SRB MEASURES AND YOUNG TOWERS FOR SURFACE DIFFEOMORPHISMS

VAUGHN CLIMENHAGA, STEFANO LUZZATTO, AND YAKOV PESIN

Abstract. We give geometric conditions that are necessary and sufficient for the existence of Sinai–Ruelle–Bowen (SRB) measures for $C^{1+\alpha}$ surface diffeomorphisms, thus proving a version of the Viana conjecture. As part of our argument we give an original method for constructing first return Young towers, proving that every hyperbolic measure, and in particular every SRB measure, can be lifted to such a tower. This method relies on a new general result on hyperbolic branches and shadowing for pseudo-orbits in nonuniformly hyperbolic sets which is of independent interest.

Contents

Date: November 28, 2020.
2010 *Mathematics Subject Classification.* 37D25 (primary); 37C40, 37E30 (secondary).
Key words and phrases. Nonuniform hyperbolicity, Young towers, SRB measures, Pesin theory, decay of correlations.

Part I. Statements of Results

The purpose of this paper is to study the connection between analytic, geometric, dynamical, and statistical properties of surface diffeomorphisms. In particular, we are interested in the way that certain analytic properties, such as hyperbolicity, imply non-trivial geometric structures which in turn produce non-trivial dynamics and statistical behavior. Although we consider only the two-dimensional case, our results on pseudo-orbits and shadowing (Theorems E and F) should extend to higher dimensions as well.

In §1 we discuss the general philosophy and theoretical framework of our study, define *Sinai-Ruelle-Bowen (SRB) measure* and recall the *Viana conjecture* on the existence of SRB measures. In §1.7 we state Theorem A which, roughly speaking, says that under some mild recurrence condition,

a fat (nonuniformly) hyperbolic set supports an SRB measure,

thus proving a version of the Viana conjecture in the two-dimensional setting. We note that, unlike most previous results in this direction, our assumptions are also *necessary*, thus giving an interesting geometric characterization of SRB measures. We give a more detailed review of existing results in §1.8.

Our construction of the SRB measure uses the technique of *Young towers*, which gives additional information about the geometry and structure of the measure. In §2.4 we state Theorem B which, roughly speaking, says that under some mild recurrence condition, a (nonuniformly) hyperbolic set supports a "topological" Young tower and, more specifically,

a fat (nonuniformly) hyperbolic set supports a Young tower.

This result implies Theorem A but is of independent interest. In §3 we state Theorem C, which says that the assumptions of Theorems A and B are necessary for the existence of an SRB measure. The following consequences are worth highlighting here:

- every SRB measure, and more generally every *hyperbolic* measure, is *liftable* to a topological Young tower (Corollary C.1);
- the towers we produce have the first return property (Theorem B);
- we formulate explicit conditions under which the decay rate of the tail of the tower can be controlled (Corollary C.2).

Our construction of a Young tower works in a general setting and differs from other constructions in the literature.[1] The starting point is a measurable subset A of a (non-invariant) "uniformly" hyperbolic set bounded by a *nice domain*. Using an abstract argument we extend A to a *rectangle* Γ – a subset with product structure of local stable and unstable curves – which is maximal in a sense, allowing us to build a tower. The key step in producing Γ is Theorem D in §4, which states that to every *almost return* to A one can associate a *hyperbolic branch*; the total collection of such branches "saturates" A to the desired rectangle Γ. The proof of Theorem D is based on two general results, which we state as Theorems E and F in §5.3 and §5.4 respectively. Theorems D, E, and F are new results in non-uniform hyperbolicity theory

[1] These typically use specific geometric characteristics of the system under consideration.

of independent interest, with Theorem E providing a new version of Katok's closing lemma and Theorem F giving a new version of the shadowing property.

In Part I of the paper we state all our results. In Part II we state and prove Theorem E and F which, as mentioned above, are general results in the theory of nonuniform hyperbolicity. Part III is devoted to the proofs of the remaining results in our more specific setting. These results have a clear logical interdependence as follows:

$$\text{E} \implies \text{D} \implies \text{B} \implies \text{A(1)} \quad \text{and} \quad \text{F} \implies \text{C} \implies \text{A(2)}.$$

The letters above refer to the corresponding Theorems. Theorem A has two parts: A(1) states our sufficient conditions for the existence of an SRB measure, and A(2) states that these conditions are necessary. More details on organization and the relations between the various results are given at the beginning of Parts II and III. In an appendix we provide a list of terminology together with references to the relevant definitions.

See §1.8 and §2.5 for a discussion of related prior work, especially that of Young [69, 70] and Sarig [63].

Acknowledgments. The authors wish to thank Stefano Bianchini who helped keep this project alive with many useful comments and suggestions when we had almost given up, Dima Dolgopyat for pointing out a missing argument in a previous version, and Vilton Pinheiro for helping us fill in the missing argument. V. C. is partially supported by NSF grants DMS-1362838 and DMS-1554794. Ya. P. was partially supported by NSF grant DMS-1400027.

1. SRB measures and the Viana conjecture: Theorem A

Throughout this paper, let M be a surface – by which we mean a compact boundaryless smooth 2-dimensional Riemannian manifold – and let $f\colon M \to M$ be a $C^{1+\alpha}$ diffeomorphism, where $\alpha \in (0,1]$. Let $d(\cdot,\cdot)$ denote the distance function on M, and let m denote Lebesgue measure on M; that is, the area form induced by the Riemannian metric. Given a curve $W \subset M$, we write m_W for the one-dimensional Lebesgue measure on W defined by the induced Riemannian metric. By "measurable" we always mean "Borel measurable".

1.1. Physical measures.

The first step in the statistical description of the diffeomorphism f is the notion of the "statistical basin of attraction" of a probability measure μ:

$$\mathcal{B}_\mu := \left\{ x \in M : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k x) = \int \phi \, d\mu \text{ for all continuous } \phi\colon M \to \mathbb{R} \right\}.$$

Equivalently, $\mathcal{B}_\mu$ consists of all points for which $\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k(x)}$ converges to μ in the weak* topology, where δ_y is the Dirac delta measure on y. If μ is f-invariant and ergodic, then $\mu(\mathcal{B}_\mu) = 1$ by the Birkhoff ergodic theorem, but since Lebesgue measure is the most natural reference measure, we are most interested in finding μ for which $\mathcal{B}_\mu$ is large in the following sense.

Definition 1.1. μ is a *physical measure* for f if $m(\mathcal{B}_\mu) > 0$.

Thus a physical measure is a probability measure which describes the asymptotic statistical behavior of a significant (positive Lebesgue measure) subset of the phase space. Not all dynamical systems admit physical measures,[2] so it is a basic problem to establish the class of dynamical systems which have physical measures before going on to investigate further questions related to the possible number of such measures and their structure and properties.

The simplest example of a physical measure is given by the Dirac-delta measure δ_p at an attracting fixed point p. This easily generalizes to the case when p is an attracting periodic point. At the other extreme, if $\mu \ll m$ is ergodic, then $\mu(\mathcal{B}_\mu) = 1$ gives $m(\mathcal{B}_\mu) > 0$, hence μ is physical. Unfortunately it is relatively rare for such absolutely continuous measures to exist, and thus the problem of the existence of a physical measure is quite non-trivial.

1.2. **Hyperbolic measures.** In the 1970's, Sinai, Ruelle, and Bowen established existence, as well as geometric and statistical properties, of physical measures for *uniformly hyperbolic* systems. The theory of *Sinai–Ruelle–Bowen measures*, or *SRB measures*, has since been extended to non-uniform hyperbolicity, in which setting we need the following definition.

Definition 1.2 (Hyperbolic measures and non-zero Lyapunov exponents)**.** An invariant probability measure μ is *hyperbolic* if there exists a set $\Lambda \subseteq M$ with $f(\Lambda) = \Lambda$ and $\mu(\Lambda) = 1$ which has *non-zero Lyapunov exponents*, i.e. there exists a measurable Df-invariant decomposition $T_xM = E^s_x \oplus E^u_x$ such that for every $x \in \Lambda$ and unit vectors $e^s \in E^s_x, e^u \in E^u_x$ we have:

(1) $\displaystyle\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df^n_x(e^s)\| =: \lambda^s_x < 0 < \lambda^u_x := \lim_{n\to\pm\infty} \frac{1}{n} \log \|Df^n_x(e^u)\|$;

(2) $\displaystyle\lim_{n\to\pm\infty} \frac{1}{n} \log \measuredangle(E^s_{f^n(x)}, E^u_{f^n(x)}) = 0$.

The heart of this definition is that the *Lyapunov exponents* λ^s_x and λ^u_x are *non-zero* and have *opposite signs*;[3] the fact that the limits exist is guaranteed by Oseledets' Multiplicative Ergodic Theorem, which also guarantees that although the angle between the two subspaces is not in general bounded away from zero, it cannot degenerate at an exponential rate along any given orbit, as stated in condition (2). We point out that in the 2-dimensional case the Margulis–Ruelle inequality implies that every measure with positive entropy is hyperbolic.

1.3. **Sinai-Ruelle-Bowen (SRB) measures.** A fundamental and crucial property of sets with non-zero Lyapunov exponents is that every point $x \in \Lambda$ has a *local stable curve* V^s_x and a *local unstable curve* V^u_x satisfying certain

[2]Consider the identity map, or see [46, p. 140] for a more interesting example, sometimes referred to as "Bowen's eye".

[3]If both Lyapunov exponents are negative or both are positive, then it can be shown that the corresponding ergodic component of the measure μ is supported on an attracting or repelling periodic orbit respectively; we exclude this trivial situation.

properties which we describe in Definition 1.10 below.[4] For the moment we use these curves to give the formal definition of SRB measure.

Definition 1.3 (Fat sets)**.** A set $A \subseteq \Lambda$ is *fat* if

$$m\left(\bigcup_{x \in A} V^s_x \right) > 0. \tag{1.1}$$

We can now give the definition of SRB measure that we use in this paper.

Definition 1.4 (SRB measures)**.** An invariant probability measure μ is an *SRB measure* if it is hyperbolic and every set $X \subset \Lambda$ with $\mu(X) = 1$ is fat.

One of the key properties of V^s_x is that $d(f^n(y), f^n(x)) \to 0$ as $n \to \infty$ for every $y \in V^s_x$. This implies that if $x \in \mathcal{B}_\mu$ for some measure μ then also $y \in \mathcal{B}_\mu$ for every $y \in V^s_x$. Therefore the fatness condition (1.1) together with Birkhoff's Ergodic Theorem implies that any ergodic SRB measure is a physical measure.[5] In his plenary lecture at the ICM in Berlin in 1998, Viana formulated the following natural conjecture.

Conjecture (Viana [67])**.** *If a smooth map has only non-zero Lyapunov exponents at Lebesgue almost every point, then it admits some SRB measure.*

Remark 1.5. In the non-uniform hyperbolicity theory a common definition of SRB measure μ requires only that "the conditional measures generated by μ on unstable manifolds are absolutely continuous with respect to the leaf volume on these manifolds" [47, 71]. Such measures may have some zero Lyapunov exponents in directions transversal to unstable manifolds and hence need not be hyperbolic. The advantage of this more general definition is that SRB measures are characterized as the only measures that satisfy the entropy formula. On the other hand, these "non-hyperbolic" SRB measures may not be physical, and some authors adopt a different convention beyond uniform hyperbolicity, in which "SRB measure" simply means "physical measure" [19, Chapter 11]; this appears to be the intent of the Viana conjecture.

We stress that our definition is in fact equivalent to the requirement of being "hyperbolic with absolutely continuous conditionals along unstable manifolds" [66, Theorem C] although we emphasize the fatness condition which is easier to state and is ultimately the crucial property for proving physicality of the measure.

In this paper we use the fatness condition to prove a version of the Viana conjecture for surface diffeomorphisms under the hyperbolicity conditions (**H1**), (**H2**), and (**H3**) (see Definition 1.6 below; these conditions are weaker than the non-zero Lyapunov exponents condition in Definition 1.2) but with the addition of a mild *recurrence* condition. In §§1.4–1.6 we give the exact

[4] In fact, the existence of local stable and unstable curves can be proved under weaker conditions than those of non-zero Lyapunov exponents, see Definition 1.6 and Theorem 1.12.

[5] The converse is not true: for example, if p is a hyperbolic fixed point whose stable and unstable curves form a figure-eight, then δ_p is a hyperbolic physical measure which is not SRB [46, p. 140].

definitions we need for the formal statement of our result in §1.7; then in §1.8 we discuss previous literature on the topic.

1.4. **Hyperbolic sets.** The requirement that the limits in Definition 1.2 exist is somewhat unnatural in a setting where we do not *a priori* have an invariant probability measure, and also obscures a crucial feature of sets with non-zero Lyapunov exponents, which is the fact that the convergence to the limit can be very *non-uniform.* We give here an alternative formulation of hyperbolicity, which is slightly more technical, but is more general and explicit about the intrinsic non-uniformity.

Definition 1.6 (Hyperbolic set)**.** Given $\chi, \epsilon > 0$, we say that an f-invariant measurable set Λ is *(χ, ϵ)-hyperbolic* if there exists a measurable Df-invariant splitting $T_xM = E^s_x \oplus E^u_x$ for all $x \in \Lambda$, and measurable positive functions $C, K\colon \Lambda \to (0, \infty)$ satisfying

$$\textbf{(H1)} \qquad e^{-\epsilon} \leq K(f(x))/K(x) \leq e^{\epsilon} \quad \text{and} \quad e^{-\epsilon} \leq C(f(x))/C(x) \leq e^{\epsilon}$$

such that for every $x \in \Lambda$

$$\textbf{(H2)} \qquad \measuredangle(E^s_x, E^u_x) \geq K(x)$$

and for all unit vectors $e^s_x \in E^s_x, e^u_x \in E^u_x$ and for all $n \geq 1$,

$$\textbf{(H3)} \qquad \begin{aligned} \|Df^n_x(e^s_x)\| &\leq C(x)e^{-\chi n}, & \|Df^n_x(e^u_x)\| &\geq C(x)^{-1}e^{\chi n}, \\ \|Df^{-n}_x(e^s_x)\| &\geq C(x)^{-1}e^{\chi n}, & \|Df^{-n}_x(e^u_x)\| &\leq C(x)e^{-\chi n}. \end{aligned}$$

A set Λ is *χ-hyperbolic* if it is (χ, ϵ)-hyperbolic for all $\epsilon > 0$, and *hyperbolic* if it is a union of χ-hyperbolic sets over all $\chi > 0$.

We will always assume that both E^s_x and E^u_x are non-trivial (hence one-dimensional) and we stress that our definition of hyperbolicity is inherently *non-uniform* and the set Λ is not in general closed. Moreover, observe that if Λ is (χ, ϵ')-hyperbolic for some $0 < \epsilon' < \epsilon$, then it is (χ, ϵ)-hyperbolic.

Remark 1.7. It can be shown that a set Λ with *non-zero Lyapunov exponents* (as in Definition 1.2) is *hyperbolic* (as in Definition 1.6), Indeed, if Λ has non-zero Lyapunov exponents then it is a union of f-invariant sets on which the Lyapunov exponents λ^s, λ^u are uniformly bounded away from 0. Each such set is then (χ, ϵ)-hyperbolic for some $\chi > 0$ and for *every* $\epsilon > 0$, where the functions $K = K_\epsilon$ and $C = C_\epsilon$ clearly depend on ϵ, see [13, §3.3].[6]

A first advantage of formulating hyperbolicity as above is that we can write Λ as a union of nested sets on which we have uniform estimates.

Definition 1.8 (Regular sets)**.** Given a (χ, ϵ)-hyperbolic set Λ, for each $\ell \geq 1$ we define the *regular level set*

$$\Lambda_{\chi,\epsilon,\ell} := \{x \in \Lambda : C(x) \leq e^{\epsilon\ell} \text{ and } K(x) \geq e^{-\epsilon\ell}\}. \tag{1.2}$$

[6]The converse is not true; the limits in the definition of non-zero Lyapunov exponents need not exist at every point (only almost every), even in uniform hyperbolicity. Although existence of these limits is not necessary for our results, the slow variation condition (**H1**) still plays a crucial role in Theorem 1.12, and it seems unlikely that it can be removed.

One can assume without loss of generality that the regular level sets $\Lambda_{\chi,\epsilon,\ell}$ are closed by replacing them with their closures, which still carry the hyperbolic structure in (**H2**) and (**H3**) with $C(x) = e^{\epsilon\ell}$ and $K(x) = e^{-\epsilon\ell}$, and then taking the union over all ℓ to get a (χ,ϵ)-hyperbolic set containing Λ. When the values of χ,ϵ are clear from context, we will often suppress them in the notation and simply write $\Lambda_\ell = \Lambda_{\chi,\epsilon,\ell}$. Clearly $\Lambda_\ell \subseteq \Lambda_{\ell+1} \subseteq \cdots$ and $\Lambda = \bigcup_{\ell\geq 1} \Lambda_\ell$ and, by (**H1**),

$$f^{\pm k}(\Lambda_\ell) \subseteq \Lambda_{\ell+k} \quad \text{for all } \ell, k \in \mathbb{N}. \tag{1.3}$$

Note that for a χ-hyperbolic set Λ the regular sets Λ_ℓ are not defined until we fix a value of $\epsilon > 0$; changing the value of ϵ changes the functions C, K and hence changes the sets Λ_ℓ. On the other hand, we can introduce a useful notation for regular sets independent of an a-priori choice of Λ.

Definition 1.9 ((χ,ϵ,ℓ)-regular sets)**.** A set $\Gamma \subset M$ is *(χ,ϵ,ℓ)-regular* if there exists a (χ,ϵ)-hyperbolic set Λ such that $\Gamma \subset \Lambda_\ell$.

1.5. **Local stable and unstable curves.** One of the fundamental consequences of hyperbolicity is the existence of stable and unstable curves.

Definition 1.10 (Local stable and unstable curves)**.** Given $C, \lambda > 0$, we say that a C^1 curve V^s in M is a *local (C,λ)-stable curve* if for every $y, z \in V^s$ and $n \geq 0$, we have

$$d(f^n(y), f^n(z)) \leq Ce^{-\lambda n} d(y,z) \tag{1.4}$$

and if in addition the curve can be written as $V^s = \exp_x(\gamma)$ for some $x \in M$ and $\gamma \subset T_xM$ satisfying the following conditions:

(1) there is a splitting $T_xM = E \oplus F$, an interval $B \subset E$, and a C^1 function $\psi \colon B \to F$ such that $\gamma := \{v + \psi(v) : v \in B\}$;
(2) γ lies in the ball around the origin of radius $\operatorname{inj}(M)$, the injectivity radius of the manifold.

Replacing f^n with f^{-n} in (1.4) gives the definition of a *local (C,λ)-unstable curve V^u*. We will sometimes omit C, λ from the notation when their precise values are unimportant, and refer simply to *local stable and unstable curves.*

Definition 1.11 (Brackets)**.** A point $z \in M$ is said to be the *bracket* of two points $x, y \in M$ if there is a local stable curve V_x^s containing x and a local unstable curve V_y^u containing y and such that $z \in V_x^s \cap V_y^u$, and if moreover z is the *only* point with this property and the intersection is transversal. In this case we write $[x,y] = z$.

The above definitions are slightly more general than those commonly used in the literature because we make no a priori assumption that the points involved lie in a (χ,ϵ)-hyperbolic set; this makes certain bookkeeping tasks more convenient (see Remark 11.4). With that said, we will still restrict our attention to points satisfying the usual picture provided by the following classical result; see [13, §7.1] or [54] for a more precise and technical statement and the proof.

Theorem 1.12 (Local Stable and Unstable Manifold Theorem)**.** *For all $\chi > \lambda > 0$ there exists $\epsilon_0 > 0$ such that, for any $\epsilon \in (0, \epsilon_0)$ and any $\ell \in \mathbb{N}$ there exist constants $C_\ell, \delta_\ell > 0$ such that if Λ is a (χ, ϵ)-hyperbolic set, then there are families $\{V^s_x \ni x\}_{x \in \Lambda_\ell}$ and $\{V^u_x \ni x\}_{x \in \Lambda_\ell}$ of $C^{1+\text{Hölder}}$ local (C_ℓ, λ)-stable and (C_ℓ, λ)-unstable curves, respectively, such that*

(1) V^s_x and V^u_x depend continuously on $x \in \Lambda_\ell$ in the C^1 topology;
(2) if $x, y \in \Lambda_\ell$ satisfy $d(x, y) < \delta_\ell$, then the bracket $[x, y]$ exists and is the unique point in $V^s_x \cap V^u_y$.

Remark 1.13. In fact Theorem 1.12 follows from Theorem F and Corollary 5.10 below, with the small caveat that there we only guarantee that $V^{s/u}_x$ are local $(C_\ell, \lambda/3)$-stable and unstable curves. The proof of Theorem 1.12 could be modified to improve $\lambda/3$ to λ by being more careful with the bounds in (1.6), (1.7), and (5.9) below, but this is not necessary for our purposes.

1.6. **Definition of constants.** Before proceeding further, we give an explicit bound on how small ϵ must be in the (χ, ϵ)-hyperbolic sets we consider. The precise form of this bound is technical and can be omitted at a first reading; the important thing is that as with ϵ_0 in Theorem 1.12, the quantity ϵ_1 depends only on f, χ, and λ. Fix constants $c_1 < 0 < c_2, c_3$ such that

$$c_1 = -c_2 \le \min_{x\in M}\{\log \|D_x f^{-1}\|^{-1}\} \le \max_{x\in M}\{\log \|D_x f\|\} \le c_2 < \frac{c_3}{1+\alpha}. \tag{1.5}$$

Given $\chi > \lambda > 0$, let ϵ_0 be as in Theorem 1.12. Define the following auxiliary constants:

$$\begin{aligned} \gamma &:= \frac{\chi - c_1}{2\chi}, \quad \beta := \frac{2\chi}{c_3 + \chi}\alpha, \qquad \iota := \frac{2(\chi - \lambda)}{6\epsilon_0\gamma\alpha + (2 + \alpha\beta)c_2 + 2\chi}, \\ &\qquad \eta := 6\gamma\alpha\iota + 2, \qquad \zeta := \alpha\beta\iota. \end{aligned} \tag{1.6}$$

Notice that $\gamma, \eta > 1$ and $\beta, \iota, \zeta \in (0, 1)$. Let $\epsilon_1 = \epsilon_1(f, \chi, \lambda)$ be given by

$$\epsilon_1 := \min\left\{\frac{\lambda\alpha}{18}, \frac{\lambda\beta}{7\gamma}, \frac{\lambda\zeta}{\eta - 1}, \frac{\lambda}{2(1 + 1/\alpha)}, \frac{\lambda^3\alpha}{9(\lambda - 4c_1)(\lambda + 4c_3)}, \epsilon_0\right\}. \tag{1.7}$$

Throughout the paper, we will consider (χ, ϵ)-hyperbolic sets where $\epsilon \in (0, \epsilon_1)$ and ϵ_1 is given by (1.7).

Remark 1.14. The first bound in (1.7) is used in (5.9), while the next three are used in the proof of Theorem H in §8, when we produce a constant $\delta > 0$ such that for every $\ell \in \mathbb{N}$ and every $x, y \in \Lambda_\ell$ with $d(x, y) \le \delta e^{-\lambda\ell}$, the corresponding *Lyapunov charts* (§5.1) are *overlapping* (Definition 8.1). More precisely, the second and third bound in (1.7) are used in Lemmas 8.13 and 8.14, and the fourth is used in (8.45). The fifth bound is used in §11.1 when we prove Theorem B. The last bound, $\epsilon_1 \le \epsilon_0$ is used to guarantee that we can apply the Local Stable and Unstable Manifold Theorem 1.12. The bounds in (1.7) also imply that $\epsilon_1 < \frac{\lambda}{12}$, which is used in (9.4).

Remark 1.15. The precise value of $\lambda \in (0, \chi)$ is not important and the reader wishing to reduce the number of constants may as well consider $\lambda = \chi/2$, although choosing a different value of λ might yield a larger value of ϵ_1.

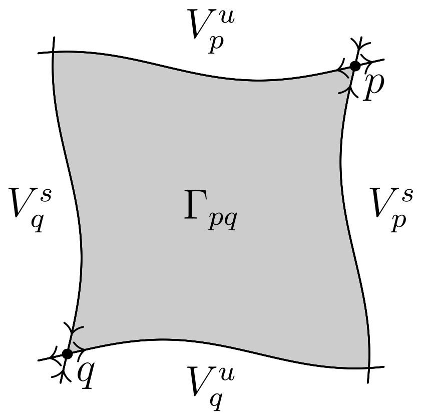

FIGURE 1.1. A nice domain.

1.7. **Nice domains, recurrent sets, and SRB measures.** We are now ready to introduce the key definitions we need to state our main result. To simplify the notation, for a positive integer T we let $T\mathbb{N}$ denote the set of positive integer multiples of T. We also use the notation $V_{p/q}^{s/u}$ to refer simultaneously to $V_p^s, V_q^s, V_p^u, V_q^u$.

Definition 1.16 (Nice Domain)**.** Given $\chi, \epsilon > 0$, $\ell \in \mathbb{N}$, and $r \in (0, \delta_\ell)$, a $(\chi, \epsilon, \ell, r)$-*nice domain* is a topological disk Γ_{pq} satisfying $\operatorname{diam}(\Gamma_{pq}) < r$ whose boundary is formed by (pieces of) the local stable and unstable curves of (χ, ϵ, ℓ)-regular periodic points p, q; see Figure 1.1. We let $T = T(\Gamma_{pq})$ denote the smallest positive integer such that:

(1) T is a common multiple of the periods of p, q, so that $f^T(p) = p$ and $f^T(q) = q$;
(2) T is even, so that all eigenvalues of Df_p^T and Df_q^t are positive; and
(3) T is large enough that $8\sqrt{2}(1 + e^{2(\lambda-\chi)})^{-1/2} e^{2\epsilon\ell} e^{-\lambda T/3} < 1$.

Remark 1.17. The "niceness" condition implies that for all $n \in T\mathbb{N}$,

$$f^n(V_{p/q}^s \cap \Gamma_{pq}) \cap \operatorname{Int}\Gamma_{pq} = \emptyset \quad \text{and} \quad f^{-n}(V_{p/q}^u \cap \Gamma_{pq}) \cap \operatorname{Int}\Gamma_{pq} = \emptyset; \tag{1.8}$$

that is, the stable (respectively, unstable) boundary of Γ_{pq} never intersects the interior of Γ_{pq} at iterates which are multiples of T in forward (respectively, backward) time. This can be thought of as a two-dimensional version of the notion of "nice interval" in one-dimensional dynamics, which refers to an interval whose boundary points never enter the interval in forward time.[7] Here (1.8) will play a crucial role by ensuring that certain regions are necessarily nested or disjoint, see e.g. Remark 4.16 and Lemma 11.11. There is no obvious generalization of this condition to higher dimensions and this is essentially the main reason for which our main results are restricted to surface diffeomorphisms.

[7]After this paper was completed we learned of recent work by Chen, Wang, and Zhang that uses a similar notion for systems with singularities; see Definition 9 in [26, §5.3].

Remark 1.18. The condition that T be even is used in the proof of Lemma 11.11 to guarantee that the 'branches' forming our eventual Markov structure are disjoint. The condition that T be sufficiently large is used in §11.4 to guarantee that the tower we build satisfies the contraction property required in [69].

Definition 1.19 (Nice regular set)**.** Given $\chi, \epsilon > 0$, $\ell \in \mathbb{N}$, and $r \in (0, \delta_\ell]$, a $(\chi, \epsilon, \ell, r)$-*nice regular set* is a set A that is (χ, ϵ, ℓ)-regular and is contained in some $(\chi, \epsilon, \ell, r)$-nice domain Γ_{pq}.

The simplest way to produce a $(\chi, \epsilon, \ell, r)$-nice regular set is to fix a $(\chi, \epsilon, \ell, r)$-nice domain Γ_{pq} and then put $A = \Lambda_{\chi,\epsilon,\ell} \cap \Gamma_{pq}$, but we do not require that the set A has this exact form.

Remark 1.20. The definitions above depend on constants χ, ϵ, ℓ, r which, for simplicity, are not always reflected in the notation; we will sometimes refer to a 'nice domain' or a 'nice regular set' without naming the constants. We stress however that any reference to nice domains or nice regular sets which does not explicitly refer to a choice of constants means that such a choice is implicitly clear from the context. The same applies to the constant T which is always associated to a nice domain and therefore to a nice regular set (as for example in the following definition in which T is implicitly given by the choice of a nice regular set A).

Definition 1.21 (Recurrence)**.** A nice regular set $A \subseteq \Lambda$ is

(1) *recurrent* if for all $x \in A$ there exist $i, j \in T\mathbb{N}$ such that
$$f^i(x), f^{-j}(x) \in A;$$
(2) *Lebesgue-strongly recurrent* if it is recurrent and there is a local unstable curve V^u and a set $E \subset V^u \cap A$ of positive 1-dimensional Lebesgue measure[8] such that for every $x \in E$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \#\{i \in T\mathbb{N}, 1 \le i \le n : f^i(x) \in A\} > 0. \tag{1.9}$$

Observe that a Lebesgue-strongly recurrent set is fat in the sense of (1.1). We are now ready to state our main result.

Theorem A. *Let f be a $C^{1+\alpha}$ surface diffeomorphism.*

(1) *For every $\chi > \lambda > 0$, $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and $\ell \in \mathbb{N}$, there exists $r > 0$ such that if there exists a $(\chi, \epsilon, \ell, r)$-nice regular set A that is Lebesgue-strongly recurrent, then f admits an SRB measure.*
(2) *Conversely, if f admits an SRB measure, then for every sufficiently small $\chi > 0$ and every $\epsilon > 0$, there exists $\ell \in \mathbb{N}$ such that for every $r > 0$, there exists an $(\chi, \epsilon, \ell, r)$-nice regular set A that is Lebesgue-strongly recurrent.*

[8]One could imagine studying other equilibrium measures besides SRB by replacing Lebesgue measure here with a reference measure such as those studied in [**?**, **?**], but we do not pursue this here.

In addition to the explicit bounds in (1.7) on ϵ, we refer to equation (10.1) for relatively explicit conditions on r.

Remark 1.22. Theorem A provides (albeit in the two-dimensional case) natural geometric conditions which are necessary and sufficient for the existence of SRB measures, which are, on the other hand, a highly non-trivial feature of a dynamical system. Part (1) of the theorem can be interpreted as a version of Viana's conjecture. The condition of Lebesgue-strongly recurrent includes a 'fatness' requirement, which is definitely necessary as it is built into the definition of SRB measures (see also Remark 2.12). Beyond this, the theorem requires 'niceness' and '(strong) recurrence'. While our proof uses the niceness property in an essential way, we believe that with some new ideas it should be possible to remove this technical requirement. In addition, part (2) of the theorem indicates that this requirement is not too strong. This leaves us with the requirement of strong recurrence which is not explicit in the statement of Viana's conjecture but we believe is needed to establish existence of an SRB measure. Here again part (2) of the theorem indicates that strong recurrence is not too much to ask for.

We stress that the proof of Theorem A includes a number of new results in non-uniform hyperbolicity theory (such as results about almost returns and shadowing) which are of independent interest. We will discuss these in more details in the next sections.

1.8. **Historical background.** We review here some of the main results on the existence of SRB measures for diffeomorphisms. To avoid getting too technical we will not be overly specific about the precise technical assumptions, emphasizing instead the general ideas. We refer the reader to [33] for more details and a discussion of the various techniques which have been used in different settings. Most results mentioned below hold in arbitrary dimension.

1.8.1. *Uniformly hyperbolic sets.* In the 1970's, Sinai, Ruelle, and Bowen constructed (in fact *invented!*) SRB measures for *fat, uniformly hyperbolic* sets Λ (attractors) under the additional assumption that Λ has a dense set of periodic points (*Axiom A*) or, equivalently, that Λ has local product structure or that Λ is locally maximal (see [65] for a proof that these three properties of uniformly hyperbolic sets are equivalent).

1.8.2. *Partially hyperbolic attractors.* In 1982, Pesin and Sinai [57] developed a new "push-forward" technique for constructing what they called "u-Gibbs" measures (also called simply u-measures), which share a lot of geometric characteristics of SRB measures but are not necessarily physical measures. They applied their construction to *partially hyperbolic* attractors. In 2008, Burns, Dolgopyat, Pesin, and Pollicott [25] showed that under some "transitivity" assumptions, a u-measure that has negative Lyapunov exponents for vectors in the central direction on a set of positive measure is a unique SRB measure. In 2000, Bonatti and Viana [20] considered a variation of this setting with a continuous splitting $E^s \oplus E^{uu}$ with uniform expansion estimates in E^{uu} and nonuniform contraction estimates in E^s, and proved the existence of genuine SRB measures.

Around the same time, Alves, Bonatti, and Viana [1] considered the more difficult setting of a continuous splitting $E^{ss} \oplus E^u$ with uniform contraction and nonuniform expansion estimates. The construction of the SRB measure in this case required a significantly more sophisticated version of the push-forward technique. An alternative construction of the SRB measure using Young towers was carried out more recently in [5] under some slightly weaker expansivity assumptions.

1.8.3. *Nonuniformly hyperbolic sets.* Relaxing the continuity of the splitting $E^s \oplus E^u$ and the uniform lower bound on the angle $\measuredangle(E^s_x, E^u_x)$ seems to bring the level of difficulty of the problem to another level. The first result for a system with this kind of hyperbolic set is due to Benedicks and Young who constructed SRB measures for certain two-dimensional "Hénon" maps, first using the push-forward technique [17] and then using Young towers [69, 18], in both cases taking significant advantage of the specific geometric and analytic properties of the maps. These results were extended to more general "Hénon-like" maps in [68] but still only apply to some quite restrictive classes of systems.

More recently, [30] significantly generalizes the techniques of [1] to successfully construct SRB measures for systems in which the splitting $E^s \oplus E^u$ is only measurable, with angle $\measuredangle(E^s_x, E^u_x)$ not bounded away from zero, and with non-uniform contraction and expansion estimates which, however, need to satisfy a non-trivial "synchronization" assumption.

1.8.4. *Necessary and sufficient conditions.* Theorem A is in some sense an optimal result, because the conditions stated there are both necessary and sufficient for the existence of an SRB measure. In contrast, the prior results mentioned above all established sufficient conditions for existence of SRB measures, without addressing the question of whether the conditions are also necessary.[9]

We mention three notable exceptions, where some set of conditions was demonstrated to be both necessary and sufficient for existence of an SRB measure. First is work of Tsujii [66], in which a point x is called 'regular' if its empirical measures $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k x}$ weak*-converge to an ergodic measure μ whose Lyapunov exponents agree with the exponents of x; then existence of an SRB measure is shown to be equivalent to existence of a positive volume set of regular points with nonzero Lyapunov exponents. More recently, Burguet [23] has proved a similar result, using a different definition of 'regular' that is related to the 'tempered' property, and that does not require one to determine in advance whether empirical measures from a point converge to a given measure.

[9] As we will see in the next section, for surfaces Young's tower conditions from [69] turn out to be necessary as well as sufficient, but this was not proved in that paper.

Finally, we mention a preprint of Snir Ben Ovadia [16] that appeared after our results were completed. This work is most closely related to ours as it establishes existence of an SRB measure under conditions of hyperbolicity, fatness and strong recurrence which are very similar to ours. Moreover, Ben Ovadia's proof works in any dimension, not just for surfaces, and does not require the use of nice domains. However, while our method is pretty much geometrical, his approach is completely different and is based on Markov coding which he developed in his earlier work on countable Markov partitions [14], building on work of Sarig [63]. This Markov coding is in general, non-uniformly bounded-to-one, and thus the tower which can be obtained by inducing on a single state need not be a first T-return Young tower for any T, in contrast to the one we build; see §11.5 for a discussion of how the first return property is related to the *saturation* property of the rectangle that we build as the base of the Young tower. Moreover, as we will see in Corollary C.2, our construction permits us to formulate a condition under which the tower has exponential tails, which appears to have no analogue in the abstract Markov partition approach. We believe that this opens up a possibility to study various properties of SRB measures such as the decay of correlations and the Central Limit Theorem using recently available techniques involving Young towers.

2. Young towers: Theorem B

Our strategy for producing the SRB measure in Theorem A consists of a new technique for the construction of a *Young tower*. The latter has a non-trivial geometric and dynamical structure and it is rather remarkable that its existence can be deduced using only a nice regular set that is Lebesgue-strongly recurrent.

2.1. Rectangles.

Recall that Theorem 1.12 guarantees that given $\chi > \lambda > 0$ and a (χ, ϵ)-hyperbolic set Λ, any two sufficiently close points $x, y \in \Lambda_\ell$ have a unique bracket $[x, y] = V_x^s \cap V_y^u$. This point need not be contained in Λ_ℓ, or even in Λ; for example, this occurs when Λ consists of two hyperbolic periodic orbits that pass sufficiently close to each other.

Nevertheless, we will see in Theorem F and Corollary 5.10 below that $z = [x, y]$ always belongs to *some* hyperbolic set, and indeed more precisely to a $(\lambda/4, 2\epsilon, \ell + \ell')$-regular set, where ℓ' depends only on χ, λ, ϵ. One could then produce local stable and unstable curves V_z^s and V_z^u by applying Theorem 1.12; however, *a priori* these may be shorter than V_x^s and V_y^u. We will generally write $V_z^s = V_x^s$ and $V_z^u = V_y^u$ for the longer curves, even though these may not be "centered" at z. Indeed, we can write $V_z^s = V_x^s$ for any $z \in V_x^s$, and similarly for unstables.

Brackets play a very important role in describing the geometry of hyperbolic sets. In particular we will be interested in sets which are *closed* with respect to the bracket operation.

Definition 2.1 (Rectangles)**.** Given $C, \lambda > 0$, a (C, λ)-*rectangle* is a set Γ such that for every $x, y \in \Gamma$ there is a (C, λ)-local stable curve V_x^s containing

x and a (C,λ)-local unstable curve V_y^u containing y with the property that $[x,y] := V_x^s \cap V_y^u$ is a single point, and $[x,y] \in \Gamma$. When the precise values of C, λ are unimportant, we will refer to Γ simply as a *rectangle.*

Remark 2.2. Rectangles are also sometimes referred to in the literature as sets with *local product structure* or *hyperbolic product structure*, though this is usually in the more restrictive uniformly hyperbolic setting. The discussion above shows that rectangles are very natural structures also in our more general (nonuniformly) hyperbolic setting. Indeed, if A is a (χ,ϵ,ℓ)-regular set of sufficiently small diameter then the bracket $[x,y]$ is well defined and consists of a single point for every pair of points $x,y \in A$. In this case then $[[x,y],[x',y']] = [x,y'] \in \Gamma$ for all $x,y,x',y' \in A$ and therefore the set

$$\Gamma := \{[x,y] : x,y \in A\} \tag{2.1}$$

is closed under the bracket operation, so Γ is a (C_ℓ,λ)-rectangle.

2.2. **Topological Young Towers.** To describe the notion of Young tower that we use, we first recall some standard and some slightly non-standard definitions.

Definition 2.3 (Nice Rectangles)**.** A rectangle Γ is a *nice rectangle* if $\Gamma \subset \Gamma_{pq}$ for some nice domain Γ_{pq} and if the local stable and unstable curves of every point $x \in \Gamma$ are "full length" in Γ_{pq} in the following sense: V_x^u intersects both V_p^s and V_q^s, and V_x^s intersects both V_p^u and V_q^u.

A nice rectangle is not assumed a priori to be a nice regular set (although it is assumed to have stable and unstable curves through every point); the word 'nice' in both cases emphasizes the geometry of the situation. We want to restrict our attention to the pieces of local stable and unstable curves which are inside the nice domain and so, for any $x \in \Gamma \cup \{p,q\}$, let

$$W_x^{s/u} := V_x^{s/u} \cap \Gamma_{pq}.$$

Then a nice rectangle $\Gamma \subset \Gamma_{pq}$ has the structure

$$\Gamma = C^s \cap C^u \quad \text{where} \quad C^s := \bigcup_{x\in\Gamma} W_x^s \quad \text{and} \quad C^u := \bigcup_{x\in\Gamma} W_x^u. \tag{2.2}$$

Definition 2.4 (s-subsets and u-subsets)**.** If Γ is a rectangle, we say that $\Gamma^s \subset \Gamma$ is an *s-subset* of Γ if $x \in \Gamma^s$ implies $V_x^s \cap \Gamma \subset \Gamma^s$ and $\Gamma^u \subset \Gamma$ is a *u-subset* of Γ if $x \in \Gamma^u$ implies $V_x^u \cap \Gamma \subset \Gamma^u$.

Definition 2.5 (T-return times)**.** For a T-recurrent set $A \subset \Gamma_{pq}$ and $x \in A$, let

$$\tau(x) := \min\{i \in T\mathbb{N} : f^i(x) \in A\}$$

be the *first T-return time* to A.

Definition 2.6 (Topological Young Tower)**.** A nice T-recurrent rectangle Γ supports a *First T-Return Topological Young Tower* if for each $i \in T\mathbb{N}$ we can subdivide $\Gamma_i^s := \{x \in \Gamma : \tau(x) = i\}$ into pairwise disjoint s-subsets

$$\Gamma_{i1}^s, \dots, \Gamma_{im_i}^s \quad \text{such that} \quad \Gamma_{ij}^u := f^i(\Gamma_{ij}^s) \tag{Y0}$$

is a u-subset of Γ for every $j = 1, \ldots, m_i$. If Γ_i^s is empty we put $m_i = 0$.

This definition in particular requires that Γ_i^s is an s-subset, so that for each $x \in \Gamma$ with $\tau(x) = i$ we have $f^i(W_x^s \cap \Gamma) \subset W_{f^i x}^s \cap \Gamma$; this Markov condition, and the analogous one for u-subsets, is often formulated explicitly as part of the definition of a Young tower.

Remark 2.7. We remark that the "First T-Return" part of the definition comes from the fact that $\tau(x)$ is the first T-return time. A similar definition can be used for a general return time function that is constant on local stable leaves to give a more general Topological Young Tower.

Remark 2.8. As mentioned in Remark 1.20 above, several constants, including T, are implicit in these definitions. We will sometimes include T explicitly in the associated terminology, for example in the notions of "T-return times" and of "First T-Return" in Definitions 2.5 and 2.6 above, when it helps maintain clarity.

2.3. **Young Towers.** We call the above a *Topological* Young Tower because **(Y0)** captures only the topological structure of a Young tower which is often, including in the present setting, the most difficult part of the construction. To state the other properties of a Young tower we need to introduce the induced map to Γ.

Definition 2.9 (Induced map)**.** For a rectangle Γ which supports a First T-Return Topological Young Tower, we define the *induced map*

$$F\colon \Gamma \to \Gamma \quad \text{by} \quad F|_{\Gamma_i^s} := f^i$$

and refer to Γ as the *base* of the tower. We also let $\mathrm{Jac}^u F(x) := |\det Df^{\tau(x)}|_{E_x^u}|$ denote the unstable Jacobian of this induced map, where $E_x^u = T_x W_x^u$.

Definition 2.10 (Young Tower)**.** Let Γ be a rectangle which supports a First T-Return Topological Young Tower. We say that Γ supports a *First T-Return Young Tower* if there exist constants $\kappa_1, \kappa_2 \in (0,1)$ and $c > 0$ such that

(Y1) for every $i \in T\mathbb{N}$, $j \in \{1, ..., m_i\}$, $x \in \Gamma_{ij}^s$:
  (a) $d(F(x), F(y)) \le \kappa_1 d(x,y)$ for every $y \in V_x^s$
  (b) $d(x,y) \le \kappa_1 d(F(x), F(y))$ for every $y \in V_x^u \cap \Gamma_{ij}^s$.

(Y2) for every $x \in \Gamma$ and $n \ge 0$:
  (a) For all $y \in V_x^s$, we have

$$\left| \log \frac{\mathrm{Jac}^u F(F^n(x))}{\mathrm{Jac}^u F(F^n(y))} \right| \le c\kappa_2^n;$$

  (b) For all $y \in V_x^u$ such that $F^k(x), F^k(y) \in \Gamma_{i_k j_k}^s$ for some $i_k \in T\mathbb{N}, j_k \in \{1, ..., m_{i_k}\}$ for all $0 \le k \le n$, we have

$$\left| \log \frac{\mathrm{Jac}^u F(F^{n-k}(x))}{\mathrm{Jac}^u F(F^{n-k}(y))} \right| \le c\kappa_2^k.$$

Definition 2.11 (Integrable return times)**.** A *fat* rectangle Γ which supports a First T-Return Young Tower has *integrable return times* if there is $x \in \Gamma$ such that

$$\int_{V_x^u \cap \Gamma} \tau \, dm_{V_x^u} < \infty.$$

2.4. **Existence of Young Towers.** Young towers are non-trivial geometric structures and their construction generally requires substantial work. A key part of our argument is to show that they are part of the intrinsic structure of hyperbolic sets.

Theorem B. *Let $f\colon M \to M$ be a $C^{1+\alpha}$ surface diffeomorphism. For every $\chi > \lambda > 0$, every $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and every $\ell \in \mathbb{N}$ there exists $r > 0$ such that if A is a $(\chi, \epsilon, \ell, r)$-nice regular set that is recurrent, then*

1. *A is contained in a nice rectangle $\Gamma \subset \Gamma_{pq}$ that is recurrent and supports a First T-Return Topological Young Tower;*
2. *if the set Γ of part (1) is Lebesgue-strongly recurrent, then Γ supports a First T-Return Young Tower with integrable return times.*

Remark 2.12. Since $A \subset \Gamma$, it follows from Definition 1.21 that if A is Lebesgue-strongly recurrent, then Γ is as well. However, it is possible that the set A is not fat (and hence not Lebesgue-strongly recurrent) but still produces a Lebesgue-strongly recurrent (hence fat) rectangle Γ. In particular the assumptions in part (1) of Theorem A can be relaxed to require that the construction which we will use in the proof of part (1) of Theorem B produces a Lebesgue-strongly recurrent rectangle. A simple setting in which this distinction may be seen to be potentially relevant is that of a two-dimensional Anosov diffeomorphism. Then we can easily find a small nice domain Γ_{pq} and may choose A as a countable dense set of points in Γ_{pq} which means in particular that A is not fat, but we shall see from the proof of Theorem B that this gives rise to a Lebesgue-strongly recurrent rectangle Γ.

Remark 2.13. We will see in the proof (see Proposition 11.3) that the rectangle Γ produced by Theorem B(1) is $(\lambda/4, 2\epsilon, \ell + \ell')$-regular for a value of ℓ' that depends only on χ, λ, ϵ.

2.5. **Historical background.** The Young tower approach for constructing invariant measures is a particular case of a general and classical method in ergodic theory, that of *inducing*. This is based on the construction of a return map to some subset of the phase space which is simpler to study and more amenable to the construction of an invariant measure; this invariant measure can then be extended to the whole phase space by an elementary argument.

The specific inducing structure defined above, which is pertinent to the study of systems with hyperbolic behavior, was introduced by Young [69] as a framework for studying the existence and ergodic properties of SRB measures.[10] Since then, it has been applied to a variety of cases, including billiards [11, 27, 28, 50, 69], certain Hénon maps [18], and partially hyperbolic diffeomorphisms [5, 7, 9, 10]. A non-invertible version of a Young tower was also introduced by Young in [70] and has proved extremely powerful in studying non-invertible maps satisfying nonuniform expansivity conditions [6, 8, 22, 34, 37, 39, 44, 59].

A quite remarkable feature of the method of Young towers is that it associates, by construction, a non-trivial geometric structure to the measure; a structure which moreover can be used effectively to study several statistical and ergodic properties of the measure, see [2, 3, 29, 36, 40, 41, 45, 51, 52, 53, 61, 69, 70] as well as the other references mentioned above. This leads naturally to the question of the domain of applicability of this method: if it implies so much structure then maybe it can only be applied to a limited number of special cases which have this structure? This legitimate doubt is partly supported by the observation that all the constructions of Young towers so far have relied heavily on specific, assumed a-priori, geometric properties of the systems under consideration. This question has therefore led to the so-called *liftability* problem, which is the question of which measures have, or "lift to", a Young tower structure, see Definition 3.1 below, and therefore can in principle be obtained by the method of Young towers.

For the non-invertible non-uniformly expanding case this has been addressed in several papers and in particular it is shown in [4] that essentially every invariant probability measure which is absolutely continuous with respect to Lebesgue lifts to a Young tower. One consequence of our results, stated formally as Corollary C.1 below, is that for surface diffeomorphisms, every hyperbolic measure (in particular, every SRB measure) lifts to a Young tower, which means that the geometric structure of Young towers is intrinsic to all hyperbolic (and in particular, all SRB) measures. Moreover, we show that every hyperbolic measure lifts to a *first return* Young tower for some iterate of the map, which is perhaps surprising because in most other applications of Young's work the towers do not have the first return property. Using this, we obtain a way of controlling the tail of the tower, see Corollary C.2.

A related result in this direction is Sarig's result [63] that every hyperbolic measure for a surface diffeomorphism can be lifted to a countable-state topological Markov shift, which is closely related to a Young tower. The results of [63] have been extended to higher dimensions [14], to three-dimensional flows [49] and to non-uniformly hyperbolic surface maps with discontinuities [48]. It seems natural to expect that our results admit similar extensions, although for the extension to higher dimensions one must confront a major technical obstruction, which is that the nice domains we use are inherently two-dimensional.

[10] A more general inducing structure was introduced in [55]; it can be used to study the existence and ergodic properties of equilibrium measures, which include SRB measures.

3. Lifting hyperbolic measures to towers: Theorem C

Statement (1) of Theorem A follows immediately from Statement (2) of Theorem B and the results in [69] on the existence of SRB measures for Young towers. Statement (2) of Theorem A follows immediately from our next result, which states that nice Lebesgue-strongly recurrent rectangles are *necessary* for the existence of hyperbolic (SRB) measures.

Theorem C. *Let f be a $C^{1+\alpha}$ surface diffeomorphism. If μ is a non-atomic ergodic hyperbolic measure for f, then for every sufficiently small $\chi > 0$ and every $\epsilon > 0$, there is an integer $\ell \in \mathbb{N}$ such that for every $r > 0$, there exists a $(\chi, \epsilon, \ell, r)$-nice regular set A with $\mu(A) > 0$ such that every $x \in A$ has positive frequencies of returns as in* (1.9)*, and continues to satisfy this property with f^i replaced by f^{-i}. In particular, A is recurrent, and if μ is an SRB measure, then A is Lebesgue-strongly recurrent.*

The rest of the paper is therefore devoted to the proofs of Theorems B and C, which we will carry out through the development of some non-trivial technical results of independent interest. We first conclude this section with two almost immediate corollaries of Theorems B and C, also of independent interest. The first is an observation concerning the geometric structure of hyperbolic measures, formalized in the notion of *liftability* which is relevant in many applications and studies of the ergodic properties of invariant measures, see e.g. [56].

Definition 3.1 (Liftable measures)**.** An invariant probability measure μ is *liftable* (to a Topological Young Tower) if there exists a recurrent rectangle Γ which supports a Topological Young Tower and a probability measure $\hat{\mu}$ on Γ which is invariant for the corresponding induced map $F : \Gamma \to \Gamma$, such that

$$E_{\hat{\mu}} := \int_\Gamma \tau \, d\hat{\mu} < \infty \quad \text{and} \quad \mu = \frac{1}{E_{\hat{\mu}}} \sum_{i=0}^{\infty} f_*^i(\hat{\mu}|\{\tau > i\}). \tag{3.1}$$

The following is an immediate consequence of Theorem C and [72].

Corollary C.1. *Let f be a $C^{1+\alpha}$ surface diffeomorphism. Then every ergodic non-atomic hyperbolic (invariant) probability measure lifts to a First T-Return Topological Young Tower for some $T > 0$.*

Before stating our second corollary, we emphasize that we have a first T-return tower which is, to all intents and purposes, essentially as good as if it was a first return time tower. Existence of a first T-return tower in this generality is a rather surprising result, and has important consequences. For example, because every return to $\Gamma_{pq} \cap \Lambda_\ell$ is a return to the base of the tower, we can relate the size of the tail of the tower to the return rate to regular level sets.

Corollary C.2. *Let f be a $C^{1+\alpha}$ surface diffeomorphism, μ an ergodic non-atomic χ-hyperbolic measure, and Λ a χ-hyperbolic set with $\mu(\Lambda) > 0$. Suppose that for every $\epsilon > 0$ there is $\ell \in \mathbb{N}$ such that for every open set $U \subset M$ with*

$\mu(U \cap \Lambda_\ell) > 0$ *and every* $T \in \mathbb{N}$*, we have*

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mu\big\{x \in U \cap \Lambda_\ell : f^{kT}(x) \notin U \cap \Lambda_\ell \text{ for every } 1 \le k \le n\big\} < 0.$$

Then μ *lifts to a First* T*-Return Topological Young Tower for some* $T > 0$*, whose tail decays exponentially in the sense that*

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \hat{\mu}\{x \in \Gamma : \tau(x) > n\} < 0.$$

Proof of Corollary C.2. Corollary C.2 follows from Theorems B and C by putting $A = \Gamma_{pq} \cap \Lambda_\ell$, $U = \mathrm{Int}(\Gamma_{pq})$, and using the fact that the tower has the first T-return property. □

Remark 3.2. A similar result can easily be formulated for subexponential rates of decay. If μ is a hyperbolic measure whose log Jacobian along unstable leaves is sufficiently regular, then control of the tail of the tower implies results on decay of correlations and other statistical properties; see [69, 70, 51, 35, 55, 64] for upper bounds on correlations, [62, 38, 64] for lower bounds, and [52, 53] for some other statistical properties (this list is far from comprehensive).

The condition on regularity of the log Jacobian is satisfied when μ is an equilibrium measure for a sufficiently regular potential function. We remark that the results in [31, 32, 24] provide techniques for studying equilibrium states for some classes of non-uniformly hyperbolic systems, which suggest a method for establishing the hypothesis of Corollary C.2. (Note that those systems are not surface diffeomorphisms, so the results here do not apply directly.)

4. Hyperbolic branches from almost returns: Theorem D

A remarkable feature of Theorem B is that its hypotheses contain only a minimal amount of structure whereas the conclusions produce a great deal of non-trivial structure. The fundamental ingredient which we will use to build this structure is that of a *hyperbolic branch* which we proceed to define. The main result of this section, Theorem D, is then a statement on the existence of hyperbolic branches under very mild recurrence conditions.

We assume throughout this section that Γ_{pq} is a nice domain as in Definition 1.16 and Figure 1.1, and that $A \subset \Gamma_{pq}$ satisfies the following condition (which in particular is true if A is a nice regular set or a nice rectangle).

Definition 4.1. Say that $A \subset \Gamma_{pq}$ *has full length local stable and unstable curves* if every $x \in A$ is contained in a local stable curve $W^s_x \subset \Gamma_{pq}$ whose endpoints are on W^u_p and W^u_q, and in a local unstable curve W^u_x whose endpoints are on W^s_p and W^s_q.

The following definitions will allow us to discuss analogues of stable and unstable curves through points in Γ_{pq} that do not necessarily lie in A.

Definition 4.2 (Cones)**.** Given a normed vector space V, a *cone* in V is any subset $K \subset V$ for which there is a decomposition $V = E \oplus F$ and $\omega > 0$ such

that
$$K = \{0\} \cup \{v + w : v \in E, w \in F, \|w\| < \omega \|v\|\}.$$
Say that two cones K and K' are *transverse* if $K \cap K' = \{0\}$. Note that we do not require K and K' to be defined in terms of the same decomposition $E \oplus F$.

Definition 4.3 (Conefields)**.** A *conefield* over Γ_{pq} is a family of cones $\mathcal{K} = \{K_x \subset T_x M\}_{x \in \Gamma_{pq}}$; it is continuous if K_x can be defined via decompositions $E_x \oplus F_x$ and widths ω_x such that E_x, F_x, ω_x vary continuously in x.

Definition 4.4 ($\mathcal{K}$-admissible curves)**.** Given a conefield $\mathcal{K} = \{K_x\}_{x \in \Gamma_{pq}}$, we say that a C^1 curve $\gamma \subset \Gamma_{pq}$ is $\mathcal{K}$*-admissible* if for every $x \in \gamma$ we have $T_x \gamma \subset K_x$.

Definition 4.5 (Adapted conefields)**.** Let $A \subset \Gamma_{pq}$ have full length stable and unstable curves. We say that two transverse continuous conefields $\mathcal{K}^s = \{K^s_x\}$ and $\mathcal{K}^u = \{K^u_x\}$ over Γ_{pq} are *adapted* for A if W^s_x is $\mathcal{K}^s$-admissible and W^u_x is $\mathcal{K}^u$-admissible for every $x \in A \cup \{p, q\}$.

Now we can define the broader family of curves that generalizes the "true" stable and unstable curves through points in A.

Definition 4.6 (Admissible curves in a nice domain)**.** Given adapted conefields $\mathcal{K}^s$ and $\mathcal{K}^u$ for a set $A \subset \Gamma_{pq}$ as in the previous definition, we refer to a $\mathcal{K}^s$-admissible curve as a *stable admissible curve*, and a $\mathcal{K}^u$-admissible curve as an *unstable admissible curve*. We say that a stable admissible curve is *full length* if its endpoints lie on W^u_p and W^u_q, and an unstable admissible curve is full length if its endpoints lie on W^s_p and W^s_q.

Definition 4.7 (Strips in a nice domain)**.** Given $A \subset \Gamma_{pq}$ and $\mathcal{K}^{s/u}$ as above, a closed region $\widehat{C}^s \subset \Gamma_{pq}$ is a *stable strip* if it is bounded by two pieces of W^u_p, W^u_q and two full length stable admissible curves; see Figure 4.1. Similarly, a closed region $\widehat{C}^u \subset \Gamma_{pq}$ is an *unstable strip* if it is bounded by two pieces of W^s_p, W^s_q and two full length unstable admissible curves.

The stable and unstable strips we consider will typically be bounded by admissible curves of the form $W^{s/u}_x$ for some x, but we do not require this as part of the definition.

Definition 4.8 (Hyperbolic Branches)**.** Let $\mathcal{K}^s, \mathcal{K}^u$ be adapted conefields for $A \subset \Gamma_{pq}$. Let $\widehat{C}^s, \widehat{C}^u \subseteq \Gamma_{pq}$ be a stable and an unstable strip respectively and suppose that there exists $i > 0$ such that $f^i(\widehat{C}^s) = \widehat{C}^u$. The map
$$f^i : \widehat{C}^s \to \widehat{C}^u \tag{4.1}$$
is a (C, κ)*-hyperbolic branch* if for every $x \in \widehat{C}^s$ and $y = f^i(x) \in \widehat{C}^u$ we have
$$Df^i_x(K^u_x) \subset K^u_y \quad \text{and} \quad Df^{-i}_y(K^s_y) \subset K^s_x \tag{4.2}$$
and if for every $v^u \in K^u_x$ and $v^s \in K^s_y$, the vectors defined for $0 \le j \le i$ by $v^u_j := Df^j_x(v^u)$ and $v^s_j := Df_y^{-(i-j)}(v^s)$ satisfy
$$\|v^u_j\| \le C e^{-\kappa(i-j)} \|v^u_i\| \quad \text{and} \quad \|v^s_j\| \le C e^{-\kappa j} \|v^s_0\|. \tag{4.3}$$

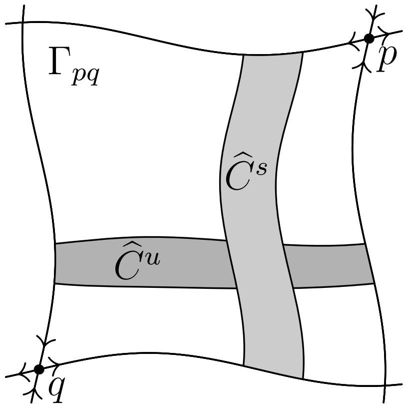

FIGURE 4.1. Stable strips and unstable strips in Γ_{pq}.

We call i the *order* of the hyperbolic branch.

Remark 4.9. Stable and unstable strips which form a hyperbolic branch have the special property that their boundaries are pieces of the global stable and unstable manifolds of the points p, q.

The following result plays no role in Theorem D, but is natural to state at this point and is important in §11.2. See §9.2 and Remark 9.1 for its proof.

Proposition 4.10. *Suppose that $I \subset \mathbb{N}$ is infinite and that $\{f^i \colon \widehat{C}^s_i \to \widehat{C}^u_i\}_{i\in I}$ is a family of (C,κ)-hyperbolic branches whose stable strips are nested in the sense that $\widehat{C}^s_i \supset \widehat{C}^s_j$ for all $i, j \in I$ with $j \geq i$. Then $\bigcap_{i\in I} \widehat{C}^s_i$ is a local (C,κ)-stable curve that has full length in Γ_{pq}.*

Similarly, if the unstable strips are nested in the sense that $\widehat{C}^u_i \supset \widehat{C}^u_j$ for all $i, j \in I$ with $j \geq i$, then $\bigcap_{i\in I} \widehat{C}^u_i$ is a local (C,κ)-unstable curve that has full length in Γ_{pq}.

Definition 4.11. If $f^i \colon \widehat{C}^s_1 \to \widehat{C}^u_1$ and $f^j \colon \widehat{C}^s_2 \to \widehat{C}^u_2$ are hyperbolic branches, then $\widehat{C}^s := f^{-i}(\widehat{C}^u_1 \cap \widehat{C}^s_2)$ is the stable strip for a hyperbolic branch $f^{i+j} \colon \widehat{C}^s \to \widehat{C}^u := f^j(\widehat{C}^u_1 \cap \widehat{C}^s_2)$. We refer to this new branch as the *concatenation* of the original two branches. Observe that $\widehat{C}^s$ consists of the points in the stable strip of the first branch that are mapped by f^i to the stable strip of the second branch, as shown in Figure 4.2.

If we start with two (C,κ)-hyperbolic branches, then the concatenation of the two has the property that for every $x \in \widehat{C}^s$ and $v^s \in Df^{-(i+j)}_{f^{i+j}(x)} K^s_{f^{i+j}(x)}$, the vectors $v^s_k := Df^k_x(v^s)$ for $0 \leq k \leq i+j$ satisfy $\|v^s_k\| \leq Ce^{-\kappa k}\|v^s_0\|$ when $0 \leq k \leq i$, and

$$\|v^s_k\| \leq Ce^{-\kappa(k-i)}\|v^s_i\| \leq C^2 e^{-\kappa k}\|v^s_0\|$$

when $k > i$, with similar estimates on v^u_k; thus $f^{i+j} \colon \widehat{C}^s \to \widehat{C}^u$ is a (C^2,κ)-hyperbolic branch. The fact that C is replaced with C^2 can be a problem as the constant would continue to grow with each repeated concatenation.

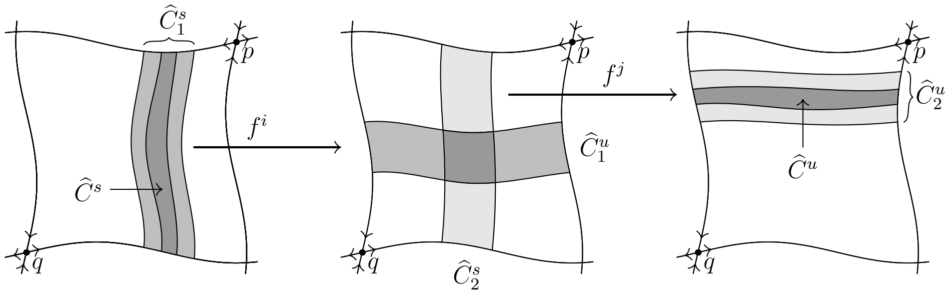

FIGURE 4.2. Concatenating hyperbolic branches.

In the uniformly hyperbolic setting, one can deal with this problem by passing to an *adapted metric*, or *Lyapunov metric*, in which $C = 1$. In non-uniform hyperbolicity this procedure does not give a continuous metric on M, and one must restrict to regular level sets. In the next section we use this idea to produce a family of branches that has the following property.

Definition 4.12 (Concatenation property of hyperbolic branches)**.** We say that a collection of (C, κ)-hyperbolic branches has the *concatenation property* if any finite concatenation of these branches is still a (C, κ)-hyperbolic branch.

Given a nice regular set $A \subset \Gamma_{pq}$, it would be a fairly routine exercise in non-uniform hyperbolicity theory (and the geometry of nice domains) to show that for every $x \in \Gamma$ and $i \in \mathbb{N}$ such that $f^i(x) \in \Gamma$, there is a hyperbolic branch $f^i\colon \widehat{C}^s \to \widehat{C}^u$ such that $x \in \widehat{C}^s$, and that the collection of branches associated to all returns to A has the hyperbolic branch property. However, when we prove Theorem B in §11, it will become important to work with a larger collection of branches associated to a weaker kind of return.[11]

Definition 4.13 (Almost Returns)**.** Let $A \subset \Gamma_{pq}$ have full length stable and unstable curves. A point $x \in A$ has an *almost return* to A at time $i \in T\mathbb{Z}$ (see Figure 4.3) if $f^i(x) \in \Gamma_{pq}$ and there is $y \in A$ such that

$$[x, y, i] \neq \emptyset \quad \text{where} \quad [x, y, i] := \begin{cases} f^i(W^s_x) \cap W^u_y & \text{if } i > 0 \\ f^i(W^u_x) \cap W^s_y & \text{if } i < 0. \end{cases}$$

Note that actual returns, where $x, f^i(x) \in A$, are special cases of almost returns and thus if A is recurrent, then every $x \in A$ has infinitely many almost returns in both forward and backward time. In §11 we will construct a rectangle using hyperbolic branches associated to almost returns, for which we need the following definition and result. (These become vacuous for a set without almost returns.)

Definition 4.14 (Hyperbolic branch property)**.** The set $A \subset \Gamma_{pq}$ has the (C, κ)*-hyperbolic branch property* if there exist adapted conefields $\mathcal{K}^{s/u}$ such that the following are true:

[11]We need the larger collection to guarantee that the rectangle we build is "saturated", see Remark 11.10 and §11.5.

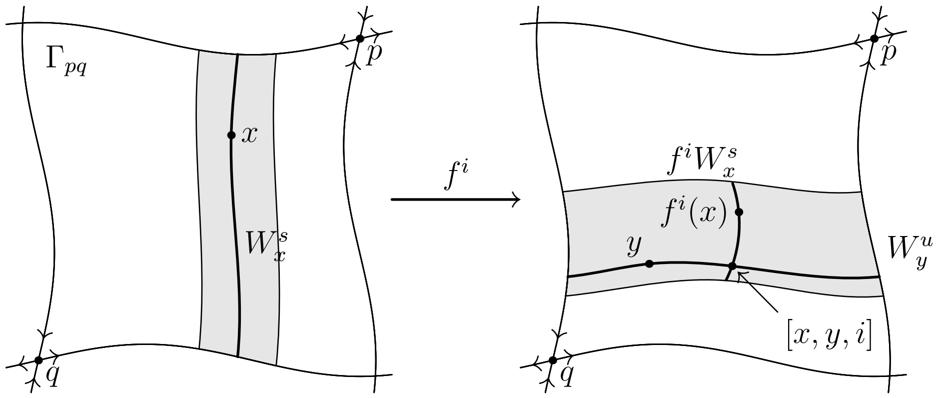

FIGURE 4.3. An almost return and the associated hyperbolic branch.

(1) whenever $x \in A$ has an almost return to A at a time $i \in T\mathbb{N}$, there exists a (C, κ)-hyperbolic branch $f^i : \widehat{C}^s \to \widehat{C}^u$ such that $x \in \widehat{C}^s$;
(2) the collection of such branches has the concatenation property.

Theorem D. *Let f be a $C^{1+\alpha}$ surface diffeomorphism. For every $\chi > \lambda > 0$, every $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and every $\ell \in \mathbb{N}$, there is $r > 0$ such that every $(\chi, \epsilon, \ell, r)$-nice regular set has the $(C, \lambda/3)$-hyperbolic branch property for $C = 8\sqrt{2}(1 - e^{2(\lambda-\chi)})^{-1/2} e^{2\epsilon\ell}$.*

The constant C in Theorem D is equal to $\widehat{Q} e^{2\epsilon\ell}$ where $\widehat{Q}$ is as in (5.3) below. The constant r determined by Theorem D is the same constant which appears in the statements of Theorem B and Theorem A, and is given relatively explicitly in (10.1). In §11 we use Theorem D, together with some regularity estimates from §5.4 and §8.2, to prove Theorem B, and thus part (1) of Theorem A. Most of that proof only relies on Theorem D and can be read independently of the other sections; the regularity results are only needed for the proof of part (2) of Theorem B.

Theorem D implies that every almost return corresponds to a hyperbolic branch. Such branches can be concatenated as in Definition 4.11, and we point out that the concatenated branches themselves may or may not correspond to almost returns; see Remark 11.1.

Remark 4.15. Our construction of a first T-return topological Young tower in §11 is solely based on the hyperbolic branch property. Theorem D ensures this property in the setting of a global invariant hyperbolic set Λ although one could imagine verifying this property in other situations using alternate arguments. In this case one would still require some extra information on bounded distortion in order to get the analogue of Theorem B(2) about a genuine Young tower.

Remark 4.16. The proof of Theorem D in §10 relies on the notion of "regular branch" that we introduce in the next section. The hyperbolic branches we seek are restrictions of these regular branches to a nice domain Γ_{pq} (see Figure 10.1), and Lemmas 10.3 and 10.4 use the niceness property in a crucial way to show that the restricted branch has stable and unstable strips that are actually contained in Γ_{pq}. One could attempt to mimic the restriction procedure for a domain Γ_{pq} without the niceness property, but in that case there would be no way to guarantee this containment.

5. Pseudo-orbits: Theorems E and F

The proofs of Theorems C and D are ultimately based on a new and non-trivial result, Theorem E below, in the general theory of (nonuniformly) hyperbolic sets, which is a generalization of the Katok closing lemma. This result is of independent interest and we expect it to have further applications beyond those presented here. For simplicity, and in view of the setting of our main theorems, we state it for two-dimensional diffeomorphisms but it should generalize in a relatively straighforward way to arbitrary dimension.

In §5.1 we introduce the basic notion of a Lyapunov chart, following the approach of Sarig in [63], and use this to define the notion of stable and unstable strips in §5.2 and of *regular* branch for pseudo-orbits in §5.3. In §5.3 we state our main result about regular branches for pseudo-orbits, and in §5.4 a useful consequence about shadowing orbits.

5.1. **Lyapunov charts.** Let Λ be a (χ,ϵ)-hyperbolic set. We fix some

$$b > 0$$

sufficiently small to be determined in the course of the proof. This is the only constant which we cannot give *explicitly* in terms of properties of f and the constants associated to the hyperbolic set Λ, however see equations (5.15), (7.13), (7.14), and Lemma 8.15 for the key places in which conditions on b appear. For $x \in \Lambda$, let $e^s_x \in E^s_x$, $e^u_x \in E^u_x$ be unit vectors and define $s,u\colon \Lambda \to [1,\infty)$ by

$$\begin{aligned} s(x)^2 &:= 2\sum_{n=0}^{\infty} e^{2n\lambda}\|Df^n_x e^s_x\|^2, \\ u(x)^2 &:= 2\sum_{n=0}^{\infty} e^{2n\lambda}\|Df^{-n}_x e^u_x\|^2. \end{aligned} \tag{5.1}$$

By (**H3**) the sums above converge and therefore $s(x), u(x)$ are well defined, though note that they are not uniformly bounded in x. Letting $e_1 = (1,0), e_2 = (0,1)$ denote the standard basis vectors in $\mathbb{R}^2$, define the linear map $L_x\colon \mathbb{R}^2 \to T_xM$, by letting

$$L_x(e_1) := u(x)^{-1}e^u_x \quad \text{and} \quad L_x(e_2) := s(x)^{-1}e^s_x \tag{5.2}$$

and extending to $\mathbb{R}^2$ by linearity. We call L_x the *Lyapunov change of coordinates* at x. In Lemma 6.4 we prove the following standard relation between the Riemannian metric and the metric induced by the Lyapunov coordinates:

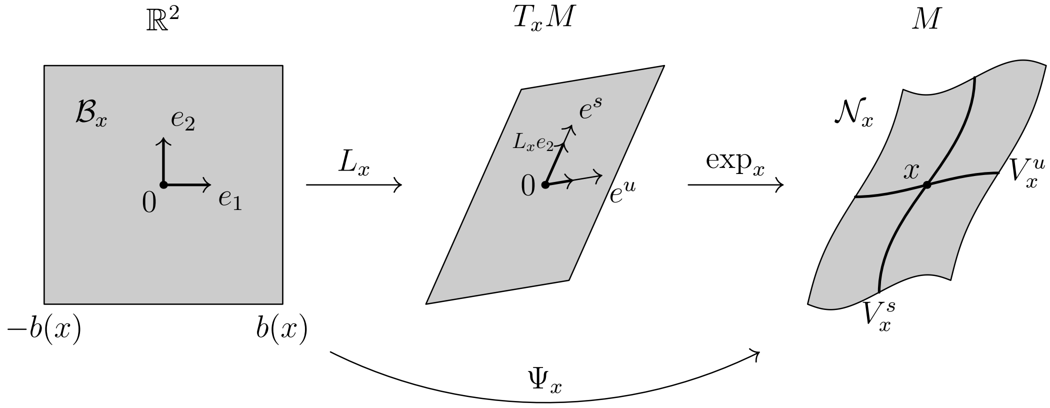

FIGURE 5.1. The Lyapunov chart at x.

for any $0 < \lambda < \chi$ let[12]

$$(5.3)\qquad Q_0 := 1/8 \quad \text{and} \quad \widehat{Q} := Q_0^{-1}\left(2\sum_{i=0}^{\infty} e^{2(\lambda-\chi)i}\right)^{1/2} = \frac{8\sqrt{2}}{\sqrt{1-e^{2(\lambda-\chi)}}},$$

then, for every $\ell \in \mathbb{N}$ and $x \in \Lambda_\ell$, we have

$$(5.4)\qquad 1 \le \|L_x^{-1}\| \le 3Q_0\widehat{Q}e^{2\epsilon\ell}.$$

For every $x \in \Lambda$ let

$$(5.5)\qquad b(x) := b\left(\sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon}\|L_{f^k(x)}^{-1}\|\right)^{-\frac{1}{\alpha}}.$$

Notice that the sum converges by (1.3) and (5.4). For $\ell \ge 1$ let

$$(5.6)\qquad b_\ell := b\left(3Q_0\widehat{Q}e^{2\epsilon\ell}\sum_{k=-\infty}^{\infty} e^{-|k|\epsilon}\right)^{-\frac{1}{\alpha}}$$

It follows immediately from the definition that $b_{\ell+1}/b_\ell = e^{-2\epsilon/\alpha}$ and $b \ge b(x) \ge b_\ell > 0$, and it follows from (5.4) (we give a formal proof in Lemma 6.5) that $e^{-3\epsilon/\alpha} < b(x)/b(f(x)) < e^{3\epsilon/\alpha}$. Then, for every $x \in \Lambda_\ell$ we define

$$(5.7)\qquad \mathcal{B}_x^{(\ell)} := [-b_\ell, b_\ell]^2 \subseteq [-b(x), b(x)]^2 =: \mathcal{B}_x \subset \mathbb{R}^2.$$

Letting $\exp_x \colon T_xM \to M$ be the exponential map, define $\Psi_x \colon \mathcal{B}_x \to M$ by

$$\Psi_x := \exp_x \circ L_x.$$

The map Ψ_x is called a *Lyapunov chart* at x; see [13] for a more general notion.[13] We write

$$(5.8)\qquad \mathcal{N}_x^{(\ell)} := \Psi_x(\mathcal{B}_x^{(\ell)}) \subseteq \Psi_x(\mathcal{B}_x) =: \mathcal{N}_x.$$

[12]We could of course replace Q_0 in the expression for $\widehat{Q}$ by its explicit value but various calculations to be given below will be easier and clearer by keeping track of Q_0 as an independent constant.

[13]In [13] the Lyapunov change of coordinates L_x is required to be *tempered*, but we do not require this condition.

Notice that $\Psi_x(0) = x$ and therefore $\mathcal{N}_x^{(\ell)}, \mathcal{N}_x$ are neighbourhoods of x, which we call respectively the *regular neighbourhood of level* ℓ of x and the *regular neighbourhood* of x.

5.2. **Stable and unstable strips.** We want to define stable and unstable cones and other objects related to these regular neighbourhoods and Lyapunov charts, these will be analogous to the corresponding definitions for nice domains in §4. For this we need to introduce an additional small constant ω which, for completeness, we define precisely.

Let $\chi > \lambda > 0$ and $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and Λ a (χ, ϵ)-hyperbolic set. Notice that (1.7) gives $\epsilon < \alpha\lambda/18$ and so $3\epsilon/\alpha < \lambda/6$, which gives $e^{-\lambda/2}e^{3\epsilon/\alpha} < e^{-\lambda/3}$. Thus we can choose $\omega > 0$ sufficiently small that the following inequalities all hold:

$$
\begin{aligned}
e^{-\lambda/2}e^{3\epsilon/\alpha} + e^{-\lambda}\omega &< e^{-\lambda/3}, & (1-\omega) &\geq \sqrt{2(1+\omega^2)}/2, \\
e^{\lambda}/\sqrt{1+\omega^2} &\geq e^{2\lambda/3}, & 2\omega &< 1 - e^{\lambda/24}e^{-\lambda/3}, \\
(e^{-\lambda/24} - \omega e^{\lambda/24})/\sqrt{1+\omega^2} &> e^{-\lambda/4}, & e^{-\lambda/24} &< 1/\sqrt{1+e^{-2\lambda}\omega^2}.
\end{aligned}
\tag{5.9}
$$

From now on we always assume that ω satisfies (5.9). The following three definitions are analogues of Definitions 4.5–4.7 for nice domains.

Definition 5.1 (Cones in regular neighbourhoods)**.** For any $\lambda > 0$ we fix the following cones defined in terms of standard Euclidean coordinates with $v = (v_1, v_2) \in \mathbb{R}^2$:

$$
\begin{aligned}
\tilde{K}^s &:= \{v : |v_1| < e^{-\lambda}\omega|v_2|\} \ \subset \{v : |v_1| < \omega|v_2|\} := K^s, \\
\tilde{K}^u &:= \{v : |v_2| < e^{-\lambda}\omega|v_1|\} \ \subset \{v : |v_2| < \omega|v_1|\} := K^u.
\end{aligned}
\tag{5.10}
$$

Given $x \in \Lambda$ and $y \in \mathcal{N}_x$, we write

$$
K^{s/u}_{x,y} := D_{\Psi_x^{-1}(y)}\Psi_x(K^{s/u}) \quad \text{and} \quad \tilde{K}^{s/u}_{x,y} := D_{\Psi_x^{-1}(y)}\Psi_x(\tilde{K}^{s/u}) \tag{5.11}
$$

for the cones in T_yM that correspond to the cones over $\mathcal{B}_x$.

Definition 5.2 (Admissible curves in regular neighbourhoods)**.** Let $x \in \Lambda$. A C^1 curve $\gamma^s \subset \mathcal{B}_x$ is a *stable admissible curve* (resp. *strongly stable admissible curve*) if its tangent directions lie in K^s (resp. $\tilde{K}^s$); similarly for unstable admissible curves $\gamma^u \subset \mathcal{B}_x$. In particular, the horizontal and vertical boundaries of $\mathcal{B}_x$ or $\mathcal{B}_x^{(\ell)}$, are stable and unstable admissible curves respectively and so we denote them by $\partial^s\mathcal{B}_x, \partial^s\mathcal{B}_x^{(\ell)}$ and $\partial^u\mathcal{B}_x, \partial^u\mathcal{B}_x^{(\ell)}$ respectively. A stable admissible curve γ^s is a *full length stable admissible curve*, with respect to $\mathcal{B}_x$ or $\mathcal{B}_x^{(\ell)}$ if its endpoints lie on distinct components of $\partial^u\mathcal{B}_x$ or $\partial^u\mathcal{B}_x^{(\ell)}$ respectively. We define *full length unstable admissible curves* similarly. For a C^1 curve $\gamma^{s/u} \subset \mathcal{N}_x$ we use the same terminology according to the geometry of the corresponding curve $\Psi_x^{-1}(\gamma^{s/u}) \subset \mathcal{B}_x$.

Definition 5.3 (Strips in regular neighbourhoods)**.** Let $x \in \Lambda$. A region $\mathcal{B}_x^s \subseteq \mathcal{B}_x^{(\ell)}$ is a *(strongly) stable strip* if its boundary is formed by two full length (strongly) stable admissible curves and two pieces of $\partial^u\mathcal{B}_x^{(\ell)}$. Full length unstable admissible curves and *(strongly) unstable strips* $\mathcal{B}_x^u \subseteq \mathcal{B}_x^{(\ell)}$ are defined similarly. Moreover, for subsets $\mathcal{N}_x^{s/u} \subseteq \mathcal{N}_x^{(\ell)}$, we use the same terminology depending on the geometry of the corresponding sets $\Psi_x^{-1}(\mathcal{N}_x^{s/u}) \subseteq \mathcal{B}_x^{(\ell)}$.

5.3. **Pseudo-orbits and regular branches.**

Definition 5.4 (Pseudo-orbits)**.** Given constants $\delta, \lambda > 0$ and a finite sequence $\bar{\ell} = (\ell_0, \dots, \ell_k)$ of positive integers satisfying $|\ell_j - \ell_{j-1}| \leq 1$ for all $1 \leq j \leq k$, we say that a finite sequence of points $\bar{x} = (x_0, \dots, x_k)$ is a *finite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit* if for every $0 \leq j \leq k$ we have

$$x_j \in \Lambda_{\ell_j} \quad \text{and} \quad d(f(x_{j-1}), x_j) \leq \delta e^{-\lambda \ell_j}.$$

Notice that the true orbit of the point $x \in \Lambda_\ell$ is trivially an $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit for $\delta = 0$ and for the sequence $\ell_j = \ell + j$, recall (1.3). The results we present here are essentially already known for true orbits, but their generalizations to pseudo-orbits is non-trivial and constitute a crucial step in our arguments. Given a finite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit we write

$$\mathcal{N}_{\bar{x}}^0 := \bigcap_{i=0}^{k} f^{-i} \mathcal{N}_{x_i}^{(\ell_i)} \quad \text{and} \quad \mathcal{N}_{\bar{x}}^j := f^j(\mathcal{N}_{\bar{x}}^0) \text{ for } 1 \leq j \leq k$$

for the sets of points corresponding to orbit segments that go through all the Lyapunov neighborhoods of the points x_i; the size of these sets depends on the choice of $\bar{\ell}$, but we suppress this in the notation.

Remark 5.5. We point out that one must carefully distinguish between $\mathcal{N}_{\bar{x}}^j$, where the subscript denotes a finite pseudo-orbit and the superscript indicates which point in the pseudo-orbit to consider, and $\mathcal{N}_x^{(\ell)}$, where the subscript denotes a point and the superscript denotes the level of a regular set containing that point. A similar distinction should be made when $\mathcal{N}$ is replaced by $\mathcal{B}$ as in the following.

In Lyapunov coordinates, for every $0 \leq j \leq k$, we write

$$\mathcal{B}_{\bar{x}}^j := \Psi_{x_j}^{-1}(\mathcal{N}_{\bar{x}}^j) \subseteq \mathcal{B}_{x_j}^{(\ell_j)} \tag{5.12}$$

and for every $0 \leq i, j \leq k$,

$$f_{\bar{x}}^{i,j} := \Psi_{x_j}^{-1} \circ f^{j-i} \circ \Psi_{x_i} \colon \mathcal{B}_{\bar{x}}^i \to \mathcal{B}_{\bar{x}}^j, \tag{5.13}$$

which is a diffeomorphism between $\mathcal{B}_{\bar{x}}^i$ and $\mathcal{B}_{\bar{x}}^j$ by definition. We will show that such maps are hyperbolic branches in a suitable sense.

Definition 5.6 ($\bar{\ell}$-regular branch)**.** A finite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit $\bar{x}$ *determines an $\bar{\ell}$-regular branch for f* if the following are true:

(i) $\mathcal{B}_{\bar{x}}^0$, $\mathcal{B}_{\bar{x}}^k$ are stable and unstable strips in $\mathcal{B}_{x_0}^{(\ell_0)}$, $\mathcal{B}_{x_k}^{(\ell_k)}$, respectively;
(ii) given $0 \leq i < j \leq k$, $\underline{y} \in \mathcal{B}_{\bar{x}}^i$, $\underline{z} \in \mathcal{B}_{\bar{x}}^j$, $\underline{v}^u \in K^u$, and $\underline{v}^s \in K^s$, we have

$$\begin{aligned} D_{\underline{y}} f_{\bar{x}}^{i,j} \underline{v}^u &\in K^u, & \|D_{\underline{y}} f_{\bar{x}}^{i,j} \underline{v}^u\| &\geq e^{\lambda(j-i)/3} \|\underline{v}^u\|, \\ D_{\underline{z}} f_{\bar{x}}^{j,i} \underline{v}^s &\in K^s, & \|D_{\underline{z}} f_{\bar{x}}^{j,i} \underline{v}^s\| &\geq e^{\lambda(j-i)/3} \|\underline{v}^s\|. \end{aligned} \tag{5.14}$$

In the specific case where $\ell_0 = \ell_k = \ell$ and $\ell_j = \min(\ell + j, \ell + k - j)$ for $0 < j < k$, we refer to this as an ℓ-regular branch.

Here and in what follows, we use underlined variables such as $\underline{y}, \underline{z}$ to represent coordinates in $\mathbb{R}^2$, while undecorated variables such as y, z will represent points in M.

We now state our main result in the general setting of hyperbolic sets.

Theorem E. *Let $f\colon M \to M$ be a $C^{1+\alpha}$ surface diffeomorphism. For every $\chi > \lambda > 0$ and every $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, there are $b, \delta > 0$ such that if Λ is a (χ, ϵ)-hyperbolic set then every finite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit in Λ determines an $\bar{\ell}$-regular branch.*

Recall from Definition 4.11 and the discussion following it that two (C, κ)-hyperbolic branches can be concatenated, but the resulting branch is only guaranteed to be (C^2, κ)-hyperbolic. The crucial advantage we gain by considering $\bar{\ell}$-regular branches is that the concatenation of two $\bar{\ell}$-regular branches is again a $\bar{\ell}$-regular branch; this is because the hyperbolicity estimates in (5.14) are given at the level of the Lyapunov charts, and not on the surface itself. The relationship between the two types of estimates is given by the following computation.

Since $D_0 \exp_x$ is the identity map, we can choose b small enough that $\Psi_x\colon \mathcal{B}_x \to \mathcal{N}_x$ is a diffeomorphism, in particular injective, for every $x \in \Lambda$, and such that (5.4) gives

$$\|D_{\underline{y}}\Psi_x\| \leq 2 \quad \text{and} \quad \|D_y\Psi_x^{-1}\| \leq 4Q_0\widehat{Q}e^{2\epsilon\ell} \tag{5.15}$$

for every $\ell \in \mathbb{N}$, $x \in \Lambda_\ell$, $\underline{y} \in \mathcal{B}_x$, and $y \in \mathcal{N}_x$. The conclusion of the following result should be compared to Definition 4.8.

Proposition 5.7. *Suppose $\bar{x} = (x_0, \dots, x_k)$ determines an $\bar{\ell}$-regular branch. Given $y \in \mathcal{N}_{\bar{x}}^0$ and $v^u \in K_{x_0,y}^u$, let $v_j^u = D_y f^j(v^u)$ for $0 \leq j \leq k$. Then $v_j^u \in K_{x_j, f^j y}^u$ and*

$$\|v_j^u\| \leq \widehat{Q}e^{2\epsilon\ell_k}e^{-\lambda(k-j)/3}\|v_k^u\|. \tag{5.16}$$

Similarly, given $z \in \mathcal{N}_{\bar{x}}^k$ and $v^s \in K_{x_k,z}^s$, let $v_j^s = D_z f^{-(k-j)}(v^s)$ for $0 \leq j \leq k$; then $v_j^s \in K_{x_j, f^{-(k-j)}z}^s$ and

$$\|v_j^s\| \leq \widehat{Q}e^{2\epsilon\ell_0}e^{-\lambda j/3}\|v_0^s\|. \tag{5.17}$$

Proof. The inclusions follow immediately from (5.11) and (5.14). For (5.16) it suffices to observe that (5.14) gives

$$\|v_j^u\| = \|D_{f^k y}(\Psi_{f^j y} \circ f_{\bar{x}}^{k,j} \circ \Psi_{f^k y}^{-1})(v_k^u)\| \leq 2 \cdot e^{-\lambda(k-j)/3} \cdot 4Q_0\widehat{Q}e^{2\epsilon\ell_k}\|v_k^u\|$$

and recall that $Q_0 = 1/8$ from (5.3). The proof of (5.17) is analogous. □

In §10 we use Theorem E and Proposition 5.7 to prove Theorem D.

Remark 5.8. In the theory of *uniformly* hyperbolic systems, it is well-known that every pseudo-orbit segment determines a regular branch as in Definition 5.6; that is, there is $\delta > 0$ such that $x_0, \dots, x_k$ determines a regular branch whenever $d(f(x_j), x_{j+1}) < \delta$ for all $0 \le j < k$. In non-uniform hyperbolicity, various versions of Theorem E have been obtained. The first result of this type is the well-known Katok closing lemma [46]. Other versions were obtained by Hirayama [42], and by Sarig [63] in his construction of countable Markov partitions for surface diffeomorphisms with positive topological entropy; this latter result was generalized to higher dimensions by Ben Ovadia [15]. What makes our Theorem E different is the *explicit* relationship between ℓ and the pseudo-orbit scale $\delta e^{-\lambda\ell}$; this is absolutely crucial for our arguments, particularly our construction of hyperbolic branches associated with almost T-returns.

5.4. **Shadowing.** We conclude this section by stating a consequence of Theorem E for the problem of *shadowing* in the non-uniformly hyperbolic setting. It is classical problem in hyperbolic dynamics whether every pseudo-orbit is "shadowed" by a real orbit of the system [58]. In uniform hyperbolicity theory a positive answer to this question is a fundamental result, with many applications and in particular, is a key ingredient in the construction of Markov partitions. Extending it to the setting of non-uniform hyperbolicity has proved challenging and few results have been obtained, see [13]. Here we give a powerful shadowing result, which follows relatively easily from Theorem E. In addition to the existence of a shadowing orbit, this result gives an explicit estimate on the hyperbolicity constants associated to this shadowing orbit and also proves Theorem 1.12 on the existence of local stable and unstable curves (up to optimizing the rate of contraction, see Remark 1.13).

Definition 5.9. Given a bi-infinite sequence $\bar{\ell} = (\ell_n)_{n\in\mathbb{Z}}$ with $|\ell_{n+1} - \ell_n| \le 1$ for all n, a *bi-infinite* $(\bar{\ell}, \delta, \lambda)$*-pseudo-orbit* is a bi-infinite sequence $(x_n)_{n\in\mathbb{Z}}$ such that for every $n \in \mathbb{Z}$, we have $x_n \in \Lambda_{\ell_n}$ and $d(f(x_{n-1}), x_n) \le \delta e^{-\lambda\ell_n}$. Replacing "$n \in \mathbb{Z}$" with "$n \ge 0$" and "$n \le 0$" gives the definitions of *forward* $(\bar{\ell}, \delta, \lambda)$*-pseudo-orbit* and *backward* $(\bar{\ell}, \delta, \lambda)$*-pseudo-orbit*, respectively.

Theorem F. *Let $f\colon M \to M$ be a $C^{1+\alpha}$ surface diffeomorphism. For every $\chi > \lambda > 0$ and every $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, there are $b, \delta > 0$ such that if Λ is a (χ, ϵ)-hyperbolic set, then the following are true.*

(1) *If $(x_n)_{n\ge 0}$ is a forward $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit in Λ, then $\bigcap_{n\ge 0} f^{-n}(\mathcal{N}_{x_n}^{(\ell_n)})$ is a $C^{1+\text{Hölder}}$ full length local $(\widehat{Q}e^{2\epsilon\ell_0}, \lambda/3)$-stable curve.*
(2) *If $(x_n)_{n\le 0}$ is a backward $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit in Λ, then $\bigcap_{n\le 0} f^{-n}(\mathcal{N}_{x_n}^{(\ell_n)})$ is a $C^{1+\text{Hölder}}$ full length local $(\widehat{Q}e^{2\epsilon\ell_0}, \lambda/3)$-unstable curve.*
(3) *If $(x_n)_{n\in\mathbb{Z}}$ is a bi-infinite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit in Λ, then there is a unique shadowing point $y \in M$ such that $f^n(y) \in \mathcal{N}_{x_n}^{(\ell_n)}$ for all $n \in \mathbb{Z}$. Moreover, the point y is $(\lambda/4, 2\epsilon, \ell_0 + \ell')$-regular for $\ell' = \lceil \frac{1}{2\epsilon} \log \widehat{Q} \rceil$.*

Let σ denote the left shift on $\mathbb{N}^{\mathbb{Z}}$ (which contains the possible sequences $\bar{\ell}$) and $\Lambda^{\mathbb{Z}}$ (which contains the possible pseudo-orbits $(x_n)_{n\in\mathbb{Z}}$). If $\bar{x}$ is a bi-infinite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit, then $\sigma\bar{x}$ is a bi-infinite $(\sigma\bar{\ell}, \delta, \lambda)$-pseudo-orbit. It follows from uniqueness of the shadowing point that the map $\bar{x} \mapsto y$ intertwines σ and f: if y is the unique shadowing point for $\bar{x}$, then $f(y)$ is the unique shadowing point for $\sigma\bar{x}$. Moreover, if the pseudo-orbit is periodic in the sense that $\sigma^n\bar{\ell} = \bar{\ell}$ and $\sigma^n\bar{x} = \bar{x}$ for some $n \in \mathbb{N}$, then $f^n(y) = y$. In particular, this allows us to deduce the Katok Closing Lemma as a specific case of Theorem F: if $x, f^n(x) \in \Lambda_\ell$ and $d(x, f^n(x)) < \delta e^{-\lambda\ell}$, then we can take $x_k = f^k(x)$ and $\ell_k = \ell + \min(k, n-k)$ for $0 \le k < n$, and repeat periodically mod n, to get a bi-infinite $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit whose unique shadowing point is the periodic point from the Katok Closing Lemma.

Corollary 5.10. *If $x, y \in \Lambda_\ell$ and $d(x,y) \le \delta e^{-\lambda\ell}$, then writing*

$$\ell_n = \ell + |n| \quad \text{and} \quad z_n = \begin{cases} f^n(y) & n < 0, \\ f^n(x) & n \ge 0 \end{cases}$$

gives a $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit $\bar{z}$, whose unique shadowing point z is the bracket $[x,y] = V^s_x \cap V^u_y$. Thus $\{[x,y] : x, y \in \Lambda_\ell, d(x,y) \le \delta e^{-\lambda\ell}\}$ is a $(\lambda/4, 2\epsilon, \ell + \ell')$-regular set, where $\ell' = \lceil \frac{1}{2\epsilon} \log \widehat{Q} \rceil$.

Remark 5.11. The δ in Corollary 5.10 is the same as in Theorem F. This corollary in particular shows that the bracket $[x,y]$ exists, so that $\delta e^{-\lambda\ell}$ plays the role of δ_ℓ from Theorem 1.12. In general the δ_ℓ used there could be larger than $\delta e^{-\lambda\ell}$, because that result made no claims about regularity of $[x,y]$.

Part II. Hyperbolic Theory

In this part of the paper we develop all the general hyperbolic theory needed for the proofs of our main results, and prove Theorems E and F. For simplicity we state and prove everything in the two-dimensional setting, as required by our applications of these results. The results in this part, as well as Theorems E and F, do not contain any inherently two-dimensional ideas, and we expect that they hold as stated in higher dimensions as well; we have not written down detailed proofs, but there do not appear to be any conceptual obstructions. This should be contrasted with Theorems A–D and their proofs in Part III, where a crucial role is played by the notion of "nice domain", which has no obvious higher-dimensional analogue.

The contents of this part are completely self-contained, with no reference to existing results in the literature, and follow directly from the definition of (χ, ϵ)-hyperbolic set.

In §6 we give some basic estimates related to Lyapunov charts. In §7 we state and prove Theorem G, on the hyperbolicity of f in Lyapunov charts, which is a fundamental result in the theory of hyperbolic sets and the key motivation behind the introduction of Lyapunov charts. In §8 we state and prove Theorem H which gives some conditions guaranteeing that Lyapunov charts of nearby points are "overlapping" in a suitable sense. Some qualitative versions of this result are known but we give a quantitative version which is not available in the existing literature and which is crucial for our arguments; this is the most involved technical step in the paper. In §9.1 we combine these two results to prove Theorem E, and in §9.2 we deduce Theorem F.

Throughout the proofs, we will write Q_j for various constants that depend only on $f, \chi, \lambda, \epsilon$ and are independent of x, y, ℓ, n, k, etc.

6. Lyapunov chart estimates

Recall the definition of Lyapunov charts in §5.1 and in particular the functions $s(x), u(x)$ defined in (5.1), the map L_x defined in (5.2), and the quantities b_ℓ and $b(x)$ defined in (5.6) and (5.5) respectively. We follow here the basic approach of Sarig in [63] and prove some properties of these objects, including (5.4). We start with a couple of simple estimates showing that although the functions $s(x), u(x)$ are not slowly varying along orbits, they are uniformly bounded on each Λ_ℓ and satisfy a bounded variation property along orbits.

Lemma 6.1. *For every $\ell \geq 1$ and $x \in \Lambda_\ell$, we have*

$$\sqrt{2} \leq s(x) \leq Q_0 \widehat{Q} e^{\epsilon \ell} \quad \text{and} \quad \sqrt{2} \leq u(x) \leq Q_0 \widehat{Q} e^{\epsilon \ell}. \tag{6.1}$$

Proof. The lower bounds follow immediately from the definition of $s(x), u(x)$ in (5.1). Using the hyperbolicity property (**H3**) and the definition of Q_0, $\widehat{Q}$ in (5.3), we have

$$s(x)^2 \leq 2 \sum_{n \geq 0} e^{2n\lambda} C(x)^2 e^{-2n\chi} = C(x)^2 (Q_0 \widehat{Q})^2,$$

and then the definition of Λ_ℓ in (1.2) gives the upper bound for $s(x)$. The upper bound for $u(x)$ is similar. $\square$

Lemma 6.2. *There exists a constant $Q_1 = Q_1(c_1, c_2, \lambda) > 0$ such that for every $x \in \Lambda$ and unit vectors $e^s_x \in E^s_x, e^u_x \in E^u_x$, we have*

$$\begin{aligned} Q_1^{-1} \leq e^\lambda \|D_x f(e^s_x)\| \leq s(x)/s(f(x)) \leq \sqrt{1 + e^{2\lambda} \|D_x f(e^s_x)\|^2} \leq Q_1, \\ Q_1^{-1} \leq e^\lambda \|D_x f(e^u_x)\|^{-1} \leq u(f(x))/u(x) \leq \sqrt{1 + e^{2\lambda} \|D_x f(e^u_x)\|^{-2}} \leq Q_1. \end{aligned}$$

Proof. For $s(x)$ we have

$$\frac{s(x)^2}{2} := \sum_{k=0}^{\infty} e^{2k\lambda} \|Df^k_x e^s_x\|^2 = 1 + e^{2\lambda} \|Df_x e^s_x\|^2 \sum_{k=1}^{\infty} e^{2(k-1)\lambda} \|Df^{k-1}_{f(x)} e^s_{f(x)}\|^2,$$

and

$$\sum_{k=1}^{\infty} e^{2(k-1)\lambda} \|Df^{k-1}_{f(x)} e^s_{f(x)}\|^2 = \sum_{k=0}^{\infty} e^{2k\lambda} \|Df^k_{f(x)} e^s_{f(x)}\|^2 = s(f(x))^2,$$

which gives $s(x)^2 = 2\big(1 + e^{2\lambda} \|Df_x e^s_x\|^2 s(f(x))^2\big)$. Dividing both sides by $s(f(x))^2$ and taking a square root, we obtain

$$\frac{s(x)}{s(f(x))} = \sqrt{2(s(f(x))^{-2} + e^{2\lambda} \|Df_x e^s_x\|^2)} \geq e^\lambda \|Df_x e^s_x\|, \tag{6.2}$$

and the upper bound for $s(x)/s(f(x))$ also follows since $s(f(x)) \geq \sqrt{2}$ so $2(s(f(x)))^{-2} \leq 1$. For uniformity of Q_1 it suffices to note that $\|Df^{\pm 1}_x\| \leq$

$\max(e^{-c_1}, e^{c_2})$ by (1.5). A similar computation for $u(x)$, using f^{-1} in place of f, gives the following analogue of (6.2):

$$\frac{u(x)}{u(f^{-1}(x))} = \sqrt{2(u(f^{-1}(x)))^{-2} + e^{2\lambda}\|Df_x^{-1}e_x^u\|^2} \geq e^{\lambda}\|Df_x^{-1}e_x^u\|,$$

with the upper bound again coming from $u(f(x)) \geq \sqrt{2}$, so that

$$e^{\lambda}\|Df_x^{-1}e_x^u\| \leq \frac{u(x)}{u(f^{-1}(x))} \leq \sqrt{1 + e^{2\lambda}\|Df_x^{-1}e_x^u\|^2}.$$

Applying this to $f(x)$ and $f^{-1}(f(x)) = x$, we get

$$e^{\lambda}\|Df_{f(x)}^{-1}e_{f(x)}^u\| \leq \frac{u(f(x))}{u(x)} \leq \sqrt{1 + e^{2\lambda}\|Df_{f(x)}^{-1}e_{f(x)}^u\|^2}.$$

Using that $\|Df_{f(x)}^{-1}e_{f(x)}^u\| = \|D f_x e_x^u\|^{-1}$ we get the bounds for u, and uniformity via Q_1 comes as it did for s. $\square$

An almost immediate, but extremely important, consequence of Lemma 6.2, and therefore of the way the functions $s(x), u(x)$ are defined, is the following fundamental result originally proved in [54].[14]

Theorem 6.3 (Oseledets–Pesin Reduction Theorem)**.** *For every $x \in \Lambda$,*

$$L_{f(x)}^{-1} \circ D_x f \circ L_x = \begin{pmatrix} A_x & 0 \\ 0 & B_x \end{pmatrix}, \tag{6.3}$$

where L_x is given by (5.2) *and $A_x, B_x \in \mathbb{R}$ satisfy*

$$0 < Q_1^{-1} < B_x \leq e^{-\lambda} < 1 < e^{\lambda} \leq A_x < Q_1. \tag{6.4}$$

Proof. The diagonal form is a consequence of the invariance of the stable and unstable subspaces E_x^s, E_x^u and the fact that $L_x, L_{f(x)}$ map the coordinate axes to these subspaces, recall (5.2). Thus, by linearity of L_x we have

$$A_x e_1 = L_{f(x)}^{-1} \circ D_x f \circ L_x(e_1) = L_{f(x)}^{-1} \circ D_x f\Big(\frac{e_x^u}{u(x)}\Big) = \frac{u(f(x))}{u(x)}\|D_x f(e_x^u)\|,$$

$$B_x e_2 = L_{f(x)}^{-1} \circ D_x f \circ L_x(e_2) = L_{f(x)}^{-1} \circ D_x f\Big(\frac{e_x^s}{s(x)}\Big) = \frac{s(f(x))}{s(x)}\|D_x f(e_x^s)\|,$$

and the statement then follows from Lemma 6.2. $\square$

Lemma 6.4. *For every $x \in \Lambda$, we have*

$$1 \leq \frac{\sqrt{s(x)^2 + u(x)^2}}{\sqrt{2}\sin\measuredangle(E_x^s, E_x^u)} \leq \|L_x^{-1}\| \leq \frac{\sqrt{s(x)^2 + u(x)^2}}{\sin\measuredangle(E_x^s, E_x^u)}. \tag{6.5}$$

In particular, for every $\ell \geq 1$, $x \in \Lambda_\ell$, and $k \in \mathbb{Z}$, we have

$$1 \leq \|L_{f^k(x)}^{-1}\| \leq 3Q_0\widehat{Q}e^{2\epsilon\ell}e^{2\epsilon|k|}. \tag{6.6}$$

[14] In [54] it is required that L_x is tempered, but this is not necessary for our formulation.

Proof. Let $\theta(x) = \measuredangle(E^s_x, E^u_x)$. Consider the orthonormal basis $\{e^u_x, (e^u_x)^\perp\}$ in T_xM, oriented so that $e^s_x = \cos\theta(x) e^u_x + \sin\theta(x)(e^u_x)^\perp$. From (5.2), we have $L_x e_1 = u(x)^{-1} e^u_x$ and $L_x e_2 = s(x)^{-1} e^s_x$, so the matrices of $L_x^{\pm 1}$ relative to the orthonormal bases $\{e_1, e_2\}$ and $\{e^u_x, (e^u_x)^\perp\}$ have the form

$$L_x = \begin{pmatrix} u(x)^{-1} & s(x)^{-1}\cos\theta(x) \\ 0 & s(x)^{-1}\sin\theta(x) \end{pmatrix} \text{ and } L_x^{-1} = \begin{pmatrix} u(x) & -u(x)/\tan\theta(x) \\ 0 & s(x)/\sin\theta(x) \end{pmatrix}.$$

The norm of $A = L_x^{-1}$ is the square root of the largest eigenvalue of

$$A^T A = \frac{1}{\sin^2\theta(x)} \begin{pmatrix} u(x)^2 & -s(x)u(x)\cos\theta(x) \\ -s(x)u(x)\cos\theta(x) & s(x)^2 \end{pmatrix},$$

and thus a routine computation with the quadratic formula gives

$$\|L_x^{-1}\|^2 = \frac{u(x)^2 + s(x)^2 + \sqrt{\big(u(x)^2 + s(x)\big)^2 - 4s(x)^2 u(x)^2 \sin^2\theta(x)}}{2\sin^2\theta(x)}.$$

The square root term lies between 0 and $u(x)^2 + s(x)^2$, which proves (6.5).

For (6.6), we first observe that $\sin\theta \geq \frac{2\theta}{\pi}$ for all $\theta \in [0, \frac{\pi}{2}]$. From (1.3), we see that if $x \in \Lambda_\ell$ then $f^k(x) \in \Lambda_{\ell+|k|}$ and so $\measuredangle(E^s_{f^k(x)}, E^u_{f^k(x)}) \geq e^{-\epsilon(\ell+|k|)}$ and, by (6.1), $s(f^k(x)) \leq Q_0 \widehat{Q} e^{\epsilon(\ell+|k|)}$ and $u(f^k(x)) \leq Q_0 \widehat{Q} e^{\epsilon(\ell+|k|)}$. Substituting these bounds into the upper bound from (6.5) gives

$$\|L^{-1}_{f^k(x)}\| \leq \frac{\sqrt{2} Q_0 \widehat{Q} e^{\epsilon(\ell+|k|)}}{2e^{-\epsilon(\ell+|k|)}/\pi} = \frac{\pi}{\sqrt{2}} Q_0 \widehat{Q} e^{2\epsilon\ell} e^{2\epsilon|k|},$$

which proves (6.6) since $\pi/\sqrt{2} < 3$. □

Lemma 6.5. *For every $\ell \geq 1$ and $x \in \Lambda_\ell$*

$$b \geq b(x) \geq b_\ell > 0 \quad \textit{and} \quad e^{-3\epsilon/\alpha} < b(x)/b(f(x)) < e^{3\epsilon/\alpha}.$$

Proof. From (6.6), the sum in the definition of $b(x)$ converges. Therefore, for ϵ small and $x \in \Lambda_\ell$ we have

$$1 \leq \|L_x^{-1}\| \leq \sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon} \|L^{-1}_{f^k(x)}\| \leq 3Q_0\widehat{Q} e^{2\epsilon\ell} \sum_{k=-\infty}^{\infty} e^{-|k|\epsilon}$$

and thus $b \geq b(x) \geq b_\ell > 0$ as in the first part of the statement. Moreover

$$\begin{aligned} b(f(x)) &:= b\left(\sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon} \|L^{-1}_{f^{k+1}(x)}\|\right)^{-1/\alpha} \\ &= b\left(\sum_{k=-\infty}^{\infty} e^{-3|k-1|\epsilon} \|L^{-1}_{f^k(x)}\| e^{3(-|k|+|k|)\epsilon}\right)^{-1/\alpha} \\ &= b\left(\sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon} \|L^{-1}_{f^k(x)}\| e^{3(-|k-1|+|k|)\epsilon}\right)^{-1/\alpha}. \end{aligned}$$

Notice that $-|k-1|+|k|$ can only take the values $+1$ or -1, depending on the value of k, and therefore

$$b(x)e^{-3\epsilon/\alpha} \leq b\left(\sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon}\|L^{-1}_{f^k(x)}\|e^{3(-|k-1|+|k|)\epsilon}\right)^{-1/\alpha} \leq b(x)e^{3\epsilon/\alpha}.$$

This completes the proof. □

7. Hyperbolicity in Lyapunov charts: Theorem G

The key motivation for introducing Lyapunov charts is to show that the map f restricted to some neighbourhood of points of Λ is uniformly hyperbolic in Lyapunov coordinates and to study the way f maps such neighbourhoods to each other. To state this precisely, we consider $x \in \Lambda$ and recall from Figure 5.1 that the Lyapunov chart Ψ_x maps $\mathcal{B}_x \subset \mathbb{R}^2$ onto $\mathcal{N}_x \subset M$, so that a point $y \in \mathcal{N}_x$ is represented by coordinates $\underline{y} = \Psi_x^{-1}(y) \in \mathcal{B}_x$. (As in §5, we will use underlined variables to represent coordinates in $\mathcal{B}_x \subset \mathbb{R}^2$.) To represent the map f in these coordinates, we write

$$\mathcal{B}_x^{s,1} := \Psi_x^{-1}(\mathcal{N}_x \cap f^{-1}\mathcal{N}_{f(x)}) \quad \text{and} \quad \mathcal{B}_{f(x)}^{u,1} := \Psi_{f(x)}^{-1}(f(\mathcal{N}_x) \cap \mathcal{N}_{f(x)})$$

and denote by

$$f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x \colon \mathcal{B}_x^{s,1} \to \mathcal{B}_{f(x)}^{u,1}$$

the corresponding diffeomorphism; see Figure 7.1. Similarly, if $x \in \Lambda_{\ell_0}$, and $f(x) \in \Lambda_{\ell_1}$ for some ℓ_1 with $|\ell_1 - \ell_0| \leq 1$ we let $\bar{\ell} = (\ell_0, \ell_1)$ and write

$$\mathcal{B}_{x,\bar{\ell}}^{s,1} := \Psi_x^{-1}(\mathcal{N}_x^{(\ell_0)} \cap f^{-1}\mathcal{N}_{f(x)}^{(\ell_1)}) \quad \text{and} \quad \mathcal{B}_{f(x),\bar{\ell}}^{u,1} := \Psi_{f(x)}^{-1}(f\mathcal{N}_x^{(\ell_0)} \cap \mathcal{N}_{f(x)}^{(\ell_1)}) \tag{7.1}$$

By (5.7), $\mathcal{B}_{x,\bar{\ell}}^{s,1} \subseteq \mathcal{B}_x^{s,1}$ and $\mathcal{B}_{f(x),\bar{\ell}}^{u,1} \subseteq \mathcal{B}_{f(x)}^{u,1}$ and therefore f_x restricts to a map

$$f_x \colon \mathcal{B}_{x,\bar{\ell}}^{s,1} \to \mathcal{B}_{f(x),\bar{\ell}}^{u,1}$$

For simplicity we will just use the same notation f_x in both cases. We also mention that in the definitions of the sets $\mathcal{B}_{x,\bar{\ell}}^{s,1}, \mathcal{B}_x^{s,1}, \mathcal{B}_{f(x),\bar{\ell}}^{u,1}, \mathcal{B}_{f(x)}^{u,1}$ we implicitly mean the *connected component*, containing x or $f(x)$ respectively, of these sets which *a priori* may not be connected. The main result of this section states that the map f_x is uniformly hyperbolic and the sets just defined have a certain specific geometry, see Figure 7.1.

A key part of the statement of Theorem G below is that certain sets are stable and unstable strips which are *strictly* contained in the sets $\mathcal{B}_x^{(\ell)}$. It is convenient to introduce the following sets.

Definition 7.1. Given $r > 0$, consider the sets

$$\widetilde{\mathcal{B}}_r^s := [-e^{-\lambda/3}r, e^{-\lambda/3}r] \times [-r, r] \subset \mathbb{R}^2,$$
$$\widetilde{\mathcal{B}}_r^u := [-r, r] \times [-e^{-\lambda/3}r, e^{-\lambda/3}r] \subset \mathbb{R}^2$$

and let

$$\widetilde{\mathcal{N}}_x^{s/u} = \Psi_x(\widetilde{\mathcal{B}}_{b(x)}^{s/u}) \quad \text{and} \quad \widetilde{\mathcal{N}}_{x,\ell}^{s/u} = \Psi_x(\widetilde{\mathcal{B}}_{b_\ell}^{s/u}). \tag{7.2}$$

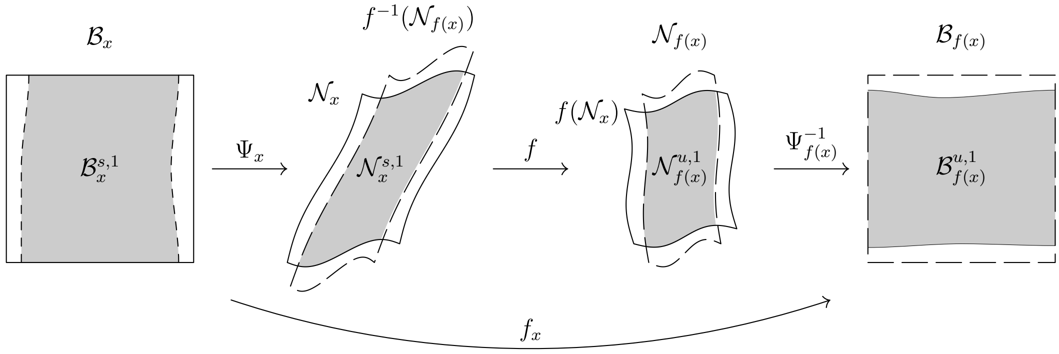

FIGURE 7.1. The map f in Lyapunov coordinates.

Theorem G. *Let $f\colon M \to M$ be a $C^{1+\alpha}$ surface diffeomorphism. Fix $\chi > \lambda > 0$, $0 < \epsilon < \epsilon_1(f,\chi,\lambda)$, and ω satisfying* (5.9). *Then there is $b > 0$ such that for every (χ,ϵ)-hyperbolic set Λ, $\bar{\ell} = (\ell_0, \ell_1)$ with $|\ell_0 - \ell_1| \leq 1$ and $x \in \Lambda_{\ell_0}$ with $f(x) \in \Lambda_{\ell_1}$, the following holds: for every $\underline{y} \in \mathcal{B}_x^{s,1}$, $\underline{z} \in \mathcal{B}_{f(x)}^{u,1}$, $\underline{v}^{u/s} \in K^{u/s}$,*

$$
\begin{aligned}
D_{\underline{y}} f_x(\underline{v}^u) &\in \widetilde{K}^u, \qquad \|D_{\underline{y}} f_x(\underline{v}^u)\| \geq e^{\lambda/2}\|\underline{v}^u\|, \\
D_{\underline{z}} f_x^{-1}(\underline{v}^s) &\in \widetilde{K}^s, \qquad \|D_{\underline{z}} f_x^{-1}(\underline{v}^s)\| \geq e^{\lambda/2}\|\underline{v}^s\|.
\end{aligned}
\tag{7.3}
$$

Moreover, the sets $\mathcal{B}_{x,\bar{\ell}}^{s,1}, \mathcal{B}_{f(x),\bar{\ell}}^{u,1}$ are strongly stable and unstable strips, satisfying

$$
\mathcal{B}_{x,\bar{\ell}}^{s,1} \subseteq \widetilde{\mathcal{B}}_{b_{\ell_0}}^s \quad \text{and} \quad \mathcal{B}_{f(x),\bar{\ell}}^{u,1} \subseteq \widetilde{\mathcal{B}}_{b_{\ell_1}}^u. \tag{7.4}
$$

Remark 7.2. Notice that the estimates (7.3) hold in particular for all $\underline{y} \in \mathcal{B}_{x,\bar{\ell}}^{s,1}$, $\underline{z} \in \mathcal{B}_{f(x),\bar{\ell}}^{u,1}$ but they *do not depend on* ℓ and give *one-step hyperbolicity* in the sense that the expansion and contraction is exhibited immediately after one iteration. This is in contrast with the fact that if $x \in \Lambda_\ell$ for large ℓ we have very poor hyperbolicity estimates on the surface, cf. (**H3**) and (1.2). This is of course the effect of the Lyapunov change of coordinates which has controlled, but very large, distortion, and applies to very small neighbourhoods of x when $x \in \Lambda_\ell$ for large ℓ, recall (5.4) and (5.5). The crucial advantage of writing the estimates as in (7.3) is that we can iterate the map any number of times without loss of hyperbolicity. We only need to worry about the effect of the distortion at the beginning and end of any arbitrarily long piece of orbit in order to recover the actual hyperbolicity estimates for the original map f on the surface.

In the rest of this section we prove Theorem G. In §7.1 we establish the derivative estimates (7.3), in §7.2 we use these to prove the invariance property of the cones, and in §7.3 we prove that $\mathcal{B}_{x,\bar{\ell}}^{s,1}, \mathcal{B}_{x,\bar{\ell}}^{u,1}$ are stable and unstable strips and satisfy (7.4).

7.1. **Derivative estimates.** Here we prove the hyperbolicity estimates (7.3). We start with the special case $\underline{y} = 0$.

Lemma 7.3. *For every* $x \in \Lambda$, $\underline{v}^u \in K^u$, $\underline{v}^s \in K^s$,

$$\|D_0 f_x(\underline{v}^u)\| \geq e^{2\lambda/3}\|\underline{v}^u\| \quad \text{and} \quad \|D_0 f_x^{-1}(\underline{v}^s)\| \geq e^{2\lambda/3}\|\underline{v}^s\|. \tag{7.5}$$

Before proving Lemma 7.3, we set up some notation. Consider the map

$$\hat{f}_x := \exp_{f(x)}^{-1} \circ f \circ \exp_x \colon T_xM \to T_{f(x)}M. \tag{7.6}$$

Then since $f_x := \Psi_{f(x)}^{-1} \circ f \circ \Psi_x = L_{f(x)}^{-1} \circ \exp_{f(x)}^{-1} \circ f \circ \exp_x \circ L_x$, we have

$$f_x = L_{f(x)}^{-1} \circ \hat{f}_x \circ L_x. \tag{7.7}$$

Given $v \in T_xM$, we have

$$D_v \hat{f}_x = D_{f \circ \exp_x(v)} \exp_{f(x)}^{-1} \circ D_{\exp_x(v)} f \circ D_v \exp_x. \tag{7.8}$$

Also $f_x \colon \mathcal{B}_x^{s,1} \to \mathcal{B}_{f(x)}^{u,1}$ and for every $\underline{y} \in \mathcal{B}_x^{s,1}$, $D_{\underline{y}} f_x = D_{\underline{y}}(L_{f(x)}^{-1} \circ \hat{f}_x \circ L_x)$. Using that $L_x, L_{f(x)}$ are linear maps, we have

$$D_{\underline{y}} f_x = L_{f(x)}^{-1} \circ D_{L_x(\underline{y})} \hat{f}_x \circ L_x \tag{7.9}$$

With this notation we can easily prove Lemma 7.3.

Proof of Lemma 7.3. For $\underline{y} = 0$ we have $L_x(\underline{y}) = L_x(0) = 0$, so (7.9) gives

$$D_0 f_x = L_{f(x)}^{-1} \circ D_0 \hat{f}_x \circ L_x.$$

Now $\exp_x(0) = x$ gives $f \circ \exp_x(0) = f(x)$, and the exponential function is tangent to the identity at 0, i.e., $D_0 \exp_x = \mathrm{Id}$ and $D_{f(x)} \exp_{f(x)}^{-1} = \mathrm{Id}$, so (7.8) implies $D_0 \hat{f}_x = D_x f$, which means that we have

$$D_0 f_x = L_{f(x)}^{-1} \circ D_x f \circ L_x. \tag{7.10}$$

Writing $\underline{v}^u = v_1 e_1 + v_2 e_2$, (6.3) and (7.10) give $D_0 f_x(\underline{v}^u) = A_x v_1 e_1 + B_x v_2 e_2$. Since $\underline{v}^u \in K^u$ implies $|v_2| \leq \omega |v_1|$, we conclude that

$$\frac{\|D_0 f_x(\underline{v}^u)\|^2}{\|\underline{v}^u\|^2} = \frac{A_x^2 v_1^2 + B_x^2 v_2^2}{v_1^2 + v_2^2} \geq \frac{A_x^2}{1 + (v_2/v_1)^2} \geq \frac{e^{2\lambda}}{1 + \omega^2} \geq e^{4\lambda/3},$$

where the last two inequalities use (6.4) and (5.9), respectively. This proves the first half of (7.5); the second is similar. $\square$

By Lemma 7.3 and continuity of the differential, the expansion estimates in (7.3) hold in some neighbourhood of $0 \in \mathcal{B}_x^{s,1}$. We show that this neighbourhood contains $B_x^{s,1}$, which is the key part of the proof of Theorem G. The main step is the following estimate for the derivatives of $\hat{f}_x$.

Lemma 7.4. *There exists* $Q_2 > 0$ *such that for all* $x \in \Lambda$ *and* $\underline{y}, \underline{z} \in \mathcal{B}_x^{s,1}$,

$$\|D_{L_x(\underline{y})} \hat{f}_x - D_{L_x(\underline{z})} \hat{f}_x\| \leq Q_2 \|\underline{y} - \underline{z}\|^\alpha.$$

Proof. For simplicity we write $v := L_x(\underline{y}), u := L_x(\underline{z})$. Since L_x is a contraction and $\underline{y}, \underline{z} \in \mathcal{B}_x^{s,1} \subseteq [-b(x), b(x)]^2 \subset [-b, b]^2$, we have $\|v - u\| \leq \|\underline{y} - \underline{z}\| \leq 2\sqrt{2}b$. In particular it is sufficient to prove that $\|D_v \hat{f}_x - D_u \hat{f}_x\| \leq Q\|v - u\|^\alpha$ for all $u, v \in T_xM$ with $\|u\|, \|v\| \leq b$. Since M is a C^2 Riemannian manifold, there is $Q_3 > 0$ such that $\|D_v \exp_x - D_u \exp_x\| \leq Q_3 \|u - v\|$,

$\|D_{f\circ\exp_x v}\exp_x^{-1} - D_{f\circ\exp_x u}\exp_x^{-1}\| \le Q_3\|u-v\|$ for all $x \in M$ and $u, v \in T_xM$ with $\|u\|, \|v\| \le 1$. Moreover, there is $Q_4 > 0$ such that $d(\exp_x(u), \exp_x(v)) \le Q_4\|u-v\|$ for all such x, u, v, and hence since Df is Hölder continuous we have $\|D_{\exp_x(v)}f - D_{\exp_x(u)}f\| \le |Df|_\alpha Q_4^\alpha \|u-v\|^\alpha$. Then the definition of $D_v\hat{f}_x, D_u\hat{f}_x$ in (7.8) gives the result. □

Lemma 7.5. *There exists $Q_5 > 0$ such that for all $x \in \Lambda$ and $\underline{y}, \underline{z} \in \mathcal{B}_x^{s,1}$,*

$$\|D_{\underline{y}}f_x - D_{\underline{z}}f_x\| \le Q_5 b^\alpha.$$

Proof. By (7.9), using $\|L_x\| \le 1$ and Lemma 7.4, for every $\underline{y}, \underline{z} \in \mathcal{B}_x^{s,1}$,

$$\begin{aligned}\|D_{\underline{y}}f_x - D_{\underline{z}}f_x\| &= \|L_{f(x)}^{-1} \circ D_{L_x(\underline{y})}\hat{f}_x \circ L_x - L_{f(x)}^{-1} \circ D_{L_x(\underline{z})}\hat{f}_x \circ L_x\| \\ &= \|L_{f(x)}^{-1} \circ (D_{L_x(\underline{y})}\hat{f}_x - D_{L_x(\underline{z})}\hat{f}_x) \circ L_x\| \\ &\le \|L_{f(x)}^{-1}\| \cdot \|D_{L_x(\underline{y})}\hat{f}_x - D_{L_x(\underline{z})}\hat{f}_x\| \\ &\le Q_2\|L_{f(x)}^{-1}\| \; \|\underline{y} - \underline{z}\|^\alpha\end{aligned} \tag{7.11}$$

Moreover, $\underline{y}, \underline{z} \in \mathcal{B}_x^{s,1} \subset [-b(x), b(x)]^2$ implies $\|\underline{y} - \underline{z}\| \le 2b(x)$ and therefore, by Lemma 6.5,

$$\|\underline{y} - \underline{z}\|^\alpha \le 2^\alpha b(x)^\alpha \le 2^\alpha e^{3\epsilon} b(f(x))^\alpha. \tag{7.12}$$

Moreover, from (5.5) we have

$$b(f(x))^\alpha := b^\alpha \left(\sum_{k=-\infty}^{\infty} e^{-3|k|\epsilon}\|L_{f^{k+1}(x)}^{-1}\|\right)^{-1} \le b^\alpha \|L_{f(x)}^{-1}\|^{-1}$$

and therefore, substituting into (7.12) and then into (7.11) we get

$$\|D_{\underline{y}}f_x - D_{\underline{z}}f_x\| \le Q_2\|L_{f(x)}^{-1}\| \; \|\underline{y} - \underline{z}\|^\alpha \le 2^\alpha e^{3\epsilon} Q_2 b^\alpha,$$

which completes the proof. □

Now we can prove the expansion estimates in (7.3) for all $\underline{y}$. By Lemmas 7.3 and 7.5, for every $x \in \Lambda$, $\underline{y} \in \mathcal{B}_x^{s,1}$, and $\underline{v}^u \in K^u$, we have

$$\|D_{\underline{y}}f_x(\underline{v}^u)\| \ge \|D_0 f_x(\underline{v}^u)\| - \|D_{\underline{y}}f_x - D_0 f_x\| \cdot \|\underline{v}^u\| \ge (e^{2\lambda/3} - Q_5 b^\alpha)\|\underline{v}^u\|.$$

Choose $b > 0$ small enough that

$$e^{2\lambda/3} - Q_5 b^\alpha \ge e^{\lambda/2}; \tag{7.13}$$

then we get $\|D_{\underline{y}}f_x(\underline{v}^u)\| \ge e^{\lambda/2}\|\underline{v}^u\|$. A similar argument gives $\|D_{\underline{y}}f_{f(x)}^{-1}(\underline{v}^s)\| \ge e^{\lambda/2}\|\underline{v}^s\|$ for every $\underline{v}^s \in K^s$, and so we have the expansion estimates in (7.3).

7.2. **Conefield invariance.** We now prove the conefield invariance from (7.3). Fix $\eta > 0$ small enough that if $z = z_1e_1 + z_2e_2 \in \mathbb{R}^2$ has $\|z\| = 1$ and $|z_2| < e^{-2\lambda}\omega|z_1|$, then every $v \in \mathbb{R}^2$ with $\|v - z\| < \eta$ is contained in $\tilde{K}^u$. By homogeneity we see that if the assumption on $\|z\|$ is removed and we have $\|v - z\| < \eta\|z\|$, then once again $v \in \tilde{K}^u$. Given $x \in \Lambda$ and $\underline{v} = v_1e_1 + v_2e_2 \in K^u$, (6.3) gives $D_0 f_x(\underline{v}) = A_x v_1 e_1 + B_x v_2 e_2$, and we have

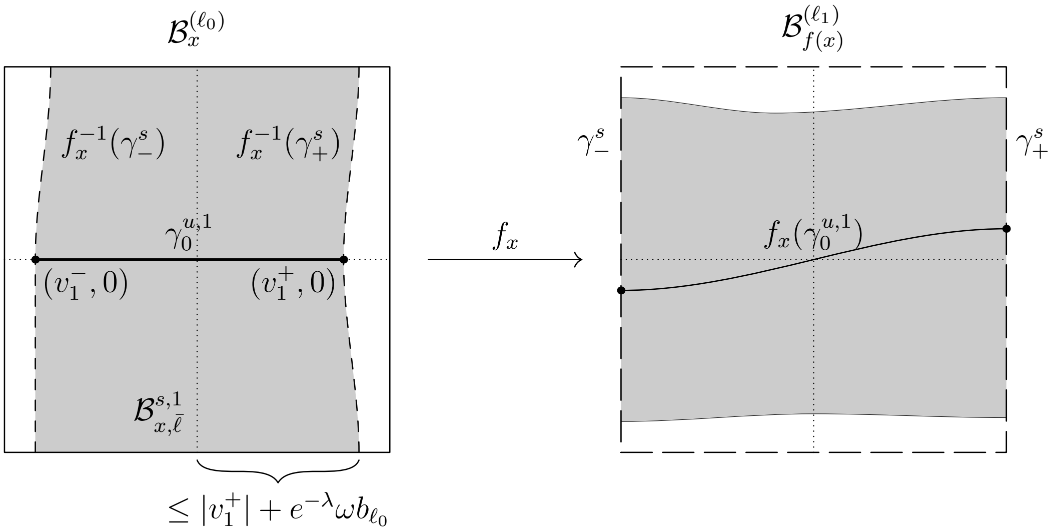

FIGURE 7.2. $\mathcal{B}_x^{s,1}$ is a strongly stable strip.

$|B_x v_2| < e^{-\lambda}\omega|v_1| < e^{-2\lambda}\omega|A_x v_1|$, so $D_0 f_x(\underline{v})$ satisfies the assumption on z mentioned above. Now choose b small enough that

$$Q_5 b^\alpha < \eta. \tag{7.14}$$

Then for every $\underline{y} = \Psi_x^{-1}(y) \in \mathcal{B}_x^{s,1}$, Lemma 7.5 gives $\|D_{\underline{y}} f_x(\underline{v}) - D_0 f_x(\underline{v})\| \leq Q_5 b^\alpha \|\underline{v}\| < \eta\|\underline{v}\|$, and by our choice of η we conclude that $D_{\underline{y}} f_x(\underline{v}) \in \tilde{K}^u$. A completely symmetric argument applies to the stable cones and f^{-1}.

7.3. **Stable and unstable strips.** To complete the proof of Theorem G we show that $\mathcal{B}_{x,\bar\ell}^{s,1}$ and $\mathcal{B}_{f(x),\bar\ell}^{u,1}$ are strongly stable and unstable strips in $\widetilde{\mathcal{B}}_{b_{\ell_0}}^s$, $\widetilde{\mathcal{B}}_{b_{\ell_1}}^u$ respectively, recall (7.1) and (7.4). We begin by proving the statement for $\mathcal{B}_{x,\bar\ell}^{s,1}$. Let $\gamma_0^{u,1} := \{(v_1, 0) \in \mathcal{B}_x^{(\ell_0)} : f_x(v_1, 0) \in \mathcal{B}_{f(x)}^{(\ell_1)}\} \subseteq \mathcal{B}_{x,\bar\ell}^{s,1}$. Notice that $f_x(0) = 0$ and therefore $\mathcal{B}_{x,\bar\ell}^{s,1}$ contains a neighbourhood of 0 and therefore $\gamma_0^{u,1}$ is a non-trivial horizontal segment, and in particular its tangent vectors are contained in the unstable cones K^u. Therefore, by (7.3), the images of the tangent vectors to $\gamma_0^{u,1}$ are contained in the strong unstable cones $\tilde{K}^u$ and in particular the slope of the curve $f_x(\gamma_0^{u,1})$ always has absolute value $< e^{-\lambda}\omega < 1$.

Since $f_x(\gamma_0^{u,1})$ goes through the origin, has slope < 1 in absolute value, and has both endpoints on the boundary of the square $\mathcal{B}_{f(x)}^{(\ell_1)} = [-b_{\ell_1}, b_{\ell_1}]^2$ these endpoints must both lie on the stable boundaries $\gamma_\pm^s := \{(\pm b_{\ell_1}, v_2), v_2 \in [-b_{\ell_1}, b_{\ell_1}]\}$ of $\mathcal{B}_{f(x)}^{(\ell_1)}$, so $f_x(\gamma_0^{u,1})$ is a *full length* strongly unstable admissible curve in $\mathcal{B}_{f(x)}^{(\ell_1)}$ as shown in Figure 7.2. Therefore the preimages $f_x^{-1}(\gamma_\pm^s) \cap \mathcal{B}_x^{(\ell_0)}$ are, by (7.3), strongly stable admissible curves through the endpoints of $\gamma_0^{u,1}$, which are points on the horizontal axis with coordinates $(v_1^\pm, 0)$ with $|v_1^\pm| \leq e^{-\lambda/2} b_{\ell_1}$. Note that this last estimate comes from the fact that every tangent vector to $f_x(\gamma_0^{u,1})$ is contracted by a factor of at least $e^{-\lambda/2}$ under the action of

f_x^{-1}. Since $f_x^{-1}(\gamma^s_\pm)$ are strongly stable admissible curves in $\mathcal{B}_x^{(\ell_0)}$, we conclude as shown in Figure 7.2 that their horizontal distance from the y-axis in $\mathbb{R}^2$ is always bounded above by

$$|v_1^\pm| + e^{-\lambda}\omega b_{\ell_0} \le e^{-\lambda/2}b_{\ell_1} + e^{-\lambda}\omega b_{\ell_0} < (e^{-\lambda/2}e^{3\epsilon/\alpha} + e^{-\lambda}\omega)b_{\ell_0},$$

where the last inequality uses Lemma 6.5. By (5.9), this is $< e^{-\lambda/3}b_{\ell_0}$, which proves that $\mathcal{B}^{s,1}_{x,\bar\ell}$ is a strongly stable strip in $\widetilde{\mathcal{B}}^s_{b_{\ell_0}}$. Similarly, $\mathcal{B}^{u,1}_{f(x),\bar\ell}$ is a strongly unstable strip in $\widetilde{\mathcal{B}}^u_{b_{\ell_1}}$.

8. Overlapping charts: Theorem H

8.1. **Overlapping charts.** The parameters defining Lyapunov charts vary slowly along orbits but in general only measurably with the point $x \in \Lambda$. On each regular level set Λ_ℓ, the dependence is continuous, and it is well-known that "if points $x, y \in \Lambda_\ell$ are close, then their Lyapunov charts are close". The condition on how close x, y need to be depends on ℓ; we need an explicit quantitative estimate, which is provided by (8.3) in Theorem H below. This is the core technical result of the paper, whose proof demands the largest share of our efforts.

First we make precise what it means for two Lyapunov charts to be close. Let $\chi > \lambda > 0$ be fixed, ϵ_1 given by (1.7), and $\epsilon \in (0, \epsilon_1)$. Let Λ be a (χ, ϵ)-hyperbolic set. Given $\ell \ge 1$ and $x, y \in \Lambda_\ell$, recall that $\mathcal{N}_x^{(\ell)}, \mathcal{N}_y^{(\ell)}$ are defined in (5.8) and $\widetilde{\mathcal{N}}^{s/u}_{x,\ell}, \widetilde{\mathcal{N}}^{s/u}_{y,\ell}$ in (7.2).

Definition 8.1 (Overlapping charts)**.** We say that $\mathcal{N}_x^{(\ell)}$ and $\mathcal{N}_y^{(\ell)}$ *are overlapping* if $x, y \in \mathcal{N}_x^{(\ell)} \cap \mathcal{N}_y^{(\ell)}$ and the following conditions hold:

A) Overlapping derivative estimates: for every $\underline{z} \in \Psi_x^{-1}(\mathcal{N}_x^{(\ell)} \cap \mathcal{N}_y^{(\ell)})$ and every $\underline{v}^u \in \tilde{K}^u$, $\underline{v}^s \in \tilde{K}^s$, we have

$$D_{\underline{z}}(\Psi_y^{-1} \circ \Psi_x)(\underline{v}^u) \subset K^u \text{ and } \|D_{\underline{z}}(\Psi_y^{-1} \circ \Psi_x)(\underline{v}^u)\| \ge e^{-\lambda/24}\|\underline{v}^u\|, \tag{8.1}$$

$$D_{\underline{z}}(\Psi_x^{-1} \circ \Psi_y)(\underline{v}^s) \subset K^s \text{ and } \|D_{\underline{z}}(\Psi_x^{-1} \circ \Psi_y)(\underline{v}^s)\| \ge e^{-\lambda/24}\|\underline{v}^s\|, \tag{8.2}$$

and similarly with the roles of x and y reversed.

B) Overlapping stable and unstable strips: Every full length strongly stable curve $\gamma^s \subset \widetilde{\mathcal{N}}^s_{x,\ell}$ (resp. $\widetilde{\mathcal{N}}^s_{y,\ell}$) completely crosses $\widetilde{\mathcal{N}}^u_{y,\ell}$ (resp. $\widetilde{\mathcal{N}}^s_{x,\ell}$) and every full length strongly unstable curve $\gamma^u \subset \widetilde{\mathcal{N}}^u_{x,\ell}$ (resp. $\widetilde{\mathcal{N}}^u_{y,\ell}$) completely crosses $\widetilde{\mathcal{N}}^s_{y,\ell}$ (resp. $\widetilde{\mathcal{N}}^u_{x,\ell}$).

Theorem H. *Let f be a $C^{1+\alpha}$ surface diffeomorphism. For every $\chi > \lambda > 0$, and every $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, there exists $\delta > 0$ such that given any (χ, ϵ)-hyperbolic set Λ, any integer $\ell \in \mathbb{N}$, and $x, y \in \Lambda_\ell$ with*

$$d(x, y) \le \delta e^{-\lambda\ell}, \tag{8.3}$$

the Lyapunov charts $\mathcal{N}_x^{(\ell)}$ and $\mathcal{N}_y^{(\ell)}$ are overlapping.

In the rest of this section we prove Theorem H. The proof depends on two intermediate results about the Hölder continuity of $E_x^{s/u}$ and of $s(x), u(x)$ which we prove in §8.2 and 8.3 respectively. In §8.4 and §8.5 we combine these to prove parts A) and B) respectively in the definition of overlapping charts, thus completing the proof of Theorem H.

Remark 8.2. A similar 'overlapping charts' condition plays a central role in Sarig's construction of countable Markov partitions; see [63, Definition 3.1] for the formal definition and [63, Proposition 3.2] for the key properties, which are similar to our definition above, although it is not immediately clear whether one definition implies the other. The crucial ingredient of our approach here is the explicit distance criterion (8.3) that guarantees overlapping charts, which has no analogue that we are aware of in [63].

8.2. **Hölder continuity of the splitting.**

Proposition 8.3. *There is $Q_6 > 0$ such that for any $\ell \in \mathbb{N}$ and $x, y \in \Lambda_\ell$, we have*

$$d(E_x^s, E_y^s) \leq Q_6 e^{6\epsilon\gamma\ell} d(x,y)^\beta \quad \text{and} \quad d(E_x^u, E_y^u) \leq Q_6 e^{6\epsilon\gamma\ell} d(x,y)^\beta$$

where $d(\cdot,\cdot)$ represents distance in the Grassmannian of M.

Recall that β is given in (1.6). For generality we prove Proposition 8.3 as a special case of the following result which does not require x, y to belong to a (χ, ϵ)-hyperbolic set. More specifically the Hölder continuity only depends on the angle and hyperbolicity estimates at the points x, y and not on how these vary along the orbits of x, y.

Proposition 8.4. *Let M be a compact smooth Riemannian manifold and $f\colon M \to M$ a $C^{1+\alpha}$ diffeomorphism. There is a constant Q_6, depending only on M, $\|Df^{\pm 1}\|$, α, $|Df^{\pm 1}|_\alpha$, and χ, such that if $C, K > 0$ and $x, y \in M$ are such that*

$$\|Df_x^n e_x^s\| \leq Ce^{-\chi n} \text{ and } \|Df_y^n e_y^s\| \leq Ce^{-\chi n} \text{ for all } n \geq 0, \tag{8.4}$$

$$\|Df_x^n e_x^u\| \geq C^{-1}e^{\chi n} \text{ and } \|Df_y^n e_y^u\| \geq C^{-1}e^{\chi n} \text{ for all } n \geq 0, \tag{8.5}$$

for some unit vectors $e_x^{s/u} \in T_xM$, $e_y^{s/u} \in T_yM$ for which the corresponding subspaces $E_{x/y}^{s/u}$ satisfy $\measuredangle(E_x^s, E_x^u) \geq K$, $\measuredangle(E_y^s, E_y^u) \geq K$, then we have

$$d(E_x^s, E_y^s) \leq Q_6 (C^2 K^{-1})^{2\gamma} d(x,y)^\beta, \tag{8.6}$$

where γ, β are as in (1.6). The same bound holds for E_x^u, E_y^u if we have

$$\|Df_x^{-n} e_x^u\| \leq Ce^{-\chi n} \text{ and } \|Df_y^{-n} e_y^u\| \leq Ce^{-\chi n} \text{ for all } n \geq 0, \tag{8.7}$$

$$\|Df_x^{-n} e_x^s\| \geq C^{-1}e^{\chi n} \text{ and } \|Df_y^{-n} e_y^s\| \geq C^{-1}e^{\chi n} \text{ for all } n \geq 0. \tag{8.8}$$

In particular, if (8.4), (8.5), (8.7), and (8.8) all hold, then

$$|\measuredangle(E_x^s, E_x^u) - \measuredangle(E_y^s, E_y^u)| \leq Q_6 (C^2 K^{-1})^{2\gamma} d(x,y)^\beta. \tag{8.9}$$

To see that Proposition 8.4 implies Proposition 8.3, it suffices to observe that given $\epsilon > 0$ and $\ell \in \mathbb{N}$, the conditions (8.4), (8.5), (8.7), (8.8) are satisfied for all $x, y \in \Lambda_\ell$, with $C = e^{\epsilon\ell}$ and $K = e^{-\epsilon\ell}$.

We will give the explicit calculations only for the stable subspaces, leading to the proof of (8.6), as the situation for the unstable subspaces is completely symmetrical. We follow Brin's approach in [12, Appendix A] (see also [13, §5.3]; the main idea of the argument goes back to [21]). The only thing we need that is not given there is the computation for how much vectors in $(E^s_x)^\perp$ are expanded, depending on C and K, given in Lemma 8.8 below. The rest of the proof of Proposition 8.4 is taken nearly verbatim from [12, Appendix A], with notation adjusted to fit our current setting.

First note that by the Whitney embedding theorem [43], we can choose $N \in \mathbb{N}$ such that M can be smoothly embedded in $\mathbb{R}^N$. By compactness of M, the Riemannian metric is uniformly equivalent to the distance induced by the embedding and therefore it suffices to prove the result under the assumption that $M \subset \mathbb{R}^N$. Then, for each $x \in M$ write $E^\perp(x)$ for the orthogonal complement to $T_xM \subset \mathbb{R}^N$; since $E^\perp$ is smooth it suffices to prove the result with E^s_x replaced by $\tilde{E}^s_x := E^s_x \oplus E^\perp(x)$.

Definition 8.5. Given $x \in M \subset \mathbb{R}^N$ and $n \in \mathbb{N}$, let $D^{(n)}_x$ be the $N \times N$ matrix representing the linear map that takes $v \mapsto D_xf^n(v)$ for $v \in T_xM$ and $v \mapsto 0$ for $v \in E^\perp(x)$.

Since we embed M in $\mathbb{R}^N$, we can treat Grassmannian distance between subspaces as follows. Given a subspace $E \subset \mathbb{R}^N$, we define the distance of a non-zero vector v from the subspace E by considering the unique decomposition $v = v^E + v^\perp$ where $v^E \in E$ and $v^\perp \perp E$ and letting $d(v, E) := \|v^\perp\|/\|v\|$. We can then define the distance between two subspaces $E, E' \subset \mathbb{R}^N$ by

$$d(E, E') := \sup\{d(v, E) : v \in E' \setminus \{0\}\} = \sup\{d(v, E') : v \in E \setminus \{0\}\}. \tag{8.10}$$

The strategy of the proof is based on the following general result.

Lemma 8.6. [12, Lemma A.1] *Let $N \geq 2$ and let $\{A_k\}, \{B_k\}$, be two sequences of real $N \times N$ matrices satisfying the following properties*

(1) there are $\Delta \in (0, 1)$ and $c_3 > 0$ such that

$$\|A_k - B_k\| \leq \Delta e^{c_3 k} \text{ for all } k \geq 0; \tag{8.11}$$

(2) there are subspaces $E_A, E_B \subset \mathbb{R}^N$, $\chi > 0$, and $C' > 1$ such that

$$\|A_k v_A\| \leq C' e^{-\chi k}\|v_A\| \quad \text{and} \quad \|A_k v_A^\perp\| \geq (C')^{-1} e^{\chi k}\|v_A^\perp\| \tag{8.12}$$

for every $v_A \in E_A, v_A^\perp \perp E_A$, $k \geq 0$, and

$$\|B_k v_B\| \leq C' e^{-\chi k}\|v_B\| \quad \text{and} \quad \|B_k v_B^\perp\| \geq (C')^{-1} e^{\chi k}\|v_B^\perp\| \tag{8.13}$$

for every $v_B \in E_B, v_B^\perp \perp E_B$, $k \geq 0$.

Then

$$d(E_A, E_B) \leq 3(C')^2 e^{2\chi} \Delta^{\frac{2\chi}{c_3+\chi}}. \tag{8.14}$$

Proof. We start by fixing $q := -(\chi + c_3)$, and let $k_0 := \lfloor \frac{\log \Delta}{q} \rfloor$, so that

$$(k_0 + 1)q < \log \Delta \leq k_0 q \leq \log \Delta - q$$

(recall that $q < 0$). In particular,

$$\Delta e^{c_3 k_0} \leq e^{k_0 q} e^{a k_0} = e^{-\chi k_0} \tag{8.15}$$

and, letting $\xi := \frac{2\chi}{c_3+\chi} = -\frac{2\chi}{q}$,

$$e^{-2\chi k_0} = (e^{q k_0})^{\frac{-2\chi}{q}} = (e^{q k_0})^{\xi} \leq e^{-q\xi} \Delta^{\xi} = e^{2\chi} \Delta^{\xi}, \tag{8.16}$$

where the inequality uses $k_0 q < -q + \log \Delta$. Then, by (8.11), for every $v_B \in E_B$ and every $k \geq 1$ we have

$$\|A_k v_B\| \leq \|B_k v_B\| + \|A_k - B_k\| \cdot \|v_B\| \leq C' e^{-\chi k} \|v_B\| + \Delta e^{c_3 k} \|v_B\|$$

and therefore, in particular, for $k = k_0$, by (8.15),

$$\|A_{k_0} v_B\| \leq C' e^{-\chi k_0} \|v_B\| + \Delta e^{c_3 k_0} \|v_B\| \leq 2C' e^{-\chi k_0} \|v_B\|.$$

This implies that

$$E_B \subset R_A := \{v \in \mathbb{R}^N : \|A_{k_0} v\| \leq 2C' e^{-\chi k_0} \|v\|\}.$$

Clearly we also have $E_A \subset R_A$ and therefore it is sufficient to estimate the "width" of R_A. For $v \in R_A$, write $v = v_A + v_A^{\perp}$, where $v_A \in E_A$ and $v_A^{\perp} \perp E_A$. Then by (8.12), for any $k \geq 1$ we have

$$\|A_k v\| \geq \|A_k v_A^{\perp}\| - \|A_k v_A\| \geq (C')^{-1} e^{\chi k} \|v_A^{\perp}\| - C' e^{-\chi k} \|v_A\|,$$

and therefore, for $k = k_0$, using also that $\|v_A\| \leq \|v\|$ since the the splitting of v is orthogonal, we get

$$\|v_A^{\perp}\| \leq C' e^{-\chi k_0} (\|A_{k_0} v\| + C' e^{-\chi k_0} \|v_A\|) \leq 3(C')^2 e^{-2\chi k_0} \|v\|,$$

which, by (8.16), implies $d(v, E_A) \leq 3(C')^2 e^{-2\chi k_0} \leq 3(C')^2 e^{2\chi} \Delta^{\xi}$ and therefore, from the definition of ξ, the conclusions of the Lemma. □

The following two Lemmas give the estimates we need to apply Lemma 8.6. Recall that c_1, c_2, c_3 are as in (1.5), that M is embedded in $\mathbb{R}^N$, and that $D_x^{(n)}$ are the matrices defined in Definition 8.5.

Lemma 8.7. [12, Lemma A.2] *There is $Q_7 \geq 1$ such that for all $x, y \in M$ and every $n \geq 1$, we have*

$$\|D_x^{(n)} - D_y^{(n)}\| \leq Q_7 e^{c_3 n} \|x - y\|^{\alpha}.$$

Proof. We prove the Lemma by induction on n. For $n = 1$, since f is $C^{1+\alpha}$ we have $\|D_x^{(1)} - D_y^{(1)}\| \leq |Df|_{\alpha} \|x - y\|^{\alpha}$ and therefore the statement holds for any $Q_7 \geq |Df|_{\alpha}$. Then, by the chain rule, we have

$$D_x^{(n+1)} - D_y^{(n+1)} = D_{f^n(x)}^{(1)} D_x^{(n)} - D_{f^n(y)}^{(1)} D_y^{(n)};$$

by adding and subtracting $D_{f^n(x)}^{(1)} D_y^{(n)}$, taking norms, using the inductive assumption and the fact that $\|f^n x - f^n y\| \leq e^{c_2 n} \|x - y\|$ for all $x, y \in M$, we get

$$\|D_x^{(n+1)} - D_y^{(n+1)}\| \leq \|D_{f^n(x)}^{(1)}\| \cdot \|D_x^{(n)} - D_y^{(n)}\| + \|D_{f^n(x)}^{(1)} - D_{f^n(y)}^{(1)}\| \cdot \|D_y^{(n)}\|$$

$$\le e^{c_2} Q_7 e^{c_3 n} \|x-y\|^\alpha + |Df|_\alpha e^{c_2 n\alpha} \|x-y\|^\alpha e^{c_2 n}$$
$$\le Q_7 e^{c_3(n+1)} \|x-y\|^\alpha \big(e^{c_2-c_3} + |Df|_\alpha Q_7^{-1} e^{(1+\alpha)c_2 n} e^{-c_3(n+1)} \big).$$

Since by (1.5), $c_3 > (1+\alpha)c_2$, we can choose Q_7 sufficiently large so that the quantity inside the brackets is less than 1 for every n, which completes the proof. □

Lemma 8.8. *Suppose A_k is a sequence of $N \times N$ matrices and $\mathbb{R}^N = E^s \oplus E^u$ is a splitting such that $\measuredangle(E^s, E^u) \ge K > 0$, and for every $k \ge 1$ and $v^u \in E^u$, $v^s \in E^s$ we have*

$$\|A_k v^s\| \le C e^{-\chi k} \|v^s\| \quad \text{and} \quad \|A_k v^u\| \ge C^{-1} e^{\chi k} \|v^u\|. \tag{8.17}$$

Then for every $w \perp E^s$ and every $k \ge 1$, we have

$$\|A_k w\| \ge (2C^2 K^{-1})^\gamma e^{\chi k} \|w\|, \tag{8.18}$$

where $\gamma := \frac{\chi - c_1}{2\chi}$ as in (1.6).

Proof. Writing $w = w^u + w^s$ where $w^u \in E^u, w^s \in E^s$, from (8.17) we get

$$\|A_k w\| \ge \|A_k w^u\| - \|A_k w^s\| \ge C^{-1} e^{\chi k} \|w^u\| - C e^{-\chi k} \|w^s\|.$$

Let $\theta = \measuredangle(w^s, w^u)$ and note that $\theta \ge K$ and since $w \perp w^s$, we have $\|w\| = \|w^u\| \sin\theta \le \|w^u\|$ and $\|w\| = \|w^s\| \tan\theta \ge \|w^s\| \tan K \ge \|w^s\| K$. Plugging this into the equation above, gives

$$\|D_x f^k w\| \ge (C^{-1} e^{\chi k} - C K^{-1} e^{-\chi k}) \|w\|. \tag{8.19}$$

Now fix $k_0 := \lfloor (2\chi)^{-1} \log(2C^2 K^{-1}) \rfloor$, Then for $k \le k_0$ we have

$$\frac{\|A_k w\|}{e^{\chi k}\|w\|} \ge \frac{e^{c_1 k}}{e^{\chi k}} \ge e^{(c_1-\chi)k_0} \ge e^{(c_1-\chi)(2\chi)^{-1}\log(2C^2K^{-1})} = (2C^2 K^{-1})^{\frac{c_1-\chi}{2\chi}}$$

where we recall that $c_1 < 0$ (see (1.5)). The formula for γ gives the required estimate.

It remains to treat $k > k_0$. In this case we have

$$e^{-2\chi k} \le e^{-2\chi(k_0+1)} \le e^{-\log(2C^2K^{-1})} = \tfrac{1}{2} C^{-2} K,$$

which gives

$$C K^{-1} e^{-\chi k} \le \frac{1}{2} C^{-1} e^{\chi k}$$

and hence, (8.19) gives

$$\|A_k w\| \ge (2C)^{-1} e^{\chi k} \|w\|.$$

Since $\gamma > 1$, $C \ge 1$, and $K \le 1$, we have

$$(2C^2K^{-1})^{-\gamma} \le (2C^2K^{-1})^{-1} = (2C)^{-1}(CK^{-1})^{-1} \le (2C)^{-1},$$

and thus we get the result in this case also, thus completing the proof. □

To complete the proof of Proposition 8.4, we will apply Lemma 8.6 with

$$A_k = D_x^{(k)}, \quad B_k = D_y^{(k)}, \quad \Delta = Q_7\|x-y\|^\alpha, \quad C' = (2C^2K^{-1})^\gamma.$$

To verify the conditions of Lemma 8.6, first note that Lemma 8.7 gives

$$\|A_k - B_k\| = \|D_x^{(k)} - D_y^{(k)}\| \le Q_7 e^{c_3 k}\|x-y\|^\alpha = \Delta e^{c_k k},$$

so that (8.11) holds. For (8.12), we observe that the hypotheses (8.4) and (8.5) of Proposition 8.4 give

$$\|A_k e_x^s\| \le Ce^{-\chi k} \quad \text{and} \quad \|A_k e_x^u\| \ge C^{-1}e^{\chi k}$$

for all $k \ge 0$, so that Lemma 8.8 can be applied with $v^s = e_x^s$ and $v^u = e_x^u$ to obtain

$$\|A_k v_A^\perp\| \ge (2C^2K^{-1})^\gamma e^{\chi k}\|v_A^\perp\| = C' e^{\chi k}\|v_A^\perp\|$$

for all $v_A^\perp \perp E_x^s$. This establishes (8.12), and (8.13) follows similarly. Thus Lemma 8.6 applies, and using (1.6) to write $\frac{2\chi}{c_3+\chi} = \frac{\beta}{\alpha}$, we have

$$d(E_x^s, E_y^s) \le 3(C')^2 e^{2\chi}\Delta^{\frac{\beta}{\alpha}} = 3(2C^2K^{-1})^{2\gamma}e^{2\chi}(Q_7 d(x,y)^\alpha)^{\frac{\beta}{\alpha}}, \tag{8.20}$$

which completes the proof of Proposition 8.4 by taking $Q_6 = 3(2)^{2\gamma}e^{2\chi}Q_7^{\beta/\alpha}$.

8.3. **Hölder continuity of Lyapunov coordinates.** In this section we prove that $s,u\colon \Lambda_\ell \to [\sqrt{2}, Q_0\widehat{Q}e^{\epsilon\ell}]$ are Hölder continuous with exponent ζ and constant given in terms of $e^{\epsilon\eta\ell}$, where ζ,η are given in (1.6). Observe that the Hölder exponent ζ depends on $\chi-\lambda$, and decays to 0 as $\lambda\to\chi$ where χ is the decay rate associated to the (χ,ϵ)-hyperbolic set Λ and $\lambda<\chi$ is the rate used in the definition of $s(x)$, $u(x)$.

Proposition 8.9. *There is $Q_8>0$ such that for any $\ell\in\mathbb{N}$ and $x,y\in\Lambda_\ell$, we have*

$$|s(x)-s(y)| \le Q_8 e^{\epsilon\eta\ell}d(x,y)^\zeta \text{ and } |u(x)-u(y)| \le Q_8 e^{\epsilon\eta\ell}d(x,y)^\zeta. \tag{8.21}$$

We give the argument for the upper bound for $|s(x)-s(y)|$; the argument for $|u(x)-u(y)|$ is analogous. Recall first that by definition,

$$s(x)^2 = 2\sum_{n\ge0} e^{2\lambda n}\|Df_x^n e_x^s\|^2.$$

Since $s(x), s(y) \ge 1$, we have

$$|s(x)-s(y)| \le \frac{|s(x)^2 - s(y)^2|}{2} \le \sum_{n\ge 0} e^{2\lambda n}\big|\|Df_x^n e_x^s\|^2 - \|Df_y^n e_y^s\|^2\big|. \tag{8.22}$$

Notice that $x,y\in\Lambda_\ell$ gives $\|Df_x^n e_x^s\| \le e^{\epsilon\ell}e^{-\chi n}$, $\|Df_y^n e_y^s\| \le e^{\epsilon\ell}e^{-\chi n}$, and so

$$\Delta_n = \Delta_n(x,y) := \big|\|Df_x^n e_x^s\|^2 - \|Df_y^n e_y^s\|^2\big| \le 2e^{2\epsilon\ell}e^{-2\chi n}. \tag{8.23}$$

Plugging this into (8.22) gives a uniform bound for $|s(x)-s(y)|$ but is not sufficient for our purposes since it does not include $d(x,y)$ and does not therefore imply Hölder continuity. It will nevertheless be useful to bound the tail of the sum for large values of n. For small n we need a more sophisticated estimate on Δ_n, as follows.

Lemma 8.10. *There is $Q_9 > 0$ such that for all $\ell \in \mathbb{N}$ and $x, y \in \Lambda_\ell$ we have*

$$\Delta_n \leq Q_9 e^{(2+\alpha\beta)c_2 n} e^{6\epsilon\gamma\alpha(\ell+n)} d(x,y)^{\alpha\beta}. \tag{8.24}$$

The proof of Lemma 8.10 uses the Hölder continuity of the hyperbolic splitting from Proposition 8.3 and so for clarity we isolate the specific estimate in which this property is used.

Sublemma 8.11. *There is $Q_{10} > 0$ such that for all $\ell \in \mathbb{N}$, $x, y \in \Lambda_\ell$, and $k \geq 0$, we have*

$$\big| \|Df_{f^k x} e^s_{f^k x}\| - \|Df_{f^k y} e^s_{f^k y}\| \big| \leq Q_{10} e^{6\epsilon\gamma\alpha(\ell+k)} d(f^k x, f^k y)^{\alpha\beta}. \tag{8.25}$$

Proof. Since Df is Hölder on TM we have

$$\big| \|Df_{f^k x} e^s_{f^k x}\| - \|Df_{f^k y} e^s_{f^k y}\| \big| \leq |Df|_\alpha d(e^s_{f^k x}, e^s_{f^k y})^\alpha$$

By (1.3), $f^k x, f^k y \in \Lambda_{\ell+k}$ and therefore, by Proposition 8.3,

$$d(e^s_{f^k x}, e^s_{f^k y})^\alpha \leq |Df|_\alpha (Q_6 e^{6\epsilon\gamma(\ell+k)} d(f^k x, f^k y)^\beta)^\alpha$$

which gives the result. □

Proof of Lemma 8.10. By (1.5) the norm of $\|Df\|$ is bounded above by e^{c_2}, and therefore, using the formula for the difference of two squares,

$$\Delta_n \leq 2e^{c_2 n} \big| \|Df^n_x e^s_x\| - \|Df^n_y e^s_y\| \big|. \tag{8.26}$$

Moreover, by the chain rule we have

$$\big| \|Df^n_x e^s_x\| - \|Df^n_y e^s_y\| \big| = \left| \prod_{k=0}^{n-1} \|Df_{f^k x} e^s_{f^k x}\| - \prod_{k=0}^{n-1} \|Df_{f^k y} e^s_{f^k y}\| \right| \tag{8.27}$$

and therefore, applying the standard equality for the difference of two products $\left|\prod_{k=0}^{n-1} a_k - \prod_{k=0}^{n-1} b_k\right| = \left|\sum_{k=0}^{n-1} a_0 ... a_{k-1}(a_k - b_k) b_{k+1} ... b_{n-1}\right|$, and using that the absolute value of each individual term is bounded by e^{c_2}, we get

$$\left| \prod_{k=0}^{n-1} \|Df_{f^k x} e^s_{f^k x}\| - \prod_{k=0}^{n-1} \|Df_{f^k y} e^s_{f^k y}\| \right| \leq e^{c_2 n} \sum_{k=0}^{n-1} \big| \|Df_{f^k x} e^s_{f_k x}\| - \|Df_{f^k y} e^s_{f^k y}\| \big|.$$

Substituting this into (8.27) and (8.26) and using (8.25), we get

$$\begin{aligned} \Delta_n &\leq 2e^{2c_2 n} \sum_{k=0}^{n-1} \big| \|Df_{f^k x} e^s_{f^k x}\| - \|Df_{f^k y} e^s_{f^k y}\| \big| \\ &\leq 2Q_{10} e^{2c_2 n} \sum_{k=0}^{n-1} e^{6\epsilon\gamma\alpha(\ell+k)} d(f^k x, f^k y)^{\alpha\beta}. \end{aligned} \tag{8.28}$$

Using the bound e^{c_2} for the derivative we get $d(f^k x, f^k y) \leq e^{kc_2} d(x,y)$ and therefore, plugging this into (8.28) and rearranging the terms we get

$$\Delta_n \leq 2Q_{10} e^{2c_2 n} e^{6\epsilon\gamma\alpha\ell} d(x,y)^{\alpha\beta} \sum_{k=0}^{n-1} e^{(6\epsilon\gamma + c_2\beta)\alpha k}.$$

To bound the geometric sum, we write

$$\sum_{k=0}^{n-1} e^{(6\epsilon\gamma+c_2\beta)\alpha k} = \frac{e^{(6\epsilon\gamma+c_2\beta)\alpha n}-1}{e^{(6\epsilon\gamma+c_2\beta)\alpha}-1} \le \frac{e^{(6\epsilon\gamma+c_2\beta)\alpha n}}{e^{(6\epsilon\gamma+c_2\beta)\alpha}-1}$$

and so we conclude that

$$\Delta_n \le \frac{2Q_{10}}{e^{(6\epsilon\gamma+c_2\beta)\alpha}-1} e^{2c_2 n} e^{6\epsilon\gamma\alpha\ell} d(x,y)^{\alpha\beta} e^{(6\epsilon\gamma+c_2\beta)\alpha n}$$

which gives the result. □

Proof of Proposition 8.9. We want to use (8.23) for large n, and (8.24) for small n; the transition happens at the point where the two bounds are roughly equal. Thus we choose N such that

$$e^{2\epsilon\ell}e^{-2\chi N} \approx e^{(2+\alpha\beta)c_2 N} e^{6\epsilon\gamma\alpha(\ell+N)} d(x,y)^{\alpha\beta};$$

more precisely, we take

$$N = \left\lfloor \frac{2\epsilon\ell - 6\epsilon\gamma\alpha\ell - \alpha\beta\log d(x,y)}{2\chi + (2+\alpha\beta)c_2 + 6\epsilon\gamma\alpha} \right\rfloor, \tag{8.29}$$

so that

$$\begin{aligned}(2\chi + (2+\alpha\beta)c_2 + 6\epsilon\gamma\alpha)N &\le 2\epsilon\ell - 6\epsilon\gamma\alpha\ell - \alpha\beta\log d(x,y) \\ &\le (2\chi + (2+\alpha\beta)c_2 + 6\epsilon_0\gamma\alpha)(N+1).\end{aligned} \tag{8.30}$$

Note that there is a number $\rho > 0$ which depends only on α, β, ℓ, and ϵ such that the numerator in (8.29) is positive provided $d(x,y) < \rho$. Continuing with this assumption our choice of N and the bound in (8.23) give

$$\begin{aligned}\sum_{n=N}^{\infty} e^{2\lambda n}\Delta_n &\le \sum_{n=N}^{\infty} 2e^{2\lambda n}e^{2\epsilon\ell}e^{-2\chi n} \\ &\le \frac{2}{1-e^{-2(\chi-\lambda)}} e^{2\epsilon\ell}e^{2(-\chi+\lambda)N} = Q_{11}e^{2\epsilon\ell}e^{2(-\chi+\lambda)(N+1)}\end{aligned}$$

where $Q_{11} = 2(1-e^{-2(\chi-\lambda)})^{-1}e^{2(\chi-\lambda)}$. Then the second inequality in (8.30) and the definitions of the constants ι, η, ζ in (1.6) give

$$\begin{aligned}\sum_{n=N}^{\infty} e^{2\lambda n}\Delta_n &\le Q_{11}e^{2\epsilon\ell}e^{-2(\chi-\lambda)\frac{2\epsilon\ell-6\epsilon\gamma\alpha\ell-\alpha\beta\log d(x,y)}{6\epsilon_0\gamma\alpha+(2+\alpha\beta)c_2+2\chi}} \\ &= Q_{11}e^{2\epsilon\ell}e^{-\iota(2\epsilon\ell-6\epsilon\gamma\alpha\ell-\alpha\beta\log d(x,y))} \\ &= Q_{11}e^{(2(1-\iota)+6\gamma\alpha\iota)\epsilon\ell}d(x,y)^{\alpha\beta\iota} \le Q_{11}e^{\epsilon\eta\ell}d(x,y)^{\zeta},\end{aligned} \tag{8.31}$$

where the last inequality uses the fact that $2(1-\iota) + 6\gamma\alpha\iota \le 2 + 6\gamma\alpha\iota = \eta$. Turning our attention to the finite part of the sum, (8.24) gives

$$\begin{aligned}\sum_{n=0}^{N-1} e^{2\lambda n}\Delta_n &\le Q_9 \sum_{n=0}^{N-1} e^{2\lambda n}e^{(2+\alpha\beta)c_2 n}e^{6\epsilon\gamma\alpha(\ell+n)}d(x,y)^{\alpha\beta} \\ &\le Q_{12}e^{6\epsilon\gamma\alpha\ell}e^{(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\lambda)N}d(x,y)^{\alpha\beta} \\ &= Q_{12}e^{6\epsilon\gamma\alpha\ell}e^{(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\chi)(1-\iota)N}d(x,y)^{\alpha\beta}\end{aligned}$$

for some constant Q_{12} independent of ℓ, x, y. Applying the first inequality in (8.30) gives

$$e^{6\epsilon\gamma\alpha\ell}e^{(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\chi)N}d(x,y)^{\alpha\beta} \le e^{2\epsilon\ell},$$

and thus

$$\sum_{n=0}^{N-1} e^{2\lambda n}\Delta_n \le Q_{12}e^{2\epsilon\ell}e^{-\iota(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\chi)N}$$
$$= Q_{13}e^{2\epsilon\ell}e^{-\iota(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\chi)(N+1)}$$

for $Q_{13} = Q_{12}e^{\iota(6\epsilon\gamma\alpha+(2+\alpha\beta)c_2+2\chi)}$. Now the second inequality in (8.30) gives

$$\sum_{n=0}^{N-1} e^{2\lambda n}\Delta_n \le Q_{13}e^{2\epsilon\ell}e^{-\iota(2\epsilon\ell-6\epsilon\gamma\alpha\ell-\alpha\beta\log d(x,y))}$$
$$= Q_{13}e^{(2(1-\iota)+6\gamma\alpha\iota)\epsilon\ell}d(x,y)^{\alpha\beta\iota} \le Q_{13}e^{\eta\epsilon\ell}d(x,y)^{\zeta},$$

where the last inequality again uses $2(1-\iota)+6\gamma\alpha\iota \le \eta$. Adding (8.31) and using (8.22) completes the proof of Proposition 8.9 in the case $d(x,y) < \rho$. When $d(x,y) \ge \rho$, (8.23) gives

$$\Delta_n \le 2e^{2\epsilon\ell}e^{-2\chi n}\rho^{-\zeta}d(x,y)^{\zeta} \tag{8.32}$$

and thus (8.22) gives

$$|s(x)-s(y)| \le \sum_{n\ge 0} e^{2\lambda n}\Delta_n \le \frac{2e^{2\epsilon\ell}\rho^{-\zeta}}{1-e^{-2(\chi-\lambda)}}d(x,y)^{\zeta},$$

which completes the proof because $\eta \ge 1$, so $e^{2\epsilon\ell} \le e^{2\epsilon\eta\ell}$. □

8.4. **Overlapping derivative estimates.** We are now ready to begin the proof of Theorem H. We consider two points $x, y \in \Lambda_\ell$ with the property that $d(x,y) \le \delta e^{-\lambda\ell}$, as in (8.3), and prove that the corresponding regular neighbourhoods at level ℓ are overlapping, subject to certain conditions on δ. Crucially, these conditions will not depend on ℓ.

In this section we prove the derivative estimates involved in the definition of overlapping charts. In fact, we will prove here a slightly stronger version of (8.1) by showing that (8.1) holds for all $\underline{z} \in \Psi_x^{-1}(\mathcal{N}_x \cap \mathcal{N}_y)$. The analogous statement (8.2) is completely symmetric.

For $x, y \in \Lambda_\ell$ with $\mathcal{N}_x \cap \mathcal{N}_y \ne \emptyset$, let $z \in \mathcal{N}_x \cap \mathcal{N}_y$ and denote

$$\underline{z}_x := \Psi_x^{-1}(z) \in \mathcal{B}_x, \quad \text{and} \quad \underline{z}_y := \Psi_y^{-1}(z) \in \mathcal{B}_y.$$

Then $(\Psi_y^{-1} \circ \Psi_x)(\underline{z}_x) = \underline{z}_y$ and $D_{\underline{z}_x}(\Psi_y^{-1} \circ \Psi_x)\colon T_{\underline{z}_x}\mathcal{B}_x \to T_{\underline{z}_y}\mathcal{B}_x$. We consider the standard coordinates given by the orthogonal basis (e_1, e_2) in $T_{\underline{z}_x}\mathcal{B}_x$ and $T_{\underline{z}_y}\mathcal{B}_x$ and consider an unstable vector

$$\underline{v}_x^u \in \tilde{K}^u \subset T_{\underline{z}_x}\mathcal{B}_x$$

which we assume is normalized, so that $\|\underline{v}_x^u\| = 1$, and which we write as

$$\underline{v}_x^u = \underline{v}_{1,x}^u e_1 + \underline{v}_{2,x}^u e_2.$$

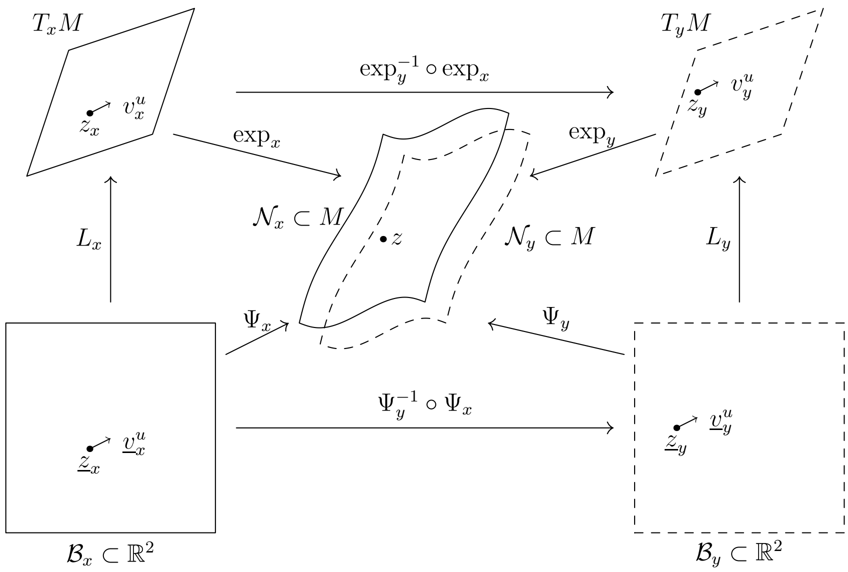

FIGURE 8.1. Applying the transition map.

Then let

$$\underline{v}_y^u := D_{\underline{z}_x}(\Psi_y^{-1} \circ \Psi_x)(\underline{v}_x^u) = \underline{v}_{1,y}^u e_1 + \underline{v}_{2,y}^u e_2.$$

We will estimate the absolute values of $\underline{v}_{1,y}^u, \underline{v}_{2,y}^u$ in order to prove (8.1). Consider unit vectors $e_x^u, e_x^s \in T_x^M$ and $e_y^u, e_y^s \in T_y^M$, in the directions given by the hyperbolic splitting. Throughout this section, we write $d_{x,y} := d(x,y)$, $u_x = u(x), s_x = s(x)$ to make our computations more compact and easier to read. Observe that

$$\Psi_y^{-1} \circ \Psi_x = L_y^{-1} \circ \exp_y^{-1} \circ \exp_x \circ L_x.$$

Use $\{e_x^u, e_x^s\}$ as a basis for each tangent space to T_xM in the obvious way, and similarly for T_y^M. With respect to these bases and the standard basis $\{e_1, e_2\}$, the derivatives of the maps L_y^{-1}, $\exp_y^{-1} \circ \exp_x$, and L_x are represented by the matrices

$$\begin{pmatrix} u_y & 0 \\ 0 & s_y \end{pmatrix}, \quad \begin{pmatrix} \xi_1^u & \xi_1^s \\ \xi_2^u & \xi_2^s \end{pmatrix}, \quad \begin{pmatrix} u_x^{-1} & 0 \\ 0 & s_x^{-1} \end{pmatrix},$$

respectively, where $\xi_{1/2}^{s/u} \in \mathbb{R}$ are determined by

$$\begin{aligned} D_{L_x(\underline{z}_x)}(\exp_y^{-1} \circ \exp_x) e_x^u &= \xi_1^u e_y^u + \xi_2^u e_y^s, \\ D_{L_x(\underline{z}_x)}(\exp_y^{-1} \circ \exp_x) e_x^s &= \xi_1^s e_y^u + \xi_2^s e_y^s. \end{aligned} \tag{8.33}$$

Thus $D_{\underline{z}_x}(\Psi_y^{-1} \circ \Psi_x)$ has matrix (with respect to $\{e_1, e_2\}$) given by the product of these matrices, which is

$$\begin{pmatrix} u_y & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} u_x^{-1}\xi_1^u & s_x^{-1}\xi_1^s \\ u_x^{-1}\xi_2^u & s_x^{-1}\xi_2^s \end{pmatrix} = \begin{pmatrix} u_y u_x^{-1}\xi_1^u & u_y s_x^{-1}\xi_1^s \\ s_y u_x^{-1}\xi_2^u & s_y s_x^{-1}\xi_2^s \end{pmatrix},$$

and we conclude that

$$\underline{v}_{1,y}^u = u_y u_x^{-1} \xi_1^u \underline{v}_{1,x}^u + u_y s_x^{-1} \xi_1^s \underline{v}_{2,x}^u, \tag{8.34}$$

$$\underline{v}^u_{2,y} = s_y u_x^{-1} \xi^u_2 \underline{v}^u_{1,x} + s_y s_x^{-1} \xi^s_2 \underline{v}^u_{2,x}. \tag{8.35}$$

We now collect the various estimates which we will plug into these equations to estimate the norms of $\underline{v}^u_{1,y}, \underline{v}^u_{2,y}$.

Lemma 8.12. *There exists a constant $Q_{14} > 0$ such that for every $\ell \in \mathbb{N}$ and $x, y \in \Lambda_\ell$ satisfying* (8.3)*, we have*

$$|\underline{v}^u_{1,x}| \geq 1/\sqrt{2} \geq 1/2 \quad \text{and} \quad |\underline{v}^u_{2,x}| \leq e^{-\lambda}\omega|\underline{v}^u_{1,x}|, \tag{8.36}$$

$$|\xi^u_2|, |\xi^s_1| \leq Q_{14} e^{6\epsilon\gamma\ell} d^\beta_{x,y} \quad \text{and} \quad |1 - \xi^u_1|, |1 - \xi^s_2| \leq Q_{14} e^{6\epsilon\gamma\ell} d^\beta_{x,y}, \tag{8.37}$$

$$Q_{14} e^{-\epsilon\ell} \leq u_y s_x^{-1} \leq Q_{14} e^{\epsilon\ell} \quad \text{and} \quad Q_{14} e^{-\epsilon\ell} \leq s_y u_x^{-1} \leq Q_{14} e^{\epsilon\ell}, \tag{8.38}$$

$$1 - Q_{14} e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y} \leq u_y u_x^{-1} \leq 1 + Q_{14} e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y}, \tag{8.39}$$

$$1 - Q_{14} e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y} \leq s_y s_x^{-1} \leq 1 + Q_{14} e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y}. \tag{8.40}$$

Proof. Equation (8.36) follows immediately from the fact that $\underline{v}^u_x$ is a unit vector and $\underline{v}^u_x \in \tilde{K}^u$. Equation (8.37) follows from (8.33) and Proposition 8.3 which gives quantitative control on the Hölder dependence of the stable and unstable directions on the base point in Λ_ℓ. Equation (8.38) follows immediately from (6.1). Finally, by (6.1) and Proposition 8.9,

$$\frac{u_y}{u_x} = 1 + \frac{u_y - u_x}{u_x} \geq 1 - \frac{|u_y - u_x|}{u_x} \geq 1 - \frac{Q_8 e^{\epsilon\eta\ell} d^\zeta_{x,y}}{Q_0 \widehat{Q} e^{\epsilon\ell}},$$

which gives the first half of (8.39). The upper bound and (8.40) are similar. □

We are now ready to start estimating the two components $\underline{v}^u_{1,y}, \underline{v}^u_{2,y}$ of $\underline{v}^u_y$. We estimate each one separately. Once we have proved these lemmas, we will be in a position to give the conditions on $\delta > 0$, which we stress will be independent of ℓ.

Lemma 8.13. *For every $\ell \in \mathbb{N}$ and $x, y \in \Lambda_\ell$ satisfying* (8.3)*, we have*

$$|\underline{v}^u_{1,y}| \geq (1 - Q_{14}\delta^\beta)(1 - Q_{14}\delta^\zeta)|\underline{v}^u_{1,x}| - Q^2_{14}\delta^\beta.$$

Proof. Using (8.37) and (8.39), we have the following estimate for the first term of (8.34):

$$|u_y u_x^{-1} \underline{v}^u_{1,x} \xi^u_1| \geq (1 - Q_{14} e^{6\epsilon\gamma\ell} d^\beta_{x,y})(1 - Q_{14} e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y})|\underline{v}^u_{1,x}|.$$

Now (8.3) and the bounds on ϵ in (1.7) give

$$\begin{aligned} e^{6\epsilon\gamma\ell} d^\beta_{x,y} &\leq e^{6\epsilon\gamma\ell}\delta^\beta e^{-\beta\lambda\ell} = \delta^\beta e^{(6\epsilon\gamma - \beta\lambda)\ell} \leq \delta^\beta, \\ e^{(\eta-1)\epsilon\ell} d^\zeta_{x,y} &\leq e^{(\eta-1)\epsilon\ell}\delta^\zeta e^{-\zeta\lambda\ell} = \delta^\zeta e^{((\eta-1)\epsilon - \zeta\lambda)\ell} \leq \delta^\zeta, \end{aligned} \tag{8.41}$$

and thus

$$|u_y u_x^{-1} \xi^u_1 \underline{v}^u_{1,x}| \geq (1 - Q_{14}\delta^\beta)(1 - Q_{14}\delta^\zeta)|\underline{v}^u_{1,x}|. \tag{8.42}$$

For the second term of (8.34), by (8.37) and (8.38), and using $|\underline{v}^u_{2,x}| \leq 1$, we have

$$|u_y s_x^{-1} \xi^s_1 \underline{v}^u_{2,x}| \leq Q^2_{14} e^{\epsilon\ell} e^{6\epsilon\gamma\ell} d^\beta_{x,y} \leq Q^2_{14} e^{7\epsilon\gamma\ell} d^\beta_{x,y},$$

where the second inequality uses the fact that $\gamma \geq 1$. As in (8.41) above,

$$e^{7\epsilon\gamma\ell} d_{x,y}^{\beta} \leq \delta^{\beta} e^{(7\epsilon\gamma - \beta\lambda)\ell} \leq \delta^{\beta}, \tag{8.43}$$

and thus $|u_y s_x^{-1} \xi_1^s \underline{v}_{2,x}^u| \leq Q_{14}^2 \delta^{\beta}$. Subtracting this from (8.42) and recalling (8.34) proves the lemma. ☐

Lemma 8.14. *For every $\ell \in \mathbb{N}$ and $x, y \in \Lambda_\ell$ satisfying* (8.3)*, we have*

$$|\underline{v}_{2,y}^u| \leq (1 + Q_{14}\delta^{\beta})(1 + Q_{14}\delta^{\zeta})|\underline{v}_{2,x}^u| + Q_{14}^2\delta^{\beta}.$$

Proof. The proof is nearly identical to the proof of the previous Lemma. Let us bound the first term in (8.35) as follows:

$$\begin{aligned} |s_y s_x^{-1} \xi_2^s \underline{v}_{2,x}^u| &\leq (1 + Q_{14} e^{(\eta-1)\epsilon\ell} d_{x,y}^{\zeta})(1 + Q_{14} e^{6\epsilon\gamma\ell} d_{x,y}^{\beta})|\underline{v}_{2,x}^u| \\ &\leq (1 + Q_{14}\delta^{\zeta})(1 + Q_{14}\delta^{\beta})|\underline{v}_{2,x}^u|. \end{aligned}$$

Similar computations for the second term of (8.35) give

$$|s_y u_x^{-1} \xi_2^u \underline{v}_{1,x}^u| \leq Q_{14}^2 e^{\epsilon\ell} e^{6\epsilon\gamma\ell} d_{x,y}^{\beta} \leq Q_{14}^2 \delta^{\beta}.$$

Adding these estimates together proves the lemma. ☐

Proof of Theorem H (derivative estimates). We can now prove the first part of Theorem H concerning properties (8.1) and (8.2) in the definition of overlapping charts. Start by requiring that $\delta > 0$ is sufficiently small that $Q_{14}\delta^{\beta} < 1$ and $Q_{14}\delta^{\zeta} < 1$. (Further conditions will come later.) Lemmas 8.13 and 8.14 give

$$\begin{aligned} \frac{|\underline{v}_{2,y}^u|}{|\underline{v}_{1,y}^u|} &\leq \frac{(1 + Q_{14}\delta^{\beta})(1 + Q_{14}\delta^{\zeta})|\underline{v}_{2,x}^u| + Q_{14}^2\delta^{\beta}}{(1 - Q_{14}\delta^{\beta})(1 - Q_{14}\delta^{\zeta})|\underline{v}_{1,x}^u| - Q_{14}^2\delta^{\beta}} \\ &\leq \frac{(1 + Q_{14}\delta^{\beta})(1 + Q_{14}\delta^{\zeta})e^{-\lambda}\omega|\underline{v}_{1,x}^u| + Q_{14}^2\delta^{\beta}}{(1 - Q_{14}\delta^{\beta})(1 - Q_{14}\delta^{\zeta})|\underline{v}_{1,x}^u| - Q_{14}^2\delta^{\beta}}, \end{aligned}$$

where the second inequality uses the fact that $\underline{v}_x^u \in \tilde{K}^u$. Note that the function $t \mapsto \frac{at+b}{ct+d}$ is decreasing in t when $ad - bc < 0$, which is the case for the expression above, and thus we can obtain an upper bound by observing that (8.36) gives $|\underline{v}_{1,x}^u| \geq \frac{1}{2}$, so monotonicity gives

$$\frac{|\underline{v}_{2,y}^u|}{|\underline{v}_{1,y}^u|} \leq \frac{(1 + Q_{14}\delta^{\beta})(1 + Q_{14}\delta^{\zeta})e^{-\lambda}\omega/2 + Q_{14}^2\delta^{\beta}}{(1 - Q_{14}\delta^{\beta})(1 - Q_{14}\delta^{\zeta})/2 - Q_{14}^2\delta^{\beta}}. \tag{8.44}$$

For sufficiently small $\delta > 0$, the right-hand side is $< \omega$, which implies that $\underline{v}_y^u \in K^u$ as required by the first part of (8.1).

For the second part of (8.1), we observe that

$$\frac{\|\underline{v}_y^u\|}{\|\underline{v}_x^u\|} \geq \frac{|\underline{v}_{1,y}^u|}{\sqrt{|\underline{v}_{1,x}^u|^2 + |\underline{v}_{2,x}^u|^2}} \geq \frac{\big((1 - Q_{14}\delta^{\beta})(1 - Q_{14}\delta^{\zeta}) - Q_{14}^2\delta^{\beta}\big)|\underline{v}_{1,x}^u|}{\sqrt{1 + e^{-2\lambda}\omega^2}|\underline{v}_{1,x}^u|},$$

where the second inequality uses Lemma 8.13 and (8.36) for the numerator and the fact that $\underline{v}_x^u \in \tilde{K}^u$ for the denominator. Recall from (5.9) that

$1/\sqrt{1+e^{-2\lambda}\omega^2} > e^{-\lambda/24}$; thus we can choose δ small enough that the right-hand side of the above expression is $> e^{-\lambda/24}$, which proves the second half of (8.1). Condition (8.2) follows by analogous arguments. $\square$

8.5. **Overlapping stable and unstable strips.** We now complete the proof of Theorem H by showing that if $\gamma \subset \widetilde{\mathcal{N}}^s_{x,\ell}$ is a full length strongly stable admissible curve, then it completely crosses $\widetilde{\mathcal{N}}^u_{y,\ell}$. The other three required conditions obtained by interchanging x/y and stable/unstable are proved analogously.

We start with a couple of simple Lemmas relating the distance $d(x,y)$ between two points and the amount of overlap of their regular neighbourhoods.

Lemma 8.15. *There exists a constant $Q_{15} > 0$ such that for every $\ell \geq 1$ and every $x, y \in \Lambda_\ell$ satisfying $d(x,y) \leq Q_{15}e^{-2\epsilon\ell}b_\ell$, we have $x, y \in \mathcal{N}^{(\ell)}_x \cap \mathcal{N}^{(\ell)}_y$.*

Proof. Recall that the regular neighbourhood $\mathcal{N}^{(\ell)}_x$ is given by $\mathcal{N}^{(\ell)}_x = \Psi_x(\mathcal{B}^{(\ell)}_x) = \exp_x(L_x(\mathcal{B}^{(\ell)}_x))$, where $\mathcal{B}^{(\ell)}_x = [-b_\ell, b_\ell]^2 \subset \mathbb{R}^2$ and L_x is given by (5.2). In particular, $L_x(\mathcal{B}^{(\ell)}_x)$ is a parallelogram in T_xM centered at 0 with side lengths $2b_\ell/u(x)$ and $2b_\ell/s(x)$. Using the upper bounds in (6.1) we see that both of these side lengths are at least $2(Q_0\widehat{Q})^{-1}e^{-\epsilon\ell}b_\ell$. Moreover, by condition (**H2**) and the definition of Λ_ℓ in (1.2) we have $\measuredangle(E^s_x, E^u_x) \geq e^{-\epsilon\ell}$.

Consider a parallelogram whose sides have length A, B and meet at an angle $\theta \in (0, \pi/2]$. The distance between each pair of opposite edges is at least $\min(A,B)\cdot\sin\theta \geq \min(A,B)\cdot 2\theta/\pi$, and thus the parallelogram contains a ball of radius $\min(A,B)\theta/\pi$ around its center. In the setting of the previous paragraph, this shows that $L_x(\mathcal{B}^{(\ell)}_x)$ contains a ball in T_xM centered at the origin with radius $2(Q_0\widehat{Q})^{-1}\pi^{-1}e^{-2\epsilon\ell}b_\ell$.

Choose $r > 0$ such that given any $x \in M$ and $v \in T_xM$ with $\|v\| \geq r$, we have $d(x, \exp_x(v)) \geq r/2$. Then choosing $Q_{15} > 0$ sufficiently small that $2Q_{15}b_0 < r$ and $Q_{15} \leq (Q_0\widehat{Q})^{-1}\pi^{-1}$, we see that $L_x(\mathcal{B}^{(\ell)}_x)$ contains a ball in T_xM centered at the origin with radius $Q_{15}2Q_{15}e^{-2\epsilon\ell}b_\ell$, and since $e^{-2\epsilon\ell}b_\ell \leq b_0$, the image of this ball under $\exp_x$ contains $B(x, Q_{15}e^{-2\epsilon\ell}b_\ell)$. This shows that $B(x, Q_{15}e^{-2\epsilon\ell}b_\ell) \subset \exp_x(L_x(\mathcal{B}^{(\ell)}_x)) = \Psi_x(\mathcal{B}^{(\ell)}_x) = \mathcal{N}^{(\ell)}_x$, which completes the proof. $\square$

Lemma 8.16. *There exists $Q_{16} > 0$ such that for $\ell \geq 1$ and $x, y \in \Lambda_\ell$,*

$$\textit{if } y \in \mathcal{N}^{(\ell)}_x \quad \textit{then} \quad \|\Psi_x^{-1} \circ \Psi_y(0)\| \leq Q_{16}e^{2\epsilon\ell}d(x,y).$$

Proof. If $y \in \mathcal{N}^{(\ell)}_x$ then $\Psi_x^{-1} \circ \Psi_y(0) = \Psi_x^{-1}(y)$ is well defined. Therefore $\|\Psi_x^{-1} \circ \Psi_y(0)\| = \|\Psi_x^{-1}(y)\| = \|\Psi_x^{-1}(y) - 0\| = \|\Psi_x^{-1}(y) - \Psi_x^{-1}(x)\|$ and so we just need to estimate the Lipschitz constant of Ψ_x^{-1}. By definition we have $\Psi_x^{-1} = L_x^{-1} \circ \exp_x^{-1}$ and the result follows using (6.6) in Lemma 6.4 and the fact that $\exp_x^{-1}$ is close to an isometry. $\square$

Proof of Theorem H (stable and unstable strips). We can now complete the proof of Theorem H. Let $x, y \in \Lambda_\ell$ with $d(x,y) \leq \delta e^{-\lambda\ell}$ as in (8.3). Then

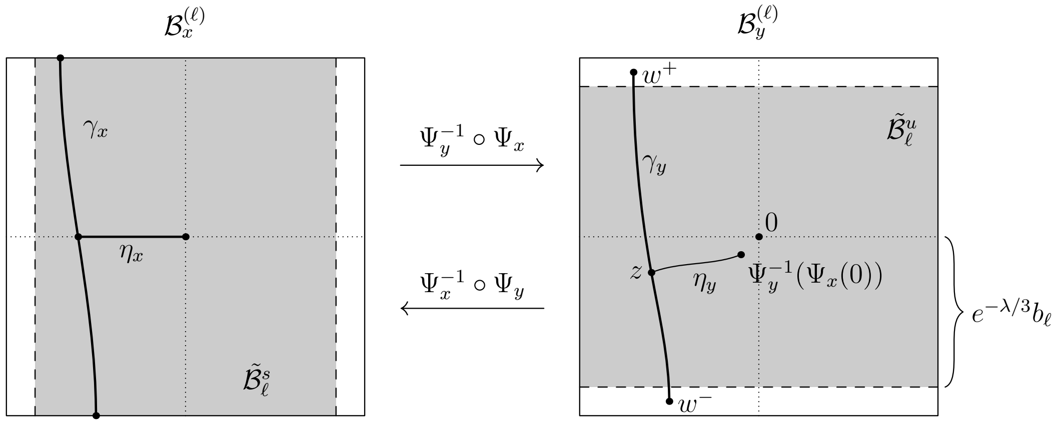

FIGURE 8.2. Proving Theorem H.

since $b_\ell = Ce^{-2\epsilon\ell/\alpha}$ for some constant $C > 0$, we have

$$d(x,y)e^{2\epsilon\ell}b_\ell^{-1} \le \delta e^{-\lambda\ell}e^{2\epsilon\ell}C^{-1}e^{2\epsilon\ell/\alpha} = \delta C^{-1}e^{(2\epsilon(1+\frac{1}{\alpha})-\lambda)\ell} \le \delta C^{-1}, \tag{8.45}$$

where the last inequality uses (1.7). By making δ sufficiently small that $\delta C^{-1} \le Q_{15}$, we guarantee that the hypothesis of Lemma 8.15 is satisfied.

Now let γ be a full length strongly stable admissible curve in $\tilde{\mathcal{N}}_{x,\ell}^s$, and consider the curves $\gamma_x = \Psi_x^{-1}(\gamma) \subset \tilde{\mathcal{B}}_\ell^s$ and $\gamma_y = \Psi_y^{-1}(\gamma) \subset \mathcal{B}_y^{(\ell)} \subset \mathcal{B}_y^{(\ell)}$, as shown in Figure 8.2. Let η_x be the segment of the x-axis in $\mathcal{B}_x^{(\ell)}$ that connects γ_x to 0. Let $\eta_y = \Psi_y^{-1} \circ \Psi_x(\eta_x)$. Let $z \in \mathcal{B}_y^{(\ell)}$ be the intersection point of η_y and γ_y, and let $w^\pm$ be the endpoints of γ_y. Writing $z = z_1e_1 + z_2e_2$ and similarly for $w^\pm$, our goal is to show that $|w_2^\pm| \ge e^{-\lambda/3}b_\ell$.

We give the proof for w^+; the proof for w^- is similar. By Lemma 8.16 and (8.45), the point $v := \Psi_y^{-1}(x) = \Psi_y^{-1}(\Psi_x(0))$ has

$$\|v\| \le Q_{16}e^{2\epsilon\ell}d(x,y) \le Q_{16}C^{-1}\delta b_\ell. \tag{8.46}$$

By (8.1), η_y is an unstable admissible curve connecting z and v with length $\le e^{\lambda/24}e^{-\lambda/3}b_\ell$ (using the fact that η_x has length at most $e^{-\lambda/3}b_\ell$). Since γ_y is a stable admissible curve, we have

$$|w_1^+ - z_1| \le 2\omega b_\ell,$$

and thus

$$|w_1^+| \le |v_1| + |z_1 - v_1| + |w_1^+ - z_1| \le Q_{16}C^{-1}\delta b_\ell + e^{\lambda/24}e^{-\lambda/3}b_\ell + 2\omega b_\ell.$$

By (5.9) we can choose $\delta > 0$ small enough that

$$Q_{16}C^{-1}\delta + e^{\lambda/24}e^{-\lambda/3} + 2\omega < 1, \tag{8.47}$$

and we conclude that $|w_1^+| < b_\ell$, so w^+ is not on a vertical boundary of $\mathcal{B}_y^{(\ell)}$.

If w^+ is on the top boundary of $\mathcal{B}_y^{(\ell)}$, then there is nothing to prove, so we can assume that the part of γ_y running from z to w^+ is the image of the top half of γ_x under the transition map. By (8.1), this part of γ_y is a stable admissible curve with length $\ge e^{-\lambda/24}b_\ell$ (using the fact that the top half of γ_x has length at least b_ℓ). Since the length of this part of γ_y is at most $\sqrt{1+\omega^2}|w_2^+ - z_2|$, we conclude that

$$|w_2^+ - z_2| \ge \frac{e^{-\lambda/24} b_\ell}{\sqrt{1+\omega^2}}. \tag{8.48}$$

As argued above, η_y has length $\le e^{\lambda/24} b_\ell$; on the other hand since it is a stable admissible curve, its length is at least

$$\sqrt{|z_1 - v_1|^2 + |z_2 - v_2|^2} \ge \sqrt{\omega^{-2}+1}|z_2 - v_2|,$$

and we conclude that

$$|z_2 - v_2| \le \frac{\omega e^{\lambda/24} b_\ell}{\sqrt{1+\omega^2}}. \tag{8.49}$$

Combining (8.46), (8.48), and (8.49) gives

$$\begin{aligned} |w_2^+| \ge |w_2^+ - z_2| - |z_2 - v_2| - |v_2| &\ge \frac{e^{-\lambda/24} b_\ell}{\sqrt{1+\omega^2}} - \frac{\omega e^{\lambda/24} b_\ell}{\sqrt{1+\omega^2}} - Q_{16} C^{-1} \delta b_\ell \\ &= \Big(\frac{e^{-\lambda/24} - \omega e^{\lambda/24}}{\sqrt{1+\omega^2}} - Q_{16} C^{-1} \delta \Big) b_\ell \ge (e^{-\lambda/4} - Q_{16} C^{-1} \delta) b_\ell, \end{aligned}$$

where the last inequality uses (5.9). As long as δ is small enough that

$$e^{-\lambda/4} - Q_{16} C^{-1} \delta > e^{-\lambda/3}, \tag{8.50}$$

this gives $|w_2^+| \ge e^{-\lambda/3} b_\ell$, completing the proof. □

9. Pseudo-orbits, branches, shadowing: Proofs of Theorems E and F

We are now ready to prove Theorems E and F.

9.1. **Regular branches: Proof of Theorem E.**

The two fundamental ingredients in the proof of Theorem E are Theorem H and Theorem G, which is essentially the special case of Theorem E where $k = 1$ and the pseudo-orbit is in fact a real orbit.

Fix $\delta > 0$ sufficiently small so that the conclusions of Theorem H hold. To prove the first part of Theorem E, we observe that if $\bar{x} = (x_0, \dots, x_k)$ is an $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit, then by Theorem H, the Lyapunov charts $\mathcal{N}_{x_j}^{(\ell_j)}$ and $\mathcal{N}_{f(x_{j-1})}^{(\ell_j)}$ are overlapping for every $1 \le j \le k$. By Theorem G, for each $0 \le j < k$, the sets

$$\begin{aligned} \mathcal{B}^{s,1}_{x_j,(\ell_j,\ell_{j+1})} &= \Psi_{x_j}^{-1}(\mathcal{N}_{x_j}^{(\ell_j)} \cap f^{-1} \mathcal{N}_{f(x_j)}^{\ell_{j+1}}), \\ \mathcal{B}^{u,1}_{f(x_j),(\ell_j,\ell_{j+1})} &= \Psi_{f(x_j)}^{-1}(f(\mathcal{N}_{x_j}^{(\ell_j)}) \cap \mathcal{N}_{f(x_j)}^{(\ell_{j+1})}) \end{aligned}$$

from (7.1) are strongly stable and strongly unstable strips in $\widetilde{\mathcal{B}}^s_{b_{\ell_j}}$ and $\widetilde{\mathcal{B}}^u_{b_{\ell_{j+1}}}$, respectively, and f_{x_j} is a diffeomorphism between them that satisfies the inclusions and estimates in (7.3). Since the Lyapunov charts $\mathcal{N}_{x_{j+1}}^{(\ell_{j+1})}$ and $\mathcal{N}_{f(x_j)}^{(\ell_{j+1})}$ are overlapping, we conclude that $\Psi_{x_{j+1}}^{-1} \circ \Psi_{f(x_j)}(\mathcal{B}^{u,1}_{f(x_j),(\ell_j,\ell_{j+1})})$ is an unstable strip in $\mathcal{B}_{x_{j+1}}^{(\ell_{j+1})}$, and thus its preimage under $\Psi_{x_{j+1}}^{-1} \circ f \circ \Psi_{x_j}$ is a stable strip in $\mathcal{B}_{x_j}^{(\ell_j)}$. Thus we have

$$\mathcal{B}_{x_0}^{(\ell_0)} \xrightarrow{\Psi_{x_0}^{-1}\circ f\circ\Psi_{x_1}} \mathcal{B}_{x_1}^{(\ell_1)} \xrightarrow{\Psi_{x_1}^{-1}\circ f\circ\Psi_{x_2}} \mathcal{B}_{x_2}^{(\ell_2)} \to \cdots \to \mathcal{B}_{x_{k-1}}^{(\ell_{k-1})} \xrightarrow{\Psi_{x_{k-1}}^{-1}\circ f\circ\Psi_{x_k}} \mathcal{B}_{x_k}^{(\ell_k)}$$

where the maps are not defined on the entirety of the indicated domain, but only on a stable strip, and the corresponding image is an unstable strip. In particular, taking the composition of all the maps we see that (5.12) with $j = 0$ defines a stable strip $\mathcal{B}_{\bar{x}}^0 \subset \mathcal{B}_{x_0}^{(\ell_0)}$ that is mapped to an unstable strip $\mathcal{B}_{\bar{x}} \subset \mathcal{B}_{x_k}^{(\ell_k)}$ by $\Psi_{x_k}^{-1} \circ f^k \circ \Psi_{x_0}$. This proves the first property in the definition of an $\bar{\ell}$-regular branch. For the second, we observe that by (7.3), each $f_{x_{j-1}} = \Psi_{f(x_{j-1})}^{-1} \circ f \circ \Psi_{x_{j-1}}$ has a derivative that maps K^u into $\widetilde{K}^u$, and that the transition map $\Psi_{x_j}^{-1} \circ \Psi_{f(x_{j-1})}$ maps this into K^u by the definition of overlapping charts; moreover, the first map above expands each vector in K^u by a factor of at least $e^{\lambda/2}$, and so after composing with the transition map, the derivative of $\Psi_{x_j}^{-1} \circ f \circ \Psi_{x_{j-1}}$ expands each vector in K^u by a factor of at least $e^{\lambda/2}e^{-\lambda/24}$. Iterating completes the proof of Theorem E.

9.2. **Shadowing: Proof of Theorem F.** We prove Theorem F using Theorem E and the definition of regular branch in Definition 5.6, together with the hyperbolicity estimates from Proposition 5.7. Start by choosing the constants as in the assumption of Theorem F.

First we prove that $V^s := \bigcap_{n\geq 0} f^{-n}(\mathcal{N}_{x_n}^{(\ell_n)})$ is a $C^{1+\text{Hölder}}$ full length local $(\widehat{Q}e^{2\epsilon\ell_0}, \lambda/3)$-stable curve. Given $n \in \mathbb{N}$, let $\bar{\ell}^{n,+} := (\ell_0, \dots, \ell_n)$ and $\bar{x}^{n,+} := (x_0, \dots, x_n)$, so that $\bar{x}^{n,+}$ is an $(\bar{\ell}^{n,+}, \delta, \lambda)$-pseudo-orbit, which by Theorem E determines an $\bar{\ell}^{n,+}$-regular branch. Let $\mathcal{B}_{\bar{x}}^0(n) \subset \mathcal{B}_{x_0}^{(\ell_0)}$ be the corresponding stable strip.

The boundary curves of the strips $\{\mathcal{B}_{\bar{x}}^0(n)\}_{n\in\mathbb{N}}$ can be represented as the graphs of a uniformly Lipschitz family of functions $\operatorname{span} e_2 \to \operatorname{span} e_1$ (recall Figure 5.1 and Definitions 5.1–5.3). This family contains a sequence that converges uniformly to a Lipschitz functions whose graph is a continuous curve $\tilde{V}^s \subset \bigcap_{n\geq 0} \mathcal{B}_{\bar{x}}^0(n) \subset \mathcal{B}_x$ with endpoints on $[-b(x), b(x)] \times \{-b(x), b(x)\} \subset \mathcal{B}_x$. Moreover, the expansion estimates in Definition 5.6 imply that any full length unstable admissible curve in $\mathcal{B}_x$ intersects $\mathcal{B}_{\bar{x}}^0(n)$ in a curve of length $\leq 2b_0 e^{-\lambda n/3}$, and thus intersects $\bigcap_{n\geq 0} \mathcal{B}_{\bar{x}}^0(n)$ in a single point. It follows that $\tilde{V}^s = \bigcap_{n\geq 0} \mathcal{B}_{\bar{x}}^0(n)$, and thus $V^s = \bigcap_{n\geq 0} f^{-n}(\mathcal{N}_{x_n}^{(\ell_n)}) = \Psi_x(\bigcap_{n\geq 0} \mathcal{B}_{\bar{x}}^0(n))$ is a continuous curve satisfying the geometric conditions in Definition 1.10; moreover, V^s is full length in $\mathcal{N}_x$. To prove that it is a local $(\widehat{Q}e^{2\epsilon\ell_0}, \lambda/3)$-stable curve, we need to prove that it is C^1 (in fact we will prove that it is $C^{1+\text{Hölder}}$) and that for every $y, z \in V^s$ and $j \geq 0$ we have

$$d(f^j(y), f^j(z)) \leq Ce^{-\lambda j/3} d(y, z) \quad \text{for } C = \widehat{Q}e^{2\epsilon\ell_0}. \tag{9.1}$$

We start by proving (9.1). Given $n \geq 0$, let W_n be an arbitrary full length stable admissible curve in the unstable strip $f^n(\Psi_{x_n}(\mathcal{B}_{\bar{x}}^0(n))$, and let $V_n = f^{-n}(W_n)$. Then for each $y, z \in V^s$ there are $y_n, z_n \in V_n$ such that $y_n \to y$

and $z_n \to z$. Moreover, (5.17) from Proposition 5.7 gives $d(f^j(y_n), f^j(z_n)) \leq Ce^{-\lambda j/3} d(y_n, z_n)$ for all $0 \leq j \leq n$. Thus for every $j \geq 0$ we can send $n \to \infty$ and deduce that (9.1) holds.

It remains to show that V^s is $C^{1+\text{Hölder}}$. To this end, let $x \in V^s$ be arbitrary and choose a unit tangent vector $v_n \in (Df_x^n)^{-1} K^u_{x_n, f^n(x)} \subset T_x M$, where the cones are as defined in (5.11). Then (5.17) from Proposition 5.7 gives

$$\|Df_x^j v_n\| \leq Ce^{-\lambda j/3} \tag{9.2}$$

for all $0 \leq j \leq n$. By compactness of the unit tangent space there is a subsequence $v_{n_k} \to e_x^s \in T_x M$. For every $j \in \mathbb{N}$, (9.2) holds for all v_{n_k} with $n_k \geq j$, and thus it holds for e_x^s as well:

$$\|Df_x^j e_x^s\| \leq Ce^{-\lambda j/3} \text{ for all } j \geq 0. \tag{9.3}$$

We will show that $x \mapsto e_x^s$ is Hölder continuous and that e_x^s is tangent to V^s at x; this will complete the proof that V^s is a local $(C, \lambda/3)$-stable curve.

Let e_x^u be an arbitrary unit tangent vector in $K^u_{x_0,x}$. From (5.16) in Proposition 5.7, we have

$$\|Df_x^j e_x^u\| \geq \widehat{Q}^{-1} e^{-2\epsilon \ell_j} e^{\lambda j/3}$$

for all $j \geq 0$. Observe that

$$-2\epsilon \ell_j + \frac{\lambda}{3} j \geq -2\epsilon \ell_0 - 2\epsilon j + \frac{\lambda}{3} j \geq -2\epsilon \ell_0 + \frac{\lambda}{4} j, \tag{9.4}$$

where the last inequality uses the fact that $\epsilon < \frac{\lambda}{3} - \frac{\lambda}{4} = \frac{\lambda}{12}$, which follows from (1.7). We conclude that

$$\|Df_x^j e_x^u\| \geq \widehat{Q}^{-1} e^{-2\epsilon \ell_0} e^{\lambda j/4} = C^{-1} e^{\lambda j/4} \text{ for all } j \geq 0. \tag{9.5}$$

Using (9.3) and (9.5), we can apply Proposition 8.4 and conclude that $x \mapsto e_x^s$ is Hölder continuous.

It remains to show that e_x^s is tangent to V^s at x. Equivalently, we must show that for any $\omega > 0$, the cone

$$K_\omega := \{0\} \cup \{a e_x^u + b e_x^s : |a| < \omega |b|\} \subset T_x M$$

has the following property: there is a neighbourhood $U \subset T_x M$ containing the origin such that $(\exp_x|_U)^{-1}(V^s)$ lies in K_ω. Once we verify this, we can conclude that V^s is differentiable at x and that e_x^s is in its tangent space. To find the neighbourhood U, we first observe that given $\omega > 0$, every $v \in T_x M \setminus K_\omega$ has the form $v = a e_x^u + b e_x^s$ for some $|a| \geq \omega |b|$ with $|a| > 0$. Thus for all $j \geq 0$ we have

$$\|Df_x^j v\| \geq |a| \|Df_x^j e_x^u\| - |b| \|Df_x^j e_x^s\| \geq |a| (C^{-1} e^{\lambda j/4} - \omega^{-1} C e^{-\lambda j/3}).$$

Moreover, $\|v\| \leq |a| + |b| \leq |a|(1 + \omega^{-1})$, and thus

$$\frac{\|Df_x^j v\|}{\|v\|} \geq \frac{C^{-1} e^{\lambda j/4} - \omega^{-1} C e^{-\lambda j/3}}{1 + \omega^{-1}}. \tag{9.6}$$

Now we fix a choice of j sufficiently large that the right-hand side of (9.6) is at least $3Ce^{-\lambda j/3}$, and let $U \subset T_xM$ be a sufficiently small neighbourhood of 0 that for all $v \in U$ we have

$$\frac{d(f^j(\exp_x v), f^j x)}{d(\exp_x v, x)} \geq \frac{1}{2}\frac{\|D_x f^j v\|}{\|v\|}.$$

Combining this with (9.6) and our choice of j, we see that for all $v \in U \setminus K_\omega$ we have

$$\frac{d(f^j(\exp_x v), f^j x)}{d(\exp_x v, x)} \geq \frac{3}{2}Ce^{-\lambda j/3},$$

and comparing this to (9.1) we conclude that $\exp_x v$ does not lie on the curve V^s. It follows that $(\exp_x|_U)^{-1}(V^s) \subset K_\omega$, and since $\omega > 0$ was arbitrary, this shows that V^s is differentiable at x and that e^s_x is in its tangent space. This completes the proof of the first part of Theorem F. The second part of the theorem follows by an identical argument.

Remark 9.1. The same proof given above can also be applied to a family of (C,κ)-hyperbolic branches as in Proposition 4.10; it suffices to replace $\lambda/3$ and $\lambda/4$ in the above arguments with κ, and to replace the cone families from (5.11) with the cone families associated to the hyperbolic branches. Thus we have also proved Proposition 4.10.

It remains to prove the third part of Theorem F. Existence and uniqueness of the shadowing point follows immediately from the first two parts: indeed, a bi-infinite pseudo-orbit determines a local stable curve $V^s = \bigcap_{n\geq 0} f^{-n}(\mathcal{N}^{(\ell_n)}_{x_n})$ and a local stable curve $V^u = \bigcap_{n\leq 0} f^{-n}(\mathcal{N}^{(\ell_n)}_{x_n})$, and thus $\bigcap_{n\in\mathbb{Z}} f^{-n}(\mathcal{N}^{(\ell_n)}_{x_n}) = V^s \cap V^u$. Since the curves determined in the first two parts are both full length, then have a unique point of intersection, which is the unique shadowing point.

Now we establish regularity of y. Fix $\delta, \lambda > 0$ and let Λ' be the collection of unique shadowing points for $(\bar\ell, \delta, \lambda)$-pseudo-orbits in Λ, where $\bar\ell \in \mathbb{N}^{\mathbb{Z}}$ can be arbitrary. Observe that Λ' is invariant since as remarked after the statement of Theorem F, if y is the unique shadowing point for $\bar x$, then $f(y)$ is the unique shadowing point for $\sigma\bar x$. Given $y \in \Lambda'$ we have local stable and unstable curves $V^{s/u}_y$ as in the first two parts of the theorem. Moreover, these curves satisfy $f(V^s_y) \subset V^s_{f(y)}$ (since shadowing the orbit for $n \geq 0$ implies shadowing for $n \geq 1$) and $f^{-1}(V^u_y) \subset V^u_{f^{-1}(y)}$. Thus the subspaces defined by $E^{s/u}_y = T_y V^{s/u}_z$ produce a measurable Df-invariant splitting on Λ'.

It remains to prove that this invariant splitting satisfies (**H1**)–(**H3**). We start by letting $e^{s/u}_y$ be unit vectors in $E^{s/u}_y$ and observing that (9.3) and (9.5) imply that writing $C(y) = \widehat{Q}e^{2\epsilon\ell_0}$, we have

$$\|Df^n_y e^s_y\| \leq C(y)e^{-\lambda n/3} \text{ and } \|Df^n_y e^u_y\| \geq C(y)^{-1}e^{\lambda n/4}$$

for all $n \geq 0$, where we observe that (9.5) applies to a broader class of unit tangent vectors (anything in the cone $K^u_{x_0,y}$), and in particular applies to our choice of e^u_y. This establishes two of the four inequalities in (**H3**); the other two follow from the analogues of (9.3) and (9.5) for the backwards pseudo-orbit and corresponding unstable curve.

Remark 9.2. The fact that (9.5) contains $\lambda/4$ instead of $\lambda/3$ is due to the fact that we must use (9.4) to compare the level of regularity at the two ends of the orbit segment. One consequence of this appears later in Proposition 11.3 and its proof in §11.2, when we build a nice rectangle for which $\lambda/3$ appears in the hyperbolic branch property, but $\lambda/4$ appears in the regularity. Roughly speaking one may say that the hyperbolic branch property controls contraction/expansion from the endpoints to somewhere in the middle of the corresponding orbit segment, but regularity requires us also to control it when going between two points in the middle of the segment, and for this we must weaken $\lambda/3$ to $\lambda/4$.

To prove (**H1**) for C, observe that $f(y)$ is the shadowing orbit for $\sigma\bar{x}$ with regularity sequence $\bar{\ell}$, so that $C(f(y)) = \widehat{Q}e^{2\epsilon\ell_1}$, and thus

$$e^{-2\epsilon} \leq C(f(y))/C(y) \leq e^{2\epsilon}. \tag{9.7}$$

Now we need to estimate $\measuredangle(E^s_y, E^u_y)$. First observe that $\measuredangle(E^s_y, E^u_y) \geq \|e^s_y - e^u_y\|$ since the angle represents the length of the arc of the unit circle joining the endpoints of e^s_y and e^u_y, while the right-hand side is the length of the straight line joining them. Let $\underline{v}^{u/s} := D_y\Psi^{-1}_{x_0}e^{u/s}_y$ and observe that $\|\underline{v}^{u/s}\| \geq \frac{1}{2}$ by the first estimate in (5.15). Moreover, $\underline{v}^{u/s} \in K^{s/u}$, so the endpoint of $\underline{v}^u$ lies in the region of $\mathbb{R}^2$ given by

$$\{(x,y) \in \mathbb{R}^2 : |y| \leq \omega|x| \text{ and } x^2 + y^2 \geq 1/4\},$$

while the endpoint of $\underline{v}^u$ lies in the region

$$\{(x,y) \in \mathbb{R}^2 : |x| \leq \omega|y| \text{ and } x^2 + y^2 \geq 1/4\},$$

Thus, by (5.9), $\|\underline{v}^u - \underline{v}^s\| \geq 1/2$,[15] and we conclude that

$$\frac{1}{2} \leq \|\underline{v}^u - \underline{v}^s\| \leq \|D_y\Psi^{-1}_{x_0}\|\|e^u_y - e^s_y\| \leq 4Q_0\widehat{Q}e^{2\epsilon\ell_0}\measuredangle(E^s_y, E^u_y)$$

using the second inequality in (5.15). Thus

$$\measuredangle(E^s_y, E^u_y) \geq \widehat{Q}^{-1}e^{-2\epsilon\ell_0},$$

so putting $K(y) := \widehat{Q}^{-1}e^{-2\epsilon\ell_0}$ establishes (**H2**), and (**H1**) for K follows just as it did for C. We conclude that the set Λ' of shadowing points is $(\lambda/4, 2\epsilon)$-hyperbolic. Moreover, to find which Λ'_ℓ contains y, we write

$$C(y) = \widehat{Q}e^{2\epsilon\ell_0} \leq e^{2\epsilon\ell}$$

and find that this holds as soon as $\ell \geq \ell_0 + \frac{1}{2\epsilon}\log\widehat{Q}$. A similar computation with $K(y)$ shows that y is $(\lambda/4, 2\epsilon, \ell_0 + \lceil\frac{1}{2\epsilon}\log\widehat{Q}\rceil)$-regular.

[15] An elementary computation shows that the optimal lower bound is $(1-\omega)/\sqrt{2(1+\omega^2)}$.

Part III. Nice Rectangles and Young Towers

In this third and final part of the paper, we apply the general results stated in Theorem E and F, and proved in Part II above, to our particular setting in order to prove Theorems B, C and D. As mentioned above, Theorem A follows directly from Theorems B and C, and therefore this completes the proofs of all our results. The sections are organized as follows: in §10 we prove Theorem D, which is essentially a reformulation of Theorem E in the setting of almost returns to nice domains, in §11 we show that Theorem D implies Theorem B, and in §12 we prove Theorem C.

10. Hyperbolic branches in nice domains: Proof of Theorem D

The proof of Theorem D consists of two parts. First we show that every almost return gives rise to a pseudo-orbit and thus, by Theorem E, to a regular branch, which satisfies the hyperbolicity estimates given in Proposition 5.7. Then we show that this regular branch can be "restricted" to give a hyperbolic branch in the nice domain Γ_{pq}. This second part of the proof does not explicitly require Theorem E, it only uses the existence of a regular branch, but does use in an essential way the fact that Γ_{pq} is a nice domain.

To begin, let $C_\ell \geq 1$ be the constant given in Theorem 1.12, c_2 as given in (1.5), and $\delta > 0$ as in Theorem E. Then we let

$$r := \frac{\delta e^{-\lambda(\ell+1)} e^{-c_2}}{2C_\ell} \leq \delta. \tag{10.1}$$

For generality we state the following Lemma for almost returns in a slightly more general setting than that of Theorem D, without any explicit references to rectangles or nice domains.

Lemma 10.1. *If $x, y \in \Lambda_\ell$ and $k \geq 1$ are such that $f^k(V_x^s) \cap V_y^u \neq \emptyset$, and $z \in f^k(V_x^s) \cap V_y^u$ satisfies $d(z, y) < r$ and $d(f^{-k}(z), x) < r$, then the sequence $\bar{x} = (x_0, \dots, x_k)$ given by $x_j = f^j(x)$ for $0 \leq j \leq k/2$ and $x_j = f^{j-k}(y)$ for $k/2 < j \leq k$ is an $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit for $\ell_j = \min(\ell + j, \ell + k - j)$.*

Proof. Write $i = \lfloor k/2 \rfloor$. By assumption $z \in V_y^u$ and $f^{-k}(z) \in V_x^s$ and therefore by the assumptions of the Lemma and Theorem 1.12 we have

$$d(f^i(x), f^{i-k}(z)) \leq C_\ell e^{-\lambda i} d(x, f^{-k}(z)) \leq C_\ell e^{-\lambda i} r \leq \frac{1}{2} \delta e^{-\lambda(\ell+i)},$$

and

$$d(f^{i-k}(z)), f^{i-k}(y)) \leq C_\ell e^{-\lambda(k-i)} d(z, y) \leq C_\ell e^{-\lambda(k-i)} r \leq \frac{1}{2} \delta e^{-\lambda(\ell+i)},$$

and thus $d(f(x_i), x_{i+1}) = d(f^i(x), f^{i-k}(y)) \leq \delta e^{-\lambda(\ell+i)}$. Since $f(x_j) = x_{j+1}$ for all $j \neq i$, this completes the proof. □

Consider now the setting of Theorem D: suppose Γ is a nice regular set with $\operatorname{diam}(\Gamma_{pq}) < r$ and suppose $x \in \Gamma$ has an almost return to Γ at time $k \in T\mathbb{N}$. Then the assumptions of Lemma 10.1 are satisfied and there is an

$(\bar{\ell}, \delta, \lambda)$-pseudo-orbit $\bar{x} = (x_0, \ldots, x_k)$ as in the lemma starting and ending inside Γ_{pq}. Moreover, notice that

$$d(f(p), x_1) = d(f(p), f(x_0)) \le e^{c_2} d(p, x_0) \le e^{c_2} \operatorname{diam}(\Gamma_{pq}) \le e^{c_2} r < \delta e^{-\lambda(\ell+1)},$$

where the first inequality uses the general fact from (1.5) that $d(f(x), f(y)) \le e^{c_2} d(x, y)$ for all $x, y \in M$, the second inequality is immediate since $p, x_0 \in \Gamma_{pq}$, the third uses our assumption that $\operatorname{diam}(\Gamma_{pq}) < r$, and the fourth uses (10.1).

By a similar calculation, $d(f(x_{k-1}), p) < \delta e^{-\lambda \ell}$, and therefore the sequence

$$\bar{p} := (p, x_1 \ldots, x_{k-1}, p)$$

is also an $(\bar{\ell}, \delta, \lambda)$-pseudo-orbit. Considering the Lyapunov chart $\Psi_p \colon \mathcal{B}_p^{(\ell)} \to \mathcal{N}_p^{(\ell)}$, by Theorem E there is an ℓ-regular branch from $\mathcal{B}_p^{(\ell)}$ to itself associated to this pseudo-orbit, and we have the corresponding maps

$$f_{\bar{p}}^{0,k} \colon \mathcal{B}_{\bar{p}}^0 \to \mathcal{B}_{\bar{p}}^k \quad \text{and} \quad f^k \colon \mathcal{N}_{\bar{p}}^0 \to \mathcal{N}_{\bar{p}}^k \tag{10.2}$$

at the level of Lyapunov charts and of the manifold respectively, recall (5.13). For this branch Proposition 5.7 gives the hyperbolicity estimates required in Definition 4.8 for a $(\widehat{Q} e^{2\epsilon \ell}, \lambda/3)$-hyperbolic branch. Moreover, concatenating any finite sequence of such branches gives a new ℓ-regular branch that is associated to the concatenated pseudo-orbit, and thus has the same hyperbolicity estimates given by Proposition 5.7. Thus the collection of such branches satisfies the concatenation property.

Remark 10.2. We emphasize that these are not yet the hyperbolic branches we require for Γ_{pq} as in Definition 4.8. Indeed, these branches are constructed on the scale of the Lyapunov chart which *a priori* may be significantly bigger than the scale of the nice domain Γ_{pq}. The strips $\mathcal{N}_{\bar{p}}^0, \mathcal{N}_{\bar{p}}^k$ intersect Γ_{pq} but may extend across the boundary of Γ_{pq}. We therefore need to "restrict" these branches to Γ_{pq} and produce Γ_{pq}-strips $\widehat{C}^s \subset \mathcal{N}_{\bar{p}}^0$ and $\widehat{C}^u \subset \mathcal{N}_{\bar{p}}^k$ such that f^k maps $\widehat{C}^s$ onto $\widehat{C}^u$. Since these are subsets of the larger strips $\mathcal{N}_{\bar{p}}^0, \mathcal{N}_{\bar{p}}^k$ and the cones $K_{p,y}^{s/u} \subset T_y M$ defined in (5.11) give conefields over Γ_{pq} that are adapted to the set Γ, the restricted strips will automatically inherit the hyperbolicity and concatenation properties.

The remaining part of the argument is essentially topological, and this is where the niceness assumption plays a crucial role. Indeed, the crucial consequence of niceness is formalized in the following statement.

Lemma 10.3. *Let Γ be a nice regular set and suppose that some $x \in \Gamma$ has an almost return to Γ at a time $k \in \mathbb{N}T$. Then $f^k(W_x^s) \subset \Gamma_{pq}$.*

Proof. Suppose by contradiction that the conclusion does not hold. Then $f^k(W_x^s)$ must intersect one of $W_{p/q}^u$, but this implies that the image under f^{-k} of this intersection point lies in the interior of Γ_{pq}, which is forbidden by niceness. $\square$

As shown in Figure 10.1, let

$$\gamma_p^u := W_p^u \cap \mathcal{N}_{\bar{p}}^0 \quad \text{and} \quad \gamma_q^u := W_q^u \cap \mathcal{N}_{\bar{p}}^0.$$

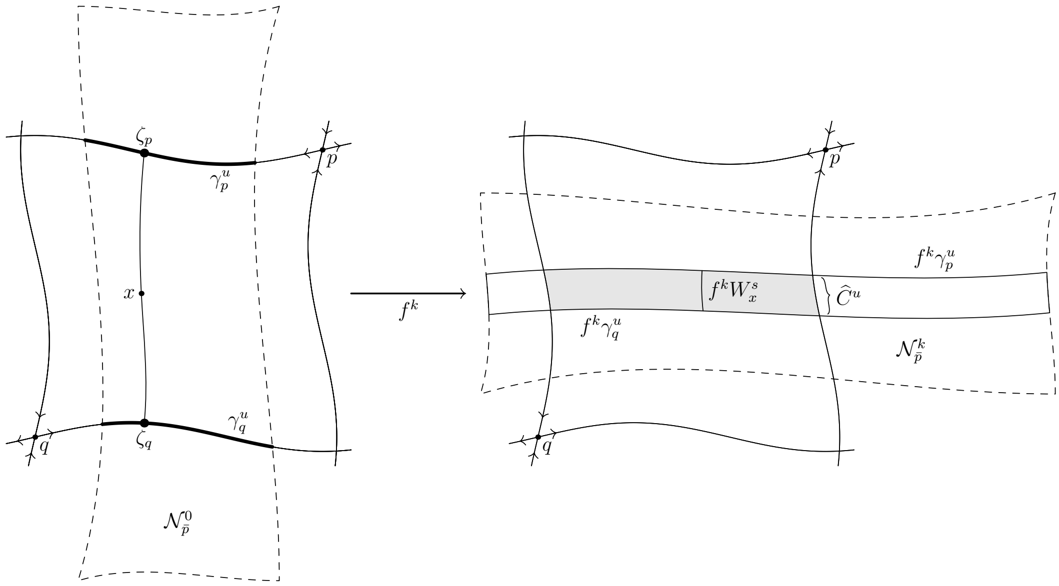

FIGURE 10.1. Proving Theorem D.

Lemma 10.4. *The curves* $f^k(\gamma_p^u)$ *and* $f^k(\gamma_q^u)$ *are full length unstable admissible curves in* Γ_{pq}.

Proof. We prove the statement for γ_p^u, the same argument applies to γ_q^u. Observe that each endpoint of γ_p^u lies on either W_p^s, W_q^s, or the stable boundary of $\mathcal{N}_{\bar p}^0$. Since p and q are fixed by f^k, we have $f^k(W_p^s) \subset W_p^s$ and $f^k(W_q^s) \subset W_q^s$; it follows that each endpoint of $f^k(\gamma_p^u)$ lies on either W_p^s, W_q^s, or the stable boundary of $f^k(\mathcal{N}_{\bar p}^0) = \mathcal{N}_{\bar p}^k$. Since $f^k(\gamma_p^u)$ must intersect Γ_{pq} by Lemma 10.3, these restrictions on its endpoints guarantee that it is a full length unstable admissible curve in Γ_{pq}. $\square$

Proof of Theorem D. By Lemma 10.4, $f^k(\gamma_p^u), f^k(\gamma_q^u)$ are full length unstable admissible curves in Γ_{pq} and therefore they define an unstable strip $\widehat{C}^u$ in Γ_{pq}, whose preimage $\widehat{C}^s := f^{-k}(\widehat{C}^u)$ is a stable strip in Γ_{pq}, thus yielding the desired hyperbolic branch. The required hyperbolicity estimates are inherited from the regular branch of which this hyperbolic branch is a subset. $\square$

11. Building a tower out of hyperbolic branches: Proof of Theorem B

In this section we prove Theorem B. In §11.1 we introduce some definitions and notation and reduce the proof of Theorem B to three Propositions 11.3, 11.5, and 11.6. We then prove each Proposition in its own subsection.

11.1. **Saturation and Young Towers.** Before formulating the propositions, we need to establish some notation and to introduce the notion of *saturation*.

Let Γ_{pq} be a nice domain and suppose that $A \subset \Gamma_{pq}$ is such that every point $x \in A$ has full length stable and unstable curves $W_x^{s/u}$. Suppose moreover that A has the (C, κ)-hyperbolic branch property for some $C, \kappa > 0$ (see Definition

4.14). Let $\mathfrak{C}(A)$ denote the set of hyperbolic branches associated to almost returns to A, and let $\mathfrak{C}_0(A) \subset \mathfrak{C}(A)$ be the subset consisting of those branches associated to (true) returns. Let $\mathfrak{C}^*(A)$ denote the set of branches obtained by concatenating finitely many members of $\mathfrak{C}(A)$. Thus we have

$$\mathfrak{C}_0(A) \subset \mathfrak{C}(A) \subset \mathfrak{C}^*(A).$$

Remark 11.1. Both inclusions can be proper. For the first one, A can have almost returns without having any true returns, for example, if the set A consists of one non-periodic point which returns to Γ_{pq}, then there is no return to A but almost returns may exist. For the second, $\mathfrak{C}^*(A)$ may even contain branches $f^k\colon \widehat{C}^s \to \widehat{C}^u$ such that $\widehat{C}^s$ and $\widehat{C}^u$ are disjoint from A. This can occur if two branches $f^i\colon \widehat{C}^s_1 \to \widehat{C}^u_1$ and $f^j\colon \widehat{C}^s_2 \to \widehat{C}^u_2$ generated by almost returns have a concatenated branch $f^{i+j}\colon \widehat{C}^s \to \widehat{C}^u$ (recall Definition 4.11 and Figure 4.2) with the property that the part of A in $\widehat{C}^s_1$ lies entirely outside of $\widehat{C}^s$, and similarly for $\widehat{C}^u_2$.

Since f is a diffeomorphism of a compact manifold the derivative of f is bounded and so for each $i \geq 1$ there can be at most a finite number of hyperbolic branches of order i. Thus we can index $\mathfrak{C}_0(A)$ by

$$I_0(A) := \{ij : i \in T\mathbb{N}, j \in \{1, ..., m_i\}\},$$

where i gives the order (return time) of the hyperbolic branch and j indexes the m_i hyperbolic branches with order i. Therefore we obtain

$$\mathfrak{C}_0(A) = \{f^i\colon \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}\}_{ij \in I_0(A)}. \tag{11.1}$$

We consider the particular case when A is a rectangle.

Definition 11.2 (Saturated Rectangle)**.** Let $\Gamma \subset \Gamma_{pq}$ be a nice rectangle with the hyperbolic branch property[16] and let $I_0(\Gamma)$ and $\mathfrak{C}_0(\Gamma)$ be as above (see (11.1)). The rectangle Γ is called *saturated* if for all $ij \in I_0(\Gamma)$, we have

$$C^s_{ij} := f^{-i}(\widehat{C}^u_{ij} \cap C^s) \subset C^s \quad \text{and} \quad C^u_{ij} := f^i(\widehat{C}^s_{ij} \cap C^u) \subset C^u, \tag{11.2}$$

where $C^{s/u} = \bigcup_{x\in\Gamma} W^{s/u}_x$ as in (2.2).

See §11.5 for a general discussion of the saturation property and its connection to the first return property of the Young tower that is eventually constructed, as well as to multiplicity of the associated countable-state Markov coding. That section also includes an example of a non-saturated rectangle.

Given $\chi > \lambda > 0$, $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and $\ell \in \mathbb{N}$, let $r > 0$ be given by Theorem D, so that every $(\chi, \epsilon, \ell, r)$-nice regular set has the $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-hyperbolic branch property.

Proposition 11.3. *With $\chi, \lambda, \epsilon, \ell, r$ as above, let A be a $(\chi, \epsilon, \ell, r)$-nice recurrent set. Then there exists a nice $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-rectangle $\Gamma \subset \Gamma_{pq}$ such that the following are true.*

[16]In fact one does not need the full strength of the hyperbolic branch property to make this definition; it suffices to have a hyperbolic branch associated to each (true) return.

(1) $A \subset \Gamma \subset \Gamma_{pq}$.
(2) Γ *is* T*-recurrent.*
(3) Γ *satisfies the* $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$*-hyperbolic branch property.*
(4) $\mathfrak{C}_0(\Gamma) = \mathfrak{C}(\Gamma) = \mathfrak{C}^*(\Gamma) = \mathfrak{C}^*(A)$.
(5) Γ *is saturated.*
(6) Γ *is* $(\lambda/4, 2\epsilon, \ell + \ell')$*-regular, where* $\ell' = \lceil \frac{1}{2\epsilon} \log \widehat{Q} \rceil$.

Remark 11.4. See Remark 9.2 for a discussion of why $\lambda/3$ appears in the hyperbolic branch property (3) and $\lambda/4$ appears in the regularity property (6). This regularity property is used in the proof of Theorem B(2) to establish a Hölder continuity estimate ((11.3) in Proposition 11.6) leading to bounded distortion (condition **(Y2)** in the definition of Young tower), but the construction in Proposition 11.5 of the topological Young tower for Theorem B(1) only uses the hyperbolic branch property (3).

At first glance it may appear redundant to state both conclusions (3) and (6) in Proposition 11.3, since Theorem D guarantees that every nice regular set satisfies the hyperbolic branch property, so one might reasonably expect to deduce (3) as a consequence of (6). The problem is that A is (χ, ϵ, ℓ)-regular while Γ is only $(\lambda/4, 2\epsilon, \ell + \ell')$-regular (which is weaker), and so in order to apply Theorem D and deduce that Γ satisfies the (C, κ)-hyperbolic branch property for some $\kappa \in (0, \lambda/4)$, we would need to make more careful choices in our original setup, choosing $0 < \epsilon < \frac{1}{2}\epsilon_1(f, \lambda/4, \kappa)$ and then $r = r(\lambda/4, \kappa, 2\epsilon)$. This can be done but would create a more complicated set of conditions for ϵ and r in our main theorems, which we prefer to avoid; after deducing conclusion (6) from Theorem F (see the paragraph following Lemma 11.9), we deduce conclusion (3) by finding an almost return to A itself, to which Theorem D can be applied, see Lemma 11.13 and the sentence following it.

Proposition 11.5. *Let* $\Gamma \subset \Gamma_{pq}$ *be a* T*-recurrent nice* (C_0, κ)*-rectangle satisfying the* (C_0, κ)*-hyperbolic branch property, and suppose that* Γ *is saturated. Then* Γ *supports a First* T*-return Topological Young Tower.*

Replacing T-recurrence with Lebesgue-strong T-recurrence allows us to upgrade the topological Young Tower to a fully fledged Young Tower satisfying the required distortion estimates and having integrable return times.

Proposition 11.6. *Let* $\Gamma \subset \Gamma_{pq}$ *be a Lebesgue-strongly* T*-recurrent nice* (C_0, κ)*-rectangle satisfying the* (C_0, κ)*-hyperbolic branch property, and suppose that* $C_0 e^{-\kappa T} < 1$. *Suppose moreover that* Γ *is saturated and that there are* $C_1, \beta_1, \gamma_1 > 0$ *such that* $\gamma_1 < \beta_1 \kappa$ *and for all* $a \in \mathbb{Z}$ *and* $x, y \in f^a(\Gamma)$*, we have*

$$d(E^u_x, E^u_y) \le C_1 e^{\gamma_1 |a|} d(x, y)^{\beta_1}, \tag{11.3}$$

Then Γ *supports a First* T*-return Young Tower with integrable return times.*

To prove Theorem B, we will apply Propositions 11.5 and 11.6 to the rectangle constructed in Proposition 11.3, with $C_0 = \widehat{Q}e^{2\epsilon\ell}$ and $\kappa = \lambda/3$. Propositions 11.3, 11.5, and 11.6 will be proved in §11.2, §11.3, and §11.4, respectively.

Remark 11.7. The only hyperbolic branches that are required in the proof of Propositions 11.5 and 11.6 are those associated to true returns (of Γ); so these results remain true if the hyperbolic branch property from Definition 4.13 is weakened to only require that the set of true returns produces a collection of hyperbolic branches with the concatenation property, rather than requiring such a collection for the (larger) set of almost returns. Branches associated to almost returns (of A) play a crucial role in the proof of Proposition 11.3 for establishing the saturation property; roughly speaking, once this property is obtained it is enough to consider true returns.

Proof of Theorem B. As stated before Proposition 11.3, we choose r depending on $\chi, \lambda, \epsilon, \ell$ to satisfy the conditions of Theorem D, and thus given a $(\chi, \epsilon, \ell, r)$-nice recurrent set A as in the assumptions of Theorem B, Proposition 11.3 yields a nice rectangle $\Gamma \subset \Gamma_{pq}$ that contains A, is T-recurrent, satisfies the $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-hyperbolic branch property, and is saturated. Thus Proposition 11.5 applies to Γ, and we conclude that Γ supports a First T-return Topological Young Tower, which proves the first part of Theorem B. For the second part of Theorem B, we will apply Proposition 11.6 to Γ, which requires us to verify Lebesgue-strong recurrence, the (C_0, κ)-hyperbolic branch property with $C_0 e^{-\kappa T} < 1$, and (11.3) with $\gamma_1 < \beta_1 \kappa$.

To this end, first observe that since A is Lebesgue-strongly T-recurrent, so is any set containing A, including Γ. Moreover, with $C_0 = \widehat{Q}e^{2\epsilon\ell}$ and $\kappa = \lambda/3$, we have $C_0 e^{-\kappa T} = \widehat{Q}e^{2\epsilon\ell}e^{-\lambda T/3} < 1$ by our choice of T in Definition 1.16. Finally, since Γ is $(\lambda/4, 2\epsilon, \ell + \ell')$-regular by Proposition 11.3(6), we see that $f^a(\Gamma)$ is $(\lambda/4, 2\epsilon, \ell + \ell' + |a|)$-regular for every $a \in \mathbb{Z}$. We can deduce (11.3) from Proposition 8.4, and specifically (8.6) (with s replaced by u), where we replace χ with $\lambda/4$. More precisely, regularity guarantees that for all $x, y \in f^a(\Gamma)$ we have (8.7) and (8.8) with $C = e^{2\epsilon(\ell+\ell'+|a|)}$ and χ replaced by $\lambda/4$, and moreover $\measuredangle(E^s_z, E^u_z) \geq C^{-1}$ for $z = x, y$, so that Proposition 8.4 gives $Q_6' > 0$ depending only on M, $\|Df^{\pm 1}\|$, α, $|Df^{\pm 1}|_\alpha$, and $\lambda/4$, such that

$$d(E^u_x, E^u_y) \leq Q_6'(e^{2\epsilon(\ell+\ell'+|a|)})^{3\cdot 2\gamma'} d(x,y)^{\beta'} = Q_6' e^{12\epsilon\gamma'(\ell+\ell')} e^{12\epsilon\gamma'|a|} d(x,y)^{\beta'},$$

where β' and γ' are obtained by replacing χ by $\lambda/4$ in (1.6) to get

$$\gamma' = \frac{\frac{\lambda}{4} - c_1}{2\frac{\lambda}{4}} = \frac{\lambda - 4c_1}{2\lambda} \quad \text{and} \quad \beta' = \frac{2\frac{\lambda}{4}}{c_3 + \frac{\lambda}{4}}\alpha = \frac{2\lambda\alpha}{4c_3 + \lambda}. \tag{11.4}$$

It follows that (11.3) holds with $C_1 = Q_6' e^{12\epsilon\gamma'(\ell+\ell')}$, $\gamma_1 = 12\epsilon\gamma'$, and $\beta_1 = \beta'$. Thus the condition that $\gamma_1 < \beta_1\kappa$ is equivalent to $12\epsilon\gamma' < \beta'\lambda/3$. Using (11.4), this is equivalent to

$$6\epsilon\frac{\lambda - 4c_1}{\lambda} < \frac{2\lambda\alpha}{4c_3 + \lambda} \cdot \frac{\lambda}{3},$$

which holds by the fifth inequality in (1.7). This confirms that we can apply Proposition 11.6 to Γ, and completes the proof of Theorem B. $\square$

11.2. **Proof of Proposition 11.3.** We will define Γ as the "maximal invariant set" for the dynamics generated by the hyperbolic branches associated to almost returns to A (via Theorem D). This can be thought of as a generalization of the standard horseshoe where we define a maximal invariant set as consisting of the points which remain in the strips for all forward and backward iterations. The key difference is that in the horseshoe setting we have at most a finite number of branches with pairwise disjoint stable strips and pairwise disjoint unstable strips, all of which have the same return time. In our setting, a point x may belong to infinitely many stable strips with varying return times and therefore we need a more involved construction. With this in mind, we make the following general definition. Let

$$\mathfrak{C} = \{f^i \colon \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}\}_{ij\in I}$$

be a collection of hyperbolic branches indexed by the set I. Note that so far we do not assume that these come from almost returns to a nice regular set; for now we allow an arbitrary collection of hyperbolic branches.

Definition 11.8 (Hyperbolic sequences)**.** A sequence $\mathfrak{h}^+ = \{i_m j_m\}_{m=0}^\infty$ with $i_m j_m \in I$ is a *forward hyperbolic sequence* for $x \in \Gamma_{pq}$ if for all $m \geq 0$ we have

$$f^{i_0+i_1+\cdots+i_{m-1}}(x) \in \widehat{C}^s_{i_m j_m}. \tag{11.5}$$

Similarly, $\mathfrak{h}^- = \{i_m j_m\}_{m=-\infty}^{-1}$ is a *backward hyperbolic sequence* for $x \in \Gamma_{pq}$ if for all $m < 0$ we have

$$f^{-(i_m+\cdots+i_{-1})}(x) \in \widehat{C}^s_{i_m j_m}. \tag{11.6}$$

If $\mathfrak{h}^-$ and $\mathfrak{h}^+$ are backward and forward hyperbolic sequences for x, their concatenation $\mathfrak{h} = \{i_m j_m\}_{m\in\mathbb{Z}}$ is called a *hyperbolic sequence* for x (associated to the collection of hyperbolic branches $\mathfrak{C}$).

A point x may or may not admit a forward or backward hyperbolic sequence and, if it does, these sequences need not be uniquely defined. Consider the sets

$$\begin{aligned} C^+ &:= \{x \in \Gamma_{pq} \mid x \text{ has a forward hyperbolic sequence } \mathfrak{h}^+\}, \\ C^- &:= \{x \in \Gamma_{pq} \mid x \text{ has a backward hyperbolic sequence } \mathfrak{h}^-\}. \end{aligned}$$

Lemma 11.9. *Given an arbitrary collection $\mathfrak{C}$ of (C,κ)-hyperbolic branches, the following set is a nice (C,κ)-rectangle and is T-recurrent:*

$$\Gamma := C^+ \cap C^- = \{x \in \Gamma_{pq} \mid x \textit{ has a hyperbolic sequence } \mathfrak{h}\}. \tag{11.7}$$

Moreover, $C^+ = \bigcup_{x\in\Gamma} W^s_x$ and $C^- = \bigcup_{x\in\Gamma} W^u_x$, so that $C^+ = C^s(\Gamma)$ and $C^- = C^u(\Gamma)$ in the notation of (2.2).

We think of Γ as the "maximal invariant set" of $\mathfrak{C}$.

Proof of Lemma 11.9. To prove that Γ is a rectangle, we use arguments similar to those in the standard horseshoe setting. Let $\mathfrak{h} = \{i_m j_m\}_{m\in\mathbb{Z}}$ denote a hyperbolic sequence for the point x, and $\mathfrak{h}^\pm$ its forward and backward parts. For each $n \geq 0$, the branches $f^{i_m} \colon \widehat{C}^s_{i_m j_m} \to \widehat{C}^u_{i_m j_m}$ with $m = 0, 1, \ldots, n$ can be concatenated as in Definition 4.12 to produce a branch of order $i_0+i_1+\cdots+i_n$; by Proposition 4.10, the intersection (over all $n \geq 0$) of the resulting branches is a local (C, κ)-stable curve that has full length in Γ_{pq}. It can be characterized as

$$W^s_x = \{y \in \Gamma_{pq} : \mathfrak{h}^+ \text{ is a forward hyperbolic sequence for } y\}. \tag{11.8}$$

A completely analogous argument shows that every $x \in \Gamma$ has a full length local (C, κ)-unstable curve

$$W^u_x = \{y \in \Gamma_{pq} : \mathfrak{h}^- \text{ is a backward hyperbolic sequence for } y\}. \tag{11.9}$$

This implies that for any $x, y \in \Gamma$ the intersection $W^s_x \cap W^u_y$ consists of a single point z. Moreover, x has a forward hyperbolic sequence $\mathfrak{h}^+_x$ and y has a backward hyperbolic sequence $\mathfrak{h}^-_y$; writing $\mathfrak{h}$ for the concatenation of these two sequences, it follows from (11.8) and (11.9) that $\mathfrak{h}$ is a hyperbolic sequence for z, so $z \in \Gamma$. This shows that Γ is a rectangle, as claimed.

To see that $C^+ = C^s(\Gamma)$ it suffices to observe that C^+ is a union of curves W^s_x by (11.8), so $C^+ = \bigcup_{x \in C^+} W^s_x \supset \bigcup_{x \in \Gamma} W^s_x$, while given an arbitrary $x \in C^+$ and $y \in \Gamma$ we have $z = [x, y] \in \Gamma$ and $x \in W^s_z$, so $C^+ \subset \bigcup_{z \in \Gamma} W^s_z$. A similar argument gives $C^- = C^u(\Gamma)$.

Finally, the fact that Γ is recurrent follows almost immediately from the definition; if $\mathfrak{h} = \{i_m j_m\}_{m \in \mathbb{Z}}$ is a hyperbolic sequence for $x \in \Gamma$, then $f^{i_0}(x)$ and $f^{-i_{-1}}(x)$ have hyperbolic sequences given by shifting $\mathfrak{h}$ one index in either direction, hence i_0 and $-i_{-1}$ are return times to Γ. □

Now we restrict ourselves to the case when $\mathfrak{C}$ arises from almost returns. Given $\chi > \lambda > 0$, $0 < \epsilon < \epsilon_1(f, \chi, \lambda)$, and $\ell \in \mathbb{N}$, let $r > 0$ be given by Theorem D, so that every $(\chi, \epsilon, \ell, r)$-nice regular set has the $(\widehat{Q} e^{2\epsilon\ell}, \lambda/3)$-hyperbolic branch property. Let A be a $(\chi, \epsilon, \ell, r)$-nice recurrent set and let $\mathfrak{C} = \mathfrak{C}(A)$ be the corresponding collection of hyperbolic branches associated to almost returns via Theorem D. Then Lemma 11.9 shows that (11.7) defines a nice $(\widehat{Q} e^{2\epsilon\ell}, \lambda/3)$-rectangle $\Gamma \subset \Gamma_{pq}$ that is T-recurrent, verifying conclusion (2) of Proposition 11.3. Conclusion (6) on regularity follows from Theorem F. For conclusion (1), we observe that since A is recurrent, given $x \in A$ there are $i_0, i_1, i_2, \dots \in \mathbb{N}$ such that $x_k := f^{i_0 + i_1 + \cdots + i_{k-1}}(x) \in A$ for all $k \geq 0$; then $f^{i_k}(x_k) = x_{k+1}$ is a return to A that produces a corresponding hyperbolic branch in $\mathfrak{C}$, and the sequence of branches obtained this way is a forward hyperbolic sequence for x. A backward hyperbolic sequence is produced similarly, and thus $x \in \Gamma$, verifying (1).

Remark 11.10. Everything in the previous paragraph remains true if we replace $\mathfrak{C}(A)$ with the (potentially smaller) collection $\mathfrak{C}_0(A)$ of branches associated to true returns. However, in the remainder of the proof it will be essential that we define Γ using the collection of branches corresponding to almost returns.

It remains to prove conclusions (3), (4), and (5). In order to prove these conclusions, we need the following result about hyperbolic branches. We remark that this is the one and only place in the paper where we use the assumption from Definition 1.16 that T is even.

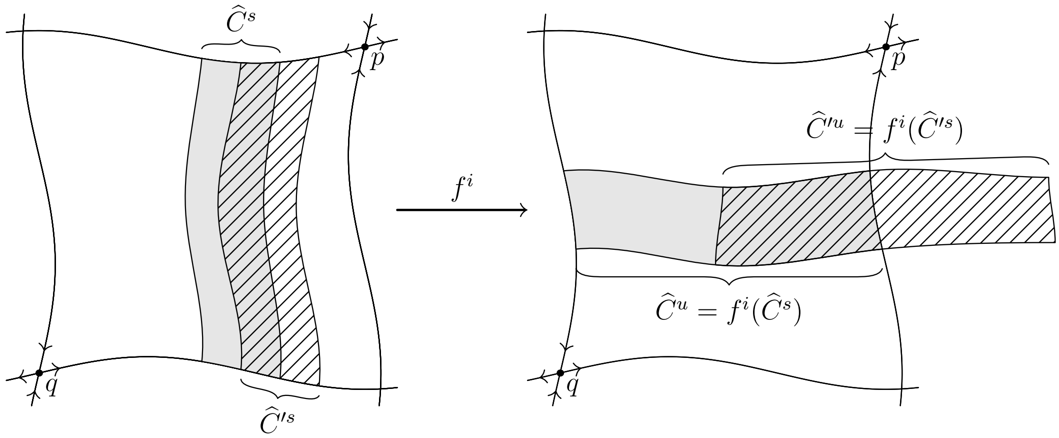

FIGURE 11.1. Two overlapping branches leads to a contradiction (Lemma 11.11).

Lemma 11.11. *Let Γ_{pq} be a nice domain. Then for any hyperbolic branch $f^i \colon \widehat{C}^s \to \widehat{C}^u$ we have $\mathrm{Int}(f^k(\widehat{C}^s)) \cap \partial\Gamma_{pq} = \emptyset$ for all $k = 0, \dots, i$ that are multiples of T. Moreover, if $f^{i'} \colon \widehat{C}'^s \to \widehat{C}'^u$ is any other hyperbolic branch, then the corresponding stable (resp. unstable) strips are either nested or disjoint.*

Proof. The first statement is automatic for $k = 0, i$. Suppose that there exists some $k \in \{1, .., i-1\}$ such that $\mathrm{Int}(f^k(\widehat{C}^s)) \cap \partial\Gamma_{pq} \neq \emptyset$. Then we must have $\mathrm{Int}(f^k(\widehat{C}^s)) \cap (W^s_q \cup W^s_p) \neq \emptyset$ or $\mathrm{Int}(f^k(\widehat{C}^s)) \cap (W^u_q \cup W^u_p) \neq \emptyset$ (or both). In the first case, iterating forward by $i-k$ iterates, this would imply that $\mathrm{Int}(\widehat{C}^u) \cap f^{i-k}(W^s_q \cup W^s_p) \neq \emptyset$, contradicting the niceness property of Γ_{pq}. Similarly, in the second case, iterating backwards by k iterates, this would imply $\mathrm{Int}(\widehat{C}^s) \cap f^{-k}(W^u_q \cup W^u_p) \neq \emptyset$, contradicting niceness.

For the second statement, assume without loss of generality that $i \leq i'$. Suppose by contradiction that the two stable strips $\widehat{C}^s, \widehat{C}'^s \neq \emptyset$ are neither nested nor disjoint (the argument for unstable strips is exactly the same). Recall that the stable boundaries of $\widehat{C}^s, \widehat{C}'^s$ are pieces of the global stable curves of p, q (see Remark 4.9) and therefore cannot intersect unless they agree; thus the intersection $\widehat{C}^s \cap \widehat{C}'^s$ is a single stable admissible curve or a non-empty stable strip.

In the first case, we see that f^i maps this stable admissible curve γ^s to the local stable manifold of either p or q. Due to our choice of t even, $D_p f^T$ and $D_q f^T$ have positive eigenvalues, and it follows that since $i' \geq i$ are multiples of T, there is an open region $U \subset \widehat{C}'^s$ adjacent to γ^s such that $f^{i'}(U) \cap \Gamma_{pq} = \emptyset$, contradicting the assumption that $f^{i'} \colon \widehat{C}'^s \to \widehat{C}'^u$ is a hyperbolic branch.

In the second case, each stable strip has one of the components of its stable boundary inside the interior of the other stable strip, see first figure in Figure 11.1. It follows that $f^i(\widehat{C}'^s)$ contains a piece of the stable boundary of Γ_{pq} in its interior, contradicting the first statement proved above. $\square$

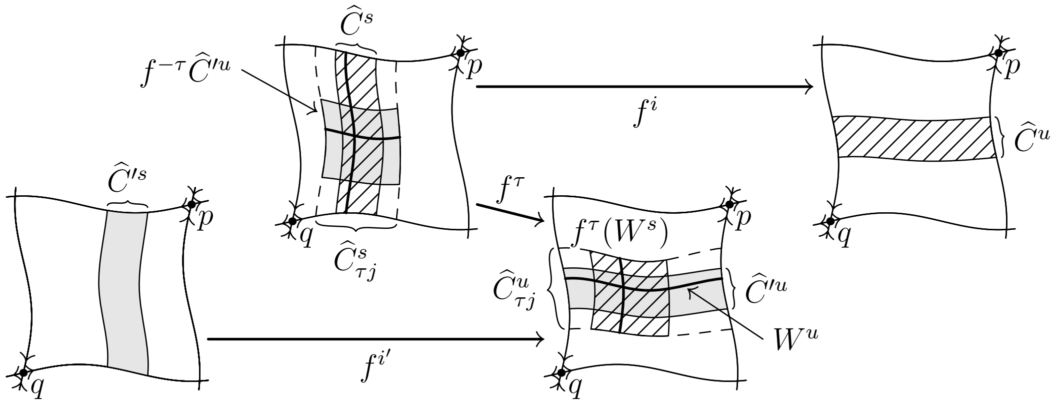

Figure 11.2. Proving Lemma 11.12.

To prove conclusion (3), that Γ has the $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-hyperbolic branch property, we start by proving the following lemma, which is illustrated in Figure 11.2.

Lemma 11.12. *Suppose that W^s and W^u are full length local stable and unstable curves in Γ_{pq} and that $\tau \in T\mathbb{N}$ is such that $f^\tau(W^s) \cap W^u \neq \emptyset$. Suppose moreover that $f^i \colon \widehat{C}^s \to \widehat{C}^u$ and $f^{i'} \colon \widehat{C}'^s \to \widehat{C}'^u$ are hyperbolic branches with $i, i' \geq \tau$ such that $W^s \subset \widehat{C}^s$ and $W^u \subset \widehat{C}'^u$. Then for any full length local stable and unstable curves $W^s_* \subset \widehat{C}^s$ and $W^u_* \subset \widehat{C}'^u$, we have $f^\tau(W^s_*) \cap W^u_* \neq \emptyset$.*

In particular, if $A \cap \widehat{C}^s \neq \emptyset$ and $A \cap \widehat{C}'^u \neq \emptyset$, then there exists $\tau j \in I(A)$ such that the hyperbolic branch $f^\tau \colon \widehat{C}^s_{\tau j} \to \widehat{C}^u_{\tau j}$ from the collection $\mathfrak{C}(A)$ has $W^s \subset \widehat{C}^s_{\tau j}$ and $W^u \subset \widehat{C}^u_{\tau j}$.

Proof. Observe that $f^\tau(\widehat{C}^s) \cap \widehat{C}'^u \supset f^\tau(W^s) \cap W^u \neq \emptyset$. Applying $f^{-\tau}$ gives

$$\widehat{C}^s \cap f^{-\tau}(\widehat{C}'^u) \neq \emptyset. \tag{11.10}$$

This implies in particular that $f^{-\tau}(\widehat{C}'^u) \cap \Gamma_{pq} \neq \emptyset$; on the other hand, Lemma 11.11 gives $\operatorname{Int} f^{-\tau}(\widehat{C}'^u) \cap \partial\Gamma_{pq} = \emptyset$, and it follows that

$$f^{-\tau}(\widehat{C}'^u) \subset \Gamma_{pq}. \tag{11.11}$$

Now we claim that

$$f^{-\tau}(\partial^s \widehat{C}'^u) \cap \operatorname{Int}(\widehat{C}^s) = \emptyset, \tag{11.12}$$

as shown in Figure 11.2. To see this, observe that

$$\begin{aligned} f^i(f^{-\tau}(\partial^s \widehat{C}'^u) \cap \operatorname{Int}(\widehat{C}^s)) &= f^{i-\tau}(\partial^s \widehat{C}'^u) \cap \operatorname{Int}(\widehat{C}^u) \\ &\subset f^{i-\tau}(W^s_p \cap W^s_q) \cap \operatorname{Int}\Gamma_{pq} = \emptyset \end{aligned}$$

since Γ_{pq} is a nice domain. Applying f^{-i} gives (11.12), and together with (11.10) and (11.11) this implies that $f^{-\tau}(\widehat{C}'^u)$ fully crosses $\widehat{C}^s$ in the unstable direction, as shown in Figure 11.2. The conclusion about W^u_* and W^s_* follows; these appear in Figure 11.2 in the same configuration as W^u and W^s do.

For the final claim in the lemma, choose $x \in A \cap \widehat{C}^s$ and $y \in \widehat{C}'^u$; then apply the first part of the lemma to deduce that $f^\tau(W^s_x) \cap W^u_y \neq \emptyset$. By Theorem D, the collection $\mathfrak{C}(A)$ contains a hyperbolic branch $f^\tau \colon \widehat{C}^s_{\tau j} \to \widehat{C}^u_{\tau j}$ associated to this almost return, which has $\widehat{C}^s_{\tau j} \supset \widehat{C}^s \supset W^s$ and $\widehat{C}^u_{\tau j} \supset \widehat{C}'^u \supset W^u$ as in Figure 11.2. $\square$

Observe that the previous two lemmas did not refer to the rectangle Γ; they are general facts about nice domains and hyperbolic branches. Now we once again consider the rectangle Γ and prove the hyperbolic branch property in conclusion (3).

Lemma 11.13. *Let Γ be the rectangle constructed in* (11.7) *using the collection $\mathfrak{C}(A)$ of $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-hyperbolic branches associated to almost returns of A by Theorem D. Let $z, z' \in \Gamma$ and $\tau \in T\mathbb{N}$ be such that $f^\tau(W^s_z) \cap W^u_{z'} \neq \emptyset$. Then there is a branch $f^\tau \colon \widehat{C}^s \to \widehat{C}^u$ in the collection $\mathfrak{C}^*(A)$ such that $z \in \widehat{C}^s$.*

Proof. A natural idea is to try to apply Lemma 11.12 to W^s_z and $W^u_{z'}$ and thus produce an almost return of A that gives the desired hyperbolic branch. However, the lemma requires us to choose branches for z and z' with order at least τ; we can find such branches in $\mathfrak{C}^*(A)$ but not necessarily in $\mathfrak{C}(A)$, and there is no guarantee that these branches will contain any elements of A (see Remark 11.1).

Thus before applying Lemma 11.12, we first observe that z has a forward hyperbolic sequence $\mathfrak{h}^+ = \{i_m j_m\}_{m=0}^\infty$, and let $n \geq 0$ be such that $\sum_{m=0}^{n-1} i_m \leq \tau < \sum_{m=0}^{n} i_m$. Let $k = \sum_{m=0}^{n-1} i_m$ and observe that by the definition of a forward hyperbolic sequence, we can concatenate the branches $f^{i_m} \colon \widehat{C}^s_{i_m j_m} \to \widehat{C}^u_{i_m j_m}$ for $0 \leq m \leq n-1$ to get a hyperbolic branch $f^k \colon \widehat{C}^s_z \to \widehat{C}^u_z$ from $\mathfrak{C}^*(A)$ such that $z \in \widehat{C}^s_z$.

Similarly, z' has a backward hyperbolic sequence $\mathfrak{h}^- = \{i_{-m} j_{-m}\}_{m=1}^\infty$, and choosing $n' \geq 1$ such that $\sum_{m=1}^{n'-1} i_{-m} \leq \tau - k < \sum_{m=1}^{n'} i_{-m}$, we can write $\ell = \sum_{m=1}^{n'-1} i_{-m}$ and do a similar concatenation to get a hyperbolic branch $f^\ell \colon \widehat{C}^s_{z'} \to \widehat{C}^u_{z'}$ from $\mathfrak{C}^*(A)$ such that $z' \in \widehat{C}^u_{z'}$.

Now we observe that

$$f^{\tau-k-\ell}(W^s_{f^k z}) \cap W^u_{f^{-\ell} z'} \supset f^{\tau-\ell}(W^s_z) \cap f^{-\ell}(W^u_{z'}) = f^{-\ell}(f^\tau(W^s_z) \cap W^u_{z'}) \neq \emptyset$$

by the invariance properties of local stable and unstable curves, and so the point $f^k(z) \in \Gamma$ has an almost return at time $\tau - k - \ell$ via $f^{-\ell}(z') \in \Gamma$. Moreover, $f^k(z) \in \widehat{C}^s_{i_n j_n}$ and $f^{-\ell}(z') \in \widehat{C}^u_{i'_{-n'} j'_{-n'}}$, where $i_n, i'_{-n'} \geq \tau - k - \ell$ by our choice of n and n'. The corresponding branches are in $\mathfrak{C}(A)$, so each of $\widehat{C}^s_{i_n j_n}$ and $\widehat{C}^u_{i'_{-n'} j'_{-n'}}$ contains an element of A, and we can apply Lemma 11.12 to get a hyperbolic branch $f^{\tau-k-\ell} \colon \widehat{C}^s_{(\tau-k-\ell)j} \to \widehat{C}^u_{(\tau-k-\ell)j}$ such that $z \in \widehat{C}^s_{(\tau-k-\ell)j}$. Concatenating this with the branches $f^k \colon \widehat{C}^s_z \to \widehat{C}^u_z$ and $f^\ell \colon \widehat{C}^s_{z'} \to \widehat{C}^u_{z'}$ gives a hyperbolic branch from $\mathfrak{C}^*(A)$ whose stable strip contains z. $\square$

Lemma 11.13 shows that every almost return of Γ is associated to a hyperbolic branch from the collection $\mathfrak{C}^*(A)$, and thus any concatenation of such branches lies in this collection as well. Since these branches are all $(\widehat{Q}e^{2\epsilon\ell}, \lambda/3)$-hyperbolic by Theorem D, we have proved conclusion (3).

In fact, we have also proved that

$$\mathfrak{C}_0(\Gamma) \subset \mathfrak{C}(\Gamma) \subset \mathfrak{C}^*(\Gamma) \subset \mathfrak{C}^*(A).$$

To prove conclusion (4) it suffices to show that $\mathfrak{C}^*(A) \subset \mathfrak{C}_0(\Gamma)$. For this, observe that any branch in $\mathfrak{C}^*(A)$ is obtained by concatenating a finite sequence of branches from $\mathfrak{C}(A)$, and repeating this finite sequence periodically in both directions produces a bi-infinite hyperbolic sequence. This hyperbolic sequence determines a unique point $x \in \Gamma$, which is periodic, and the hyperbolic branch associated to this periodic return is exactly the branch from $\mathfrak{C}^*(A)$ that we began with. This establishes conclusion (4), and then the proof of Proposition 11.3 is completed by the following result, which establishes conclusion (5).

Lemma 11.14. Γ *is saturated.*

Proof. Given a branch $f^i \colon \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}$ from $\mathfrak{C}_0(\Gamma) = \mathfrak{C}^*(A)$, for every $w \in f^{-i}(\widehat{C}^u_{ij} \cap C^s)$ we have $f^i(w) \in C^s$ and therefore $f^i(w)$ has a forward hyperbolic sequence $\{i_m j_m\}_{m\geq 0}$. The branch $f^i \colon \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}$ is a concatenation of finitely many branches from $\mathfrak{C}(A)$; adding these branches to the start of the forward hyperbolic sequence for $f^i(w)$ produces a forward hyperbolic sequence for w. It follows that $w \in C^s$, thus proving the first inclusion in (11.2). A completely analogous argument works for the unstable leaves to show that Γ is saturated. $\square$

11.3. **Proof of Proposition 11.5.** Let $\Gamma = C^s \cap C^u$ be a nice T-recurrent saturated rectangle satisfying the (C,κ)-hyperbolic branch property and let $\mathfrak{C}_0 = \mathfrak{C}_0(\Gamma)$ denote the collection of hyperbolic branches associated to returns to Γ, indexed by I_0. Recall the definition of s-subsets and u-subsets from Definition 2.4. To each branch $f^i \colon \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}$ in $\mathfrak{C}_0$ we will associate an s-subset $\Gamma^s_{ij} \subset \widehat{C}^s_{ij} \cap \Gamma$ and a u-subset $\Gamma^u_{ij} \subset \widehat{C}^u_{ij} \cap \Gamma$ such that $f^i \colon \Gamma^s_{ij} \to \Gamma^u_{ij}$ is a bijection. We stress that both inclusions are in general proper; roughly speaking, the reason for this is that there may be some $x \in \widehat{C}^s_{ij} \cap \Gamma$ for which $f^i(x) \notin \Gamma$, and such points must be excluded from Γ^s_{ij}; similarly for $x \in \widehat{C}^u_{ij}$ and $f^{-i}(x)$. To define $\Gamma^{s,u}_{ij}$, first recall from (11.2) that for $ij \in I_0$ we write

$$\begin{aligned} C^s_{ij} &= f^{-i}(\widehat{C}^u_{ij} \cap C^s) = \widehat{C}^s_{ij} \cap f^{-i}(C^s), \\ C^u_{ij} &= f^{i}(\widehat{C}^s_{ij} \cap C^u) = \widehat{C}^u_{ij} \cap f^{i}(C^u); \end{aligned} \tag{11.13}$$

then let

$$\Gamma^s_{ij} := C^s_{ij} \cap C^u, \qquad \Gamma^u_{ij} := C^u_{ij} \cap C^s. \tag{11.14}$$

Notice that C^s_{ij}, C^u_{ij} are collections of stable and unstable leaves respectively, whereas $\Gamma^s_{ij}, \Gamma^u_{ij}$ may be Cantor sets.

Lemma 11.15. *For every $ij \in I_0$, $\Gamma^s_{ij}, \Gamma^u_{ij}$ are s-subsets and u-subsets respectively of Γ and $f^i(\Gamma^s_{ij}) = \Gamma^u_{ij}$. Moreover, if $x \in \Gamma$ and $i \in T\mathbb{N}$ are such that $f^i(x) \in \Gamma$, then $x \in \Gamma^s_{ij}$ for some $ij \in I_0$; in particular this implies that $\Gamma = \bigcup_{ij\in I_0} \Gamma^s_{ij} = \bigcup_{ij\in I_0} \Gamma^u_{ij}$.*

Proof. By the saturation assumption, $C^s_{ij} \subseteq C^s, C^u_{ij} \subseteq C^u$ and therefore $\Gamma^s_{ij} := C^s_{ij} \cap C^u \subseteq C^s \cap C^u = \Gamma$ and $\Gamma^u_{ij} := C^u_{ij} \cap C^s \subseteq C^u \cap C^s = \Gamma$ and so $\Gamma^s_{ij}, \Gamma^u_{ij} \subseteq \Gamma$. Since C^s_{ij} is a union of stable leaves and C^u_{ij} is a union of unstable leaves, the sets $\Gamma^s_{ij} := C^s_{ij} \cap C^u$ and $\Gamma^u_{ij} := C^u_{ij} \cap C^s$ are s-subsets and u-subsets, respectively, of Γ. Moreover, directly from the definitions we have

$$f^i(\Gamma^s_{ij}) = f^i(\widehat{C}^s_{ij}) \cap C^s \cap f^i(C^u) = f^i(\widehat{C}^s_{ij} \cap C^u) \cap C^s = \Gamma^u_{ij}.$$

For the second statement, let $x \in \Gamma$ and $i \in T\mathbb{N}$ be such that $f^i(x) \in \Gamma$. Since Γ has the (C, κ)-hyperbolic branch property, there is a hyperbolic branch $f^i : \widehat{C}^s_{ij} \to \widehat{C}^u_{ij}$ in $\mathfrak{C}_0$ such that $x \in \widehat{C}^s_{ij} \cap \Gamma$ and $f^i(x) \in \widehat{C}^u_{ij} \cap \Gamma$. Since $\Gamma \subset C^s$, the definition of C^s_{ij} gives

$$x = f^{-i}(f^i(x)) \in f^{-i}(\widehat{C}^u_{ij} \cap \Gamma) \subset f^{-i}(\widehat{C}^u_{ij} \cap C^s) = C^s_{ij}.$$

Since we also have $x \in \Gamma = C^s \cap C^u \subseteq C^u$ it follows that $x \in C^s_{ij} \cap C^u = \Gamma^s_{ij}$. The final assertion follows because Γ is recurrent and so every $x \in \Gamma$ has some $i, i' \in T\mathbb{N}$ such that $f^i(x), f^{-i'}(x) \in \Gamma$. □

Lemma 11.15 shows that there exists a *cover* of Γ by s-subsets and another by u-subsets satisfying the Markov property **(Y0)** required in the definition of topological Young tower. It does not however claim that this cover is a *partition* of Γ as required by the definition, i.e. that the s-subsets of the cover are disjoint which in fact they may not be. The following Lemma shows however that they are pairwise either nested or disjoint, and this will then allow us to choose a sub-cover made up of pairwise disjoint sets (see the proof of Proposition 11.5 below).

Lemma 11.16. *Let $k\ell \in I_0$ and suppose there exists $x \in \Gamma^s_{k\ell}$ and $0 < i < k$ such that $f^i(x) \in \Gamma$. Then there exist $ij \in I_0$ such that $\Gamma^s_{k\ell} \subseteq \Gamma^s_{ij}$. In particular all $\{\Gamma^s_{ij}\}_{ij\in I_0}$ are pairwise either nested or disjoint.*

We will prove this lemma momentarily.

Remark 11.17. Notice that the last statement in Lemma 11.16 does not follow directly from Lemma 11.11. Indeed, the fact that two stable strips $\widehat{C}^s_{ij}, \widehat{C}^s_{k\ell}$ are nested does not *a priori* imply that the corresponding sets $C^s_{ij}, C^s_{k\ell}$ are either disjoint or nested, recall (11.2), and therefore also does not *a priori* imply that $\Gamma^s_{ij}, \Gamma^s_{k\ell}$ are either disjoint or nested, recall (11.14).

Sublemma 11.18. *In the setting of Lemma 11.16, letting $m = k - i$, there exist $ij, mn \in I_0$ such that $\widehat{C}^s_{k\ell} \subseteq \widehat{C}^s_{ij}$, $\widehat{C}^u_{k\ell} \subseteq \widehat{C}^u_{mn}$, and such that $f^i(\widehat{C}^s_{k\ell}) = \widehat{C}^u_{ij} \cap \widehat{C}^s_{mn} = f^{-m}(\widehat{C}^u_{k\ell})$.*

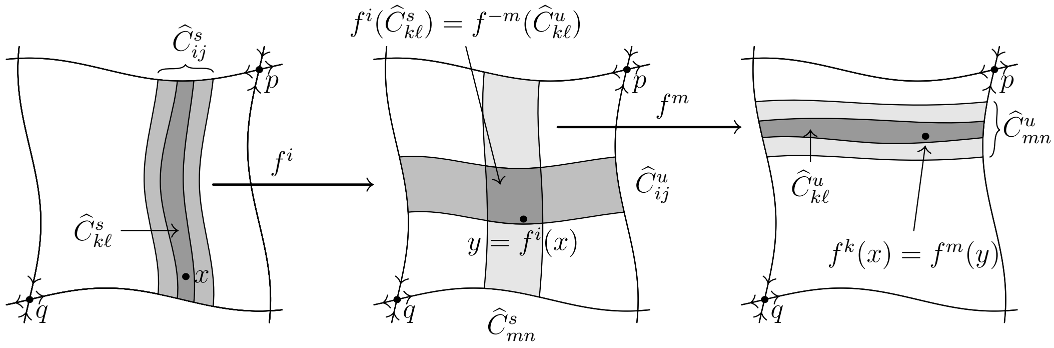

FIGURE 11.3. Proof of Sublemma 11.18.

Proof. From Lemma 11.15 we have $x \in \Gamma^s_{ij}$ for some $ij \in I_0$. Therefore $x \in \Gamma^s_{ij} \cap \Gamma^s_{k\ell}$ and so $x \in \widehat{C}^s_{ij} \cap \widehat{C}^s_{k\ell}$ and in particular $\widehat{C}^s_{ij} \cap \widehat{C}^s_{k\ell} \neq \emptyset$ and therefore, by Lemma 11.11, $\widehat{C}^s_{k\ell} \subseteq \widehat{C}^s_{ij}$, as shown in Figure 11.3. Also, from Lemma 11.15, $x \in \Gamma^s_{k\ell}$ implies $x \in \Gamma$ and $f^k(x) \in \Gamma$ and therefore, letting $y = f^i(x) \in \Gamma$ we have $f^m(y) = f^m(f^i(x)) = f^k(x) \in \Gamma$. Thus there exists $mn \in I_0$ such that $y \in \widehat{C}^s_{mn}$. We therefore have $f^m(y) \in f^m(\widehat{C}^s_{mn}) = \widehat{C}^u_{mn}$ and also $f^m(y) = f^k(x) \in \Gamma^u_{k\ell} \subseteq \widehat{C}^u_{k\ell}$, and therefore $\widehat{C}^u_{mn} \cap \widehat{C}^u_{k\ell} \neq \emptyset$ and thus, since $m < k$, $\widehat{C}^u_{k\ell} \subseteq \widehat{C}^u_{mn}$. Then, since $\widehat{C}^s_{k\ell} \subseteq \widehat{C}^s_{ij}$ are both full height vertical (stable) strips and $\widehat{C}^u_{k\ell} \subseteq \widehat{C}^u_{mn}$ are both full length horizontal (unstable) strips, it follows that $f^i(\widehat{C}^s_{k\ell})$ is "full height" relative to the horizontal strip $\widehat{C}^u_{ij}$ and $f^{-m}(\widehat{C}^u_{k\ell})$ is "full width" relative to the vertical strip $\widehat{C}^s_{mn}$. Since $f^i(\widehat{C}^s_{k\ell}) = f^{-m}(\widehat{C}^u_{k\ell})$ we complete the proof. □

Proof of Lemma 11.16. Let $m = k - i$. Directly from the definitions,

$$C^s_{k\ell} = f^{-k}(\widehat{C}^u_{k\ell} \cap C^s) = f^{-i}(f^{-m}(\widehat{C}^u_{k\ell} \cap C^s)). \tag{11.15}$$

From Sublemma 11.18 we have $f^{-m}(\widehat{C}^u_{k\ell}) = \widehat{C}^u_{ij} \cap \widehat{C}^s_{mn}$ and thus (11.13) gives

$$f^{-m}(\widehat{C}^u_{k\ell} \cap C^s) = \widehat{C}^u_{ij} \cap \widehat{C}^s_{mn} \cap f^{-m}(C^s) = \widehat{C}^u_{ij} \cap C^s_{mn}.$$

Substituting this into (11.15) and using the saturation condition which implies $C^s_{mn} \subseteq C^s$, we get $C^s_{k\ell} = f^{-i}(\widehat{C}^u_{ij} \cap C^s_{mn}) \subseteq f^{-i}(\widehat{C}^u_{ij} \cap C^s) =: C^s_{ij}$, which implies the statement since $\Gamma^s_{k\ell} = C^s_{k\ell} \cap C^u \subseteq C^s_{ij} \cap C^u = \Gamma^s_{ij}$ by (11.14). □

Proof of Proposition 11.5. Since the family of sets $\{\Gamma^s_{ij}\}_{ij \in I_0}$ are pairwise either nested or disjoint, they are partially ordered by inclusion. We can therefore define the set $I'_0 \subset I_0$ of indices ij which are *maximal* with respect to this partial order. We then let $\mathcal{P} := \{\Gamma^s_{ij}\}_{ij \in I'_0}$. By Lemma 11.15, every point $x \in \Gamma^s$ belongs to some Γ^s_{ij} for some $ij \in I_0$ and therefore must also belong to some maximal element Γ^s_{ij} for some $ij \in I'_0$. Thus $\mathcal{P}$ is a partition of Γ into *pairwise disjoint* s-subsets whose images are u-subsets. This gives the Markov–Young structure. To see that it is a First Return Topological Young Tower we suppose by contradiction that there exists some $k\ell \in I'_0$, $x \in \Gamma^s_{k\ell}$ and

$0 < i < k$ such that $f^i(x) \in \Gamma$. Then Lemma 11.16 implies that there exists some $ij \in I_0$ such that $\Gamma^s_{k\ell} \subset \Gamma^s_{ij}$, contradicting the maximality of $\Gamma^s_{k\ell}$. $\square$

11.4. **Proof of Proposition 11.6.** We split the proof into two independent parts, one to prove the hyperbolicity and distortion conditions **(Y1)**–**(Y2)** using the hyperbolic branch property and Hölder estimate (11.3), and the second to prove the integrability of the return times, which follows from the Lebesgue-strong T-return property.

11.4.1. *Hyperbolicity and distortion properties of the tower.* We will verify Conditions **(Y1)** and **(Y2)** in Definition 2.10. Fix $i \in T\mathbb{N}$, $j \in \{1, \dots, m_i\}$, and $x \in \Gamma^s_{ij}$. Note that the map $F = f^i \colon \Gamma^s_{ij} \to \Gamma^u_{ij}$ has the (C, κ)-hyperbolic branch property with constant $C > 0$ and $\kappa > 0$ independent of x. Since the number T is large enough, (see Condition 3 in Definition 1.16) it ensures that $Ce^{-\kappa T} < 1$ and Condition **(Y1)**(a) follows with $\kappa_1 = Ce^{-\kappa T}$. Condition **(Y1)**(b) can be shown by a similar argument.

We now prove Condition **(Y2)**(a), the proof of Condition **(Y2)**(b) is similar. It suffices to show that there are constants $c > 0$ and $\alpha_1 > 0$ such that for any $z \in \Gamma$ and $w \in V^s_z$ we have

$$\Big|\log \frac{\mathrm{Jac}^u F(z)}{\mathrm{Jac}^u F(w)}\Big| \le c d(z, w)^{\alpha_1}. \tag{11.16}$$

Indeed, setting $z = F^n(x)$ and $w = F^n(y)$, the desired bounded distortion estimate follows from (11.16) and Condition **(Y1)**(a).

To show (11.16) notice that $z \in \Gamma^s_{ij}$ for some $i \in T\mathbb{N}$ and $j \in \{1, \dots, m_i\}$ and hence,

$$\Big|\log \frac{\mathrm{Jac}^u F(z)}{\mathrm{Jac}^u F(w)}\Big| = \Big|\sum_{a=0}^{i-1} \log \frac{\mathrm{Jac}^u f(f^a(z))}{\mathrm{Jac}^u f(f^a(w))}\Big|. \tag{11.17}$$

Since Γ is a (C_0, κ)-rectangle, we see that for any $z \in \Gamma$, $w \in V^s_z$, and $0 \le a \le i - 1$, we have $f^a(w) \in V^s_{f^a(z)}$ and

$$d(f^a(z), f^a(w)) \le C_0 e^{-\kappa a} d(z, w). \tag{11.18}$$

Moreover, using the assumption that Γ satisfies (11.3), we see that

$$\begin{aligned} d(E^u_{f^a(z)}, E^u_{f^a(w)}) &\le C_1 e^{\gamma_1 a} d(f^a(z), f^a(w))^{\beta_1} \\ &\le C_1 e^{\gamma_1 a} C_0^{\beta_1} e^{-\kappa a \beta_1} d(z, w)^{\beta_1} = C_1 C_0^{\beta_1} e^{(\gamma_1 - \kappa\beta_1)a} d(z, w)^{\beta_1}. \end{aligned} \tag{11.19}$$

Since f is $C^{1+\alpha}$, it follows that there is $C_2 > 0$ such that

$$\begin{aligned} \Big|\frac{\mathrm{Jac}^u f(f^a(z))}{\mathrm{Jac}^u f(f^a(w))} - 1\Big| &= \Big|\frac{\mathrm{Jac}^u f(f^a(z)) - \mathrm{Jac}^u f(f^a(w))}{\mathrm{Jac}^u f(f^a(w))}\Big| \\ &\le C_2 d(E^u_{f^a(z)}, E^u_{f^a(w)})^\alpha \\ &\le C_2 C_1^\alpha C_0^{\beta_1 \alpha} e^{\alpha(\gamma_1 - \kappa\beta_1)a} d(z, w)^{\beta_1 \alpha}. \end{aligned} \tag{11.20}$$

If $\mathrm{Jac}^u f(f^a(z)) \ge \mathrm{Jac}^u f(f^a(w))$ this gives

$$\log \frac{\mathrm{Jac}^u f(f^a(z))}{\mathrm{Jac}^u f(f^a(w))} \le \frac{\mathrm{Jac}^u f(f^a(z))}{\mathrm{Jac}^u f(f^a(w))} - 1 \le C_2 C_1^\alpha C_0^{\beta_1 \alpha} e^{\alpha(\gamma_1 - \kappa\beta_1)a} d(z, w)^{\beta_1 \alpha}.$$

The same bound holds with z and w exchanged, by exchanging their roles in (11.20). Then by (11.17) we get

$$\Big|\log\frac{\mathrm{Jac}^u F(z)}{\mathrm{Jac}^u F(w)}\Big| \le \sum_{a=0}^{i-1} C_2 C_1^{\alpha} C_0^{\beta_1\alpha} e^{\alpha(\gamma_1-\kappa\beta_1)a} d(z,w)^{\beta_1\alpha}. \tag{11.21}$$

Since we assumed $\gamma_1 < \beta_1\kappa$, we see that $\gamma_1 - \kappa\beta_1 < 0$, and thus (11.21) implies (11.16), which proves Condition **(Y2)**(a).

11.4.2. *Integrability of the return times.* Let $F\colon \Gamma \to \Gamma$ be the induced map to the base of the Topological First T-Return Young Tower where $F(x) = f^{\tau(x)}(x)$ and $\tau(x)$ is the first return time to Γ which is a multiple of T.

For every $n \ge 0$ let

$$R_n(x) := \sum_{j=0}^{n-1} \tau(F^j(x)) \quad \text{and} \quad v_n(x) := \#\{0 < i \le n/T : f^{iT}(x) \in \Gamma\}$$

where we define $R_0(x) = 0$ by convention. Notice that R_n, v_n are on quite different time scales, the index n in v_n refers to the iterates of the original map f, whereas in R_n it refers to the iterates of the induced map F. The following relation between the two quantities is not surprising but neither is it completely trivial, we thank Vilton Pinheiro for explaining it to us.

Lemma 11.19 (Pinheiro [60])**.** *Let $x \in \Gamma$ and suppose $\lim_{n\to\infty} R_n(x)/n$ exists. Then*

$$\lim_{n\to\infty} \frac{v_n(x)}{n} = \left(\lim_{n\to\infty} \frac{R_n(x)}{n}\right)^{-1}.$$

In particular $\lim_{n\to\infty} v_n(x)/n$ exists and is equal to 0 if $\lim_{n\to\infty} R_n(x)/n = \infty$.

Proof of Lemma 11.19. For any $x \in \Gamma$, by definition of $v_n(x)$ we have

$$v_n(x) := \#\{0 < i \le n/T : f^{iT}(x) \in \Gamma\} = \max\Big\{k \ge 0 : \sum_{j=0}^{k-1} \tau(F^j(x)) \le n\Big\}$$

and so, for every $n \ge \tau(x)$, so that $v_n(x) \ge 1$, we have

$$R_{v_n(x)}(x) := \sum_{j=0}^{v_n(x)-1} \tau(F^j(x)) \le n < \sum_{j=0}^{v_n(x)} \tau(F^j(x)) =: R_{v_n(x)+1}(x).$$

Dividing through by $v_n(x)$ this gives

$$\frac{R_{v_n(x)}(x)}{v_n(x)} \le \frac{n}{v_n(x)} < \frac{v_n(x)+1}{v_n(x)} \frac{R_{v_n(x)+1}(x)}{v_n(x)+1} \tag{11.22}$$

Since $(v_n(x)+1)/v_n(x) \to 1$ as $n \to \infty$ and the limit of the sequence $R_n(x)/n$ exists, the subsequences on the left and right hand side of (11.22) also converge to the same limit. It follows that the $n/v_n(x)$ converges and therefore also $v_n(x)/n$ to a limit as in the statement. $\square$

Lemma 11.19 leads to the integrability of the return times stated in Definition 2.11. Indeed, by the results in [69] the induced map $F\colon \Gamma \to \Gamma$ admits an SRB measure $\hat\mu$ whose conditional measures $\hat\mu_z$ on unstable curves of points of Γ are equivalent to the Lebesgue measure $m_{V^u_z}$ on these same curves. It is therefore sufficient to show the integrability with respect to one of these conditional measures. For $\hat\mu$-a.e. $z \in \Gamma$ there is $E_z \subset V^u_z$ with $\hat\mu_z(E^c_z) = 0$ such that

$$\lim_{n\to\infty} \frac{R_n(x)}{n} = \int_\Gamma \tau \, d\hat\mu = \int_\Gamma \int \tau \, d\hat\mu_z \, d\hat\mu(z) \text{ for all } x \in E_z, \tag{11.23}$$

where a priori these quantities may be infinite. However, since Γ is Lebesgue-strongly T-recurrent, there is a local unstable curve V^u and a set $E \subset V^u \cap \Gamma$ of positive 1-dimensional Lebesgue measure such that for every $x \in E$ we have $\limsup_{n\to\infty} v_n(x)/n > 0$ (see (1.9)). Now Lemma 11.19 implies that if $x \in E$ is such that $\lim_{n\to\infty} R_n(x)/n$ exists, then this limit is finite.

Observe that for each $n \geq 1$ and $y \in \Gamma$, the function $x \mapsto R_n(x)/n$ is constant on $V^s_y \cap \Gamma$. In particular, fixing $z \in \Gamma$ and writing $E'_z = \bigcup_{y\in E} V^s_y \cap V^u_z$, we see that $E'_z \subset V^u_z$ has $\hat\mu_z(E'_z) > 0$ and that $\lim_{n\to\infty} R_n(x)/n$ is finite for every $x \in E'_z$ such that the limit exists.

Choosing $z \in \Gamma$ such that $\hat\mu_z(E^c_z) = 0$, we get $\hat\mu_z(E'_z \cap E_z) > 0$. For any point $x \in E'_z \cap E_z$, (11.23) gives $\int_\Gamma \tau \, d\hat\mu < \infty$, and it follows that $\int \tau \, d\hat\mu_z < \infty$ for $\hat\mu$-a.e. z. This completes the proof of the integrability of the return times and thus the proof of Proposition 11.6.

11.5. **Discussion of the saturation condition.** The saturation condition in Definition 11.2 is crucial to our construction of a first T-return Young tower, but may appear mysterious at first glance, so we give some general discussion here of the role it plays; this section is not part of the proofs.

The strategy in the proof of Theorem B was to use a collection of hyperbolic branches $\mathfrak{C}$ to produce a rectangle Γ (Proposition 11.3), and then to build a Young tower that has Γ as its base (Propositions 11.5 and 11.6). In §11.2, Proposition 11.3 was proved by taking $\mathfrak{C}$ to be the collection of branches associated to almost returns of some nice regular set $A \subset \Gamma_{pq}$, but Lemma 11.9 shows that the construction of a nice rectangle Γ that is recurrent works if we begin with *any* collection $\mathfrak{C}$ of hyperbolic branches.

The proof of conclusions (3), (4), and (5) in Proposition 11.3, however, required us to work with the specific choice of $\mathfrak{C}$ given by the collection of almost returns. To illustrate why this was important, consider the following example, which shows how (4) and (5) can fail for a general collection $\mathfrak{C}$, where the collection of hyperbolic branches associated to returns of Γ may contain some branches that do not appear as concatenations of branches from the generating set $\mathfrak{C}$.

Example 11.20. Suppose that p, q are fixed points and that Γ_{pq} contains two hyperbolic branches $f\colon \widehat{C}^s_0 \to \widehat{C}^u_0$ and $f\colon \widehat{C}^s_1 \to \widehat{C}^u_1$. Define a map $\pi\colon \{0,1\}^{\mathbb{Z}} \to \Gamma_{pq}$ by $\pi(x) = \bigcap_{n\in\mathbb{Z}} f^{-n}(\widehat{C}^s_{x_n})$. To each finite word $w = w_0 w_1 \cdots w_{n-1} \in \{0,1\}^n$ we can associate the cylinder $[w] = \{x \in \{0,1\}^{\mathbb{Z}} :$

$x_j = w_j$ for all $0 \le j < n\}$, and $\pi[w]$ is the stable strip for a hyperbolic branch of order n (recall Definition 4.8) associated to the word w.

Let $\mathfrak{C}$ be the collection of two hyperbolic branches associated to the words $v = 01$ and $w = 10$. Applying Lemma 11.9 to $\mathfrak{C}$ produces the rectangle Γ consisting of all points $\pi(x)$ for which $x_{2n}x_{2n+1} \in \{v, w\}$ for all $n \in \mathbb{Z}$. Observe that for $x = \overline{10}.\overline{10} \in \{0,1\}^{\mathbb{Z}}$ (where the overline denotes infinite repetition to either the left or the right), we have $\pi(x) \in \Gamma$ and $f(\pi(x)) \in \Gamma$ (this is a period-2 orbit), but the hyperbolic branch corresponding to this return is $f\colon \widehat{C}^s_1 \to \widehat{C}^u_1$, which is not a concatenation of branches from $\mathfrak{C}$, so conclusion (4) fails. Moreover, the rectangle Γ is not saturated, because $z = \pi(\overline{01}.\overline{10}) \in \Gamma$ has

$$f^{-1}(\widehat{C}^u_1 \cap W^s_z) = \{\pi(y) : y_0 y_1 y_2 \cdots = 1\overline{10}\},$$

which is a full-length local stable curve that is disjoint from Γ and thus does not lie in C^s, violating (11.2). Then the construction of a topological Young tower in §11.3 fails: Lemma 11.15 does not go through because the sets defined in (11.14) are not contained in Γ.

In fact one can still build a "not first return" topological Young tower over Γ by writing $\Gamma^s_v = \{\pi(y) : y_0y_1 = v\}$ and $\Gamma^s_w = \{\pi(y) : y_0y_1 = w\}$, then setting the inducing time τ to be equal to 2 everywhere so that $\Gamma^u_v = f^2(\Gamma^s_v) = \{\pi(y) : y_{-2}y_{-1} = v\}$, and similarly for Γ^u_w. This gives the Markov structure in **(Y0)**, but is not a first return tower because some points, such as the point $\pi(x)$ given in Example 11.20, return before time 2.

One might reasonably object that the rectangle in Example 11.20 is indeed saturated if we take $T = 2$ and only consider returns at even times, since then every T-return is associated to a concatenation of the branches in $\mathfrak{C}$, and thus Γ supports a first 2-return topological Young tower. This phenomenon occurs if we start with any finite collection of branches $\mathfrak{C}$: taking T to be any common multiple of their lengths will guarantee that the rectangle is saturated (for T-returns) and produces a first T-return topological Young tower. This simple solution fails when $\mathfrak{C}$ is infinite, as in the next example.

Example 11.21. Given three hyperbolic branches $\{f\colon \widehat{C}^s_j \to \widehat{C}^u_j\}_{j\in\{0,1,2\}}$, define $\pi\colon \{0,1,2\}^{\mathbb{Z}} \to \Gamma_{pq}$ as in Example 11.20, and let $\mathfrak{C}$ be the infinite collection of hyperbolic branches associated to words of the form 210^n1 and 10^n12 for $n \ge 0$. Then given any $i \ge 4$, we observe that $x = \overline{210^{i-4}1}.\overline{210^{i-4}1}$ has the property that $\pi(x), f^i(\pi(x)) \in \Gamma$, so for $w = 210^{i-4}12 \in \{0,1,2\}^i$, the branch $f^i\colon \widehat{C}^s_w = \widehat{C}^u_w$ is associated to a return to Γ, but is not a concatenation of branches from $\mathfrak{C}$. Putting $z = \pi(\overline{10^{i-4}12}.\overline{210^{i-4}1})$ one can check that $f^{-i}(\widehat{C}^u_w \cap W^s_z) \not\subset C^s$, so (11.2) fails and Γ is not saturated for any choice of T. Because the stable strips associated to the branches in $\mathfrak{C}$ are disjoint, one can still build a topological Young tower (with infinitely many branches) as in the paragraph following Example 11.20, but it will not be a first T-return tower for any choice of T.

The reason we can avoid these problems in Proposition 11.3 is that the collection of branches we use there is not arbitrary: rather, it is the collection of *all* hyperbolic branches associated to almost returns to a particular nice regular set A. For such a collection, Lemmas 11.12 and 11.13 guarantee that any T-return for the rectangle Γ corresponds to an almost return for the original set A, and thus to a branch from the original collection. Thus one may interpret the saturation condition as the requirement that the original collection of branches be "large enough" that no new branches are created by "accidental returns".

For a discussion of a related issue in the setting of coded shift spaces, see [**?**, §3], and specifically condition [III*] of that paper, which plays the same role there as saturation does here, guaranteeing that a certain Markov coding is 1-1. Example 3.6 of that paper could be translated into the setting of Examples 11.20 and 11.21 by letting $\mathfrak{C}$ be the collection of four hyperbolic branches associated to the words $\{0, 10, 01, 101\}$; then the associated rectangle is not saturated, and moreover the procedure described after Example 11.20 does not produce a topological Young tower at all since the s-sets Γ^s_w are not disjoint. One can still obtain a Markov coding but it is uncountable-to-1. Together with the previous two examples, this illustrates how important the saturation condition is for our results.

12. Hyperbolic measures have nice regular sets: Proof of Theorem C

In this section we prove Theorem C. The non-trivial part of the proof is to show that we can find arbitrarily small domains Γ_{pq} with $\mu(\Gamma_{pq} \cap \Lambda_\ell) > 0$ where $p, q \in \Lambda_\ell$ are periodic points. Then letting $T > 0$ be any common multiple of the periods of p and q, it follows that p, q are fixed points for f^T and therefore Γ_{pq} is a nice domain with $T(\Gamma_{pq}) = T$. Moreover, μ is also f^T-invariant and therefore μ-a.e. $x \in A := \Gamma_{pq} \cap \Lambda_\ell$ returns to A with positive frequency for iterates which are multiples of T, in both forward and backward time. Thus A is T-recurrent, and if μ is an SRB measure it is Lebesgue-strongly T-recurrent. Thus Theorem C follows from the statement below.

Proposition 12.1. *Let f be a $C^{1+\alpha}$ diffeomorphism, μ an ergodic non-atomic χ-hyperbolic measure, and Λ a χ-hyperbolic set. Fix $\lambda \in (0, \chi)$ and $\epsilon \in (0, \epsilon_1(f, \chi, \lambda))$. Let $U \subset M$ be an open set and $\ell \in \mathbb{N}$ such that $\mu(U \cap \Lambda_\ell) > 0$. Then with ℓ' as in Theorem F, there are $(\lambda/4, 2\epsilon, \ell + \ell')$-regular periodic points p, q such that Γ_{pq} is defined, contained in U, and satisfies $\mu(\Gamma_{pq} \cap \Lambda_\ell) > 0$. Since $\Gamma_{pq} \subset U$, $\operatorname{diam} \Gamma_{pq}$ can be made arbitrarily small.*

Before proving Proposition 12.1, we use Theorem F to establish a result reminiscent of the Katok Closing Lemma. Say that $y \in B(x, \delta) \cap \Lambda_\ell$ is Λ_ℓ-*nonwandering* if there is a sequence $n_k \to \infty$ and $y_k \in \Lambda_\ell \cap f^{-n_k} \Lambda_\ell$ such that $y_k, f^{n_k}(y_k) \xrightarrow{k\to\infty} y$. Observe that by Poincaré Recurrence, every point in $\operatorname{supp}(\mu|\Lambda_\ell)$ is Λ_ℓ-nonwandering.

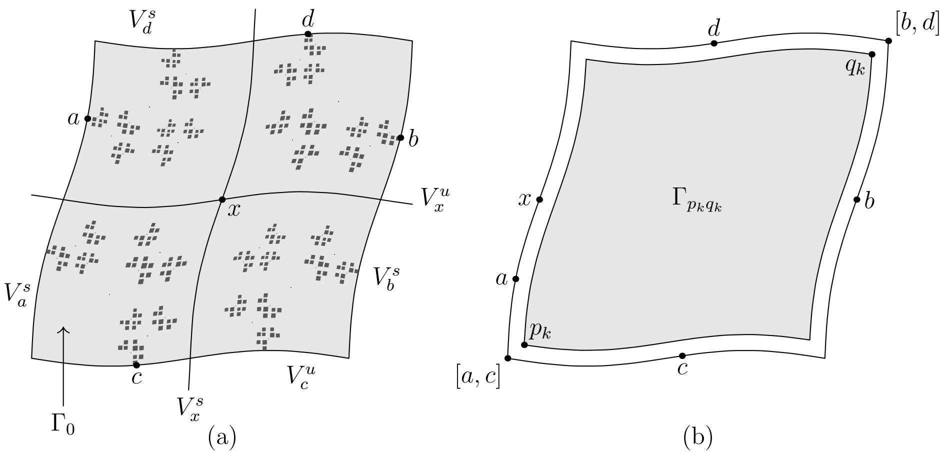

FIGURE 12.1. Proving Proposition 12.1.

Lemma 12.2. *Given $\delta > 0$ as in Theorem E, $\ell' \in \mathbb{N}$ as in Theorem F, and any $\ell \in \mathbb{N}$, for all Λ_ℓ-nonwandering points $y, z \in B(x, \delta e^{-\lambda\ell}/3) \cap \Lambda_\ell$ there is a sequence of $(\lambda/4, 2\epsilon, \ell+\ell')$-regular periodic points $p_k \xrightarrow{k\to\infty} V_y^s \cap V_z^u$.*

Proof. Suppose y, z are as in the hypothesis. Choose $n_k \to \infty$ and $y_k \in \Lambda_\ell \cap f^{-n_k}\Lambda_\ell$ such that $y_k, f^{n_k}(y_k) \to y$. Choose m_k, z_k similarly for z. For suitably large k, we have

$$y_k, f^{n_k}(y_k), z_k, f^{m_k}(z_k) \in B(x, \delta e^{-\lambda\ell}/2)$$

and thus in particular

$$d(f^{n_k}(y_k), z_k) \le \delta e^{-\lambda\ell} \quad \text{and} \quad d(f^{m_k}(z_k), y_k) \le \delta e^{-\lambda\ell}.$$

It follows that $y_k, f(y_k), \dots, f^{n_k-1}(y_k), z_k, f(z_k), \dots, f^{m_k-1}(z_k), y_k$ is a $(\bar\ell, \delta, \lambda)$-pseudo-orbit with

$$\ell_i = \ell + \begin{cases} \min(i, n_k - i) & 0 \le i \le n_k, \\ \min(i - n_k, n_k + m_k - i) & n_k \le i \le n_k + m_k. \end{cases}$$

Repeating this finite pseudo-orbit $\bar x$ periodically gives a periodic bi-infinite pseudo-orbit to which we can apply Theorem F and obtain a $(\lambda/4, 2\epsilon, \ell+\ell')$-regular periodic shadowing point p_k. Note that $p_k \in \mathcal{N}_{\bar x}^0 \cap \mathcal{N}_{\bar x}^{n_k+m_k}$, and that the intersections converge to $V_y^s \cap V_z^u$ as $k \to \infty$ because $y_k \in \mathcal{N}_{\bar x}^0$ and $f^{m_k} z_k \in \mathcal{N}_{\bar x}^{n_k+m_k}$. Thus $p_k \to V_y^s \cap V_z^u$. $\square$

Proof of Proposition 12.1. Fix $x \in U \cap \operatorname{supp}(\mu|\Lambda_\ell)$. Since Λ_ℓ is closed, we have $\operatorname{supp}(\mu|\Lambda_\ell) \subset \Lambda_\ell$. Choose $\delta', \delta > 0$ sufficiently small that $B(x, \delta') \subset U$, and such that for every $y, z \in \overline{B(x,\delta)} \cap \Lambda_\ell$, the intersection $V_y^s \cap V_z^u$ is a single point and lies in $B(x, \delta') \cap \Lambda_{\ell'}$. Assume also that δ is chosen small enough and ℓ' large enough to satisfy Lemma 12.2.

Let $Z := \overline{B(x,\delta)} \cap \operatorname{supp}(\mu|\Lambda_\ell)$. Observe that Z is compact, and that $\mu(Z) > 0$ by our choice of x. Let $\pi^s \colon Z \to V_x^u$ and $\pi^u \colon Z \to V_x^s$ be projection along local stable and unstable leaves, respectively. Since $V_x^{s,u}$ are one-dimensional we can equip each with a total order, and by compactness we can choose $a, b, c, d \in Z$ such that

$$\pi^s(a) = \inf \pi^s(Z), \quad \pi^s(b) = \sup \pi^s(Z),$$
$$\pi^u(c) = \inf \pi^u(Z), \quad \pi^u(d) = \sup \pi^u(Z).$$

Let Γ_0 be the region bounded by V_a^s, V_b^s, V_c^u, and V_d^u, as shown in Figure 12.1. Observe that $\Gamma_0 \supset Z$ and thus $\mu(\Gamma_0 \cap \Lambda_\ell) > 0$. By Lemma 12.2 there are periodic points $p_k, q_k \in \Lambda_{\ell'}$ such that $p_k \to V_a^s \cap V_c^u$ and $q_k \to V_b^s \cap V_d^u$. It is possible that none of the domains $\Gamma_{p_k q_k}$ contains x (this can occur, for example, if $a \in V_x^u$, as in Figure 12.1(b)); on the other hand, the union $\bigcup_n \Gamma_{p_n q_n}$ covers all of Z except possibly for $Z \cap V_a^s \cup V_b^s \cup V_c^u \cup V_d^u$. Since μ is non-atomic, a single local stable or unstable curve always has zero measure, thus this subset is μ-null. Using the fact that $\mu(Z) > 0$, we conclude that there is some n such that $\mu(\Gamma_{p_n q_n} \cap Z) > 0$. This completes the proof of the proposition. □

Appendix A. List of terminology and notation

(1) almost returns, Definition 4.13 on page 22
(2) brackets, Definition 1.11 on page 7
(3) branch
  – (C, κ)-hyperbolic, Definition 4.8 on page 20
  – $\hat{\ell}$-regular, Definition 5.6 on page 27
  – (C, κ)-hyperbolic branch property, Definition 4.14 on page 23
(4) overlapping charts, Definition 8.1 on page 39
(5) concatenation property, Definition 4.12 on page 22
(6) cones, Definition 4.2 on page 19
  – in regular neighbourhoods, Definition 5.10 on page 26
(7) conefield, Definition 4.3 on page 20
  – adapted, Definition 4.5 on page 20
(8) curves
  – local (C, λ)-stable (unstable), Definition 1.10 on page 7
  – $\mathcal{K}$-admissible, Definition 4.4 on page 20
  – stable and unstable admissible, Definition 4.6 on page 20
  – in regular neighbourhoods, Definition 5.2 on page 26
  – full length stable and unstable admissible, Definition 4.6 and Definition 5.2
(9) $(\chi, \varepsilon, \ell, r)$-nice domain, Definition 1.16 on page 9
(10) measure
  – hyperbolic, Definition 1.2 on page 4
  – physical, Definition 1.1 on page 3
  – SRB, Definition 1.4 on page 5
(11) nice
  – domain, Definition 1.16 on page 9
  – regular set, Definition 1.19 on page 10
  – rectangle, Definition 2.3 on page 14

References

1. José F. Alves, Christian Bonatti, and Marcelo Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), no. 2, 351–398. MR 1757000
2. José F. Alves, Maria Carvalho, and Jorge Milhazes Freitas, *Statistical stability and continuity of SRB entropy for systems with Gibbs-Markov structures*, Comm. Math. Phys. **296** (2010), no. 3, 739–767. MR 2628821
3. ______, *Statistical stability for Hénon maps of the Benedicks-Carleson type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **27** (2010), no. 2, 595–637. MR 2595193
4. José F. Alves, Carla L. Dias, and Stefano Luzzatto, *Geometry of expanding absolutely continuous invariant measures and the liftability problem*, Ann. Inst. H. Poincaré Anal. Non Linéaire **30** (2013), no. 1, 101–120. MR 3011293
5. José F. Alves, Carla L. Dias, Stefano Luzzatto, and Vilton Pinheiro, *SRB measures for partially hyperbolic systems whose central direction is weakly expanding*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 10, 2911–2946. MR 3712997

6. José F. Alves, Jorge M. Freitas, Stefano Luzzatto, and Sandro Vaienti, *From rates of mixing to recurrence times via large deviations*, Adv. Math. **228** (2011), no. 2, 1203–1236. MR 2822221
7. José F. Alves and Xin Li, *Gibbs-Markov-Young structures with (stretched) exponential tail for partially hyperbolic attractors*, Adv. Math. **279** (2015), 405–437. MR 3345188
8. José F. Alves, Stefano Luzzatto, and Vilton Pinheiro, *Markov structures and decay of correlations for non-uniformly expanding dynamical systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 6, 817–839. MR 2172861
9. José F. Alves and Vilton Pinheiro, *Slow rates of mixing for dynamical systems with hyperbolic structures*, J. Stat. Phys. **131** (2008), no. 3, 505–534. MR 2386574
10. ———, *Gibbs-Markov structures and limit laws for partially hyperbolic attractors with mostly expanding central direction*, Adv. Math. **223** (2010), no. 5, 1706–1730. MR 2592507
11. Péter Bálint and Imre Péter Tóth, *Exponential decay of correlations in multi-dimensional dispersing billiards*, Ann. Henri Poincaré **9** (2008), no. 7, 1309–1369. MR 2453251
12. L. Barreira and Ya. Pesin, *Lectures on Lyapunov exponents and smooth ergodic theory*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, Appendix A by M. Brin and Appendix B by D. Dolgopyat, H. Hu and Pesin, pp. 3–106. MR 1858534
13. Luis Barreira and Yakov Pesin, *Nonuniform hyperbolicity*, Encyclopedia of Mathematics and its Applications, vol. 115, Cambridge University Press, Cambridge, 2007, Dynamics of systems with nonzero Lyapunov exponents. MR 2348606
14. Snir Ben Ovadia, *Symbolic dynamics for non-uniformly hyperbolic diffeomorphisms of compact smooth manifolds*, J. Mod. Dyn. **13** (2018), 43–113. MR 3918259
15. ———, *Generalized SRB measures, physical properties, and thermodynamic formalism of smooth hyperbolic systems*, arXiv:1904.10074 (2019).
16. ———, *Hyperbolic srb measures and the leaf condition*, Preprint (2020).
17. Michael Benedicks and Lai-Sang Young, *Sinaĭ-Bowen-Ruelle measures for certain Hénon maps*, Invent. Math. **112** (1993), no. 3, 541–576. MR 1218323
18. ———, *Markov extensions and decay of correlations for certain Hénon maps*, Astérisque (2000), no. 261, xi, 13–56, Géométrie complexe et systèmes dynamiques (Orsay, 1995). MR 1755436
19. Christian Bonatti, Lorenzo J. Díaz, and Marcelo Viana, *Dynamics beyond uniform hyperbolicity*, Encyclopaedia of Mathematical Sciences, vol. 102, Springer-Verlag, Berlin, 2005, A global geometric and probabilistic perspective, Mathematical Physics, III. MR 2105774
20. Christian Bonatti and Marcelo Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly contracting*, Israel J. Math. **115** (2000), 157–193. MR 1749677
21. M. Brin and Yu. Kifer, *Dynamics of Markov chains and stable manifolds for random diffeomorphisms*, Ergodic Theory Dynam. Systems **7** (1987), no. 3, 351–374. MR 912374
22. Henk Bruin, Stefano Luzzatto, and Sebastian Van Strien, *Decay of correlations in one-dimensional dynamics*, Ann. Sci. École Norm. Sup. (4) **36** (2003), no. 4, 621–646. MR 2013929
23. David Burguet, *Entropy of physical measures for C^∞ smooth systems*, arXiv:1812.04308 (2018).
24. K. Burns, V. Climenhaga, T. Fisher, and D. J. Thompson, *Unique equilibrium states for geodesic flows in nonpositive curvature*, Geom. Funct. Anal. **28** (2018), no. 5, 1209–1259. MR 3856792
25. Keith Burns, Dmitry Dolgopyat, Yakov Pesin, and Mark Pollicott, *Stable ergodicity for partially hyperbolic attractors with negative central exponents*, J. Mod. Dyn. **2** (2008), no. 1, 63–81. MR 2366230

26. Jianyu Chen, Fang Wang, and Hong-Kun Zhang, *Markov partition and thermodynamic formalism for hyperbolic systems with singularities*, arXiv:1709.00527 (2017).
27. N. Chernov, *Decay of correlations and dispersing billiards*, J. Statist. Phys. **94** (1999), no. 3-4, 513–556. MR 1675363
28. N. Chernov and H.-K. Zhang, *Billiards with polynomial mixing rates*, Nonlinearity **18** (2005), no. 4, 1527–1553. MR 2150341
29. Yong Moo Chung, *Large deviations on Markov towers*, Nonlinearity **24** (2011), no. 4, 1229–1252. MR 2776118
30. Vaughn Climenhaga, Dmitry Dolgopyat, and Yakov Pesin, *Non-stationary non-uniform hyperbolicity: SRB measures for dissipative maps*, Comm. Math. Phys. **346** (2016), no. 2, 553–602. MR 3535895
31. Vaughn Climenhaga, Todd Fisher, and Daniel J. Thompson, *Unique equilibrium states for Bonatti-Viana diffeomorphisms*, Nonlinearity **31** (2018), no. 6, 2532–2570. MR 3816730
32. ———, *Equilibrium states for Mañé diffeomorphisms*, Ergodic Theory Dynam. Systems **39** (2019), no. 9, 2433–2455. MR 3989124
33. Vaughn Climenhaga, Stefano Luzzatto, and Yakov Pesin, *The geometric approach for constructing Sinai-Ruelle-Bowen measures*, J. Stat. Phys. **166** (2017), no. 3-4, 467–493. MR 3607577
34. Vaughn Climenhaga and Yakov Pesin, *Building thermodynamics for non-uniformly hyperbolic maps*, Arnold Math. J. **3** (2017), no. 1, 37–82. MR 3646530
35. Van Cyr and Omri Sarig, *Spectral gap and transience for Ruelle operators on countable Markov shifts*, Comm. Math. Phys. **292** (2009), no. 3, 637–666. MR 2551790
36. Mark F. Demers, *Functional norms for Young towers*, Ergodic Theory Dynam. Systems **30** (2010), no. 5, 1371–1398. MR 2718898
37. K. Díaz-Ordaz, M. P. Holland, and S. Luzzatto, *Statistical properties of one-dimensional maps with critical points and singularities*, Stoch. Dyn. **6** (2006), no. 4, 423–458. MR 2285510
38. Sébastien Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math. **139** (2004), 29–65. MR 2041223
39. ———, *Decay of correlations for nonuniformly expanding systems*, Bull. Soc. Math. France **134** (2006), no. 1, 1–31. MR 2233699
40. Chinmaya Gupta, Mark Holland, and Matthew Nicol, *Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps*, Ergodic Theory Dynam. Systems **31** (2011), no. 5, 1363–1390. MR 2832250
41. N. T. A. Haydn and Y. Psiloyenis, *Return times distribution for Markov towers with decay of correlations*, Nonlinearity **27** (2014), no. 6, 1323–1349. MR 3215837
42. Michihiro Hirayama, *Periodic probability measures are dense in the set of invariant measures*, Discrete Contin. Dyn. Syst. **9** (2003), no. 5, 1185–1192. MR 1974422
43. Morris W. Hirsch, *Differential topology*, Graduate Texts in Mathematics, vol. 33, Springer-Verlag, New York, 1994, Corrected reprint of the 1976 original. MR 1336822
44. Mark Holland, *Slowly mixing systems and intermittency maps*, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 133–159. MR 2122916
45. Mark Holland, Matthew Nicol, and Andrei Török, *Extreme value theory for non-uniformly expanding dynamical systems*, Trans. Amer. Math. Soc. **364** (2012), no. 2, 661–688. MR 2846347
46. A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math. (1980), no. 51, 137–173. MR 573822
47. F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula*, Ann. of Math. (2) **122** (1985), no. 3, 509–539. MR 819556

48. Yuri Lima and Carlos Matheus, *Symbolic dynamics for non-uniformly hyperbolic surface maps with discontinuities*, Ann. Sci. Éc. Norm. Supér. (4) **51** (2018), no. 1, 1–38. MR 3764037
49. Yuri Lima and Omri Sarig, *Symbolic dynamics for three-dimensional flows with positive topological entropy*, J. Eur. Math. Soc. (JEMS) **21** (2019), no. 1, 199–256. MR 3880208
50. Roberto Markarian, *Billiards with polynomial decay of correlations*, Ergodic Theory Dynam. Systems **24** (2004), no. 1, 177–197. MR 2041267
51. Véronique Maume-Deschamps, *Projective metrics and mixing properties on towers*, Trans. Amer. Math. Soc. **353** (2001), no. 8, 3371–3389. MR 1828610
52. Ian Melbourne and Matthew Nicol, *Almost sure invariance principle for nonuniformly hyperbolic systems*, Comm. Math. Phys. **260** (2005), no. 1, 131–146. MR 2175992
53. ———, *Large deviations for nonuniformly hyperbolic systems*, Trans. Amer. Math. Soc. **360** (2008), no. 12, 6661–6676. MR 2434305
54. Ja. B. Pesin, *Families of invariant manifolds that correspond to nonzero characteristic exponents*, Izv. Akad. Nauk SSSR Ser. Mat. **40** (1976), no. 6, 1332–1379, 1440. MR 0458490
55. Y. Pesin, S. Senti, and K. Zhang, *Thermodynamics of towers of hyperbolic type*, Trans. Amer. Math. Soc. **368** (2016), no. 12, 8519–8552. MR 3551580
56. Ya. B. Pesin, S. Senti, and K. Zhang, *Lifting measures to inducing schemes*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 553–574. MR 2408392
57. Ya. B. Pesin and Ya. G. Sinaĭ, *Gibbs measures for partially hyperbolic attractors*, Ergodic Theory Dynam. Systems **2** (1982), no. 3-4, 417–438 (1983). MR 721733
58. Sergei Yu. Pilyugin, *Shadowing in dynamical systems*, Lecture Notes in Mathematics, vol. 1706, Springer-Verlag, Berlin, 1999. MR 1727170
59. Vilton Pinheiro, *Expanding measures*, Ann. Inst. H. Poincaré Anal. Non Linéaire **28** (2011), no. 6, 889–939. MR 2859932
60. Vilton Pinheiro, *Lift and synchronization*, work in progress, 2019.
61. Luc Rey-Bellet and Lai-Sang Young, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 587–612. MR 2408394
62. Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653. MR 1946554
63. Omri M. Sarig, *Symbolic dynamics for surface diffeomorphisms with positive entropy*, J. Amer. Math. Soc. **26** (2013), no. 2, 341–426. MR 3011417
64. Farruh Shahidi and Agnieszka Zelerowicz, *Thermodynamics via inducing*, J. Stat. Phys. **175** (2019), no. 2, 351–383. MR 3968859
65. Michael Shub, *Stabilité globale des systèmes dynamiques*, Astérisque, vol. 56, Société Mathématique de France, Paris, 1978, With an English preface and summary. MR 513592
66. Masato Tsujii, *Regular points for ergodic Sinaĭ measures*, Trans. Amer. Math. Soc. **328** (1991), no. 2, 747–766. MR 1072103
67. Marcelo Viana, *Dynamics: a probabilistic and geometric perspective*, Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998), no. Extra Vol. I, 1998, pp. 557–578. MR 1648047
68. Qiudong Wang and Lai-Sang Young, *Toward a theory of rank one attractors*, Ann. of Math. (2) **167** (2008), no. 2, 349–480. MR 2415378
69. Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR 1637655
70. ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR 1750438
71. ———, *What are SRB measures, and which dynamical systems have them?*, vol. 108, 2002, Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays, pp. 733–754. MR 1933431

72. Roland Zweimüller, *Invariant measures for general(ized) induced transformations*, Proc. Amer. Math. Soc. **133** (2005), no. 8, 2283–2295. MR 2138871

Department of Mathematics, University of Houston, Houston, Texas 77204, USA

Email address: climenha@math.uh.edu

Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, Trieste, Italy

Email address: luzzatto@ictp.it

Department of Mathematics, McAllister Building, Pennsylvania State University, University Park, PA 16802, USA

Email address: pesin@math.psu.edu