SHADOWING AND STRUCTURAL STABILITY OF ANOSOV DIFFEOMORPHISMS

DINO ZAVATTINI

Abstract. The shadowing property of hyperbolic dynamical systems is vital to understanding their behavior. Unfortunately, many proofs of this result use techniques which do not intuitively demonstrate the link between hyperbolicity and shadowing. In order to address this issue, I offer a direct approach to the proof. Then, in order to highlight its importance, I use the shadowing lemma to prove the structural stability of Anosov diffeomorphisms.

Contents

1. Introduction

We define a *dynamical system* to be a homeomorphism $f : X \to X$ of a compact space X. Of interest to us is the differentiable case where f is a diffeomorphism of a smooth manifold. We would like to use the smooth structure of X to describe the dynamical behavior of f. In particular, we will consider smooth dynamical systems with a splitting of the tangent bundle into sub-bundles along which f contracts and expands. This property is called hyperbolicity. When hyperbolicity holds on the whole space X we say that f is an Anosov Diffeomorphism.

One of the interesting features of hyperbolic systems is how "stable" they are. The goal of this paper is to introduce two important results that illustrate this stability. First we prove the shadowing lemma which shows that the orbit structure of an Anosov diffeomorphism is resilient to compounded error over time. We will then use this lemma to prove that Anosov diffeomorphisms are “structurally stable”, meaning their dynamical behavior is preserved under C^1-perturbation.

2. Stability

Our interest in dynamical systems lies in their effects under iteration, their ‘dynamical behavior’. In this section we will discuss what it means to preserve this behavior, and as a consequence what it means for a dynamical system to be ‘structurally stable’.

2.1. Topological Conjugacy.

The focus on the iterative behavior of f motivates the following rephrasing of the dynamical system definition: A *dynamical system* is a compact $\mathbb{Z}$-space. The $\mathbb{Z}$-action associated to f is given by $n \cdot x = f^n(x)$ for any $n \in \mathbb{Z}$ and $x \in X$ and conversely, f is given by $f(x) = 1 \cdot x$.

This alternate definition is useful because it clearly indicates what the proper notion of equivalence should be for dynamical systems:

Definition 2.1. We say dynamical systems $f : X \to X$ and $g : Y \to Y$ are *topologically conjugate* if there exists a homeomorphism $h : X \to Y$ such that $hfh^{-1} = g$.

This homeomorphism can equivalently be thought of as an isomorphism of $\mathbb{Z}$-spaces since it preserves the action: $h(f^n(x)) = g^n(h(x))$ for any $n \in \mathbb{Z}$.

To illustrate this notion of equivalence, we consider the simplest useful example—homeomorphisms of $[0, 1]$. In particular, we restrict our attention to orientation preserving homeomorphisms of $[0, 1]$.

Lemma 2.2. *Let f and g be orientation preserving homeomorphisms from $[0, 1]$ to itself. If f and g have no fixed points on $(0, 1)$ then they are topologically conjugate.*

Moreover, the conjugating homeomorphism h is orientation preserving if and only if $f - \mathrm{id}$ and $g - \mathrm{id}$ have the same sign on $(0, 1)$.[1]

Proof. First we construct h. We will assume $f - \mathrm{id}$ and $g - \mathrm{id}$ are positive for now. Fix any point $p \in (0, 1)$. The orbits of p with respect to f and g are strictly increasing $\mathbb{Z}$-indexed sequences and therefore partition $[0, 1]$ into countably many intervals. Let

$$I_n = [f^n(p), f^{n+1}(p)] \qquad J_n = [g^n(p), g^{n+1}(p)]$$

for all $n \in \mathbb{Z}$. We now fix some orientation preserving homeomorphism $h_0 : I_0 = [p, f(p)] \to J_0 = [p, g(p)]$. Since $I_n = f^n(I_0)$ and $J_n = g^n(J_0)$ we can pushforward h_0 along the following diagram

$$\begin{array}{ccc} I_0 & \xrightarrow{h_0} & J_0 \\ {\scriptstyle f^n}\downarrow & & \downarrow{\scriptstyle g^n} \\ I_n & \dashrightarrow & J_n \end{array}$$

to get a homeomorphism $h_n = g^n \circ h_0 \circ f^{-n} : I_n \to J_n$. Since h_0 is orientation preserving it maps $p \mapsto p$ and $f(p) \mapsto g(p)$. Therefore

$$h_n(f^{n+1}(p)) = g^n \circ h_0 \circ f(p) = g^{n+1}(p)$$

---

[1] Wen[1] theorem 1.8, page 11.

and

$$h_{n+1}(f^{n+1}(p)) = g^{n+1} \circ h_0(p) = g^{n+1}(p).$$

In other words, the h_n agree on the boundaries of their domains and so we can paste them together into a map $h : (0,1) \to (0,1)$. This map must be strictly increasing and surjective since all h_n are. Thus h extends to an orientation preserving homeomorphism from $[0,1]$ to itself.

Since f, g, and h all fix 0 and 1, conjugacy holds on the boundary of the unit interval. Any x in the interior is contained in some I_n in which case $f(x) \in I_{n+1}$ and so

$$hf(x) = h_{n+1}f(x) = g^{n+1} \circ h_0 \circ f^n(x) = gh_n(x) = gh(x).$$

Thus h is a topological conjugacy of f and g.

Now suppose more generally that $f - \mathrm{id}$ and $g - \mathrm{id}$ are not both positive. We repeat the above construction but account for orientations of intervals as necessary. For example, if $f - \mathrm{id} \leq 0$ but $g - \mathrm{id} \geq 0$ our I_0 and J_0 become $[f(p), p]$ and $[p, g(p)]$. This requires us to choose an orientation reversing h_0 so that $p \mapsto p$ and $f(p) \mapsto g(p)$. If both $f - \mathrm{id}$ and $g - \mathrm{id}$ are negative h_0 returns to being orientation preserving.

Simply put, if the signs are different then we need an orientation reversing h_0 to make sure we can paste the h_n together. This orientation reversing h_0 makes h itself orientation reversing

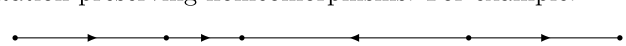
We have constructed an h which is orientation preserving if and only if $f - \mathrm{id}$ and $g - \mathrm{id}$ have the same sign. We now want to show that this holds for any conjugating map h.

Suppose first that $f - \mathrm{id} \geq 0$. Consider any point $x \in (0,1)$. The sequence $x, f(x), f^2(x), \ldots$ is strictly increasing because $f - \mathrm{id}$ is positive on $(0,1)$. The limit of this sequence must be a fixed point by continuity and therefore the only option is 1. Similarly, if $f - \mathrm{id} \leq 0$ all points in the interior of the interval approach 0. Orientation preserving and reversing homeomorphisms of $[0,1]$ can be characterized by whether they preserve or reverse the endpoints. Conjugacy preserves orbits and thus by continuity preserves limit points of orbits. Therefore if $f^n|_{(0,1)}$ and $g^n|_{(0,1)}$ pointwise approach $p, q \in \{0,1\}$, any conjugacy h of the two must associate p and q. We have h orientation preserving if and only if $p = q$ if and only if $f - \mathrm{id}$ and $g - \mathrm{id}$ have the same sign. □

Visually, if $f - \mathrm{id}$ is positive on $(0,1)$ we can think of this dynamical system using the following diagram:

In this case all the points in the interior of the interval move monotonically towards the limit 1. If $f - \mathrm{id}$ is negative we simply reverse the direction of the arrow.

Following lemma 2.2, it is natural to ask what happens when we string together multiple orientation preserving homeomorphisms. For example:



where the dots indicate fixed points. The answer turns out to be a fairly quick consequence of lemma 2.2. We first more precisely define the graphs we drew above:

Definition 2.3. Suppose $f : [0,1] \to [0,1]$ is an orientation preserving homeomorphism with finitely many fixed points. We construct the *graph representation* of f as follows:

Take $\text{Fix}(f) = \{x_0 = 0, x_1, \ldots, x_n = 1\}$ to be the vertex set. If $f - \text{id}$ is positive on the interval (x_k, x_{k+1}) add an edge from x_k to x_{k+1}. If $f - \text{id}$ is negative on this interval, add and edge from x_{k+1} to x_k.

We now have the following characterization of topological conjugacy:

Theorem 2.4. *Suppose f and g are orientation preserving homeomorphisms of $[0,1]$ each with finitely many fixed points. Then f and g are topologically conjugate if and only if their graph representations are isomorphic as directed graphs.*

Proof. Any conjugacy h of f and g restricts to an order preserving or reversing bijection of $\text{Fix}(f)$ and $\text{Fix}(g)$. Therefore any conjugacy h must also define an order preserving or reversing bijection of the intervals bounded by the fixed points. This gives us an isomorphism of the undirected graphs with the vertices as the fixed points and edges as the intervals between. Initial and terminal vertices can then be determined by taking 'negative' and 'positive' limits $\lim_{n\to-\infty} f^n(x)$ and $\lim_{n\to+\infty} f^n(x)$ of points x on the edge in question. Since h conjugates these limits in f to the analogous limits in g, it must also preserve the orientations of the edges.

Conversely, suppose the graphs representing f and g are isomorphic as directed graphs. For each pair of associated edges choose a map which conjugates f and g on those intervals. Because the conjugating map is orientation preserving if and only if the edges point in the same direction, the piecewise conjugacies will paste together into a global one. □

Note that by replacing fixed points with intervals we can strengthen this result to characterize topological conjugacies for all orientation preserving f and g on $[0,1]$ where $\text{Fix}(f)$ and $\text{Fix}(g)$ have finitely many components. While this theorem does not classify all order preserving homeomorphisms of $[0,1]$, it hopefully provides some geometric insight into the meaning of topological conjugacy.

As mentioned previously, this paper's focus is on differentiable dynamical systems. To that end, one might expect that we would need to upgrade the regularity of our conjugating maps. However, even when we work with differentiable structures, *topological* conjugacy is usually sufficient to preserve dynamical behavior. Equivariance allows us to associate the group theoretic aspects of the dynamical systems— orbits, fixed points, stabilizers, and so on— while continuity preserves the topological data of these objects and their limiting behavior. Moreover, differentiable conjugacy is very strong and in practice much more difficult to obtain.

To illustrate this, take $f, g, h : M \to M$ to be C^1-diffeomorphisms with $hf = gh$. If f has a fixed point $p \in M$ and g has a corresponding fixed point $h(p) = q$ then $D_p h D_p f D_p h^{-1} = D_q g$. In other words, the Jacobian matrices $f'(p)$ and $g'(q)$ must be similar. For example, suppose $f, g : [0,1] \to [0,1]$ are orientation preserving diffeomorphisms with $f - \text{id} > 0$ and $g - \text{id} > 0$ on $(0,1)$. These systems seem to have the same dynamical behavior and they are topologically conjugate, but if $f'(0) \neq g'(0)$ or $f'(1) \neq g'(1)$ then they cannot even be C^1-conjugate. Since we will consider only topological conjugacy in this paper, 'conjugacy' will indicate topological conjugacy.

2.2. **Structural Stability.** In order to determine whether a given dynamical system is 'stable', we need to consider perturbations of it in the C^k-topology. We begin with the C^k topology for maps of Euclidean spaces.

Definition 2.5. Suppose $U \subseteq \mathbb{R}^n$ is open with $K = \overline{U}$ compact. For any C^k map $f : K \to \mathbb{R}^m$ we define the C^k *norm* by

$$\|f\|_{C^k} = \sup_{x \in K}\{|f(x)|, \|D_x f\|, \ldots, \|D_x^k f\|\}$$

where $\|.\|$ indicates the operator norm.

This norm defines the C^k *topology* on $C^k(K, \mathbb{R}^m)$

This definition generalizes to manifolds.

Definition 2.6. We say a chart (U, φ) is *admissible* if there exists (U', φ') such that $\overline{U} \subseteq U'$ and $\varphi = \varphi'|_U$.

Suppose M and N are compact manifolds. Using compactness we can choose finite covers $\{(U_i, \varphi_i)\}$ and $\{(V_j, \psi_j)\}$ of M and N by admissible charts. For any C^k maps $f, g : M \to N$ define

$$d(f, g) = \sum_{i,j} \|\psi_j f \varphi_i^{-1} - \psi_j g \varphi_i^{-1}\|_{C^k}.$$

This metric depends on the covers we choose, but the resulting topology does not since the transition functions are diffeomorphisms. We define this to be the C^k *topology* on $C^k(M, N)$.

Note that to take C^k norms in the definition of $d(f, g)$ we needed to be working over a compact domain, which the $\varphi_i(U_i)$ are not. Extending to $\overline{\varphi_i(U_i)}$ solves this problem which is why we require admissible charts.

For $k \geq 1$ the C^k topology on $C^k(M, M)$ induces the subspace topology on $\mathrm{Diff}^k(M)$. Similarly, the C^0 topology on $C^0(M, M)$ induces the subspace topology on $\mathrm{Homeo}(M)$.

The C^0 topology is equivalently the uniform topology given by any Riemannian metric on N. More generally we can define a topology on $C^0(X, Y)$ for any compact X and metric space Y as the uniform topology. The metric on $\mathrm{Homeo}(X)$ for a metric space X is often defined as

$$d_{\mathrm{Homeo}}(f, g) = \sup_{x \in X}\{d(f(x), g(x)), d(f^{-1}(x), g^{-1}(x))\}.$$

The resulting topology is equivalent to the subspace topology induced by $C^0(X, X)$, but this metric has the added benefit of being complete.

The C^k topology allows us to perturb maps which leads to the notion of stability:

Definition 2.7. Suppose $f : M \to M$ is a C^k-diffeomorphism for some $k \geq 1$. We say that f is a C^k *structurally stable* dynamical system if there exists a neighborhood $\mathcal{U} \subseteq \mathrm{Diff}^k(M)$ of f such that every $g \in \mathcal{U}$ is topologically conjugate to f.

Intuitively, f is C^k structurally stable if small C^k perturbations do not change its dynamical structure. For large k these perturbations can become hard to visualize. We will only discuss C^1 structural stability in this paper. The unit interval dynamical systems are again a useful example:

Theorem 2.8. *Suppose $f : [0,1] \to [0,1]$ is an orientation preserving C^1 diffeomorphism without fixed points in $(0,1)$. Then f is C^1 structurally stable if and only if $f'(0) \neq 1$ and $f'(1) \neq 1$.*[2]

Proof. First suppose $f'(0) = 1$. For any $\varepsilon > 0$ we can choose a g which is within the C^1 topology ε-ball of f but agrees with id on some interval $[0, \alpha)$. Such a g would not be topologically conjugate to f. An analogous argument works for $f'(1) = 1$.

For the converse direction suppose $f'(0) \neq 1 \neq f'(1)$. For simplicity we will also assume $f - \text{id} > 0$ on the interior of the interval but the other case is analogous. We must have $f'(0) > 1 > f'(1)$. For some $\delta > 0$ and $\alpha > 0$ we can require $f'(x) > 1 + 2\alpha$ for all $x \in [0, \delta)$ and $f'(y) < 1 - 2\alpha$ for all $y \in (1 - \delta, 0]$. By choosing a small enough C^1 neighborhood of f, we can require that all g in that neighborhood satisfy $f'(x) > 1+\alpha$ and $f'(y) < 1-\alpha$ for the same x and y as above. Call this neighborhood U. We can now choose a C^0 (and thus C^1) neighborhood V of f such that for any $g \in V$, $g(x) > x$ for all $x \in [\delta, 1 - \delta]$. By lemma 2.2 it suffices to prove that any $g \in U \cap V$ has no fixed points on $(0, \delta) \cup (1 - \delta, 1)$. Since $g'(x) > 1$ for all $x \in [0, \delta]$, $g(x) > x$ for all $x \in (0, \delta]$ otherwise we would have $(g(x) - g(0))/(x - 0) = 1$ for some x on that interval which would contradict the mean value theorem. A similar argument shows that $g(x) > x$ for $x \in [1 - \delta, 1)$ and so we are done. □

For any $m \geq n$ the C^m topology is finer than the C^n topology. It follows that C^n perturbations are larger and thus that C^n stability is stronger. The strongest useful notion of structural stability is C^1 structural stability. The reason we have not defined C^0 structural stability is because it is too strong to be of any use. This is demonstrated by the following:

Example 2.9. Consider any dynamical system $f : [0,1] \to [0,1]$ with at least one isolated fixed point p. Any ε-neighborhood of f in the uniform topology will contain a homeomorphism g which is fixed on a neighborhood of p.

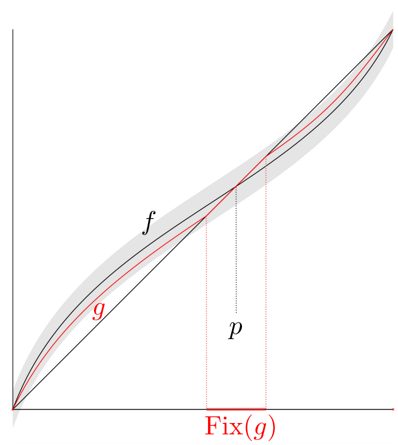


FIGURE 1.

In fact we can require g to be fixed on a neighborhood of each of its fixed points. Since g has no isolated fixed points, it cannot be conjugate to f. Conversely, if f

[2]Wen[1] theorem 1.9, page 13.

had only fixed points which were not isolated, we could use this same technique to isolate them.

This argument generalizes. Suppose X is a compact metric space and $f : X \to X$ has an isolated fixed point $p \in X$. Fix $\varepsilon > 0$. There is some neighborhood $U \subseteq X$ of p on f is within $\varepsilon/2$ of p. If we also require $U \subseteq B(p, \varepsilon/2)$ it then follows that $d(f, \mathrm{id}) < \varepsilon$ on U. If X is a manifold we can then extend $\mathrm{id}|_U$ to a function over X which stays within ε of f.

As this example demonstrates, structural stability under the C^0 topology is too stringent of a criterion to be useful. However, we still would like a notion of (partial) preservation of dynamical behavior under topological perturbation. The case of homeomorphisms on the unit interval in the above example hints at how to do this: We could contract and expand fixed points at will, but not fully do away with them. While we have lost the bijective correspondence between fixed points which topological conjugacy would give us, if $f : [0, 1] \to [0, 1]$ has non-isolated fixed points, then $\mathrm{Fix}(g)$ retracts onto $\mathrm{Fix}(f)$ for g sufficiently close to f. This indicates that it might be useful to replace the homeomorphism h in our definition of conjugacy with a surjection.

2.3. **Topological Semi-conjugacy.** Occasionally two dynamical systems may have some similar characteristics without being fully conjugate. For such purposes we define a weaker version of conjugacy:

Definition 2.10. We say that a dynamical system $f : X \to X$ is *topologically semi-conjugate* to $g : Y \to Y$ if there is a surjective map $h : X \to Y$ such that $hf = gh$. If such an h exists we say that g is a *factor* of f.

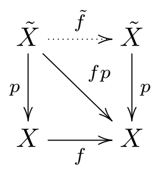
The map h can be thought of as a surjective $\mathbb{Z}$-equivariant map from the $\mathbb{Z}$-space X to the $\mathbb{Z}$-space Y.

The first thing to note is that this relation is not symmetric and therefore should not be thought of as an equivalence. Instead, we think of semi-conjugacy as conjugacy with a loss of information from g to f. We introduce a sequence of examples to illustrate what this means:

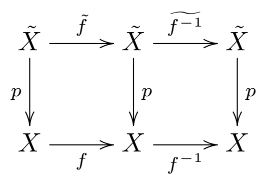
Example 2.11. Consider a map with a single fixed point $g : \{p\} \to \{p\}$. For any $f : X \to X$, the constant map $h : X \to Y = \{p\}$ gives us $hf(x) = p = g(p) = gh(x)$. In other words the dynamical system on a point is a factor of every other dynamical system. This is intuitively reasonable in the sense that any $f : X \to X$ fixes the entire space X as an invariant set. Collapsing X to a point erases everything interesting about the map f, but maintains this overall 'fixed' structure. In other words, g is a simplified version of f. The loss in detail from f to g corresponds to the failure of h to be injective.

Example 2.12. A more interesting example of this effect is covering spaces. Suppose $p : \tilde{X} \to X$ is the universal cover of X and $f : X \to X$ is a homeomorphism. If we want to be strict about our definition of dynamical system, we can require that X and $\tilde{X}$ are compact (for example, taking $X = \mathbb{R}P^n$). Fix $x_0 \in X$ and $\tilde{x}_0 \in p^{-1}(x_0)$. Let $x_1 = f(x_0)$ and choose some $\tilde{x}_1 \in p^{-1}(x_1)$. We can lift $fp : \tilde{X} \to X$ to get

$\tilde{f} : \tilde{X} \to \tilde{X}$ making the diagram

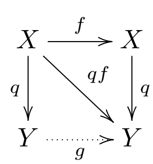


commute and satisfying $\tilde{f}(\tilde{x}_0) = \tilde{x}_1$. We can now lift f^{-1} in the same manner, requiring that $\widetilde{f^{-1}}(\tilde{x}_1) = \tilde{x}_0$. Lifting is unique up to basepoint. In particular, since basepoints have been properly accounted for, the composite $f \circ f^{-1} = \mathrm{id}$ in the diagram



must lift uniquely to the identity on $\tilde{X}$ and so $\tilde{f} \circ \widetilde{f^{-1}} = \mathrm{id}$. The same logic shows that the other composition is also the identity and that $\tilde{f}$ is a homeomorphism on $\tilde{X}$. We thus have that $fp = p\tilde{f}$ where both f and $\tilde{f}$ are homeomorphisms. In other words, $\tilde{f}$ is semi-conjugate to f. Just as in the fixed point example, the factor of the semi-conjugacy encodes the target homeomorphism along with some extra information. In this case that extra information can be thought of as the choice of basepoint $\tilde{x}_0$. Along with the fixed point example, covering spaces emphasize that semi-conjugacy is a way of collapsing one dynamical system into another.

Example 2.13. We will now generalize both of the previous examples. Suppose $f : X \to X$ is a homeomorphism on a compact space X. Consider any identification $\sim$ on X such that $x \sim x'$ if and only if $f(x) \sim f(x')$. Let $q : X \to Y := X/\sim$ be the quotient map. The map $qf : X \to Y$ is constant on equivalence classes and thus gives us a map $g : Y \to Y$ such that



commutes. Since qf^{-1} is constant on equivalence classes it too descends to a map on Y. This map will be g^{-1} by the same uniqueness argument we made in the covering space examples. It follows that g is a homeomorphism. We thus have a semi-conjugacy from f to g defined by q.

Suppose $\Lambda \subseteq X$ is an invariant set of f. Then the equivalence relation corresponding to the quotient map $X \to X/\Lambda$ is preserved by f and f^{-1}. Thus the induced homeomorphism $g : X/\Lambda \to X/\Lambda$ is a factor of f. In other words, semi-conjugacy allows us to contract invariant sets into fixed points.

At the end of the previous subsection we motivated semi-conjugacy by implying that we wanted a surjection between the fixed points of the dynamical systems in question. As example 2.11 illustrates, this is not true in general for semi-conjugacy since invariant subsets may be collapsed to fixed points. We have the following compromise:

Proposition 2.14. *Suppose $h : X \to Y$ defines a semi-conjugacy from $f : X \to X$ to $g : Y \to Y$.*

(1) The semi-conjugacy h induces a surjection from invariant sets of f to invariant subsets of g.
(2) $h(\mathrm{Fix}(f)) \subseteq \mathrm{Fix}(g)$.

Proof. Suppose $\Lambda \subseteq X$ is an invariant set of f. Any $\mu = h(\lambda) \in h(\Lambda)$ satisfies $g(\mu) = gh(\lambda) = hf(\lambda) = h(f(\lambda)) \in h(\Lambda)$ since $f(\lambda) \in \Lambda$. Thus $h(\Lambda)$ is an invariant set of g.

Now suppose $M \subseteq Y$ is an invariant set of g. For any $\lambda \in h^{-1}(M)$ we have $hf(\lambda) = gh(\lambda) \in M$ since $h(\lambda) \in M$ which is g-invariant. It follows that $f(\lambda) \in h^{-1}(M)$. Thus $h^{-1}(M)$ is an invariant set of f. Now note that since h is surjective, $h(h^{-1}(M)) = M$ and so we get the desired surjection of invariant sets.

We now move on to (2). Suppose p is a fixed point of f and $q = h(f)$. Then $g(q) = gh(p) = hf(p) = h(p) = q$ and so q is a fixed point of g. Thus h restricts to a map from $\mathrm{Fix}(f)$ to $\mathrm{Fix}(g)$. □

Example 2.15. While $h : \mathrm{Fix}(f) \to \mathrm{Fix}(g)$ is not a surjection in general, it does become one when we consider orientation preserving homeomorphisms with finitely many (ie. isolated) fixed points on the unit interval. To demonstrate, first note that the minimal invariant subsets of f are the fixed points and the intervals between them. Moreover, each preimage under h of a fixed point of g is a closed invariant subset of f. Any f-invariant set without fixed points will be a union of open intervals between fixed points of f. Therefore any closed invariant set of f must contain fixed points. In particular, the preimage under h of any fixed point of g must contain a fixed point of f.

2.4. Topological Semi-Stability.

We can now offer a useful replacement to the notion of "C^0 structural stability".

Definition 2.16. Suppose $f : X \to X$ is a homeomorphism of a compact metric space. We say that f is a *topologically semi-stable* dynamical system if there exists a neighborhood $U \subseteq \mathrm{Homeo}(X)$ of f such that f is a factor of every $g \in U$. That is to say, for every $g \in U$ there exists a surjection $h : X \to X$ such that $fh = hg$.

Semi-stability is often referred to as topological stability since there is little risk of confusion with the unused notion of C^0 structural stability. We will stick to "semi-stability" for consistency with "semi-conjugacy".

3. Hyperbolicity

3.1. Hyperbolic Sets.

For the remainder of the paper we will take $f : M \to M$ to be a C^1-diffeomorphism of a closed smooth Riemannian manifold unless stated otherwise.

Definition 3.1. We say an f-invariant subset $\Lambda \subseteq M$ is *hyperbolic* if there exists an f-invariant splitting of the tangent space $T_xM = E^s(x) \oplus E^u(x)$ for each $x \in \Lambda$ which satisfies the following property:

There exist constants $C > 0$ and $0 < \lambda < 1$ such that for any $x \in \Lambda$ and all $n \in \mathbb{N}$,

$$\begin{aligned} |Df^n(v)| &\leq C\lambda^n |v| && \text{for all } v \in E^s(x) \subseteq T_x M \\ |Df^{-n}(w)| &\leq C\lambda^n |w| && \text{for all } w \in E^u(x) \subseteq T_x M \end{aligned}$$

where $Df : TM \to TM$ is the basepoint free derivative of f. The subbundles E^s and E^u are respectively referred to as *stable* and *unstable*.

Note that any hyperbolic set Λ of a diffeomorphism $f : M \to M$ is also a hyperbolic set of f^{-1}. The splitting for f^{-1} is the same but with the stable and unstable subbundles swapped. Note also that we did not specify the regularity of the splitting $E^s \oplus E^u$ in the definition of hyperbolicity. This is justified by the following result.

Proposition 3.2. *If $f : M \to M$ has a hyperbolic set $\Lambda \subseteq M$ then the splitting $E^s \oplus E^u$ on Λ is continuous. It follows that the dimensions of E^s and E^u are locally constant.*[3]

The simplest example of a hyperbolic set is a hyperbolic fixed point.

Example 3.3. Suppose $f : M \to M$ has a hyperbolic fixed point p. The hyperbolicity of this fixed point is wholly determined by the linear automorphism $D_p f : T_p M \to T_p M$. For convenience let $T = D_p f$ and $V = T_p M$. We have an invariant splitting $V = E^s \oplus E^u$ and constants $C > 0$ and $0 < \lambda < 1$ such that

$$\begin{aligned} |T^n(v)| &\leq C\lambda^n |v| && \text{for all } v \in E^s \\ |T^{-n}(w)| &\leq C\lambda^n |w| && \text{for all } w \in E^u \end{aligned}$$

for all $n \in \mathbb{N}$. A linear automorphism which satisfies this property is called a *hyperbolic linear automorphism.*

We can think of hyperbolic sets as hyperbolic linear automorphisms with a moving basepoint. It is therefore intuitively valuable to characterize these linear automorphisms.

Proposition 3.4. *A linear automorphism $T : V \to V$ is hyperbolic if and only if it has no eigenvalue of absolute value 1.*

Proof. It is clear that if T has an eigenvalue of absolute value 1 then it cannot be hyperbolic. For the converse direction we claim that we can take E^s to be the direct sum of the generalized eigenspaces with eigenvalues of modulus less than 1 and E^u to be the direct sum of the generalized eigenspaces with eigenvalues of modulus greater than 1.

Since all norms on a finite dimensional vector space are equivalent, we note that the hyperbolicity of T is invariant under change of basis. It therefore suffices to assume T is a $d \times d$ Jordan block with an eigenvalue $\lambda \in \mathbb{C}$ of modulus not equal to 1:

$$T = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} = \lambda I + J$$

[3] Wen[1] theorem 4.3, page 80.

where J is the matrix of superdiagonal 1's. Since J is a nilpotent with $J^d = 0$ we have

$$T^n = \lambda^n I + \binom{n}{1}\lambda^{n-1}J + \cdots + \binom{n}{d-1}\lambda^{n-(d-1)}J^{d-1}.$$

First we consider the case $|\lambda| < 1$. It follows from the above identity that

$$\|T^n\| \leq |\lambda|^n \cdot (1 + n|\lambda|^{-1}\|J\| + \cdots + n^{d-1}|\lambda|^{-(d-1)}\|J\|^{d-1}) = |\lambda|^n \cdot p(n)$$

where $p(n)$ is a polynomial in n. For any $\mu \in (|\lambda|, 1)$,

$$\|T^n\| \leq \mu^n \cdot \left(\frac{|\lambda|}{\mu}\right)^n \cdot p(n) \leq C\mu^n$$

for some positive constant C uniform over $n \in \mathbb{N}$ since $|\lambda|/\mu < 1$. It follows that any matrix with all eigenvalues of modulus less than 1 is contracting up to a constant C.

If we now take $|\lambda| > 1$ we immediately have that all eigenvalues of T^{-1} are $1/\lambda$ and in particular have modulus less than 1. It follows from the treatment of the previous case that for any $\mu \in (1/|\lambda|, 1)$ there exists $C > 0$ such that

$$\|T^{-n}\| \leq C\mu^n$$

for all $n \in \mathbb{N}$. □

The underlying theme of the above proof is that diagonalizable matrices with no modulus 1 eigenvalues immediately satisfy hyperbolicity. That is to say, they exponentially dilate and contract along E^u and E^s without the need for a constant C. We would more generally like to be able to do away with the constant C for any hyperbolic dynamical system. Fortunately, since all norms on a finite dimensional vector space are equivalent, we can adjust the norm as desired:

Proposition 3.5. *Suppose $f : M \to M$ has a hyperbolic set Λ. There exist a norm $\|.\|$ on the tangent space and a constant $\tau \in (0, 1)$ such that*

$$\|Df^n(v)\| \leq \tau^n\|v\| \qquad \text{for all } v \in E^s(x) \subseteq T_xM$$
$$\|Df^{-n}(w)\| \leq \tau^n\|w\| \qquad \text{for all } w \in E^u(x) \subseteq T_xM$$

for all $n \in \mathbb{N}$.[4]

Proof. Suppose $|.|$ is the norm induced by the metric on M and $\lambda \in (0, 1)$, $C > 1$ are the constants which give us hyperbolicity on Λ. Let $n \in \mathbb{N}$ be such that $C\lambda^n < 1$ and define

$$\|v\| = |v| + |Df(v)| + |Df^2(v)| + \cdots + |Df^{n-1}|.$$

If $a = 1 + C\lambda + C\lambda^2 + \cdots + C\lambda^{n-1} \geq 1$ we have

$$\|v\| \leq a|v| \qquad \text{for all } v \in E^s$$
$$\|w\| \leq a|Df^{n-1}(w)| \qquad \text{for all } w \in E^u.$$

It follows that for any $v \in E^s$

$$\begin{aligned}\|Df(v)\| &= \|v\| - |v| + |Df^n(v)| \leq \|v\| - (1 - C\lambda^n)|v| \\ &\leq \|v\| - a^{-1}(1 - C\lambda)\|v\| = (1 - a^{-1}(1 - C\lambda^n))\|v\|\end{aligned}$$

[4]Wen[1] theorem 2.3, page 28.

since $1 - C\lambda^n > 1$ by assumption. Similarly, for any $w \in E^u$

$$\begin{aligned}\|Df^{-1}(w)\| &= \|w\| - |Df^{n-1}(w)| + |Df^{-1}(w)| \\ &\leq \|w\| - (1 - C\lambda^n)|Df^{n-1}(w)| \\ &\leq \|w\| - a^{-1}(1 - C\lambda)\|w\| = (1 - a^{-1}(1 - C\lambda^n))\|w\|.\end{aligned}$$

Since $a \geq 1$ and $C\lambda^n < 1$ we have that $1 - a^{-1}(1 - C\lambda^n) \in (0, 1)$. Therefore let $\tau = 1 - a^{-1}(1 - C\lambda^n)$ and we are done. □

This norm is called the *adapted norm* and the minimal τ is called the *skewness* of f with respect to the norm $\|.\|$. Since f is at least C^1 this norm varies continuously by basepoint. Moreover, because all norms on a finite dimensional vector space are equivalent and M is a compact manifold, it follows that $\|.\|$ induces a metric on the manifold comparable to the original one. In other words, though we require a Riemannian metric to state hyperbolicity, the property is independent of which metric we choose. In other words it is purely dependent on the differential topology of f and its manifold M.

We can make one further adjustment to the adapted norm by defining the norm of any vector v to be $\max(\|\pi_s(v)\|, \|\pi_u(v)\|)$ where π_s and π_u are projection onto E^s and E^u respectively. This new norm is called the *box adapted norm*. It has the same skewness as the adapted norm it was derived from and can be more convenient to work with.

3.2. **Anosov Diffeomorphisms.** Once we have defined hyperbolic sets, it seems obvious to question what happens when hyperbolicity extends over the whole manifold:

Definition 3.6. Suppose $f : M \to M$ is a C^1-diffeomorphism of a closed smooth Riemannian manifold. We say f is an *Anosov diffeomorphism* if the whole manifold M is a hyperbolic set.

Anosov diffeomorphisms can be hard to come by due to the complexity of the hyperbolicity criterion. For example, no Anosov diffeomorphism exists on S^2 because the hairy ball theorem does not allow the existence of a global splitting $TM = E^s \oplus E^u$. Fortunately, we have a class of Anosov diffeomorphisms that are easily constructed:

Example 3.7. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a hyperbolic linear automorphism with determinant ± 1 and suppose the matrix representation of T has integer entries. Since T has integer entries, it maps $\mathbb{Z}^2 \to \mathbb{Z}^2$. This means that the kernel of the composite $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2 = \mathbb{T}^2$ contains $\mathbb{Z}^2$ and thus that T descends to a smooth map $f : \mathbb{T}^2 \to \mathbb{T}^2$ on the torus. The 2×2-matrix inverse formula tells us that T^{-1} also has integer entries and so it too induces a map on the torus which is the inverse of f. Thus f is a diffeomorphism.

The derivative of f at any point can be identified with $T : \mathbb{R}^2 \to \mathbb{R}^2$. Since we have already required that T be hyperbolic, f must be Anosov. Diffeomorphisms induced in this way by hyperbolic linear isomorphisms are called *Anosov toral automorphisms*.

4. Shadowing and Expansiveness

Hyperbolicity describes functions which locally stretch and compress the manifold in a way which is globally consistent. As alluded to by the notation, the duality between stability and instability is the underlying intuition. We can talk about this duality between controlled and chaotic behavior in a more concrete topological sense by looking at approximations of orbits and how disruptive these approximations are to the orbit structure.

Definition 4.1. Let $f : X \to X$ be a homeomorphism on a compact metric space (X, d). If $\delta > 0$ δ-*pseudo orbit* is a $\mathbb{Z}$-indexed sequence $\{x_n\}_{n\in\mathbb{Z}} \subseteq X$ such that $d(f(x_n), x_{n+1}) \leq \delta$ for all $n \in \mathbb{Z}$.

Definition 4.2. We say that a pseudo orbit $\{x_n\}_{n\in\mathbb{Z}}$ is ε-*shadowed* by $y \in X$ if $d(x_n, f^n(y)) \leq \varepsilon$ for all $n \in \mathbb{Z}$.

It turns out that the question of which pseudo orbits are shadowed is closely related to hyperbolicity. Our first instinct to try to force shadowing of pseudo orbits might be to require f to be contracting. For example, consider a differentiable map $f : M \to M$ with a degenerate form of hyperbolicity where E^u is trivial. In other words, there exist $\lambda \in (0, 1)$ and $C > 0$ such that $|Df^n(v)| \leq C\lambda^n|v|$ for all $v \in TM$. Under the metric induced by the adapted norm from proposition 3.5 f becomes a contraction map with constant τ, the skewness. Now consider a positive δ-pseudo semi-orbit $x_0, x_1, x_2, \dots$ with $d(f(x_n), x_{n+1}) \leq \delta$. We have $d(f(x_0), x_1) \leq \delta$ and

$$\begin{aligned} d(f^n(x_0), x_n) &\leq d(f(f^{n-1}(x_0)), f(x_{n-1})) + d(f(x_{n-1}), x_n) \\ &\leq \tau d(f^{n-1}(x_0), x_{n-1}) + \delta. \end{aligned}$$

By induction we get

$$d(f^n(x_0), x_n) \leq \delta + \tau\delta + \cdots + \tau^{n-1}\delta \leq \frac{\delta}{1-\tau}.$$

It follows that for any positive semi-orbit $y, f(y), f^2(y), \dots$

$$d(f^n(y), x_n) \leq d(f^n(y), f^n(x_0)) + d(f^n(x_0), x_n) \leq d(y, x_0) + \frac{\delta}{1-\tau}.$$

In other words, if we require $\delta/(1-\tau) < \varepsilon$ we get that any positive δ-pseudo semi-orbit is ε-shadowed by any y sufficiently close to x_0.

The issues with this example map arise when we consider the other half of the orbit. If f is invertible, the stable contracting properties in the positive directions become unstable exponential expansion in the negative. Shadowing in the negative direction over whole pseudo orbits or even negative pseudo semi-orbits becomes impossible. Note that by a symmetrical argument, the properties of the unstable space E^u give us shadowing over negative pseudo semi-orbits. However, such maps are likewise unable to account for the positive direction.

Fortunately, if we require f to be an Anosov diffeomorphism of a compact manifold M, it is impossible for either E^u or E^s to be trivial. If E^u were trivial then for $C\lambda^n < 1$ the map f^n would be locally measure decreasing. As a result, the compactness of M would tell us that the finite measure of $\operatorname{im} f^n$ is less than that of M, contradicting surjectivity.

It is a surprising fact that when working with Anosov diffeomorphisms, shadowing is possible. In a sense, the positive shadowing and negative shadowing we would expect if it was fully stable or fully unstable are synthesized. Moreover, since the stable and unstable properties operate within complementary subspaces, the positive and negative semi-shadowing points we get from each intersect transversely to give us a unique solution. This is somewhat unexpected since the shadowing semi-orbits we found for contracting f were resistant to small perturbation. In other words, although we might expect some amount of topological control when approximating like this, the shadowing of Anosov diffeomorphisms is rigid:

Theorem 4.3. ***The Shadowing Lemma.*** *Suppose $f : M \to M$ is a diffeomorphism of a closed manifold and $\Lambda \subseteq M$ is a hyperbolic set. For every $\varepsilon > 0$ there exists $\delta > 0$ such that every δ-pseudo orbit is ε-shadowed.*

Moreover, if ε is sufficiently small then δ can be chosen so that any ε-shadowing orbit is unique. In other words, there exist $\varepsilon_0 > 0$ and $\delta_0 > 0$ dependent on f such that every δ_0-pseudo orbit is ε_0-shadowed by at most one orbit.

The shadowing lemma states that as δ approaches zero, δ-pseudo orbits uniformly approach actual orbits. The conditions with ε_0 and δ_0 are necessary because large ε make uniqueness of a shadowing impossible irrespective of how small δ is.

Before proving the shadowing lemma we require a basic extension result:

Lemma 4.4. *Suppose $f : U \to \mathbb{R}^n$ is a local diffeomorphism defined on a bounded neighborhood U of 0 such that $f(0) = 0$. Then for any $\varepsilon > 0$ there exists $\delta > 0$ and an extension $\bar{f} : \mathbb{R}^n \to \mathbb{R}^n$ such that:*

(1) $\|\bar{f} - D_0 f\|_{C^1} \leq \varepsilon$
(2) $\bar{f} = f$ *on* $B(0, \delta)$.[5]

Proof. For any $\delta > 0$ define $\varphi : \mathbb{R}^n \to [0, 1]$ to be a smooth function which is 1 on $B(0, \delta)$ and 0 outside of $B(0, 2\delta)$. Let

$$\bar{f} = \varphi \cdot f + (1 - \varphi) \cdot D_0 f$$

so that $\bar{f} - D_0 f = \varphi \cdot (f - D_0 f)$. By differentiability $\|\varphi \cdot (f - D_0 f)\|_{C^0} = o(\delta)$. In other words, as δ becomes small, the difference vanishes with respect to δ. For the derivative we use the product rule

$$\|D(\bar{f} - D_0 f)\| \leq \|(D\varphi) \cdot (f - D_0 f)\| + \|\varphi \cdot (Df - D_0 f)\|.$$

Since f is C^1 the second term becomes small for small δ. Since $\|D\varphi\| \leq C/\delta$ for small $\delta > 0$ and some constant C, and since we have already shown that $f - D_0 f$ is $o(\delta)$ with respect to δ, it follows that $\|(D\varphi) \cdot (f - D_0 f)\| = o(1)$ as $\delta \to 0$. Thus, $\|D(\bar{f} - D_0 f)\|$ approaches 0 when δ does. Since we have bounded both $\bar{f}$ and its derivative to $D_0 f$ for small δ we are done. □

Note that as a consequence of this lemma, if we require ε small enough we can force $\bar{f}$ to be invertible.

Proof. (Shadowing Lemma) We will prove the case where M is a 2-manifold with $\dim E^s = \dim E^u = 1$. The same method generalizes with virtually no changes to higher dimensions but it is harder to visualize and some of the details become less streamlined.

Though the proof is long, the main ideas are intuitive. We first lift to the tangent spaces $T_{x_n} M$ via the exponential maps $\exp_{x_n}$ at the points in our pseudo

[5] Katok and Hasselblatt[2] lemma 6.2.7, page 242.

orbit. This preserves the composition structure of f as well as the radial distance from the points x_n.

The second step is to prove that for any n we can find orbits which ε-shadow the finite segment $x_0, \ldots, x_n$ of the pseudo orbit. We do this by considering the generating points of these orbits in $T_{x_0}M$. In particular, we will show that the solution set to the ε-shadowing problem over $x_0, \ldots, x_n$ contains the image Γ_n of some path in $T_{x_0}M$.

In the third step we prove that Γ_n are actually the graphs of functions α_n which map the stable axis to the unstable axis in $T_{x_0}M$. Moreover, these functions are uniformly Lipschitz with constant $C < 1$.

Step four is to show that α_n converge uniformly to some α and that the graph Γ of α is the solution set of orbits which ε-shadow $\{x_n\}_{n\geq 0}$.

Finally, in our fifth step, we note that by the same argument applied in the negative direction, the solution set of orbits which ε-shadow $\{x_n\}_{n\leq 0}$ is the graph of some function β from the unstable axis to the stable axis in $T_{x_0}M$. The graphs of α and β must intersect, and because they are both Lipschitz this intersection is unique.

1. Preliminaries. Just like hyperbolicity, the shadowing lemma is topological. In other words, there is no loss of generality if we choose an equivalent norm to work with on M. We fix the box adapted norm on M and let $\tau \in (0, 1)$ be its skew.

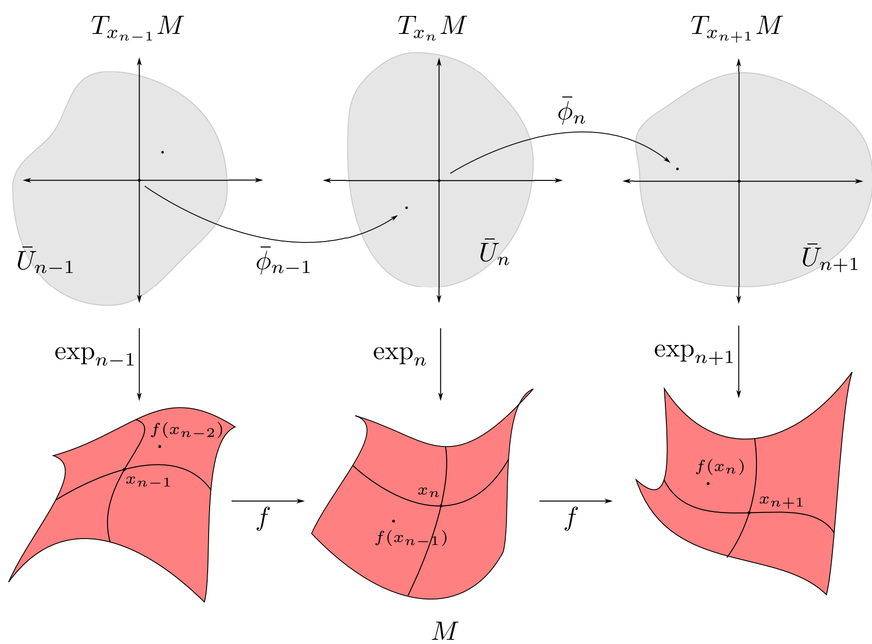


FIGURE 2.

Our general approach will be to think of f as some C^1-perturbation of Df. We will do this by lifting to the tangent space at each x_n using the Riemannian exponential. Lifting f gives us $\phi_n : U_n \subseteq T_{x_n}M \to T_{x_{n+1}}M$ defined by the commutative diagram

$$\begin{CD} U_n \subseteq T_{x_n}M @>{\phi_n}>> T_{x_{n+1}}M \\ @V{\exp_{x_n}}VV @AA{\exp_{x_{n+1}}^{-1}}A \\ M @>{f}>> M \end{CD}$$

for U_n a sufficiently small neighborhood around $0 \in T_{x_n}M$ that $\exp_{x_{n+1}}$ is invertible on $f \circ \exp_{x_n}(U_n)$. While the exponential map must be derived from the original Riemannian metric on the manifold, it does not make much difference since the adapted box norm is continuous and therefore globally comparable to the original metric.

We now apply lemma 4.4 to get $\bar{\phi}_n$ defined over all of $T_{x_n}M$ which agrees with ϕ_n on some neighborhood $\bar{U}_n$ and is C^1-close to $\phi_n(0) + D_0\phi_n$. As noted previously we can also require $\bar{\phi}_n$ to be a invertible. Since $\bar{\phi}_n$ will be a diffeomorphism, we can also require $\bar{\phi}_n^{-1}$ to be C^1-close to $(\phi_n(0) + D_0\phi_n)^{-1}$ by making $\bar{U}_n$ small enough. Note that by compactness of the manifold, the size of $\bar{U}_n$ is uniform over n.

By the chain rule,

$$\begin{CD} T_0T_{x_n}M @>{D_0\phi_n}>> T_{\phi_n(0)}T_{x_{n+1}}M \\ @V{D_0 \exp_{x_n}}VV @AA{D_{f(x_n)} \exp_{x_{n+1}}^{-1}}A \\ T_{x_n}M @>{D_{x_n}f}>> T_{f(x_n)}M \end{CD}$$

commutes. We want $D_0\phi_n$ to look like a hyperbolic derivative, for example $D_{x_n}f$. Under the canonical identification $T_0T_{x_n}M \cong T_{x_n}M$ the derivative $D_0 \exp_{x_n}$ in the composition is the identity. Therefore the problematic factor is the $D_{f(x_n)} \exp_{x_{n+1}}^{-1}$. We know that the exponential is continuous as a map $TM \to M$. We also have that $\|.\|$ and $E^s \oplus E^u$ vary continuously by propositions 3.2 and 3.5. Therefore by making $\|f(x_n) - x_{n+1}\| < \delta$ small we can require that $D_{f(x_n)} \exp_{x_{n+1}}^{-1}$ is close to preserving the norm $\|.\|$ and the splitting $E^s \oplus E^u$. We can thus perturb $D_{f(x_n)} \exp_{x_{n+1}}^{-1}$ to a linear map

$$L_n : T_{f(x_n)}M \to T_{\phi_n(0)}T_{x_{n+1}}M \cong T_{x_{n+1}}M$$

which is norm and splitting preserving. To do this take a basis $v \in E^s$ and $w \in E^u$. There are two vectors in $T_{x_{n+1}}M$ which are of norm $\|v\|$. Let $L_n(v)$ be whichever of them is closer to $D_{f(x_n)} \exp_{x_{n+1}}^{-1}(v)$. Define $L_n(w)$ similarly. Since $D_{f(x_n)} \exp_{x_{n+1}}^{-1}$ is close to preserving the norm and the splitting, we must have that $\|D_{f(x_n)} \exp_{x_{n+1}}^{-1} - L_n\|$ is small under the operator norm. Now consider the linear map Θ_n defined by the diagram

$$\begin{CD} T_0T_{x_n}M @>{\Theta_n}>> T_{\phi_n(0)}T_{x_{n+1}}M \\ @V{D_0 \exp_{x_n}}VV @AA{L_n}A \\ T_{x_n}M @>{D_{x_n}f}>> T_{f(x_n)}M \end{CD} \quad .$$

We will think of Θ_n as a map $T_{x_n}M \to T_{x_{n+1}}M$. Since L_n was constructed to preserve norm and splitting, we know that the hyperbolic properties of $D_{x_n}f$ are pushed forward to Θ_n. In other words, Θ_n preserves the splitting and

$$\|\Theta_n(v)\| \leq \tau\|v\| \qquad \text{for all } v \in E^s$$
$$\|\Theta_n^{-1}(w)\| \leq \tau\|w\| \qquad \text{for all } w \in E^u$$

where $\tau \in (0,1)$ is the skew of our box adapted norm. Now since L_n is close to $D_{f(x_n)} \exp_{x_{n+1}}^{-1}$ in a way which is controlled by δ we know the same is true of Θ_n and $D_0\phi_n$. Because Θ_n and $D_0\phi_n$ are close under the operator norm, they must be C^1-close within bounded regions of $T_0T_{x_n}M \cong T_{x_n}M$. In particular, since $\bar{U}_n$ are assumed to be bounded and $\phi_n(0) + D_0\phi_n$ is close to $\bar{\phi}_n$, it follows that for any $\iota > 0$ we can require, by taking δ small, that

$$\|\bar{\phi}_n - \Phi_n\|_{C^1(\bar{U}_n)} \leq \iota$$
$$\|\bar{\phi}_n^{-1} - \Phi_n^{-1}\|_{C^1(\bar{U}_{n+1})} \leq \iota$$

where $\Phi_n = \phi_n(0) + \Theta_n$.

For notational convenience we define

$$\phi : \coprod_{n\in\mathbb{Z}} U_n \to \coprod_{n\in\mathbb{Z}} T_{x_n}M \qquad \bar{\phi}, \Theta, \Phi : \coprod_{n\in\mathbb{Z}} T_{x_n}M \to \coprod_{n\in\mathbb{Z}} T_{x_n}M$$

as ϕ_n, $\bar{\phi}_n$, Θ_n, and Φ_n respectively on each $T_{x_n}M$. We will also write B_n for the closed ball of radius ε centered at 0. Since $\|.\|$ is a box norm, these balls look like boxes. We will assume ε is small enough so that $B_n \subseteq \bar{U}_n$. It is actually not obvious why we can make this assumption but it will be simpler to postpone the explanation until the end of the proof.

2. Constructing Γ_n. Any shadowing orbit $\{y_n\}_{n\in\mathbb{N}}$ with $\bar{\phi}(y_n) = y_{n+1}$ and $y_n \in B_n$ is uniquely determined by

$$y_0 \in \bigcap_{n\in\mathbb{Z}} \bar{\phi}^{-n}(B_n).$$

Therefore it suffices to show that this intersection is non-empty. For our purposes it will be convenient to split it into the positive and negative directions

$$\left(\bigcap_{n=0}^{\infty} \bar{\phi}^{n}(B_{-n})\right) \cap \left(\bigcap_{n=0}^{\infty} \bar{\phi}^{-n}(B_n)\right).$$

We will begin by working with the positive direction $\bar{\phi}^{-n}(B_n)$ for $n \geq 0$. The idea will be to use Φ to analyze the behavior of $\bar{\phi}$.

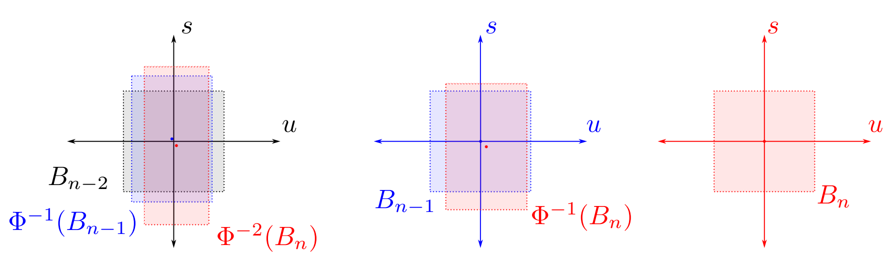


FIGURE 3.

Since Θ is a hyperbolic linear transformation, pulling back the B_n along Φ will dilate them by a factor of τ^{-1} along the stable axis and contract them by τ along the unstable axis.

We will begin by showing that for ι, δ chosen uniformly over n we can make each $\bar{\phi}^{-n}(B_n)$ intersect B_0. We take coordinates u and s for the unstable and stable axes in each tangent space. Since $\bar{\phi}$ is close to Φ we would expect B_n to approach a vertical line the further we pull it back. Therefore to capture this limiting behavior we consider $\ell = \{u = 0\} \cap B_n$. We will show by induction that

$$\Gamma_n = B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(B_{n-1}) \cap \bar{\phi}^{-n}(\ell)$$

is non-empty.

We begin with $\bar{\phi}^{-1}(\ell) = \bar{\phi}_{n-1}^{-1}(\ell)$ which we know is a path in $T_{x_{n-1}}M$. We would like to show that this intersects B_{n-1}. Since everything expands rapidly along the stable axis when we pull back, we would like to remove this axis from the picture. To do this we will choose δ and ι to be small enough that the path $\bar{\phi}^{-1}(\ell)$ crosses the horizontal lines $s = \varepsilon$ and $s = -\varepsilon$. In the ideal hyperbolic model Φ, the pullback $\Phi^{-1}(\ell)$ is a vertical line of length $2\varepsilon\tau^{-1}$ with midpoint at $\bar{\phi}^{-1}(x_n)$. In this case we can easily get $\bar{\phi}^{-1}(\ell)$ to pass $s = \varepsilon$ and $s = -\varepsilon$ by making $\varepsilon\tau^{-1} - \delta > \varepsilon$.

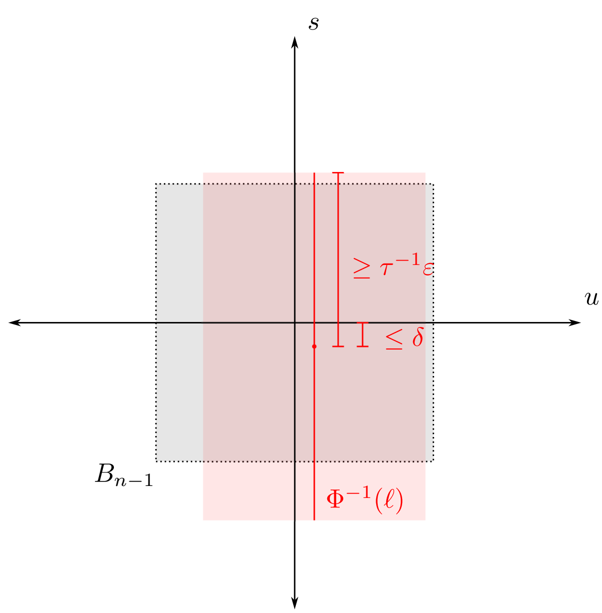


Figure 4.

Now since $\bar{\phi}$ is within ι of Φ we can simply require $\varepsilon\tau^{-1} - \delta - \iota > \varepsilon$. In other words, $\delta + \iota < \varepsilon \cdot (\tau^{-1} - 1)$ which we can do because $\tau \in (0, 1)$ and ι is controlled by δ.

Now note that the same computations uniformly give us that if γ is a path in $T_{x_n}M$ with initial point on $s = -\varepsilon$ and terminal point on $s = \varepsilon$ then $\bar{\phi}^{-1}(\gamma)$ still crosses those lines in $T_{x_{n-1}}M$. Let $S_n = \{s \in [-\varepsilon, \varepsilon]\} \subseteq T_{x_n}M$. We have already shown that $S_{n-1} \cap \bar{\phi}^{-1}(\ell)$ is a path crossing S_{n-1}. Thus taking $\gamma = S_{n-1} \cap \bar{\phi}^{-1}(\ell)$ and reapplying $\bar{\phi}^{-1}$ we get a path

$$\bar{\phi}^{-1}(\gamma) = \bar{\phi}^{-1}(S_{n-1} \cap \bar{\phi}^{-1}(\ell)) = \bar{\phi}^{-1}(S_{n-1}) \cap \bar{\phi}^{-2}(\ell)$$

which crosses $s = -\varepsilon$ and $s = \varepsilon$. Continuing inductively we see that by forcing $\bar{\phi}^{-1}(\ell)$ to cross $s = -\varepsilon$ and $s = \varepsilon$ in $T_{x_{n-1}}M$, we have actually guaranteed that

$$\begin{aligned}&\bar{\phi}^{-1}(S_1 \cap \bar{\phi}^{-1}(S_2 \cap \ldots \bar{\phi}^{-1}(S_{n-1} \cap \bar{\phi}^{-1}(\ell)) \ldots)) \\ &\quad = \bar{\phi}^{-1}(S_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(S_{n-1}) \cap \bar{\phi}^{-n}(\ell)\end{aligned}$$

crosses $s = -\varepsilon$ and $s = \varepsilon$ in $T_{x_0}M$. Each intersection with a $\bar{\phi}^{-m}(S_m)$ restricts the domain of the path to a subinterval of ℓ so that no $\bar{\phi}^m(\gamma)$ extends beyond S_m. This is necessary because $\bar{\phi}^{-1}$ is only close to Φ^{-1} within the $\bar{U}_n$ and since $\bar{\phi}^n(\ell)$ lengthens exponentially it quickly exits those bounds.

Since $\bar{\phi}^{-1}(S_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(S_{n-1}) \cap \bar{\phi}^{-n}(\ell)$ spans S_0 we have reduced the problem of showing that $\bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(B_{n-1}) \cap \bar{\phi}^{-n}(\ell)$ intersects B_0 to showing that the u-coordinates of $\bar{\phi}^{-1}(S_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(S_{n-1}) \cap \bar{\phi}^{-n}(\ell)$ always lie within $[-\varepsilon, \varepsilon]$.

To resolve the u-coordinate direction we again start by considering Φ. Since Φ is affine it suffices to consider points $w \in E^u$. Since $\Phi_n = \phi_n(0) + \Theta_n$ we have

$$\Phi_n^{-1}(w) = \Theta_n^{-1}(w - \phi_n(0)).$$

Now considering the norm of the horizontal coordinate we have

$$\|\Theta_n^{-1}(w - \phi_n(0))\|_u = \tau\|w - \phi_n(0)\|_u \leq \tau\|w\|_u + \tau\delta.$$

Accounting for the ι error we are left with

$$\|\Phi_n^{-1}(w)\|_u \leq \iota + \tau\|w\|_u + \tau\delta.$$

Since $\ell = \{u = 0\} \cap B_n$ we begin with $\|w\|_u = 0$. Applying our recursive rule we get the sequence

$$0, \quad \iota + \tau\delta, \quad \iota + \tau\delta + \tau\iota + \tau^2\delta, \quad \ldots, \quad \iota \cdot (1 + \tau + \cdots + \tau^{n-1}) + \delta \cdot (1 + \tau + \cdots + \tau^n).$$

Therefore we see that the horizontal deviation from the origin is bounded by the infinite sum

$$\iota \cdot (1 + \tau + \tau^2 + \ldots) + \delta \cdot (1 + \tau + \tau^2 + \ldots) = \frac{\iota + \delta}{1 - \tau}.$$

Thus we want to require that $\iota + \delta < \varepsilon \cdot (1 - \tau)$. As before, note that our ι estimate only works within the $\bar{U}_n$. Since $\bar{\phi}^{-n}(\ell)$ blows up in the vertical direction, the estimate cannot hold over the whole path. However, if after each application of $\bar{\phi}^{-1}$ we restrict to the segment of the path going between the lines $s = -\varepsilon$ and $s = \varepsilon$ then there is no issue. Let this iteratively truncated version of $\bar{\phi}^{-n}(\ell)$ be called

$$\Gamma_n = B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-(n-1)}(B_{n-1}) \cap \bar{\phi}^{-n}(\ell).$$

Since the bounds we have set are uniform, we know by induction that each

$$B_m \cap \bar{\phi}^{-1}(B_{m+1}) \cap \cdots \cap \bar{\phi}^{m-n}(\ell)$$

is non-empty and is the image of a path in B_m from $s = -\varepsilon$ to $s = \varepsilon$. In particular, this holds for Γ_n. It is immediate from the definition that have

$$\Gamma_n \subseteq \bigcap_{m=0}^{n} \bar{\phi}^{-m}(B_m).$$

In summary, within each partial intersection we have found a set Γ_n which is the image of a path with initial point lying on $s = -\varepsilon$, terminal point on $s = \varepsilon$, and whose u-coordinates remain within $[-\varepsilon, \varepsilon]$.

3. The Γ_n are graphs of Lipschitz functions α_n. We now want to use the fact that $\|\bar{\phi}^{-1} - \Phi^{-1}\|_{C^1(U_n)} \leq \iota$ to show that Γ_n is the graph of a function in s. We already know Γ_n is the image of a C^1 path $\gamma_n : I \to B_0$ where I is some subinterval of $[-\varepsilon, \varepsilon] \cong \ell$. Suppose Γ_n were not the graph of a function. In other words, that there were $a, b \in I$ distinct such that the stable coordinates $(\gamma_n)_s(a) = (\gamma_n)_s(b)$ agreed. Then by Rolle's theorem we would have some $t \in I$ at which $(\dot{\gamma}_n)_s(t) = 0$. We will use induction to find ι to make this impossible and thus give us that Γ_n is the graph of some function $\alpha_n(s)$. Our induction will also simultaneously prove that the Lipschitz constant of $\alpha_n(s)$ is less than 1. Since $\alpha_n(s)$ is C^1 this second claim is equivalently $\|\alpha_n'\|_\infty = \sup \|(\dot{\gamma}_n)_u\|/\|(\dot{\gamma}_n)_s\| < 1$.

The base case of ℓ is a vertical line which is of course a function in s with Lipschitz constant less than 1. Now suppose $\operatorname{im} \gamma_n$ is the graph of a function, ie. $(\dot{\gamma}_n)_s(t) \neq 0$, and $\|(\dot{\gamma}_n)_u\| < \|(\dot{\gamma}_n)_s\|$. Then $(\gamma_{n+1})_s = \bar{\phi}_s^{-1} \circ \gamma_n(t)$ and $(\gamma_{n+1})_u = \bar{\phi}_u^{-1} \circ \gamma_n(t)$. Therefore, using the fact that $\|\bar{\phi}^{-1} - \Phi^{-1}\|_{C^1(\bar{U}_n)} \leq \iota$ and Φ^{-1} is a diagonal affine transformation with respect to s and u we get

$$\begin{aligned}
\|(\dot{\gamma}_{n+1})_s\| &= \left\| \frac{d}{dt} \bar{\phi}_s^{-1} \circ \gamma_n \right\| = \left\| \frac{d\bar{\phi}_s^{-1}}{ds} \cdot (\dot{\gamma}_n)_s + \frac{d\bar{\phi}_s^{-1}}{du} \cdot (\dot{\gamma}_n)_u \right\| \\
&\geq \left\| \frac{d\bar{\phi}_s^{-1}}{ds} \right\| \cdot \|(\dot{\gamma}_n)_s\| - \left\| \frac{d\bar{\phi}_s^{-1}}{du} \right\| \cdot \|(\dot{\gamma}_n)_u\| \\
&\geq \left(\left\| \frac{d\Phi_s^{-1}}{ds} \right\| - \iota \right) \cdot \|(\dot{\gamma}_n)_s\| - \left(\left\| \frac{d\Phi_s^{-1}}{du} \right\| + \iota \right) \cdot \|(\dot{\gamma}_n)_u\| \\
&= (\tau^{-1} - \iota)\|(\dot{\gamma}_n)_s\| - \iota\|(\dot{\gamma}_n)_u\| > (\tau^{-1} - 2\iota)\|(\dot{\gamma}_n)_s\|.
\end{aligned}$$

Note that we used the Lipschitz part of the inductive hypothesis in the second line. Since we also have $\|(\dot{\gamma}_n)_s\| > 0$ by the inductive hypothesis, requiring $\iota < \tau^{-1}/2$ gives us $(\dot{\gamma}_{n+1})_s \neq 0$. Now for the Lipschitz part of this induction we compute the derivative of the other component

$$\begin{aligned}
\|(\dot{\gamma}_{n+1})_u\| &= \left\| \frac{d}{dt} \bar{\phi}_u^{-1} \circ \gamma_n \right\| = \left\| \frac{d\bar{\phi}_u^{-1}}{du} \cdot (\dot{\gamma}_n)_u + \frac{d\bar{\phi}_u^{-1}}{ds} \cdot (\dot{\gamma}_n)_s \right\| \\
&\leq (\tau + \iota)\|(\dot{\gamma}_n)_u\| + \iota\|(\dot{\gamma}_n)_s\| < (\tau + 2\iota)\|(\dot{\gamma}_n)_s\|.
\end{aligned}$$

Dividing by the inequality for the stable component we get

$$\|(\dot{\gamma}_{n+1})_u\|/\|(\dot{\gamma}_{n+1})_s\| \leq \frac{\tau + 2\iota}{\tau^{-1} - 2\iota}.$$

Therefore not only can we force the Lipschitz constant to be less than 1, we can also, for any $C \in (\tau^2, 1)$, choose ι small enough that the Lipschitz constant of any α_n is less than C.

Note that our inductive notation "γ_{n+1}" was somewhat of an abuse of notation: The running assumption has been that every γ_n is a path in B_0 corresponding to a segment of $\bar{\phi}^{-n}(\ell)$. From our usage, $\gamma_{n+1} = \bar{\phi}^{-1} \circ \gamma_n$. However this poses no issue since everything we have done holds uniformly over n. In other words the important part of the induction was how many times we pulled back, not which chart we ended up in. We therefore get a sequence of functions $\alpha_n(s)$ whose graphs lie in B_0 and which are C^1 with Lipschitz constants bounded by $C < 1$. The graph of each $\alpha_n(s)$ is Γ_n which is contained in the partial intersection $B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-n}(B_n)$.

4. The positive solution set is the limit of Γ_n. Let

$$I_n = \bigcap_{m=0}^{n} \bar{\phi}^{-m}(B_m)$$

be the partial intersections. This region in B_0 is bounded on either side by the paths

$$B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-n}(\{u = -\varepsilon\}) \qquad B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-n}(\{u = \varepsilon\}).$$

The same arguments made for Γ_n prove that these two paths are actually graphs of functions $\varphi_n(s)$ and $\psi_n(s)$ respectively with Lipschitz constants less than 1. I claim that the width of I_n given by $\omega(I_n) = \|\psi_n - \varphi_n\|_\infty$ converges to 0. We will prove this by induction. Suppose R is some region in B_n bounded by $\varphi(s)$ on the left and $\psi(s)$ on the right. The derivative of Φ^{-1} along any $R \cap \{s = r\} = [\varphi(r), \psi(r)] \times \{r\}$ is τ. Therefore the derivative of $\bar{\phi}^{-1}$ along $R \cap \{s = r\}$ is no greater than $\tau + \iota$. It follows that the length $\ell(R \cap \{s = r\})$ is no greater than

$$(\tau + \iota) \cdot (\psi(r) - \varphi(r)) \leq (\tau + \iota) \cdot \omega(R)$$

where $\omega(R) = \|\psi - \varphi\|_\infty$ is the width of R.

Now look at any slice $\bar{\phi}^{-1}(R) \cap \{s = r\} \subseteq T_{x_{n-1}}M$ whose length we would like to compute. Parametrize $\bar{\phi}^{-1}(R) \cap \{s = r\} \subseteq T_{x_{n-1}}M$ as a path ρ. Suppose p is the left endpoint of ρ and let $r' = \bar{\phi}_s(p)$. Define σ to be the path given by $\bar{\phi}^{-1}(R \cap \{s = r'\})$ so that ρ and σ intersect at their left endpoint p. Let h be the vertical distance between the right endpoints of ρ and σ. Suppose q and q' are respectively the endpoints of ρ and σ. Since both q and q' lie on the right boundary of $\bar{\phi}^{-1}(R)$, which is Lipschitz with constant less than 1, we know that $\|q - q'\|_u < \|q - q'\|_s = h$.

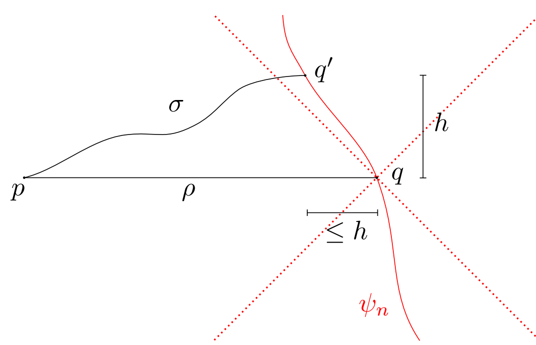


FIGURE 5.

If we parametrize σ by $\bar{\phi}^{-1}|_{R \cap \{s = r'\}}$ according to its definition then we have that $\dot{\sigma}_u$ is within ι of τ and $\dot{\sigma}_s$ is within ι of 0. Therefore by requiring $2\iota < \tau$ we can force $\|\dot{\sigma}_s\| < \|\dot{\sigma}_u\|$. Since we are working with a box norm it will follow that $\|\dot{\sigma}\| = \|\dot{\sigma}_u\|$ and so the length $\ell(\sigma)$ is equal to the horizontal distance $\|q - q'\|_u$. We then have $|\ell(\rho) - \ell(\sigma)| \leq h$.

Now note that

$$h \leq (\ell(\rho) + h) \cdot \frac{\|\dot{\sigma}_s\|}{\|\dot{\sigma}_u\|} \leq (\ell(\rho) + h) \cdot \frac{\iota}{\tau - \iota}.$$

Isolating for h in the inequality we get

$$h \leq \frac{\iota}{\tau - 2\iota} \ell(\rho).$$

Thus we have

$$|\ell(\rho) - \ell(\sigma)| \leq \frac{\iota}{\tau - 2\iota} \ell(\rho).$$

We have already shown that $\ell(\sigma) \leq (\tau + \iota) \cdot \omega(R)$ and so

$$\ell(p) \leq (\tau + \iota) \cdot \omega(R) + \frac{\iota}{\tau - 2\iota} \ell(\rho).$$

This finally leaves us with the inequality

$$\ell(p) \leq \frac{\tau - 2\iota}{\tau - 3\iota} \cdot (\tau + \iota) \cdot \omega(R).$$

By taking ι small, $(\tau - 2\iota)/(\tau - 3\iota)$ can be made close to 1 and $\tau + \iota$ can be made close to τ. Therefore for some number λ close to τ we have $\ell(p) \leq \lambda \omega(R)$ and thus

$$\omega(\bar{\phi}^{-1}(R)) \leq \lambda \omega(R).$$

It follows that $\omega(B_0 \cap \bar{\phi}^{-1}(B_1) \cap \cdots \cap \bar{\phi}^{-n}(B_n)) \leq \lambda^n \cdot 2\varepsilon$ which converges to 0.

Since any I_m contains the graphs of α_n for all $n \geq m$, the convergence of the width $\omega(I_m) \to 0$ tells us that α_n have a uniform limit α. Because α_n are all C-Lipschitz for some uniform $C < 1$, their limit α must also be C-Lipschitz. Moreover, since we took the balls B_m to be closed, we know that I_m are closed. It follows that the graph Γ of α is contained in I_m for all m and therefore

$$\Gamma \subseteq \mathcal{I} = \bigcap_{n=0}^{\infty} I_n = \bigcap_{n=0}^{\infty} \bar{\phi}^{-n}(B_n).$$

Finally, for each $r \in [-\varepsilon, \varepsilon]$ we have

$$\mathcal{I} \cap \{s = r\} = \bigcap_{n=0}^{\infty} I_n \cap \{s = r\}$$

which is an intersection of nested intervals with lengths converging to 0. Therefore each $\mathcal{I} \cap \{s = r\}$ is a point and so we actually have $\mathcal{I} = \Gamma$.

5. Finding a unique solution. So far we have described the points solving the shadowing problem in the positive direction as the graph of some C-Lipschitz function α in s. An analogous argument shows that the shadowing problem in the negative direction is solved by a C-Lipschitz function β in u.

The shadowing orbits are in bijection with the intersections of the graphs of α and β so we want to show that these graphs have a unique intersection. First we show that an intersection exists. At an intersection point we have

$$(\alpha(s), s) = (u, \beta(u)) = (\alpha(s), \beta(\alpha(s))).$$

In other words the intersection points are the fixed points of $\beta \circ \alpha : [-\varepsilon, \varepsilon] \to [-\varepsilon, \varepsilon]$. We have $\beta \circ \alpha(-\varepsilon) \geq -\varepsilon$ and $\beta \circ \alpha(\varepsilon) \leq \varepsilon$. Thus by the intermediate value theorem applied to $\beta \circ \alpha(s) - s$ we have a fixed point. This fixed point must be unique since

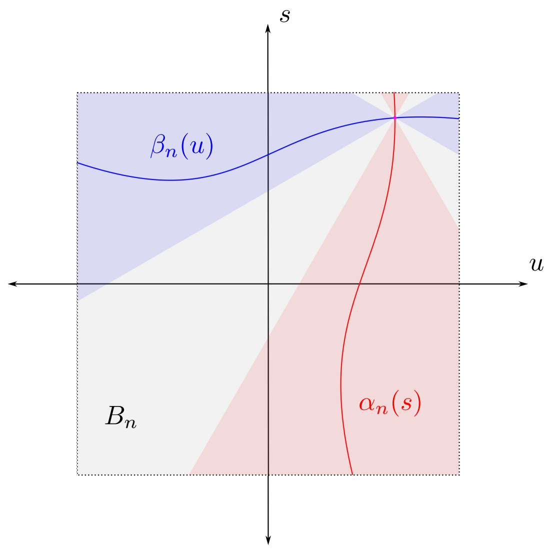


FIGURE 6. Since α and β are Lipschitz with constant $C < 1$ there is a unique intersection between their graphs.

$\beta \circ \alpha$ is C^2-Lipschitz and $C^2 < 1$.

6. Concluding remarks. In summary, we have found a point $y \in B_0$ such that $\bar{\phi}^n(y) \in B_n$ for any $n \in \mathbb{Z}$. Since $\bar{\phi}$ agrees with ϕ on $\bar{U}_n$ and $B_n \subseteq \bar{U}_n$ this means

$$\phi^n(y) \in B_n \quad \text{for all } n \in \mathbb{Z}.$$

Since all norms on a finite dimensional vector space are comparable, we can assume that $\phi^n(y)$ is within ε of $0 \in T_{x_n}M$ under the original metric. Explicitly, since

$$\phi_m = \exp_{x_{m+1}}^{-1} \circ f \circ \exp_{x_m} \qquad \phi_m^{-1} = \exp_{x_m}^{-1} \circ f^{-1} \circ \exp_{x_{m+1}}$$

we get that

$$\phi^n(y) = \phi_{n-1} \circ \cdots \circ \phi_0(y) = \exp_{x_n}^{-1} \circ f^n \circ \exp_{x_0}$$

and

$$\phi^{-n}(y) = \phi_{-n} \circ \cdots \circ \phi_{-1}(y) = \exp_{x_{-n}}^{-1} \circ f^{-n} \circ \exp_{x_0}$$

for all $n \geq 0$. In other words $\phi^n(y) = \exp_{x_n}^{-1} \circ f^n \circ \exp_{x_0}$ for all $n \in \mathbb{Z}$. Since $\exp_{x_n}^{-1}$ is radially distance preserving, $\phi^n(y)$ being within ε of 0 implies that $f^n(\exp_{x_0}(y))$ is within ε of x_n. In other words, $\exp_{x_0}(y)$ generates the ε-shadowing orbit. This demonstrates that existence projects down from our tangent charts to the manifold.

Before we get to uniqueness we must address the way our variables ε, δ and ι were chosen in the proof. We restricted ι and δ based on ε, but then we also required that the ε-balls B_n be contained in $\bar{U}_n$, whose size was dependent on ι.

To resolve this circular dependence, we consider all of the restrictions placed on ι. In step 2 of the proof we required

$$\delta + \iota < \varepsilon \cdot (\tau^{-1} - 1) \quad \text{and} \quad \delta + \iota < \varepsilon \cdot (1 - \tau).$$

In step 3 we required

$$2\iota < \tau^{-1} \quad \text{and} \quad \frac{\tau + 2\iota}{\tau^{-1} - 2\iota} \le C \in (\tau^2, 1).$$

In step 4 we required

$$2\iota < \tau \quad \text{and} \quad \frac{\tau - 2\iota}{\tau - 3\iota} \cdot (\tau + \iota) \le \lambda \in (\tau, 1).$$

The restrictions from steps 3 and 4 pose no issues since they are not reliant on ε. When we restrict ι in step 2, we only make use of the C^0 aspect of this bound. In other words, step 2 only requires

$$\|\bar{\phi}_n^{-1} - \Phi_n^{-1}\|_{C^0(\bar{U}_n)} \le \iota$$

for which we only need

$$\|\bar{\phi}_n^{-1} - (\phi_n(0) + D_0\phi_n)^{-1}\|_{C^0(\bar{U}_n)} \le \iota/2.$$

From the proof of the extension lemma, or more directly from the definition of the derivative, we see that

$$\|\bar{\phi}_n^{-1} - (\phi_n(0) + D_0\phi_n)^{-1}\|_{C^0(\bar{U}_n)} = o(\operatorname{diam} \bar{U}_n).$$

In other words we can choose $\bar{U}_n$ based on ι such that $\iota/\operatorname{diam} \bar{U}_n \to 0$ as $\iota \to 0$. Since the restrictions in step 2 are linear in ε, we see that as $\varepsilon \to 0$, we can choose $\iota \to 0$ to be proportional to ε. But then if we choose $\bar{U}_n$ so that $\operatorname{diam} \bar{U}_n$ blows up with respect to ι, it will also blow up with respect to ε. In summary, as ε is made arbitrarily small, we can eventually choose $\bar{U}_n$ such that ι satisfies the restrictions and $B_n \subseteq \bar{U}_n$.

Now for uniqueness, suppose ε_0 is small enough that we can require $B_n \subseteq \bar{U}_n$ as described above and δ_0 is the δ which the proof assigns to ε_0. If $\{x_n\}_{n\in\mathbb{Z}}$ is a δ_0-pseudo orbit which is ε_0-shadowed by some orbit $\{y_n\}_{n\in\mathbb{Z}}$ then the lifts $\exp_{x_n}^{-1}(y_n)$ must lie within B_n and therefore within $\bar{U}_n$. It follows that $\exp_{x_n}^{-1}(y_n)$ are an orbit of $\bar{\phi}$ and that

$$y_0 \in \bigcap_{n\in\mathbb{Z}} \bar{\phi}^{-n}(B_n)$$

which we have already showed contains precisely one point. Thus $\{y_n\}_{n\in\mathbb{Z}}$ is the unique ε_0-shadowing orbit and we have obtained uniqueness in the shadowing lemma. □

The shadowing lemma as we have stated it is often thought of as a combination of two properties. The first of these is the shadowing property, that is, existence of a shadowing orbit. The uniqueness conclusion is separated and rephrased as follows.

Corollary 4.5. ***Expansiveness.*** *Suppose $f : M \to M$ is an Anosov diffeomorphism. There exists $\varepsilon > 0$ such that if, for any $x, y \in \Lambda$, $d(f^n(x), f^n(y)) \le \varepsilon$ for all $n \in \mathbb{Z}$ then $x = y$.*

Proof. Let $\varepsilon = \varepsilon_0$ with ε_0 as in the shadowing lemma. The orbits of x and y are δ-pseudo orbits for any $\delta > 0$. In particular, they are δ_0-orbits and are therefore each ε_0-shadowed by at most one point. However, x and y each shadow their own orbits, and each ε_0-shadow each others' orbits. The uniqueness clause of the shadowing lemma therefore guarantees that $x = y$. □

For our purposes, we only care about the shadowing and expansiveness theorems in the case where f is an Anosov diffeomorphism. However, they also hold over any hyperbolic set $\Gamma \subseteq M$ of a diffeomorphism $f : M \to M$.

It is interesting to note that though we used the uniqueness part of the shadowing lemma to prove expansiveness, the converse implication also holds: Expansiveness is sufficient to guarantee uniqueness of shadowing for some $\varepsilon_0 > 0$ and $\delta_0 > 0$. To see why this is true note that if the orbits of x and y both ε-shadow a pseudo orbit $\{x_n\}_{n\in\mathbb{Z}}$ then the orbits 2ε-shadow each other.

In essence, the shadowing lemma as we have stated it is a combination of two properties of hyperbolic sets: Expansiveness and (non-unique) shadowing. We precisely define these properties as follows:

Definition 4.6. Let $f : X \to X$ be a homeomorphism of a compact metric space (X, d). We say that f has the *shadowing property* (sometimes also referred to as the *pseudo orbit tracing property*) if for any $\varepsilon > 0$ there exists $\delta > 0$ such that every δ-pseudo orbit of X is ε-shadowed by at least one point.

We say that f is *expansive* if there exists $\varepsilon_0 > 0$ such that for any $x, y \in X$, $d(f^n(x), f^n(y)) \leq \varepsilon_0$ for all $n \in \mathbb{Z}$ implies $x = y$.

If $f : X \to X$ is expansive and has shadowing then we will call f an *Anosov homeomorphism*.

Anosov homeomorphisms generalize Anosov diffeomorphisms. For example, we will see that most of the structural stability of Anosov diffeomorphisms can be derived using shadowing and expansiveness.

5. Stability of Anosov Diffeomorphisms

We return to our original purpose of demonstrating the stability of Anosov Diffeomorphisms. Shadowing and expansiveness will play a vital role.

5.1. **Topological Semi-Stability of Anosov Homeomorphisms.**

For Anosov homeomorphisms in general we require an even weaker version of stability which we will then strengthen in the case of f acting on a closed manifold M.

Definition 5.1. A dynamical system $f : X \to X$ is *weakly topologically semi-conjugate* to $g : Y \to Y$ if there is a map $h : X \to Y$ such that $hf = gh$.

Definition 5.2. A dynamical system $f : X \to X$ is *weakly topologically semi-stable* if there exists some C^0 neighborhood $\mathcal{U} \subseteq \text{Homeo}(X)$ of f such that every $g \in \mathcal{U}$ is weakly topologically semi-conjugate to f.

To begin proving stability results we need the following finite version of expansiveness:

Lemma 5.3. *Suppose X is a compact metric space and $f : X \to X$ is an expansive homeomorphism under the constant ε_0. Then for any $\lambda > 0$ there exists $N \in \mathbb{N}$ such that*

$$d(f^n(x), f^n(y)) \leq \varepsilon_0 \quad \textit{for all } |n| \leq N$$

implies $d(x, y) < \lambda$.[6]

[6] Walters[3] lemma 2, page 235.

Proof. If this lemma did not hold for some $\lambda > 0$ we could find x_N, y_N such that

$$d(f^n(x_N), f^n(y_N)) \le \varepsilon_0 \quad \text{for all } |n| \le N$$

and $d(x_N, y_N) \ge \lambda$. Since X is a compact metric space we can extract convergent subsequences $x_{N_i} \to x$ and $y_{N_i} \to y$. For arbitrarily large N_i we have

$$\begin{aligned} d(f^n(x),&f^n(y)) \\ &\le d(f^n(x), f^n(x_{N_i}) + d(f^n(x_{N_i}), f^n(y_{N_i})) + d(f^n(y_{N_i}), f^n(y)) \\ &\le d(f^n(x), f^n(x_{N_i}) + \varepsilon_0 + d(f^n(y_{N_i}), f^n(y)) \end{aligned}$$

for all $|n| \le N_i$. For any fixed n the error terms $d(f^n(x), f^n(x_{N_i}) + d(f^n(y_{N_i}), f^n(y))$ become arbitrarily small as $N_i \to \infty$. It therefore follows that

$$d(f^n(x), f^n(y)) \le \varepsilon_0 \quad \text{for all } n.$$

But since $d(x_{N_i}, y_{N_i}) \ge \lambda$ we also have $d(x, y) \ge \lambda > 0$ which contradicts expansiveness. □

We now address the most general case of stability.

Theorem 5.4. *Any Anosov homeomorphism $f : X \to X$ of a compact metric space X is weakly topologically semi-stable. In fact, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any g satisfying $d(f, g) < \delta$*

(1) There exists a unique weak semi-conjugacy h from g to f
(2) $d(h, \mathrm{id}) < \varepsilon$.[7]

Proof. Suppose f is expansive under the constant ε_0. Let $0 < \varepsilon < \varepsilon_0/3$ and $\delta > 0$ correspond to ε as given to us by the shadowing property. We take our neighborhood of stability $\mathcal{U}$ to be the ball of radius δ around $f \in \mathrm{Homeo}(X)$.

Suppose $g \in \mathcal{U}$, in other words $d(f, g) < \delta$. Fix $x \in X$ and let $x_n = g^n(x)$. Note that

$$d(f(x_n), x_{n+1}) = d(f(x_n), g(x_n)) \le d(f, g) < \delta.$$

In other words, $\{x_n\}_{n\in\mathbb{Z}}$ is a δ-pseudo orbit of f. We therefore find that it is ε-shadowed by some unique orbit $\{f^n(y)\}_{n\in\mathbb{Z}}$. Let $h(x) = y$.

For any $x \in X$ we have found $h(x) \in X$ such that

$$d(f^n(h(x)), g^n(x)) \le \varepsilon$$

for all $n \in \mathbb{Z}$. For $n = 0$ we see that $d(h, \mathrm{id}) < \varepsilon$ so we can make h uniformly close to id as desired. By substituting x for $g(x)$ we see that

$$d(f^n(h(g(x))), g^n(g(x))) \le \varepsilon$$

and by replacing n with $n + 1$ we get

$$\text{and} \quad d(f^{n+1}(h(x)), g^{n+1}(x)) \le \varepsilon.$$

Adding these two together leaves us with

$$\begin{aligned} d(f^n(hg(x)),&f^n(fh(x))) \\ &\le d(f^n(h(x)), g^n(x)) + d(f^n(h(g(x))), g^n(g(x))) < 2\varepsilon < \varepsilon_0 \end{aligned}$$

for all $n \in \mathbb{Z}$. By expansiveness it follows that $hg(x) = fh(x)$ for all $x \in X$.

[7] Walters[3] theorem 4, page 236.

We now address the continuity of h. Fix $\lambda > 0$ and choose a corresponding $N \in \mathbb{N}$ as given by lemma 5.3. Since g is continuous we can find μ such that $d(x, y) < \mu$ implies $d(g^n(x), g^n(y)) < \varepsilon$ for all $|n| \le N$. Then for any such x, y it also follows that

$$\begin{aligned} d(f^n(h(x)), &f^n(h(y))) \\ &\le d(f^n(h(x)), g^n(x)) + d(g^n(x), g^n(y)) + d(f^n(h(y)), g^n(y)) \\ &< 3\varepsilon < \varepsilon_0 \end{aligned}$$

which implies $d(h(x), h(y)) < \lambda$ and so we have continuity.

Finally for uniqueness, suppose k were another weak semi-conjugacy within ε of id. Then

$$\begin{aligned} d(f^n(h(x)), f^n(k(x))) &= d(h(g^n(x)), k(g^n(x))) \\ &\le d(h(g^n(x)), g^n(x)) + d(g^n(x), k(g^n(x))) < 2\varepsilon < \varepsilon_0 \end{aligned}$$

for all $n \in \mathbb{Z}$ and so it follows by expansiveness that $h(x) = k(x)$. □

If X is a closed manifold then we get semi-stability as promised.

Corollary 5.5. *If $f : M \to M$ is a Anosov homeomorphism on a closed manifold M then f is topologically semi-stable.*[8]

Proof. We want to show that $h : M \to M$ is surjective. In general, any surjective homeomorphism between two manifolds without boundary has a C^0 neighborhood in which every map is also surjective.[9] Using theorem 5.4 we can make h as close to id as is necessary to force it to be surjective. □

5.2. **Structural Stability of Anosov Diffeomorphisms.** We can now use hyperbolicity to strengthen this result for Anosov diffeomorphisms.

In order to prove that h is injective we need to first show that C^1-perturbations of Anosov diffeomorphisms are expansive. This fact follows from the proof of the shadowing lemma.

Lemma 5.6. ***Persistence of Expansiveness.*** *Suppose $f : M \to M$ is an Anosov diffeomorphism. There is a neighborhood $\mathcal{U} \subseteq \mathrm{Diff}^1(M)$ of f such that every $g \in \mathcal{U}$ is expansive.*

Proof. In the proof we gave of the shadowing lemma (4.3) the hyperbolicity of f was only used to show $\|\bar{\phi}_n - \Phi_n\|_{C^1(\bar{U}_n)} \le \iota$ and the analogous statement for the inverses. Now suppose we want to repeat the proof but for g. We can construct $\bar{\psi}_n$ for g in the same way that $\bar{\phi}_n$ were constructed for f, and by making $\mathcal{U} \subseteq \mathrm{Diff}^1(M)$ and $\bar{U}_n$ sufficiently small, we can guarantee that

$$\|\bar{\phi}_n - \bar{\psi}_n\|_{C^1(\bar{(U_n)})} \le \iota/2 \quad \text{and} \quad \|\bar{\phi}_n - \Phi_n\|_{C^1(\bar{U}_n)} \le \iota/2.$$

In other words we get the same restriction

$$\|\bar{\psi}_n - \Phi_n\|_{C^1(\bar{U}_n)} \le \iota$$

which allows us to complete the same proof but for g. This does not prove that g satisfies the conclusion of the shadowing lemma because we chose $\mathcal{U}$ based on ι.

[8] Walters[3] remark, page 237.

[9] Munkres[4] lemma 3.11, page 36.

However, it does prove that for any fixed $\varepsilon > 0$ there exists a neighborhood $\mathcal{U} \subseteq \text{Diff}^1(M)$ such that any $g \in \mathcal{U}$ satisfies the shadowing lemma for ε. In particular, taking $\mathcal{U} \subseteq \text{Diff}^1(M)$ corresponding to $\varepsilon = \varepsilon_0$, we get the uniqueness part of the shadowing lemma for all $g \in \mathcal{U}$.

More concretely, the uniqueness part of the proof of the shadowing lemma was dependent on our ability to make α_n uniformly C-Lipschitz and to force $\omega(I_n) \to 0$. We brought about these two conditions in steps 3 and 4 of the proof respectively. As noted in the concluding remarks of the shadowing lemma proof, steps 3 and 4 put bounds on ι that were dependent only on τ and not ε. This is why we only need to make $\mathcal{U}$ small enough so that ι satisfies those restrictions in terms of τ. □

Note from this proof we see that the expansiveness constant ε_0 which the shadowing lemma calculates for f will also hold for all $g \in \mathcal{U}$. This is by no means the maximal constant for either f or g, however it proves that the elements of $\mathcal{U}$ are *uniformly* expansive. Structural stability follows quickly.

Theorem 5.7. ***Structural Stability.*** *Anosov diffeomorphisms are C^1-structurally stable.*[10] [11]

Proof. Suppose $f : M \to M$ is our Anosov diffeomorphism. Using corollary 5.5 and lemma 5.6 choose a C^1 neighborhood $\mathcal{U} \subseteq \text{Diff}^1(M)$ of f over which topological semi-stability and persistence of expansiveness both hold. In other words, for every $g \in \mathcal{U}$ we expect g to be expansive and semi-conjugate to f. As remarked upon above, the elements of $\mathcal{U}$ are *uniformly* expansive. In other words there is some ε_0 over which they all satisfy expansiveness. We can now further restrict $\mathcal{U}$ according to theorem 5.4 so that for any $g \in \mathcal{U}$ the associated semi-conjugacy h to f satisfies $d(h, \text{id}) < \varepsilon_0/2$.

Finally, fix any $g \in \mathcal{U}$ and its associated semi-conjugacy h. If $h(x) = h(y)$ then

$$\begin{aligned} d(g^n(x),&g^n(y)) \\ &\leq d(g^n(x), h(g^n(x))) + d(h(g^n(x)), h(g^n(y))) + d(h(g^n(y)), g^n(y)) \\ &= d(g^n(x), h(g^n(x))) + d(g^n(h(x)), g^n(h(y))) + d(h(g^n(y)), g^n(y)) \\ &< \varepsilon_0/2 + 0 + \varepsilon_0/2 = \varepsilon_0 \end{aligned}$$

for all $n \in \mathbb{Z}$ and so by the expansiveness of g we find $x = y$. It follows that h is a bijection. Since M is compact Hausdorff this means h is a homeomorphism and a conjugacy. □

Acknowledgments

I would like to firstly thank my mentor Meg Doucette whose guidance made this paper possible. I would also like to thank Amie Wilkinson for her advice on proving the shadowing lemma. Finally, I am very thankful to Peter May for organizing the program.

---

[10] Walters[3] theorem 5, page 237.

[11] Hyperbolicity is not preserved under topological conjugacy (see [5]) so structural stability does not imply that maps C^1-close to an Anosov diffeomorphism are Anosov. However, "persistence of hyperbolicity" is independently true. See Wen[1] theorem 4.6, page 87.

References

[1] Lan Wen. Differentiable Dynamical Systems. American Mathematical Society. 2016.
[2] Anatole Katok and Boris Hasselblatt. Introduction to the Modern Theory of Dynamical Systems. Cambridge University Press. 1995.
[3] Peter Walters. On the Pseudo Orbit Tracing Property and its Relationship to Stability. In *The Structure of Attractors in Dynamical Systems*, edited by A. Dold and B. Eckmann. Springer-Verlag Berlin Heidelberg. 1978.
[4] James Munkres. Elementary Differential Topology. Princeton University Press. 1966.
[5] Andrey Gogolev. Diffeomorphisms H older Conjugate to Anosov Diffeomorphisms. Ergodic Theory and Dynamical Systems, 30(2). 2010. https://people.math.osu.edu/gogolyev.1/index_files/tangency.pdf
[6] Clark Robinson. Stability Theorems and Hyperbolicity in Dynamical Systems. Rocky Mountain Journal of Mathematics, 7(3). 1977.