_n]) \leq C\theta^n$ for every finite path $R_{-n} \to \cdots \to R_n$ on $\widehat{\mathscr{G}}$. This allows us to make the following definition.

SYMBOLIC DYNAMICS FOR $f$: Let $\widehat{\pi} : \Sigma(\widehat{\mathscr{G}}) \to \Lambda$ be defined by

$$\widehat{\pi}(\underline{R}) := \text{ The unique point in } \bigcap_{n=0}^{\infty} \overline{{}_{-n}[R_{-n}, \ldots, R_n]}.$$

**Theorem 5.6.** *The following holds for all $\varepsilon$ small enough.*

(1) $\widehat{\pi} \circ \sigma = f \circ \widehat{\pi}$.
(2) $\widehat{\pi} : \Sigma(\widehat{\mathscr{G}}) \to \Lambda$ *is Hölder continuous.*
(3) $\widehat{\pi}[\Sigma^\#(\widehat{\mathscr{G}})]$ *has full $\mu_\Lambda$–measure.*
(4) *Every $x \in \widehat{\pi}[\Sigma^\#(\widehat{\mathscr{G}})]$ has finitely many pre-images in $\Sigma^\#(\widehat{\mathscr{G}})$. If $\mu$ is ergodic, this number is equal a.e. to a constant.*
(5) *Moreover, there exists $N : \mathscr{R} \times \mathscr{R} \to \mathbb{N}$ s.t. if $x = \widehat{\pi}(\underline{R})$ where $R_i = R$ for infinitely many $i < 0$ and $R_i = S$ for infinitely many $i > 0$, then $\#\{\underline{S} \in \Sigma^\#(\widehat{\mathscr{G}}) : \pi(\underline{S}) = x\} \leq N(R, S)$.*

*Proof.* (1) If $\underline{R} \in \Sigma(\widehat{\mathscr{G}})$, then $\widehat{\pi}[\sigma(\underline{R})] = f[\widehat{\pi}(\underline{R})]$:

$$\begin{aligned}\{\widehat{\pi}[\sigma(\underline{R})]\} &= \bigcap_{n\geq 0} \overline{{}_{-n}[R_{-n+1},\ldots,R_{n+1}]} \supseteq \bigcap_{n\geq 0} \overline{{}_{-(n+2)}[R_{-n-1},\ldots,R_{n+1}]} \\ &\overset{!}{=} \bigcap_{n\geq 0} \overline{f\big({}_{-n-1}[R_{-n-1},\ldots,R_{n+1}]\big)} \overset{!}{\supset} \bigcap_{n\geq 0} f\big(\overline{{}_{-n-1}[R_{-n-1},\ldots,R_{n+1}]}\big) \\ &\overset{!}{=} f\left(\bigcap_{n\geq 0} \overline{{}_{-n-1}[R_{-n-1},\ldots,R_{n+1}]}\right) = \{f[\widehat{\pi}(\underline{R})]\}.\end{aligned}$$

The equalities $\overset{!}{=}$ are because $f$ is invertible. To justify $\overset{!}{\supset}$, it is enough to show that $f$ is continuous on an open neighborhood of $C := \overline{{}_{-n-1}[R_{-n-1},\ldots,R_{n+1}]}$. Fix some $v_0 = \Psi_{x_0}^{p_0^u,p_0^s}$ s.t. $Z(v_0) \supset R_0$, then $C \subset \overline{R_0} \subset \overline{Z(v_0)} \subset \Psi_{x_0}([-Q_\varepsilon(x_0), Q_\varepsilon(x_0)]^2) \subset \Lambda \setminus \mathfrak{S}$. So $f$ is continuous on $C$.

(2) is because of the inequality $\mathrm{diam}({}_{-n}[R_{-n},\ldots,R_n]) \leq C\theta^n$ mentioned above.

(3) is because for every $x \in \mathrm{NUH}_\chi^\#(f)$, $x = \widehat{\pi}(\underline{R})$ where $R_i :=$ unique element of $\mathscr{R}$ which contains $f^i(x)$. Clearly $\underline{R} \in \Sigma(\widehat{\mathscr{G}})$. To see that $\underline{R} \in \Sigma^\#(\widehat{\mathscr{G}})$, we use Lemma 5.4 to construct a regular gpo $\underline{v} \in \Sigma^\#(\mathscr{G})$ s.t. $x = \pi(\underline{v})$, with $R_i \subset Z(v_i)$. Every $Z(v)$ contains at most finitely many elements of $\mathscr{R}$ (Proposition 5.2). Therefore, the regularity of $\underline{v}$ implies the regularity of $\underline{R}$.

(4) follows from (5) and the $f$–invariance of $x \mapsto \#\{\underline{R} \in \Sigma^\#(\widehat{\mathscr{G}}) : \widehat{\pi}(\underline{R}) = x\}$.

(5) is proved using Bowen's method [Bow78, pp. 13–14], see also [PP90, p. 229]. The proof is the same as in [Sar13], but since the presentation there has an error, we decided to include the complete details in the appendix. □

The next lemma is used in [LLS16]. Recall from Lemma 2.6 that there is a set $\Lambda_\chi^*$ of full $\mu_\Lambda$–measure s.t. every $x \in \Lambda_\chi^*$ has tangent unit vectors $\vec{v}_x^s, \vec{v}_x^u \in T_x\Lambda$ s.t. $\lim_{n\to\infty} \frac{1}{n}\log\|df_x^n \vec{v}_x^s\|_{f^n(x)} < -\chi$ and $\lim_{n\to\infty} \frac{1}{n}\log\|df_x^n \vec{v}_x^u\|_{f^n(x)} > \chi$. The maps $x \in \Lambda_\chi^* \mapsto \vec{v}_x^s, \vec{v}_x^u$ are not necessarily Hölder continuous with respect to the Riemannian metric (they may not even be globally defined). But the symbolic metric is so much stronger than the Riemannian metric that the following holds.

**Lemma 5.7.** *The maps $\underline{R} \in \Sigma(\widehat{\mathscr{G}}) \mapsto \vec{v}^s_{\widehat{\pi}(\underline{R})}, \vec{v}^u_{\widehat{\pi}(\underline{R})}$ are Hölder continuous with respect to the symbolic metric.*

Lemma 5.7 is a version of [Sar13, Prop. 12.6] in our setup, and is proved similarly.

**Step 5: Symbolic coding for $\varphi : M \to M$.** Let $\widehat{\pi} : \Sigma(\widehat{\mathscr{G}}) \to \Lambda$ be the symbolic coding for $f : \Lambda \to \Lambda$ given by Theorem 5.6. Recall that $R : \Lambda \to (0,\infty)$ denotes the roof function of $\Lambda$. By the choice of $\Lambda$, $R$ is bounded away from zero and infinity, and there is a global constant $\mathfrak{C}$ s.t. $\sup_{x\in\Lambda\setminus\mathfrak{S}} \|dR_x\| < \mathfrak{C}$, see Lemma 2.5. Let

$$r : \Sigma(\widehat{\mathscr{G}}) \to (0,\infty),\ r := R \circ \widehat{\pi}.$$

This function is also bounded away from zero and infinity, and since $\widehat{\pi} : \Sigma(\widehat{\mathscr{G}}) \to \Lambda$ is Hölder and Pesin charts are connected subsets of $\Lambda \setminus \mathfrak{S}$, $r$ is Hölder continuous.

Let $\sigma_r : \widehat{\Sigma}_r \to \widehat{\Sigma}_r$ denote the topological Markov flow with roof function $r$ and base map $\sigma : \Sigma(\widehat{\mathscr{G}}) \to \Sigma(\widehat{\mathscr{G}})$ (see page 2 for definition). Recall that the *regular part* of $\widehat{\Sigma}_r$ is $\widehat{\Sigma}_r^\# := \{(\underline{x}, t) : \underline{x} \in \Sigma^\#(\widehat{\mathscr{G}}),\ 0 \leq t < r(\underline{x})\}$. This is a $\sigma_r$–invariant set, which contains all the periodic orbits of $\sigma_r$. By the Poincaré recurrence theorem, $\widehat{\Sigma}_r^\#$ has full measure with respect to every $\sigma_r$–invariant probability measure. Let

$$\widehat{\pi}_r : \widehat{\Sigma}_r \to M\ ,\ \widehat{\pi}_r(\underline{x}, t) := \varphi^t[\widehat{\pi}(\underline{x})].$$

The following claims follow directly from Theorem 5.6:

(1) $\widehat{\pi}_r \circ \sigma_r^t = \varphi^t \circ \widehat{\pi}_r$ for all $t \in \mathbb{R}$.
(2) $\widehat{\pi}_r[\widehat{\Sigma}_r^\#]$ has full measure with respect to $\mu$.
(3) Every $p \in \widehat{\pi}_r[\widehat{\Sigma}_r^\#]$ has finitely many pre-images in $\widehat{\Sigma}_r^\#$. In case $\mu$ is ergodic, $p \mapsto \#(\widehat{\pi}_r^{-1}(p) \cap \widehat{\Sigma}_r^\#)$ is $\varphi$–invariant, whence constant almost everywhere.
(4) Moreover, there exists $N : \mathscr{R} \times \mathscr{R} \to \mathbb{N}$ s.t. if $p = \widehat{\pi}_r(\underline{x}, t)$ where $x_i = R$ for infinitely many $i < 0$ and $x_i = S$ for infinitely many $i > 0$, then $\#\{(\underline{y}, s) \in \widehat{\Sigma}_r^\# : \widehat{\pi}_r(\underline{y}, s) = p\} \leq N(R, S)$.

This proves all parts of Theorem 1.3, except for the Hölder continuity of $\widehat{\pi}_r$.

**Step 6: Hölder continuity of $\widehat{\pi}_r$.** Every topological Markov flow is continuous with respect to a natural metric, introduced by Bowen and Walters. We will show that $\widehat{\pi}_r : \widehat{\Sigma}_r \to M$ is Hölder continuous with respect to this metric. First we recall the definition of the Bowen-Walters metric. Let $\sigma_r : \Sigma_r \to \Sigma_r$ denote a general topological Markov flow (cf. page 2). Suppose first that $r \equiv 1$ (constant suspension). Let $\psi : \Sigma_1 \to \Sigma_1$ be the suspension flow, and make the following definitions [BW72]:

- *Horizontal segments:* ordered pairs $[z, w]_h \in \Sigma_1 \times \Sigma_1$ where $z = (\underline{x}, t)$ and $w = (\underline{y}, t)$ have the same height $0 \leq t < 1$. The *length* of a horizontal segment $[z, w]_h$ is defined to be $\ell([z, w]_h) := (1-t)d(\underline{x}, \underline{y}) + td(\sigma(\underline{x}), \sigma(\underline{y}))$, where $d$ is the metric on $\Sigma$ given by $d(\underline{x}, \underline{y}) := \exp[-\min\{|n| : x_n \neq y_n\}]$.
- *Vertical segments:* ordered pairs $[z, w]_v \in \Sigma_1 \times \Sigma_1$ where $w = \psi^t(z)$ for some $t$. The *length* of a vertical segment $[z, w]_v$ is $\ell([z, w]_v) := \min\{|t| > 0 : w = \psi^t(z)\}$.
- *Basic paths* from $z$ to $w$: $\gamma := (z_0 = z \xrightarrow{t_0} z_1 \xrightarrow{t_1} \cdots \xrightarrow{t_{n-2}} z_{n-1} \xrightarrow{t_{n-1}} z_n = w)$ with $t_i \in \{h, v\}$ s.t. $[z_i, z_{i+1}]_{t_i}$ is a horizontal segment when $t_i = h$, and a vertical segment when $t_i = v$. Define $\ell(\gamma) := \sum_{i=0}^{n-1} \ell([z_i, z_{i+1}]_{t_i})$.
- *Bowen-Walters Metric on $\Sigma_1$:* $d_1(z, w) := \inf\{\ell(\gamma)\}$ where $\gamma$ ranges over all basic paths from $z$ to $w$.

Next we consider the general case $r \not\equiv 1$. The idea is to use a canonical bijection from $\Sigma_r$ to $\Sigma_1$, and declare that it is an isometry.

Bowen-Walters Metric on $\Sigma_r$ [BW72]: $d_r(z, w) := d_1(\vartheta_r(z), \vartheta_r(w))$, where $\vartheta_r : \Sigma_r \to \Sigma_1$ is the map $\vartheta_r(\underline{x}, t) := (\underline{x}, t/r(\underline{x}))$.

**Lemma 5.8.** *Assume $r$ is bounded away from zero and infinity, and Hölder continuous. Then $d_r$ is a metric, and there are constants $C_1, C_2, C_3 > 0$, $0 < \kappa < 1$ which only depend on $r$ such that for all $z = (\underline{x}, t), w = (\underline{y}, s)$ in $\Sigma_r$:*

(1) $d_r(z, w) \leq C_1[d(\underline{x}, \underline{y})^\kappa + |t - s|]$.
(2) *Conversely:*
   (a) *If* $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| \leq \frac{1}{2}$ *then* $d(\underline{x}, \underline{y}) \leq C_2 d_r(z, w)$ *and* $|s - t| \leq C_2 d_r(z, w)^\kappa$.
   (b) *If* $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$ *then* $d(\sigma(\underline{x}), \underline{y}) \leq C_2 d_r(z, w)$ *and* $|t - r(x)|, s \leq C_2 d_r(z, w)$.
(3) *For all* $|\tau| < 1$, $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq C_3 d_r(z, w)^\kappa$.

See the appendix for a proof.

**Lemma 5.9.** *The map $\widehat{\pi}_r : \widehat{\Sigma}_r \to M$ is Hölder continuous with respect to the Bowen-Walters metric.*

*Proof.* Fix $(\underline{x}, t), (\underline{y}, s) \in \widehat{\Sigma}_r$. If $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| \leq \frac{1}{2}$ then

$$\begin{aligned}
&\mathrm{dist}_M(\widehat{\pi}_r(\underline{x}, t), \widehat{\pi}_r(\underline{y}, s)) = \mathrm{dist}_M(\varphi^t(\widehat{\pi}(\underline{x})), \varphi^s(\widehat{\pi}(\underline{y})))\\
&\leq \mathrm{dist}_M(\varphi^t(\widehat{\pi}(\underline{x})), \varphi^s(\widehat{\pi}(\underline{x}))) + \mathrm{dist}_M(\varphi^s(\widehat{\pi}(\underline{x})), \varphi^s(\widehat{\pi}(\underline{y})))\\
&\leq \max_{p \in M} \|X_p\| \cdot |t - s| + \mathrm{Lip}(\varphi^s)\,\mathrm{Höl}(\widehat{\pi}) d(\underline{x}, \underline{y})^\delta
\end{aligned}$$

where $X_p$ is the vector field of $\varphi$, and $\delta$ is the Hölder exponent of $\widehat{\pi} : \Sigma(\widehat{\mathscr{G}}) \to \Lambda$. The first summand is bounded by $\mathrm{const}\, d_r(z, w)^\kappa$, by Lemma 5.8(2)(a). The second summand is bounded by $\mathrm{const}\, d_r(z, w)^\delta$, because $\varphi$ is a flow of a Lipschitz (even $C^{1+\beta}$) vector field, therefore there are global constants $a, b$ s.t. $\mathrm{Lip}(\varphi^s) \leq be^{a|s|}$ [AMR88, Lemma 4.1.8] and so $\mathrm{Lip}(\varphi^s) \leq b^{a \sup R} = O(1)$. It follows that $\mathrm{dist}_M(\widehat{\pi}_r(\underline{x}, t), \widehat{\pi}_r(\underline{y}, s)) \leq \mathrm{const}\, d_r((\underline{x}, t), (\underline{y}, s))^{\min\{\kappa, \delta\}}$.

Now assume that $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$. Since $\varphi^{r(\underline{x})}[\widehat{\pi}(\underline{x})] = \widehat{\pi}[\sigma(\underline{x})]$, we have

$$\begin{aligned}
&\mathrm{dist}_M(\widehat{\pi}_r(\underline{x}, t), \widehat{\pi}_r(\underline{y}, s)) = \mathrm{dist}_M(\varphi^t[\widehat{\pi}(\underline{x})], \varphi^s[\widehat{\pi}(\underline{y})])\\
&\leq \mathrm{dist}_M(\varphi^t[\widehat{\pi}(\underline{x})], \varphi^{r(\underline{x})}[\widehat{\pi}(\underline{x})]) + \mathrm{dist}_M(\widehat{\pi}[\sigma(\underline{x})], \widehat{\pi}[\underline{y}]) + \mathrm{dist}_M(\widehat{\pi}[\underline{y}], \varphi^s[\widehat{\pi}(\underline{y})])\\
&\leq \max_{p \in M} \|X_p\| \cdot (|t - r(\underline{x})| + |s|) + \mathrm{Höl}(\widehat{\pi}) d(\sigma(\underline{x}), \underline{y})^\delta \leq \mathrm{const}\, d_r((\underline{x}, t), (\underline{y}, s))^\delta,
\end{aligned}$$

by Lemma 5.8(2)(b).

In both cases we find that $\mathrm{dist}_M(\widehat{\pi}_r(\underline{x}, t), \widehat{\pi}_r(\underline{y}, s)) \leq \mathrm{const}\, d_r((\underline{x}, t), (\underline{y}, s))^\gamma$, where $\gamma := \min\{\kappa, \delta\}$. □

## Part 3. Applications

### 6. Measures of maximal entropy

We use the symbolic coding of Theorem 1.3 to show that a geodesic flow on a closed smooth surface with positive topological entropy can have at most countably many ergodic measures of maximal entropy. This application requires dealing with non-ergodic measures.

**Lemma 6.1.** *Let $\varphi$ be a continuous flow on a compact metric space $X$. If $\varphi$ has uncountably many ergodic measures of maximal entropy, then $\varphi$ has at least one measure of maximal entropy with non-atomic ergodic decomposition.*

*Proof.* Let $M_\varphi(X)$ denote the space of $\varphi$–invariant probability measures, together with the weak star topology. This is a compact metrizable space [Wal82, Thm 6.4]. The following claims are standard, but we could not find them in the literature.

Claim 1. *Suppose $E \subset X$ is Borel measurable, then $\mu \mapsto \mu(E)$ is Borel measurable.*

*Proof.* Let $\mathscr{M} := \{E \subset X : E \text{ is Borel, and } \mu \mapsto \mu(E) \text{ is Borel measurable}\}$. Let $\mathscr{A}$ denote the collection of Borel sets $E$ for which there are $f_n \in C(X)$ s.t. $0 \leq f_n \leq 1$ and $f_n(x) \xrightarrow[n\to\infty]{} 1_E(x)$ everywhere.

- $\mathscr{A}$ is an algebra, because if $0 \leq f_n, g_n \leq 1$ and $f_n \to 1_A, g_n \to 1_B$, then $f_n g_n \to 1_{A\cap B}$, $(1 - f_n) \to 1_{X\setminus A}$, and $(f_n + g_n - f_n g_n) \wedge 1 \to 1_{A\cup B}$.

- $\mathscr{A}$ generates the Borel $\sigma$–algebra $\mathscr{B}(X)$, because it contains every open ball $B_r(x_0)$: take $f_n(x) := \varphi_n[d(x,x_0)]$ where $\varphi_n \in C(\mathbb{R})$ and $1_{[0,r-\frac{1}{n}]} \leq \varphi_n \leq 1_{[0,r)}$.
- $\mathscr{M} \supset \mathscr{A}$: if $A \in \mathscr{A}$, then by the dominated convergence theorem $\mu(A) = \lim\limits_{n\to\infty} \int f_n d\mu$ for the $f_n \in C(X)$ s.t. $0 \leq f_n \leq 1$ and $f_n \to 1_A$. Since $\mu \mapsto \int f_n d\mu$ is continuous, $\mu \mapsto \mu(A)$ is Borel measurable.
- $\mathscr{M}$ is closed under increasing unions and decreasing intersections.

By the monotone class theorem [Sri98, Prop. 3.1.14], $\mathscr{M}$ contains the $\sigma$–algebra generated by $\mathscr{A}$, whence $\mathscr{M} = \mathscr{B}(X)$. The claim follows.

Claim 2. *$E_\varphi(X) := \{\mu \in M_\varphi(X) : \mu$ is ergodic$\}$ is a Borel subset of $M_\varphi(X)$.*

*Proof.* Fix a countable dense collection $\{f_n\}_{n\geq 1} \subset C(X)$, $0 \leq f_n \leq 1$, then $\mu$ is ergodic iff $\limsup_{k\to\infty} \int |\frac{1}{k}\int_0^k f_n \circ \varphi^t dt - \int f_n d\mu| d\mu = 0$ for every $n$. This is a countable collection of Borel conditions.

Claim 3. *The entropy map $\mu \mapsto h_\mu(\varphi)$ is Borel measurable.*

*Proof.* Let $T := \varphi^1$ (the time-one map of the flow $\varphi$), then $h_\mu(\varphi) = h_\mu(T)$. Thus $h_\mu(\varphi) = h_\mu(T) = \lim\limits_{n\to\infty} h_\mu(T, \alpha_n)$ for any sequence of finite Borel partitions $\alpha_n$ s.t. $\max\{\mathrm{diam}(A) : A \in \alpha_n\} \xrightarrow[n\to\infty]{} 0$ [Wal82, Thm 8.3]. The claim follows, since it easily follows from claim 1 that $\mu \mapsto h_\mu(T, \alpha_n)$ is Borel measurable.

Let $E_{\max}(X)$ denote the set of ergodic measures with maximal entropy. By claims 2 and 3, this is a Borel subset of $M_\varphi(X)$. By the assumptions of the lemma, $E_{\max}(X)$ is uncountable. Every uncountable Borel subset of a compact metric space carries a non-atomic Borel probability measure, because it contains a subset homeomorphic to the Cantor set [Sri98, Thm 3.2.7]. Let $\nu$ be a non-atomic Borel probability measure s.t. $\nu[E_{\max}(X)] = 1$, and let $m := \int_{E_{\max}(X)} \mu \, d\nu(\mu)$. This is a $\varphi$–invariant measure with non-atomic ergodic decomposition. Since the entropy map is affine [Wal82, Thm 8.4], $m$ has maximal entropy. $\square$

**Theorem 6.2.** *Suppose $\varphi$ is a $C^{1+\beta}$ flow with positive speed and positive topological entropy on a $C^\infty$ closed three dimensional manifold, then $\varphi$ has at most countably many ergodic measures of maximal entropy.*

*Proof.* Let $h :=$ topological entropy of $\varphi$, and assume by way of contradiction that $\varphi$ has uncountably many ergodic measures of maximal entropy. By Lemma 6.1, $\varphi$ has a measure of maximal entropy $\mu$ with a non-atomic ergodic decomposition.

By the variational principle [Wal82, Thm 8.3], $h_\mu(\varphi) = h$. By the affinity of the entropy map [Wal82, Thm 8.4], almost every ergodic component $\mu_x$ of $\mu$ has entropy $h$. Fix some $0 < \chi_0 < h$. By the Ruelle entropy inequality [Rue78], a.e. ergodic component $\mu_x$ is $\chi_0$–hyperbolic. Consequently, $\mu$ is $\chi_0$–hyperbolic.

This places us in the setup considered in part 2, and allows us to apply Theorem 1.3 to $\mu$. We obtain a coding $\pi_r : \Sigma_r \to M$ s.t. $\mu[\pi(\Sigma_r^\#)] = 1$ and $\pi_r : \Sigma_r^\# \to M$ is finite-to-one (though not necessarily bounded-to-one).

Lifting Procedure: *Define a measure $\widehat{\mu}$ on $\Sigma_r$ by setting for $E \subset \Sigma_r$ Borel*

$$\widehat{\mu}(E) := \int_{\pi_r(\Sigma_r^\#)} \left( \frac{1}{|\pi_r^{-1}(p) \cap \Sigma_r^\#|} \sum_{\pi_r(\underline{x},t)=p} 1_E(\underline{x}, t) \right) d\mu(p), \tag{6.1}$$

*then $\widehat{\mu}$ is a $\sigma_r$–invariant measure, $\widehat{\mu} \circ \pi_r^{-1} = \mu$, and $h_{\widehat{\mu}}(\sigma_r) = h_\mu(\varphi)$.*

*Proof.* We start by clearing away all the measurability concerns. Let $X := \Sigma_r$ and $Y := M \uplus X$ (disjoint union). Define $f : \Sigma_r \to Y$ by $f\restriction_{\Sigma_r^\#} = \pi_r$ and $f\restriction_{\Sigma_r \setminus \Sigma_r^\#} = \mathrm{Id}$, then $f : X \to Y$ is a countable-to-one Borel map between polish spaces. Such maps send Borel sets to Borel sets [Sri98, Thm 4.12.4], so $\pi_r(\Sigma_r^\#) = f(\Sigma_r^\#)$ is Borel.

Next we show that the integrand in (6.1) is Borel. Let $B := \{(x, f(x)) : f(x) \in M\}$. This is a Borel subset of $X \times Y$, because the graph of a Borel function is Borel [Sri98, Thm 4.5.2]. For every $y \in Y$, $B_y := \{x \in X : (x, y) \in B\}$ is countable, because either $y \in M$ and $B_y := \pi_r^{-1}(y) \cap \Sigma_r^\#$, or $y \notin M$ and then $B_y = \varnothing$. By Lusin's theorem [Sri98, Thm 5.8.11], there are countably many partially defined Borel functions $\varphi_n : M_n \to X$ s.t. $B = \bigcup_{n=1}^\infty \{(\varphi_n(y), y) : y \in M_n\}$. Write $B = \biguplus_{n=1}^\infty \{(\varphi_n(y), y) : y \in M_n'\}$, $M_n' := \{y \in M_n : k < n, y \in M_k \Rightarrow \varphi_n(y) \neq \varphi_k(y)\}$. Then for every $y \in M$,

$$\pi_r^{-1}(y) = \{\varphi_n(y) : n \geq 1, y \in M_n'\}, \text{ and } m \neq n \Rightarrow \varphi_m(y) \neq \varphi_n(y).$$

Thus, the integrand in (6.1) equals $\sum_{n=1}^\infty 1_{M_n'}(p) 1_E(\varphi_n(p)) \big/ \sum_{n=1}^\infty 1_{M_n'}(p)$, a Borel measurable function.

Now that we know that (6.1) makes sense it is a trivial matter to see that it defines a measure $\widehat{\mu}$. This measure is $\sigma_r$–invariant because of the $\varphi$–invariance of $\mu$ and the commutation relation $\pi_r \circ \sigma_r = \varphi \circ \pi_r$. It has the same entropy as $\mu$, because finite-to-one factor maps preserve entropy [AR62].

Projection Procedure: *Every $\sigma_r$–invariant probability measure $\widehat{m}$ on $\Sigma_r$ projects to a $\varphi$–invariant probability measure $m := \widehat{m} \circ \pi_r^{-1}$ on $M$ with the same entropy.*

*Proof.* By the Poincaré recurrence theorem, every $\sigma_r$–invariant probability measure is carried by $\Sigma_r^\#$, therefore $\pi_r : (\Sigma_r, \widehat{m}) \to (M, m)$ is a finite-to-one factor map. Such maps preserve entropy.

Combining the lifting procedure and the projection procedure we see that the supremum of the entropies of $\varphi$–invariant measures on $M$ equals the supremum of the entropies of $\sigma_r$–invariant measures on $\Sigma_r$, and therefore $\widehat{\mu}$ given by (6.1) is a measure of maximal entropy for $\sigma_r$.

Claim. *$\sigma_r$ has at most countably many ergodic measures of maximal entropy.*

*Proof.* We recall the well-known relation between measures of maximal entropy for $\sigma_r$ and equilibrium measures for the shift map $\sigma : \Sigma \to \Sigma$ [BR75]: $S := \Sigma \times \{0\}$ is a Poincaré section for $\sigma_r : \Sigma_r \to \Sigma_r$, therefore every measure of maximal entropy $\widehat{\mu}$ for $\sigma_r$ can be put in the form $\widehat{\mu} = \frac{1}{\int_\Sigma r d\widehat{\mu}_\Sigma} \int_\Sigma \int_0^{r(\underline{x})} \delta_{(\underline{x},t)} dt\, d\widehat{\mu}_\Sigma(\underline{x})$ where $\widehat{\mu}_\Sigma$ is a shift-invariant measure on $\Sigma$. The denominator is well-defined, because $r$ is bounded away from zero and infinity. If $\widehat{\mu}$ is ergodic, then $\widehat{\mu}_\Sigma$ is ergodic.

By the Abramov formula, $h_{\widehat{\mu}}(\sigma_r) = h_{\widehat{\mu}_\Sigma}(\sigma) \big/ \int_\Sigma r d\widehat{\mu}_\Sigma$. Similar formulas hold for all other $\sigma_r$–invariant probability measures $m$ and the measures $m_\Sigma$ they induce on $\Sigma$. Since $\widehat{\mu}$ is a measure of maximal entropy, $h_{m_\Sigma}(\sigma) \big/ \int_\Sigma r dm_\Sigma = h_m(\sigma_r) \leq h$ (the maximal possible entropy) for all $\sigma$–invariant measures $m_\Sigma$. This is equivalent to saying that $h_{m_\Sigma}(\sigma) + \int_\Sigma (-hr) dm_\Sigma \leq 0$, with equality iff $h_m(\sigma_r) = h$. Thus, if $\widehat{\mu}$ is a measure of maximal entropy for $\sigma_r$, then $\widehat{\mu}_\Sigma$ is an equilibrium measure for $-hr$, where $h$ is the value of the maximal entropy. Also, the topological pressure $P(-hr) := \sup\{h_\nu(\sigma) - h \int_\Sigma r d\nu\} = 0$, where the supremum ranges over all $\sigma$–invariant probability measures $\nu$ on $\Sigma$.

Recall that $r : \Sigma \to \mathbb{R}$ is Hölder continuous. By [BS03], a Hölder continuous potential on a topologically transitive countable Markov shift has at most one equilibrium measure. If the condition of topological transitivity is dropped, then there are at most countably many such measures, one for each transitive component with maximal topological entropy [Gur69] (see the proof of [Sar13, Thm. 5.3]). It follows that there are at most countably many possibilities for $\widehat{\mu}_\Sigma$, and therefore at most countably many possibilities for $\widehat{\mu}$.

We can now obtain the contradiction which proves the theorem. Consider the ergodic decomposition of $\widehat{\mu}$ defined by (6.1). Almost every ergodic component is a measure of maximal entropy (because the entropy function is affine). By the claim there are at most countably many different such measures. Therefore the ergodic decomposition of $\widehat{\mu}$ is atomic: $\widehat{\mu} = \sum p_i \widehat{\mu}_i$ with $\widehat{\mu}_i$ ergodic and $p_i \in (0,1)$ s.t. $\sum p_i = 1$. Projecting to $M$, and noting that factors of ergodic measures are ergodic, we find that $\mu = \sum p_i \mu_i$ where $\mu_i := \widehat{\mu}_i \circ \pi_r^{-1}$ are ergodic. This is an atomic ergodic decomposition for $\mu$. But the ergodic decomposition is unique, and we assumed that $\mu$ has a non-atomic ergodic decomposition. □

## 7. Mixing for equilibrium measures on topological Markov flows

Let $\sigma_r : \Sigma_r \to \Sigma_r$ be a topological Markov flow, together with the Bowen-Walters metric. Let $\Phi : \Sigma_r \to \mathbb{R}$ be bounded and continuous.

The topological pressure of $\Phi$: $P(\Phi) := \sup\{h_\mu(\sigma_r) + \int \Phi d\mu\}$, where the supremum ranges over all $\sigma_r$–invariant probability measures $\mu$ on $\Sigma_r$.

Equilibrium measure for $\Phi$: A $\sigma_r$–invariant probability measure $\mu$ on $\Sigma_r$ s.t. $h_\mu(\sigma_r) + \int \Phi d\mu = P(\Phi)$.

**Theorem 7.1.** *Suppose $\mu$ is an equilibrium measure of a bounded Hölder continuous potential for a topological Markov flow $\sigma_r : \Sigma_r \to \Sigma_r$. If $\sigma_r$ is topologically transitive, then the following are equivalent:*

(1) *If $e^{i\theta r} = h/h \circ \sigma$ for some Hölder continuous $h : \Sigma \to S^1$ and $\theta \in \mathbb{R}$, then $\theta = 0$ and $h = \text{const}$.*
(2) *$\sigma_r$ is weak mixing.*
(3) *$\sigma_r$ is mixing.*

(3) ⇒ (2) ⇒ (1) because if $e^{i\theta r} = h/h \circ \sigma$, then $F(x,t) = e^{-i\theta t} h(x)$ is an eigenfunction of the flow. (1) ⇒ (2) ⇒ (3) are known in the special case when $\Sigma$ is a subshift of finite type: Parry and Pollicott proved (1) ⇒ (2) [PP90], and Ratner proved (2) ⇒ (3) ⇒ Bernoulli [Rat74], [Rat78]. Dolgopyat showed us a different proof of (2) ⇒ (3) (private communication). These proofs can be pushed through to the countable alphabet case with some effort, using the thermodynamic formalism for countable Markov shifts [BS03]. The details can be found in [LLS16, Thm 4.6].

The following theorem is a symbolic analogue of Plante's necessary and sufficient condition for a transitive Anosov flow to be a constant suspension of an Anosov diffeomorphism [Pla72], see also [Bow73].

**Theorem 7.2.** *Let $\sigma_r : \Sigma_r \to \Sigma_r$ be a topologically transitive topological Markov flow. Either every equilibrium measure of a bounded Hölder continuous potential is mixing, or there is $\Sigma_r' \subset \Sigma_r$ of full measure s.t. $\sigma_r : \Sigma_r' \to \Sigma_r'$ is topologically conjugate to a topological Markov flow with constant roof function.*

*Proof.* If $\Sigma$ is a finite set, then $\Sigma_r$ equals a single closed orbit, and the claim is trivial. From now on assume that $\Sigma$ is infinite.

Assume $\sigma_r$ is not mixing, then $\exp[i\theta r] = h/h \circ \sigma$ with $h : \Sigma \to S^1$ Hölder continuous and $\theta \neq 0$. Write $\theta = 2\pi/c$ and put $h$ in the form $h = \exp[i\theta U]$, where $U : \Sigma \to \mathbb{R}$ is Hölder continuous. Necessarily $r + U \circ \sigma - U \in c\mathbb{Z}$. We are free to change $U$ on every partition set by a constant in $c\mathbb{Z}$ to make sure $U$ is bounded and positive. Fix $N > 2\|U\|_\infty / \inf(r)$.

CONSTRUCTION: *There is a cylinder* $A =_{-m} [y_{-m}, \ldots, y_n]$ *s.t.:*

(i) $m, n > 0$ *and* $y_{-m} = y_n$.
(ii) $n_A(\cdot) > N$ *on* $A$*, where* $n_A(\underline{x}) := \inf\{n \geq 1 : \sigma^n(\underline{x}) \in A\}$.
(iii) $\underline{x}, \underline{x}' \in A \Rightarrow |U(\underline{x}) - U(\underline{x}')| < N \inf(r)$.

To find $A$, take $\underline{y} \in \Sigma$ with dense orbit. Since $\Sigma$ is infinite, $\sigma^k(\underline{y})$ are distinct. Therefore, $\underline{y}$ has a cylindrical neighborhood $C$ s.t. $\sigma^k(C) \cap C = \varnothing$ for $k = 1, \ldots, N$. Choose $m, n > 0$ so large that ${}_{-m}[y_{-m}, \ldots, y_n] \subset C$, and $|U(\underline{x}) - U(\underline{x}')| < N \inf(r)$ for all $\underline{x}, \underline{x}' \in C$. Since $\underline{y}$ has a dense orbit, every symbol appears in $\underline{y}$ infinitely often in the past and in the future, therefore we can choose $m, n$ so that $y_{-m} = y_n$. The cylinder $A = {}_{-m}[y_{-m}, \ldots, y_n]$ satisfies (i), (ii) and (iii), because $A \subset C$.

Since $\mu$ is ergodic and globally supported, the following set has full $\mu$–measure: $\Sigma_r' := \{z \in \Sigma_r : \sigma_r^t(z) \in A \times \{0\}$ infinitely often in the past and in the future$\}$.

STEP 1: *$\sigma_r : \Sigma_r' \to \Sigma_r'$ is topologically conjugate to a topological Markov flow $\sigma_{r^*} : \Sigma^*_{r^*} \to \Sigma^*_{r^*}$ whose roof function $r^*$ takes values in $c\mathbb{Z}$.*

*Proof.* $A \times \{0\}$ is a Poincaré section for $\sigma_r : \Sigma_r' \to \Sigma_r'$. The roof function is $r_A := r + r \circ \sigma + \cdots + r \circ \sigma^{n_A - 1}$. By (ii), $\inf(r_A) > N \inf(r)$, so $0 < U < \inf(r_A)$.

Let $S := \{\sigma_r^{U(\underline{x})}(\underline{x}, 0) : (\underline{x}, 0) \in \Sigma_r'\}$. This is a Poincaré section for $\sigma_r : \Sigma_r' \to \Sigma_r'$, and its roof function is $r_A^* := r_A + U \circ \sigma^{n_A} - U$ (this is always positive because $U < \inf(r_A)$). All the values of $r_A^*$ belong to $c\mathbb{Z}$, as can be seen from the identity $r_A^* = \sum_{k=0}^{n_A - 1} (r + U \circ \sigma - U) \circ \sigma^k$. We claim that the section map of $S$ is topologically conjugate to a topological Markov shift. Let $V$ denote the collection of sets of the form

$$\langle \underline{B} \rangle := \{\sigma_r^{U(\underline{x})}(\underline{x}, 0) : \underline{x} \in {}_{-m}[\underline{A}, \underline{B}, \underline{A}]\},$$

where $\underline{A} = (y_{-m}, \ldots, y_n)$ is the word defining $A$, and $\underline{B}$ is any other word s.t. ${}_{-m}[\underline{A}, \underline{B}, \underline{A}] \neq \varnothing$ for which the only appearances of $\underline{A}$ in $(\underline{A}, \underline{B}, \underline{A})$ are at the beginning and at the end.

It is easy to see that $\sigma_r^{U(\underline{x})}(\underline{x}, 0) \in S$ iff $\underline{x} = (\ldots, \underline{A}, \underline{B}^1, \underline{A}, \underline{B}^2, \underline{A}, \ldots)$ with $\langle \underline{B}^i \rangle \in V$, and that any sequence $\{\langle \underline{B}^i \rangle\}_{i \in \mathbb{Z}} \in V^{\mathbb{Z}}$ appears this way. Let $\pi : S \to V^{\mathbb{Z}}$ be the map $\pi(\underline{x}) = \{\langle \underline{B}^i \rangle\}_{i \in \mathbb{Z}}$. Since $\underline{A}$ appears in $(\underline{A}, \underline{B}^i, \underline{A})$ only at the beginning and the end, $\pi \circ \sigma_r^{r_A^*} = \sigma \circ \pi$, with $\sigma =$ the left shift on $V^{\mathbb{Z}}$. So the section map of $S$ is topologically conjugate to the shift on $V^{\mathbb{Z}}$. Let $\Sigma^* := V^{\mathbb{Z}}$. The roof function with respect to this new coding is $r^* := r_A^* \circ \pi^{-1}$. Direct calculations show that the Hölder continuity of $r$ implies the Hölder continuity of $r^*$. So $\sigma_{r^*} : \Sigma^*_{r^*} \to \Sigma^*_{r^*}$ is a TMF, and $\sigma_r : \Sigma_r' \to \Sigma_r'$ is topologically conjugate to $\sigma_{r^*}$.

STEP 2: *$\sigma_{r^*} : \Sigma^*_{r^*} \to \Sigma^*_{r^*}$ is topologically conjugate to a topological Markov flow $\sigma_{\widetilde{r}} : \widetilde{\Sigma}_{\widetilde{r}} \to \widetilde{\Sigma}_{\widetilde{r}}$ where $\widetilde{r}$ takes values in $c\mathbb{Z}$, and $\widetilde{r}(\underline{x}) = \widetilde{r}(x_0)$.*

*Proof.* Since $r^*$ is Hölder continuous and takes values in $c\mathbb{Z}$, there must be some $n_0 > 0$ s.t. $r^*$ is constant on every cylinder of the form ${}_{-n_0}[a_{-n_0}, \dots, a_{n_0}]$. Take $\widetilde{\pi}(\underline{x}, t) := (\{\underline{x}^i\}_{i\in\mathbb{Z}}, t)$, where $\underline{x}^i := (x_{-n_0+i}, \dots, x_{n_0+i})$. The reader can check that the collection of $\{\underline{x}^i\}_{i\in\mathbb{Z}}$ thus obtained is a topological Markov shift $\widetilde{\Sigma}$, and that $\widetilde{r}(\{\underline{x}^i\}_{i\in\mathbb{Z}})$ only depends on the first symbol $\underline{x}^0$.

Step 3: *$\sigma_{\widetilde{r}} : \widetilde{\Sigma}_{\widetilde{r}} \to \widetilde{\Sigma}_{\widetilde{r}}$ is topologically conjugate to a topological Markov flow $\sigma_{\widehat{r}} : \widehat{\Sigma}_{\widehat{r}} \to \widehat{\Sigma}_{\widehat{r}}$ where $\widehat{r}$ is constant equal to $c$.*

*Proof.* The set $\{(\underline{x}, kc) : \underline{x} \in \widetilde{\Sigma}, k \in \mathbb{Z},\ 0 \leq kc < \text{value of } \widetilde{r} \text{ on } {}_0[x_0]\}$ is a Poincaré section for the suspension flow with constant roof function (equal to $c$). The section map is conjugate to a topological Markov shift $\widehat{\Sigma}$ which we now describe.

Let $\widetilde{\mathscr{G}} = \mathscr{G}(\widetilde{V}, \widetilde{E})$ be the graph of $\widetilde{\Sigma}$. Let $\widehat{\Sigma} = \Sigma(\widehat{\mathscr{G}})$, where $\widehat{\mathscr{G}}$ has the set of vertices $\widehat{V} := \{\binom{v}{k} : v \in \widetilde{V}, 0 \leq kc < \text{value of } \widetilde{r} \text{ on } {}_0[v]\}$ and edges $\binom{v}{k} \to \binom{v}{k+1}$ when $\binom{v}{k+1} \in \widetilde{V}$, and $\binom{v}{k} \to \binom{w}{0}$ when $\binom{v}{k+1} \notin \widetilde{V}$ and $v \to w$ in $\widetilde{E}$. The conjugacy $\widehat{\pi} : \widetilde{\Sigma}_{\widetilde{r}} \to \widehat{\Sigma}_{\widehat{r}}$ is $\widehat{\pi}(\underline{x}, t) := (\sigma^{\lfloor t/c \rfloor}(\underline{y}), t - \lfloor t/c \rfloor c)$, where $\underline{y}$ is given by $(\dots; \binom{x_0}{0}, \binom{x_0}{1}, \dots, \binom{x_0}{\widetilde{r}(x_0)/c-1}; \binom{x_1}{0}, \binom{x_1}{1}, \dots, \binom{x_1}{\widetilde{r}(x_1)/c-1}; \dots)$ with $\binom{x_0}{0}$ at the zeroth coordinate. ☐

## 8. Counting simple closed orbits

Let $\pi(T) := \#\{[\gamma] : \gamma \text{ is a simple closed geodesic s.t. } \ell[\gamma] \leq T\}$. In this section we prove the following generalization of Theorem 1.1.

**Theorem 8.1.** *Suppose $\varphi$ is a $C^{1+\beta}$ flow with positive speed and positive topological entropy $h$ on a $C^\infty$ closed three dimensional manifold $M$. If $\varphi$ has a measure of maximal entropy, then $\pi(T) \geq C\frac{e^{hT}}{T}$ for all $T$ large enough and $C > 0$.*

This implies Theorem 1.1, because every $C^\infty$ flow admits a measure of maximal entropy. Indeed, by a theorem of Newhouse [New89], $\varphi^1 : M \to M$ admits a measure of maximal entropy $m$, and $\mu := \int_0^1 m \circ \varphi^t dt$ has maximal entropy for $\varphi$.

**Discussion.** Theorem 8.1 strengthens Katok's bound $\liminf_{T\to\infty} \frac{1}{T} \log \pi(T) \geq h$, see [Kat80, Kat82] for general flows, and it improves Macarini and Schlenk's bound $\liminf_{T\to\infty} \frac{1}{T} \log \pi(T) > 0$ for the class of Reeb flows in [MS11]. If one assumes more on the flow, then much better bounds for $\pi(T)$ are known:

(1) *Geodesic flows on closed hyperbolic surfaces:* $\pi(T) \sim e^t/t$ [Hub59].
(2) *Topologically weak mixing Anosov flows* (e.g. geodesic flows on closed surfaces with negative curvature): $\pi(T) \sim Ce^{hT}/T$ [Mar69] where $C = 1/h$ (C. Toll, unpublished). See [PS98] for estimates of the error term. The earliest estimates for $\pi(T)$ in variable curvature are due to Sinaĭ [Sin66].
(3) *Topologically weak mixing Axiom A flows:* $\pi(T) \sim e^{hT}/hT$ [PP83]. See [PS01] for an estimate of the error term.
(4) *Geodesic flows on compact rank one manifolds:* $C_1\frac{e^{hT}}{T} \leq \pi_0(T) \leq C_2\frac{e^{hT}}{T}$ for some $C_1, C_2 > 0$, where $\pi_0(T)$ counts the homotopy classes of simple closed geodesics with length less than $T$ [Kni97, Kni02].
(5) *Geodesic flows for certain non-round spheres:* for certain metrics constructed by [Don88, BG89], $\pi(T) \sim e^{hT}/hT$ [Wea14].

We cannot give upper bounds for $\pi(T)$ as in (1)–(5), because in the general setup we consider there can be compact invariant sets with lots of closed geodesics but zero topological entropy (e.g. embedded flat cylinders). Such sets have zero measure for any ergodic measure with positive entropy, and they lie outside the "sets of full measure" that we can control using the methods of this paper. Adding to our pessimism is the existence of $C^r$ $(1 < r < \infty)$ surface diffeomorphisms with super-exponential growth of periodic points [Kal00]. The suspension of these examples gives $C^r$ flows with super-exponential growth of closed orbits. To the best of our knowledge, the problem of doing this in $C^\infty$ is still open.

**Preparations for the proof of Theorem 8.1.** Fix an ergodic measure of maximal entropy for $\varphi$, and apply Theorem 1.2 with this measure. The result is a topological Markov flow $\sigma_r : \Sigma_r \to \Sigma_r$ together with a Hölder continuous map $\pi_r : \Sigma_r \to M$, satisfying (1)–(6) in Theorem 1.2.

We saw in the proof of Theorem 6.2 (see page 31) that if $\varphi$ has a measure of maximal entropy, then $\sigma_r$ has a measure of maximal entropy. By the ergodic decomposition, $\sigma_r$ has an ergodic measure of maximal entropy. Fix such a measure $\mu$, and write $\mu = \frac{1}{\int_\Sigma r d\nu} \int_\Sigma \left( \int_0^{r(\underline{x})} \delta_{(\underline{x},t)} dt \right) d\nu(x)$. The induced measure $\nu$ is an ergodic shift invariant measure on $\Sigma$. When we proved Theorem 6.2, we saw that $\nu$ is an equilibrium measure for $\phi = -hr$. Like all ergodic shift invariant measures, $\nu$ is supported on a topologically transitive topological Markov shift $\Sigma' \subseteq \Sigma$ [ADU93]. There is no loss of generality in assuming that $\sigma : \Sigma \to \Sigma$ is topologically transitive (otherwise we work with $\Sigma'$).

**Proof of Theorem 8.1 when $\mu$ is mixing.** Fix $0 < \varepsilon < 10^{-1} \inf(r)$. Since $r$ is Hölder, there are $H > 0$ and $0 < \alpha < 1$ s.t. $|r(\underline{x}) - r(\underline{y})| \leq Hd(\underline{x}, \underline{y})^\alpha$. Recall that $d(\underline{x}, \underline{y}) = \exp[-\min\{|n| : x_n \neq y_n\}]$. For every $\ell \geq 1$, if $x_{-n_0}^{n_0+\ell} = y_{-n_0}^{n_0+\ell}$ then

$$|r_\ell(\underline{x}) - r_\ell(\underline{y})| \leq \sum_{i=0}^{\ell-1} Hd(\sigma^i(\underline{x}), \sigma^i(\underline{y}))^\alpha \leq H \sum_{i=0}^{\ell-1} e^{-\alpha \min\{n_0+i, n_0+\ell-i\}} < \frac{2He^{-\alpha n_0}}{1 - e^{-\alpha}}.$$

Choose $n_0$ s.t. $\sup\{|r_\ell(\underline{x}) - r_\ell(\underline{y})| : x_{-n_0}^{n_0+\ell} = y_{-n_0}^{n_0+\ell}, \ell \geq 1\} < \varepsilon$. Fix some cylinder $A := {}_{-n_0}[a_{-n_0}, \ldots, a_{n_0}]$ s.t. $\nu(A) \neq 0$, and let $\Upsilon(T) := \biguplus_{n=1}^\infty \Upsilon(T, n)$, where

$$\Upsilon(T, n) := \{(\underline{y}, n) : \underline{y} \in A, \sigma^n(\underline{y}) = \underline{y}, |r_n(\underline{y}) - T| < 2\varepsilon\}.$$

Given $(\underline{y}, n) \in \Upsilon(T, n)$, let $\gamma_{\underline{y},n} : [0, r_n(\underline{y})] \to M$, $\gamma_{\underline{y},n}(t) = \pi_r[\sigma_r^t(\underline{y}, 0)]$. This is a closed orbit with length $\ell(\gamma_{\underline{y},n}) = r_n(\underline{y}) \in [T - 2\varepsilon, T + 2\varepsilon]$. But $\gamma_{\underline{y},n}(t)$ is not necessarily simple, because $\pi$ is not injective. Let $\gamma_{\underline{y},n}^s := \gamma_{\underline{y},n} \restriction_{[0, \ell(\gamma_{\underline{y},n})/N]}$, where $N = N(\underline{y}, n) := \#\{0 \leq t < \ell(\gamma_{\underline{y},n}) : \gamma_{\underline{y},n}(t) = \gamma_{\underline{y},n}(0)\}$. Then $\gamma_{\underline{y},n}^s$ is a simple closed orbit. We have $N = 1$ iff $\gamma_{\underline{y},n}$ is simple, and $N < \ell(\gamma_{\underline{y},n})/\inf(r)$, because an orbit with length less than $\inf(r)$ cannot be closed.

We obtain a map $\Theta : \Upsilon(T) \to \{[\gamma]$: $\gamma$ is a simple closed orbit s.t. $\ell(\gamma) \leq T + 2\varepsilon\}$,

$$\Theta : (\underline{y}, n) \mapsto [\gamma_{\underline{y},n}^s].$$

The map $\Theta$ is not one-to-one, but there is a uniform bound on its non-injectivity:

$$1 \leq \frac{\#\Theta^{-1}([\gamma_{\underline{y},n}^s])}{n} \leq c_0. \tag{8.1}$$

Here is the proof. Suppose $(\underline{y}, n), (\underline{z}, m) \in \Upsilon(T)$ and $[\gamma_{\underline{y},n}^s] = [\gamma_{\underline{z},m}^s]$, then:

- $N(\underline{y},n) = N(\underline{z},m)$: $\left[\frac{T-2\varepsilon}{N(\underline{y},n)}, \frac{T+2\varepsilon}{N(\underline{y},n)}\right]$ and $\left[\frac{T-2\varepsilon}{N(\underline{z},m)}, \frac{T+2\varepsilon}{N(\underline{z},m)}\right]$ both contain $\ell = \ell(\gamma^s_{\underline{y},n}) = \ell(\gamma^s_{\underline{z},m})$. But $N(\underline{y},n), N(\underline{z},m) < \frac{T+2\varepsilon}{\inf(r)}$, and $\left[\frac{T-2\varepsilon}{j}, \frac{T+2\varepsilon}{j}\right]$ are pairwise disjoint for $j = 1, \ldots, [\frac{T+2\varepsilon}{\inf(r)}]$, because $\varepsilon < \frac{1}{10}\inf(r)$.
- $[\gamma_{\underline{y},n}] = [\gamma_{\underline{z},m}]$, because $[\gamma^s_{\underline{y},n}] = [\gamma^s_{\underline{z},m}]$ and $N(\underline{y},n) = N(\underline{z},m)$.
- $n = m$, because $n$, $m$ are the number of times $\gamma_{\underline{y},n}$, $\gamma_{\underline{z},m}$ enter $\Lambda_0 := \pi_r(\Sigma \times \{0\})$, and equivalent closed orbits enter $\Lambda_0$ the same number of times.
- $\pi_r(\underline{z},0) = \pi_r(\sigma^k(\underline{y}),0)$ for some $k = 0, \ldots, n-1$, because $\pi_r(\underline{z},0) \in \gamma_{\underline{y},n} \cap \Lambda_0$.
- $\underline{y}, \underline{z} \in \Sigma^\#$, and $y_i = a_0$ for infinitely many $i < 0$ and infinitely many $i > 0$.

By Theorem 1.2(5) there is a constant $c_0 := N(a_0, a_0)$ s.t. if $x_i = a_0$ for infinitely many $i > 0$ and infinitely many $i < 0$, then $\#\{\underline{z} \in \Sigma^\# : \widehat{\pi}(\underline{z}) = \widehat{\pi}(\underline{x})\} \leq c_0$. Thus $\#\Theta^{-1}([\gamma^s_{\underline{y},n}]) \leq \#[\Sigma^\#_r \cap \bigcup_{k=0}^{n-1} \pi_r^{-1}\{\pi_r(\sigma^k(\underline{y}),0)\}] \leq c_0 n$. Also $\#\Theta^{-1}([\gamma^s_{\underline{y},n}]) \geq n$, because $[\gamma^s_{\sigma^k(\underline{y}),n}] = [\gamma^s_{\underline{y},n}]$ for $k = 0, \ldots, n-1$. This proves (8.1).

By the inequality (8.1) and the fact shown above that $[\gamma^s_{\underline{y},n}] = [\gamma^s_{\underline{z},m}] \Rightarrow m = n$,

$$
\begin{aligned}
&\#\{[\gamma] : \gamma \text{ simple closed orbit s.t. } \ell(\gamma) \leq T + 2\varepsilon\} \geq \\
&\geq \#\{[\gamma^s_{\underline{y},n}] : (\underline{y},n) \in \Upsilon(T)\} = \sum_{n=1}^{\infty} \#\{[\gamma^s_{\underline{y},n}] : (\underline{y},n) \in \Upsilon(T,n)\} \\
&\asymp \sum_{n=1}^{\infty} \frac{\#\Upsilon(T,n)}{n} \text{ , where } A_n \asymp B_n \text{ means } \exists C, N_0 \text{ s.t. } \forall n > N_0,\ C^{-1} \leq \frac{A_n}{B_n} \leq C \\
&= \sum_{n=1}^{\infty} \left( \frac{1}{n} \sum_{\sigma^n(\underline{y})=\underline{y}} 1_A(\underline{y}) 1_{[-2\varepsilon,2\varepsilon]}(r_n(\underline{y}) - T) \right) \\
&\asymp \frac{e^{hT}}{T} \sum_{n=1}^{\infty} \sum_{\sigma^n(\underline{y})=\underline{y}} 1_A(\underline{y}) 1_{[-2\varepsilon,2\varepsilon]}(r_n(\underline{y}) - T) e^{-hr_n(\underline{y})} \\
&= \frac{e^{hT}}{T} S(T), \text{ where } S(T) := \sum_{n=1}^{\infty} \sum_{\sigma^n(\underline{y})=\underline{y}} 1_A(\underline{y}) 1_{[-2\varepsilon,2\varepsilon]}(r_n(\underline{y}) - T) e^{-hr_n(\underline{y})}.
\end{aligned}
$$

To prove the theorem, it is enough to show that $\liminf S(T) > 0$.

Recall that $\nu$ is an equilibrium measure for $\phi = -hr$ and $P(-hr) = 0$ (see the claim on page 31). The structure of such measures was found in [BS03]. We will not repeat the characterization here, but we will simply note that it implies the following uniform estimate [BS03, page 1387]: $\exists C(a) > 1$ s.t. for every cylinder of the form ${}_0[\underline{b}] = {}_0[\xi_0, \ldots, \xi_n]$ with $\xi_n = a$, $C(a)^{-1} \leq \frac{\nu({}_0[\underline{b}])}{\exp(-hr_n(\underline{y}))} \leq C(a)$ for all $\underline{y} \in {}_0[\underline{b}]$. It follows that there is a constant $G = G(A)$ s.t.

$$G^{-1} \leq \frac{\exp[-hr_n(\underline{y})]}{\nu({}_{-n_0}[a_{-n_0}, \ldots, a_{-1}; y_0, \ldots, y_{n-1}; a_0, \ldots, a_{n_0}])} \leq G$$

for every $\underline{y} \in {}_{-n_0}[a_{-n_0}, \ldots, a_{-1}; y_0, \ldots, y_{n-1}; a_0, \ldots, a_{n_0}]$. Let

$$U_T := \bigcup_{(\underline{y},n) \in \Upsilon(T)} {}_{-n_0}[a_{-n_0}, \ldots, a_{-1}; y_0, \ldots, y_{n-1}; a_0, \ldots, a_{n_0}],$$

then $S(T) \asymp \nu[U_T] = (\varepsilon^{-1} \int r d\nu) \mu(U_T \times [0, \varepsilon])$.

We claim that

$$U_T \times [0,\varepsilon] \supset (A \times [0,\varepsilon]) \cap \sigma_r^{-T}(A \times [0,\varepsilon]). \tag{8.2}$$

Once this is shown, we can use the mixing of $\mu$ to get $\liminf \mu(U_T \times [0,\varepsilon]) > 0$, whence $\liminf S(T) > 0$. Suppose $(\underline{x},t), \sigma_r^T(\underline{x},t) \in A \times [0,\varepsilon]$, and write $\sigma_r^T(\underline{x},t) = (\sigma^n(\underline{x}), t+T-r_n(\underline{x}))$. Since $\underline{x} \in A \cap \sigma^{-n}(A)$, there exists $\underline{y} \in A$ s.t. $\sigma^n(\underline{y}) = \underline{y}$ and $y_{-n_0}^{n+n_0} = x_{-n_0}^{n+n_0}$. By the choice of $n_0$, $|r_n(\underline{x}) - r_n(\underline{y})| < \varepsilon$. Since $t, t+T-r_n(\underline{x}) \in [0,\varepsilon]$, $|r_n(\underline{x}) - T| < \varepsilon$, whence $|r_n(\underline{y}) - T| < 2\varepsilon$. So $(\underline{x},t) \in U_T \times [0,\varepsilon]$. This proves (8.2).

**Proof of Theorem 8.1 when $\mu$ is not mixing.** In this case, Theorem 7.2 gives us a set of full measure $\Sigma_r' \subset \Sigma_r$ s.t. $\sigma_r : \Sigma_r' \to \Sigma_r'$ is topologically conjugate to a constant suspension over a topologically transitive topological Markov shift $\sigma_c : \widetilde{\Sigma} \times [0,c) \to \widetilde{\Sigma} \times [0,c)$. Let $\vartheta : \widetilde{\Sigma} \times [0,c) \to \Sigma_r'$ denote the topological conjugacy, and let $\widetilde{p} : \widetilde{\Sigma} \times [0,c) \to M$ be the map $\widetilde{p} := \pi_r \circ \vartheta$. The map $\widetilde{p}$ has the same finiteness-to-one properties of $\pi_r$, because looking carefully at the proof of Theorem 7.2, we can see that if $\underline{x} \in \widetilde{\Sigma}$ contains some symbol $v$ infinitely many times in its future (resp. past), then $\vartheta(\underline{x},t) = (\underline{y},s)$ where $\underline{y}$ contains some symbol $a = a(v)$ infinitely many times in its future (resp. past).

Since $\sigma_r : \Sigma_r' \to \Sigma_r'$ has a measure of maximal entropy, $\sigma_c : \widetilde{\Sigma} \times [0,c) \to \widetilde{\Sigma} \times [0,c)$ has a measure of maximal entropy. By the Abramov formula, $\sigma : \widetilde{\Sigma} \to \widetilde{\Sigma}$ has a measure of maximal entropy, and the value of this entropy is $hc$. Gurevich characterized the countable state topological Markov shifts which possess measures of maximal entropy [Gur69, Gur70]. His work shows that there are $p \in \mathbb{N}$, $C > 0$, and a vertex $v$ s.t. $\#\{\underline{x} \in \widetilde{\Sigma} : x_0 = v, \sigma^{np}(\underline{x}) = \underline{x}\} \asymp e^{np \cdot hc}$. Such $\underline{x}$ determines a simple closed orbit $\gamma^s_{\underline{x},np} : \left[0, \frac{npc}{N(\underline{x},np)}\right] \to M$, where $\gamma^s_{\underline{x},np}(t) = \widetilde{p}[\widetilde{\sigma}_c^t(\underline{x},0)]$ and $N(\underline{x},np) := \#\{0 \leq t < npc : \widetilde{p}[\widetilde{\sigma}_c^t(\underline{x},0)] = \widetilde{p}[(\underline{x},0)]\}$. We therefore get a map $\Theta : \{(\underline{x},n) : x_0 = v, \sigma^{np}(\underline{x}) = \underline{x}\} \to \{[\gamma] : \gamma \text{ simple s.t. } \ell(\gamma) \leq npc\}$,

$$\Theta(\underline{x},n) := [\gamma^s_{\underline{x},np}].$$

Again $\Theta$ is not one-to-one, but again one can show that $1 \leq \frac{\#\Theta^{-1}([\gamma^s_{\underline{x},np}])}{np} \leq C(v)$, where $C(v) = N(a(v),a(v))$. Thus $\#\{[\gamma^s_{\underline{y},np}] : \underline{y} \in \widetilde{\Sigma}, y_0 = v, \sigma^{np}(\underline{y}) = \underline{y}\} \asymp \frac{e^{np \cdot hc}}{n}$. Since $\ell(\gamma^s_{\underline{y},np}) \leq npc$, $\#\{[\gamma] : \gamma \text{ is simple s.t. } \ell(\gamma) \leq T_n\} \geq \text{const} \times \frac{e^{hT_n}}{T_n}$ for $T_n = npc$. It follows that $\#\{[\gamma] : \gamma \text{ is simple s.t. } \ell(\gamma) \leq T\} \geq \text{const} \times \frac{e^{hT}}{T}$ for large $T$. $\square$

## Appendix A: Standard proofs

**Proof of Lemma 2.1.** $M$ is closed (compact and boundaryless) and smooth, so there is a constant $r_{\text{inj}} > 0$ s.t. for every $p \in M$, $\exp_p : \{\vec{v} \in T_pM : \|\vec{v}\|_p \leq r_{\text{inj}}\} \to M$ is $\sqrt{2}$–bi-Lipschitz onto its image (see e.g. [Spi79], chapter 9). Fix $0 < r < r_{\text{inj}}$, and complete $\vec{n}_p := \frac{X_p}{\|X_p\|}$ to an orthonormal basis $\{\vec{n}_p, \vec{u}_p, \vec{v}_p\}$ of $T_pM$. Then

$$J_p(x,y) := \exp_p(x\vec{u}_p + y\vec{v}_p)$$

is a $C^\infty$ diffeomorphism from $U_r := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq r\}$ onto $S_r(p)$ for all $0 < r < r_{\text{inj}}$, proving that $S = S_r(p)$ is a $C^\infty$ embedded disc.

We claim that $\text{dist}_M(\cdot,\cdot) \leq \text{dist}_S(\cdot,\cdot) \leq 2\,\text{dist}_M(\cdot,\cdot)$. The first inequality is obvious. For the second, suppose $z_1, z_2 \in S$. There are $\vec{v}_1, \vec{v}_2 \perp X_p$ s.t. $\|\vec{v}_i\|_p \leq r$ and $z_i = \exp_p(\vec{v}_i)$. Let $\gamma(t) := \exp_p[t\vec{v}_2 + (1-t)\vec{v}_1]$, $t \in [0,1]$. Clearly $\gamma \subset S$, whence

$\mathrm{dist}_S(z_1,z_2) \leq \mathrm{length}(\gamma)$. Since $\exp_p$ has bi-Lipschitz constant $\sqrt{2}$, $\mathrm{dist}_S(z_1,z_2) \leq \sqrt{2}\|\vec{v}_1 - \vec{v}_2\|_p \leq (\sqrt{2})^2\,\mathrm{dist}_M(z_1,z_2)$.

We bound $\measuredangle(X_q, T_qS)$ for $q \in S$. If $\vec{u},\vec{v},\vec{w} \in \mathbb{R}^3 \setminus \{\underline{0}\}$, then $|\measuredangle(\vec{u}, \mathrm{span}\{\vec{v},\vec{w}\})| \geq |\sin\measuredangle(\vec{u}, \mathrm{span}\{\vec{v},\vec{w}\})| \geq \left|\frac{\langle \vec{u},\vec{v},\vec{w}\rangle}{\|\vec{u}\|\cdot\|\vec{v}\|\cdot\|\vec{w}\|}\right|$, where $\langle \vec{u},\vec{v},\vec{w}\rangle$ is the signed volume of the parallelepiped with sides $\vec{u},\vec{v},\vec{w}$. So for every $q \in S_r(p)$,

$$|\measuredangle(X_q, T_qS)| \geq A(p,\underline{x}) := \left| \frac{\left\langle X_{J_p(\underline{x})}, (dJ_p)_{\underline{x}}\frac{\partial}{\partial x}, (dJ_p)_{\underline{x}}\frac{\partial}{\partial y}\right\rangle_{J_p(\underline{x})}}{\|X_{J_p(\underline{x})}\|_{J_p(\underline{x})} \cdot \|(dJ_p)_{\underline{x}}\frac{\partial}{\partial x}\|_{J_p(\underline{x})} \cdot \|(dJ_p)_{\underline{x}}\frac{\partial}{\partial y}\|_{J_p(\underline{x})}} \right|,$$

where $\underline{x} = \underline{x}(q)$ is characterized by $q = J_p(\underline{x})$. By definition $A(p,\underline{0}) = 1$, so there is an open neighborhood $V_p$ of $p$ and $\delta_p > 0$ s.t. $A(q,\underline{x}) > \frac{1}{2}$ on $W_p := V_p \times B_{\delta_p}(\underline{0})$. Working in $M \times \mathbb{R}^3$, we cover $K := M \times \{\underline{0}\}$ by a finite collection $\{W_{p_1}, \ldots, W_{p_N}\}$, and let $r_{\mathrm{leb}}$ be a Lebesgue number. Then $A(p,\underline{x}) > \frac{1}{2}$ for every $p \in M$ and $\|\underline{x}\| < r_{\mathrm{leb}}$. The lemma follows with $\mathfrak{r}_s := \frac{1}{2}\min\{1, r_{\mathrm{inj}}, r_{\mathrm{leb}}\}$. $\square$

**Uniform Inverse Function Theorem.** *Let $F : U \to V$ be a differentiable map between two open subsets of $\mathbb{R}^d$ s.t. $\det(dF_{\underline{x}}) \neq 0$ for all $\underline{x} \in U$. Suppose there are $K, H, \beta$ s.t. $\|dF_{\underline{x}}\|, \|(dF_{\underline{x}})^{-1}\| \leq K$ and $\|dF_{\underline{x}_1} - dF_{\underline{x}_2}\| \leq H\|\underline{x}_1 - \underline{x}_2\|^\beta$ for all $\underline{x}_1, \underline{x}_2 \in U$. If $\underline{x} \in U$, $B_\varepsilon(\underline{x}) \subset U$, and $0 < \varepsilon < 2^{-\frac{\beta+1}{\beta}}(KH)^{-\frac{1}{\beta}}$, then:*

(1) *$F^{-1}$ is a well-defined differentiable open map on $W := B_\delta(F(\underline{x}))$, $\delta := \frac{\varepsilon}{2K}$.*
(2) *$\|(dF^{-1})_{\underline{y}_1} - (dF^{-1})_{\underline{y}_2}\| \leq H^*\|\underline{y}_1 - \underline{y}_2\|^\beta$ for all $\underline{y}_1, \underline{y}_2 \in W$, with $H^* := K^3H$.*

*Proof.* Track the constants in the fixed point theorem proof of the inverse function theorem (see e.g. [Sma74]). $\square$

**Proof of Lemma 2.2.** Let $B := \{\underline{x} \in \mathbb{R}^3 : \|\underline{x}\| < 1\}$.

Let $\mathscr{V}$ be a finite open cover of $M$ such that for every $V \in \mathscr{V}$:

(1) $V = C_V(B)$ where $C_V : B \to V$ is a $C^2$ diffeomorphism.
(2) $C_V$ extends to a bi-Lipschitz $C^2$ map from a neighborhood of $\overline{B}$ onto $\overline{V}$.
(3) $(dC_V)_{\underline{x}}\frac{\partial}{\partial x}, (dC_V)_{\underline{x}}\frac{\partial}{\partial y}, X_{C_V(\underline{x})}$ are linearly independent for $\underline{x} \in \overline{B}$.

Since $M$ is compact and $X$ has no zeroes, $\|X_p\|$ is bounded from below. This, together with the $C^{1+\beta}$ regularity of $X$, implies that $\vec{n}_p := X_p/\|X_p\|$ is Lipschitz on $M$. Apply the Gram-Schmidt procedure to $\vec{n}_{C_V(\underline{x})}, (dC_V)_{\underline{x}}\frac{\partial}{\partial x}$, $(dC_V)_{\underline{x}}\frac{\partial}{\partial y}$ for $\underline{x} \in \overline{B}$. The result is a Lipschitz orthonormal frame $\{\vec{n}_p, \vec{u}_p, \vec{v}_p\}$ for $T_pM$, $p \in \overline{V}$.

For every $p \in \overline{V}$, define the function $F_p(x,y,t) := \varphi^t[\exp_p(x\vec{u}_p + y\vec{v}_p)]$. Then $(dF_p)_{\underline{0}}$ is non-singular for every $p \in \overline{V}$. Since $p \mapsto \det(dF_p)_{\underline{0}}$ is continuous and $\overline{V}$ is compact, $\det(dF_p)_{\underline{0}}$ is bounded away from zero for $p \in \overline{V}$. Since $(p,x,y,t) \mapsto \det(dF_p)_{(x,y,t)}$ is uniformly continuous on $\overline{V} \times \overline{B}$, $\exists \delta(V) > 0$ s.t. $\det(dF_p)_{(x,y,t)}$ is bounded away from zero on $\{(p,x,y,t) : p \in \overline{V}, x^2+y^2 \leq \delta(V)^2, |t| \leq \delta(V)\}$.

Fix $0 < \delta < \min\{\delta(V) : V \in \mathscr{V}\}$ s.t. $\delta < r_{\mathrm{leb}}/2S_0$ where $S_0 := 1 + \max_{p\in M}\|X_p\|$ and $r_{\mathrm{leb}}$ is a Lebesgue number for $\mathscr{V}$. For every $p \in M$, $F_p\big(\{(x,y,t) : x^2+y^2 \leq \delta^2, |t| \leq \delta\}\big) \subset B_{r_{\mathrm{leb}}}(p)$, so $\exists V \in \mathscr{V}$ s.t. $F_p\big(\{(x,y,t) : x^2+y^2 \leq \delta^2, |t| \leq \delta\}\big) \subset V = \mathrm{dom}(C_V^{-1})$. For this $V$,

$$G = G_{p,V} := C_V^{-1} \circ F_p : \{(x,y,t) : x^2+y^2 \leq \delta^2, |t| \leq \delta\} \to \mathbb{R}^3$$

is a well-defined map, with Jacobian uniformly bounded away from zero. A direct calculation shows that $\|dG_{(x,y,t)}\|, \|(dG_{(x,y,t)})^{-1}\|$ and the $\beta$–Hölder norm of $dG$ are uniformly bounded by constants that do not depend on $p, V$.

By the uniform inverse function theorem, for every $0 < \delta' \leq \delta$, the image $G\big(\{(x,y,t) : x^2+y^2 \leq (\delta')^2, |t| \leq \delta'\}\big)$ contains a ball $B^*$ of some fixed radius $\mathfrak{d}'(\delta')$ centered at $C_V^{-1}(p)$, and $G$ can be inverted on $B^*$. So $F_p^{-1}$ is well-defined and smooth on $C_V(B^*)$. Since $C_V$ is bi-Lipschitz, there is a constant $\mathfrak{d}(V,\delta')$ s.t. $C_V(B^*) \supset B_{\mathfrak{d}(V,\delta')}(p)$, so $F_p^{-1}$ is well-defined and smooth on $B_{\mathfrak{d}(V,\delta')}(p)$. The $C^{1+\beta}$ norm of the $F_p^{-1}$ there is uniformly bounded by a constant which only depends on $V$. Thus $(q,t) \mapsto \varphi^t(q)$ can be inverted with bounded $C^{1+\beta}$ norm on $B_{\mathfrak{d}(V,\delta')}(p)$. Let $K(V)$ denote a bound on the Lipschitz constant of the inverse function, and let $\rho(V) := \delta/2K(V)$, then $(q,t) \mapsto \varphi^t(q)$ is a diffeomorphism from $S_{\rho(V)}(p) \times [-\rho(V), \rho(V)]$ onto $\mathrm{FB}_{\rho(V)}(p)$. Let $\mathfrak{r}_f := \min\{\rho(V) : V \in \mathscr{V}\}$, then $(q,t) \mapsto \varphi^t(q)$ is a diffeomorphism from $S_{\mathfrak{r}_f}(q) \times [-\mathfrak{r}_f, \mathfrak{r}_f]$ onto $\mathrm{FB}_{\mathfrak{r}_f}(p)$. The lemma follows with this $\mathfrak{r}_f$, and with $\mathfrak{d} := \min\{\mathfrak{d}(V, \frac{1}{2}\mathfrak{r}_f) : V \in \mathscr{V}\}$. □

**Proof of Lemma 2.3.** We use the notation of the previous proof. Invert the function $F_p(x,y,t) := \varphi^t[\exp_p(x\vec{u}_p + y\vec{v}_p)]$ on $B_{\mathfrak{d}}(p)$:

$$F_p^{-1}(z) = (x_p(z), y_p(z), t_p(z)) \quad (z \in B_{\mathfrak{d}}(p)).$$

By the uniform inverse function theorem, the $C^{1+\beta}$ norm of $G^{-1}$ is bounded by some constant independent of $p, V$. Since $F_p^{-1} = G^{-1} \circ C_V^{-1}$, $C_V$ is bi-Lipschitz, and $\mathscr{V}$ is finite, $x_p(\cdot), y_p(\cdot), t_p(\cdot)$ have uniformly bounded Lipschitz constants (independent of $p$), and the differentials of $x_p, y_p, t_p$ are $\beta$–Hölder with uniformly bounded Hölder constants (independent of $p$). Clearly $\mathfrak{t}_p(z) := t_p(z)$ and $\mathfrak{q}_p(z) := \exp_p[x_p(z)\vec{u}_p + y_p(z)\vec{v}_p]$ are the unique solutions for $z = \varphi^{\mathfrak{t}_p(z)}[\mathfrak{q}_p(z)]$. Thus $\mathfrak{t}_p, \mathfrak{q}_p$ are Lipschitz functions with Lipschitz constant bounded by some $\mathfrak{L}$ independent of $p$, and $C^{1+\beta}$ norm bounded by some $\mathfrak{H}$ independent of $p$. □

**Proof of Lemma 2.7.** Cover $M$ by a finite number of flow boxes $\mathrm{FB}_r(z_i)$ with radius $r$. The union of $S_r(z_i)$ is a Poincaré section, but this section is not necessarily standard, because $S_r(z_i)$ are not necessarily pairwise disjoint. To solve this problem we approximate each $S_r(z_i)$ by a finite "net" of points $z^i_{jk}$, and shift each $z^i_{jk}$ up or down along the flow to points $p^i_{jk} = \varphi^{\theta^i_{jk}}(z^i_{jk})$ in such a way that $S_{R_0}(p^i_{jk})$ are pairwise disjoint for some $R_0 < r$ which is still large enough to ensure that $\bigcup S_{R_0}(z^i_{jk})$ is a Poincaré section.

We begin with the choice of some constants. Let:

- ◦ $h_0 > 0$ small, $K_0 > 1$ large (given to us). Without loss of generality, $0 < h_0 < \mathfrak{r}_f$.
- ◦ $r_{\mathrm{inj}} \in (0,1)$ s.t. $\exp_p : \{\vec{v} \in T_pM : \|\vec{v}\| \leq r_{\mathrm{inj}}\} \to M$ is $\sqrt{2}$–bi-Lipschitz for all $p \in M$.
- ◦ $S_0 := 1 + \max\|X_p\|$ and $\mathfrak{r}, \mathfrak{d}, \mathfrak{L}$ are as in Lemmas 2.1–2.3. Recall that $\mathfrak{r}, \mathfrak{d} \in (0,1)$ and $\mathfrak{L} > 1$.
- ◦ $r_0 := \frac{1}{9}\mathfrak{r}\mathfrak{d}h_0 r_{\mathrm{inj}}/(K_0 + S_0)$. Notice that $r_0 < \frac{1}{9}\mathfrak{r}, \frac{1}{9}\mathfrak{d}, \frac{1}{9}h_0, \frac{1}{9}r_{\mathrm{inj}}$.

By Lemma 2.2 and the compactness of $M$, it is possible to cover $M$ by finitely many flow boxes $\mathrm{FB}_{r_0}(z_1), \ldots, \mathrm{FB}_{r_0}(z_N)$. With this $N$ in mind, let:

- ◦ $\rho_0 := r_0(10K_0S_0N\mathfrak{L})^{-20}$. This is smaller than $r_0$.
- ◦ $R_0 := K_0\rho_0$. This is larger than $\rho_0$, but still much smaller than $r_0$.
- ◦ $\delta_0 := \rho_0/(8\mathfrak{L}^2)$. This is much smaller than $r_0$.

○ $\kappa_0 := \lceil 10^2 K_0 \mathfrak{L}^4 \rceil$, a big integer.

For every $i$, complete $\vec{n}_i := X_{z_i}/\|X_{z_i}\|$ to an orthonormal basis $\{\vec{u}_i, \vec{v}_i, \vec{n}_i\}$ of $T_{z_i}M$, and let $J_i : \mathbb{R}^2 \to M$ be the map

$$J_i(x,y) = \exp_{z_i}(x\vec{u}_i + y\vec{v}_i),$$

then $S_{r_0}(z_i) = J_i\big(\{(x,y) : x^2+y^2 \leq r_0^2\}\big)$. The map $J_i$ is $\sqrt{2}$–bi-Lipschitz, because $r_0 < r_{\rm inj}$. Let $I := \{(j,k) \in \mathbb{Z}^2 : (j\delta_0)^2 + (k\delta_0)^2 \leq r_0^2\}$. Given $1 \leq i \leq N$ and $(j,k) \in I$, define

$$z_{jk}^i := J_i(j\delta_0, k\delta_0).$$

Then $\{z_{jk}^i : (j,k) \in I\}$ is a net of points in $S_{r_0}(z_i)$, and for all $(j,k) \neq (\ell, m)$:

$$\frac{1}{\sqrt{2}} \leq \frac{\mathrm{dist}_M(z_{jk}^i, z_{\ell m}^i)}{\delta_0\sqrt{(j-l)^2+(k-m)^2}} \leq \sqrt{2}. \tag{8.3}$$

We will construct points $p_{jk}^i := \varphi^{\theta_{jk}^i}(z_{jk}^i)$ with $\theta_{jk}^i \in [-r_0, r_0]$ s.t. $S_{R_0}(p_{jk}^i)$ are pairwise disjoint. The following claim will help us prove disjointness.

Claim. *Suppose $p_{jk}^i = \varphi^{\theta_{jk}^i}(z_{jk}^i)$, $p_{\ell m}^i = \varphi^{\theta_{\ell m}^i}(z_{\ell m}^i)$, where $\theta_{jk}^i, \theta_{\ell m}^i \in [-r_0, r_0]$. If $S_{R_0}(p_{jk}^i) \cap S_{R_0}(\varphi^{\tau_1}(z_{\alpha\beta}^\gamma)) \neq \varnothing$ and $S_{R_0}(p_{\ell m}^i) \cap S_{R_0}(\varphi^{\tau_2}(z_{\alpha\beta}^\gamma)) \neq \varnothing$ for the same $z_{\alpha\beta}^\gamma$ and some $\tau_1, \tau_2 \in [-r_0, r_0]$, then* $\max\{|j-\ell|, |k-m|\} < \kappa_0$.

In particular, $S_{R_0}(p_{jk}^i) \cap S_{R_0}(p_{\ell m}^i) \neq \varnothing \Rightarrow \max\{|j-\ell|, |k-m|\} < \kappa_0$ (take $z_{\alpha\beta}^\gamma = z_{\ell m}^i$, $\tau_1 = \tau_2 = \theta_{\ell m}^i$).

*Proof.* $S_{R_0}(p_{jk}^i), S_{R_0}(p_{\ell m}^i), z_{\alpha\beta}^\gamma, \varphi^{\tau_1}(z_{\alpha\beta}^\gamma), \varphi^{\tau_2}(z_{\alpha\beta}^\gamma)$ are all contained in $B_{\mathfrak{d}}(z_i)$:

○ $S_{R_0}(p_{jk}^i) \subset B_{\mathfrak{d}}(z_i)$, because if $q \in S_{R_0}(p_{jk}^i)$ then $\mathrm{dist}_M(q, z_i) \leq \mathrm{dist}_M(q, p_{jk}^i) + \mathrm{dist}_M(p_{jk}^i, z_{jk}^i) + \mathrm{dist}_M(z_{jk}^i, z_i) \leq R_0 + r_0 S_0 + r_0 < \mathfrak{d}$. Similarly, $S_{R_0}(p_{\ell m}^i) \subset B_{\mathfrak{d}}(z_i)$.

○ $z_{\alpha\beta}^\gamma \in B_{\mathfrak{d}}(z_i)$: $\mathrm{dist}_M(z_{\alpha\beta}^\gamma, z_i) \leq \mathrm{dist}_M(z_{\alpha\beta}^\gamma, \varphi^{\tau_1}(z_{\alpha\beta}^\gamma)) + \mathrm{dist}_M(\varphi^{\tau_1}(z_{\alpha\beta}^\gamma), p_{jk}^i) + \mathrm{dist}_M(p_{jk}^i, z_{jk}^i) + \mathrm{dist}_M(z_{jk}^i, z_i) \leq r_0 S_0 + 2R_0 + r_0 S_0 + r_0 < \mathfrak{d}$.

○ $\varphi^{\tau_1}(z_{\alpha\beta}^\gamma) \in B_{\mathfrak{d}}(z_i)$: $\mathrm{dist}_M(\varphi^{\tau_1}(z_{\alpha\beta}^\gamma), z_i) \leq \mathrm{dist}_M(\varphi^{\tau_1}(z_{\alpha\beta}^\gamma), p_{jk}^i) + \mathrm{dist}_M(p_{jk}^i, z_{jk}^i) + \mathrm{dist}_M(z_{jk}^i, z_i) < 2R_0 + r_0 S_0 + r_0 < \mathfrak{d}$. Similarly, $\varphi^{\tau_2}(z_{\alpha\beta}^\gamma) \in B_{\mathfrak{d}}(z_i)$.

By Lemma 2.2, the flow box coordinates $\mathfrak{t}_{z_i}(\cdot), \mathfrak{q}_{z_i}(\cdot)$ of $\varphi^{\tau_1}(z_{\alpha\beta}^\gamma)$, $\varphi^{\tau_2}(z_{\alpha\beta}^\gamma)$, $z_{\alpha\beta}^\gamma$, and of every point in $S_{R_0}(p_{jk}^i), S_{R_0}(p_{\ell m}^i)$ are well-defined.

Recall that $\mathfrak{t}_{z_i}, \mathfrak{q}_{z_i}$ have Lipschitz constants less than $\mathfrak{L}$. In the set of circumstances we consider $\mathrm{dist}_M(p_{jk}^i, \varphi^{\tau_1}(z_{\alpha\beta}^\gamma)) \leq 2R_0$ and $\mathfrak{q}_{z_i}(p_{jk}^i) = z_{jk}^i$, so

$$\mathrm{dist}_M(z_{jk}^i, \mathfrak{q}_{z_i}(z_{\alpha\beta}^\gamma)) = \mathrm{dist}_M(\mathfrak{q}_{z_i}(p_{jk}^i), \mathfrak{q}_{z_i}(\varphi^{\tau_1}(z_{\alpha\beta}^\gamma))) \leq 2\mathfrak{L}R_0.$$

Similarly, $\mathrm{dist}(z_{\ell m}^i, \mathfrak{q}_{z_i}(z_{\alpha\beta}^\gamma)) \leq 2\mathfrak{L}R_0$. It follows that $\mathrm{dist}_M(z_{jk}^i, z_{\ell m}^i) \leq 4\mathfrak{L}R_0$. By (8.3), $\max\{|j-\ell|, |k-m|\} \leq 4\sqrt{2}\mathfrak{L}R_0/\delta_0 = 4\sqrt{2}\mathfrak{L}K_0\rho_0\big/(\rho_0/8\mathfrak{L}^2) < \kappa_0$.

The claim is proved. We proceed to construct by induction $\theta_{jk}^i \in [-r_0, r_0]$ and $p_{jk}^i := \varphi^{\theta_{jk}^i}(z_{jk}^i)$ such that $\{S_{R_0}(p_{jk}^i) : 1 \leq i \leq N, (j,k) \in I\}$ are pairwise disjoint.

Basis of induction: $\exists \theta_{jk}^1 \in [-r_0, r_0]$ *s.t.* $\{S_{R_0}(p_{jk}^1)\}_{(j,k)\in I}$ *are pairwise disjoint.*

*Construction:* Let $\widehat{\sigma} : \{0, 1, \ldots, \kappa_0 - 1\} \times \{0, 1, \ldots, \kappa_0 - 1\} \to \{1, \ldots, \kappa_0^2\}$ be a bijection, and set $\sigma_{jk} := \widehat{\sigma}\big(j \bmod \kappa_0, k \bmod \kappa_0\big)$. This has the effect that

$$0 < \max\{|j-\ell|, |k-m|\} < \kappa_0 \Longrightarrow |\sigma_{jk} - \sigma_{\ell m}| \geq 1.$$

We let $\theta^1_{jk} := 2R_0\mathfrak{L}\sigma_{jk}$ and $p^1_{jk} := \varphi^{\theta^1_{jk}}(z^1_{jk})$. It is easy to check that $0 < \theta^1_{jk} < r_0$. One shows as in the proof of the claim that $S_{R_0}(p^1_{jk}) \subset B_{\mathfrak{d}}(z_1)$, therefore $\mathfrak{t}_{z_1}$ is well-defined on $S_{R_0}(p^1_{jk})$. Since $\mathrm{Lip}(\mathfrak{t}_{z_1}) \leq \mathfrak{L}$ and $\mathfrak{t}_{z_1}(p^1_{jk}) = \theta^1_{jk}$,

$$\mathfrak{t}_{z_1}\big[S_{R_0}(p^1_{jk})\big] \subset \big(\theta^1_{jk} - \mathfrak{L}R_0, \theta^1_{jk} + \mathfrak{L}R_0\big). \tag{8.4}$$

Now suppose $(j,k) \neq (\ell,m)$. If $\max\{|j-\ell|,|k-m|\} \geq \kappa_0$, then $S_{R_0}(p^1_{jk}) \cap S_{R_0}(p^1_{\ell m}) = \varnothing$, because of the claim. If $\max\{|j-\ell|,|k-m|\} < \kappa_0$, then $|\theta^1_{jk} - \theta^1_{\ell m}| \geq 2R_0\mathfrak{L}$. By (8.4), $\mathfrak{t}_{z_1}\big[S_{R_0}(p^1_{jk})\big] \cap \mathfrak{t}_{z_1}\big[S_{R_0}(p^1_{\ell m})\big] = \varnothing$, and again $S_{R_0}(p^1_{jk}) \cap S_{R_0}(p^1_{\ell m}) = \varnothing$.

Induction step: *If $\exists \theta^i_{jk} \in [-r_0, r_0]$ s.t. $\{S_{R_0}(p^i_{jk}) : 1 \leq i \leq n, (j,k) \in I\}$ are pairwise disjoint, then $\exists \theta^i_{jk} \in [-r_0, r_0]$ s.t. $\{S_{R_0}(p^i_{jk}) : 1 \leq i \leq n+1, (j,k) \in I\}$ are pairwise disjoint.*

Fix $(j,k) \in I$. We divide $\{p^i_{\ell,m} : 1 \leq i \leq n, (\ell,m) \in I\}$ into two groups:

- *"Dangerous"* (for $p^{n+1}_{jk}$): $\exists \theta \in [-r_0, r_0]$ s.t. $S_{R_0}(p^i_{\ell m}) \cap S_{R_0}(\varphi^{\theta}(z^{n+1}_{jk})) \neq \varnothing$;
- *"Safe"* (for $p^{n+1}_{jk}$): not dangerous.

Here we employ the terminology "safe" when the induction step follows directly from the basis of induction, and "dangerous" otherwise. Indeed, no matter how we define $\theta^{n+1}_{jk}$, $S_{R_0}(p^{n+1}_{jk}) \cap S_{R_0}(p^i_{\ell m}) = \varnothing$ for all safe $p^i_{\ell m}$. But the dangerous $p^i_{\ell m}$ will introduce constraints on the possible values of $\theta^{n+1}_{jk}$.

By the claim, if $p^i_{\ell_1,m_1}, p^i_{\ell_2,m_2}$ are dangerous for $p^{n+1}_{jk}$, then $|\ell_1 - \ell_2|, |m_1 - m_2| < \kappa_0$. It follows that there are at most $4\kappa_0^2 N$ dangerous points for a given $p^{n+1}_{jk}$.

Let $W_{n+1}(p^i_{\ell m}) := \mathfrak{t}_{z_{n+1}}\big[S_{R_0}(p^i_{\ell m})\big]$. Since $\mathrm{Lip}(\mathfrak{t}_{z_{n+1}}) \leq \mathfrak{L}$, $W_{n+1}(p^i_{\ell m})$ is a closed interval of length less than $w_0 := 2\mathfrak{L}R_0$.

If $p^i_{\ell m}$ is dangerous for $p^{n+1}_{jk}$, then we call $W_{n+1}(p^i_{\ell m})$ a "dangerous interval" for $p^{n+1}_{jk}$. Let $W^{n+1}_{jk}$ denote the union of all dangerous intervals for $p^{n+1}_{jk}$, and define

$$T^{n+1}(j,k) := \bigcup_{|j'-j|,|k'-k|<\kappa_0} W^{n+1}_{j'k'}.$$

This is a union of no more than $16\kappa_0^4 N$ intervals of length less than $w_0$ each.

Cut $[-r_0, r_0]$ into four equal "quarters": $Q_1 := [-r_0, -\frac{r_0}{2}], \ldots, Q_4 := [\frac{r_0}{2}, r_0]$. If we subtract $n < L/w$ intervals of length less than $w$ from an interval of length $L$, then the remainder must contain at least one interval of length $(L - nw)/(n+1)$. It follows that for every $s = 1, \ldots, 4$,

$$Q_s \setminus T^{n+1}(\kappa_0\lfloor \tfrac{j}{\kappa_0}\rfloor, \kappa_0\lfloor \tfrac{k}{\kappa_0}\rfloor) \supset \text{ an interval of length } \frac{r_0}{32\kappa_0^4 N + 2} - w_0 \gg 10\kappa_0^2 w_0.$$

Let $\tau^{n+1}(j,k)$ denote the

- center of such an interval in $Q_1$, when $(\lfloor \frac{j}{\kappa_0}\rfloor, \lfloor \frac{k}{\kappa_0}\rfloor) = (0,0) \mod 2$,
- center of such an interval in $Q_2$, when $(\lfloor \frac{j}{\kappa_0}\rfloor, \lfloor \frac{k}{\kappa_0}\rfloor) = (1,0) \mod 2$,
- center of such an interval in $Q_3$, when $(\lfloor \frac{j}{\kappa_0}\rfloor, \lfloor \frac{k}{\kappa_0}\rfloor) = (0,1) \mod 2$,
- center of such an interval in $Q_4$, when $(\lfloor \frac{j}{\kappa_0}\rfloor, \lfloor \frac{k}{\kappa_0}\rfloor) = (1,1) \mod 2$.

Define $\theta^{n+1}_{jk} := \tau^{n+1}\big(\kappa_0\lfloor \frac{j}{\kappa_0}\rfloor, \kappa_0\lfloor \frac{k}{\kappa_0}\rfloor\big) + 3w_0\sigma_{jk}$. This belongs to $[-r_0, r_0]$, because $\tau^{n+1}\big(\kappa_0\lfloor \frac{j}{\kappa_0}\rfloor, \kappa_0\lfloor \frac{k}{\kappa_0}\rfloor\big)$ is the center of an interval in $Q_s$ of radius at least $5\kappa_0^2 w_0 > 3w_0\sigma_{jk}$, and $Q_s \subset [-r_0, r_0]$. Moreover, since $\mathrm{Lip}(\mathfrak{t}_{z_{n+1}}) \leq \mathfrak{L}$ and $w_0 =$

$2\mathfrak{L}R_0$, we have $\mathfrak{t}_{z_{n+1}}[S_{R_0}(p_{jk}^{n+1})] \subset \left[\theta_{jk}^{n+1} - w_0, \theta_{jk}^{n+1} + w_0\right]$, which by the definition of $\tau^{n+1}(j,k)$, lies outside $T^{n+1}(\kappa_0\lfloor \frac{j}{\kappa_0}\rfloor, \kappa_0\lfloor\frac{k}{\kappa_0}\rfloor)$. Thus

$$\mathfrak{t}_{z_{n+1}}[S_{R_0}(p_{jk}^{n+1})] \subset [-r_0, r_0] \setminus T^{n+1}(\kappa_0\lfloor \tfrac{j}{\kappa_0}\rfloor, \kappa_0\lfloor\tfrac{k}{\kappa_0}\rfloor). \tag{8.5}$$

We use this to show that $S_{R_0}(p_{jk}^{n+1}) \cap S_{R_0}(p_{\ell m}^i) = \varnothing$ for $(\ell, m) \in I$, $i \le n$. If $p_{\ell m}^i$ is safe for $p_{jk}^{n+1}$, then there is nothing to prove. If it is dangerous, $\mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{\ell m}^i)\right] \subset W_{jk}^{n+1} \subset T^{n+1}\big(\kappa_0\lfloor \frac{j}{\kappa_0}\rfloor, \kappa_0\lfloor\frac{k}{\kappa_0}\rfloor\big)$. By (8.5), $S_{R_0}(p_{jk}^{n+1}) \cap S_{R_0}(p_{\ell m}^i) = \varnothing$.

Next we show that $S_{R_0}(p_{jk}^{n+1})$ is disjoint from every $S_{R_0}(p_{\ell m}^{n+1})$ s.t. $(\ell, m) \neq (j,k)$. There are three cases:

- $\max\{|j-\ell|, |k-m|\} \ge \kappa_0$: use the claim.
- $0 < \max\{|j-\ell|, |k-m|\} < \kappa_0$ and $(\lfloor\frac{j}{\kappa_0}\rfloor, \lfloor\frac{k}{\kappa_0}\rfloor) = (\lfloor\frac{\ell}{\kappa_0}\rfloor, \lfloor\frac{m}{\kappa_0}\rfloor)$: in this case $|\theta_{jk}^{n+1} - \theta_{\ell m}^{n+1}| \ge 3w_0$. Since $\mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{jk}^{n+1})\right] \subset [\theta_{jk}^{n+1} - w_0, \theta_{jk}^{n+1} + w_0]$ and $\mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{\ell m}^{n+1})\right] \subset [\theta_{\ell m}^{n+1} - w_0, \theta_{\ell m}^{n+1} + w_0]$, $\mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{jk}^{n+1})\right] \cap \mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{\ell m}^{n+1})\right] = \varnothing$. So $S_{R_0}(p_{jk}^{n+1}) \cap S_{R_0}(p_{\ell m}^{n+1}) = \varnothing$.
- $0 < \max\{|j-\ell|, |k-m|\} < \kappa_0$ and $(\lfloor\frac{j}{\kappa_0}\rfloor, \lfloor\frac{k}{\kappa_0}\rfloor) \neq (\lfloor\frac{\ell}{\kappa_0}\rfloor, \lfloor\frac{m}{\kappa_0}\rfloor)$: in this case
$$\max\left\{\left|\lfloor\tfrac{j}{\kappa_0}\rfloor - \lfloor\tfrac{\ell}{\kappa_0}\rfloor\right|, \left|\lfloor\tfrac{k}{\kappa_0}\rfloor - \lfloor\tfrac{m}{\kappa_0}\rfloor\right|\right\} = 1,$$
so $\tau^{n+1}\big(\kappa_0\lfloor\frac{j}{\kappa_0}\rfloor, \kappa_0\lfloor\frac{k}{\kappa_0}\rfloor\big)$, $\tau^{n+1}\big(\kappa_0\lfloor\frac{\ell}{\kappa_0}\rfloor, \kappa_0\lfloor\frac{m}{\kappa_0}\rfloor\big)$ fall in different $Q_s$. Necessarily $\mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{jk}^{n+1})\right] \cap \mathfrak{t}_{z_{n+1}}\left[S_{R_0}(p_{\ell m}^{n+1})\right] = \varnothing$, so $S_{R_0}(p_{jk}^{n+1}) \cap S_{R_0}(p_{\ell m}^{n+1}) = \varnothing$.

This concludes the inductive step, and the construction of $\theta_{jk}^i$.

Completion of the proof: *For every $r \in [\rho_0, R_0]$, $\Lambda_r := \biguplus_{i=1}^N \biguplus_{(j,k)\in I} S_r(p_{jk}^i)$ is a standard Poincaré section with roof function bounded above by $h_0$.*

We saw that the union is disjoint for $r = R_0$, therefore it is disjoint for all $r \le R_0$. We will show that the union is a Poincaré section with roof function bounded by $h_0$ for $r = \rho_0$, and then this statement will follow for all $r \ge \rho_0$.

Given $p \in M$, we must find $0 < R < h_0$ s.t. $\varphi^R(p) \in \Lambda_{\rho_0}$. Since $M \subset \bigcup_{i=1}^N \mathrm{FB}_{r_0}(z_i)$, $\exists i$ s.t. $\varphi^{4r_0}(p) \in \mathrm{FB}_{r_0}(z_i)$. Therefore $\varphi^{4r_0}(p) = \varphi^t(z)$ for some $z \in S_{r_0}(z_i)$, $|t| < r_0$, whence $\varphi^{4r_0 - t}(p) \in S_{r_0}(z_i)$.

Write $\varphi^{4r_0-t}(p) = J_i(x,y)$ for some $(x,y)$ s.t. $x^2 + y^2 \le r_0^2$, and choose $(j,k) \in I$ s.t. $|x - j\delta_0|, |y - k\delta_0| < \delta_0$. Since $J_i$ is $\sqrt{2}$–bi-Lipschitz, $\mathrm{dist}_M(\varphi^{4r_0-t}(p), z_{jk}^i) < 2\delta_0$. It follows that $\mathrm{dist}_M(\varphi^{4r_0-t}(p), p_{jk}^i) < 2\delta_0 + r_0 S_0 < \mathfrak{d}$. This places $\varphi^{4r_0-t}(p)$ inside $\mathrm{FB}_{\mathfrak{r}_f}(p_{jk}^i)$. Let $\mathrm{dist}_S$ denote the intrinsic distance on $S_{\mathfrak{r}_s}(p_{jk}^i)$. We have $\mathrm{dist}_S \le 2\,\mathrm{dist}_M$ (see Lemma 2.1), therefore, since $p_{jk}^i = \mathfrak{q}_{p_{jk}^i}(z_{jk}^i)$,

$$\begin{aligned}
&\mathrm{dist}_S(\mathfrak{q}_{p_{jk}^i}(\varphi^{4r_0-t}(p)), p_{jk}^i) = \mathrm{dist}_S(\mathfrak{q}_{p_{jk}^i}(\varphi^{4r_0-t}(p)), \mathfrak{q}_{p_{jk}^i}(z_{jk}^i)) \le \\
&\le 2\,\mathrm{dist}_M(\mathfrak{q}_{p_{jk}^i}(\varphi^{4r_0-t}(p)), \mathfrak{q}_{p_{jk}^i}(z_{jk}^i)) \le 2\mathfrak{L}\,\mathrm{dist}_M(\varphi^{4r_0-t}(p), z_{jk}^i) < 4\mathfrak{L}\delta_0 < \rho_0.
\end{aligned}$$

Thus $\varphi^R(p) \in S_{\rho_0}(p_{jk}^i) \subset \Lambda_{\rho_0}$ for $R := 4r_0 - t - \mathfrak{t}_{p_{jk}^i}[\varphi^{4r_0-t}(p)]$.

Now $|\mathfrak{t}_{p_{jk}^i}[\varphi^{4r_0-t}(p)]| \le 2r_0$, because $|\theta_{jk}^i| \le r_0$ and $|\mathfrak{t}_{p_{jk}^i}[\varphi^{4r_0-t}(p)] + \theta_{jk}^i| = |\mathfrak{t}_{p_{jk}^i}[\varphi^{4r_0-t}(p)] - \mathfrak{t}_{p_{jk}^i}[z_{jk}^i]| \le \mathfrak{L}\,\mathrm{dist}_M(\varphi^{4r_0-t}(p), z_{jk}^i) < 2\mathfrak{L}\delta_0 < r_0$. Also $|t| < r_0$. So $r_0 < R < 7r_0$. Since $r_0 < \frac{1}{9}h_0$, we conclude that $0 < R < h_0$.

**Proof of Theorem 5.6(5).** The proof is motivated by [Bow78]. Say that $R, R' \in \mathscr{R}$ are *affiliated*, if there are $Z, Z' \in \mathscr{Z}$ s.t. $R \subset Z$, $R' \subset Z'$, and $Z \cap Z' \neq \varnothing$. Let $N(R,S) := N(R)N(S)$, where

$$N(R) := \#\{(R', v') \in \mathscr{R} \times \mathscr{A} : R' \text{ is affiliated to } R \text{ and } Z(v') \supset R'\}.$$

This is finite, because of the local finiteness of $\mathscr{Z}$. Let $x = \pi(\underline{R})$ where $R_i = R$ for infinitely many $i < 0$ and $R_i = S$ for infinitely many $i > 0$. Let $N := N(R,S)$, and suppose by way of contradiction that $x$ has $N+1$ different pre-images $\underline{R}^{(0)}, \ldots, \underline{R}^{(N)} \in \Sigma^{\#}(\widehat{\mathscr{G}})$, with $\underline{R}^{(0)} = \underline{R}$. Write $\underline{R}^{(j)} = \{R_k^{(j)}\}_{k \in \mathbb{Z}}$. By Lemma 5.4 there are $\underline{v}^{(j)} \in \Sigma(\mathscr{G})$ s.t. for every $n$,

$${}_{-n}[R_{-n}^{(j)}, \ldots, R_n^{(j)}] \subset Z_{-n}(v_{-n}^{(j)}, \ldots, v_n^{(j)}) \text{ and } R_n^{(j)} \subset Z(v_n^{(j)}).$$

For every $j$, $\underline{v}^{(j)} \in \Sigma^{\#}(\mathscr{G})$, because $\underline{R}^{(j)} \in \Sigma^{\#}(\widehat{\mathscr{G}})$ and $\mathscr{Z}$ is locally finite. It follows that $\pi(\underline{v}^{(j)}) \in Z_{-n}(v_{-n}^{(j)}, \ldots, v_n^{(j)})$ for all $n$.[6]

Since $x = \widehat{\pi}(\underline{R}^{(j)}) \in \overline{{}_{-n}[R_{-n}, \ldots, R_n]} \subset \overline{Z_{-n}(v_{-n}^{(j)}, \ldots, v_n^{(j)})}$, and since the diameter of $Z_{-n}(v_{-n}^{(j)}, \ldots, v_n^{(j)})$ tends to zero as $n \to \infty$ by the Hölder continuity of $\pi$, $\pi(\underline{v}^{(j)}) = x$. Thus $Z(v_i^{(0)}), \ldots, Z(v_i^{(N)})$ all intersect (they contain $f^i(x) = \pi[\sigma^i(\underline{v}^{(j)})]$). This and the inclusion $R_i^{(j)} \subset Z(v_i^{(j)})$ give that $R_i^{(0)}, \ldots, R_i^{(N)}$ are affiliated for all $i$.

In particular, if $k, \ell > 0$ satisfy $R_{-k}^{(0)} = R$ and $R_\ell^{(0)} = S$ (there are infinitely many such $k, \ell$), then there are at most $N = N(R)N(S)$ possibilities for the quadruple $(R_{-k}^{(j)}, Z(v_{-k}^{(j)}); R_\ell^{(j)}, Z(v_\ell^{(j)}))$, $j = 0, \ldots, N$. By the pigeonhole principle, there are $0 \le j_1, j_2 \le N$ s.t. $j_1 \neq j_2$ and

$$(R_{-k}^{(j_1)}, v_{-k}^{(j_1)}) = (R_{-k}^{(j_2)}, v_{-k}^{(j_2)}) \text{ and } (R_\ell^{(j_1)}, v_\ell^{(j_1)}) = (R_\ell^{(j_2)}, v_\ell^{(j_2)}).$$

We can also guarantee that

$$(R_{-k}^{(j_1)}, \ldots, R_\ell^{(j_1)}) \neq (R_{-k}^{(j_2)}, \ldots, R_\ell^{(j_2)}).$$

To do this fix in advance some $m$ s.t. $(R_{-m}^{(j)}, \ldots, R_m^{(j)})$ $(j = 0, \ldots, N)$ are all different, and work with $k, \ell > m$.

Now let $\underline{A} := \underline{R}^{(j_1)}$, $\underline{B} := \underline{R}^{(j_2)}$, $\underline{a} := \underline{v}^{(j_1)}$, $\underline{b} := \underline{v}^{(j_2)}$. Write $A_{-k} = B_{-k} =: B$, $A_\ell = B_\ell =: A$, $a_{-k} = b_{-k} =: b$, and $a_\ell = b_\ell =: a$. Choose

$$x_A \in {}_{-k}[A_{-k}, \ldots, A_\ell] \text{ and } x_B \in {}_{-k}[B_{-k}, \ldots, B_\ell]$$

and two points $z_A, z_B$ by the equations

$$f^{-k}(z_A) := [f^{-k}(x_B), f^{-k}(x_A)] \in W^u(f^{-k}(x_B), B) \cap W^s(f^{-k}(x_A), B)$$
$$f^{\ell}(z_B) := [f^{\ell}(x_B), f^{\ell}(x_A)] \in W^u(f^{\ell}(x_B), A) \cap W^s(f^{\ell}(x_A), A).$$

This makes sense, because $f^{-k}(x_A), f^{-k}(x_B) \in B$ and $f^\ell(x_A), f^\ell(x_B) \in A$. One checks using the Markov property of $\mathscr{R}$ that $z_A \in {}_{-k}[A_{-k}, \ldots, A_\ell]$, and $z_B \in {}_{-k}[B_{-k}, \ldots, B_\ell]$. Since $(A_{-k}, \ldots, A_\ell) \neq (B_{-k}, \ldots, B_\ell)$ and the elements of $\mathscr{R}$ are pairwise disjoint, $z_A \neq z_B$. We will obtain the contradiction we are after by showing that $z_A = z_B$.

Since $f^\ell(z_A) \in A_\ell = A \subset Z(a)$ and $f^{-k}(z_B) \in B_{-k} = B \subset Z(b)$, there are $\underline{\alpha}, \underline{\beta} \in \Sigma^{\#}(\mathscr{G})$ s.t. $z_A = \pi(\underline{\alpha}), z_B = \pi(\underline{\beta})$, $\alpha_\ell = a$, $\beta_{-k} = b$. Let $\underline{c} = \{c_i\}_{i \in \mathbb{Z}}$ where

[6] At this point the proof given in [Sar13] has a mistake. There it is claimed that $\pi(\underline{v}^{(j)}) \in Z_{-n}(v_{-n}^{(j)}, \ldots, v_n^{(j)})$ without making the assumption that $\underline{R}^{(j)} \in \Sigma^{\#}(\widehat{\mathscr{G}})$.

$c_i = \beta_i$ for $i \leq -k$, $c_i = a_i$ for $-k < i < \ell$, and $c_i = \alpha_i$ for $i \geq \ell$. This belongs to $\Sigma^{\#}(\mathscr{G})$, because $\underline{\alpha}, \underline{\beta} \in \Sigma^{\#}(\mathscr{G})$ and $\beta_{-k} = b = a_{-k}$ and $a_\ell = a = \alpha_\ell$. We will show that $z_A = \pi(\underline{c}) = z_B$. Write $c_i = \Psi_{x_i}^{p_i^u, p_i^s}$ $(i \in \mathbb{Z})$. By the definition of $z_A, z_B$ and the Markov property, $f^{-k}(z_A), f^{-k}(z_B)$ both belong to $W^u(f^{-k}(x_B), B)$, thus

$$W^u(f^{-k}(z_A), B) = W^u(f^{-k}(z_B), B) = W^u(\pi(\sigma^{-k}\underline{\beta}), B) \subset V^u[(c_i)_{i \leq -k}].$$

It follows that $f^i(z_A), f^i(z_B) \in \Psi_{x_i}([-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2)$ for all $i \leq -k$.

Similarly, $f^\ell(z_A), f^\ell(z_B)$ both belong to $W^s(f^\ell(x_A), A)$, whence

$$W^s(f^\ell(z_A), A) = W^s(f^\ell(z_B), A) = W^s(\pi(\sigma^\ell \underline{\alpha}), A) \subset V^s[(c_i)_{i \geq \ell}].$$

It follows that $f^i(z_A), f^i(z_B) \in \Psi_{x_i}([-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2)$ for all $i \geq \ell$. For $-k < i < \ell$, $f^i(z_A), f^i(z_B) \in A_i \cup B_i \subset Z(a_i) \cup Z(b_i)$. The sets $Z(a_i), Z(b_i)$ intersect, because as we saw above:

- $x = \pi(\underline{a}) \in Z_{-k}(a_{-k}, \ldots, a_\ell)$, whence $f^i(x) \in Z(a_i)$.
- $x = \pi(\underline{b}) \in Z_{-k}(b_{-k}, \ldots, b_\ell)$, whence $f^i(x) \in Z(b_i)$.

By the overlapping charts property of $\mathscr{Z}$ (see §5) and since $a_i = c_i$ for $-k < i < \ell$,

$$Z(a_i) \cup Z(b_i) \subset \Psi_{x_i}([-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2) \text{ for } i = -k+1, \ldots, \ell - 1.$$

In summary, $f^i(z_A), f^i(z_B) \in \Psi_{x_i}([-Q_\varepsilon(x_i), Q_\varepsilon(x_i)]^2)$ for all $i \in \mathbb{Z}$. As shown in the proof of the shadowing lemma (Thm. 4.2), $\underline{c}$ shadows both $z_A$ and $z_B$, whence $z_A = z_B$. $\square$

**Remark.** We take this opportunity to correct a mistake in [Sar13]. Theorem 12.8 in [Sar13] (the analogue of the statement we just proved) is stated wrongly as a bound for the number of all pre-images of $x \in \widehat{\pi}[\Sigma^{\#}(\widehat{\mathscr{G}})]$. But what is actually proved there (and all that is needed for the remainder of the paper) is just a bound on the number of pre-images which belong to $\Sigma^{\#}(\widehat{\mathscr{G}})$ (denoted there by $\Sigma_\chi^{\#}$). Thus the statements of Theorems 1.3 and 1.4 in [Sar13] should be read as bounds on the number of pre-images in $\Sigma_\chi^{\#}$ (denoted here by $\Sigma^{\#}(\widehat{\mathscr{G}})$), and not as bounds on the number of pre-images in $\Sigma_\chi$ (denoted here by $\Sigma(\widehat{\mathscr{G}})$). The other results or proofs in [Sar13] are not affected by these changes, since $\Sigma_\chi \setminus \Sigma_\chi^{\#}$ does not contain any periodic orbits, and because $\Sigma_\chi \setminus \Sigma_\chi^{\#}$ has zero measure for every shift invariant probability measure (Poincaré recurrence theorem).

**Proof of Lemma 5.8.** Let $\psi : \Sigma_1 \to \Sigma_1$ be the *constant* suspension flow, then:

- For every horizontal segment $[z,w]_h$, $|\tau| < 1 \Longrightarrow \left| \frac{\ell([\psi^\tau(z), \psi^\tau(w)]_h)}{\ell([z,w]_h)} - 1 \right| \leq 2e^2|\tau|$. This uses the trivial bound $d(\underline{x}, \underline{y})/d(\sigma^k(\underline{x}), \sigma^k(\underline{y})) \in [e^{-1}, e]$ for $|k| \leq 1$ and the metric $d(\underline{x}, \underline{y}) := \exp[-\min\{|n| : x_n \neq y_n\}]$.
- For every vertical segment $[z,w]_v$, $\ell([\psi^\tau(z), \psi^\tau(w)]_v) = \ell([z,w]_v)$ for all $\tau$.
- Thus for all $z, w \in \Sigma_1$, $|\tau| < 1 \Longrightarrow (1 + 2e^2|\tau|)^{-1} \leq \frac{d_1(\psi^\tau(z), \psi^\tau(w))}{d_1(z,w)} \leq (1 + 2e^2|\tau|)$.

CLAIM: *$d_r$ is a metric on $\Sigma_r$.*

*Proof.* It is enough to show that $d_1$ is a metric. Symmetry and the triangle inequality are obvious; we show that $d_1(z,w) = 0 \Rightarrow z = w$. Let $z = (\underline{x}, t)$, $w = (\underline{y}, s)$, $\tau := \frac{1}{2} - t$. If $d_1(z,w) = 0$, then $d_1(\psi^\tau(z), \psi^\tau(w)) = 0$. Let $\gamma = (z_0, z_1, \ldots, z_n)$ be a basic path from $\psi^\tau(z)$ to $\psi^\tau(w)$ with length less than $\varepsilon$, with $\varepsilon < \frac{1}{3}$ fixed but arbitrarily small. Write $z_i = (\underline{x}_i, t_i)$, then $\psi^\tau(z) = (\underline{x}_0, t_0)$ and $\psi^\tau(w) = (\underline{x}_n, t_n)$.

Since the lengths of the vertical segments of $\widetilde{\gamma}$ add up to less than $\varepsilon$ and $t_0 = \frac{1}{2}$, $\widetilde{\gamma}$ does not leave $\Sigma\times[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]$. It follows that $|t_n - t_0| < \varepsilon$. Since $\varepsilon$ was arbitrary, $t_n = t_0$, and $\psi^\tau(z)$, $\psi^\tau(w)$ have the same second coordinate.

Since $\widetilde{\gamma}$ does not leave $\Sigma \times [\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]$, it does not cross $\Sigma \times \{0\}$. Writing a list of the horizontal segments $[(\underline{x}_{i_k}, t_{i_k}), (\underline{x}_{i_k+1}, t_{i_k+1})]_h$, we find that $\underline{x}_{i_k+1} = \underline{x}_{i_{k+1}}$. By the triangle inequality $\varepsilon > d_1(\psi^\tau(\underline{x}, t), \psi^\tau(\underline{y}, t)) \geq e^{-1}\sum d(\underline{x}_{i_k}, \underline{x}_{i_{k+1}}) \geq e^{-1}d(\underline{x}_0, \underline{x}_n)$. Since $\varepsilon$ is arbitrary, $\underline{x}_0 = \underline{x}_n$, and $\psi^\tau(z)$, $\psi^\tau(w)$ have the same first coordinate. Thus $\psi^\tau(z) = \psi^\tau(w)$, whence $z = w$.

PART (1): $d_r((\underline{x}, t), (\underline{y}, s)) \leq \mathrm{const}[d(\underline{x}, \underline{y})^\alpha + |t - s|]$, *where $\alpha$ denotes the Hölder exponent of $r$.*

*Proof.* $d_r((\underline{x}, t), (\underline{y}, s)) \equiv d_1((\underline{x}, \frac{t}{r(\underline{x})}), (\underline{y}, \frac{s}{r(\underline{y})}))$. The basic path $(\underline{x}, \frac{t}{r(\underline{x})})$, $(\underline{x}, \frac{s}{r(\underline{y})})$, $(\underline{y}, \frac{s}{r(\underline{y})})$ shows that $d_1((\underline{x}, \frac{t}{r(\underline{x})}), (\underline{y}, \frac{s}{r(\underline{y})})) \leq \left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| + ed(\underline{x}, \underline{y}) \leq \frac{1}{\inf(r)}\big[|t-s| + \mathrm{Höl}_\alpha(r)d(\underline{x}, \underline{y})^\alpha\big] + ed(\underline{x}, \underline{y}) \leq \mathrm{const}[d(\underline{x}, \underline{y})^\alpha + |t-s|]$.

PART (2): *Let $\alpha$ denote the Hölder exponent of $r$. There is a constant $C_2$ which only depends on $r$ s.t. for all $z = (\underline{x}, t)$, $w = (\underline{y}, s)$ in $\Sigma_r$:*

(a) *If* $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| \leq \frac{1}{2}$, *then* $d(\underline{x}, \underline{y}) \leq C_2 d_r(z, w)$ *and* $|s - t| \leq C_2 d_r(z, w)^\alpha$.

(b) *If* $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$, *then* $d(\sigma(\underline{x}), \underline{y}) \leq C_2 d_r(z, w)$ *and* $|t - r(x)|, s \leq C_2 d_r(z, w)$.

*Proof.* These estimates are trivial when $d_r(z, w)$ is bounded away from zero, so it is enough to prove part (2) for $z, w$ s.t. $d_r(z, w) < \varepsilon_0$, with $\varepsilon_0$ a positive constant that will be chosen later.

Suppose $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| < \frac{1}{2}$ and let $\tau := \frac{1}{2} - \frac{t}{r(\underline{x})}$ (a number in $(-\frac{1}{2}, \frac{1}{2}]$), then

$$\begin{aligned} d_r(z, w) &= d_1(\vartheta_r(z), \vartheta_r(w)) \geq (1 + 2e^2|\tau|)^{-1} d_1(\psi^\tau[\vartheta_r(z)], \psi^\tau[\vartheta_r(w)]) \\ &\geq (1 + 2e^2)^{-1} d_1((\underline{x}, \tfrac{1}{2}), (\underline{y}, \tfrac{1}{2} + \delta)), \text{ where } \delta := \tfrac{s}{r(\underline{y})} - \tfrac{t}{r(\underline{x})}. \end{aligned}$$

Notice that $(\underline{y}, \frac{1}{2} + \delta) \in \Sigma_1$, because $|\delta| < \frac{1}{2}$.

Suppose $\varepsilon_0(1 + 2e^2) < \frac{1}{4}$, then $d_1((\underline{x}, \frac{1}{2}), (\underline{y}, \frac{1}{2} + \delta)) < \frac{1}{4}$. The basic paths whose lengths approximate $d_1((\underline{x}, \frac{1}{2}), (\underline{y}, \frac{1}{2} + \delta))$ are not long enough to leave $\Sigma \times [\frac{1}{4}, \frac{3}{4}]$, and they cannot cross $\Sigma \times \{0\}$. For such paths the lengths of the vertical segments add up to at least $\delta$, and the lengths of the horizontal segments add up to at least $e^{-1}d(\underline{x}, \underline{y})$. Since $d_1((\underline{x}, \frac{1}{2}), (\underline{y}, \frac{1}{2} + \delta)) \leq (1 + 2e^2)d_r(z, w)$,

$$d(\underline{x}, \underline{y}) \leq e(1 + 2e^2)d_r(z, w) \text{ and } |\delta| \leq (1 + 2e^2)d_r(z, w).$$

In particular, $d(\underline{x}, \underline{y}) \leq \mathrm{const}\, d_r(z, w)$, and $|s-t| = \left|r(\underline{y})\frac{s}{r(\underline{y})} - r(\underline{x})\frac{t}{r(\underline{x})}\right| \leq \sup(r)|\delta| + |r(\underline{y}) - r(\underline{x})| \leq (1 + 2e^2)\sup(r)d_r(z, w) + \mathrm{Höl}_\alpha(r)d(\underline{x}, \underline{y})^\alpha \leq \mathrm{const}\, d_r(z, w)^\alpha$, where the last inequality uses our estimate for $d(\underline{x}, \underline{y})$ and the finite diameter of $d_r$. This proves part (a) when $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| < \frac{1}{2}$. If $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} = \frac{1}{2}$, repeat the previous argument with $\tau := 0.49 - \frac{t}{r(\underline{x})}$.

For part (b), suppose $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$, and let $\tau := \frac{r(\underline{x}) - t}{r(\underline{x})} + \frac{1}{2}$. Now $\psi^\tau[\vartheta_r(z)] = (\sigma(\underline{x}), \frac{1}{2})$ and $\psi^\tau[\vartheta_r(w)] = (\underline{y}, \frac{1}{2} + \delta')$, where $\delta' := 1 - \big(\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\big)$. As before,

$$d(\sigma(\underline{x}), \underline{y}) \leq e(1 + 2e^2)d_r(z, w) \text{ and } |\delta'| \leq (1 + 2e^2)d_r(z, w).$$

Using $s \leq r(\underline{y})\delta'$, $|r(\underline{x}) - t| \leq r(\underline{x})\delta'$, we see that $s, |t - r(\underline{x})| < (1+2e^2)\sup(r)d_r(z, w)$.

Part (3): *There are constants $C_3 > 0, 0 < \kappa < 1$ which only depend on $r$ s.t. for all $z, w \in \Sigma_r$ and $|\tau| < 1$, $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq C_3 d_r(z,w)^\kappa$.*

*Proof.* We will only discuss the case $\tau > 0$. The case $\tau < 0$ can be handled similarly, or deduced from the following symmetry: Let $\widehat{\Sigma} := \{\widehat{\underline{x}} : \underline{x} \in \Sigma\}$ where $\widehat{x}_i := x_{-i}$, and let $\widehat{r}(\underline{x}) := r(\widehat{\sigma\underline{x}})$ (a function on $\widehat{\Sigma}$). Then $\Theta(\underline{x}, t) = (\widehat{\sigma\underline{x}}, r(\underline{x}) - t)$ is a bi-Lipschitz map from $\Sigma_r$ to $\widehat{\Sigma}_{\widehat{r}}$, and $\Theta \circ \sigma_r^{-\tau} = \sigma_{\widehat{r}}^\tau \circ \Theta$. This symmetry reflects the representation of the flow $\sigma_r^{-t}$ with respect to the Poincaré section $\Sigma \times \{0\}$.

We will construct a constant $C_3'$ s.t. for all $z, w \in \Sigma_r$, if $0 < \tau < \frac{1}{2}\inf(r)$, then $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq C_3' d_r(z,w)^\alpha$. Part (3) follows with $\kappa := \alpha^N$, $C_3 := (C_3')^{\frac{1}{1-\alpha}}$, $N := \lceil 1/\min\{1, \frac{1}{2}\inf(r)\}\rceil$. We will also limit ourselves to the case when $C_2 d_r(z,w) < \frac{1}{2}\inf(r)$; part (3) is trivial when $d_r(z,w)$ is bounded away from zero.

Let $z := (\underline{x}, t), w := (\underline{y}, s)$. Since $\tau > 0$, $\sigma_r^\tau(z) = (\sigma^m(\underline{x}), \varepsilon r(\sigma^m(\underline{x}))$ and $\sigma_r^\tau(w) = (\sigma^n(\underline{y}), \eta r(\sigma^n(\underline{y}))$ where $0 \leq \varepsilon, \eta < 1$ and $m, n \geq 0$. Notice that $m, n \in \{0, 1\}$, (because $0 < \tau < \frac{1}{2}\inf(r)$, so $\sigma_r^t(z), \sigma_r^t(w)$ cannot cross $\Sigma \times \{0\}$ twice).

**Case 1:** $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| \leq \frac{1}{2}$ and $m = n$. Then:

$$
\begin{aligned}
&d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) = d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^m(\underline{y}), \eta)) \\
&\leq d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^m(\underline{y}), \varepsilon)) + d_1((\sigma^m(\underline{y}), \varepsilon), (\sigma^m(\underline{y}), \eta)) \\
&\leq e d(\sigma^m(\underline{x}), \sigma^m(\underline{y})) + |\varepsilon - \eta| \\
&\leq e^2 d(\underline{x}, \underline{y}) + |\varepsilon - \eta| \leq e^2 C_2 d_r(z,w) + |\varepsilon - \eta|, \text{ by part (2)(a).}
\end{aligned}
$$

Since $m = n$, $|\varepsilon - \eta| = \left|\frac{t+\tau-r_m(\underline{x})}{r(\sigma^m(\underline{x}))} - \frac{s+\tau-r_m(\underline{y})}{r(\sigma^m(\underline{y}))}\right| \leq \frac{1}{\inf(r)^2}[I_1 + I_2 + I_3]$, where:

- $I_1 = |tr(\sigma^m(\underline{y})) - sr(\sigma^m(\underline{x}))| \leq t|r(\sigma^m(\underline{x})) - r(\sigma^m(\underline{y}))| + |t-s| r(\sigma^m(\underline{x})) \leq \sup(r)[e^\alpha C_2^\alpha \operatorname{Höl}_\alpha(r) + C_2] d_r(z,w)^\alpha$ by part (2)(a).
- $I_2 = \tau|r(\sigma^m(\underline{x})) - r(\sigma^m(\underline{y}))| \leq e^\alpha C_2^\alpha \inf(r) \operatorname{Höl}_\alpha(r) d_r(z,w)^\alpha$, because $m \leq 1$.
- $I_3 = |r_m(\underline{x}) r(\sigma^m(\underline{y})) - r_m(\underline{y}) r(\sigma^m(\underline{x}))| \leq |r_m(\underline{x}) - r_m(\underline{y})| r(\sigma^m(\underline{y})) + r_m(\underline{y}) \cdot |r(\sigma^m(\underline{x})) - r(\sigma^m(\underline{y}))| \leq \text{const} \operatorname{Höl}_\alpha(r) d_r(z,w)^\alpha$, again because $m \leq 1$.

Thus $|\varepsilon - \eta| \leq \text{const}\, d_r(z,w)^\alpha$, where the constant only depends on $r$. It follows that $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq \text{const}\, d_r(z,w)^\alpha$ where the constant only depends on $r$.

**Case 2:** $\left|\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})}\right| \leq \frac{1}{2}$ and $m \neq n$. We can assume that $n = m+1$, thus:

$$
\begin{aligned}
&d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) = d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^{m+1}(\underline{y}), \eta)) \\
&\leq d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^m(\underline{y}), \varepsilon)) + d_1((\sigma^m(\underline{y}), \varepsilon), (\sigma^{m+1}(\underline{y}), \eta)) \\
&\leq e d(\sigma^m(\underline{x}), \sigma^m(\underline{y})) + 1 - \varepsilon + \eta \leq e^2 C_2 d_r(z,w) + 1 - \varepsilon + \eta, \text{ by part (2)(a).}
\end{aligned}
$$

In our scenario, $t + \tau - r_{m+1}(\underline{x})$ is negative, and $s + \tau - r_{m+1}(\underline{y})$ is non-negative. The distance between these two numbers is bounded by $|t-s| + |r_{m+1}(\underline{x}) - r_{m+1}(\underline{y})|$, whence by $\text{const}\, d_r(z,w)^\alpha$. So $|t+\tau-r_{m+1}(\underline{x})|, |s+\tau-r_{m+1}(\underline{y})| \leq \text{const}\, d_r(z,w)^\alpha$. Since $1 - \varepsilon = \frac{|t+\tau-r_{m+1}(\underline{x})|}{r(\sigma^m \underline{x})}$, $\eta = \frac{s+\tau-r_{m+1}(\underline{y})}{r(\sigma^{m+1}\underline{y})}$, and the denominators are at least $\inf(r)$, there is a constant which only depends on $r$ s.t. $1 - \varepsilon, \eta < \text{const}\, d_r(z,w)^\alpha$. It follows that $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq \text{const}\, d_r(z,w)^\alpha$.

**Case 3:** $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$ and $m = n$. We have:

$$\begin{aligned} & d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) = d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^m(\underline{y}), \eta)) \\ & \leq d_1((\sigma^m(\underline{x}), \varepsilon), (\sigma^{m+1}(\underline{x}), \eta)) + d_1((\sigma^{m+1}(\underline{x}), \eta), (\sigma^m(\underline{y}), \eta)) \\ & \leq 1 - \varepsilon + \eta + e^2 C_2 d_r(z, w), \text{ by part (2)(b), and since } m \leq 1. \end{aligned}$$

Because $t + \tau - r_{m+1}(\underline{x}) < 0 \leq s + \tau - r_m(\underline{y})$, it follows by part (2)(b) that

$$\begin{aligned} & |t + \tau - r_{m+1}(\underline{x})|, |s + \tau - r_m(\underline{y})| \leq |t - s - r_{m+1}(\underline{x}) - r_m(\underline{y})| \\ & \leq |t - r(\underline{x})| + s + |r_m(\sigma(\underline{x})) - r_m(\underline{y})| \leq 2C_2 d_r(z, w) + \text{const}\, d_r(z, w)^\alpha. \end{aligned}$$

As in case 2, this means that $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) < \text{const}\, d_r(z, w)^\alpha$.

**Case 4:** $\frac{t}{r(\underline{x})} - \frac{s}{r(\underline{y})} > \frac{1}{2}$ and $m \neq n$. Recall that $C_2 d_r(z, w), \tau < \frac{1}{2}\inf(r)$. By part (2)(b), $s \leq \frac{1}{2}\inf(r)$, thus $s + \tau < \inf(r)$. Necessarily $n = 0$, $m = 1$, $m = n + 1$, so:

$$\begin{aligned} & d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) = d_1((\sigma^{n+1}(\underline{x}), \varepsilon), (\sigma^n(\underline{y}), \eta)) \\ & \leq d_1((\sigma^{n+1}(\underline{x}), \varepsilon), (\sigma^{n+1}(\underline{x}), \eta)) + d_1((\sigma^{n+1}(\underline{x}), \eta), (\sigma^n(\underline{y}), \eta)) \\ & \leq |\varepsilon - \eta| + e d(\sigma(\underline{x}), \underline{y}) \quad (\because n = 0) \\ & \leq |\varepsilon - \eta| + e C_2 d_r(z, w), \quad \text{by part (2)(b).} \end{aligned}$$

We have $|\varepsilon - \eta| = \left|\frac{t + \tau - r(\underline{x})}{r(\sigma \underline{x})} - \frac{s + \tau}{r(\underline{y})}\right| \leq \frac{1}{\inf(r)^2}[I_1 + I_2]$, where by part (2)(b):

- $I_1 := |[t - r(\underline{x})]r(\underline{y}) - s r(\sigma \underline{x})| \leq 2\sup(r) C_2 d_r(z, w)$.
- $I_2 := \tau |r(\sigma \underline{x}) - r(\underline{y})| \leq \frac{1}{2}\inf(r)\, \text{Höl}(r) C_2^\alpha d_r(z, w)^\alpha$.

It follows that $d_r(\sigma_r^\tau(z), \sigma_r^\tau(w)) \leq \text{const}\, d_r(z, w)^\alpha$ where the constant only depends on $r$. This completes the proof of part (3). □

## 9. Acknowledgements

The authors thank Anatole Katok, François Ledrappier and Federico Rodriguez-Hertz for useful and inspiring discussions. Special thanks go to Dmitry Dolgopyat for showing us a proof that weak mixing ⇒ mixing in Theorem 7.1, Giovanni Forni for suggesting the application to Reeb flows, and Edriss Titi for giving us an elementary proof that $C^{1+\varepsilon}$ vector fields generate flows with $C^{1+\varepsilon}$ time $t$ maps.

## References

[ADU93] Jon Aaronson, Manfred Denker, and Mariusz Urbański. Ergodic theory for Markov fibred systems and parabolic rational maps. *Trans. Amer. Math. Soc.*, 337(2):495–548, 1993.

[AMR88] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, tensor analysis, and applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988.

[AR62] L. M. Abramov and V. A. Rohlin. Entropy of a skew product of mappings with invariant measure. *Vestnik Leningrad. Univ.*, 17(7):5–13, 1962.

[AW67] R. L. Adler and B. Weiss. Entropy, a complete metric invariant for automorphisms of the torus. *Proc. Nat. Acad. Sci. U.S.A.*, 57:1573–1576, 1967.

[AW70] Roy L. Adler and Benjamin Weiss. *Similarity of automorphisms of the torus*. Memoirs of the American Mathematical Society, No. 98. American Mathematical Society, Providence, R.I., 1970.

[BG89] Keith Burns and Marlies Gerber. Real analytic Bernoulli geodesic flows on $S^2$. *Ergodic Theory Dynam. Systems*, 9(1):27–45, 1989.

[Bow70] Rufus Bowen. Markov partitions for Axiom A diffeomorphisms. *Amer. J. Math.*, 92:725–747, 1970.

[Bow73] Rufus Bowen. Symbolic dynamics for hyperbolic flows. *Amer. J. Math.*, 95:429–460, 1973.

[Bow75] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin, 1975.

[Bow78] Rufus Bowen. *On Axiom A diffeomorphisms*. American Mathematical Society, Providence, R.I., 1978. Regional Conference Series in Mathematics, No. 35.

[BP07] Luis Barreira and Yakov Pesin. *Nonuniform hyperbolicity*, volume 115 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. Dynamics of systems with nonzero Lyapunov exponents.

[BR75] Rufus Bowen and David Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29(3):181–202, 1975.

[BS03] Jérôme Buzzi and Omri Sarig. Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps. *Ergodic Theory Dynam. Systems*, 23(5):1383–1400, 2003.

[BW72] Rufus Bowen and Peter Walters. Expansive one-parameter flows. *J. Diff. Equations*, 12:180–193, 1972.

[CBP02] Gonzalo Contreras-Barandiarán and Gabriel P. Paternain. Genericity of geodesic flows with positive topological entropy on $S^2$. *J. Differential Geom.*, 61(1):1–49, 2002.

[Don88] Victor J. Donnay. Geodesic flow on the two-sphere. I. Positive measure entropy. *Ergodic Theory Dynam. Systems*, 8(4):531–553, 1988.

[EM70] David G. Ebin and Jerrold Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)*, 92:102–163, 1970.

[Gur69] B. M. Gurevič. Topological entropy of a countable Markov chain. *Dokl. Akad. Nauk SSSR*, 187:715–718, 1969.

[Gur70] B. M. Gurevič. Shift entropy and Markov measures in the space of paths of a countable graph. *Dokl. Akad. Nauk SSSR*, 192:963–965, 1970.

[Hub59] Heinz Huber. Zur analytischen Theorie hyperbolischen Raumformen und Bewegungsgruppen. *Math. Ann.*, 138:1–26, 1959.

[Hut10] Michael Hutchings. Taubes's proof of the Weinstein conjecture in dimension three. *Bull. Amer. Math. Soc. (N.S.)*, 47(1):73–125, 2010.

[Kal00] Vadim Yu. Kaloshin. Generic diffeomorphisms with superexponential growth of number of periodic orbits. *Comm. Math. Phys.*, 211(1):253–271, 2000.

[Kat80] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.*, (51):137–173, 1980.

[Kat82] A. Katok. Entropy and closed geodesics. *Ergodic Theory Dynam. Systems*, 2(3-4):339–365 (1983), 1982.

[KH95] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[Kni97] G. Knieper. On the asymptotic geometry of nonpositively curved manifolds. *Geom. Funct. Anal.*, 7(4):755–782, 1997.

[Kni02] Gerhard Knieper. Hyperbolic dynamics and Riemannian geometry. In *Handbook of dynamical systems, Vol. 1A*, pages 453–545. North-Holland, Amsterdam, 2002.

[KSLP86] Anatole Katok, Jean-Marie Strelcyn, F. Ledrappier, and F. Przytycki. *Invariant manifolds, entropy and billiards; smooth maps with singularities*, volume 1222 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986.

[KU07] Svetlana Katok and Ilie Ugarcovici. Symbolic dynamics for the modular surface and beyond. *Bull. Amer. Math. Soc. (N.S.)*, 44(1):87–132, 2007.

[KW02] Gerhard Knieper and Howard Weiss. $C^\infty$ genericity of positive topological entropy for geodesic flows on $S^2$. *J. Differential Geom.*, 62(1):127–141, 2002.

[LLS16] François Ledrappier, Yuri Lima, and Omri Sarig. Ergodic properties of equilibrium measures for smooth three dimensional flows. *Comment. Math. Helv.*, 91(1):65–106, 2016.

[LS82] François Ledrappier and Jean-Marie Strelcyn. A proof of the estimation from below in Pesin's entropy formula. *Ergodic Theory Dynam. Systems*, 2(2):203–219 (1983), 1982.

[Mar69] G. A. Margulis. Certain applications of ergodic theory to the investigation of manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 3(4):89–90, 1969.

[MS11] Leonardo Macarini and Felix Schlenk. Positive topological entropy of Reeb flows on spherizations. *Math. Proc. Cambridge Philos. Soc.*, 151(1):103–128, 2011.

[New89] Sheldon E. Newhouse. Continuity properties of entropy. *Ann. of Math. (2)*, 129(2):215–235, 1989.

[Pes76] Ja. B. Pesin. Families of invariant manifolds that correspond to nonzero characteristic exponents. *Izv. Akad. Nauk SSSR Ser. Mat.*, 40(6):1332–1379, 1440, 1976.

[Pla72] Joseph F. Plante. Anosov flows. *Amer. J. Math.*, 94:729–754, 1972.

[PP83] William Parry and Mark Pollicott. An analogue of the prime number theorem for closed orbits of Axiom A flows. *Ann. of Math. (2)*, 118(3):573–591, 1983.

[PP90] William Parry and Mark Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.

[PS98] Mark Pollicott and Richard Sharp. Exponential error terms for growth functions on negatively curved surfaces. *Amer. J. Math.*, 120(5):1019–1042, 1998.

[PS01] Mark Pollicott and Richard Sharp. Error terms for closed orbits of hyperbolic flows. *Ergodic Theory Dynam. Systems*, 21(2):545–562, 2001.

[Rat69] M. E. Ratner. Markov decomposition for an U-flow on a three-dimensional manifold. *Mat. Zametki*, 6:693–704, 1969.

[Rat73] M. Ratner. Markov partitions for Anosov flows on $n$-dimensional manifolds. *Israel J. Math.*, 15:92–114, 1973.

[Rat74] M. Ratner. Anosov flows with Gibbs measures are also Bernoullian. *Israel J. Math.*, 17:380–391, 1974.

[Rat78] M. Ratner. Bernoulli flow over maps of the interval. *Israel J. Math.*, 31(3-4):298–314, 1978.

[Rue78] David Ruelle. An inequality for the entropy of differentiable maps. *Bol. Soc. Brasil. Mat.*, 9(1):83–87, 1978.

[Sar13] Omri M. Sarig. Symbolic dynamics for surface diffeomorphisms with positive entropy. *J. Amer. Math. Soc.*, 26(2):341–426, 2013.

[Ser81] Caroline Series. Symbolic dynamics for geodesic flows. *Acta Math.*, 146(1-2):103–128, 1981.

[Ser87] Caroline Series. Symbolic dynamics for geodesic flows. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, pages 1210–1215, Providence, RI, 1987. Amer. Math. Soc.

[Ser91] Caroline Series. Geometrical methods of symbolic coding. In *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, Oxford Sci. Publ., pages 125–151. Oxford Univ. Press, New York, 1991.

[Sin66] Ja. G. Sinaĭ. Asymptotic behavior of closed geodesics on compact manifolds with negative curvature. *Izv. Akad. Nauk SSSR Ser. Mat.*, 30:1275–1296, 1966.

[Sin68a] Ja. G. Sinaĭ. Construction of Markov partitionings. *Funkcional. Anal. i Priložen.*, 2(3):70–80 (Loose errata), 1968.

[Sin68b] Ja. G. Sinaĭ. Markov partitions and U-diffeomorphisms. *Funkcional. Anal. i Priložen*, 2(1):64–89, 1968.

[Sma74] D. R. Smart. *Fixed point theorems*. Cambridge University Press, London, 1974. Cambridge Tracts in Mathematics, No. 66.

[Spi79] Michael Spivak. *A comprehensive introduction to differential geometry. Vol. I*. Publish or Perish Inc., Wilmington, Del., second edition, 1979.

[Sri98] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[Wal82] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982.

[Wea14] Bryce Weaver. Growth rate of periodic orbits for geodesic flows over surfaces with radially symmetric focusing caps. *J. Mod. Dyn.*, 8(2):139–176, 2014.

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France

*E-mail address*: yurilima@gmail.com

Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, POB 26, Rehovot, Israel
*E-mail address*: omsarig@gmail.com