PRELIMINARY VERSION

DYNAMICS BEYOND UNIFORM HYPERBOLICITY: A GLOBAL GEOMETRIC AND PROBABILISTIC APPROACH

CHRISTIAN BONATTI, LORENZO J. DÍAZ, MARCELO VIANA

Contents

Date: June 10, 2003.

Introduction

About forty years ago, Smale was proposing the notion of *uniformly hyperbolic dynamical system*: the limit set, consisting of all forward or backward accumulation points of orbits, is a *hyperbolic set*: the tangent space at each point splits into two complementary subspaces that are uniformly contracted under, respectively, forward and backward iterations.

The objective was twofold: to characterize structurally stable systems, and to show that most systems (an open dense set) are structurally stable. The notion of *structural stability*, introduced by Andronov, Pontrjagin, means that the whole orbit structure remains the same when the system is slightly modified: there exists a homeomorphism of the ambient manifold mapping orbits of the initial system into orbits of the modified one, and preserving the time arrow.

The first goal was most successful, as uniform hyperbolicity indeed proved to be the key ingredient characterizing structurally stable systems. At the same time, the theory provided the definitive conceptual framework for the phenomenon of transverse *homoclinic orbits*, that is, transverse intersections of stable and unstable manifolds, discovered by Poincaré in the late XIXth century, and immediately recognized by him as a major source of complexity in the dynamics.

In the process the importance of the theory of uniformly hyperbolic systems extended much beyond the original objectives. It was part of a revolution in our vision of determinism, which was strongly motivated by observations coming from the experimental sciences, where the classical opposition between deterministic and random evolutions was broken. The hyperbolic theory provided a mathematical foundation for the fact that deterministic systems, even with a small number of degrees of freedom, often present chaotic behavior, due to *robust* mechanisms. Thus, it led to the almost paradoxical conclusion that chaos may be stable.

On the other hand, structural stability and uniform hyperbolicity were soon realized to be less universal properties than was initially thought: there exist many classes of systems that are robustly unstable and non-hyperbolic. Dynamics was, once more, left without a general paradigm. As the theory was being extended in several directions – bifurcation theory, non-uniform hyperbolicity, smooth ergodic theory, – a good deal of the extraordinary progress attained in these last decades concerned specific classes of systems, such as unimodal maps of the interval, Lorenz flows, and Hénon maps.

Building on these advances, a new point of view has emerged and several ideas and results have been put forward recently, in the direction of understanding the behavior of most dynamical systems from a global viewpoint. The present work is an attempt to put such recent developments in a unified perspective, and to point likely directions of further progress. It consists of four main parts:

I - Hyperbolicity and beyond

Section 1 is a review of fundamental facts about uniformly hyperbolic systems, together with a discussion of robust mechanisms of non-hyperbolicity and other main issues involved in extending the theory beyond the hyperbolic set-up.

II - Critical behavior, homoclinic phenomena

The situations discussed in Sections 2 through 5 are united by a common mechanism generating non-hyperbolicity: tangencies or criticalities. The global dynamics of multimodal transformations in dimension 1 is determined, fundamentally, by the presence and recurrence properties of critical points. For dissipative diffeomorphisms modelled by the Hénon family, tangencies between stable and unstable manifolds play a role similar to that of critical points in dimension 1. A theory developed over the last decade, showed that these diffeomorphisms are amenable to a very complete statistical and topological description. Here and, even more explicitly, in the general context of homoclinic bifurcations of diffeomorphisms, fractal invariants have a determinant role in controlling the geometry and the recurrence of tangencies. For surface diffeomorphisms, critical behavior (homoclinic tangencies) is the only essential obstruction to hyperbolicity.

III - Dominated behavior, dynamical decomposition

Sections 6 to 9 introduce another robust mechanism of non-hyperbolicity: cycles between periodic points with variable Morse indices (unstable dimensions). One main theme is the connection between robust indecomposability of the dynamics – transitivity, ergodicity – and weaker hyperbolicity conditions – partial hyperbolicity, projective hyperbolicity, – as well as the behavior of invariant foliations – accessibility, minimality. Another, are the relations between elementary objects of the dynamics, such as homoclinic classes and maximal transitive sets. This culminates in the extension to the greatest possible generality, *tame systems*, of Smale's classical theorem on decomposition of the non-wandering set into finitely many indecomposable pieces. New mechanisms generating generic diffeomorphisms with infinitely many homoclinic classes have been devised in recent years, but the understanding of such *wild systems* remains a major challenge.

IV - Non-uniform hyperbolicity

Sections 11 and 12 focus on ergodic aspects of the theory. The notion of non-uniform hyperbolicity is somewhat vague. Classically, one considers a probability measure invariant under the system, and tries to describe the properties of the corresponding ergodic system, assuming most orbits have only non-zero Lyapunov exponents, that is, they exhibit asymptotic exponential contraction and expansion in complementary directions. We want to view non-uniform hyperbolicity also from another angle, where no invariant probability is assumed a priori. Instead, one tries to construct such probabilities, especially SRB measures, from assuming non-uniformly hyperbolic behavior on most orbits, in the sense of Lebesgue measure. Another fundamental problem is to understand how general is the assumption of non-zero Lyapunov exponents.

1. Hyperbolicity and beyond

Let us begin by reviewing some fundamental notions and facts about uniformly hyperbolic dynamical systems which one tries to extend to great generality. The expository paper of Smale [318] is a classical general reference on uniformly hyperbolic dynamics. The ergodic theory of these systems is presented by Bowen [70]. Recent surveys may be found in the books of Palis, Takens [264, Chapter 0] and Katok, Hasselblat [171, Part 4].

1.1. Spectral decomposition.

Let M be a compact manifold, and $f : M \to M$ be a diffeomorphism.

Definition 1.1. An invariant $\Lambda \subset M$ is a *hyperbolic set* for $f : M \to M$ if the tangent bundle over Λ admits a continuous decomposition

$$T_\Lambda M = E^u \oplus E^s, \tag{1}$$

invariant under the derivative and such that $\|Df^{-1} \mid E^u\| \le \lambda$ and $\|Df \mid E^s\| \le \lambda$ for some constant $\lambda < 1$ and some choice of a Riemannian metric on the manifold.

A point z is *non-wandering for f* if for every neighborhood U of z there is $n \ge 1$ such that $f^n(U)$ intersects U. The set of non-wandering points is denoted $\Omega(f)$. It contains the set $\mathrm{Per}(f)$ of periodic points, as well as the α-limit set and the ω-limit set of every orbit.

Definition 1.2. The diffeomorphism $f : M \to M$ is *uniformly hyperbolic*, or satisfies the *Axiom A*, if $\Omega(f)$ is a hyperbolic set for f and $\mathrm{Per}(f)$ is dense in $\Omega(f)$.

The definitions for smooth flows $f^t : M \to M$, $t \in \mathbb{R}$, are analogous, except that (unless Λ consists of equilibria) the decomposition (1) becomes

$$T_\Lambda M = E^u \oplus E^0 \oplus E^s, \tag{2}$$

where E^0 is 1-dimensional and collinear to the flow direction.

The spectral decomposition theorem of Smale [318] asserts that the limit set of a uniformly hyperbolic system splits into a finite number of pairwise disjoint *basic pieces* that are compact, invariant, and dynamically indecomposable. The precise statement follows.

We say that an f-invariant set is *indecomposable*, or *transitive*, if it contains some dense orbit $\{f^n(z) : n \ge 0\}$. An f-invariant set Λ is called *isolated*, or *locally maximal*, if there exists a neighborhood U of Λ such that

$$\Lambda = \bigcap_{n \in \mathbb{Z}} f^{-n}(U). \tag{3}$$

That is, Λ coincides with the set of points whose orbits remain in U for all times.

Theorem 1.3. *The non-wandering set $\Omega(f)$ of a uniformly hyperbolic diffeomorphism f decomposes as a finite pairwise disjoint union*

$$\Omega(f) = \Lambda_1 \cup \cdots \cup \Lambda_N$$

of f-invariant transitive sets Λ_i, that are compact and isolated. Moreover, the α-limit set and the ω-limit sets of every orbit are contained in some Λ_i.

Here is a sketch of the proof. Consider the equivalence relation $\sim$ defined in $\mathrm{Per}(f)$ by $p_1 \sim p_2 \Leftrightarrow$ the stable set of the orbit of each of the points has some transverse intersection with the unstable set of the orbit of the other. The stable manifold theorem implies that there are finitely many equivalence classes, and they are open in $\mathrm{Per}(f)$. The basic pieces Λ_i in the theorem are the closures of the equivalence classes. By construction, they are compact, invariant, and open in $\Omega(f) = \overline{\mathrm{Per}(f)}$. The latter implies that they are isolated sets, because $\overline{\mathrm{Per}(f)}$ is an isolated set if it is hyperbolic. Moreover, the stable and the unstable manifold of any periodic point in some Λ_i are dense in Λ_i . This implies that Λ_i is transitive. Finally, if the α- or ω-limit set of some orbit intersected more than one Λ_i , there would be non-wandering points outside the union of the basic pieces, a contradiction.

Transitive sets and isolated sets of flows $f^t : M \to M$, $t \in \mathbb{R}$, are defined in the same way. Theorem 1.3 remains true for uniformly hyperbolic flows $\{f^t : t \in \mathbb{R}\}$.

Remark 1.4. A question dating from the late sixties asked wether every hyperbolic set is contained in an isolated one. This was recently solved by Crovisier [101], who constructed a transitive diffeomorphism of $M = T^4$ having a hyperbolic set which is not contained in any isolated hyperbolic set Λ: Crovisier shows that Λ would have to be the whole torus, which is not a hyperbolic set because the diffeomorphism has saddles with different indices.

1.2. **Structural stability.** A smooth dynamical system is called *structurally stable* [10] if it is equivalent to any other system in a C^1 neighborhood. In the discrete-time case, equivalence means conjugacy by a global homeomorphism. In the case of flows this notion is too restrictive: it forces all periods of closed orbits to be preserved under perturbation. Instead, one asks for the existence of a global homeomorphism sending orbits of one system to orbits of the other, and preserving the direction of time. More generally, replacing C^1 by C^r neighborhoods, any $r \geq 1$, one obtains the notion of C^r structural stability.

The stability conjecture of Palis-Smale [262] proposes a complete characterization of structurally stable systems. Namely, they should coincide with the hyperbolic systems having the property of strong transversality: every stable and unstable manifolds of points in the non-wandering set should be transversal. In fact their conjecture is for C^r structural stability, any $r \geq 1$.

Robbin [290], de Melo [102], and Robinson [292, 293] proved that these are sufficient conditions for structural stability. Strong transversality is also necessary [291]. These results hold in the C^r topology, any $r \geq 1$. The hardest part was to prove that stable systems must be hyperbolic. This was achieved by Mañé [211] in the mid-eighties, for C^1 diffeomorphisms, and extended about ten years later by Hayashi [150] for C^1 flows. Thus

Theorem 1.5. *A C^1 diffeomorphism (or flow) on a compact manifold is structurally stable if and only if it is uniformly hyperbolic and verifies the strong transversality condition.*

A weaker property, called Ω-*stability* is defined requiring conjugacy/equivalence only restricted to the non-wandering set. The Ω-stability conjecture in [262] proposes a characterization of the Ω-stable systems: they should be the hyperbolic systems having no *cycles*, that is, no basic pieces in their spectral decompositions cyclically related by intersections of the corresponding stable and unstable sets.

The Ω-stability theorem of Smale [318] states that these properties are sufficient. The proof uses the following notion:

Definition 1.6. A *filtration* for a diffeomorphism $f : M \to M$ is a finite family $M_1, M_2, \dots, M_k$ of submanifolds with boundary and with the same dimension as M, such that

- $M_1 = M$ and M_{i+1} is contained in the interior of M_i for every $1 \leq i < k$.
- $f(M_i)$ is contained in the interior of M_i for all $1 \leq i \leq k$.

The open sets $L_i = \operatorname{int}(M_i \setminus M_{i+1})$ are the *levels of the filtration* (set $M_{k+1} = \emptyset$).

The first step is top show that if f is hyperbolic and has the no-cycles property, then it admits a filtration such that each basic piece coincides with the set of orbits contained in some level: up to reordering,

$$\Lambda_i = \bigcap_{n \in \mathbb{Z}} f^n(L_i) \qquad \text{for all } i.$$

Let g be any diffeomorphism C^r-close to f. Then $M_1, \ldots, M_k$ is also a filtration for g. Therefore, $\Omega(g)$ is contained in $\Lambda_1(g) \cup \cdots \cup \Lambda_k(g)$, where

$$\Lambda_i(g) = \bigcap_{n \in \mathbb{Z}} g^n(L_i).$$

Stability of hyperbolic sets gives that each $f \mid \Lambda_i$ is conjugate to $g \mid \Lambda_i(g)$. Then, every $\Lambda_i(g)$ is contained in $\Omega(g)$, and $f \mid \Omega(f)$ is conjugate to $g \mid \Omega(g)$. Thus f is Ω-stable, in the C^r sense.

Palis [258] proved that the no-cycles condition is necessary for Ω-stability, in any C^r topology. Necessity of hyperbolicity for Ω-stability was proved by Palis [259], based on [211], for C^1 diffeomorphisms, and extended to C^1 flows by Hayashi [150].

1.3. **Sinai-Ruelle-Bowen theory.** A basic piece Λ_i is a *hyperbolic attractor* if the stable set

$$W^s(\Lambda_i) = \{x \in M : \omega(x) \subset \Lambda_i\}$$

contains a neighborhood of Λ_i. In this case we call $W^s(\Lambda_i)$ the *basin* of the attractor Λ_i, and denote it $B(\Lambda_i)$. When the system is of class C^2, a basic piece is an attractor if and only if its stable set has positive Lebesgue measure. Thus, the union of the basins of all attractors is a full Lebesgue measure subset of M. This remains true for a residual (dense G_δ) subset of C^1 uniformly hyperbolic diffeomorphisms and flows. See Bowen [70, 71].

The following fundamental result, due to Sinai [316], Ruelle, Bowen [70, 73, 300] says that, no matter how complicated it may be, the behavior of typical orbits in the basin of a hyperbolic attractor is completely well-defined at the statistical level:

Theorem 1.7. *Every attractor Λ of a C^2 uniformly hyperbolic diffeomorphism (or flow) supports a unique invariant probability measure μ such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(z)) = \int \varphi \, d\mu \tag{4}$$

for every continuous function φ and Lebesgue almost every point $x \in B(\Lambda)$.

One way to construct μ is starting with normalized Lebesgue measure m_D over a compact domain D inside any leaf of the unstable foliation of the attractor. A distortion argument, using the fact that f is C^2, gives that the sequence of iterates $\{f_*^n(m_D) : n \geq 1\}$ has a property of *uniform* absolute continuity along the unstable foliation. Contracting behavior in the transverse direction, together with minimality of the unstable foliation, are used to show that the limit $\mu = \lim_{n\to\infty} f_*^n(m_D)$ exists and is an ergodic measure whose support coincides with Λ. Very important, because of the previous observation μ disintegrates into conditional measures along unstable foliation (see Appendix C) that are equivalent to the Lebesgue measure of each leaf.

Ergodicity gives (4) for μ-almost every point in Λ. Using conditional measures we get it for a full Lebesgue measure subset L of some unstable leaf. To prove the much more interesting fact that (4) is true for Lebesgue almost every point in the basin of attraction, a whole open set, one uses the observation that time-averages of continuous functions are constant on stable manifolds. The stable manifolds of points in the attractor foliate the whole $B(\Lambda)$. Absolute continuity of this stable foliation (see Appendix C) ensures that the stable manifolds of the points in L cover a full Lebesgue measure of the basin. Theorem 1.7 follows.

Property (4) means that the Sinai-Ruelle-Bowen measure μ may be explicitly computed, meaning that the weights of subsets may be found with any degree of precision, as the sojourn-time of any orbit picked "at random" in the basin of attraction:

$$\mu(V) = \text{ fraction of time the orbit of } z \text{ spends in } V$$

for any typical subset V of M (the boundary of V should have zero μ-measure), and for Lebesgue almost any point $z \in B(\Lambda)$. For this reason μ is called a *physical measure*.

There is another sense in which this measure is "physical" and that is that μ is the zero-noise limit of the stationary measures associated to the stochastic processes obtained by adding small random noise to the system. This property is called *stochastic stability*; a formal definition will appear later. For uniformly hyperbolic systems it is due to Sinai [316], Kifer [177, 179], and Young [351]. The model of small stochastic perturbations to represent external influences, too small or too complex to express in deterministic terms, goes back to Kolmogorov and Sinai.

1.4. **Heterodimensional cycles.** Although uniform hyperbolicity was originally intended to encompass a residual, or at least dense subset of all dynamical systems, it was soon realized that this is not true. There are two main mechanisms that yield robustly non-hyperbolic behavior, that is, whole open sets of non-hyperbolic systems. Not surprisingly, they are at the heart of recent developments that we are going to review in the next sections.

Historically, the first one was the coexistence of periodic points with different Morse indices (dimensions of the unstable manifolds) inside the same transitive set. See Figure 1. This is how the first examples of C^1-open subsets of non-hyperbolic diffeomorphisms were obtained by Abraham, Smale [2, 315] on manifolds of dimension $d \geq 3$. It was also the key in the constructions by Shub [312] and Mañé [205] of non-hyperbolic yet robustly transitive diffeomorphisms, that is, such that every diffeomorphism in a C^1 neighborhood has a dense orbit. The examples of Shub and Mañé are outlined in Section 7.3.

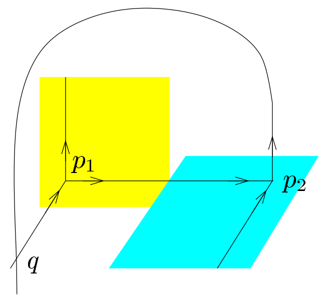

FIGURE 1. A heterodimensional cycle

For flows this mechanism may assume a novel form, because of the interplay between regular periodic orbits and singularities (equilibrium points). That is, robust non-hyperbolicity may stem from the coexistence of regular and singular orbits in the same transitive set. The first and very striking example was the geometric

Lorenz attractor proposed by Afraimovich, Bykov, Shil'nikov [3] and Guckenheimer, Williams [143, 344] to model the behavior of the Lorenz equations [195]. This is a main topic in Section 9.

1.5. **Homoclinic tangencies.** Heterodimensional cycles may exist only in dimension 3 or higher. The first robust examples of non-hyperbolic diffeomorphisms on surfaces were constructed by Newhouse [240], exploiting the second of the mechanisms we mentioned: homoclinic tangencies, or non-transverse intersections between the stable and the unstable manifold of the same periodic point. See Figure 2.

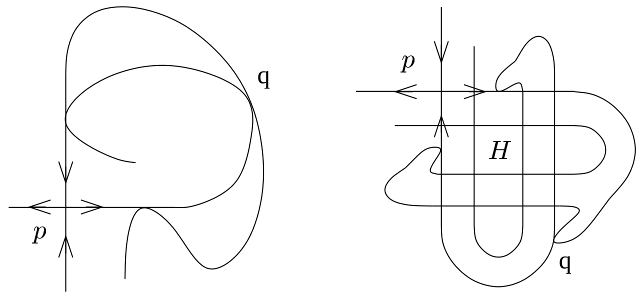

FIGURE 2. Homoclinic tangencies

It is important to observe that individual homoclinic tangencies are easily destroyed by small perturbations of the invariant manifolds. To construct open examples of surface diffeomorphisms with *some* tangency, Newhouse started from systems where the tangency is associated to a periodic point inside an invariant hyperbolic set with rich geometric structure. See the right hand side of Figure 2. His argument requires a very delicate control of distortion, as well as of the dependence of the fractal dimension on the dynamics. Actually, for this reason, his construction is restricted to the C^r topology for $r \geq 2$. By comparison, robust examples of heterodimensional cycles in higher dimensions are obtained by much more elementary transversality and hyperbolicity arguments.

1.6. **Attractors and physical measures.** Several attempts were made, specially in the seventies, to weaken the definitions of hyperbolicity and structural stability, while keeping their topological flavor, so that they could encompass a residual set of dynamical systems. One such extension was the notion of Ω-stability that we mentioned before.

In parallel, a more probabilistic approach was being proposed, especially by Sinai, Ruelle, Eckmann, where one focus on the statistical behavior of typical orbits, and its stability under perturbations. See [126] for a detailed exposition.

Definition 1.8. A set $\Lambda \subset M$ is an *attractor* for a diffeomorphism (or a flow) on a manifold M if it is invariant and transitive, and the *basin of attraction*,

$$B(\Lambda) = \{x \in M : \omega(x) \subset \Lambda\}$$

has positive Lebesgue measure.

Definition 1.9. A *physical measure*, or *SRB measure*, of a diffeomorphism on a manifold M is an invariant probability measure μ on M, such that the time average of every continuous function $\varphi : M \to \mathbb{R}$ coincides with the corresponding space-average with respect to μ,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(z)) = \int \varphi \, d\mu \tag{5}$$

for a set of initial points z with positive Lebesgue measure. We call this set the *basin* of μ, and denote it $B(\mu)$.

For flows (5) is replaced by

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi(f^t(z)) \, dt = \int \varphi \, d\mu.$$

Problem 1.10. For most dynamical systems, do time averages converge Lebesgue almost everywhere ? Are there SRB measures whose basins cover almost all M ?

This is the case for C^2 hyperbolic systems, cf. Theorem 1.7. But the answer can not always be affirmative. A simple counter-example, due to Bowen, is described in Figure 3: time averages diverge over any of the spiraling orbits in the region bounded by the saddle connections. However, no robust counter-example is known.

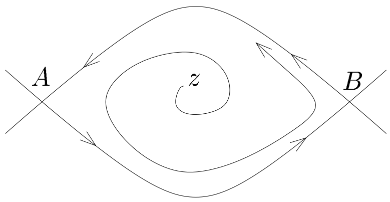

FIGURE 3. A planar flow with divergent time averages

The following semi-global version of the previous problem also goes back to Sinai and Ruelle:

Problem 1.11. (basin problem) Let Λ be an attractor for a diffeomorphism, or a flow, supporting a unique SRB measure μ. Does

$$B(\mu) = B(\Lambda) \quad \text{up to a zero Lebesgue measure set ?} \tag{6}$$

More generally, does $B(\Lambda)$ coincide with the union of the basins of the SRB measures supported in Λ, up to zero Lebesgue measure ?

Again, the answer can not always be positive: [31] shows how Bowen's example can be inserted into an attractor, so that the basin property (6) is violated. Note also that general (transitive) attractors may support more than one SRB measure. See Kan [168] and Section 11. But no robust examples of either of these phenomena have been found to date.

1.7. **A conjecture on finitude of attractors.** A program towards a global understanding of complex dynamical behavior has been proposed by Palis [260]. Here we quote some of the conjectures embodying his program, others will appear later.

Conjecture 1.12. (finitude of attractors [260]) There is a C^r, $r \geq 1$, dense set subset $\mathcal{D}$ of dynamical systems on any compact manifold that exhibit a finite number of attractors whose basins cover Lebesgue almost all of the manifold.

Conjecture 1.13. (physical measures and stochastic stability [260]) For any element of $\mathcal{D}$ all the attractors support SRB measures and have the basin property. Moreover, any element of $\mathcal{D}$ is stochastically stable on the basin of each of the attractors.

Stochastic stability means that adding small random noise to the system has little effect on its statistical behavior. For discrete-time systems $f : M \to M$, one considers *random orbits* $\{z_j : j \geq 0\}$ where each z_{j+1} is chosen at random in the ε-neighborhood of $f(z_j)$. Then, for ε small, the time averages of continuous functions over almost every random orbit should be close to the corresponding time averages over typical orbits of the original f. There is a corresponding notion for flows, where noise takes place at infinitesimal intervals of time: random orbits are solutions of a stochastic differential equation. Precise definitions are given in Appendix D.

Conjecture 1.14. (metric stability of basins of attraction [260]) For any element of $\mathcal{D}$ and any of its attractors, almost all C^r small perturbations along generic k-parameter families of dynamical systems exhibit finitely many attractors supporting physical measures and whose basins cover most (a fraction close to 1 in volume) of the original basin of attraction.

A key novelty in the formulation of this conjecture, and in the whole scenario proposed by Palis, is to allow the existence of pathological phenomena, e.g. related to cycles, occupying a small volume in the ambient space. Indeed, cycles have been a main obstruction to the realization of previous global scenarios for Dynamics.

For one-dimensional systems these conjectures take a stronger form: finiteness of attractors, with the nice properties above and with their basins containing Lebesgue almost all points, should correspond to full Lebesgue measure in parameter space, for generic parametrized families. This has been fully verified for C^2 unimodal maps of the interval, as we shall see in Section 2. In this case the attractor is unique. There is also substantial partial progress in higher dimensions, some of which is reported throughout the text.

2. One-dimensional dynamics

Here we only review certain results, especially about typical dynamics inside families of maps, that relate more closely to the general topics of hyperbolicity and homoclinic phenomena. Surveys of much recent progress in one-dimensional dynamics have been written by Graczyk, Swiatek [141] and Lyubich [203], and the book of de Melo, van Strien [103] remains the fundamental expository text on the subject.

Throughout this section we consider C^r maps $f : M \to M$, where M is either the circle S^1 or the compact interval, and $r \geq 1$. In the interval case we assume that $f(\partial M) \subset \partial M$ and any fixed point on the boundary is repelling.

2.1. **Hyperbolicity.** The map f is called *hyperbolic* if the periodic attractors are all hyperbolic, and the complement $\Sigma(f)$ of the basins of attraction is a hyperbolic set for f. The latter means that there are $C > 0$ and $\sigma > 1$ such that $|Df^n(x)| \geq C\sigma^n$ for every $x \in \Sigma(f)$ and $n \geq 1$. The definition is in terms of $\Sigma(f)$, instead of the non-wandering set $\Omega(f)$, because the latter may not be fully invariant when the map is not a diffeomorphism.

For C^2 maps without critical points, a satisfactory characterization of hyperbolicity is provided by the following

Theorem 2.1 (Mañé [208]). *Let $f : M \to M$ be a C^2 map of the circle or the interval, without critical points and whose periodic points are all hyperbolic. Then either f is hyperbolic or f is (conjugate to) an irrational rotation of $M = S^1$.*

Hyperbolicity of all periodic points is, of course, a necessary condition for hyperbolicity of the map. Among circle diffeomorphisms, the hyperbolic ones form an open dense subset. In contrast, by a celebrated theorem of Herman [155], irrational rotations correspond to positive Lebesgue measure inside generic parametrized families. This kind of dichotomy is not uncommon in one-dimensional dynamics, as we shall see.

Example 2.2. The Arnold family $h_{a,b} : S^1 \to S^1$ of circle maps is defined by [23]

$$h_{a,b}(\theta) = \theta + a + b \sin 2\pi\theta \mod \mathbb{Z}. \tag{7}$$

For $|b| < 1/2\pi$, the map $h_{a,b}$ is a diffeomorphism. There is an open dense subset of values of a for which the rotation number is rational. However, by Herman's theorem, for a positive Lebesgue measure subset of values of a, the diffeomorphism $h_{a,b}$ is differentiably conjugate to an irrational rotation.

For $|b| = 1/2\pi$ the map develops a critical point and for $|b| > 1/2\pi$ it is no longer a homeomorphism. These situations have also been intensively studied, and found applications in higher dimensions, e.g. in [113, 246]. See [103, Section I.7] for references and more information on the Arnold family.

If f is hyperbolic and C^2, the set $\Sigma(f)$ has zero Lebesgue measure, unless it coincides with the whole ambient space. The second case is possible only if $M = S^1$ and f has no critical points. Moreover, f is topologically conjugate to a linear covering map of the circle [311], and there is a unique SRB measure, which is equivalent to Lebesgue measure. In the first case, the typical dynamics is really very simple: almost every orbit just converges to some periodic attractor; the SRB measures are the Dirac measures supported on those attractors.

In general, hyperbolic maps are structurally stable, that is, they are topologically conjugate to all nearby maps. See [103, Chapter 3]. The converse is not known in general, but it may be deduced in some situations, e.g., whenever the subset of hyperbolic maps is known to be dense.

We call $f : M \to M$ a *multimodal map* if it has a finite, non-zero number of critical points, all non-flat (finite-order contact). *Unimodal* means that the critical point is unique. *Infinite-modal* maps are studied in [257].

Example 2.3. The main model of unimodal maps is the quadratic family

$$q_a : [0,1] \to [0,1], \quad q_a(x) = ax(1-x),$$

where a is a parameter in $(0, 4]$. For all a, the critical point $c = 1/2$ is quadratic: $D^2 q_a(c) \neq 0$. Quadratic maps are sometimes written in different, but conjugate, forms e.g. $x \mapsto 1 - ax^2$ or $x \mapsto x^2 + a$.

The dynamics of multimodal maps is strongly influenced by the behavior of critical orbits. This is already apparent in the following remarkable result, that extends Theorem 2.1:

Theorem 2.4 (Mañé [208]). *Let $f : M \to M$ be a C^2 multimodal map of the circle or the interval. If all the periodic points are hyperbolic and every critical point is contained in the basin of a periodic attractor, then f is hyperbolic.*

In fact, Mañé proves a stronger statement: A compact invariant set of a C^2 multimodal map is hyperbolic if (and only if) it contains neither critical points nor non-repelling periodic points. See Theorem 2.5 below.

In the special case of maps with negative Schwarzian derivative, including quadratic maps, every non-hyperbolic periodic point is either an attractor or a semi-attractor and, in either case, the basin of attraction contains some critical point. See Singer [317] and [103, Section III.3]. Therefore, the map is hyperbolic if and only if every critical orbit is contained in the basin of some hyperbolic attractor. Moreover, if the map is unimodal, hyperbolicity is equivalent to existence of a hyperbolic attractor.

2.2. **Non-critical behavior.** Theorems 2.1 and 2.4 may be seen as predecessors to results of Pujals, Sambarino [286] about surface diffeomorphisms, that we are going to review in Section 5. We sketch their proofs, because ideas from them proved to be fruitful in higher dimensions too. The following statement contains both theorems:

Theorem 2.5 (Mañé [208]). *Let f be a C^2 map of the circle or the interval, and Λ be an f-invariant compact set containing no critical points of f. Then*

(a) *the periods of the non-hyperbolic periodic points of f in Λ are bounded above, and all such points are contained in finitely many periodic intervals, restricted to which f is a diffeomorphism (permuting the intervals);*
(b) *if Λ does not contain any non-hyperbolic periodic point then Λ is a hyperbolic set, or else $\Lambda = S^1$ and f is conjugate to an irrational rotation.*

We start the sketch of the proof with an alternative characterization of hyperbolicity of an invariant set Λ of a C^1 one-dimensional map $f : M \to M$, in terms of the *natural extension*

$$\hat{f} : \hat{\Lambda} \to \hat{\Lambda}, \quad \hat{f}(\ldots, x_n, \ldots, x_0) = (\ldots, x_n, \ldots, x_0, f(x_0))$$

of f restricted to Λ. Here $\hat{\Lambda}$ is the set of all pre-orbits of points in Λ, that is, all sequences $(x_n)_{n \geq 0}$ in Λ such that $f(x_{n+1}) = x_n$ for all $n \geq 0$.

Lemma 2.6. *Let $f : M \to M$ be a C^1 map and Λ be a compact f-invariant set such that*

$$\lim_{n \to \infty} |(f^n)'(x_n)| = \infty$$

for all $\hat{x} \in \hat{\Lambda}$. Then Λ is a hyperbolic set for f.

A *pre-orbit* of an open interval $I \subset M$ is a sequence $(\phi_n)_n$ where each ϕ_n is a branch of $f^{-n} \mid I$ satisfying $f \circ \phi_{n+1} = \phi_n$ for all n. The pre-orbit is *associated* to the invariant set Λ if there is $x \in I \cap \Lambda$ such that $\phi_n(x) \in \Lambda$ for all $n \geq 0$. An interval $I \subset M$ is *adapted to* Λ if, for any $x \in I \cap \Lambda$ and any pre-orbit $(x_n)_n$ of $x = x_0$ there is a pre-orbit $(\phi_n)_n$ of I such that $\phi_n(x) = x_n$, the intervals $\phi_n(I)$ are uniformly far from the critical points of f, and each $\phi_n(I)$ is either contained or disjoint from I.

Lemma 2.7. *Consider a C^2 map $f : M \to M$ which is not topologically equivalent to an irrational rotation. Let Λ be a compact invariant subset that does not meet the boundary of M nor contains critical points of f.*

If $\cap_{n\geq 0} f^n(\Lambda)$ contains non-periodic points, then there is an adapted interval J and a constant $K > 0$ such that every pre-orbit $(\phi_n)_n$ of J associated to Λ satisfies

$$\sum_{i=1}^{\infty} |\phi_i'(x)| \leq K \quad \textit{and} \quad |\phi_i'(x)| \leq K \ |\phi_i'(y)|$$

for any points x and y in J and every $i \geq 1$. Moreover, for every non-periodic point $x \in \cap_{n\geq 0} f^n(\Lambda)$ there is an interval J adapted to Λ and containing x.

The two lemmas above imply that every f-invariant compact set Λ that does not contain critical points, sinks or non-hyperbolic periodic points is hyperbolic or f is equivalent to an irrational rotation. The proof is by contradiction as follows.

Assume f is not a rotation and Λ is not hyperbolic. Consider the family $\mathcal{N}$ of non-hyperbolic compact f-invariant subsets of Λ, ordered by inclusion. The intersection of the sets of any totally ordered sequence in $\mathcal{N}$ is also non-hyperbolic, thus it belongs to $\mathcal{N}$. So, by Zorn's lemma, the family $\mathcal{N}$ has some minimal element Λ_0. The compact set Λ_0 can not be a union of periodic orbits: otherwise, it would be hyperbolic. We also claim that $\Lambda_0 = f(\Lambda_0)$: if $f(\Lambda_0)$ were properly contained in Λ_0, then it would be hyperbolic, which would imply the hyperbolicity of Λ_0. This shows that the set $\cap_n f^n(\Lambda_0) = \Lambda_0$ contains some non-periodic point. Let J be an interval adapted to Λ_0 as given by Lemma 2.7. From the existence of such an interval, we are going to deduce that Λ_0 is hyperbolic. This contradiction will show that the family $\mathcal{N}$ must be empty, and so Λ must be hyperbolic.

By Lemma 2.6, it suffices to prove that $\lim_{n\to\infty} |(f^n)'(x_n)| = \infty$ for every pre-orbit $(x_n)_n \in \hat{\Lambda}_0$ of a point $x_0 \in \Lambda_0$. Denote by Γ the closure of the sequence $(x_n)_n$. This is a compact invariant set. If Γ is properly contained in Λ_0 then it is hyperbolic, and so $\lim_n |(f^n)'(x_n)| = \infty$. Otherwise, $\Gamma = \Lambda_0$ and so there is $k \geq 0$ such that $x_k \in J$. Replacing x by x_k, we may suppose $k = 0$. Since J is adapted to Λ_0, it has an (associated) pre-orbit $(\phi_n)_n$ such that $\phi_n(x_0) = x_n$ for all n. The summability property in Lemma 2.7, implies that $\lim_n |\phi_n'(x_0)| \to 0$. Since ϕ_n is an inverse branch of f^n with $\phi_n(x_0) = x_n$, this means that $\lim_n |(f^n)'(x_n)| \to \infty$. This ends our sketch of the proof of Theorem 2.5.

2.3. **Density of hyperbolicity.** The set of hyperbolic maps is C^r open in the space of C^r multimodal maps. For $r = 1$, it was proved by Jakobson [165] that this set is also dense. Whether this remains true for $r \geq 2$ has been one of the main problems in one-dimensional dynamics:

Conjecture 2.8. The set of hyperbolic maps is C^r dense in the space of multimodal maps, for any $r \geq 2$.

A classical transversality argument, combined with Theorem 2.4, shows that any multimodal map whose critical points are all contained in basins of hyperbolic periodic attractors is C^r approximated by a hyperbolic one. So, Conjecture 2.8 is equivalent to showing that the set of multimodal maps whose critical points are all in basins of hyperbolic attractors is C^r dense.

A major step was the solution of this problem for quadratic maps:

Theorem 2.9 (Graczyk, Swiatek [140], Lyubich[201]). *The set of parameters a for which $q_a(x) = ax(1-x)$ is hyperbolic is dense in $(0, 4]$.*

The core of the proofs was to show that any two quadratic maps whose periodic points are all repelling are quasi-symmetrically conjugate if they are topologically conjugate. A homeomorphism h is *quasi-symmetric* if there exists $K > 0$ such that any two adjacent intervals of equal length are mapped by h onto intervals whose lengths differ by a factor $\leq K$. The arguments also relied on substantial information specific to the quadratic family, especially the fact that the topological entropy or, equivalently, the kneading invariant [219] of q_a depends monotonically on the parameter a.

So, it came as a surprise when Kozlovski announced, only short afterwards, that he could extend the conclusion to the general context of unimodal maps:

Theorem 2.10 (Kozlovski [181, 182, 183]). *The set of hyperbolic maps is C^r dense in the space of all unimodal maps, for any $r \geq 3$.*

Thus, Smale's original conjecture that hyperbolicity should be topologically prevalent was confirmed in this low-dimensional setting! Density of structural stability among unimodal maps is easier and was known before. See [103, section III.2].

Concerning general multimodal maps, Shen [309] recently proved C^2 density of hyperbolicity. At the time of writing the cases $r > 2$ remain open.

2.4. **Chaotic behavior.** If a multimodal map f is hyperbolic and C^2, almost every orbit converges to some periodic attractor, and the whole dynamics is quite simple. A theorem of Jakobson states that maps with complicated (non-hyperbolic) behavior are also abundant, in the probabilistic sense, in one-dimensional dynamics:

Theorem 2.11 (Jakobson [165]). *For an open class of families of C^2 unimodal maps, including the quadratic family, the set of parameters for which the map has an absolutely continuous invariant probability μ has positive Lebesgue measure.*

This remarkable result opened the way to much progress in non-uniformly hyperbolic dynamics.

A general strategy to construct an absolutely continuous invariant measure is to consider the forward iterates $f_*^n m$ of Lebesgue measure m. Any Cesaro weak limit of a subsequence is an invariant measure. To ensure that some of these limit measures is absolutely continuous, one must show that iterates of f do not distort Lebesgue measure too much. For multimodal maps, distortion occurs, primarily, close to the critical points.

Jakobson exploited the method of *inducing*. He showed how to find, for Lebesgue almost every point x, an integer $n(x)$ such that during the time interval $[0, n(x)]$ iterates $f^i(x)$ close to the critical point c are compensated for by iterates far from c, so that $Df^{n(x)}(x)$ is larger than 1 in norm. More precisely, he constructed a countable partition $\{I_j\}$ of a full Lebesgue measure subset of M into intervals, such that $n(\cdot)$ is constant on each I_j, and every $f^{n(I_j)} \mid I_j$

- maps I_j onto some fixed interval J, which coincides with the union of a subfamily of I_j's (Markov property);
- is uniformly expanding and has bounded distortion: the logarithm of the derivative is uniformly Lipschitz.

The construction is by induction, and requires parameter exclusions at each step. A crucial part of the proof is to show that the exclusions decrease sufficiently fast, so that a set of parameters with positive Lebesgue measure remains at the end.

The restriction g of $f^{n(\cdot)}$ to J is a *Markov map*. A well-known result states that every Markov map has an absolutely continuous invariant measure ν, which is the limit of forward iterates of Lebesgue measure. Then

$$\mu_0 = \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} f_*^i(\nu | \{n(\cdot) = k\})$$

defines an f-invariant measure which is also absolutely continuous. To ensure that μ_0 is finite, one needs the set of points with large inducing time $n(\cdot)$ to have small measure:

$$\sum_{k=1}^{\infty} k\, \nu(\{n(\cdot) = k\}) < \infty. \tag{8}$$

Actually, it suffices to check the corresponding property for Lebesgue measure m, because ν has bounded density. Finally, $\mu = \mu_0/\mu_0(M)$ is the absolutely continuous invariant probability of f.

We call a unimodal map *chaotic* if it has an absolutely continuous invariant probability. For unimodal maps with negative Schwarzian derivative, such a probability is unique and ergodic, and so it is an SRB measure. See Blokh, Lyubich [48] and [103, Section V.1].

Jakobson's theorem was much extended, and alternative approaches have been proposed for proving chaotic dynamics in dimension 1. See [103, Chapter V] and references therein. The first of these, due to Collet, Eckmann [92], focus on proving that the map is exponentially expanding along the critical orbit: there are $C > 0$ and $\sigma > 1$ such that

$$|Df^n(c_1)| \geq C\sigma^n \quad \text{for all } n \geq 1, \tag{9}$$

where $c_1 = f(c)$ is the critical value. Indeed, in [93] they showed that, for unimodal maps with negative Schwarzian derivative, condition (9) implies the existence of absolutely continuous invariant probability. As a matter of fact, Collet, Eckmann needed a second condition, involving the backward orbit of the critical point, but this was shown by Nowicki [248] to be a consequence of the first one. This was further improved by Nowicki, van Strien [249]: if ℓ denotes the order of the critical point, the summability condition

$$\sum_n |Df^n(c_1)|^{-1/\ell} < \infty, \tag{10}$$

suffices for existence of an absolutely continuous invariant probability.

2.5. **The renormalization theorem.** Much of the impressive development of the theory of unimodal maps over the last years had to do with understanding the partition of the space into classes of topological conjugacy, and how this relates with finer equivalence classes, like real-analytic, differentiable, or quasi-symmetric conjugacy (rigidity phenomena). Two main tools were the kneading theory of Milnor, Thurston [219], which reduces essentially all the topological dynamics to symbolic terms, and renormalization theory, which provides an extremely efficient way to deal with strongly non-linear behavior in small scales.

The idea of renormalization comes from Physics, and was proposed by Coullet, Tresser [97] and by Feigenbaum [130] to explain the so-called universality phenomena they observed in bifurcation cascades of unimodal maps. Renormalization operators proved to be powerful tools to reveal the fine geometric structure of phase space, not only for these systems but also in other situations of Dynamics, like critical circle maps, the Herman-Yoccoz theory of circle diffeomorphisms, or rational transformations of the Riemann sphere. In the unimodal setting, they led to the proof of universal geometries in phase and in parameter spaces. More recently, generalized renormalization techniques have been very fruitful in the analysis of statistical properties of these maps.

A unimodal map f is *renormalizable* if there is $p > 1$ and a family $I_0, \ldots, I_{p-1}$, $I_p = I_0$ of compact disjoint intervals such that I_0 contains the critical point in its interior, and $f(I_{j-1}) \subset I_j$ for every $1 \le j \le p$. The *renormalization operator* $f \mapsto \mathcal{R}(f)$ is defined, in the subset of renormalizable maps, by

$$\mathcal{R}(f) = \phi \circ f^p \circ \phi^{-1}, \tag{11}$$

where one chooses p and I_0 smallest as above, and $\phi : I_0 \to M$ is the affine rescaling of I_0 to the original interval M. Then $\mathcal{R}(f)$ is also a unimodal map. We say that f is *infinitely-renormalizable* if $\mathcal{R}^n(f) = (\mathcal{R} \circ \cdots \circ \mathcal{R})(f)$ is defined for all $n \ge 1$.

The renormalization operator is much better behaved in the real-analytic case: in particular, it is differentiable and injective. In addition, successive renormalizations of any C^2 map come closer and closer to real-analytic maps. See [103, Theorem VI.1.1]. For these reasons, one often considers the restriction of $\mathcal{R}$ to real-analytic unimodal maps. In what follows we also restrict ourselves to the case of quadratic critical point. The following statement is a global reformulation of the universality conjectures of Coullet, Tresser and Feigenbaum:

Conjecture 2.12 (Renormalization Conjecture)**.** The renormalization operator $\mathcal{R}$ has an invariant uniformly hyperbolic set $\mathcal{H}$ consisting of infinitely renormalizable maps, with one-dimensional expanding direction and topological conjugacy classes as (codimension-one) stable manifolds, and such that the restriction of $\mathcal{R}$ to the set $\mathcal{H}$ is a homeomorphism conjugate to the full two-sided shift with countably many symbols.

The proof of this conjecture was a major achievement in the whole theory of one-dimensional dynamics. Sullivan [322] characterized the stable manifolds as conjugacy classes and proved sub-exponential contraction, in the case of bounded combinatorics (bounded p). McMullen [217] proved that the contraction is indeed exponential. Lyubich [200, 202] proved exponential expansion along the unstable direction, and treated also the unbounded case, thus completing the proof.

2.6. **Statistical properties of unimodal maps.** Hyperbolic sets of sufficiently smooth systems in finite dimension have zero volume, unless they coincide with the whole ambient. Thus, although $\mathcal{R}$ acts in infinite dimension, hyperbolicity of the renormalization horseshoe $\mathcal{H}$ may motivate the following (non-trivial) consequence proved by Lyubich [200]: the set of infinitely renormalizable quadratic maps has zero Lebesgue measure in parameter space.

On the other hand, Lyubich [204] and Martens, Nowicki [215] prove that Lebesgue almost all finitely renormalizable quadratic maps without periodic attractors satisfy the summability condition (10), and so they have an absolutely continuous invariant probability. Combining these two results one gets the following remarkable conclusion:

Theorem 2.13 (Lyubich [200])**.** *For Lebesgue almost every $a \in (0,4]$, the quadratic map q_a is either hyperbolic or chaotic.*

Recently, Avila, Lyubich, de Melo [26] extended the conclusion of Theorem 2.13 to most real-analytic families of unimodal maps. The key to this extension is the following structure theorem for the space of non-hyperbolic maps:

Theorem 2.14 (Avila, Lyubich, de Melo [26])**.** *Non-hyperbolic topological conjugacy classes are codimension-one submanifolds of the space of real-analytic unimodal maps, with the possible exception of countably many classes. The lamination defined by these submanifolds is transversely quasi-symmetric: its holonomy maps extend to quasi-symmetric homeomorphisms.*

The possible exceptional cases in the statement correspond to the existence of indifferent periodic points. The quadratic family is everywhere transverse to the laminae. Thus, using the holonomies, one can transfer information from the quadratic family to any other that is fairly transverse to the lamination. In this way, Avila, Lyubich, de Melo [26] prove that for every *non-trivial* family $(f_a)_a$, that is, any family such that no parameter interval is contained in the same non-hyperbolic conjugacy class, almost every map f_a is either hyperbolic or chaotic.

Let us point out that this extension of Theorem 2.13 is not a direct consequence of the statement of Theorem 2.14, one must appeal to the proof itself. One reason is that *the lamination is not transversely absolutely continuous:* the holonomies do not preserve the class of zero Lebesgue measure sets. We shall return to this point in a while.

In a series of very recent works, Avila, Moreira enhance this strategy, to provide a very complete picture of the statistical properties of unimodal maps. To begin with, they prove

Theorem 2.15 (Avila, Moreira [25, 28, 29])**.** *Let $(f_a)_a$ be a non-trivial real-analytic family of unimodal maps. Then, for Lebesgue almost every parameter a, either f_a is hyperbolic or it satisfies the Collet-Eckmann condition* (9) *together with slow recurrence of the critical orbit: given any $\delta > 0$,*

$$|f_a^n(c) - c| > n^{-(1+\delta)} \quad \textit{for every large } n. \tag{12}$$

They deduce that almost every f_a has a finite number of SRB measures (exactly one if the Schwarzian derivative is negative), after proving that f_a has a renormalization conjugate to a quadratic map. For hyperbolic parameters, almost every point is in the basin of some periodic attractor. In the non-hyperbolic case, there exist $N \geq 1$ and an interval J such that $F_a = f_a^N$ defines a non-renormalizable unimodal map from J to J, and there is an absolutely continuous invariant probability μ_a supported inside the orbit of J. Almost every point in J is in the basin of μ_a, and almost every orbit that never enters J is in the basin of some periodic attractor.

Combining Theorem 2.15 and results of Keller, Nowicki [176] and Young [352], Avila, Moreira conclude that (F_a, μ_a) has exponential decay of correlations in the space of observables with bounded variation, for almost all non-hyperbolic parameter a. Actually, [176, 352] assume negative Schwarzian derivative, but this can be by-passed using the Collet-Eckmann condition.

It had been shown by Baladi, Viana [33] that the Collet-Eckmann condition and a recurrence condition weaker than (12) imply stochastic stability, in a strong sense: L^1 convergence of densities. They considered additive absolutely continuous noise. Weak stochastic stability was proved in [44, 172], under somewhat different hypotheses. In view of the easy fact that hyperbolic maps are stochastically stable, Theorem 2.15 and [33] imply that almost all maps f_a are strongly stochastically stable. Thus, the conjectures of Palis from Section 1.7 are fully verified in the context of unimodal maps!

In fact, these results extend to generic C^r families of unimodal maps, any $r \geq 2$. The description of the generic set is less explicit in this case, but Avila, Moreira still get that, for such families and for Lebesgue almost every parameter, either f_a is hyperbolic or it satisfies the Collet-Eckmann condition and has sub-exponential recurrence of the critical orbit. See [25]. Then, f_a has finitely many SRB measures, with stochastic stability (strong if $r \geq 3$) and exponential decay of correlations.

The next theorem answers another old question:

Theorem 2.16 (Avila, Moreira [27]). *Let $(f_a)_a$ be a non-trivial real-analytic family of unimodal maps. For Lebesgue almost every non-hyperbolic parameter a, the critical point is in the basin of the absolutely continuous invariant measure μ_a, and*

$$\lim_n \frac{1}{n} \log |Df^n(c_1)| = \int \log |Df| \, d\mu_a,$$

that is, the Lyapunov exponent of the critical orbit coincides with that of μ_a.

This leads to the following surprising conclusion [27]: for almost every map in any non-trivial real-analytic family, *the eigenvalue of every repelling periodic orbit in the support of μ_a is determined, through an explicit formula, by the itineraries (kneading sequences) of the periodic orbit and the critical orbit.*

Here is a quick explanation of the connection between eigenvalues and itineraries. For a periodic point p in the support of μ_a one considers neighborhoods V_k of p such that each V_k is mapped to V_{k-1} by the per(p):th iterate. On the one hand, each $\mu_a(V_k)$ may be computed as the frequency of visits of $f^n(c)$ to V_k, because the critical point c is in the basin of μ_a, and that is readily translated as the frequency inside the itinerary of the critical orbit of periodic blocks duplicating the itinerary of p. On the other hand, using absolute continuity of the SRB measure, one shows that $\mu_a(V_k)^{-1/k}$ converges to the eigenvalue of p when $k \to \infty$. This would follow from elementary arguments if $\mu_a(J)$ were proportional to length $|J|$ for every interval J in the support. However, the density is not bounded from above and, in general, we only have $c_1|J| \leq \mu_a(J) \leq c_2|J|^{1/2}$, which is not enough for the convergence statement. In short terms, one uses certain intervals related to gaps of hyperbolic Cantor sets, satisfying $\mu_a(J) \leq C_\delta |J|^{1-\delta}$ with $\delta > 0$ arbitrarily small.

Topologically conjugate unimodal maps are real-analytically conjugate if they have the same eigenvalues on corresponding periodic orbits. Therefore, the previous result implies that, for a set $\mathcal{E}$ of unimodal maps intersecting every non-trivial family on a full measure subset, *any two maps in $\mathcal{E}$ are topologically conjugate if and* *only if they are real-analytically conjugate on the support of their SRB measures.* This striking rigidity statement means that the "full measure" subset $\mathcal{E}$ intersects each topological conjugacy class in a subset with infinite codimension (contained in a single real-analytic conjugacy class!). This shows that the lamination into topological conjugacy classes is highly non-absolutely continuous.

3. Homoclinic tangencies

A detailed presentation of the theory of homoclinic bifurcations at the beginning of the nineties is given in the book of Palis, Takens [264]. Here we recall some fundamental facts and focus on developments that took place over the last decade.

A point is called *homoclinic* if it is in the intersection of the stable manifold and the unstable manifold of the same hyperbolic periodic point p. One speaks of *homoclinic tangency* if the intersection is not transverse. See Figure 2 for two examples; in the second one p is part of a non-trivial hyperbolic set H.

A first fundamental question is:

Problem 3.1. (frequency of hyperbolicity) How common are hyperbolic diffeomorphisms close to one exhibiting a homoclinic tangency ?

A geometric approach developed by Newhouse, Palis, Takens over the seventies and the eighties, led to a surprisingly precise answer to this problem for surface diffeomorphisms, formulated in terms of fractal invariants such as the Hausdorff dimension.

3.1. Homoclinic tangencies and Cantor sets.

The basic strategy is to use invariant foliations to translate the dynamical problem, usually with some degree of simplification, into questions about certain dynamically defined Cantor sets.

Let $f : M \to M$ be a diffeomorphism with a homoclinic tangency associated to a hyperbolic periodic point p, and H be a hyperbolic basic set of f containing p. That is, H is transitive, isolated, and periodic points are dense in it. In the simplest case H is just the orbit of p. We take all maps to be at least C^2. Moreover, for the time being *we suppose that the ambient manifold M is a surface.* This is crucial to ensure that the families

$$\mathcal{W}^s = \{W^s_{loc}(x) : x \in H\} \quad \text{and} \quad \mathcal{W}^u = \{W^u_{loc}(x) : x \in H\}$$

of local stable manifolds and local unstable manifolds extend to C^1 foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ defined on a neighborhood of H. In fact, the extensions may be chosen so that tangent bundles of $\mathcal{F}^s$ and $\mathcal{F}^u$ are generated by C^1 vector fields. See [264, Appendix 1]. Consequently, the *holonomies*, i.e. the projections along leaves between any pair of cross-sections of $\mathcal{F}^s$ or $\mathcal{F}^u$ are C^1 maps.

We may extend these foliations – by negative iteration of $\mathcal{W}^s$, $\mathcal{F}^s$ and positive iteration of $\mathcal{W}^u$, $\mathcal{F}^u$ – so that their domains contain the homoclinic point q. We are going to assume that the tangency is quadratic, meaning that the curvatures of the stable manifold and the unstable manifold at q are different. Then the leaves of $\mathcal{F}^s$ and $\mathcal{F}^u$ intersect transversely near q, except for a well-defined curve ℓ of mutual tangencies containing q. See Figure 4. This follows from an implicit function argument applied to the tangent bundles of $\mathcal{F}^s$ and $\mathcal{F}^u$, which also shows that the line of tangencies ℓ is of class C^1.

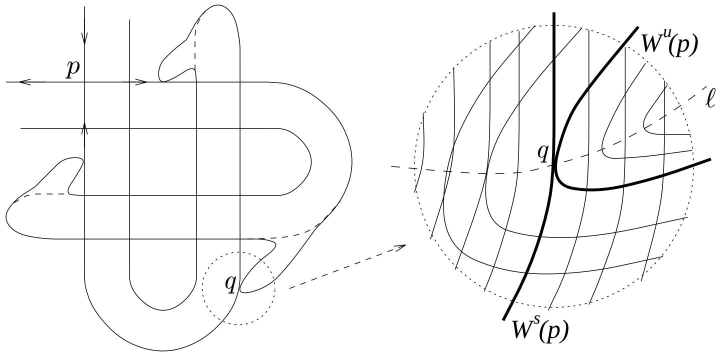

FIGURE 4. Invariant foliations close to a quadratic homoclinic tangency

Let K^s and K^u be the sets of intersections of ℓ with the leaves of $\mathcal{W}^s$ and $\mathcal{W}^u$, respectively. K^s and K^u are Cantor sets, as long as the hyperbolic set H is non-trivial. Indeed, K^s is the image of $W^u_{loc}(p) \cap H$ under $\mathcal{W}^s$-holonomy, and similarly for K^u. Recall that these holonomies are C^1 diffeomorphisms [1]. Clearly, q is in the intersection of K^s and K^u. In the example of Figure 4 this is the only point of intersection (K^s is to the left of $W^s(p)$ and K^u is to the right of $W^u(p)$), but this needs not be so in general.

Now we analyze the evolution of this picture as the parameter varies. Let g be a diffeomorphism close to f. By robustness of hyperbolic objects, g has a hyperbolic basic set H_g close to $H = H_f$ and containing the continuation p_g of the saddle point $p = p_f$. Then we may perform for H_g the same construction as above for H. Denote by $\mathcal{W}^*_g$, $\mathcal{F}^*_g$, K^*_g, $* = s, u$, the corresponding objects.

Observe that every $z \in K^s_g \cap K^u_g$ is a point of tangency between some stable manifold $W^s(g, x)$ and some unstable manifold $W^u(g, y)$ of points $x, y \in H_g$. It is not difficult to see that any point in the intersection of $W^s(H_g)$ with $W^u(H_g)$ is non-wandering, e.g. using the fact that such a point is accumulated by the stable manifold and the unstable manifold of any periodic point in H_g .

Thus, *when such a tangency does exist the diffeomorphism g can not be hyperbolic.* A much deeper converse was observed in [245, 263]: *if the two Cantor sets are sufficiently far away from each other then the dynamics of g related to the tangency is hyperbolic.* This will be made more precise later. These two observations show that *the hyperbolicity problem is directly related to determining how frequently do K^s_g and K^u_g have points of (quasi) intersection.*

3.2. **Persistent tangencies and coexistence of attractors.** It was a remarkable discovery of Newhouse [240] that the Cantor sets K^s and K^u may intersect each other in a robust fashion, that is for a whole open set of diffeomorphisms close to the original f. Here openness is relative to the C^2 topology, for reasons that will become apparent from the explanation that follows. Perhaps even more striking, assuming the initial map is area-dissipative at the periodic point p, *the generic map in such an open set of persistent tangencies has an infinite number of coexisting periodic attractors.*

[1]This fact, together with the local product property, implies that the hyperbolic set H is locally C^1-*diffeomorphic* to a product $\left(W^u_{loc}(x) \cap H\right) \times \left(W^s_{loc}(x) \cap H\right)$ close to any point $x \in H$.

For the construction, Newhouse begins by introducing the following notion of thickness of a Cantor set [2]. Let $K \subset \mathbb{R}$ be a Cantor set and $\hat{K}$ be its convex hull, i.e. the smallest compact interval containing K. Let $\{G_n : n \geq 1\}$ be the gaps of K, i.e. the connected components of $\hat{K} \setminus K$, ordered by nonincreasing length. The *thickness* of K is

$$\tau(K) = \inf \left\{ \frac{\text{length}\,(B(u))}{\text{length}\,(G_n)} : u \in \partial G_n \text{ and } n \geq 1 \right\}$$

where $B(u)$ denotes the connected component of $\hat{K} \setminus (G_1 \cup \cdots \cup G_n)$ containing the point u. To interpret this definition one should think of the Cantor set as obtained from $\hat{K}$ by successive removal of gaps. Then K has large thickness if every gap G_n is much smaller than both adjacent connected components of the set $\hat{K} \setminus (G_1 \cup \cdots \cup G_n)$ that remains after G_n is removed. What makes this concept so useful is the following

Lemma 3.2 (gap lemma [240]). *Let $K^1, K^2 \subset \mathbb{R}$ be two Cantor sets such that $\hat{K}^1 \cap \hat{K}^2 \neq \emptyset$ and neither of the two sets is contained in a gap of the other. If $\tau(K^1)\tau(K^2) > 1$ then $K^1 \cap K^2$ is non-empty.*

Proof. Let $\{G^i_n : n \geq 1\}$ be the gaps of K^i, $i = 1, 2$, in nonincreasing order of length. We say that two open intervals are linked if each one contains exactly one endpoint of the other. The proof of the lemma is by contradiction. Assuming the intersection is empty, we construct sequences $G^1_{m_i}, G^2_{n_i}$ of gaps such that every $G^1_{m_i}$ is linked to $G^2_{n_i}$ and the indices m_i and n_i converge to ∞. The latter implies that $\text{diam}\,(G^1_{m_i} \cup G^2_{n_i})$ converges to zero. Then any accumulation point of $G^1_{m_i} \cup G^2_{n_i}$ belongs to K_1 and to K_2, and so the intersection is not empty after all.

To begin the construction, let G^1_m be any gap of K^1 that intersects K^2. Each endpoint of G^1_m is either outside $\hat{K}^2$ or in some gap of K^2, because we suppose that $K_1 \cap K_2$ is empty. They can not be both outside the convex hull, because K^2 is not contained in G^1_m, and they can not be in the same gap of K^2, because G^1_m intersects K^2. So, we can always find a gap G^2_n linked to G^1_m and such that the component $B(v^2_n)$ is contained in G^1_m for some endpoint v^2_n of G^2_n.

Let u^1_m be the endpoint of G^1_m contained in G^2_n. The component $B(u^1_m)$ can not be contained in G^2_n: otherwise,

$$\frac{\text{length}\,(B(u^1_m))}{\text{length}\,(G^1_m)} \frac{\text{length}\,(B(v^2_n))}{\text{length}\,(G^2_n)} < 1,$$

which would imply $\tau(K^1)\tau(K^2) < 1$. Therefore, $B(u^1_m)$ must contain the other endpoint u^2_n of G^2_n. It follows that u^2_n is contained in some gap G^1_s of K^1 with $s > n$. Then the intervals G^2_n and G^1_s are linked, and the component $B(v^1_s)$ is contained in G^2_n for some endpoint v^1_s of G^1_s. This means that we recovered the situation in the previous paragraph, with G^2_n, G^1_s, v^1_s in the place of G^1_m, G^2_n, v^2_n, respectively. Iterating this procedure we construct sequences $G^1_{m_i}$ and $G^2_{n_i}$ as announced. □

Going back to the context of homoclinic tangencies, we may now define thicknesses of K^s and K^u, by fixing a differentiable parametrization of the line of tangencies ℓ so that the two Cantor sets may be considered as subsets of $\mathbb{R}$. In the

[2] A rediscovery really: the same notion hade been used by Hall [146] in a rather different context of number theory.

same way we may define thicknesses of $W^u_{loc}(p) \cap H$ and $W^s_{loc}(p) \cap H$, by means of parametrizations of the stable manifold and the unstable manifold. These definitions depend on the choice of the parametrizations, because the thickness is preserved a diffeomorphism up to a factor only, but that is not important for what follows [3].

Homoclinic tangencies as in Figure 4 may be found with $\tau(W^s_{loc}(p) \cap H)$ and $\tau(W^u_{loc}(p) \cap H)$ arbitrarily large, and the same for $\tau(K^s)$ and $\tau(K^u)$. In particular, we may have such examples satisfying the hypothesis of

Theorem 3.3 (Newhouse [240, 241])**.** *Suppose that* $\tau(K^s)\tau(K^u) > 1$. *Then*

(1) *f belongs to the closure of an open subset $\mathcal{U}$ of $\mathrm{Diff}^2(M^2)$ such that every $g \in \mathcal{U}$ may be approximated by a diffeomorphism exhibiting a homoclinic tangency associated to the continuation of p; in particular no element of $\mathcal{U}$ can be uniformly hyperbolic;*
(2) *if f is area-dissipative at p, meaning that $|\det Df^{\mathrm{per}(p)}(p)| < 1$, then there exists a residual subset $\mathcal{R} \subset \mathcal{U}$ such that every $g \in \mathcal{R}$ has an infinite number of periodic attractors.*

The first step is to unfold the homoclinic tangency in order to obtain an open set $\mathcal{U}$ of diffeomorphisms g, containing f in its closure, for which the interiors $\hat{K}^s_g$ and $\hat{K}^u_g$ are linked. Then neither K^s_g nor K^u_g are contained in a gap of the other. Another crucial ingredient in the proof, besides the gap lemma, is the continuity of thickness as a function of the dynamical system:

Lemma 3.4. *The functions $g \mapsto \tau(K^s_g)$ and $g \mapsto \tau(K^u_g)$ are continuous on a neighborhood of f, relative to the C^2 topology.*

So, reducing $\mathcal{U}$ if necessary, $\tau(K^s_g)\tau(K^u_g) > 1$ and so $K^s_g \cap K^u_g \neq \emptyset$ for every $g \in \mathcal{U}$. Thence, there exists a tangency between some stable manifold and some unstable manifold of points in H_g. Using that the stable manifold and the unstable manifold of p_g are dense in $W^s(H_g)$ and $W^u(H_g)$, respectively, one gets the first part of Theorem 3.3.

The second part is a consequence, using that the unfolding of a homoclinic tangency associated to a dissipative saddle always yields periodic attractors:

Lemma 3.5. *Suppose $f : M \to M$ has a homoclinic tangency q associated to an area-dissipative saddle point p. For every $\varepsilon > 0$ there exist diffeomorphisms g arbitrarily close to f exhibiting a periodic attractor contained in the ε-neighborhood of the f-orbit of q.*

The subset R_n of diffeomorphisms $g \in \mathcal{U}$ with at least n periodic attractors is open, since periodic attractors are robust. Using Lemma 3.5 one gets that every R_{n+1} is dense in R_n. Hence, $\mathcal{R} = \cap_n R_n$ is residual, as claimed.

If f is area-expansive at p, meaning that $|\det Df^{\mathrm{per}(p)}| > 1$, the same argument applied to f^{-1} produces a residual subset of diffeomorphisms in $\mathcal{U}$ with an infinite number of periodic repellers.

Another remarkable result of Newhouse shows that the phenomena in parts (1) and (2) in Theorem 3.3 are actually much more general: they occur close to *every*

[3] One way to avoid this dependence is to consider instead *local thickness*, which is invariant under diffeomorphisms. See [242] or [264, Chapter IV].

surface diffeomorphism with a homoclinic tangency, it is not necessary to assume that p is part of a thick hyperbolic set.

Theorem 3.6 (Newhouse [242]). *Let $f : M \to M$ be any surface diffeomorphism with a homoclinic tangency associated to a saddle-point p. Then*

(1) *there exists an open subset $\mathcal{U}$ of $\mathrm{Diff}^2(M^2)$ containing f in its closure, such that every $g \in \mathcal{U}$ may be approximated by a diffeomorphism with a homoclinic tangency associated to the continuation of p;*

(2) *if f is area-dissipative, respectively area-expansive, at p then there exists a residual subset $\mathcal{R} \subset \mathcal{U}$ such that every $g \in \mathcal{R}$ has an infinite number of periodic attractors, respectively periodic repellers.*

For the proof one successively perturbs the initial map in order to reach a situation similar to Theorem 3.3. A first step allows us to suppose that p belongs to a non-trivial hyperbolic basic set H_1: otherwise we just perturb f so as to create transverse homoclinic orbits together with a new homoclinic tangency associated to p. The thickness $\tau^u(H_1) := \tau(W^s_{loc}(p) \cap H_1)$ is always positive, and further perturbations will be small enough so that H_1 and $\tau^u(H_1)$ remain essentially unchanged. Up to an additional perturbation we may suppose that the tangency is quadratic, and the Jacobian of f on the orbit of p is different from 1. We consider the area-dissipative case; for the area-expanding case it suffices to replace f by its inverse.

The crucial step in the proof of Theorem 3.6 consists in showing that *the generic unfolding of a homoclinic tangency always generates (small) thick hyperbolic sets which are involved in new homoclinic tangencies.* One way such thick hyperbolic sets can be constructed is through the renormalization scheme that we outline in the next couple of paragraphs. See Section III.4 and Chapter VI of [264] for details.

Let $(f_\mu)_\mu$ be a curve of diffeomorphisms with $f_0 = f$ and such that the stable manifold and the unstable manifold of the continuation p_μ of p move with non-zero relative velocity close to the tangency as the parameter μ varies. One finds parameter values $\mu_n \to 0$ and small domains $Q_n \subset M$ converging to the point of homoclinic tangency, such that Q_n returns close to itself under $f^n_{\mu_n}$ and, in convenient *n-dependent* coordinates, the return maps $f^n_{\mu_n} \mid Q_n$ converge in the C^2 topology to

$$\phi(x, y) = (1 - 2x^2, x)$$

as $n \to \infty$. The fact that the limit ϕ has zero Jacobian is no surprise: most iterates of the return map take place close to the saddle p, where our diffeomorphisms are area-dissipative.

The most important feature of ϕ in this context is that it admits invariant hyperbolic Cantor sets with arbitrarily large thickness in the x-direction. This can be seen, for instance, from the fact that ϕ is smoothly conjugate to the piecewise affine map $\psi(x, y) = (1 - 2|x|, x)$, for which such thick invariant sets may be exhibited explicitly. By hyperbolic continuation and continuity of the thickness, it follows that for every large n the map $f^n_{\mu_n} \mid Q_n$ admits invariant hyperbolic sets H_2 with large thickness transversely to the stable direction: we may take H_2 and a periodic point $p_2 \in H_2$ so that $\tau^s(H_2) := \tau(W^u_{loc}(p_2) \cap H_2)$ satisfies

$$\tau^u(H_1)\tau^s(H_2) > 1. \tag{13}$$

The hyperbolic basic sets H_1 and H_2 are *heteroclinically related,* i.e. there are transverse intersections between their stable manifolds and unstable manifolds, and

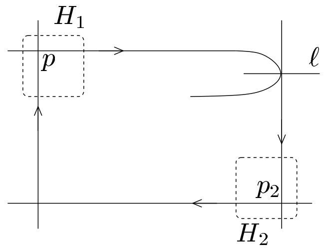

FIGURE 5. Heteroclinic tangencies

a tangency between $W^u(p)$ and $W^s(p_2)$ may be created by another small perturbation of the diffeomorphism. See Figure 5. This is possible because the map ϕ itself has homoclinic tangencies: the unstable set [4] of the fixed point $P = (-1, 0)$ has non-transverse intersections with the stable set of P, for instance at $q = (1, 0)$. The continuation of P for the diffeomorphism $f^n_{\mu_n}$ is heteroclinically related to both H_1 and H_2 and it is close to having a homoclinic tangency. Then the stable and unstable manifolds of p and p_2 accumulate on the corresponding invariant manifolds of that continuation, and a heteroclinic tangency associated to p and p_2 may be created by a small perturbation, as claimed.

At this point we may use the gap lemma and (13) in much the same way as for Theorem 3.3. We conclude that there exists an open set $\mathcal{U} \subset \mathrm{Diff}^2(M)$ with persistent heteroclinic tangencies involving H_1 and H_2 , and so also with persistent homoclinic tangencies associated to p. The same Baire type argument as in Theorem 3.3 gives that maps in a residual subset of $\mathcal{U}$ have infinitely many coexisting attractors.

Remark 3.7. This renormalization scheme also gives a quick proof of Lemma 3.5: there exist $\phi_a(x, y) = (1 - ax^2, x)$ with a arbitrarily close to 2 admitting hyperbolic periodic attractors; by robustness, this remains true for f^n_μ with μ close to μ_n. So f_μ has hyperbolic periodic attractors for μ arbitrarily close to zero.

3.2.1. *Open sets with persistent tangencies.* Some three decades after their existence was proved, the phenomena of persistent tangencies and coexistence of infinitely many sinks or sources remain a mystery, and we shall be referring to them often as we attempt to the develop a global picture of dynamics beyond hyperbolicity. Here we review some recent results and extensions that also give a measure of their inherent complexity.

We begin by noticing that the two phenomena are intimately related. We have already seen how persistence of tangencies leads to coexistence of sinks or sources. A converse follows from results of Pujals, Sambarino [286] that we review in Section 5: if an open set of $\mathrm{Diff}^1(M^2)$ contains a dense subset corresponding to infinitely many coexisting sinks or sources, then it also contains a dense subset corresponding to homoclinic tangencies.

Let us also mention that Newhouse's argument can not work in the C^1 topology. Indeed, Ures [329] shows that hyperbolic sets of a C^1 residual subset of diffeomorphisms have thickness equal to zero, in both stable direction and unstable directions.

[4]The unstable set is the union of all $\phi^n(V)$, $n \geq 1$ for some small neighborhood V of P, and the stable set is the union of all pre-images $\phi^{-n}(P)$, $n \geq 0$.

He also deduces that for generic C^1 families unfolding a homoclinic tangency, the set of parameters corresponding to hyperbolicity has full Lebesgue density at the tangency parameter. Indeed, the following basic problem remains wide open:

Problem 3.8. Are uniformly hyperbolic systems dense among all C^1 diffeomorphisms on a surface ?

On the other hand, Bonatti, Díaz [57] constructed C^1 open sets with coexistence of infinitely many sinks for a residual subset, on manifolds of dimension $d \geq 3$. See Section 10.6.2 below.

In addition, it is interesting to point out that the phenomenon of coexistence of periodic attractors has been extended by Colli [94] for strange attractors. Indeed, he constructs open sets of diffeomorphisms containing dense subsets for which there are infinitely many coexisting Hénon-like attractors. Observe that the argument must be much more delicate than in the periodic case, because Hénon-like attractors are persistent in a measure-theoretical sense only.

Bonatti, Díaz [53] also obtain coexistence of infinitely many non-trivial attractors, hyperbolic or not, in C^1 generic subset of diffeomorphisms on manifolds of dimension $d \geq 3$.

Conjecture 3.9 ([325, 260])**.** In the unfolding of a homoclinic tangency associated to an area-dissipative (sectionally dissipative in higher dimensions, see Section 3.5), the set of parameter values corresponding to coexistence of infinitely many periodic or Hénon-like attractors has Lebesgue measure zero.

Also very interesting, Theorem 3.6 has been extended to the area-preserving maps by Duarte [123, 124]: any map with a homoclinic tangency is C^2 approximated by an open domain with persistent tangencies; the generic diffeomorphism on such a domain has infinitely many "independent" elliptic islands (which take the role of attractors or repellers in the non-conservative case).

Another striking property of Newhouse's domains with persistent tangencies has been discovered recently by Kaloshin: super-exponential growth of the number of periodic orbits. Indeed, he proves

Theorem 3.10 (Kaloshin [167])**.** *Let $\mathcal{U} \subset \mathrm{Diff}^2(M^2)$ be an open set with persistent tangencies. Given any sequence $(\eta_n) \to \infty$ there exists a residual subset of $\mathcal{U}$ for which the number of periodic points of period n is larger than η_n for all large n.*

In a few words, Kaloshin exploits the fact that diffeomorphisms with highly degenerate saddle-nodes (high contact with the identity along the central direction) are dense in the Newhouse domain $\mathcal{U}$ (Gonchenko, Shil'nikov, Turaev [139]), to produce abnormal rates of growth of the number of periodic points.

Quite in contrast, Hunt, Kaloshin [161] announce that the number of periodic points of period n is bounded by const $\exp(n^{1+\delta})$, for a "full probability" subset of diffeomorphisms.

3.3. **Hyperbolicity and fractal dimensions.** We are going to see that a refinement of the previous analysis leads to a deep connection between frequency of hyperbolicity in the unfolding a homoclinic tangency and fractal dimensions of invariant sets.

Let $(f_\mu)_\mu$ be a smooth parametrized family of diffeomorphisms such that $f = f_0$ has a quadratic homoclinic tangency q associated to a periodic point p, and this tangency is *generically unfolded* as the parameter μ varies: the stable manifold and the unstable manifold of the continuation p_μ of p move with respect to each other with non-zero velocity near the point q, so that the tangency gives rise to a pair of transverse intersections when $\mu > 0$. As before, we suppose that p is part of a hyperbolic basic H, which may or may not be reduced to the orbit of p.

We are going to suppose that $\mu = 0$ is a *first bifurcation*, meaning that f_μ is uniformly hyperbolic for every $\mu < 0$ close to zero. This is for simplicity only: the conclusions hold in general, restricted to the dynamics of f_μ related to the homoclinic tangency of f. More precisely, in the general case one must refer to the maximal invariant set

$$\Sigma_\mu = \bigcap_{n \in \mathbb{Z}} f_\mu^n(U \cup V_\mu). \tag{14}$$

inside a neighborhood U of H union a neighborhood V_μ of the orbit of tangency. See [264, Chapter V]. In the first bifurcation case the non-wandering set $\Omega(f_\mu)$ is contained in the union of Σ_μ with a finite number of hyperbolic basic sets, so that the diffeomorphism is uniformly hyperbolic if Σ_μ is.

As we have seen, hyperbolicity is directly related to existence of intersections, or almost intersections, between the Cantor sets $K_\mu^* := K_{f_\mu}^*$, $* = s, u$. Let us again look at the way these Cantor sets vary with the parameter μ. Refer to Figure 4. The stable foliation $\mathcal{F}_\mu^s$ and the unstable foliation $\mathcal{F}_\mu^u$ may be chosen such that their tangent bundles vary in a C^1 fashion with μ. Then so does the curve of tangencies ℓ_μ. We fix parametrizations of these curves, jointly C^1 in phase space and parameter space. Of course, we may choose the point of $W^u(p_\mu) \cap \ell_\mu$ close to the tangency to lie at the origin of the coordinates induced in ℓ_μ by this parametrization.

The main effect on the variation of the two Cantor sets comes from the generic unfolding of the tangency: K_μ^s is translated with respect to K_μ^u when the parameter varies, so that their convex hulls become linked when $\mu > 0$. With our choices of coordinates, and rescaling the parameter μ is necessary,

$$K_\mu^u \approx K^u \quad \text{and} \quad K_\mu^s \approx K^s + \mu$$

for all μ close to zero. Thus, K_μ^u and K_μ^s are close to intersecting each other if and only if K^u and $K^s + \mu$ are close to intersecting each other. In this way, one is led to studying the *arithmetic difference*

$$\begin{aligned} K^u - K^s &= \{\mu \in \mathbb{R} : \mu = \kappa_u - \kappa_s \text{ for some } \kappa_u \in K^u \text{ and } \kappa_s \in K^s\} \\ &= \{\mu \in \mathbb{R} : K^u \cap (K^s + \mu) \text{ is non-empty}\}. \end{aligned}$$

And Problem 3.1 directly points to:

Problem 3.11. Let K^1 and K^2 be two Cantor sets in the real line, especially dynamically defined ones. How big is their arithmetic difference $K^1 - K^2$?

A very precise answer to this problem may be given in terms of metric invariants of K^1 and K^2. To begin with, it follows from the gap lemma 3.2 that if $\tau(K_1)\tau(K_2) > 1$ then the arithmetic difference is big: it contains at least the intervals $\hat{K}^1 - \hat{K}^2$ for which the corresponding convex hulls are linked. More refined statements are in terms of fractal dimensions, like the Hausdorff dimension and the limit capacity (box dimension).

Let X be a compact metric space and $N(X,\varepsilon)$ denote the minimum number of ball of radius $\varepsilon > 0$ required to cover X. The *limit capacity* of X is

$$c(X) = \inf\{d > 0 : N(X,\varepsilon) \le \varepsilon^{-d} \text{ for every small } \varepsilon > 0\}.$$

Now we define the Hausdorff dimension of X. Given any $d > 0$ the *Hausdorff d-measure* of X is

$$m_d(X) = \lim_{\varepsilon\to 0}\left(\inf \sum_{U\in\mathcal{U}} \operatorname{diam}(U)^d\right)$$

where the infimum is taken over all coverings $\mathcal{U}$ of X by sets U with diameter less than $\varepsilon > 0$. It is easy to see that there is a unique number $0 \le \mathrm{HD}(X) \le \infty$, the *Hausdorff dimension* of X, such that

$$m_d(X) = \infty \text{ if } d < \mathrm{HD}(X) \quad \text{and} \quad m_d(X) = 0 \text{ if } d > \mathrm{HD}(X).$$

Detailed information about these notions can be found in [127, 271] and [264, Chapter IV].

It is easy to show that $\mathrm{HD}(X) \le c(X)$, and $0 \le \mathrm{HD}(X) \le c(X) \le d$ for every compact subset of a d-dimensional manifold. Moreover, $\mathrm{HD}(Y) \le \nu^{-1}\mathrm{HD}(X)$ and $c(Y) \le \nu^{-1}c(X)$ if Y is the image of X by a ν-Hölder continuous transformation. In particular, both dimensions are preserved by bi-Lipschitz homeomorphisms. Since stable and unstable holonomies are C^1, it follows that

$$c(K^s_\mu) = c\big(W^u_{loc}(p_\mu)\cap H_\mu\big) \quad \text{and} \quad \mathrm{HD}(K^s_\mu) = \mathrm{HD}\big(W^u_{loc}(p_\mu)\cap H_\mu\big)$$

and analogously for K^u_μ and $W^s_{loc}(p_\mu)\cap H_\mu$.

In fact, our dynamical sets are much more regular than arbitrary Cantor subsets of $\mathbb{R}$. In particular, $c(K^*_\mu) = \mathrm{HD}(K^*_\mu) \in (0,1)$ for $* = s,u$ and

$$c(H_\mu) = \mathrm{HD}(H_\mu) = \mathrm{HD}(K^s_\mu) + \mathrm{HD}(K^u_\mu) \in (0,2).$$

Moreover, the fractal dimension $c(K^*_\mu) = \mathrm{HD}(K^*_\mu)$ varies continuously with the diffeomorphism f_μ, in the C^1 topology. See [264, Chapter IV] and [213, 265].

The following simple lemma shows that the arithmetic difference of sets with small limit capacity is a zero Lebesgue measure set. The proof is an exercise.

Lemma 3.12. *Let $K^1, K^2 \subset \mathbb{R}$ be compact sets. Then*

$$c(K^1 - K^2) \le c(K^1 \times K^2) \le c(K^1) + c(K^2).$$

In particular, if $c(K^1) + c(K^2) < 1$ then $m(K^1 - K^2) = 0$.

Starting from this observation, Palis, Takens show that *if H has fractal dimension less than 1 then most values of μ close to $\mu = 0$ correspond to hyperbolicity.* The case when the set H reduces to the orbit of p had been treated in [245].

Theorem 3.13 (Palis, Takens [263])**.** *If $c(H) < 1$ then $\mu = 0$ is a Lebesgue density point for the set $\mathcal{H}$ of parameter values for which f_μ is uniformly hyperbolic:*

$$\lim_{\varepsilon\to 0}\frac{m(\mathcal{H}\cap[0,\varepsilon])}{\varepsilon} = 1.$$

Conversely, Palis, Yoccoz prove that *if the fractal dimension is bigger than 1 then non-hyperbolic parameters correspond to a sizable portion of parameter space near $\mu = 0$.* The statement is slightly more involved than in the case of Theorem 3.13. They consider smooth two-parameter families of surface diffeomorphisms $(f_{\mu,\nu})_{\mu,\nu}$ such that $(f_{\mu,0})_\mu$ unfolds generically a homoclinic tangency associated to a hyperbolic set H of $f_{0,0}$. For every ν small the family $(f_{\mu,\nu})_\mu$ unfolds generically a homoclinic tangency associated to the continuation of H; we may suppose that the tangency takes place at $\mu = 0$. Let $\Sigma_{\mu,\nu}$ be the maximal invariant set of $f_{\mu,\nu}$ defined as in (14). Then

Theorem 3.14 (Palis, Yoccoz [267]). *If* $\mathrm{HD}(H) > 1$ *then*

$$\liminf_{\varepsilon\to 0} \frac{m(\mathcal{H}_\nu \cap [0,\varepsilon])}{\varepsilon} < 1, \text{ for Lebesgue-almost every } \nu \text{ close to zero,}$$

where $\mathcal{H}_\nu$ *is the set of values of* μ *for which* $f_{\mu,\nu}$ *is uniformly hyperbolic.*

At the heart of the proof is the following deep result which provides a converse to Lemma 3.12. The need for the extra parameter ν in Theorem 3.14 is related to the presence of the parameter λ in Marstrand's theorem:

Theorem 3.15 (Marstrand [214]). *Let* K^1 *and* K^2 *be compact subsets of* $\mathbb{R}$ *with* $\mathrm{HD}(K^1) + \mathrm{HD}(K^2) > 1$. *Then for Lebesgue almost every* $\lambda \in \mathbb{R}$ *there exists a positive Lebesgue measure set* T_λ *such that for every* $t \in T_\lambda$

$$\mathrm{HD}(K^1 \cap (\lambda K^2 + t)) \geq d > 0, \qquad d = \mathrm{HD}(K^1) + \mathrm{HD}(K^2) - 1.$$

In particular, $T_\lambda \subset K^1 - \lambda K^2$ *and so* $m(K^1 - \lambda K^2) > 0$ *for almost every* $\lambda \in \mathbb{R}$.

This admits the following geometric interpretation. Let $\pi_\lambda : \mathbb{R}^2 \to \mathbb{R}$ be the projection along co-slope λ: $\pi_\lambda(x_1, x_2) = x_1 - \lambda x_2$. For almost every λ the image $\pi_\lambda(K^1 \times K^2) = K^1 - \lambda K^2$ has positive Lebesgue measure. Actually, Marstrand proves that *for almost every* λ *the projection* $(\pi_\lambda)_*(m_{\mathrm{HD}})$ *of the Hausdorff measure* m_{HD} *on* $K^1 \times K^2$ *is absolutely continuous, with* L^2 *density.* Moreover, there even exists a positive Lebesgue measure subset T_λ such that for $t \in T_\lambda$ the fiber

$$\pi_\lambda^{-1}(t) \cap (K^1 \times K^2) \approx K^1 \cap (\lambda K^2 + t)$$

has Hausdorff dimension at least d.

Remark 3.16. The hypothesis of Theorem 3.14 is more general than that of Theorem 3.3: thick Cantor sets always have Hausdorff dimension close to 1, in particular,

$$\tau(K^s)\tau(K^u) > 1 \quad \Rightarrow \quad \mathrm{HD}(H) = \mathrm{HD}(K^s) + \mathrm{HD}(K^u) > 1.$$

3.4. **Stable intersections of regular Cantor sets.** In the context of these results, Palis has proposed the following ambitious answer to Problem 3.11:

Conjecture 3.17. The arithmetic difference $K^1 - K^2$ of generic dynamically defined Cantor sets either has zero Lebesgue measure or contains an interval.

This conjecture was proved by Moreira, Yoccoz [237, 236], in a very strong form. From it they deduced a statement about homoclinic tangencies that improves Theorem 3.14 in more than one way. These outstanding results are the subject of the present section. Notice that the statement above can not be true for *all* pairs of dynamically defined Cantor sets, by [308].

A Cantor set $K \subset \mathbb{R}$ is called *dynamically defined*, or *regular*, if it is the limit set

$$K = \bigcap_{n=0}^{\infty} \bigcup f_{j_{-n}} \circ \cdots \circ f_{j_{-1}}(D_{j_1}) \tag{15}$$

of a Markov family of contractions, that is, a finite family $f_j : D_j \to R_j$ of C^r, $r \geq 1$ diffeomorphisms between compact subintervals D_j and R_j of $I = [0,1]$, with derivative bounded by some constant $\lambda < 1$, and such that every domain D_j is the convex hull of some subset of ranges R_i. The union in (15) is over all *admissible sequences* $j_{-n}, \ldots, j_{-1}$, that is, such that the composition is defined.

The Cantor sets K^s, K^u, $W^u_{loc}(p)$, $W^s_{loc}(p)$ introduced above are regular with $r > 1$, that is, the contractions f_j may be chosen so that their derivatives are at least Hölder continuous. See [264, Chapter IV]. In what follows we fix $r > 1$.

The space of regular Cantor sets inherits a natural topology from the C^r topology in the space of such families (with fixed number of elements N). This induces a natural C^r topology in the space $\mathcal{K}$ of pairs of regular Cantor sets (K^1, K^2).

Definition 3.18 (Moreira [234]). K^1 and K^2 have *stable intersection* if for every pair $(\tilde{K}_1, \tilde{K}_2) \in \mathcal{K}$ close to (K_1, K_2) we have $\tilde{K}_1 \cap \tilde{K}_2 \neq \emptyset$.

Translates of a regular Cantor set K are also regular, and they are C^r close to K if the translation is small. So, if K^1 and K^2 have stable intersection or, more generally, if K^1 and $K^2 + t$ have stable intersection for some $t \in \mathbb{R}$, then $K^1 - K^2$ contains an interval. Moreira, Yoccoz prove that *an open dense subset of pairs of Cantor sets whose sum of the Hausdorff dimensions is larger than* 1 *have stable intersection after convenient translation.* In view of Lemma 3.12, this result implies Palis conjecture 3.17.

Theorem 3.19 (Moreira, Yoccoz [237]). *Let Ω be the set of $(K^1, K^2) \in \mathcal{K}$ such that $\mathrm{HD}(K^1) + \mathrm{HD}(K^2) > 1$. Then there exists an open and dense subset $\mathcal{W}$ of Ω such that for every $(K^1, K^2) \in \mathcal{W}$ the set*

$$S(K^1, K^2) := \{t \in \mathbb{R} : K^1 \text{ and } K^2 + t \text{ have stable intersection } \}$$

is open and dense in $K_1 - K_2$ and $\mathrm{HD}\big((K^1 - K^2) \setminus S(K^1, K^2)\big)$ is less than 1.

Let $(f_\mu)_\mu$ be a smooth family of surface diffeomorphisms generically unfolding a homoclinic tangency associated to a hyperbolic set H of $f = f_0$. Adapting the proof of Theorem 3.19 and using results from [234], Moreira, Yoccoz show that, if the Hausdorff dimension of H is larger than 1, then persistent tangencies correspond to a definite fraction of parameters close to zero and the union of such parameters with those corresponding to hyperbolicity has full Lebesgue density at the bifurcation. More formally,

Theorem 3.20 (Moreira, Yoccoz [236]). *Assume $\mathrm{HD}(H) > 1$. Then there exists an open set $\mathcal{N}$ in parameter space corresponding to persistent tangencies such that*

$$\liminf_{\varepsilon \to 0} \frac{m(\mathcal{N} \cap [0, \varepsilon])}{\varepsilon} > 0 \quad \text{and} \quad \liminf_{\varepsilon \to 0} \frac{m((\mathcal{N} \cup \mathcal{H}) \cap [0, \varepsilon])}{\varepsilon} = 1,$$

where $\mathcal{H}$ is the set of parameters for which f_μ is uniformly hyperbolic.

The key to proving Theorem 3.19 is the fact, remarkable in itself, that for most pairs of regular Cantor sets the same geometric pattern is repeated over and over again at arbitrarily small scales. More precisely, Moreira, Yoccoz introduce a family of renormalization (or scale refinement) operators $\mathcal{R}_{\alpha, \tilde{\alpha}}$ and prove that, up to small perturbation of the Cantor sets, this family admits a compact invariant set.

In the sequel we are going to describe the main steps in the proof in more detail. We begin by explaining the meaning and giving the precise definition of the renormalization operators.

3.4.1. *Renormalization operators and pattern recurrence.* For simplicity we consider Cantor sets dynamically defined by contractions $f_j : I \to I$ defined on the whole $I = [0, 1]$. Let Σ be the space of sequences $\theta = (\dots, \theta_{-n}, \dots, \theta_{-1})$. For every $n \geq 1$ denote $f_\theta^n = f_{\theta_{-n}} \circ \cdots \circ f_{\theta_{-1}}$. This is defined also when θ is a finite sequence with length at least n. Endow Σ with the metric

$$d(\theta, \theta') = \text{length}\,(f_\beta^k(I)), \tag{16}$$

where $\beta = (\theta_{-k}, \dots, \theta_{-1})$ if $\theta_{-i} = \theta'_{-i}$ for $1 \leq i \leq k$ and k is maximum.

Let α and $\tilde{\alpha}$ be two finite admissible sequences of length k and $\tilde{k}$, respectively. Given regular Cantor sets (K, f) and $(\tilde{K}, \tilde{f})$, admissible sequences θ and $\tilde{\theta}$, and integers m and $\tilde{m}$, the interval $f_\theta^m(I)$ is a connected component of the m:th stage of the construction of K, and analogously for $\tilde{K}$, $\tilde{f}$, $\tilde{\theta}$, $\tilde{m}$. Roughly, the renormalization operator is meant to assign

$$\mathcal{R}_{\alpha,\tilde{\alpha}} : \left(f_\theta^m(I), \tilde{f}_{\tilde{\theta}}^{\tilde{m}}(I)\right) \mapsto \left(f_{\theta\alpha}^{m+k}(I), \tilde{f}_{\tilde{\theta}\tilde{\alpha}}^{\tilde{m}+\tilde{k}}(I)\right) \quad \text{after rescaling.}$$

Observe that the intervals on the right hand side are connected components of deeper stages of the construction of the two Cantor sets.

For the formal definition one takes a kind of limit when m, $\tilde{m}$ go to ∞. For each $\theta \in \Sigma$ and $n \geq 1$ let h_θ^n be the orientation preserving affine map from $f_\theta^m(I)$ to the interval I. The limit

$$\kappa_\theta = \lim h_\theta^n \circ f_\theta^n$$

exists for every $\theta \in \Sigma$ and depends Hölder continuously on θ, because our maps are differentiable and the derivatives are Hölder continuous. The images $\kappa_\theta(K)$ of K under the maps κ_θ are the *limit geometries* of the Cantor set. A direct important consequence of the definition is that for any admissible sequence α with length $k \geq 1$, the set $\kappa_\theta(f_\alpha^k(K)) = \kappa_\theta(K) \cap \kappa_\theta(f_\alpha^k(I))$ is the affine image of another limit geometry, namely $\kappa_{\theta\alpha}(K)$.

For our purposes it is important to keep track also of relative positions, which are forgotten by the quotient by the affine group involved in the definition of limit geometries. The *relative position* of two oriented intervals J_1, J_2 is the pair $(\lambda, t) \in \mathbb{R}^* \times \mathbb{R}$ defined by

$$h(J_2) = \lambda h(J_1) + t$$

where h is the orientation preserving affine map sending J_1 onto $[0, 1]$, and $\lambda < 0$ if the orientations of J_1 and J_2 are opposite. Given any finite admissible sequences $\alpha, \tilde{\alpha}$, the associated *renormalization operator* is defined by

$$\mathcal{R}_{\alpha,\tilde{\alpha}} : \Sigma \times \tilde{\Sigma} \times \mathbb{R}^* \times \mathbb{R} \to \Sigma \times \tilde{\Sigma} \times \mathbb{R}^* \times \mathbb{R}, \quad (\theta, \tilde{\theta}, \lambda, t) \mapsto (\theta\alpha, \tilde{\theta}\tilde{\alpha}, \lambda_1, t_1)$$

where (λ_1, t_1) is the relative position of $\kappa_\theta(f_\alpha^k(I))$ and $\lambda\tilde{\kappa}_{\tilde{\theta}}(\tilde{f}_{\tilde{\alpha}}^{\tilde{k}}(I)) + t$. See Figure 6.

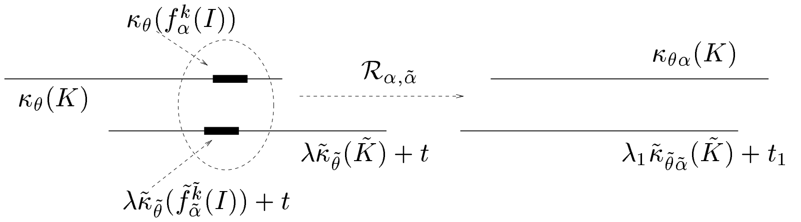

FIGURE 6. Renormalization operator $\mathcal{R}_{\alpha,\tilde{\alpha}}$

Lemma 3.21. *Suppose there exists a compact subset $\mathcal{L}$ of $\Sigma \times \tilde{\Sigma} \times \mathbb{R}^* \times \mathbb{R}$ such that for every $(\theta, \tilde{\theta}, \lambda, t) \in \mathcal{L}$ there exist α, $\tilde{\alpha}$ such that*

$$\mathcal{R}_{\alpha,\tilde{\alpha}}(\theta, \tilde{\theta}, \lambda, t) \in \operatorname{int}(\mathcal{L}).$$

Then $\kappa_\theta(K)$ and $\lambda\tilde{\kappa}_{\tilde{\theta}}(\tilde{K}) + t$ have C^r stable intersection [5] *for every $(\theta, \tilde{\theta}, \lambda, t) \in \mathcal{L}$.*

The idea of the proof is the following: by successive renormalizations one constructs a sequence $(\theta^{(n)}, \tilde{\theta}^{(n)}, \lambda^{(n)}, t^{(n)})$ inside the compact set $\mathcal{L}$. In particular, the sequence $t^{(n)}$ is bounded: the intervals involved are at bounded distance, *in terms of relative positions.* Since the diameters are going to zero as n increases, this really means that these intervals get closer and closer. In this way we get a common point of accumulation, and this point is in the intersection of the two Cantor sets.

The problem is now how to find such a *recurrent set* $\mathcal{L}$ as in the hypothesis of Lemma 3.21, for every pair of Cantor sets $(K, \tilde{K})$ in some open and dense subset $\mathcal{W}$ of Ω. Density is the main issue, of course. Fix $\varepsilon > 0$. We are going to explain that a recurrent set does exist, after some 10ε-perturbation of the original pair. The main difficulty comes from the fact that the behavior of the renormalization operators along the variable λ is mostly neutral; in contrast, the operators are strongly expanding along the t variable, corresponding to the rescalings involved in the definition. For this reason, the construction is carried out in two steps.

3.4.2. *The scale recurrence lemma.* The first step is to find a compact recurrent set $\mathcal{L}_*$ for the renormalization operators acting on $\Sigma \times \tilde{\Sigma} \times \mathbb{R}^*$, that is, taking only the scale variable λ in consideration. This is handled by the following crucial *scale recurrence lemma.*

Choose a set $\Sigma(\varepsilon)$ of finite admissible sequences α, with variable lengths k_α, such that length $(f_\alpha^{k_\alpha}(I)) \approx \varepsilon$ and the intervals $f_\alpha^{k_\alpha}(I)$ cover K. Choose $\tilde{\Sigma}(\varepsilon)$ analogously for $\tilde{K}$. Taking ε small,

$$\#\big(\Sigma(\varepsilon) \times \tilde{\Sigma}(\varepsilon)\big) \approx \varepsilon^{-\mathrm{HD}(K) - \mathrm{HD}(\tilde{K})}$$

Assume at least one of the Cantor sets, K say, is essentially non-affine: the corresponding expanding transformation ψ is C^2 and there exist θ_1, $\theta_2 \in \Sigma^-$ and $x \in K$ such that $(\kappa_{\theta_1}^{-1} \circ \kappa_{\theta_2})''(x) \neq 0$. The scale recurrence lemma says that *given a large $A > 0$ and a family $\{L_{\theta,\tilde{\theta}} : (\theta, \tilde{\theta}) \in \Sigma \times \tilde{\Sigma}\}$ of subsets of $[-A, A]$ with*

$$\operatorname{Leb}([-A, A] \setminus L_{\theta,\tilde{\theta}}) \ll 1 \quad \textit{for all } (\theta, \tilde{\theta}) \in \Sigma \times \tilde{\Sigma},$$

there exists a family $\{\mathcal{L}_{\theta,\tilde{\theta}} : (\theta, \tilde{\theta}) \in \Sigma \times \tilde{\Sigma}\}$ of compact sets, each $\mathcal{L}_{\theta,\tilde{\theta}}$ near $L_{\theta,\tilde{\theta}}$, such that

$$\operatorname{Leb}([-A, A] \setminus \mathcal{L}_{\theta,\tilde{\theta}}) \ll 1 \quad \textit{for all } (\theta, \tilde{\theta}) \in \Sigma \times \tilde{\Sigma},$$

and, for each $\lambda \in \mathcal{L}_{\theta,\tilde{\theta}}$ there is a positive fraction [6] *of pairs $(\alpha, \tilde{\alpha}) \in \Sigma(\varepsilon) \times \tilde{\Sigma}(\varepsilon)$*

$$\mathcal{R}_{\alpha,\tilde{\alpha}}(\theta, \tilde{\theta}, \lambda) \in \operatorname{int}(\mathcal{L}_{\theta\alpha,\tilde{\theta}\tilde{\alpha}}). \tag{17}$$

To use this lemma, one considers sets $L_{\theta,\tilde{\theta}}$ as follows. Notice that

$$\mathrm{HD}(\kappa_\theta(K)) + \mathrm{HD}(\tilde{\kappa}_{\tilde{\theta}}(\tilde{K})) = \mathrm{HD}(K) + \mathrm{HD}(\tilde{K}) > 1,$$

[5] For any $r > 1$ such that the regular Cantor sets are C^r.

[6] In other words, at least $c\varepsilon^{-\mathrm{HD}(K) - \mathrm{HD}(\tilde{K})}$ of them.

and so Marstrand's theorem applies to $(\kappa_\theta(K), \tilde\kappa_{\tilde\theta}(K))$. Then, for almost every λ the π_λ-projection down to $\mathbb{R}$ of the Hausdorff measure on the product of these Cantor sets has L^2 density. Consider $B \gg A \gg 0$ and let $L_{\theta,\tilde\theta}$ be the set of $\lambda \in [-A, A]$ for which the L^2 norm is less than B. Using the Cauchy-Schwarz inequality, we get that

$$B^2 \mathrm{Leb}(\pi_\lambda(Y)) \geq m_{\mathrm{HD}}(Y)^2 \quad \text{for every measurable set } Y \tag{18}$$

and any $\lambda \in L_{\theta,\tilde\theta}$. This relation will be important in the second step.

Then $\mathcal{L}_* = \{(\theta, \tilde\theta, \lambda) : \lambda \in \mathcal{L}_{\theta,\tilde\theta} \text{ and } (\theta, \tilde\theta) \in \Sigma \times \tilde\Sigma\}$ is a compact recurrent set for the renormalization operator in $\Sigma \times \tilde\Sigma \times \mathbb{R}^*$. The conclusion (17) contains

$$\lambda \in \mathcal{L}_{\theta,\tilde\theta} \quad \Longrightarrow \lambda_1 \in \mathcal{L}_{\theta\alpha,\tilde\theta\tilde\alpha}.$$

Observe that $\lambda_1 \approx \lambda$ because length $(\kappa_\theta(f_\alpha^{k_\alpha}(I))) \approx \varepsilon \approx$ length $(\tilde\kappa_{\tilde\theta}(\tilde f_{\tilde\alpha}^{k_{\tilde\alpha}}(I)))$. This corresponds to the fact we already mentioned that the renormalization operator acts in a neutral fashion along the λ variable.

Remark 3.22. The scale recurrence lemma implies the following beautiful formula:

$$\mathrm{HD}(K^1 + K^2) = \min\{1, \mathrm{HD}(K^1) + \mathrm{HD}(K^2)\}$$

whenever K^1 and K^2 are regular Cantor sets, with K^1 being C^2 and essentially non-affine. See Moreira [233].

3.4.3. *The probabilistic argument.* The hypothesis of non-affinity above is the first condition in the definition of the open dense set $\mathcal{W}$ in Theorem 3.19. The remaining (much less explicit) conditions correspond to additional perturbations to ensure, from $\mathcal{L}_*$, the existence of an invariant set $\mathcal{L}$ for the full renormalization operator acting on $\Sigma \times \tilde\Sigma \times \mathbb{R}^* \times \mathbb{R}$.

Let $(\theta, \tilde\theta, \lambda) \in \mathcal{L}_*$. By Marstrand's theorem 3.15, there exists a positive measure set $T_{\theta,\tilde\theta,\lambda}$ of values of t for which the Hausdorff dimension of

$$\kappa_\theta(K) \cap \left(\lambda \tilde\kappa_{\tilde\theta}(\tilde K) + t\right)$$

is at least $d = \mathrm{HD}(K) + \mathrm{HD}(\tilde K) - 1$. Since length $(\kappa_\theta(f_\alpha^{k_\alpha}(I))) \approx \varepsilon$, this implies (fix any $\tilde d < d$) that there are at least $c\varepsilon^{-\tilde d}$ pairs $(\alpha, \tilde\alpha) \in \Sigma(\varepsilon) \times \tilde\Sigma(\varepsilon)$ for which the corresponding intervals intersect. We shall refer to these as distinguished pairs and distinguished intervals. To be more precise, we satisfy also the conclusion (17) of the scale recurrence lemma. The reason we can do this is, essentially, the following. The pairs given by the recurrence lemma cover a definite fraction of the product Cantor set. Using (18) we see that this covered subset projects to a positive Lebesgue measure subset and, in fact, Marstrand's argument may be carried out within this subset. So we may pick $c\varepsilon^{-d}$ distinguished pairs from among those $(\alpha, \tilde\alpha)$ given by the scale recurrence lemma.

Very roughly, we now take $\mathcal{L}_\varepsilon$ to be the set of $(\theta, \tilde\theta, \lambda, t) \in \Sigma \times \tilde\Sigma \times \mathbb{R}^* \times \mathbb{R}$ such that t is in the ε-neighborhood of $T_{\theta,\tilde\theta,\lambda}$. We want to prove that this is a recurrent set: given any $(\theta, \tilde\theta, \lambda, t) \in \mathcal{L}_\varepsilon$ there is a distinguished pair $(\alpha, \tilde\alpha)$ such that

$$\mathcal{R}_{\alpha,\tilde\alpha}(\theta, \tilde\theta, \lambda, t) \in \mathcal{L}_{\varepsilon/2} \tag{19}$$

at least after a 10ε-perturbation of the Cantor sets. To this end, we embed the expanding transformation ψ of K into a parametrized family $\{\psi_\xi : \xi \in [-10\varepsilon, 10\varepsilon]^p\}$, with $p \approx \varepsilon^{-d/2}$ parameters, such that varying one parameter causes a corresponding interval $\kappa_\theta(f_\alpha^{k_\alpha}(I))$ to move with respect to the associated $\tilde{\kappa}_{\tilde{\theta}}(\tilde{f}_{\tilde{\alpha}}^{k_{\tilde{\alpha}}}(I))$, while keeping the other $p-1$ distinguished intervals unchanged [7].

This displacement causes the relative position of the pair of intervals to move across a whole interval, of uniform size, along the t-direction. Noting that the two intervals intersect for the initial Cantor sets $(K, \tilde{K})$, we conclude that the probability (Lebesgue measure in parameter space) that $\mathcal{R}_{\alpha,\tilde{\alpha}} \in \mathcal{L}_{\varepsilon/2}$ is bounded below by some uniform $\eta > 0$. Since we are dealing with $p \approx \varepsilon^{-d/2}$ distinguished pairs, and the behavior of each one is ruled by a different parameter in an independent fashion, the probability that (19) fail for all these pairs is bounded above by

$$(1-\eta)^{\varepsilon^{-d/2}} \ll 1.$$

Summarizing these steps, given any $(\theta, \tilde{\theta}, \lambda, t) \in \Sigma \times \tilde{\Sigma} \times \mathbb{R}^* \times \mathbb{R}$ the majority of values of the parameter gives rise to Cantor sets $(K_\xi, \tilde{K})$ such that $\mathcal{R}_{\alpha,\tilde{\alpha}}$ is in the interior of $\mathcal{L}_\varepsilon$ for some distinguished pair $(\alpha, \tilde{\alpha})$. We still need to find some parameter value such that this holds for all $(\theta, \tilde{\theta}, \lambda, t)$ simultaneously. This uses the following pair of observations.

On the one hand, the previous estimates have some room for accommodating variation of the point. Using continuity of the renormalization operators $\mathcal{R}_{\alpha,\tilde{\alpha}}$, one checks that (19) implies

$$\mathcal{R}_{\alpha,\tilde{\alpha}}(\theta', \tilde{\theta}', \lambda', t') \in \operatorname{int}(\mathcal{L}_\varepsilon)$$

for every $(\theta', \tilde{\theta}', \lambda', t')$ in a ball of radius ε^3. On the other hand, $\mathrm{HD}(\Sigma) = \mathrm{HD}(K)$ and analogously for $\tilde{K}$. It follows that the Hausdorff dimension of $\Sigma \times \tilde{\Sigma} \times \mathbb{R}^* \times \mathbb{R}$ is less than 4, and so compact subsets are covered by not more than ε^{-12} balls of radius ε^3. Therefore, the overall probability of the set of bad parameters is bounded above by

$$\varepsilon^{-12}(1-\eta)^{\varepsilon^{-d/2}} \ll 1.$$

This completes the outline of the proof of Theorem 3.19.

3.5. **Homoclinic tangencies in higher dimensions.** Extension of this theory to higher dimensions involves some fundamental difficulties. A first obstruction is the lack of regularity of invariant foliations with codimension bigger than 1: in general the holonomy maps are not more than Hölder continuous, regardless of the smoothness of the dynamics. Related to this, the following basic problem remains wide open, outside the codimension 1 case:

Problem 3.23. Let H be a hyperbolic transitive set. Are the local Hausdorff dimension $\mathrm{HD}(H \cap W^*_{loc}(x))$ and the local limit capacity $c(H \cap W^*_{loc}(x))$, $* \in \{u, s\}$, independent of the point $x \in H$?

Another main difficulty is that, instead of Cantor sets on the real line, one now has to deal with fractal sets living in higher dimensional spaces. In a sense, such sets are still dynamically defined, but the dynamics is usually not conformal, besides being only Hölder continuous, which makes the geometry much harder to analyze. The Hausdorff dimension and limit capacity no longer coincide, in general, and the notion of thickness does not apply to such higher dimensional fractal sets. However,

[7] Such a family exists with C^1 bounded norm. The argument extends to the C^r topology, any $r \geq 1$, considering only $p \approx \varepsilon^{-1/(r+1)}$ distinguished intervals and parameters.

Buzzard [83] has extended Theorem 3.3 to holomorphic maps of $\mathbb{C}^2$, using an astute variation of the thickness strategy.

In addition, contrary to the surface case, the Hausdorff dimension and the limit capacity of hyperbolic sets of higher dimensional systems may vary discontinuously with the dynamics. The following simple example, is from [62]:

Example 3.24. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a diffeomorphism exhibiting a horseshoe $\Sigma = \cap_{n\in\mathbb{Z}} f^n([-1,1]^2)$. For simplicity, suppose that $f^{-1}([-1,1]^2) \cap [-1,1]^2$ has two connected components, A_1 and A_2, and the restriction of f to each component is affine:

$$Df \mid A_i = \begin{pmatrix} \lambda_i & 0 \\ 0 & \sigma_i \end{pmatrix}, \qquad \|\lambda_i\|,\ \|\sigma_i^{-1}\| < 1/2.$$

Then $A_i = [-1,1] \times B_i$ for some interval B_i. Define $f_t : \mathbb{R}^3 \to \mathbb{R}^3$, $t \in [-1,1]$ by

$$f_t(x,y,z) = \begin{cases} (f(x,y), \sigma z) & \text{if } y \in B_1 \\ (f(x,y), \sigma z - t) & \text{if } y \in B_2 \end{cases}$$

Let Λ_t be the continuation for f_t of the hyperbolic set $\Lambda_0 = \Sigma \times \{0\}$ of f_0. In other words, Λ_t is the homoclinic class of the fixed point at the origin. Assuming $\sigma \in (1,2)$, one checks that the projection of Λ_t onto the z coordinate contains the interval $[0,t]$. Consequently, $c(\Lambda_t) \geq \mathrm{HD}(\Lambda_t) \geq 1$ for every $t \neq 0$. Choosing f so that $c(\Sigma) = \mathrm{HD}(\Sigma) < 1$, the Hausdorff dimension and limit capacity of Λ_t are *not* upper semi-continuous at $t = 0$.

Problem 3.25. Are the Hausdorff dimension and the limit capacity of hyperbolic sets always lower semi-continuous functions of the dynamical system ?

3.5.1. *Intrinsic differentiability of foliations.* Some of these difficulties have been dealt with by Palis, Viana [266], in their proof that Theorem 3.6 does extend to arbitrary manifolds. To find attractors, one assumes that the periodic saddle involved in the tangency is *sectionally dissipative*: the product of any two eigenvalues of the derivative at this point has norm less than 1; in particular, the unstable manifold is 1-dimensional.

Theorem 3.26 (Palis, Viana [266])**.** *Let $f : M \to M$ be a diffeomorphism on any manifold, having a homoclinic tangency associated to a sectionally dissipative saddle point p. Then there exists an open subset $\mathcal{U}$ of* $\mathrm{Diff}^2(M)$ *containing f in its closure, such that every $g \in \mathcal{U}$ may be approximated by a diffeomorphism with a homoclinic tangency associated to the continuation of p. Moreover, there exists a residual subset $\mathcal{R} \subset \mathcal{U}$ such that every $g \in \mathcal{R}$ has an infinite number of periodic attractors.*

For persistence of tangencies one does not need the dissipativeness condition. This was proved, independently, by Romero [298], who reduces the general case to the sectionally-dissipative one, and by Gonchenko, Shil'nikov, Turaev [139].

A key novelty in the proof of Theorem 3.26 is the use of a notion of *intrinsic differentiability* à la Whitney. A function $\varphi : X \mapsto \mathbb{R}^n$ on a subset $X \subset \mathbb{R}^m$ is *intrinsically* C^1 if there exists a continuous map $X \times X \ni (x,z) \mapsto \Delta\varphi(x,z) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\varphi(x) - \varphi(z) = \Delta\varphi(x,z) \cdot (x - z) \quad \text{for every } (x,z) \in X \times X.$$

In simple terms, it is shown that if the hyperbolic set admits a strong-stable foliation such that each leaf contains a unique point of the set, then the unstable holonomy maps are intrinsically C^1. A more precise statement is given below.

Let us explain in more detail how the proof of Theorem 3.26 goes. Keep in mind the 2-dimensional arguments outlined in the context of Theorem 3.6. Just as in the surface case, we may assume that p is part of a nontrivial hyperbolic set H_1. Another useful observation is that an arbitrarily small unfolding of the homoclinic tangency yields the formation of other homoclinic tangencies, associated to periodic saddle points having a unique weakest contracting eigenvalue. Thus, we may assume right from the start that the saddle p has a unique weakest contracting eigenvalue which, as a consequence, is a real number.

Then we may choose the hyperbolic set H_1 to be transverse to the strongest contracting directions of $Df^{\mathrm{per}(p)}(p)$, in the following sense: at every $x \in H_1$, the *intrinsic tangent space*

$$IT_x(H_1) = \operatorname{span}\left\{ v\colon \text{ there is } (x_n)_n \in H_1^{\mathbb{N}} \text{ so that } \frac{x_n - x}{\|x_n - x\|} \to v \right\}$$

(consider $M = \mathbb{R}^m$, for simplicity) is 2-dimensional and transverse to the codimension 2 subspace generated by the strongest contracting directions. For such an H_1 one proves that the projections $\pi\colon \Sigma_0 \cap W^u(H_1) \to \Sigma_1 \cap W^u(H_1)$ along unstable leaves of H_1 are intrinsically C^1 maps.

Now we may define the *local unstable thickness* of H_1 at a point $x \in H_1$ as

$$\tau^u(H_1, x) = \tau(\tilde{\pi}(W^u(H_1) \cap \Sigma_0), \tilde{\pi}(x)),$$

where Σ_0 is a transverse section to unstable foliation of H_1 at the point x, and $\tilde{\pi}\colon W^u(H_1) \cap \Sigma_0 \to \mathbb{R}$ is any intrinsically differentiable map admitting an intrinsic derivative $\Delta\tilde{\pi}$ such that $\Delta\tilde{\pi}(x,x) \mid IT_x(W^u(H_1) \cap \Sigma_0)$ is a bijection. Intrinsic differentiability of the unstable foliation permits to check that this definition does not depend on the choices of Σ_0 and $\tilde{\pi}$, and to prove that $\tau^u(H_1, x)$ is strictly positive and independent of $x \in H_1$. Furthermore, it varies continuously with the diffeomorphism $f \in \mathrm{Diff}^2(M)$.

Next, by arbitrarily small perturbations leaving H_1 and $\tau^u(H_1)$ essentially unchanged, one obtains another hyperbolic basic set, H_2, with codimension 1 stable foliation and large stable thickness: $\tau^u(H_1)\tau^s(H_2) > 1$. This is done by a comparatively straightforward extension of the renormalization scheme used in the surface case. As before, H_1 and H_2 may be taken to be heteroclinically related, and to exhibit a tangency q between $W^u(p_1)$ and $W^s(p_2)$, for some periodic points $p_1 \in H_1$ and $p_2 \in H_2$. The stable foliation of H_2, being of codimension 1, admits an extension to a C^1 foliation $\mathcal{F}^s(H_2)$ of a neighborhood of H_2, which we may assume to contain q.

An implicit function argument shows that there exists an intrinsically differentiable map $\pi_1\colon W^u(H_1) \cap W^s_{loc}(p_1) \to M$ such that $\pi_1(p_1) = q$ and each $\pi_1(x)$, $x \in W^u(H_1) \cap W^s_{loc}(p_1)$, is a point of tangency between leaves of $W^u(H_1)$ and $\mathcal{F}^s(H_2)$. See Figure 7. Let π_2 be the projection along the leaves of $\mathcal{F}^s(H_2)$ onto $W^u_{loc}(p_2)$, which we identify with an interval in $\mathbb{R}$. Then the theorem follows from an application of the gap lemma to the Cantor sets $\pi_2 \circ \pi_1(W^u(H_1) \cap W^s_{loc}(p_1))$ and $W^s(H_2) \cap W^u_{loc}(p_2)$. This finishes our sketch of the proof of Theorem 3.26.

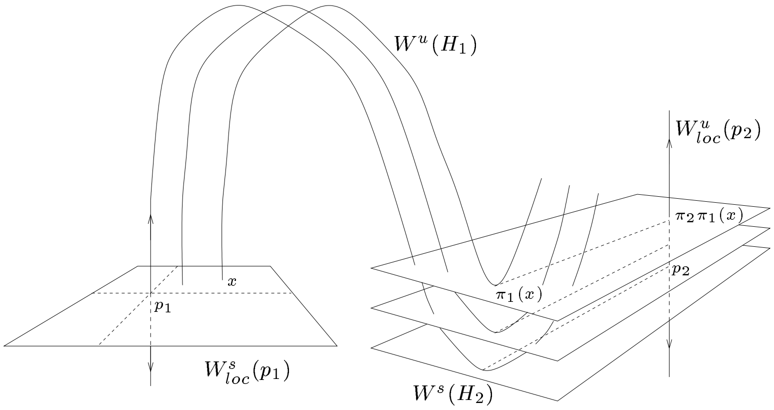

FIGURE 7. Line of tangencies in higher dimensions

3.5.2. *Frequency of hyperbolicity.* More recently, Moreira, Palis, Viana announced in [235] that the principle

hyperbolicity prevails $\Leftrightarrow$ *Hausdorff dimension is smaller than* 1

remains valid for homoclinic bifurcations in any dimension.

They consider C^2 families of diffeomorphisms $f_\mu : M \to M$, $\mu \in (-1,1)$ unfolding a generic homoclinic tangency associated to a periodic point p contained in a horseshoe Λ of $f = f_0$. It is assumed that $\mu = 0$ is the first bifurcation: f_μ is uniformly hyperbolic for all $\mu < 0$; this implies that the weakest contracting and weakest expanding eigenvalues of p are both real. The main theorem is stated in terms of the following notions of upper stable and unstable Hausdorff dimensions.

Let Λ be a horseshoe for a diffeomorphism $f : M \to M$. Let $\mathcal{V}_n$ be the set of vertical n-cylinders, that is, the subsets of Λ defined by prescribing the first n symbols in the backward itinerary with respect to some Markov partition. There is a dual notion of horizontal n-cylinder, where one considers forward itinerary instead. For each $V \in \mathcal{V}_n$ let $D_s(V) = \sup\{\mathrm{diam}(W^s_{loc}(x) \cap V) : x \in \Lambda \cap V\}$. Define λ_n by the relation

$$\sum_{V \in \mathcal{V}_n} D_s(V)^{\lambda_n} = 1.$$

The *upper stable dimension* of Λ is defined by $\bar{d}_s(\Lambda) = \lim_{n\to\infty} \lambda_n$. There is a dual notion of *upper unstable dimension* $\bar{d}_u(\Lambda)$, dealing with W^u_{loc} instead of W^s_{loc}.

It is not difficult to show that the limit always exists, and $\bar{d}_s$ is an upper semi-continuous function of φ. Moreover,

$$\mathrm{HD}(W^s_\varepsilon(x) \cap \Lambda) \leq \bar{d}_s(\Lambda) \quad \text{for each } x \in \Lambda.$$

When $\bar{d}_s(\Lambda) < 1$ we even have equality, for a residual subset $\mathcal{R}$ of diffeomorphisms g in a neighborhood of f and every point x in the continuation Λ_g of Λ. In addition, $\mathcal{R}$ may be taken so that its elements are points of continuity of $\bar{d}_s$.

Theorem 3.27 (Moreira, Palis, Viana [235]). *There are open sets $\mathcal{U}$ and $\mathcal{V}$ of families of diffeomorphisms unfolding a first tangency such that $\mathcal{U} \cup \mathcal{V}$ is dense and*

(1) *For families in $\mathcal{U}$ we have $\bar{d}_s(\Lambda)+\bar{d}_u(\Lambda) < 1$ and $\mathcal{H} = \{\mu \mid f_\mu$ is hyperbolic$\}$ has full Lebesgue density at $\mu = 0$:*
$$\lim_{\delta\to 0} \frac{m(\mathcal{H}\cap[0,\delta])}{\delta} = 1.$$
(2) *For families in $\mathcal{V}$ we have $\bar{d}_s(\Lambda)+\bar{d}_v(\Lambda) > 1$ and $\mathcal{N} = \{\mu \mid f_\mu$ has persistent homoclinic tangencies associated to the continuation Λ_μ of $\Lambda_0\}$ has positive Lebesgue density at $\mu = 0$:*
$$\liminf_{\delta\to 0} \frac{m(\mathcal{N}\cap\ [0,\delta])}{\delta} > 0.$$

Let us say a few words about the proof of this result. The main step towards proving part (1) is to reduce to the case when the horseshoe Λ_0 does not intersect the strong-stable manifold, nor the strong-unstable manifold of the point p_0:

Proposition 3.28. *Assume $\bar{d}_s(\Lambda) < 1$. For an open and dense subset of maps in some neighborhood of f we have $\Lambda\cap W^{ss}(p) = \{p\}$. Analogously for $\bar{d}_u$ and W^{uu}.*

This proposition implies that $\Lambda\cap W^s(p)$ is contained in a cuspidal region around the weak-stable direction of p. Let K^s be the projection of $\Lambda\cap W^s(p)$ to this weak-stable direction along the strong-stable foliation of p. Then, for any $\delta > 0$ there exists $\varepsilon > 0$ such that each $x \in \Lambda\cap W^s(p)$ in the ε-neighborhood of p is within distance $\varepsilon\delta$ from some $y \in K_s$. We say that $\Lambda\cap W^s(p)$ is well-represented by K_s near p.

Then a similar fact is true replacing $W^s(p)$ by any other cross-section Σ to the unstable direction at any $q \in W^u(p)$: the intersection $W^u(p)\cap\Sigma$ is well-represented by K_s near q. Using also smoothness of the unstable leaves, together with Hölder continuity of their tangent bundle, we conclude that the intersection of $W^u(\Lambda)$ with a neighborhood of q is well-represented by the product of K_s by $T_qW^u(p)$ near q.

Consider q to be the point of homoclinic tangency. The previous paragraph, together with a dual statement for $W^s(\Lambda)$, mean that the stable and unstable foliations of Λ are well-represented by affine models
$$K_s \times T_qW^u(p) \quad\text{and}\quad K_u \times T_qW^s(p)$$
near q. By continuity of the invariant foliations this remains true when the tangency is unfolded along a one-parameter family f_μ: $W^s(\Lambda_\mu)$ and $W^u(\Lambda_\mu)$ are well-represented by translates
$$[K_s \times T_qW^u(p)] + \mu v_1 \quad\text{and}\quad [K_u \times T_qW^s(p)] + \mu v_2$$
for some vectors v_1, v_2. So, to ensure that the two foliations are transverse, with angles bounded from zero, it suffices to prove a corresponding fact for these affine models. Thus, the original situation is once more reduced to a problem of arithmetic difference of Cantor sets.

The main ingredient for proving part (2) of the theorem is the construction, perhaps after perturbation of f, of strong-stable and strong-unstable foliations of codimension 1 for hyperbolic subsets of Λ with almost the same upper Hausdorff dimensions. These foliations are used to (essentially) reduce the study of the geometries of the stable and unstable foliations near the initial homoclinic tangency to the 2-dimensional case. More precisely, in case (2) of Theorem 3.27,

Proposition 3.29. *Generically, there are disjoint subsets Λ_1 , Λ_2 of Λ, invariant by some iterate f^k, with the following properties:*

(1) Λ_1 *has a globally defined strong-stable foliation, and* Λ_2 *has a globally defined strong-unstable foliation, both invariant under* f_0.
(2) $\bar{d}_s(\Lambda_1) < 1$, $\bar{d}_u(\Lambda_2) < 1$, *and* $\bar{d}_s(\Lambda_1) + \bar{d}_u(\Lambda_2) > 1$.
(3) *For each* $p_1 \in \Lambda_1$ *the projection* $\pi_{ss} \mid (\Lambda_1 \cap W^s(p_1))$ *to the weak stable direction along strong-stable leaves is injective and its inverse is intrinsically* C^1 *with Hölder continuous intrinsic derivative. Analogously for* $p_2 \in \Lambda_2$ *and the projection* $\pi_{uu} \mid (\Lambda_2 \cap W^u(p_2))$ *along strong-unstable leaves.*
(4) *Either* Λ_1 *contracts area in the weak direction or* Λ_2 *expands area in that weak direction.*

Conditions (1) to (3) imply that the projection of $\Lambda_1 \cap W^s(p_1)$ onto a weak-stable (1-dimensional) manifold, along the strong-stable foliation, is a regular $C^{1+\varepsilon}$ Cantor set K^s, whose conjugacy class does not depend on $p_1 \in \Lambda_1$ nor on the choice of the weak-stable leaf. The same holds for $\pi_{uu}(\Lambda_2 \cap W^u(p_2)) = K^u$.

The next lemma concludes the reduction to a problem of stable intersection of Cantor sets. It is assumed that the eigenvalues of f at the point p satisfy a generic non-resonance condition.

Lemma 3.30. *If there are* $\lambda > 0$ *and* $t \in \mathbb{R}$ *such that* K^s *intersects* $\lambda K^u + t$ *stably, then* $(f_\mu)_\mu$ *persistently exhibits positive density of stable tangencies at* $\mu = 0$.

In order to get the stable intersection between K^s and $\lambda K^u + t$, we first prove that condition (4) implies that either K^s or K^u is a C^2-regular Cantor set on the central leaf. This is important to show that, generically, K^s and K^u satisfy the hypothesis of the scale recurrence lemma in Section 3.4.2. Afterwards, we consider again a family with a large number of parameters of local perturbations of the diffeomorphism f and show that for most parameters there are λ and t such that K^s stably intersects $\lambda K^u + t$.

4. Hénon-like dynamics

The two-parameter Hénon family of transformations on the plane

$$H_{a,b} : (x, y) \mapsto (1 - ax^2 + y, bx) \tag{20}$$

was initially proposed in [152, 153] as the simplest model of an invertible dynamical system with a strange attractor. It has become the prototype for a whole class of dynamical behaviors, which we shall call Hénon-like. Their most distinctive features are

- non-hyperbolic behavior, of homoclinic (folding) type, in well localized regions of phase space,
- combined with a fair amount of hyperbolicity away from those so-called critical regions.

The first one is associated to the presence of homoclinic tangencies which very much influence the way the dynamics unfolds as the parameters vary.

Although here we discuss only real systems, let us mention that the complex counterpart to (20) has also been intensively studied over the last decade or so, especially by Bedford and Smillie, Fornaess and Sibony, and Hubbard and his coauthors, as a fundamental step in extending the theory of iterations of rational maps on the sphere to higher dimensional holomorphic systems.

4.1. **Hénon-like families.** Most results on the dynamics of Hénon maps are based on a perturbative approach initiated by Benedicks, Carleson [41], and have been proven when the Jacobian $-b$ is very close to zero. Notice that (20) is affinely conjugate to

$$(x, y) \mapsto (1 - ax^2, 0) + \sqrt{|b|}\,(y, \operatorname{sign}(b)\,x)$$

so that for $|b|$ small we may think of $f_a = H_{a,b}$ as a singular perturbation of the family of quadratic interval maps $\phi_a : x \mapsto 1 - ax^2$. This led Mora, Viana [223, 333] to consider more general families

$$(x, y) \mapsto f_a(x, y) = (1 - ax^2, 0) + R(a, x, y) \tag{21}$$

where R has small norm, in the C^3 topology say. One important motivation was that such families model a good part of the dynamics associated to dissipative homoclinic tangencies, as we shall comment upon later.

Other applications, like the unfolding of saddle-node cycles [113] or kick-forced invariant curves [340] required further generalization. It was observed by Díaz, Rocha, Viana [113] and by Wang, Young [339] that the properties of the quadratic family relevant in this context are shared by much more general C^3 one-parameter families $\phi_a : M \to M$ of multimodal maps of the circle or the interval. Indeed, the results we are going to discuss extend to all families $(f_a)_a$ of d-dimensional embeddings obtained by singular perturbation of such families $(\phi_a)_a$ satisfying conditions (α), (β), (γ) that follow.

The first condition in the definition of these *Hénon-like families* $(f_a)_a$ corresponds to the existence of a Misiurewicz parameter:

(α) There exists $a = \bar{a}$ such that $\phi_{\bar{a}}$ has negative Schwarzian derivative, all its critical points are quadratic, and the critical set is non-recurrent.

As a consequence, all periodic points are hyperbolic and repelling: use Singer [317] and also recall that our definition of multimodal map includes $\phi_a(\partial M) \subset \partial M$.

The second condition is that the kneading sequences of the critical orbits of $\phi_{\bar{a}}$ unfold generically with the parameter a close to $\bar{a}$:

(β) Let $c_i(a)$ be the continuation, for nearby parameters, of each critical point $c_i(\bar{a})$ of $\phi_{\bar{a}}$. Also for nearby parameters, let $x_i(a)$ be the unique point whose ϕ_a-itinerary, relative to the critical points, coincides with the $\phi_{\bar{a}}$-itinerary of the critical value $\phi_{\bar{a}}(c_i(\bar{a}))$. Then,

$$\frac{d}{da}\phi_a(c_i(a)) \neq \frac{d}{da}x_i(a) \quad \text{at } a = \bar{a}.$$

For $d \geq 2$, let B be the unit ball of dimension $d - 1$ around zero. Then we consider one-parameter families of embeddings

$$f_a : M \times B \to M \times B\,, \qquad f_a(x, y) = (\phi_a(x, y), 0) + R(a, x, y),$$

such that the C^3 norm of R is small and the Jacobian of f_a is homogeneous, in the following sense:

(γ) For all a, x, y, we have $\|R(a, x, y)\|_{C^3} \leq C\sqrt{b}$,

$$C^{-1}b \leq |\det Df_a(x, y)| \leq Cb \quad \text{and} \quad \|D(\log|\det Df_a(x, y)|)\| \leq C$$

for some sufficiently small b (depending also on the constant C, which is arbitrary).

For parameters a close to $\bar{a}$, the critical regions of the diffeomorphism f_a correspond to neighborhoods of the critical points of $\phi_{\bar{a}}$, whereas hyperbolic behavior away from the critical regions results from the kind of phenomenon in Theorem 2.5, as we shall see.

In our presentation we consider more closely the original 2-dimensional setting (21), which contains most main ingredients, referring the reader to [113, 339] for those more general formulations. Hénon-like families in higher dimensional manifolds have been studied in [333] and, more recently, [113, 238, 338].

4.1.1. *Identifying the attractor.* Let us consider $\phi_a(x) = 1 - ax^2$ with $a \in (1,2)$. The map ϕ_a has exactly two fixed points q_a and p_a , with

$$q_a < -1 < 0 < p_a < 1 < -q_a\,, \quad \lim_{a\to 2} q_a = -1, \quad \lim_{a\to 2} p_a = 1/2,$$

and they are both repelling. Denoting $M_a = [q_a, -q_a]$, we have $\phi_a(M_a) = M_a$ and $\phi_a(\partial M_a) \subset \partial M_a$ (the fact that M_a depends on the parameter is easily disposed of by rescaling). Moreover, we may find an open interval $I \subset M_a$, depending only on a lower bound for $2-a$, such that $\phi_a(\bar{I}) \subset I$: just take $I = (-r, r)$ with $1 < r < q_a$. Then given any higher-dimensional map $f_a = \phi_a + R$ and any $\varepsilon > 0$, we have $f_a(\bar{I} \times [-1,1]) \subset I \times (-\varepsilon, \varepsilon)$ if the norm $\|R\|$ is sufficiently small. This means there is some compact topological "attractor"

$$\Lambda_a = \bigcap_{n=0}^{\infty} f_a^n(\bar{I} \times [-1,1])$$

containing all the asymptotic dynamics of f_a.

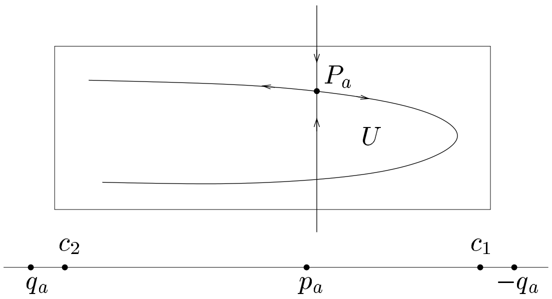

FIGURE 8. Hénon-like attractor

Let us give a more explicit description of this set. For this purpose, consider the fixed point p_a of ϕ_a and let $c_j = \phi_a^j(c)$ be the iterates of the critical point $c = 0$. The unstable set $W^u(p_a)$ of p_a , defined as the union of all forward iterates of some small neighborhood, coincides with the interval $[c_2, c_1]$. By hyperbolic continuation, p_a gives rise to a fixed hyperbolic saddle point $P_a \approx (p_a, 0)$ for the embedding f_a. The unstable manifold $W^u(P_a)$ is contained in $I \times (-1,1)$, and so its closure is contained in Λ_a . It is easy to show that this closure attracts whole open sets of orbits. The first step is to note that, by continuous dependence of unstable sets on the dynamics, $W^u(P_a)$ is close to $W^u(p_a) \times \{0\}$ in compact parts. In particular, $W^u(P_a)$ must contain a parabola-shaped curve whose legs project to intervals close to $[c_2, c_1]$ on the horizontal axis. See Figure 8. Similarly, by continuous dependence of stable sets on the dynamics, the stable manifold $W^s(P_a)$ must contain a long nearly vertical segment through the fixed point. It follows that P_a has transverse homoclinic points, and so there are open regions U bounded by segments of the stable and the unstable manifold. Using that f_a dissipates area (Jacobian close to zero) and contracts the stable manifold, we immediately get that the iterates $f_a^n(U)$ approach the unstable manifold $W^u(P_a)$ uniformly as $n \to \infty$, so that the ω-limit set of every point of U is contained in the closure of $W^u(P_a)$.

More elaborate versions of such arguments permit to prove (see [43, 87, 334]) that in most situations the basin of attraction of the closure of $W^u(P_a)$ contains the whole $\bar{I} \times [-1, 1]$, and so

$$\Lambda_a = \overline{W^u(P_a)}.$$

4.1.2. *Hyperbolicity outside the critical regions.* The starting point to trying to understand the dynamics on the attractor is the observation that the orbits of Hénon-like maps behave in a uniformly hyperbolic fashion as long as they remain far from the critical points. At the origin is the following one-dimensional lemma:

Lemma 4.1. *Let $(\phi_a)_a$ be any C^3 family of multimodal maps satisfying condition (α) above. Then there exist $\sigma > 1$ and $c > 0$ such that for every small $\delta > 0$ there exists $c(\delta) > 0$ and a neighborhood $\Omega = \Omega(\delta)$ of the parameter $\bar{a}$ such that for every $a \in \Omega$,*

1. *$|(\phi_a^n)'(x)| \geq c(\delta)\sigma^n$ for any x such that $\phi_a^i(x)$ is outside the δ-neighborhood of the critical set for all $0 \leq i < n$.*
2. *$|(\phi_a^n)'(x)| \geq c\sigma^n$ if, in addition, $\phi_a^n(x)$ is inside the δ-neighborhood of the critical set.*

For the proof, begin by considering the Misiurewicz parameter $\bar{a}$. Let $\Delta(\delta)$ denote the δ-neighborhood of the critical set. A key ingredient is that for every small $\delta > 0$ there exist $m(\delta) \geq 1$ and $\sigma(\delta) > 1$ such that

$$|(\phi_{\bar{a}}^{m(\delta)})'(x)| > \sigma(\delta)^{m(\delta)} \quad \text{for any } x \text{ whose first } m(\delta) \text{ iterates remain outside } \Delta(\delta).$$

This is similar to Theorem 2.5 and was first proved by Misiurewicz [220] in this context. The assumption that the critical points are non-recurrent also implies that the images of the monotonicity intervals of $\phi_{\bar{a}}^n$ have endpoints far from the critical set. Using that the Schwarzian derivative is negative one shows that there exists a uniform constant $c_0 > 0$ such that

$$|(\phi_{\bar{a}}^r)'(x)| > c_0 \quad \text{for any } r \geq 1 \text{ and } x \text{ such that } \phi_{\bar{a}}^r(x) \text{ is in } \Delta(\delta),$$

as long as δ is small enough. The claims in the lemma would follow directly, if we were not requiring the expansion rate to be independent of δ. To get this we proceed as follows. Fix $\delta_0 > 0$ and $1 < \sigma_0 < \sigma(\delta_0)$ and then choose a smaller $\delta_1 > 0$ such that for every $x \in \Delta(\delta_1)$ there exists some large $p = p(x) \geq 1$ such that $\phi_{\bar{a}}^i(x)$ is outside $\Delta(\delta_0)$ for all $1 \leq i < p$ and

$$|(\phi_{\bar{a}}^p)'(x)| > \sigma_0^p\,.$$

It is no restriction to suppose that $\sigma = \sigma(\delta_1)$ is smaller than $\sigma < \sigma_0$. Now, given any x and $n \geq 1$ such that both x and $\phi_{\bar{a}}^n(x)$ are in $\Delta(\delta_1)$ and no intermediate iterate is in $\Delta(\delta_1)$, we may write $n = p + qm + r$ with $0 \leq r < m$, and deduce that

$$|(\phi_{\bar{a}}^n)'(x)| > \sigma_0^p\, \sigma^{qm} c_0 > \sigma^n \tag{22}$$

(reducing δ_1 increases p without affecting σ_0 or c_0). The conclusion of the lemma now follows, for all $\delta \leq \delta_1$, by breaking the time interval $\{0, 1, \dots, n\}$ according to the passages of the orbit through $\Delta(\delta_1)$, and defining

$$c = (c_0\sigma^{-m})^2 \quad \text{and} \quad c(\delta) = c_0\sigma^{-m} \inf_{r,x} |(\phi_{\bar{a}}^r)'(x)|\, \sigma^{-r}$$

where the infimum is over all $1 \le r < m$ and x whose first iterates are outside $\Delta(\delta)$. Finally, notice that m is fixed and p is bounded by some constant that depends only on δ, since the orbit avoids $\Delta(\delta)$. Therefore, all these estimates extend by continuity to any parameter a in a neighborhood $\Omega(\delta)$ of $\bar{a}$. This completes the argument.

Consider a Hénon-like map $f_a(x,y) = (\phi_a(x), 0) + R(a,x,y)$ with b much smaller than δ. Given any x outside $\Delta(\delta)$ and any tangent vector $(1,v)$ with $|v|$ small, we have

$$Df_a(x,y)\cdot(1,v) \approx \phi_a'(x)\,(1,w)$$

where $|w|$ is again small. Thus, from Lemma 4.1 one easily gets that, over orbits avoiding the critical region $\Delta(\delta)\times B$, nearly horizontal vectors are uniformly expanded and mapped to nearly horizontal vectors:

Proposition 4.2. *Let $(\phi_a)_a$ be any C^3 family of multimodal maps satisfying condition (α). Then for every small $\delta > 0$ there exists $b(\delta) > 0$ such that if $(f_a)_a$ is a Hénon-like family as in (γ) with $b \le b(\delta)$, then for any $a \in \Omega$ we have*

(1) $\|Df_a^n(z)\| \ge c(\delta)\sigma^n$ *for any z such that $f_a^i(x) \notin \Delta(\delta)\times B$ for all $0 \le i < n$.*
(2) $\|Df_a^n(z)\| \ge c\sigma^n$ *if, in addition, $f_a^n(z)$ is in $\Delta(\delta)\times B$.*

Since these maps are area dissipative, this also implies uniform contraction in a complementary direction: tangent vectors that are mapped close to the vertical direction by the derivative are contracted by a factor $\approx (Cb/\sigma)^n$.

4.1.3. *Abundance of strange attractors.* The real difficulty with the dynamics of Hénon-like maps arises from the behavior in the critical region $\Delta(\delta)\times B$, where expanding (roughly horizontal) directions are mapped to contracting (roughly vertical), with the possible consequence that all the expansion gets cancelled out. As the unstable manifold of the fixed point P_a crosses the critical region it develops folds, that are then propagated under iteration. Since the local stable manifold of P_a is roughly vertical, one may expect that varying the parameter a will lead to the formation and unfolding of homoclinic tangencies. Indeed, such homoclinic bifurcations are quite apparent in numerical experiments. According to Lemma 3.5 and Theorem 3.6, the unfolding of homoclinic tangencies of area dissipative maps leads to the formation of periodic attractors. As pointed out by Newhouse, a periodic attractor with very high period is numerically indistinguishable from a non-periodic attractor. The possible presence of periodic attractors [8] turns out to be a major obstacle in proving Hénon's assertion that a "strange" attractor occurs in (20). In fact, this assertion remains unproven, for parameter values close to $(a,b) = (1.4, 0.3)$ as he was considering originally.

However, by the end of the eighties, Benedicks, Carleson were able to prove that some non-uniformly hyperbolic attractor does exist in the Hénon family (20), for a positive Lebesgue measure set of parameters (a,b) with b close to zero:

[8] Non-periodic hyperbolic attractors are ruled out in this dissipative setting, by the structure theorem of Plykin [276].

Theorem 4.3 (Benedicks, Carleson [41]). *For any small $b > 0$, there exists a positive Lebesgue measure set of values of a for which the Hénon map $f_a = H_{a,b}$ has a strange (non-hyperbolic) attractor $\Gamma_a = \overline{W^u(P_a)}$: there exists $z_1 \in \Gamma_a$ whose forward orbit is dense in Γ_a and there exists a constant $\sigma > 1$ such that*

$$\|Df_a^n(z_1)\| \geq \sigma^n \quad \text{for every } n \geq 1. \tag{23}$$

In particular, for these parameter values the closure of $W^u(P_a)$ contains no periodic attractors.

This remarkable result opened the way for much progress in understanding Hénon-like dynamical behavior. Shortly afterwards, Mora, Viana [223] extended Benedicks, Carleson's approach to the general setting of (21), thus freeing the arguments from any dependence on the explicit expression of the Hénon maps, and also established a connection between these systems and the unfolding of homoclinic tangencies we shall review in Section 4.5. Moreover, Viana [333] extended the conclusions of [223] to arbitrary dimension. The ergodic theory of Hénon-like systems was then developed, as we shall discuss in the next sections. At about the same time, Díaz, Rocha, Viana [113] extended [223] to the strange attractors arising from saddle-node cycles and, in doing so, observed that the original approach applies to perturbations of very general families of unimodal or multimodal maps in one dimension, besides the quadratic family. More recently, Wang, Young [339] showed that the theory extends to the full generality of Hénon-like families, in essentially the sense we consider here. Palis, Yoccoz [268] have been applying a related analysis to non-uniformly hyperbolic sets of saddle type arising from homoclinic tangencies, as we shall see in Section 4.5. Similar ideas have been used in several other related contexts: saddle-focus bifurcations of flows, by Pumariño, Rodriguez [288, 289]; saddle-node horseshoes, by Costa [96]; invariant curves subject to kick forcing, by Wang, Young [340]; Lorenz-like attractors without invariant foliations, by Luzzatto, Viana [197].

4.1.4. *Critical points of dissipative diffeomorphisms.* A detailed survey of the proofs of the results in [41, 223] has appeared recently, by Luzzatto, Viana [198]. Here we restrict ourselves to a few general comments.

The strategy introduced by Benedicks, Carleson for proving is inspired by alternative approaches to the proof of Jakobson's Theorem 2.11, especially by Collet, Eckmann [92] and Benedicks, Carleson [40] themselves, based on controlling the recurrence of the critical orbits through parameter exclusions. Let $(\phi_a)_a$ be a family of one-dimensional maps satisfying condition (α) above, and $\mathcal{C} = \{c_i\}$ represent the set of critical points of each of the maps. The aim is to find a constant $\sigma > 1$ and a positive Lebesgue measure set of parameters close to $\bar{a}$ such that

$$|(\phi_a^n)'(\phi_a(c_i))| \geq \sigma^n \quad \text{for all } n \geq 1 \text{ and } c_i \in \mathcal{C}. \tag{24}$$

In a few words, on the one hand one collects expansion during the iterates outside $\Delta(\delta)$ using Lemma 4.1. On the other, one forbids the critical set to return close to itself too fast:

$$d(\phi_a^n(c_i), c_j) \geq e^{-\alpha n} \quad \text{for all } n \geq 1 \text{ and } c_i, c_j \in \mathcal{C}, \tag{25}$$

where α is some small positive constant, to ensure that expansion is not completely lost when the critical orbit hits $\Delta(\delta)$. An important ingredient is a recovery period (bound period) following any return of a critical point c_i to $\Delta(\delta)$, at the end of which the expansion lost at the return time n has been completely compensated for. The idea is that the orbit of $\phi_a^n(c_i)$ remains close to the orbit of some critical point c_j (possibly the same) for some period of time, during which information about the latter may be transmitted to the former, by continuity and in an inductive fashion. The recurrence conditions (a large deviations type condition is needed in addition to (25)) are taken as restrictions on the parameter, and shown to hold for a positive Lebesgue measure set of values of a. For such parameters one gets (24). By Singer [317], this implies that the map has no periodic attractors. In fact, it even admits absolutely continuous invariant probability measures whose basins cover a full measure subset of phase space.

The cornerstone of Benedicks, Carleson's approach to the Hénon family is a notion of critical point for surface diffeomorphisms which they use to organize the dynamics of these maps in much the same way as before. In very simplified terms, a *critical point* corresponds to a tangency between Pesin stable and unstable manifolds: the main feature is that the same tangent direction is contracted by both forward and backward iterates of the derivative. For proving Theorem 4.3 it suffices to consider critical points inside the unstable manifold of the fixed point and, in fact, one needs only a subset $\mathcal{C}$. The precise choice of $\mathcal{C}$ and the very definition of these points turn out to be very delicate, though.

Critical points play a dual role in the one-dimensional arguments above. First, the distance to the nearest critical point is a measure of the norm of the derivative, which is why a recurrence condition like (25) can be used to bound the loss of expansion at each return. Next, as explained before, information is passed on from the orbit of c_j to that of $\phi_a^n(c_i)$ during the recovery period in which the two orbits remain close to each other. In order to duplicate this in the surface case, one must be able to associate to each return $f_a^n(c_i)$ of a critical point to the region $\Delta(\delta) \times B$ a suitable critical point c_j such that the loss of hyperbolicity taking place at the return be, in some sense, proportional to the distance from $f_a^n(c_i)$ to c_j. Indeed,

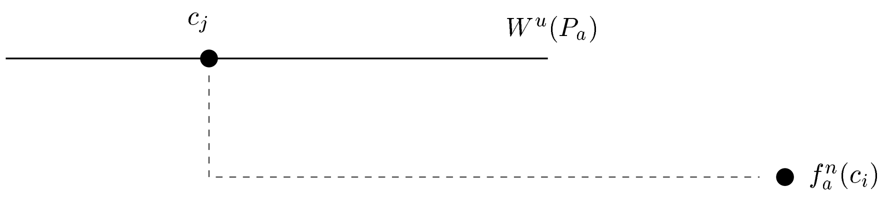

FIGURE 9. Tangential position

one key step in the proof is to *construct* a critical point c_j such that $f_a^n(c_i)$ is in tangential position to it: the horizontal distance is much larger than the vertical distance, as described in Figure 9. Then the same type of estimates as in the one-dimensional situation may be used to bound the loss of hyperbolicity. This point c_j is then used for the recovery argument, which also similar to the one-dimensional case, although technically much harder.

These considerations also mean that the critical set $\mathcal{C}$ can not be chosen too small: it must contain enough points so that a convenient c_j in tangential position may be found at each return. But $\mathcal{C}$ should not be too large either: the more critical points there are, the more parameters must be excluded to impose recurrence conditions like (25) to each one of them. The following compromise is reached: at each step n one deals with not more than $e^{\theta n}$ critical points, where $\theta \approx 1/|\log b|$. In order to control the orbit of each one of them, a subset of parameters is excluded whose measure is bounded by $e^{-\gamma n}$. Here the constant $\gamma > 0$ is independent of b. So, taking b be small enough (this is one step where this is crucial) we ensure that $\sum_n e^{\theta n} e^{-\gamma n}$ is small, so that a positive Lebesgue measure subset of parameters survives all the exclusions.

But the main difficulty, that makes the proof so subtle and involved, lies in the precise definition of a critical point. The reason is that Pesin stable and unstable manifolds are not given a priori: one must assume some amount of hyperbolicity, which is precisely what one wants to prove using these arguments. The way to bypass this is to deal with finite time *approximations to critical points,* defined by means of approximate stable directions.

Roughly speaking, for a point $z \in W^u(P_a)$ such that the horizontal direction $(1,0)$ is uniformly expanded by the first n iterates of the derivative (a hyperbolicity hypothesis), one considers the direction $e^n(z)$ that is most contracted by $Df_a^n(z)$, and one says that z is a *critical value* of order n if $e^n(z)$ is tangent to the unstable manifold of P_a. Then the pre-image of z is a critical point. Wang, Young [339] do not assume points to be inside the unstable manifold of a periodic saddle; then they use approximations for the unstable direction as well. Such order n critical approximations are used at the n:th step of the induction to obtain the hyperbolicity hypothesis at time $n+1$, after parameter exclusions. At this point, critical approximations of order $n+1$ may be defined, and the whole procedure repeated. We refer the reader to the original papers, as well as to [198], for more explanations and details.

4.2. **Sinai-Ruelle-Bowen measures.**

4.2.1. *Existence and uniqueness.*

Theorem 4.4 (Benedicks, Young [45]). *Hénon-like attractors support a unique SRB measure μ. The system (f,μ) is Bernoulli and has a unique positive Lyapunov exponent. The measure has absolutely continuous conditional measures along unstable manifolds.*

4.2.2. *Solution of the basin problem.* The following result solves the basin Problem 1.11 for Hénon-like attractors.

Theorem 4.5 (Benedicks, Viana [43]). *Let Λ be the attractor of a Hénon-like diffeomorphism f, and μ be the SRB measure. Then Lebesgue almost every point z in $B(\Lambda)$ is contained in the stable manifold of some point of the attractor. Moreover, f has the no-holes property: $B(\mu) = B(\Lambda)$ up to a zero Lebesgue measure set.*

4.3. **Markov extensions.** An ergodic system (f,μ) has *exponential decay of correlations* in a Banach space $\mathcal{H}$ of functions if iterates $f^n(\varphi\mu)$ of an initial mass distribution $\varphi\mu$ converge exponentially fast to the equilibrium $\mu \int \varphi \, d\mu$, in the following sense: there exists $\lambda < 1$ such that

$$\Big| \int \psi d f_*^n(\varphi\mu) - \int \psi \, d\mu \int \varphi \, d\mu \Big| \le C(\|\varphi\|, \|\psi\|_1)\lambda^n$$

for every $\varphi \in \mathcal{H}$, $\psi \in L^1(\mu)$, and $n \ge 1$. Replacing the right-hand side by other sequences one obtains notions of slower/faster decay of correlations, e.g. stretched exponential $C\exp(cn^\alpha)$, or polynomial Cn^α.

Theorem 4.6 (Benedicks, Young [46]). *Let Λ be the attractor of a Hénon-like map f and μ be the SRB measure. Then (f, μ) has exponential decay of correlations in the space of Hölder continuous functions.*

4.4. **Stochastic stability.** Despite being metrically persistent in parameter space, chaotic (non-uniformly hyperbolic) dynamics in Hénon-like systems is very unstable under small perturbations of the system. A good measure of this is given by the result of Ures [329] stating that all parameter values in the proof of Theorem 4.3 are approximated by other parameter values for which the map has a homoclinic tangency associated to the fixed point P_a. Consequently, under arbitrarily small perturbations one may cause the closure of $W^u(P_a)$ to contain infinitely many periodic attractors!

Quite in contrast, the main result in this section states that at the statistical level the dynamics is remarkably stable under small random noise. To state this precisely, fix a bounded open neighborhood U of the Hénon-like attractor Λ, contained in the basin of attraction $B(\Lambda)$ and such that $f(U)$ is relatively compact in U. Consider random perturbations $\{\nu_\varepsilon : \varepsilon > 0\}$ where each ν_ε is supported in the ε-neighborhood of f relative to the C^2 topology on the closure of U. Let $p_\varepsilon(\cdot \mid z)$ be the corresponding *transition probabilities*, given by

$$p_\varepsilon(E \mid z) = \nu_\varepsilon\big(\{g : g(z) \in E\}\big).$$

A probability measure μ_ε is *stationary* if, for every Borel set $E \subset U$,

$$\mu_\varepsilon(E) = \int p_\varepsilon(E \mid z)\, d\mu_\varepsilon(z).$$

We assume that there exist sets $\Lambda_{\varepsilon,z}$ containing the support of each $p_\varepsilon(\cdot \mid z)$, and there exist constants $K > 0$, $\kappa > 0$, independent of ε and z, satisfying

- (H1) every $\Lambda_{\varepsilon,z}$ admits a lamination into nearly horizontal (slope less than 10) curves such that the union of the laminae with length less than εt has Lebesgue measure bounded by $K\varepsilon^2 t^{1+\kappa}$, for every $t > 0$;
- (H2) the conditional probability of $p_\varepsilon(\cdot \mid z)$ along each lamina y is absolutely continuous with respect to arc-length m_y, with density $\psi_y = \psi_{\varepsilon,z,y}$ bounded by $K/\operatorname{length}(y)$;
- (H3) restricted to some ball of radius $\rho(\varepsilon)$ around $f(z)$, the probability $p_\varepsilon(\cdot \mid z)$ is absolutely continuous with respect to Lebesgue measure, with positive density.

Theorem 4.7 (Benedicks, Viana [42]). *Let f_a be a Hénon-like map, μ be its SRB measure, and $\{\nu_\varepsilon : \varepsilon > 0\}$ fulfill conditions (H1), (H2), (H3). Then there is a unique stationary measure μ_ε supported in the basin $B(\Lambda)$, and this measure is ergodic. Moreover, as $\varepsilon \to 0$ the measure μ_ε converges to μ in the weak* sense: $\int \varphi\, d\mu_\varepsilon \to \int \varphi\, d\mu$ for every continuous $\varphi : U \to \mathbb{R}$.*

Condition (H3) is needed only for proving uniqueness. If one drops this assumption, keeping (H1)-(H2), then one still gets stochastic stability, in the sense that the simplex of all stationary measures converges to μ when $\varepsilon \to 0$.

An easier version of these methods may be applied to certain one-dimensional maps. In this way, one gets an improved version of the results in [44]. The main improvement concerns the type of random noise allowed: since we require no lower bound on the density, we are able to treat random noise supported in very general domains, not only intervals.

Example 4.8. (Additive noise) Let Λ_ε be a neighborhood of the origin in $\mathbb{R}^2$ contained in the ball of radius ε, and let θ_ε be a probability supported on Λ_ε. Let ν_ε be the measure induced by θ_ε in the space of C^2 diffeomorphisms via the map $t \mapsto f_t = f + t$. Here $\Lambda_{\varepsilon,z} = f(z) + \Lambda_\varepsilon$. Property (H1) translates immediately to a similar condition about Λ_ε. Conditions (H2)-(H3) hold, for instance, if θ_ε is absolutely continuous with respect to Lebesgue measure (area), with density ψ_ε such that $\varepsilon^2\psi_\varepsilon$ is bounded from zero and infinity.

Example 4.9. (Noise in parameter space) Let $\phi : B \times \mathbb{R}^2 \to \mathbb{R}^2$ be a C^2 map, where $B \subset \mathbb{R}^2$ is the unit ball around the origin, such that $\phi(0,\cdot) = 0$ and $\omega \mapsto \phi(\omega, z)$ is a diffeomorphism for every z near the attractor Λ. Define $f_\omega(z) = f(z) + \phi(\omega, z)$. Let θ_ε be the normalized Lebesgue measure on the ε-ball around $\omega = 0$, and ν_ε be the probability induced by θ_ε in the space of diffeomorphisms via the map $\omega \mapsto f_\omega$. Then $\{\nu_\varepsilon : \varepsilon > 0\}$ satisfies the assumptions of Theorem 4.7.

4.5. **Chaotic dynamics near tangencies.** A generic unfolding of a homoclinic tangency of a dissipative (determinant of the Jacobian smaller than one) surface diffeomorphism yields a rather striking number of rich, intricate dynamical phenomena:

Theorem 4.10 (Mora, Viana [223, 333]). *For an open dense set of C^3 one-parameter families of diffeomorphisms $(f_t)_t \in \mathbb{R}$ going through a homoclinic bifurcation at $t = 0$, there exists a positive Lebesgue measure set of parameters t close to zero for which f_t exhibits a Hénon-like attractor or repeller.*

Theorem 4.11 (Palis, Yoccoz [268, 269]). *There exists a neighborhood D of the set $\{(a_u, a_s) : a_u + a_s < 1\}$ inside $(0,1)^2$ such that if $(d_u(\Lambda), d_s(\Lambda)) \in D$ then $\mathcal{W} = \{\mu : \Lambda_\mu$ is a non-uniformly hyperbolic horseshoe$\}$ has full density at $\mu = 0$.*

5. Non-critical behavior of surface maps

As we have seen in the previous sections, much of the richness of the dynamics of surface diffeomorphisms is due to phenomena associated to the unfolding of homoclinic tangencies, as in the Hénon family. Moreover, these phenomena have many analogies with critical endomorphisms of the interval or the circle. For instance, homoclinic tangencies correspond in the one-dimensional setting to pre-periodic critical points.

For non-critical one-dimensional maps which are sufficiently regular, one has good control of the distortion and, thus, a lot of rigidity: Theorem 2.5 asserts that non-critical maps are almost hyperbolic. Here we discuss some remarkable recent results of Pujals, Sambarino [286, 287, 285] extending this conclusion to *non-critical* surface diffeomorphisms, that is, such that one cannot create homoclinic tangencies by small C^1 perturbations of the diffeomorphism.

5.1. **Non-critical surface dynamics.** Let M be a closed Riemannian surface, $f\colon M \to M$ be a diffeomorphism, and Λ be an f-invariant set. A Df-invariant splitting $T_xM = E(x) \oplus F(x)$, $x \in \Lambda$ of the tangent bundle over Λ is called *dominated* if there is $m \in \mathbb{N}$ such that

$$|Df^n \mid E(x)| \cdot |(Df^n \mid F(x))^{-1}| \leq 1/2 \quad \text{for all } x \in \Lambda \text{ and } n \geq m.$$

See Appendix B for much more information about this notion.

In general terms, the results in this section show that surface diffeomorphisms far from homoclinic tangencies have dominated dynamics: the limit set admits a dominated splitting. Moreover, dominated diffeomorphisms on surfaces behave very much like non-critical one-dimensional maps.

Theorem 5.1 (Pujals, Sambarino [286])**.** *Let f be a C^2 surface diffeomorphism and $\Lambda \subset \Omega(f)$ be a compact f-invariant set admitting a dominated splitting $T_\Lambda M = E \oplus F$. Assume all periodic points in Λ are hyperbolic saddles. Then $\Lambda = \Lambda_1 \cup \Lambda_2$, where Λ_1 is a hyperbolic set and Λ_2 is the union of finitely many pairwise disjoint normally hyperbolic circles $C_1, \ldots, C_k$ such that $f^{m_i}(C_i) = C_i$ and $f^{m_i} \colon C_i \to C_i$ is an irrational rotation, for some $m_i \geq 1$.*

Here is a useful global reformulation of this theorem:

Theorem 5.2 (Pujals, Sambarino [287])**.** *Let f be a C^2 surface diffeomorphism whose periodic points are all hyperbolic and which can not be C^1 approximated by diffeomorphisms exhibiting homoclinic tangencies.*

Then the non-wandering set of f may be decomposed into compact f-invariant sets, $\Omega(f) = \Omega_1(f) \cup \Omega_2(f)$, such that $\Omega_1(f)$ is hyperbolic and $\Omega_2(f)$ is the union of finitely many pairwise disjoint normally hyperbolic circles $C_1, \ldots, C_k$ such that $f^{m_i}(C_i) = C_i$ and $f^{m_i} \colon C_i \to C_i$ is an irrational rotation, for some $m_i \geq 1$. In particular, Lebesgue almost every point of M is in the basin of attraction of either a hyperbolic attractor or a normally attracting periodic circle.

More recently, Pujals, Sambarino [285] extended these results by removing the assumption that all periodic points are hyperbolic: *If the limit set $L(f)$ admits a dominated splitting then the periods of the non-hyperbolic periodic orbits of f are bounded above. Moreover, $L(f) = L_1 \cup L_2 \cup L_3$, where L_1 is the union of finitely many pairwise disjoint homoclinic classes, each class containing at most finitely many non-hyperbolic periodic points; L_2 is the union of finitely many normally hyperbolic circles on which a power of f is a rotation; L_3 consists of periodic points contained in a finite union of normally hyperbolic periodic intervals.* Compare Mañé's Theorem 2.5.

A very interesting by-product of the arguments leading to the proof of this last result is the construction in [285] of surface diffeomorphisms having coexisting hyperbolic periodic orbits with normalized eigenvalues arbitrarily close to one, and such that there is no C^2 perturbations breaking the hyperbolicity of any of these points. In particular, Frank's lemma A.8 does not extend to the C^2 topology. See also Section A.5.

5.2. **Domination implies almost hyperbolicity.** To prove Theorem 5.1 it is enough to see that, under the hypotheses, either Λ is hyperbolic or it contains a periodic circle where the dynamics of a power of f is conjugate to an irrational rotation. In fact, this implies the theorem, after noting that the number of periodic circles of f must be finite. The main step of the proof is

Proposition 5.3. *Consider a C^2 surface diffeomorphism f and a nontrivial transitive f-invariant compact set Σ admitting a dominated splitting $T_\Sigma M = E \oplus F$. Suppose that every proper compact invariant subset of Σ is hyperbolic, and Σ does not contain normally hyperbolic periodic circles C such that $f^m(C) = C$ and the restriction of f^m to C is conjugate to an irrational rotation, for some $m \geq 1$. Then Σ itself is hyperbolic.*

Assuming this proposition we now sketch the proof of Theorem 5.1. Suppose Λ is not hyperbolic. Consider the family $\mathcal{K}$ of compact f-invariant subsets of Λ which are not hyperbolic. The intersection of any decreasing sequence of non-hyperbolic invariant compact sets is also a non-hyperbolic invariant compact set. So, by Zorn's lemma, the family $\mathcal{K}$ has some minimal element Σ. By hypothesis, every periodic point of f in $\Lambda \supset \Sigma$ is hyperbolic. Hence, Σ cannot be trivial (a union of periodic orbits). We claim that Σ is a transitive set. Thus, since it is also minimal, Σ verifies all the hypotheses of Proposition 5.3. It follows that it is hyperbolic, which is a contradiction.

To see that Σ is a transitive set one uses the following useful alternative characterization of hyperbolicity, which is an analogue of Lemma 2.6 in higher dimensions:

Lemma 5.4. *Let Σ be an f-invariant compact set having a dominated splitting $T_\Sigma M = E \oplus F$ such that $|Df^n(x) \mid E(x)| \to 0$ and $|Df^{-n}(x) \mid F(x)| \to 0$ as $n \to \infty$ for all $x \in \Sigma$. Then Σ is a hyperbolic set: $E = E^s$ is uniformly contracting and $F = E^u$ is uniformly expanding.*

We are going to show that there exists $x \in \Sigma$ such that either $\Sigma = \omega(x)$ or $\Sigma = \alpha(x)$, which implies transitivity. Indeed, suppose the ω-limit set of a point $x \in \Sigma$ is properly contained in Σ. Then, by minimality, $\omega(x)$ is a hyperbolic set. It follows that $Df^n(x)(v) \to 0$ for every vector v of $E(x)$. In the same way, arguing with α-limit sets, $Df^{-n}(x)(v) \to 0$ for every vector v of $F(x)$. If this were valid for every $x \in \Sigma$, Lemma 5.4 would imply that Σ is hyperbolic, which is a contradiction. So, there is some $x \in \Sigma$ such that either $\omega(x)$ or $\alpha(x)$ is the whole Σ.

Let us now present some main ingredients in the proof of Proposition 5.3. A first one is a Denjoy type property about intervals transverse to the E-direction. Existence of a dominated splitting $T_\Sigma M = E \oplus F$ is equivalent to existence of forward invariant and backward invariant cone-fields containing F and E, respectively, and defined on some neighborhood V of the set Σ. Then we say that a curve I contained in V is *transverse to the E-direction* if the tangent direction at every point $x \in I$ is contained in the corresponding forward invariant cone.

Fix an open neighborhood U of Σ whose closure is contained in V. Let Σ^+ be the maximal forward invariant set of f in the closure of U. By construction, Σ^+ has a dominated splitting and it contains Σ. We call a curve I a *(δ, E)-interval* if it is contained in Σ^+, all its forward iterates are transverse to the E-direction, and length $(f^n(I)) < \delta$ for every $n \geq 0$.

Proposition 5.5. *There is $\delta > 0$ such that the ω-limit set $\omega(I)$ of every (δ, E)-interval I satisfies one of the following two possibilities:*

- *either $\omega(I)$ is a normally hyperbolic periodic circle C where the restriction of f^m (m is the period of C) is conjugate to an irrational rotation,*
- *or $\omega(I)$ is contained in the set of periodic points of f in the neighborhood V of Λ.*

By definition, $\omega(I)$ is the union of the sets $\omega(x)$ for $x \in I$. The Denjoy type statement implies that if the set Λ in Theorem 5.1 does not contain normally hyperbolic periodic circles then $\omega(I)$ is contained in the set of periodic points, for every (δ, E)-interval I.

Another key ingredient is to control center stable and center unstable manifolds. Existence of invariant families of curves $W^{cs}_\varepsilon(x)$ and $W^{cs}_\varepsilon(x)$ tangent to the subbundles E and F follows from general normal hyperbolicity theory [158]. In the present setting these curves are of class C^2. Most important, they are dynamically defined: for every $x \in \Sigma$ there exists $\varepsilon(x) > 0$ there is such that the lengths of $f^{-n}(W^{cu}_{\varepsilon(x)}(x)$ and $f^n(W^{cs}_{\varepsilon(x)}(x))$ go to zero as $n \to \infty$. In a second stage one proves that

$$\sum_{n=0}^{\infty} \text{length}\,(f^{-n}(W^{cu}_{\varepsilon(x)}(x))) < \infty \quad \text{and} \quad \sum_{n=0}^{\infty} \text{length}\,(f^{n}(W^{cs}_{\varepsilon(x)}(x))) < \infty$$

for every x in an open subset of Σ. One deduces that $|Df^{-n}(x) \mid F(x)| \to 0$ and $|Df^n(x) \mid E(x)| \to 0$ as $n \to \infty$, for every $x \in \Sigma$. In view of Lemma 5.4 this gives the hyperbolicity of Σ.

The proof of these facts uses ideas from the proof of Theorem 2.5, together the Denjoy-like Proposition 5.5 and the following property of the unstable, or stable, separatrices of hyperbolic points p of Σ: there is κ such that if a separatrix L of $W^u(p) \setminus \{p\}$ has length less than κ then the end-point of L other than p is either a sink or a non-hyperbolic saddle.

5.3. **Homoclinic tangencies versus Axiom A.** An important consequence of Theorem 5.1 is that every diffeomorphism f that cannot be approximated by diffeomorphisms exhibiting homoclinic tangencies can be approximated by an Axiom A diffeomorphism:

Theorem 5.6 (Pujals-Sambarino [286]). *Let M be a closed surface. Then there is a dense subset $\mathcal{D}$ of* $\mathrm{Diff}^1(M)$ *of diffeomorphisms f such that either f has a homoclinic tangency or f satisfies the Axiom A.*

The proof has two main parts. The first one is the construction of a dominated splitting defined on the non-wandering set of every diffeomorphism g in an open and dense subset of a neighborhood of f. The second step is to prove that such a splitting is hyperbolic. The latter is done by approximating the diffeomorphism by a C^2 one, in the C^1 topology. This allows us to use Theorem 5.2 to decompose the dynamics into two parts, one corresponding to irrational rotations and another which is hyperbolic. Let us now comment both parts in more detail.

Consider the set $\mathcal{U}$ defined as the complement in $\mathrm{Diff}^1(M)$ of the closure of the diffeomorphisms exhibiting homoclinic tangencies. Theorem 5.6 just claims the density of the Axiom A diffeomorphisms in $\mathcal{U}$. Let $\mathcal{H}$ be the dense subset of Kupka-Smale diffeomorphisms $f \in \mathcal{U}$. In particular, every periodic point of $f \in \mathcal{H}$ is hyperbolic. For $f \in \mathcal{H}$, we let $\Sigma(f)$ be the set of (possibly infinitely many) sinks and sources, and $\Omega_0(f)$ be the complement of $\Sigma(f)$ inside the non-wandering set $\Omega(f)$. Clearly, $\Omega_0(f)$ is f-invariant and compact. The next step is to see that $\Omega_0(f)$ has a dominated splitting.

To prove this, one first observes that every point $x \in \Omega_0(f)$ is approximated by sequences $(p_n)_n$ such that each p_n is a hyperbolic periodic point of a diffeomorphism g_n arbitrarily C^1 close to f in the C^1 topology. This follows from the closing lemma, Theorem A.1. The key point here is that the natural splitting over the set of periodic points, given by the hyperbolic directions, is dominated uniformly in g_n, and so extends to the closure. The proof of this key fact is based on Franks' Lemma A.8 and arguments from the proof of the stability conjecture in [211]. In a few words, there are two cases. First, if the angles between the stable and unstable are of the periodic points p_n may be taken arbitrarily small, then we can obtain a homoclinic tangency after a small perturbation. So, this case is forbidden by the fact that $f \in \mathcal{U}$. Second, if the splitting is not dominated, a small perturbation allows us to recline the unstable direction against to the stable direction. Together with lack of domination, this leads to small angles between the two subspaces, and again we obtain a contradiction. See Section 7.1.1 for more details.

Now we move to the second and last part of the proof of Theorem 5.6. We only have to show that the diffeomorphisms in a dense subset of $\mathcal{H}$ verify the Axiom A. From the first part we know that, for every $f \in \mathcal{H}$, the set $\Omega_0(f)$ has a dominated splitting and so it verifies the hypotheses of Theorem 5.1. Since existence of normally hyperbolic curves supporting irrational rotations is a non-generic, Theorem 5.1 implies that $\Omega_0(f)$ is a hyperbolic set, for every f in a dense subset $\mathcal{D}$ of $\mathcal{H}$. Thus, the non-wandering set $\Omega(f)$ of $f \in \mathcal{D}$ is hyperbolic if, and only if, the set $\Sigma(f)$ of sinks and sources is hyperbolic.

The hyperbolicity of $\Sigma(g)$ is evident if this set is finite, and so we only have to consider the infinite case. Since $\Omega_0(f)$ is hyperbolic, it admit a neighborhood V such every f-invariant set in V is hyperbolic. Observe that every accumulation point of sinks or sources must be in $\Omega_0(f)$. Therefore, all but finitely many elements of $\Sigma(f)$ are contained in V, which is a contradiction. It follows that $\Sigma(f)$ is hyperbolic, and this concludes the proof of the hyperbolicity of $\Omega(g)$.

5.4. **Entropy and homoclinic points on surfaces.** Existence of transverse homoclinic points implies the existence of horseshoes (nontrivial hyperbolic sets), and so diffeomorphisms with transverse homoclinic points have strictly positive topological entropy. Conversely, the following deep result asserts that surface diffeomorphisms with positive topological entropy exhibit horseshoes:

Theorem 5.7 (Katok [170]). *Let f be a $C^{1+\alpha}$ diffeomorphism, $\alpha > 0$, on a closed surface. If f has positive topological entropy then it has hyperbolic periodic saddles with transverse homoclinic points.*

In view of these results, it is a natural important problem to try and characterize the diffeomorphisms having zero topological entropy, as well as to understand the dynamical consequences of the variation of the topological entropy.

It is clear that Morse-Smale diffeomorphisms have zero entropy. The following consequence of Theorem 5.6 states that these are the only systems having zero topological entropy *robustly*. See also Gambaudo, Rocha [166].

Theorem 5.8 (Pujals, Sambarino [286]). *Let M be a closed surface and* MS *denote the C^1-closure of the set of Morse-Smale diffeomorphisms of* $\operatorname{Diff}^1(M)$. *Then there is an open and dense subset of* $\operatorname{Diff}^1(M) \setminus \mathrm{MS}$ *consisting of diffeomorphisms with transverse homoclinic orbits. Thus,* MS *coincides with the closure of the interior of the set of diffeomorphisms with zero topological entropy.*

Since the set of diffeomorphisms with transverse homoclinic orbits is open, we only have to show that it is dense in $\operatorname{Diff}^1(M) \setminus \mathrm{MS}$. This is an easy consequence of Theorem 5.6. Indeed, the theorem gives that close to any $f \notin \mathrm{MS}$ there is a diffeomorphism g either having a homoclinic tangency or verifying the Axiom A. In the first case, we may perturb g to create a transverse homoclinic orbit, thence proving the claim. In the second one, we may assume that the non-wandering set of g is infinite: otherwise g would be Morse-Smale, contradicting that $f \in \mathrm{MS}$. Then, since g is Axiom A, the non-wandering set of g contains nontrivial hyperbolic sets and, thus, transverse homoclinic orbits.

As an immediate consequence of Theorem 5.2 one gets the following result, that generalizes a theorem for diffeomorphisms of the 2-sphere in [166]:

Theorem 5.9. *Let f be a C^2 surface diffeomorphism with infinitely many periodic points, all of them hyperbolic. Then f is C^1 approximated by a diffeomorphism with a homoclinic point.*

This follows observing that if f can not be approximated by diffeomorphisms with homoclinic tangencies then it verifies Theorem 5.2. Since the non-wandering set is infinite, it must contain some nontrivial hyperbolic set and, hence, transverse homoclinic points.

We mention yet another important consequence of Theorem 5.2, concerning variation of the topological entropy. A C^∞ diffeomorphism f is a *point of entropy variation* if every C^∞ neighborhood $\mathcal{U}$ of f contains diffeomorphisms g such that the topological entropies of f and g are different.

Theorem 5.10 (Pujals, Sambarino [287])**.** *Let f be a point of entropy variation. Then f is C^1 accumulated by diffeomorphisms exhibiting homoclinic tangencies.*

Suppose f is far from homoclinic tangencies. Let g be any diffeomorphism C^∞ close to f. Consider an arc $(f_t)_{t\in[0,1]}$ of C^∞ diffeomorphisms with $f_0 = f$ and $f_1 = g$, contained in a small neighborhood of f. By [319] we may take this arc such that the set S of parameter values t such that f_t is not Kupka-Smale is countable and $[0,1] \setminus S$ is a countable union of intervals. The assumption that f is far from homoclinic tangencies ensures that f_t satisfies the hypotheses of Theorem 5.2 for every $t \notin S$. Since invariant circles do not contribute for the entropy, one deduces that the entropy is constant on each interval in the complement of S. On the other hand, $h_{top}(f_t)$ is a continuous function of the parameter, because the topological entropy varies continuously with the dynamics in the C^∞ topology [348, 243]. As S is only countable, these two facts imply that the entropy is constant on the whole parameter interval $[0,1]$. Thus $h_{top}(g) = h_{top}(f)$, and so f is not a point of entropy variation.

5.5. **Non-critical behavior in higher dimensions.** Not much is known about the dynamics of invariant sets far from tangencies in higher dimensions. Recently, Wen [341] has shown that such non-critical behavior implies the existence of a dominated splitting, thus extending the first step of the proof of Theorem 5.6 to any dimension. Let us call a point x *i-pseudo-periodic* for f if there exists y arbitrarily close to x and g arbitrarily close to f in the C^1 sense, such that y is a hyperbolic periodic point for g with stable manifold of dimension i. Then he proves that f *is far from homoclinic tangencies if and only if, for every* $0 < i < d$ *the set of i-pseudo-periodic points admits a dominated splitting* $E \oplus F$ *with* $\dim E = i$.

This, and much of the progress in this direction is motivated by the following

Conjecture 5.11 (Palis [260])**.** In any dimension, the diffeomorphisms exhibiting either a homoclinic tangency or heterodimensional cycles are C^r dense in the complement of the closure of the hyperbolic ones, for any $r \geq 1$.

Of course, surface diffeomorphisms can not have heterodimensional cycles. So, in the 2-dimensional situation the case $r = 1$ of the conjecture reduces to Theorem 5.6, which had also been conjectured by Palis. A solution of Conjecture 5.11 has been announced very recently by Hayashi [148].

The following weaker version of Conjecture 5.11 has also been proposed:

Conjecture 5.12 (Palis [260])**.** The subset of dynamical systems that either have their limit set consisting of finitely many hyperbolic periodic orbits or else they have transversal homoclinic orbits, are C^r dense in the set of all dynamical systems, $r \geq 1$.

In this regard, Mora, Martin-Ribas [222] prove that every C^2 surface diffeomorphism that admits an *almost homoclinic sequence*, that is, a sequence of points that have iterates arbitrarily close to the stable manifold and to the unstable manifold of some saddle, either has transverse homoclinic points or can be C^1 approximated by diffeomorphisms with homoclinic tangencies. The proof goes as follows.

Under the hypotheses, Hayashi's connecting lemma, Theorem A.4, gives C^1 perturbations of the original map that have transverse homoclinic orbits. The closure of these homoclinic orbits is an invariant set with a dominated splitting. If the domination becomes weaker when the size of the perturbation goes to zero, methods of Mañé [206] permit to find a homoclinic tangency for a nearby system. Otherwise, the unperturbed map also has an invariant set with dominated splitting. From Theorem 5.1 it follows that the original diffeomorphism has a nontrivial hyperbolic set. This is a contradiction because we assumed that there are no transverse homoclinic points.

Recently, Arroyo, Rodriguez-Hertz [24] extended some of the previous results to vector fields on 3-dimensional manifolds. Their main result says that *every C^1 vector field may be approximated by another which either is uniformly hyperbolic with no-cycles, has a homoclinic tangency, or has a singular cycle.* By cycle we mean a finite family of equilibrium points and periodic orbits cyclically related by intersections of their stable and unstable manifolds. The cycle is called singular if it does involve equilibria (not just regular periodic orbits). This is an important step towards case $r = 1$ of the following reformulation of Conjecture 5.11 for flows:

Conjecture 5.13 (Palis [260])**.** Every vector field may be C^r approximated by another which either is uniformly hyperbolic with no-cycles, has a homoclinic tangency, or has a singular (Lorenz-like) attractor.

The basis for proving the main result is a continuous-time version of Theorem 5.1: Arroyo, Rodriguez-Hertz prove that *if invariant set of a C^2 flow on a three-dimensional manifold contains no equilibria and the only periodic orbits in it are hyperbolic saddles, then it may be decomposed as the union of a hyperbolic set and finitely many tori supporting an irrational flow.* Let us point out that they found it convenient to work with a different notion of dominated splitting, stated in terms of the linear Poincaré flow. See [24, Proposition 3.3].

We close with a few more conjectures inspired by the previous results. Partially hyperbolic systems with invariant splitting $TM = E^u \oplus E^c \oplus E^s$ with one-dimensional central direction E^c are, in a sense, closest to being hyperbolic among all non-hyperbolic diffeomorphisms. Hence, one may expect a description of the dynamics to be easier and more complete in this case. In this spirit, we formulate

Conjecture 5.14. Generically, partially hyperbolic diffeomorphisms with central direction of dimension 1 admit a filtration separating the transitive pieces of the limit (or the non-wandering) set.

A more ambitious formulation would require the same conclusion for partially hyperbolic diffeomorphisms with splitting $TM = E^u \oplus E^1 \oplus \cdots \oplus E^l \oplus E^s$, where every E^i, $i = 1, \ldots, l$ is one-dimensional and, as before, E^u is uniformly expanding and E^s is uniformly contracting. On the other hand, Conjecture 5.14 might be easier when the invariant splitting contains only two sub-bundles:

Conjecture 5.15. Let f be a C^2 diffeomorphism such that the limit set $L(f)$ has a dominated splitting $E \oplus F$ where F has dimension 1 and E is uniformly contracting. Then $L(f)$ is the union of finitely many hyperbolic circles on which f is an irrational rotation, of (possibly infinitely many) periodic points contained in a finite union of periodic intervals, and finitely many pairwise disjoint homoclinic classes, each of them containing at most finitely many non-hyperbolic periodic points.

For the general definition of dominated splitting, see Definition B.1. Crovisier [100] has obtained some progress in the direction of this conjecture, in the case when there is a unique non-hyperbolic periodic point.

6. Heterodimensional cycles and blenders

In this section we study the unfolding of heterodimensional cycles, that is, the dynamical phenomena arising from perturbing diffeomorphisms with such cycles. Cycles involving periodic points with different unstable dimensions were first considered by Newhouse, Palis in [244], and were systematically studied in the series of papers [106, 107, 109, 110, 111, 112, 115].

The notion of blender was introduced in [56], motivated by [107]. Blenders are geometric models for certain hyperbolic sets originating in the unfolding of heterodimensional cycles, and play an important role as a mechanism for creation of cycles, and a semi-local source of transitivity.

6.1. **Heterodimensional cycles.** Consider a closed manifold M and a diffeomorphism f defined on M and having a pair of hyperbolic saddles P and Q with different indices, that is, different dimensions of their unstable subspaces. Assume $W^s(P)$ and $W^u(Q)$ have non-empty intersection, and the same for $W^u(P)$ and $W^s(Q)$. We say that f has a *heterodimensional cycle* associated to P and Q. Of course, heterodimensional cycles can only exist in dimension bigger or equal to 3.

6.1.1. *Explosion of homoclinic classes.* The main result in this section depicts heterodimensional cycles as a powerful mechanism for explosion of homoclinic classes of periodic points and creation of robust non-hyperbolic transitive sets. The assumptions are not minimal, we shall see later that this kind of behavior is typical of a large class of cycles.

For simplicity, we suppose that the periodic points P and Q in the cycle are actually fixed points. In addition, we always assume

- (codimension 1) the saddles P and Q have indices p and $q = p + 1$, respectively,
- (quasi-transversality) the manifolds $W^s(P)$ and $W^u(Q)$ intersect transversely, and the intersection between $W^u(P)$ and $W^s(Q)$ is quasi-transverse: $T_x W^s(Q) \cap T_x W^u(P) = \{0\}$ for every intersection point x.

The next theorem asserts that the homoclinic classes of P and Q often explode, and become intermingled (non-empty intersection) when the cycle is unfolded.

A parametrized family $(f_t)_{t\in[-1,1]}$ of diffeomorphisms *unfolds generically* a heterodimensional cycle of $f = f_0$ if there are open disks $K^u_t \subset W^u(P_t)$ and $K^s_t \subset W^s(Q_t)$, depending continuously on t, such that $K^u_0 \cap K^s_0$ contains a point of quasi-transverse intersection, and the distance between K^s_t and K^u_t increases with positive velocity when t increases. Here P_t and Q_t denote the continuations for f_t of the periodic points P and Q.

Theorem 6.1 (Díaz [107]). *There is a non-empty open set of C^∞ parametrized families of diffeomorphisms $(f_t)_{t\in[-1,1]}$ unfolding generically a heterodimensional cycle of $f = f_0$, such that for all small positive t*

1. *the transverse intersection between $W^s(P_t)$ and $W^u(Q_t)$ is contained in the homoclinic class of Q_t;*
2. *and the homoclinic class of P_t is contained in the homoclinic class of Q_t.*

Let d be the dimension of the ambient manifold M. The key fact behind Theorem 6.1 is that, for every small positive t, the $(d-p)$-dimensional manifold $W^s(P_t)$ is contained in the closure of $W^s(Q_t)$. This makes that the stable manifold of Q_t, which has dimension $(d-p-1)$, behaves in fact as a manifold of dimension 1 unit bigger. This type of phenomenon has been synthesized in the notion of blender, which we discuss in Section 6.2.

The proof of the theorem, which we are going to sketch next, relies on a reduction of the dynamics to a family of iterated functions systems on the interval. Besides the standing hypotheses of codimension 1 and quasi-transversality above, the main conditions defining the open set in the theorem are *connectedness* of the transverse intersection γ, *non-criticality* or partial hyperbolicity of the dynamics on γ, and *small distortion* of the transition from the neighborhood of Q to that of P. This will be explained in the sequel, but let us point out right away that these conditions may hold simultaneously for a parametrized family f_t and for its inverse f_t^{-1}: one gets open sets of families for which the homoclinic classes of P_t and Q_t coincide after the bifurcation. See Remark 6.3 below.

6.1.2. *A simplified example.* To explain the ideas of the proof of Theorem 6.1 while avoiding technical difficulties, we consider here a simple situation of a cycle in $\mathbb{R}^3$, which displays the main features of the general case we are dealing with.

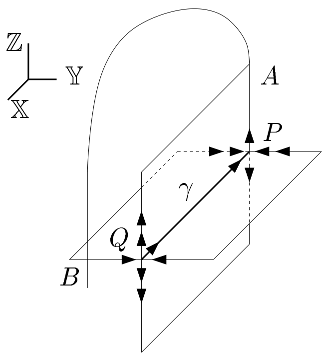

FIGURE 10. Model heterodimensional cycle

We assume that the restriction of f_t to a domain $C = [-1,4] \times [-2,2] \times [-2,2]$ satisfies (see Figure 10):

Partial hyperbolicity and product structure along the transverse intersection:

(a) On some domain $[-1,4] \times [-2,2] \times [-a,a]$ whose image contains some $[-1,4] \times [-b,b] \times [-2,2]$ the diffeomorphism f_t is independent of t and coincides with the product of its restriction to the $\mathbb{X}$-axis by a hyperbolic affine map, such that $\mathbb{Y}$ and $\mathbb{Z}$ are the strong stable and strong unstable directions: $\mathbb{Z} \oplus \mathbb{X} \oplus \mathbb{Y}$ is a dominated splitting for f_t.

(b) The points $Q_t = (0,0,0)$ and $P_t = (3,0,0)$ are hyperbolic saddles of indices 2 and 1, respectively. The *connection* $\gamma = (0,3) \times \{(0,0)\}$ is included in $W^u(Q_t) \cap W^s(P_t)$.

(c) The restrictions of f_t to the neighborhoods $U_Q = [-1,1] \times [-2,2] \times [-a,a]$ and $U_P = [2,4] \times [-2,2] \times [-a,a]$ are affine maps, and the eigenvalues of f_t at Q_t and P_t in the $\mathbb{X}$-direction are $\beta > 1$ and $0 < \lambda < 1$, respectively.

Product structure along the quasi-transverse orbit and generic unfolding of the cycle:

(d) Consider the points $A = (3,0,1) \in W^u(P_0)$ and $B = (0,-1,0) \in W^s(Q_0)$. There is $\ell > 0$ such that $f_0^\ell(A) = B$, and so A and B are heteroclinic points for $f = f_0$. Moreover, there is a small neighborhood U_A of A where f_0^ℓ coincides with the translation $(x,y,z) \mapsto (x-3, y-1, z-1)$.

(e) The restriction of f_t^ℓ to U_A coincides with the translation

$$(x,y,z) \mapsto (x-3+t, y-1, z-1) = f_0^\ell(x,y,z) + (t,0,0).$$

By the first two conditions, $(0,3) \times [-2,2] \times \{0\}$ is contained in the stable manifold of P_t, and $(0,3) \times \{0\} \times [-2,2]$ is contained in the unstable manifold of Q_t. Then A and $B_t = (t,-1,0)$ are transverse homoclinic points of P_t, and $A_t = (3-t,0,1)$ and B are transverse homoclinic points of Q_t , for every small positive t.

Key properties of the diffeomorphism f_t may be derived from properties of the dynamics it induces in the quotient space by the sum $\mathbb{Y} \oplus \mathbb{Z}$ of the strong stable and the strong unstable directions. This quotient dynamics consists of an iterated functions system with two generators: the restriction of f to the $\mathbb{X}$-axis, also denoted by f, and the quotient

$$g_t \colon [3-\delta, 3+\delta] \to \mathbb{R}, \quad g_t(x) = x - 3 + t,$$

of the restriction of f_t^ℓ to U_A, defined for small $t > 0$. Since we are interested in the dynamics in a neighborhood of the cycle, it is natural to restrict f to $[0,3]$ and g_t to $[3-t, 3]$. See Figure 11.

The dynamics of the iterated functions system is, essentially, determined by the smooth conjugacy class and by the Mather invariant [216] of the interval map f. The former is described by the pair of eigenvalues (β, λ). Let us recall the definition of the latter.

By assumption, in a neighborhood of $Q = 0$ the transformation f is the time-1 map of the flow defined by the affine vector field $X_Q(x) = (\log \beta)\, x \frac{\partial}{\partial x}$. Similarly, in a neighborhood of $P = 3$ the transformation f is the time-1 of the vector field $X_P(x) = (\log \lambda)\,(x-3) \frac{\partial}{\partial x}$. For $x \in (0,3)$ close to Q, consider n big enough so that $f^n(x)$ is close to P, and then write

$$Df^n(x)(X_Q(x)) = \mu(x)\, X_P(f^n(x)),$$

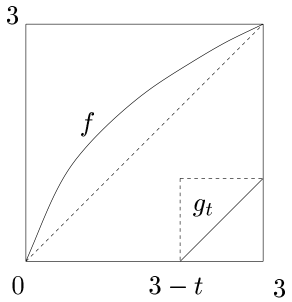

FIGURE 11. Iterated functions system with 2 generators

for some $\mu(x) \in \mathbb{R}$. Since the vector fields X_Q and X_P are f-invariant, the function $\mu(\cdot)$ does not depend on n, and it satisfies $\mu(x) = \mu(f(x))$ for all x. Hence, the *Mather invariant* $\mu(\cdot)$ is a smooth function on the circle $(0,3)/f$. It corresponds to what [107] calls the transition from Q to P.

As part of defining the open set in the statement of Theorem 6.1, we also assume that the transition has small distortion, in other words, that μ is close to 1. The precise condition appears in Lemma 6.2 below, involving also the eigenvalues β and λ. Observe that μ is constant equal to 1 if the restriction of f to $[0,3]$ is the time-1 of a smooth vector field. We now sketch the proof of Theorem 6.1 for our simplified model.

Denote by $D_P^t = [3-t, 3-\lambda t]$ the fundamental domain of f at distance t from P. Let $s_t = f^{-n_t}(3-t)$ be the unique backward iterate of $3-t$ under f in the interval $(\beta^{-1}t, t]$. Then let $D_Q^t = [s_t, \beta s_t]$ be the corresponding fundamental domain of f close to Q. We are going to construct a return map to this last domain, for the iterated function system. See Figure 12.

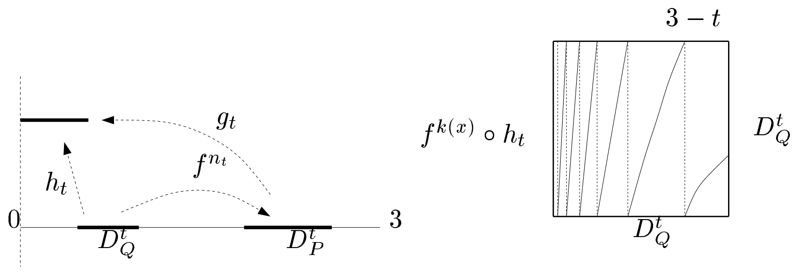

FIGURE 12. A return map to D_Q^t

By construction, $f^{n_t}(D_Q^t) = D_P^t$. Consider the increasing surjective map

$$h_t = g_t \circ f^{n_t} : D_Q^t \to [0, (1-\lambda)t].$$

Given a point $x \in D_Q^t$ there exists a unique integer $k(x)$ such that $f^{k(x)} \circ h_t(x)$ is in D_Q^t. This defines a return map to the fundamental domain D_Q^t. Due to the jumps of $k(\cdot)$, the map $f^{k(x)} \circ h_t(x)$ has a sequence of discontinuities accumulating at the point $x = \beta^{-1}s_t$. We get rid of these discontinuities, though not the one at the accumulating point, by identifying the endpoints of the domain D_Q^t, so that it becomes a circle. To make this precise, consider the projection

$$\pi\colon (0,1] \to S^1 = \mathbb{R}/\mathbb{Z}, \qquad \pi(x) = \frac{\log x}{\log \beta} \pmod 1.$$

Let $c_t = \pi(s_t) = \pi(\beta\, s_t)$, and denote by π_t^{-1} the inverse of $\pi\colon (s_t, \beta\, s_t] \to S^1$. The *return map* is given by

$$R_t = \pi \circ h_t \circ \pi_t^{-1} \colon S^1 \to S^1.$$

This circle map R_t has a unique discontinuity, at c_t. Since $h_t(s_t) = 0$ and $k(x)$ goes to infinity when x goes to s_t inside D_Q^t, we have that the derivative $DR_t(z)$ goes to $+\infty$ when $z \to c_t$ from the right. Moreover, $R_t(z)$ "goes to $-\infty$", meaning that the graph winds infinitely many times around the circle.

Lemma 6.2. *Suppose the Mather function $\mu(\cdot)$ and the eigenvalues λ, β satisfy*

$$\frac{\mu(x)\,\lambda\,|\log\lambda|}{(1-\lambda)\,\log\beta} > 1, \quad \textit{for every } x.$$

Then, for every small $t > 0$, the map R_t is uniformly expanding: $|DR_t(z)| > 1$ for all $z \neq c_t$.

This lemma implies that for any open interval $I \subset [0,3]$ there is an element ϕ of the iterated functions system $\mathcal{G}(f, g_t)$ generated by f and g_t such that $\phi(I)$ contains $Q = 0$. Indeed, consider the projection $J = \pi(I)$. Since the map R_t is expanding, there is $n \geq 1$ such that $R_t^n(J)$ contains the discontinuity c_t. Then, if ψ is the element of $\mathcal{G}(f, g_t)$ corresponding to R_t^n, we have that $\psi(I)$ contains s_t. Then it suffices to take $\phi = h_t \circ \psi$ and to recall that $h_t(s_t) = 0$.

From the previous paragraph we get that *the stable manifold of Q intersects any disk Δ transverse to the stable manifold of P*. Indeed, suppose Δ is a vertical disk $I \times \{y\} \times [-2,2]$ for some interval $I \subset [0,t]$ and some $y \in [-2,2]$. Consider the image of Δ under the map ϕ (extended to the ambient space): the previous conclusion means that this image intersects $\{0\} \times [-2,2] \times \{0\}$. Since this last segment is contained in the stable manifold of Q, it follows that $\Delta \cap W^s(Q)$ is nonempty. For a general disk Δ transverse to $W^s(P)$, the assertion follows after observing that the forward orbit of Δ contains disks arbitrarily close to vertical disks as before.

As an immediate consequence we get that the connection curve $\gamma = (0,3) \times \{(0,0)\}$ is contained in the homoclinic class of Q. The other claim in Theorem 6.1 follows from a variation of the same argument. To finish our sketch of the proof of the theorem, we give the

Proof of Lemma 6.2. Consider $z \in S^1$ and let $x = \pi_t^{-1}(z)$. Observe that $DR_t(z)$ is defined by the relation

$$D(g_t \circ f^{n_t})(X_Q(x)) = DR_t(z) \cdot X_Q(g_t \circ f^{n_t}(x)).$$

Recalling that $Df^{n_t}(X_Q(x)) = \mu(x)\, X_P(f^{n_t}(x))$ and the expressions of X_Q and X_P, one has

$$\mu(x)\,|\log\lambda|\,(f^{n_t}(x) - 3) = DR_t(z)\,(\log\beta)\,g_t(f^{n_t}(x)).$$

Since, by definition, $f^{n_t}(x) \in D_P^t = [3-t, 3-\lambda t]$, one has $(f^{n_t}(x) - 3) \in [-t, -\lambda\, t]$. Thus $g_t(f^{n_t}(x)) = f^{n_t}(x) - 3 + t \in [0, t - \lambda\, t]$. Hence

$$DR_t(z) = \frac{\mu(x)\,|\log\lambda|\,(3 - f^{n_t}(x))}{(\log\beta)\, g_t(f^{n_t}(x))} \geq \frac{\mu(x)\,|\log\lambda|\,\lambda\, t}{(\log\beta)\,(1-\lambda)\, t} \geq \frac{\mu(x)\,\lambda\,|\log\lambda|}{(\log\beta)\,(1-\lambda)} > 1.$$

This ends the proof of the lemma. □

Remark 6.3. When $\mu = 1$ and $\beta\lambda = 1$ the hypothesis of Lemma 6.2 becomes $\lambda > 1/2$ or, equivalently, $\beta < 2$. This condition does not change if we replace f_t by its inverse.

6.1.3. *Unfolding heterodimensional cycles.* The dynamics associated to the unfolding of general heterodimensional cycles is very rich, and still far from being completely understood. However, a good number of results are available, especially for the codimension-1 quasi-transversal case. Here we review some of those results.

The cycles we considered in the previous section are *connected*, meaning that the intersection $\gamma = W^s(P) \cap W^u(Q)$ has finitely many connected components (only one, actually), and *non-critical*, meaning that the closed invariant set $\gamma \cup \{P, Q\}$ admits a Df-invariant dominated splitting of the form $E^u \oplus T\gamma \oplus E^s$. In addition we assumed that the restriction to the connection γ has *small distortion*. Due to non-criticality, no homoclinic tangencies related to either P or Q may occur in the unfolding of those cycles. As we are going to see, other types of cycles may exhibit very different dynamical behavior.

Firstly, [110] shows that the unfolding of *critical* cycles may lead to the phenomenon of persistence of tangencies. An interesting point is that this construction does not involve fractal dimensions, as was the case for the results in Section 3.2.

Next, in the same paper it is shown that the unfolding of *non-connected* cycles may lead to Morse-Smale dynamics after the bifurcation: essentially, one considers cycles such that the connected components of γ are small, compared to the length of fundamental domains of f near P and Q.

Finally, [111] exhibits connected non-critical cycles whose Mather invariant μ has *large distortion*, and for which there are sets of parameters t with positive Lebesgue density at $t = 0$ (the density can be made arbitrarily close to 1) such that the limit set of f_t in the neighborhood of the cycle is hyperbolic and coincides with the disjoint union of the homoclinic classes of P_t and Q_t.

A natural question in this context is whether having both homoclinic classes of P_t and Q_t hyperbolic suffices to guarantee hyperbolicity of the limit set in a neighborhood of the cycle. The answer turns out to be negative, as shown in [109]. In simple terms, the reason is that the unfolding of the cycle may contain the creation of secondary cycles, which are not associated neither to P_t nor to Q_t, and of new homoclinic classes disjoint from those of P_t and Q_t. Also, for any fixed k, it is possible to construct parametrized families generically unfolding a heterodimensional cycle such that, for a set of parameter values with positive density at $t = 0$, the limit set of f_t in a neighborhood of the cycle is hyperbolic and has at least k different homoclinic classes.

Problem 6.4. Given a parametrized family $(f_t)_{t\in[-1,1]}$ generically unfolding a connected non-critical cycle of $f = f_0$, is there a constant K and a set T with full Lebesgue measure in parameter space close to $t = 0$ such that, for every $t \in T$, the homoclinic classes of f_t in a neighborhood of the cycle are pairwise disjoint and their number is bounded by K ?

Another pair of questions naturally raised by the previous remarks:

Problem 6.5. Are there examples of parametrized families of diffeomorphisms generically unfolding a heterodimensional cycle for which there are (infinitely many) parameter values t such that:

- the map f_t has infinitely many pairwise disjoint homoclinic classes in a neighborhood of the cycle ?
- the homoclinic classes of P_t and Q_t are distinct but not disjoint ?

If such examples exist, generically, the set of such parameter values should have zero Lebesgue measure.

6.2. **Blenders.** While the arguments in Section 6.1.2 require a fair amount of differentiability, the conclusion of Theorem 6.1 does extend to the C^1 topology, as we shall see in Section 6.3. One way to prove this is using the notion of blender, which we discuss here.

Blenders have been used to obtain persistence of cycles and transitivity, cf. Section 7.3.3. In this regard, they play a role similar to that of thick Cantor sets in the phenomenon of persistence of tangencies in Section 3.2. With the advantage that, unlike thick Cantor sets, blenders are robust already in the C^1 category.

Let us point out that what makes blenders so useful is not so much the dynamics they carry, but rather their geometry. Indeed, in simple terms, *a blender is a hyperbolic basic set such that convenient projections of its stable, or unstable, set have larger topological dimension than the stable, or unstable, set itself.*

One such model appeared already in Example 3.24. This was a hyperbolic Cantor $H \subset \mathbb{R}^3$, with 1-dimensional stable manifolds and such the projection of its stable set $W^s(H)$ onto the plane generated by the stable and the weak unstable directions contained a whole open set. In fact, the topologically 1-dimensional set $W^s(H)$ intersects any curve roughly parallel to the strong unstable direction.

6.2.1. *A simplified model.* Let us start by describing a simple example that displays the main features of blenders. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a diffeomorphism of $\mathbb{R}^2$ such that the rectangle $R = [0,1]^2$ is a Smale horseshoe for f:

- $R \cap f(R)$ is the union of the vertical sub-rectangles $A = I_a \times [0,1]$ and $B = I_b \times [0,1]$, where I_a and I_b are closed intervals contained in $(0,1)$.
- $R \cap f^{-1}(R)$ is the union of the horizontal sub-rectangles $f^{-1}(A) = [0,1] \times J_a$ and $f^{-1}(B) = [0,1] \times J_b$, where J_a and J_b are closed intervals contained in $(0,1)$.
- The restrictions of f to $f^{-1}(A)$ and to $f^{-1}(B)$ are affine maps with linear part $\begin{pmatrix} \pm\frac{1}{3} & 0 \\ 0 & \pm 3 \end{pmatrix}$. In particular, such restrictions preserve the horizontal and vertical directions.

Denote by $q = (x_0, y_0)$ the unique fixed point of f in A.

Consider now a diffeomorphism $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ such that in $[-1,1] \times R$ is of the form

$$F(t,x,y) = \begin{cases} (\frac{5}{4}t, f(x,y)) & \text{if } (x,y) \in f^{-1}(A) \text{ and } t \in [-1,1], \\ (\frac{5}{4}t - \frac{1}{2}, f(x,y)) & \text{if } (x,y) \in f^{-1}(B) \text{ and } t \in [-1,1]. \end{cases}$$

The dynamics of F in the cube $\Gamma = [-1,1] \times R$ is hyperbolic, with $\mathbb{T} \oplus \mathbb{Y}$ as unstable direction ($\mathbb{T}$ is weak unstable and $\mathbb{Y}$ is strong unstable), and $\mathbb{X}$ as stable direction. The *blender* is the maximal invariant set inside Γ. See Figure 13.

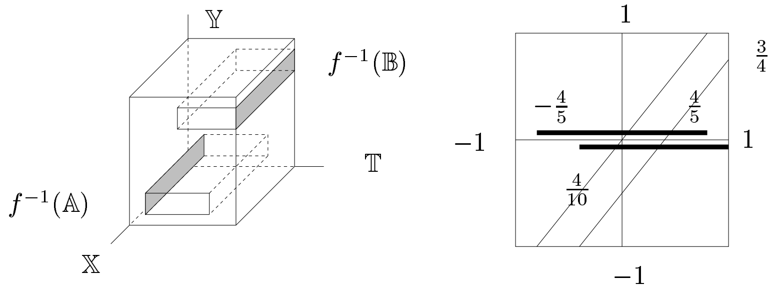

FIGURE 13. Blender with one-dimensional reduction

The distinctive geometric feature of blenders is expressed in the next lemma. For the statement we need some terminology. Observe first that $\Gamma \cap F(\Gamma)$ contains the cubes $\mathbb{A} = [-1,1] \times A$ and $\mathbb{B} = [-1, \frac{3}{4}] \times B$ and that $Q = (0, x_0, y_0) = (0, q) \in \mathbb{A}$ is a hyperbolic fixed saddle of F of index 2. Let $W^s_{loc}(Q) = \{0\} \times [0,1] \times \{y_0\}$ be the connected component of $W^s(Q) \cap \Gamma$ that contains Q.

- A *vertical segment to the right of* $W^s_{loc}(Q)$ is a segment $\sigma_{t,x} = \{t\} \times \{x\} \times [0,1]$, where $x \in [0,1]$ and $0 < t \leq 1$.
- A *vertical strip to the right of* $W^s_{loc}(Q)$ is a rectangle $\Delta = [t_1, t_2] \times \{x\} \times [0,1]$, where $x \in [0,1]$ and $0 < t_1 < t_2 \leq 1$. The number $w(\Delta) = (t_2 - t_1)$ is the *width of* Δ.

Lemma 6.6. *(blender distinctive property) Every vertical strip Δ to the right of $W^s_{loc}(Q)$ intersects $W^s(Q)$.*

Proof. We are going to show that if Δ is a vertical strip to the right of $W^s_{loc}(Q)$ then either $F(\Delta)$ intersects $W^s_{loc}(Q)$, or else $F(\Delta) \cap \Gamma$ contains a vertical strip Δ_1 to the right of $W^s_{loc}(Q)$ with $w(\Delta_1) = (5/4)\, w(\Delta)$.

To deduce the lemma from this claim, start with Δ as in the statement and consider its forward iterates. For as long as these iterates do not intersect the local stable manifold of Q, they contain vertical strips whose widths grow exponentially fast. Since the width is bounded above by 1, some iterate $F^n(\Delta)$ must intersect $W^s_{loc}(Q)$. This implies that Δ intersects the (global) stable manifold of Q. Hence, we are left to prove the claim in the previous paragraph.

Consider any vertical strip $\Delta = [t_1, t_2] \times \{x\} \times [-1,1]$ to the right of $W^s_{loc}(Q)$. Write $\Delta_A = F(\Delta) \cap \mathbb{A}$ and $\Delta_B = F(\Delta) \cap \mathbb{B}$.

- If $t_2 \leq \frac{4}{5}$ then $[\frac{5}{4} t_1, \frac{5}{4} t_2] \subset (0,1]$. It follows that Δ_A is a vertical strip to the right of $W^s_{loc}(Q)$, with width $w(\Delta_A) = \frac{5}{4}\, w(\Delta)$. Thus the second possibility in the claim holds.
- If $t_2 \in (\frac{4}{5}, 1]$ then $\frac{5}{4} t_2 - \frac{1}{2} \in (\frac{1}{2}, \frac{3}{4}]$. By the definition of F, there is x' such that
$$\Delta_B = F(\Delta) \cap \mathbb{B} = \{[\frac{5}{4} t_1 - \frac{1}{2}, \frac{5}{4} t_2 - \frac{1}{2}]\} \times \{x'\} \times [0,1].$$
There are now two subcases:

- If $\frac{5}{4}t_1 - \frac{1}{2} > 0$ then Δ_B is a vertical strip to the right of $W^s_{loc}(Q)$, with width $w(\Delta_B) = (5/4)\,w(\Delta)$. So, the second possibility in the claim holds in this case too.
- Otherwise $\frac{5}{4}t_1 - \frac{1}{2} \le 0$ and $\frac{5}{4}t_2 - 1/2 \ge 0$. Then Δ_B meets $W^s_{loc}(Q) = \{0\} \times [0,1] \times \{y_0\}$, which corresponds to the first possibility in the claim.

The proof is complete. □

As we did in the previous section, we may consider the quotient of the dynamics by the sum of the strong stable and strong unstable directions. This yields an iterated functions system $\mathcal{G}(\phi_1, \phi_2)$ with two generators, as follows:

- for $t \in [-\frac{4}{5}, \frac{4}{5}]$ consider $\phi_1(t) = \frac{5}{4}t$, corresponding to the restriction of F to the domain $F^{-1}(\mathbb{A}) = [-\frac{4}{5}, \frac{4}{5}] \times f^{-1}(A)$,
- for $t \in [-\frac{2}{5}, \frac{4}{5}]$ consider $\phi_2(t) = \frac{5}{4}t - \frac{1}{2}$, corresponding to the restriction of F to the domain $F^{-1}(\mathbb{B}) = [\frac{-2}{5}, \frac{4}{5}] \times f^{-1}(B)$.

Notice that both maps ϕ_1 and ϕ_2 are expanding, and their domains of definition overlap on the interval $[-\frac{2}{5}, \frac{4}{5}]$. This reflects the fact that given any vertical segment $\sigma = \{t\} \times \{x\} \times [0,1]$ with t in the overlapping interval, there are x_1 and x_2 such that the vertical segments $\{\phi_1(t)\} \times \{x_1\} \times [0,1]$ and $\{\phi_2(t)\} \times \{x_2\} \times [0,1]$ are the two connected components of $F(\sigma) \cap \Gamma$.

Lemma 6.6 may now be reformulated for the iterated functions system, as follows. We also translate the proof, since it is very short.

Lemma 6.7. *Given any open interval $\alpha \subset [0,1]$ there is $\phi \in \mathcal{G}(\phi_1, \phi_2)$ such that $0 \in \phi(\alpha)$.*

Proof. Since $\phi_2(\frac{2}{5}) = 0$, it is enough to exhibit $\phi \in \mathcal{G}(\phi_1, \phi_2)$ such that $\phi(\alpha)$ contains $\frac{2}{5}$. Given $\alpha \subset [0,1]$ we let $\alpha_0 = \alpha$ and inductively define subintervals α_n as follows. Suppose α_n does not contain $\frac{2}{5}$. Then there are two possibilities. If $\alpha_n \subset (0, \frac{4}{5}]$, we let $\alpha_{n+1} = \phi_1(\alpha_n)$. Otherwise, $\alpha_n \subset (\frac{2}{5}, 1]$ and we let $\alpha_{n+1} = \phi_2(\alpha_n)$. In both cases, $\alpha_{n+1} \subset (0,1]$ and $\operatorname{length}(\alpha_{n+1}) = \frac{5}{4} \operatorname{length}(\alpha_n) = (\frac{5}{4})^n \operatorname{length}(\alpha_0)$. This means that α_n must, eventually, contain $\frac{2}{5}$. At that point we are done. □

To illustrate the dynamical usefulness of blenders , we close this section with the following direct consequence of Lemma 6.6:

Lemma 6.8. *Suppose that there is a hyperbolic periodic point P of F of index 1 whose one-dimensional unstable manifold crosses Γ along a vertical segment to the right of $W^s_{loc}(Q)$. Then $W^s(P) \subset \overline{W^s(Q)}$.*

So the one-dimensional stable manifold of Q looks like a 2-dimensional manifold, as its closure contains the 2-dimensional $W^s(P)$.

6.2.2. *Relaxing the construction.* So far we discussed a very simple affine model of blender. Of course, for most applications we need more flexible versions of the previous constructions, robust under C^1 small perturbations. Just as an example, in the context of Lemma 6.8 we would like to have the conclusion when the segment is just approximately vertical. We are going to explain how this flexibilization is carried out, starting with perturbations of the affine model, and to indicate other ways in which the construction may be generalized. Motivated by this, we discuss possible definitions of the notion of blender.

In the affine model F we have been considering, $\mathbb{X} \oplus \mathbb{T} \oplus \mathbb{Y}$ is a dominated on Γ, with XX as strong stable, $\mathbb{Y}$ as strong unstable, and $\mathbb{T}$ as central (weak unstable) direction. Hence, there are (thin) cone fields $\mathcal{C}^s$, $\mathcal{C}^{uu}$, $\mathcal{C}^u$ around the directions $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{T} \oplus \mathbb{Y}$, respectively, strictly invariant under the derivative DF. These cone fields allow us to define *almost vertical segments (to the right of* $W^s_{loc}(Q)$*)*, as curve segments tangent to $\mathcal{C}^{uu}$ and crossing Γ (on the right hand side of the local stable manifold), as well as *almost vertical strips (to the right of* $W^s_{loc}(Q)$*)*, as 2-dimensional disks through Γ that are tangent to $\mathcal{C}^u$ and foliated by almost vertical segments (to the right of $W^s_{loc}(Q)$). The notion of width extends to these almost vertical strips: the length of the shortest curve transverse to $\mathcal{C}^{uu}$ joining the two boundary components inside the strip.

The families of almost vertical segments and strips are invariant under forward iteration, in the same sense as before for truly vertical segments and strips. Moreover, the widths of almost vertical strips not intersecting $W^s_{loc}(Q)$ increase exponentially under iteration, again in the same sense as before. This means that the proof of Lemma 6.6 carries through to this context, only with extra technical difficulties. We get that *every almost vertical strip at the right of* $W^s_{loc}(Q)$ *intersects* $W^s(Q)$.

Now, the cone fields remain invariant for any diffeomorphism in a C^1 neighborhood of F. Moreover, the width estimate is also robust under any small C^1 perturbation, because it only depends only on the fact that the derivative expands all vectors in the cone $\mathcal{C}^u$. This means that the conclusion of the previous paragraph is true for any map near F: *the distinctive blender property is a robust property!*

Blenders from a constructive viewpoint. One way the construction of blenders may be relaxed, with respect to the affine model in Section 6.2.1, is allowing for more complicated combinatorics: any number of dynamical components, with variable return times to the ambient box.

Consider a C^1 diffeomorphism F and a box $\Gamma = D^k \times [-1,1] \times D^l$, where $D^i = [-1,1]^i$, in the ambient space, endowed with cone fields $\mathcal{C}^s$ and $\mathcal{C}^{uu} \subset \mathcal{C}^u$ around the directions of D^k, D^l, and $[-1,1] \times D^l$, respectively. Assume that there are sub-boxes $(\Gamma_i)_{i=1}^k$ of Γ, and return times $(n_i)_{i=1}^k$ such that every return map $F^{n_i}\colon \Gamma_i \to \Gamma$ preserves both cone fields and expands uniformly all vectors in $\mathcal{C}^u$.

Define as above almost vertical (l-dimensional) disks, tangent to $\mathcal{C}^{uu}$, and almost vertical ($l+1$-dimensional) strips, tangent to $\mathcal{C}^u$ and foliated by almost vertical disks crossing Γ. Define the width $w(\mathcal{S})$ of an almost vertical strip $\mathcal{S}$, again as before. We turn the type of property in Lemma 6.6 into our first definition of a blender:

Definition 6.9. (Γ, F) is a *blender associated to a periodic saddle* Q *of index* $l+1$ if for every almost vertical strip $\mathcal{S}$ crossing Γ there is $i \in \{1, \ldots, k\}$ such that

- either $F^{n_i}(\mathcal{S} \cap \Gamma_i)$ contains an almost vertical strip $\mathcal{S}'$ intersecting $W^s_{loc}(Q)$
- or $F^{n_i}(\mathcal{S} \cap \Gamma_i)$ contains an almost vertical strip $\mathcal{S}'$ with $w(\mathcal{S}') \geq \lambda\, w(\mathcal{S})$, where $\lambda > 1$ is independent of the strip $\mathcal{S}$.

The periodic point Q in the definition does not need to belong to the cube Γ: our arguments just require that $W^s(Q)$ contain a disk Δ crossing Γ and tangent to the cone field $\mathcal{C}^s$. The reader may check that the unfolding of the heterodimensional cycle in Section 6.1.2 yields a blender inside a box close to the fixed point Q but not necessarily containing it.

In the particular case where the maps F^{n_i} preserve the D^k and D^l directions, as was the case for the affine blender, one may consider the quotient dynamics of F by the sum $D^k \oplus D^l$. In this way we obtain, for each i, a one-dimensional expanding map $\phi_i : I_i \to [0,1]$ induced by F^{n_i} on a compact $I_i \subset [0,1]$. (Γ, F) being a blender corresponds to the maps ϕ_i having the covering property: $\cup_i \operatorname{int}(I_i) = (0,1)$. A point $x \in [0,1]$ is *periodic* for the iterated functions system $\mathcal{G}(\phi_1, \ldots, \phi_k)$ if $\phi(x) = x$ for some $\phi \in \mathcal{G}$. Extending Lemma 6.7, we have

Proposition 6.10. *Given any system $\mathcal{G}(\phi_1, \ldots, \phi_k)$ of iterated expanding functions with the covering property, there is a finite set $\mathcal{F} = \{x_1, \ldots, x_s\}$, $1 \le s < k$, of periodic points such that, for any open interval $\alpha \subset [0,1]$, there is $\phi \in \mathcal{G}$ with $\phi(\alpha) \cap \mathcal{F} \neq \emptyset$.*

Blenders from an operational viewpoint. We now adopt a different point of view concerning blenders: we will think of them as tools for proving transitivity. Thus we propose another definition which takes into account their properties, rather than their construction.

Definition 6.11. Let $F : M \to M$ be a diffeomorphism and Q be a hyperbolic saddle of index $l+1$. We say that *F has a blender associated to Q* if there is a C^1 neighborhood $\mathcal{U}$ of F, and a C^1 open set $\mathcal{D}$ of embeddings of an l-dimensional disk D^l into M, such for every $G \in \mathcal{U}$, every disk $D \in \mathcal{D}$ intersects the closure of $W^s(Q_G)$, where Q_G is the continuation of the periodic point Q for G.

The set $\mathcal{D}$ in the definition is the *superposition region* of the blender. For instance, in the simplified model considered before, $l = 1$ and the superposition region is the family of all almost vertical segments (that is, tangent to a thin vertical cone field and crossing the box Γ) to the right of $W^s_{loc}(Q)$. Bearing in mind the distinctive property of blenders, one say that the blender is *activated* by a saddle P of index l if the unstable manifold of P contains a disk of the superposition region.

Lemma 6.12. *Let F be a diffeomorphism having a blender associated to a saddle Q of index $l+1$. Suppose that the blender is activated by a saddle P of index ℓ. Then, for every diffeomorphism G in a C^1 neighborhood of F, the closure of $W^s(Q_G)$ contains $W^s(P_G)$, where P_G and Q_G are the continuations of P and Q for G.*

As noted before, the point with this result is that it means that the $(d-l-1)$-dimensional manifold $W^s(Q_G)$ of Q_G looks like it contains a manifold with higher dimension, namely $W^s(P_G, G)$.

Proof. Let $\mathcal{D}$ the superposition region of the blender. As compact parts of invariant manifolds depend continuously on the diffeomorphism, $W^u(P_G)$ contains a disk $D_G \in \mathcal{D}$, for all G close to F. Let U be any open set intersecting $W^s(P_G)$. By the λ-lemma, for every n big enough, $G^n(U)$ contains a C^1-open set of disks arbitrarily close to D_G. Then these disks are in the superposition region. By definition of blender, it follows that $W^s(Q_G)$ intersects $G^n(U)$. As the stable manifold is invariant, it also intersects U. □

6.3. **Partially hyperbolic cycles.** We are going to see that the formation of intermingled homoclinic classes is a frequent phenomenon in the unfolding of cycles, even in the C^1 topology.

Consider a diffeomorphism f with a cycle associated to saddles P and Q of indices p and $p+1$, respectively. We say that the cycle is *C^k-far from homoclinic* *tangencies* if every diffeomorphism g which is C^k close to f does not have homoclinic tangencies associated either of the continuations P_g and Q_g of P and Q.

Theorem 6.13 (Díaz, Rocha [112]). *Let f be a C^1 diffeomorphism with a heterodimensional cycle associated to saddles P and Q of indices p and $q = p + 1$. Suppose that the cycle is C^1-far from homoclinic tangencies. Then there is an open set $\mathcal{U} \subset \operatorname{Diff}^1(M)$ whose closure contains f and such that the homoclinic classes of P_g and Q_g coincide for all $g \in \mathcal{U}$.*

This theorem also holds for non-critical connected cycles, without any assumption on the Mather invariant. In that case we even have the following version for one-parameter families generically unfolding the cycle: *Given $t > 0$, let $B(t)$ be the interior of the set of parameters $s \in [-t, t]$ such that the homoclinic classes of P_s and Q_s coincide. Then $B(t)$ has positive Lebesgue measure for every $t > 0$.* Moreover, if the cycle corresponds to a first bifurcation, the relative density of $B(t)$ at the bifurcation is positive:

$$\liminf_{t \to 0^+} \frac{1}{t} \operatorname{Leb}\big(B(t)\big) > 0.$$

Theorem 6.13 is a consequence of the following proposition, whose proof uses a blender construction, with several generators.

Proposition 6.14. *Let f be a diffeomorphism as in Theorem 6.13. Then there is an open set $\mathcal{V} \subset \operatorname{Diff}^1(M)$ whose closure contains f such that for every g in $\mathcal{V}$ there are blenders $\Gamma(g)$ and $\Gamma(g^{-1})$, defined for g and g^{-1}, respectively, such that:*

- *$\Gamma(g)$ is associated to a hyperbolic periodic point R_g homoclinically related to Q_g and is activated by P_g.*
- *$\Gamma(g^{-1})$ is associated to a hyperbolic periodic point S_g homoclinically related to P_g and is activated by Q_g.*

Observe that $\Gamma(g^{-1})$ is defined for g^{-1}: to be activated by Q_g means that the stable manifold of Q_g contains a (p-dimensional) disk of the superposition region. Since R_g and Q_g (respectively, S_g and P_g) are homoclinically related, Lemma 6.12 and Proposition 6.14 imply that, for every g in the open set $\mathcal{V}$,

$$W^s(P_g) \subset \overline{W^s(R_g)} = \overline{W^s(Q_g)} \quad \text{and} \quad W^u(Q_g) \subset \overline{W^u(S_g)} = \overline{W^u(P_g)}.$$

This is the main step towards showing that the homoclinic classes of P_g and Q_g coincide.

7. Robust transitivity and projective hyperbolicity

An f-invariant compact set Λ is *robustly transitive* if there is an open neighborhood U of Λ and a neighborhood $\mathcal{V}$ of f in $\operatorname{Diff}^1(M)$ such that, for every $g \in \mathcal{V}$, the maximal invariant set Λ_g of g in $\bar{U}$ is a transitive set contained in U and $\Lambda_f = \Lambda$. This notion was introduced in [108] as a natural generalization of the elementary pieces of hyperbolic systems in Theorem 1.3. A special case occurs when U is the whole manifold M. In this case we call the diffeomorphism itself robustly transitive. A weaker notion of *generically transitive set* will be discussed in Section 10.3.

Although a robustly transitive set persists after any C^1 perturbation, by definition, its dynamics needs not be stable: for instance, periodic points may bifurcate. However, such bifurcations can never produce neither sinks nor sources. In the case of surface diffeomorphisms this implies that the periodic points inside a robustly transitive set are persistently hyperbolic, in a uniform fashion. In fact,

Theorem 7.1 (Mañé [206]). *Every robustly transitive set of a surface C^1 diffeomorphism is a hyperbolic set.*

7.1. **Consequences of robust transitivity.** In arbitrary dimension, the absence of creation of sinks or sources implies the following generalization of Theorem 7.1:

Theorem 7.2 (Bonatti, Díaz, Pujals, Ures [58, 108]). *Every robustly transitive set of a C^1 diffeomorphism is volume hyperbolic.*

An invariant set Λ is *volume hyperbolic* if there is a splitting $T_\Lambda M = E^1 \oplus \cdots \oplus E^k$ into Df-invariant subbundles such that

(1) the splitting is *dominated*: there exists $m \geq 1$ such that
$$\|Df^n \mid E^i(x)\| \, \|(Df^n \mid E^j(x))^{-1}\| \leq 1/2$$
for every $n \geq m$, $i > j$, and $x \in \Lambda$.

(2) the extreme subbundles E^1 and E^k are uniformly volume expanding and volume contracting: for all $n \geq m$ and $x \in \Lambda$,
$$|\det(Df^{-n} \mid E^1(x))| \leq 1/2 \quad \text{and} \quad |\det(Df^n \mid E^k(x))| \leq 1/2$$

In dimension 2, the only possibility is the splitting to be into two 1-dimensional subbundles ($2 = 1+1$), and then volume hyperbolicity is the same as hyperbolicity. Therefore, Theorem 7.2 contains Theorem 7.1. Moreover, in dimension 3 there are two possibilities, $3 = 2 + 1$ and $3 = 1 + 1 + 1$. In either case, at least one of the subbundles is 1-dimensional and hyperbolic: either uniformly expanding or uniformly contracting. Thus, in this case robust transitivity implies partial hyperbolicity [108].

These results are best possible. Indeed, examples in [206] show that robustly transitive sets in dimension larger than 2 may not be hyperbolic. In Mañé's construction the set admits a dominated splitting $T_\Lambda M = E^1 \oplus E^2 \oplus E^3$ into three 1-dimensional subbundles, the extreme ones being hyperbolic: uniformly expanding and uniformly contracting, respectively. Then appeared examples where the finest dominated splitting consists only of two bundles, the 1-dimensional one being hyperbolic and the 2-dimensional one being volume hyperbolic but not hyperbolic. See [67]. Finally, in dimension greater than or equal to 4, robustly transitive sets may not be partially hyperbolic, that is, they may admit no hyperbolic invariant subbundles. See [67]. We review these and other examples in Section 7.3.

As mentioned before, the way to prove Theorem 7.2 is by showing that the lack of a dominated splitting leads to the creation of a sink or source, after a convenient perturbation. If the lack of dominated splitting is persistent, this yields infinitely many coexisting sinks or sources! We may apply this idea to a transitive set naturally associated to every hyperbolic saddle point p of a diffeomorphism f, its *homoclinic class*. This is defined as the closure of the transverse intersections between its invariant manifolds and denoted by $H(p, f)$. See Section **??** for properties of homoclinic classes.

Theorem 7.3 (Bonatti, Díaz, Pujals [58])). *There exists a residual subset $\mathcal{R}$ of $\mathrm{Diff}^1(M)$ such that for every diffeomorphism $f \in \mathcal{R}$ and periodic point p of f,*

- *either $H(p, f)$ is contained in the closure of infinitely many sources or sinks,*

- *or $H(p, f)$ admits a dominated splitting.*

In the rest of this section we explain the main ideas in the proofs of the theorems above. Roughly speaking, Theorem 7.2 follows from an idea in linear algebra: in the absence of domination, the eigenspaces of a linear map obtained by multiplication of several linear maps are very unstable: by small perturbation of each of the factors, one can mix the eigenvalues in order to get a homothetic transformation, which then leads to a new sink or a new source. The proof of volume expansion in the extremal subbundles uses the ergodic closing lemma [206].

7.1.1. *Lack of domination and creation of sinks or sources.* We begin the exposition with, simpler, the 2-dimensional case. The argument we are going to present, due to Mañé [206], shows the following: *if a set $\mathcal{P}$ of hyperbolic periodic saddles of a diffeomorphism of a compact surface does not admit a dominated splitting then it is possible to turn one of such periodic points to be a new sink or source, by a C^1 small perturbation.*

We begin with an elementary property of matrices in $GL(2, \mathbb{R})$. Denote by R_α the rotation of angle α in $\mathbb{R}^2$.

Lemma 7.4. *Consider an orientation preserving matrix $A \in GL(2, \mathbb{R})$ having two different real eigenvalues and whose eigenspaces form an angle less than $\varepsilon > 0$. Then there is $t \in [-\varepsilon, \varepsilon]$ such that $R_t \circ A$ has a pair of conjugate complex eigenvalues.*

A way to prove this is to consider the action h_t of $R_t \circ A$ on the projective space $\mathbb{RP}^1$, identified with the circle S^1. By hypothesis, the map h_0 has two hyperbolic fixed points, at distance less than ε. Moreover, outside the small interval between these fixed points, the sign of $h_0(x) - x$ is constant, say positive. So, $h_0(x) - x > -\varepsilon$ for every x in S^1. Consequently, $h_\varepsilon(x) - x > 0$ for all x in S^1. This implies that the rotation number of the circle map h_t is not constant in $[0, \varepsilon]$. Choosing t in this interval such that h_t has non-zero rotation number, we get that the eigenvalues of $R_t \circ A$ cannot be real, as claimed.

Now consider the natural splitting $T_pM = E^u(p) \oplus E^s(p)$ into eigenspaces, defined over $\mathcal{P}$.

Lemma 7.5. *Suppose that the angle* $\operatorname{angle}(E^s(p), E^u(p))$ *between $E^s(p)$ and $E^u(p)$ is not bounded away from 0 for $p \in \mathcal{P}$. Then there are g arbitrarily C^1-close to f and $p \in \mathcal{P}$ such that p is either a sink or a source of g.*

Lemma A.8 asserts that any small perturbation of the derivative of a diffeomorphism f along a periodic orbit p can be realized dynamically, that is, as the derivative of a C^1-nearby diffeomorphism g for which p is still a periodic point. Then, it is enough to see that there are arbitrarily small perturbations of the derivative of f along the orbit of some $p \in \mathcal{P}$ such that the resulting matrix has non-real eigenvalues of absolute value different from 1.

This is deduced from Lemma 7.4 as follows. Choose a point $p \in \mathcal{P}$ such that $\operatorname{angle}(E^s(p), E^u(p))$ is less than some small fixed ε and let n be the period of p. Then compose the derivative $Df(f^{n-1}(p))$ with a rotation $R_t \colon T_p(M) \to T_p(M)$ for an appropriate t, so that the resulting linear map

$$R_t \cdot Df(f^{n-1}(p)) \cdot Df(f^{n-2}(p)) \cdots Df(p)$$

has complex conjugate eigenvalues. Up to a final perturbation, the absolute value may be taken different from 1.

Lemma 7.5 completes the argument when the angle between E^s and E^u is not bounded away from 0. So, we are left to consider the case in which this angle is uniformly bounded from below on $\mathcal{P}$. In this case, up to a uniformly bounded change of the Riemannian metric, there are unitary vectors $e^s(p) \in E^s(p)$ and $e^u(p) \in E^u(p)$ forming an orthonormal basis of T_pM. In this basis, the derivative $Df(p)$ is given by a diagonal matrix. The set $\mathcal{P}$ is called *hyperbolic at the period* if there is $0 < \lambda < 1$ such that

$$\|Df^{n_p}(p)e^s(p)\| \le \lambda^{n_p} \quad \text{and} \quad \|Df^{-n_p}(p)e^u(p)\| \le \lambda^{n_p}$$

for every $p \in \mathcal{P}$, where n_p is the period of p.

Lemma 7.6. *Suppose $\mathcal{P}$ is not hyperbolic in the period. Then there exists a small C^1-perturbation g of f and a periodic point $p \in \mathcal{P}$ such that p is either a sink or a source of g.*

This follows by multiplying each matrix $Df(p)$ by a homothetic transformation of ratio $(1+\varepsilon)$ or $(1-\varepsilon)$, depending on which direction hyperbolicity fails, and then applying Lemma A.8.

Finally, consider the case where $\mathcal{P}$ is hyperbolic in the period. As we are assuming that $\mathcal{P}$ does not admit a dominated splitting, there are arbitrarily large n and $p \in \mathcal{P}$ such that $n < n_p$ and

$$\frac{\|Df^n(p)e^u(p)\|}{\|Df^n(p)e^s(p)\|} < 2.$$

Since $\mathcal{P}$ is hyperbolic in the period, $n_p - n$ must go to infinite as n goes to infinity.

Lemma 7.7. *Assume the splitting $E^u \oplus E^s$ over $\mathcal{P}$ is not dominated. Then there is an arbitrarily small perturbation of the derivative of f along the orbit of some point $p \in \mathcal{P}$ such that the angle between the eigenspaces of the corresponding matrix is arbitrarily small.*

This reduces the problem to the situation in Lemma 7.5, thus ending the proof of the statement made at the beginning of this section. So let us now explain the proof of this lemma.

Consider n and p as in the inequality above. Since n is large and the matrices of the derivatives are diagonal, by multiplying each $Df(f^i(p))$, $0 \le i < n$ by

$$C = \begin{pmatrix} \dfrac{1}{1+\varepsilon} & 0 \\ 0 & 1+\varepsilon \end{pmatrix}$$

we can make the quotient $\|Df^n(p)(e^u(p))\|/\|Df^n(p)(e^s(p))\|$ arbitrarily small, Denote $A(f^i(p)) = C \cdot Df(f^i(p))$. Then consider

$$B = \begin{pmatrix} 1 & 0 \\ \varepsilon & 1 \end{pmatrix}$$

and replace $A(p)$ by $A'(p) = A(p) \cdot B$. At this stage the direction $E^u(p)$ is no longer invariant, but our perturbations do keep $E^s(p)$ invariant. The linear map $A(f^{n-1}(p)) \circ \cdots \circ A(f(p)) \circ A'(p)$ sends $E^s(p)$ to $E^s(f^n(p))$, and $E^u(p)$ to some straight line $\tilde{E}(f^n(p))$ forming a small angle with $E^s(f^n(p))$, provided n is large enough. Through a final perturbation at times $> n$ we are going to recover the invariance of E^u, while keeping E^s invariant. This means they will again be the eigenspaces and so, since at time n their angle is small, the proof will be complete.

Because of hyperbolicity at the period, the image of $e^u(p)+\varepsilon e^s(p)$ under $A^{n_p}(p)$ makes a small angle with $E^u(p)$. In other words, $A(f^{n_p-1}(p))\circ\cdots A(f(p))\circ A'(p)$ maps $E^s(p)$ to itself and $E^u(p)$ to a subspace $\tilde{E}^u(p)$ forming a small angle with $E^u(p)$. Then there exists a matrix J close to the identity mapping $E^s(p)$ into $E^s(p)$ and $\tilde{E}^u(p)$ into $E^u(p)$. Thus, $E^s(p)$ and $E^u(p)$ are eigenspaces of

$$J\circ A(f^{n-1}(p)))\circ A(f^{n-2}(p)))\circ\cdot\circ A(f(p))\circ A'(p).$$

This is the final perturbation mentioned above, and so the proof of the lemma is complete.

7.1.2. *Dominated splittings versus homothetic transformations.* The previous ideas extend easily to arbitrary dimension. One gets that, given a set $\mathcal{P}$ of hyperbolic periodic saddles all with the same Morse index (the same dimension of the unstable direction), if the natural splitting induced by the stable and unstable directions is not dominated then it is possible to change the index of some saddle by a small C^1 perturbation.

However, in dimension larger than 2 such bifurcations do not break transitivity, and they do occur in robustly transitive sets. Towards proving Theorem 7.2, we are going to show that *in the absence of any dominated splitting it is possible to perturb the derivative of f along a periodic orbit in order to turn into a homothetic transformation and, after an additional small perturbation, to create a sink or a source for a diffeomorphism close to f.* An important tool is Frank's lemma A.8, according to which C^1 small perturbations of the derivative can always be realized dynamically.

The main novelty with respect to the previous arguments is the introduction of *transitions* between different periodic orbits, which permit a kind of concatenation of such orbits and corresponding multiplication of eigenvalues. In a few words, while in the 2-dimensional situation presented above we deal with a linear cocycle defined over period points and their perturbations only, the arguments that follow use strongly the dynamics on a homoclinic class and the continuity of the derivative: the transitions correspond to heteroclinic orbits.

We begin with another argument of linear algebra which is crucial in the proof. Consider a basis $\beta=\{v_1,\dots,v_n\}$ of $\mathbb{R}^n$ and let L_i be the subspace spanned by v_i. Consider a linear map $A\in GL(\mathbb{R},n)$ which is diagonal relative to the basis β, and $(n-1)$ linear maps $A_1,A_2,\dots,A_{n-1}\in GL(\mathbb{R},n)$ such that each A_j

- leaves invariant the subspaces $L_1,\dots,L_{j-1},L_{j+2},\dots,L_n$ and
- permutes L_j and L_{j+1}, that is, $A_j(L_j)=L_{j+1}$ and $A_j(L_{j+1})=L_j$.

For arbitrarily small neighborhoods $\mathcal{A}$ of A and $\mathcal{A}_i$ of each A_i, consider all possible products of matrices in $\mathcal{A}\cup(\cup_i\mathcal{A}_i)$. Then some of these products are homothetic transformations. The idea is that, as the matrices A_i generate all permutations of the eigenspaces of A, it is possible to mix the eigenvalues of A to obtaining a matrix having essentially the same rate of expansion in all directions.

In the model we have in mind, the matrices A and A_i correspond to the derivatives $Df^k(p)$, $k=\operatorname{per}(p)$ at periodic points p of a diffeomorphism f. Of course, to multiply derivatives corresponding to different points does not make dynamical sense. However, properties of homoclinic classes $H(p,f)$ allows us to give a sense to this procedure.

First of all, there is a dense subset K in $H(p,f)$ consisting of periodic points homoclinically related to p. As dominated splitting extends to the closure of its domain, absence of dominated splitting on $H(p,f)$ implies the same for K. Thus, it is enough to consider this set K of periodic orbits.

Next, given two fixed or periodic hyperbolic saddles x and y in K, and given any family $\{n_1, m_1, \ldots, n_k, m_k\}$ of arbitrarily big natural numbers, there are a natural number r (independent of n_i and m_i) and a saddle $z \in H(p,f)$ of period $t = kr + \sum_i (n_i + m_i)$ whose orbit spends alternately n_i consecutive iterates close to x and m_i consecutive iterates close to y. This means that the derivative of $Df^t(z)$ is, roughly, of the form

$$Df^t(z) = (T_2 \circ Df^{m_k}(q) \circ T_1 \circ Df^{n_k}(p)) \circ \cdots \circ (T_2 \circ Df^{m_1}(q) \circ T_1 \circ Df^{n_1}(p))$$

where T_1 and T_2 are called the *transitions* from x to y and from y to x, respectively. These transitions correspond to a bounded number of iterations, independent of n_i and m_i, so that their contribution to the product becomes negligible as n_i and m_i go to infinite.

Thus, it makes sense to multiply derivatives corresponding to different periodic points if they are homoclinically related. Besides the resulting linear maps correspond (roughly) to the derivative at another periodic saddle in the homoclinic class. Moreover, in any homoclinic class there is a dense subset of periodic points whose derivative at the period can be made diagonal, with positive eigenvalues of multiplicity 1, by a small perturbation of the derivative along the orbit.

The construction above shows that a robust transitive set always admits a dominated splitting. Then the *finest* such splitting is well-defined. To get volume hyperbolicity it remains to prove uniform volume expansion and volume contraction along the two extreme subbundles. This is an application of Mañé's ergodic closing lemma [206], as we explain next.

Since E^1 does not admit any dominated sub-splitting, the previous arguments ensure that by a small perturbation we can produce a homothetic transformation along E^1. If E_1 is not uniformly volume expanding, by using the ergodic closing lemma we get a homothetic transformation along E^1 at a periodic point whose volume contraction rate is less than $1 + \delta$. This is explained in more detail in Appendix A.2. So, maybe after a new perturbation, we get a homothetic transformation with contraction rate strictly smaller than 1. Then, by domination, implies that all the subbundles E_i are contracting. This means that the periodic point is really a source, which is a contradiction.

7.2. **On the dynamics of robustly transitive sets.** Here the is to deduce dynamical properties of robustly transitive sets, such as the kind of bifurcations they may exhibit, or the types of saddles they may contain, from the analysis of the finest dominated splitting.

The arguments in the previous section imply that *if it is not possible to change, by perturbation, the index of a periodic point in a robustly transitive set Λ, then Λ is hyperbolic.* In the same spirit, the following result explains what happens when it is not possible to increase the indices of the periodic saddle points of Λ: it relates the maximal and minimal indices with the hyperbolicity of the subbundles in the finest dominated splitting.

Theorem 7.8 (Bonatti, Díaz, Pujals, Rocha [59]). *Consider $k \geq 1$ and an open subset U of a closed manifold. Let $\mathcal{U}$ be an open subset of $\mathrm{Diff}^1(M)$ of diffeomorphisms f such that $\Lambda_f(U)$ is robustly transitive, has a dominated splitting $E_f \oplus F_f$ with $\dim E_f = k$, and does not contain saddles of Morse index less than k. Then E_f is uniformly expanding for every diffeomorphism f in $\mathcal{U}$.*

If Λ is not hyperbolic then, up to a perturbation, it contains points of different indices. What can be said about the indices in between the maximal and minimal ones are different?

Problem 7.9. Given a robustly transitive set Λ having points of indices k and m, does Λ contain a point of index l for each $l \in (k, m)$?

The answer is affirmative for most robustly transitive sets, as we are going to see. Given an open set $U \subset M$ denote by $\mathcal{R}(U)$ the subset of diffeomorphisms $f \in \mathrm{Diff}^1(M)$ for which $\Lambda_f(U) = \cap_{\mathbb{Z}} f^i(U)$ is a robustly transitive set.

Theorem 7.10 (Bonatti, Díaz, Pujals, Rocha [59]). *There is an open and dense subset $\mathcal{I}(U)$ of $\mathcal{R}(U)$ such that for all $g \in \mathcal{I}(U)$ the set of indices of the hyperbolic saddles of $\Lambda_g(U)$ is an interval in $\mathbb{N}$.*

The main ingredient in the proof is a notion of transitions associated to a heterodimensional cycle. These are similar to the transitions associated to homoclinically related periodic points in Section 7.1.2, but there are also some subtle differences that we now explain.

Given two periodic points x and y of Λ with different indices, using robust transitivity and the connecting lemma, Theorem A.5, we may create a heterodimensional cycle involving x and y. Then there is a perturbation of f having a periodic orbit which remains arbitrarily large periods close to x and close to y, with bounded transition times from one saddle point to the other. The derivative on this periodic orbit is a small perturbation of the product of arbitrarily large powers of the derivatives at x and y by a pair of bounded matrices, the *transitions*. One difference with respect to Section 7.1.2 is that here the new periodic orbits passes close to x and y only once, it closes the period already at the first return. A second difference is that in the homoclinic case the transitions are for the initial diffeomorphism, whereas here they are defined for perturbations of it.

Another important topic is the occurrence of homoclinic tangencies inside robustly transitive sets. Let us observe that this kind of bifurcation is indeed compatible with robust transitivity because, in higher dimensions, the creation of homoclinic tangencies needs not lead to the creation of sinks or sources, unless the saddle point involved in the tangency is sectionally dissipative in the sense of Section 3.5.

The heuristic principle is that *if the finest dominated splitting of $\Lambda_f(U)$ involves a non-hyperbolic subbundle of dimension bigger than 1 then there is g close to f such that $\Lambda_g(U)$ contains a homoclinic tangency.* Moreover, it is possible to determine the index of the hyperbolic saddle associated to the tangency. See [59] for the precise statements. On the other hand, if all non-hyperbolic subbundles of the finest dominated splitting are 1-dimensional then it is not possible to create homoclinic tangencies inside the robustly transitive set by small perturbations of the system. In this case the prototypical bifurcations are heterodimensional cycles and saddle-node and period-doubling bifurcations.

7.3. **Examples of robust transitivity.** The best known examples of robustly transitive diffeomorphisms are the Anosov diffeomorphisms, especially the linear ones, that is the diffeomorphisms induced on the torus T^n by a hyperbolic map in $GL(n,\mathbb{Z})$. Several questions about Anosov diffeomorphisms remain open, in particular it is not known which manifolds support such diffeomorphisms. Here we are going to focus on *non-hyperbolic* robustly transitive diffeomorphisms.

The first examples of these diffeomorphisms were constructed in the seventies by Shub [312] and Mañé [205], respectively, on T^4 and on T^3. Their examples are partially hyperbolic with a dominated splitting of the tangent bundle $TM = E^s \oplus E^c \oplus E^u$ into 3 non-trivial subbundles, the central subbundle E^c being one-dimensional and non-hyperbolic. Much later, [56] and [67] gave new classes of robustly transitive diffeomorphisms, on much more general manifolds and also satisfying weaker hyperbolicity conditions.

Let us list the three classes of known examples, which we are going to discuss in detail in the next sections:

(A) Perturbations of the time-1 of any transitive Anosov flow.

These examples are isotopic to identity, and they exist in many manifolds, other than the tori, already in dimension 3. It is not known which 3-manifolds do support such diffeomorphisms. See Section 7.3.5

(B) Skew-products : $T^n \times N \to T^n \times N$, $(x,y) \mapsto (A(x), B(x,y))$ and their perturbations, where A is a transitive Anosov diffeomorphism, N is any compact manifold, and the behavior in the y direction is dominated by that of A along its stable and unstable directions.

In these examples the non-hyperbolic central direction may be chosen tangent to the fibers $\{x\} \times N$ and so $\dim E^c = \dim N$, providing examples of robustly transitive diffeomorphism having a central direction of any dimension.

(C) Robustly transitive diffeomorphisms derived from Anosov, that is obtained by deforming Anosov torus diffeomorphisms by isotopy. They may be constructed

- (in dim ≥ 3) without any stable subbundle, that is, $TM = E^u \oplus E^{cs}$, where E^{cs} is indecomposable and non-hyperbolic. Analogously, without any unstable subbundle.
- (in dim ≥ 4) without any hyperbolic (stable or unstable) subbundle, that is, $TM = E^{cu} \oplus E^{cs}$ is a dominated splitting where E^{cs} and E^{cu} are indecomposable and non-hyperbolic.

These two last constructions correspond to the weakest possible hyperbolicity, by Theorem 7.2.

7.3.1. *An example of Shub.* Consider Anosov diffeomorphisms A and B defined on the torus T^2 such that the product map $(A,B)\colon T^2 \times T^2 \to T^2 \times T^2$ is partially hyperbolic with central bundle E^c tangent to the fibers $\{x\} \times T^2$ (for that it is enough to choose $A = B^n$ for any $n > 1$). Assume that A has fixed points p and q and B has a fixed point r.

Consider a map $f\colon T^2 \times T^2 \to T^2 \times T^2$ of the form $f(x,y) = (A(x), F_x(y))$, partially hyperbolic with E^c tangent to the fibers $\{x\} \times T^2$, and also that $F_p = B$.

Proposition 7.11. *The diffeomorphism f is robustly transitive.*

Proof. As the central foliation $\mathcal{F}^c$ (tangent to E^c) is differentiable, Theorem B.10 implies that the pair $(\mathcal{F}^c, f)$ is structurally stable. It follows that the stable manifold and the unstable manifold of the continuation of the fiber $\{p\} \times T^2$ are both dense in T^4 for every C^1 perturbation of f. One deduces that the invariant manifolds of the fixed point (p, r) are both robustly dense in T^4. This implies that f is robustly transitive: the stable manifold of the fixed point intersect any open set U of T^4, thus the iterates $f^n(U)$, $n > 0$ accumulate arbitrarily large disks in the unstable manifold, which, on their turn, intersect any open set V. That is, $f^n(U) \cap V \neq \emptyset$. □

To ensure that these examples are non-hyperbolic, it is enough to take $F_q : T^2 \to T^2$ having periodic points with different Morse indices: then f also has points with different Morse indices, and transitivity implies that f can not be hyperbolic.

7.3.2. *An example of Mañé.* Consider a linear Anosov map $A: T^3 \to T^3$ having three real eigenvalues, $0 < \lambda_1 < 1 < \lambda_2$ and $3 < \lambda_3$. Denote by $\mathcal{F}^c$ the linear foliation corresponding to the eigenvalue λ_2. Fix small domains $C' \subset C$ with C' much smaller than C. Consider a partially hyperbolic perturbation f of A preserving the foliation $\mathcal{F}^c$, having it as its central foliation, and coinciding with A outside C'.

Assume that the expansion of f in the strong-unstable direction remains bigger than 3. If C is small enough, there is a constant $L > 0$ such that, for every unstable segment σ of length$(\sigma) \geq L$, the set $(f(\sigma) \setminus C)$ contains an unstable segment of length bigger than L. On the other hand, given any unstable segment σ the lengths of its forward iterates $f^n(\sigma)$ are eventually bigger than L. This shows that every unstable segment contains some point x_0 whose forward orbit intersects C finitely many times only.

We claim that, given any small central segment α through such a point x_0, the lengths of its forward iterates increase exponentially. This is obvious while the iterates are shorter than the distance from C' to the complement of C, just because $f^n(\alpha)$ remains outside C' (discard finitely many iterates if necessary), where the central direction is actually expanding. When the segments are long just observe that the fraction of $f^n(\alpha)$ inside C' is very small: expansion of the big sub-segments outside C' suffices for expansion of the total length.

Now, given any open set $U \subset M$ consider an unstable segment $\sigma \subset U$, a point $x_0 \in \sigma$ visiting C finitely many times only, and a central segment $\alpha \subset U$ through x_0. By Theorem B.10 each leaf of $\mathcal{F}^c$ is dense in T^3, in a robust fashion. So, the fact that the length of $f^n(\alpha)$ goes to infinity implies that the orbit of α is dense, and so the same is true for the orbit of U. This proves (robust) transitivity of f.

In order to guarantee that this construction yields non-hyperbolic maps, Mañé considered a fixed point p of A and performed a saddle-node bifurcation along a central leaf $\mathcal{F}^c$ inside a small neighborhood C' of p. In this way, f has periodic points with different Morse indices 2 and 1.

7.3.3. *A local criterium for robust transitivity.* In this section we explain the method used in [56] and [112] for proving robust transitivity. Let us start with the following simple criterium, which we used in Shub's example.

Lemma 7.12. *(Basic transitivity criterium) Suppose f has a hyperbolic periodic point p whose invariant manifolds are both (robustly) dense in M. Then f is (robustly) transitive.*

For simplicity, in what follows we restrict ourselves to 3-dimensional manifolds. See Section 6.2 for details and terminology concerning blenders. In the next statement note that $W^s(p)$ and $W^u(q)$ are two-dimensional.

Proposition 7.13. *(Blender criterium) Assume that*

(1) *f is partially hyperbolic on the whole manifold and there is a dominated splitting $TM = E^u \oplus E^c \oplus E^s$ into three non-trivial subbundles such that E^u is uniformly expanding and E^s is uniformly contracting.*
(2) *There are a constant $L > 0$ and hyperbolic periodic points p and q with Morse indices 1 and 2, respectively, such that the stable manifold $W^s(p)$ of p intersects every unstable segment of length bigger than L, and the unstable manifold $W^u(q)$ of q intersects every stable segment of length bigger than L.*
(3) *The point q is homoclinically related to a center unstable blender Γ and the one-dimensional unstable manifold $W^u(p)$ of p intersects every superposition region of Γ.*

Then f is robustly transitive.

Proof. From assumption 2 it follows that $W^s(p)$ intersects all unstable segments, since the length of forward iterates of unstable segments goes to infinity, and the stable manifold is invariant. This implies that the stable manifold of p is dense in M.

By compactness of the family of unstable segments of $2L$, and the fact that the unstable foliation and the stable manifold of p depend continuously on the dynamics, one easily gets that $W^s(p, g)$ intersects every unstable segment of length $2L$ for all g close to f. As before, this implies the density of $W^s(p, g)$. In other words, the stable manifold of p is robustly dense in M. The same arguments give robust density of the unstable manifold of q.

Using the principle of superposition of the blenders from Section 6.2, we deduce that the closure of the stable manifold of q robustly contains the stable manifold of p, thus it is robustly dense in M. At this point we know that both invariant manifolds of q are robustly dense in M, and so the basic criterium in Lemma 7.12 implies that f is robustly transitive. $\square$

The blender criterium is used in [56] to prove that both

- the time-1 map of a transitive Anosov flow and
- the product of an Anosov map by the identity on a compact manifold N

are approximated by robustly transitive diffeomorphisms. For the first situation one perturbs the dynamics along a closed orbit of the flow to get a pair of periodic orbits with different Morse index. In the second situation, one uses a chain of blenders to find a periodic point of index u (the dimension of the unstable manifold of the Anosov map) such that the closure of its unstable manifold contains the unstable manifold of a point of index $\dim N + u$.

7.3.4. *Robust transitivity without hyperbolic directions.* All robustly transitive diffeomorphisms we described so far admit hyperbolic directions, that is, either uniformly contracting or uniformly expanding. Here we describe the construction given in [67] of robustly transitive diffeomorphisms without any hyperbolic directions. This construction follows closely the one in Section 7.3.2, so we just emphasize the main differences.

Consider an linear Anosov map A of the torus T^4 having 4 real eigenvalues, $0 < \lambda_1 < \lambda_2 < 1 < \lambda_3 < \lambda_4$. Fix small domains $C_1' \subset C_1$ and $C_2' \subset C_2$, with C_1 and C_2 disjoint. Fix an A-invariant cone field $\mathcal{C}^{cu}$ and an A^{-1}-invariant cone field $\mathcal{C}^{cs}$, around the planes corresponding to the expanding eigenvalues λ_3, λ_4 and the contracting eigenvalues λ_1, λ_2, respectively.

Now consider diffeomorphisms f coinciding with A outside $C_1 \cup C_2$ and satisfying

(1) Df preserves $\mathcal{C}^{cu}$ and expands area uniformly in this cone field.
(2) Df^{-1} preserves $\mathcal{C}^{cs}$ and expands area uniformly in this cone field.
(3) Df expands uniformly the norm of vectors in $\mathcal{C}^u$ outside C_1'.
(4) Df^{-1} expands uniformly the norm of vectors in $\mathcal{C}^s$ outside C_2'.

By choosing sufficiently thin cone fields, we verify that there exists a constant $L > 0$ such that every center unstable disk D_1 (tangent to $\mathcal{C}^{cu}$) of radius bigger than L intersects any center stable disk D_2 (tangent to $\mathcal{C}^{cs}$) of radius larger than L.

Arguing as in Mañé's example, but using uniform expansion of area instead of uniform expansion of norms, we show that every center unstable disk D^1 contains a point whose forward orbit meets C_1 only finitely many times. This allows us to show that $f^n(D_1)$ contains a center unstable disk of radius bigger than L, for every large $n > 0$. The same argument shows that the large negative iterates of any center stable disk D_2 contains a center stable disk of radius bigger that L.

Finally, given any pair of open sets U and V, we may choose a center unstable disk $D_1 \subset U$ and a center stable disk $D_2 \subset V$. Thus, forward iterates of U and backward iterates of V intersect eventually: $f^n(U) \cap f^{-m}(V)$ for all large n and m. This implies that f is transitive.

In order to ensure that f does not preserve any hyperbolic subbundle, we choose the diffeomorphism such that C_1 contains a fixed point p_1 of f of index 2 with non-real contracting eigenvalues together with a fixed point q_1 of index 3. This prevents the existence of uniformly contracting subbundles: the presence of p_1 implies that a stable subbundle would have to be of dimension at least 2, while q_1 implies that the dimension is at most 1. We also choose f so that C_2 contains a fixed point p_2 of index 2 with non-real expanding eigenvalues together with a fixed point q_2 of index 1. Arguing as before, this prevents the existence of any uniformly expanding subbundle.

7.3.5. *Manifolds supporting robustly transitive maps.* It is well-known that the existence of globally hyperbolic (Anosov) diffeomorphisms imposes very strong restrictions on the manifold; see [318]. In contrast, the corresponding question is wide open in the partially hyperbolic setting:

Problem 7.14. Which manifolds support partially hyperbolic diffeomorphisms, respectively diffeomorphisms admitting dominated splittings over the ambient space ?

One obvious topological condition is that the tangent bundle should admit a decomposition into continuous sub-bundles. In the case of 3-dimensional manifolds this constitutes no obstruction at all, as the tangent bundle is always trivial. In fact, Mitsumatsu [221] proves that every 3-manifold supports a diffeomorphism with a dominated splitting into three sub-bundles. However, based on the fact that robustly transitive systems are partially hyperbolic, we expect that

Conjecture 7.15. There are no robustly transitive maps on the sphere S^3.

A proof of the second part of the conjecture was given in [108], in the case of diffeomorphisms admitting a center stable or a center unstable foliation. The proof is based on Novikov's theorem [85]: such a foliation must have a compact leaf, which would separate the ambient space and, thus, would be an obstruction to transitivity.

In this direction, Brin, Burago, Ivanov [76] recently announced

Theorem 7.16. *For 3-dimensional manifolds M with Abelian fundamental group, the action of any partially hyperbolic diffeomorphism with three invariant sub-bundles on the first homology $H_1(M)$ must have some eigenvalue strictly larger than* 1.

Consequently, *there are no partially hyperbolic diffeomorphism with three invariant sub-bundles $E^u \oplus E^c \oplus E^s$ on the sphere S^3.* Note that some hypothesis on the manifold is necessary in the theorem, since the time-1 map of the geodesic flow on a hyperbolic surface is partially hyperbolic yet acts trivially on the homology.

The crucial point in the proof of the theorem is to construct a center unstable foliation W^{cu} with C^1 leaves whose tangents are close to $E^u \oplus E^c$ at every point. Roughly speaking, they do this by considering the union of curves tangent to E^c passing through each unstable manifold. Let $\tilde{f}$, $\tilde{W}^u$, $\tilde{W}^{cu}$ be lifts of f, W^u, W^{cu} to the universal cover $\tilde{M}$. Consider a small segment γ inside a leaf of $\tilde{W}^u$, and a small domain Γ around it. The argument proceeds by contradiction.

Assuming that all eigenvalues of the action of f on $H_1(M)$ are less or equal than 1, one deduces that the volume of $\tilde{f}^n(\Gamma)$ grows sub-exponentially, at most, whereas the length of the iterates of γ grows exponentially fast. Does $\tilde{f}^n(\gamma)$ must be rather curly for large n. One deduces that some leaf of $\tilde{W}^u$ accumulates in itself. Then, of course, so does the leaf of $\tilde{W}^{cu}$ that contains it. Combining this with Novikov's theorem, one constructs an $\tilde{f}$-invariant torus tangent to $E^u \oplus E^c$ and bounding a solid torus in $\tilde{M}$. This leads to a contradiction, because of the following fact: Any diffeomorphism of T^2 that admits a uniquely integrable uniformly expanding direction has non-trivial action on the homology $H_1(T^2)$ and, consequently, can not be extended to the solid torus.

Another very interesting set of results has been announced by Bennequin and Herman [47, 154] concerning manifolds that support partially hyperbolic *symplectic* diffeomorphisms. In the symplectic category, partial hyperbolicity is equivalent to existence of a dominated splitting. Moreover, the stable and unstable sub-bundles have the same dimension and they are symplectic conjugate. See [50] for a proof. Corresponding to this, Bennequin proves that the Chern class of M must admit a factorization $c(M) = \alpha\beta$ where β involves only even terms. Then he checks that this condition excludes several manifolds, such as the complex projective spaces $\mathbb{CP}^n$ and their quotients by degree 2 equations (quadric manifolds). An example of quadric manifold is $S^2 \times S^2$.

7.4. **Invariant foliations.** Throughout this section $f : M \to M$ is a diffeomorphism admitting a partially hyperbolic invariant splitting $TM = E^u \oplus E^c \oplus E^s$ into three sub-bundles, all of them with positive dimension.

7.4.1. *Pathological central foliations.* Basic facts about invariant foliations of partially hyperbolic maps are recalled in Appendix B. In particular, the strong stable and the strong unstable foliations of C^2 diffeomorphisms are absolutely continuous with respect to Lebesgue measure [78]: the holonomy maps between any two transverse sections send zero Lebesgue measure sets to zero Lebesgue measure sets.

That is not true for the central foliation, in general:

Theorem 7.17 (Shub, Wilkinson [313]). *There is an open subset of C^1 volume preserving partially hyperbolic diffeomorphisms of the 3-torus admitting a unique central foliation which, nevertheless, is not absolutely continuous.*

Even more, Ruelle, Wilkinson [304] show that for these maps the conditional probabilities of the Lebesgue measure along central leaves are supported on a finite number of points, almost everywhere.

We prove the theorem using a variation of the construction in [313]. Start from the product $f = h \times \mathrm{id}$ of a linear area preserving Anosov diffeomorphism h in T^2 by the identity on the circle T^1. Let e_s, e_c, e_u be eigenvectors of Df corresponding to the eigenvalues $\lambda < 1 < \lambda^{-1}$, respectively. Now deform f by composing it with a diffeomorphism ϕ of $T^3 = T^2 \times T^1$, of the form

$$\phi(\xi, z) = \big(\xi + \psi(z)(e_s + \alpha e_c), z\big)$$

for some $\alpha > 0$ and $\psi : T^1 \to \mathbb{R}$ close to zero in the C^1 topology. Observe that ϕ preserves volume and the center unstable foliation of f. Since the central foliation is structurally stable, by Theorem B.10, the map $g = f \circ \phi$ (and any other map C^1 close to it) has a central foliation which is a topological fibration by circles.

Lemma 7.18. *If ψ is not constant then the integrated Lyapunov exponent*

$$\lambda^c(g) = \int_{T^3} \log |Dg \mid E^c_g| \, d\mathrm{Leb}$$

of $g = f \circ \phi$ along the corresponding central direction E^c_g is strictly negative.

In this setting the central Lyapunov exponent varies continuously with the map, because the central sub-bundle E^c does. In particular, λ^c remains negative for every map in a C^1 neighborhood of g. The following argument applies to any such map.

Firstly, the central Lyapunov exponent must be strictly negative on a subset E of T^3 with positive Lebesgue measure. Now, all central leaves are compact and so their forward and backward iterates have uniformly bounded length. Therefore, each leaf may intersect E only on a zero arc-length subset. However, if the central foliation were absolutely continuous then E would have to intersect many central leaves on subsets with positive arc-length. This shows that the central foliation is not absolutely continuous.

To complete the proof of Theorem 7.17, we only have to give the

Proof of Lemma 7.18. The matrices of Df, $D\phi$ in the basis $\{e_s + \alpha e_c\,, e_c\,, e_u\}$ are

$$Df = \begin{pmatrix} \lambda & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad D\phi = \begin{pmatrix} 1 & \psi' & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $a = \alpha(1 - \lambda) > 0$. Since the perturbation preserves the original center stable bundle, we may represent the central direction E^c_g by $(\omega, 1, 0)$ with $\omega : T^3 \to \mathbb{R}$

uniformly close to zero. By invariance, $Dg \cdot (\omega, 1, 0) = J^c g\,(\omega \circ g, 1, 0)$ where $J^c g$ takes values in $\mathbb{R}$. This translates into

$$\omega \circ g = \frac{\lambda(\omega + \psi')}{1 + a(\omega + \psi')} \quad \text{and} \quad J^c g = 1 + a(\omega + \psi'). \tag{26}$$

Now we use the elementary fact that, for $t > -1/a$,

$$\frac{\lambda t}{1 + at} \leq \lambda t \quad \text{and equality holds only if } t = 0.$$

This gives that $\omega \circ g \leq \lambda(\omega + \psi')$, and equality holds only where $\omega + \psi'$ vanishes. Integrating the inequality over the 3-torus and using $\int \psi' \, d\mathrm{Leb} = 0$ and invariance of Lebesgue measure,

$$\int (\omega + \psi') \, d\mathrm{Leb} = \int (\omega \circ g) \, d\mathrm{Leb} \leq \lambda \int (\omega + \psi') \, d\mathrm{Leb}$$

and equality holds only if $\omega + \psi'$ is identically zero. In that case, using (26), both ω and ψ' would have to vanish identically. Thus, the assumption that ψ is not constant ensures a strict inequality:

$$\int (\omega + \psi') \, d\mathrm{Leb} < \lambda \int (\omega + \psi') \, d\mathrm{Leb} \quad \text{and so} \quad \int (\omega + \psi') \, d\mathrm{Leb} < 0,$$

since $\lambda < 1$. Finally, $J^c g = |Dg \mid E^c_g|$ for appropriate choice of the Riemannian metric (the Lyapunov exponent is independent of this choice). Then

$$\lambda^c(g) = \int \log\big(1 + a(\omega + \psi')\big) \, d\mathrm{Leb} \leq a \int (\omega + \psi') \, d\mathrm{Leb} < 0,$$

as claimed. □

Applying a similar argument to the inverse, one proves that f is C^1 approximated by maps with central Lyapunov exponent strictly positive.

7.4.2. *Density of accessibility.* Due to the absolute continuity of the strong stable and the strong unstable foliations, one may propagate ergodic properties along their leaves. Thereafter, the topological features of these foliations are an important tool in the ergodic theory of partially hyperbolic systems, as we shall discuss in more detail in Section 8.2.

With this in mind we consider the following notion, first introduced in Dynamics by Brin, Pesin [78], and much developed in recent years, starting from the work of Grayson, Pugh, Shub [142]:

Definition 7.19. A partially hyperbolic diffeomorphism $f : M \to M$ has the *accessibility property* if for any pair of points $x, y \in M$ there is a path joining x to y composed by finitely many segments inside strong stable or strong unstable leaves (a *us-path*). More generally, f has the *essential accessibility property* if every set saturated by us-paths has either zero or full Lebesgue measure.

An important recent result of Dolgopyat, Wilkinson asserts that accessibility holds on an open and dense subset:

Theorem 7.20 (Dolgopyat, Wilkinson [122])**.** *There is an open and dense subset of C^1 partially hyperbolic diffeomorphisms on any compact manifold having the accessibility property. Moreover, this set intersects the spaces of volume preserving, or symplectic, partially hyperbolic diffeomorphisms on open and dense subsets.*

Combined with an older result of Brin [74] this implies that among volume preserving partially hyperbolic diffeomorphisms the transitive ones contain a C^1 open and dense subset, and similarly in the symplectic case. See Section 8.5.

The proof of Theorem 7.20 has two main steps. Firstly, one proves *stable (robust) accessibility modulo a finite family of central disks.* That is, one chooses pairwise disjoint disks Δ_i, $i = 1, \dots, k$ transverse to the sum $E^u \oplus E^s$ and such that

- given any points x and y in M there exists a path joining x to y and consisting of finitely many segments contained in strong stable leaves, strong unstable leaves, or the interior of some of the disks Δ_i ;
- there exists some large $N \geq 1$ such that the iterates $f^j(\cup_i \Lambda_i)$, $j = 0, \dots, N$ are pairwise disjoint.

Notice that this property is robust under small C^1 perturbations. Secondly, one carries out local perturbations close to each of the disks Δ_i in order to obtain that

- given any points x and y in Δ_i there exists a us-path joining x to y.

Combining the two steps one gets that any pair points in M may be joined by a us-path, as claimed.

Let us outline the construction in this second step, in the case when the central sub-bundle is one-dimensional. See Figure 14.

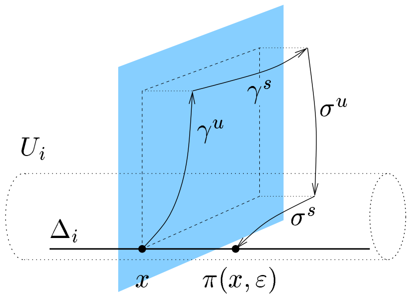

FIGURE 14. Local perturbation to create accessibility

For each disk Δ_i one considers a continuous choice of two strong stable and two strong unstable arcs $\gamma^u(x,\varepsilon)$, $\gamma^s(x,\varepsilon)$, $\sigma^u(x,\varepsilon)$, $\sigma^s(x,\varepsilon)$, for $x \in \Lambda_i$ and $0 < \varepsilon \leq \bar{\varepsilon}$, all with length $\approx \varepsilon$ and such that the concatenation

$$\gamma^u(x,\varepsilon) \cdot \gamma^s(x,\varepsilon) \cdot \sigma^u(x,\varepsilon) \cdot \sigma^s(x,\varepsilon)$$

links x to a point $\pi(x,\varepsilon) \in \Lambda_i$. Next one considers a perturbation $g = h \circ f$ of the diffeomorphism f, where h is supported on a small box U_i close to Δ_i, in such a way as to change $\gamma^u(x,\bar{\varepsilon})$, uniformly over the whole Δ_i while keeping the other arcs essentially unchanged. To show that this is possible one uses the following observation: Due to the fact that the return time to U_i is very large (the choice of N in the first step), after the perturbation the strong unstable and strong stable sub-bundles on U_i are given by

$$E^u_g \approx Dh(E^u_f) \quad \text{and} \quad E^s_g \approx E^s_f.$$

In addition, we take U_i so that it avoids the other unstable segments $\sigma^u(x,\bar{\varepsilon})$. In this way one gets that, after perturbation,

$$\pi(x,\bar{\varepsilon}) \approx x + \bar{\varepsilon}\partial_i \quad \text{for every } x \in \Delta_i$$

where ∂_i is the tangent direction to Δ_i. By considering $\pi(x,\varepsilon)$, for the perturbed diffeomorphism g and for every $0<\varepsilon\leq\bar\varepsilon$ we see that x may be joined to any points between x and $x+\bar\varepsilon\partial_i$ using a us-path formed by four arcs only. Since this holds uniformly for all x in the disk, we get that any two points in Δ_i are joined by a us-path, as claimed.

7.4.3. *Minimality of the strong invariant foliations.* A foliation $\mathcal{F}$ is *minimal* if every leaf is dense in the ambient manifold. The first result about minimality of strong invariant foliations applied to the geodesic flow of a compact hyperbolic surface: This is an Anosov flow, and its strong stable and strong unstable foliations are the horocyclic foliations. Furstenberg [134] proved that the horocyclic flow is uniquely ergodic.

More generally, we have

Theorem 7.21 (Plante [275]). *Let X be a transitive Anosov flow defined on a compact 3-manifold and $\mathcal{F}^u$ be its strong unstable foliation.*

(1) *If there is a leaf of $\mathcal{F}^u$ which is not dense, then X is conjugate to the suspension of an Anosov diffeomorphism of the torus T^2.*
(2) *If there is a leaf of $\mathcal{F}^u$ which is dense then the unstable foliation is minimal.*

Minimality and unique ergodicity of the strong foliations were crucial tools to understand the ergodic properties of Anosov flows. Motivated by results such as Theorem 11.26 below, we expect the strong invariant foliations to play a similarly important role in the general partially hyperbolic setting.

So, here we are going to discuss some sufficient conditions for minimality of the strong foliations of diffeomorphisms. For simplicity, we always restrict ourselves to diffeomorphisms on 3-dimensional manifolds, partially hyperbolic with 3 invariant sub-bundles E^u, E^c, E^s.

Theorem 7.22 (Bonatti, Díaz, Ures [61]). *There is an open and dense subset of C^1 robustly transitive partially hyperbolic diffeomorphisms on any 3-dimensional manifold for which either the strong stable foliation or the strong unstable foliation is robustly minimal.*

The main step in the proof of this theorem is the construction of a *dynamical complete section* for the strong unstable (or the strong stable) foliation $\mathcal{F}^u$ (or $\mathcal{F}^s$). By such we mean a compact, possibly non-connected, surface S with boundary which is transverse to $\mathcal{F}^u$ and satisfies:

- $f(S)$ is contained in the interior of S,
- the restriction of f to S is a Morse-Smale diffeomorphism,
- every strong unstable leaf intersects transversely the interior of S.

Let us assume, for a moment, that a dynamical complete section S for $\mathcal{F}^u$ does exist. A simple argument, using the Morse-Smale condition and the λ-lemma, gives that every strong unstable leaf accumulates the leaf through some periodic point in S. So, it suffices to prove that the latter leaves are dense in the ambient space. Moreover, by a simple perturbation, it is enough to consider the leaves through the periodic points in S which are sinks for f restricted to S, that is, which have Morse index 1 in M. For these periodic points, the unstable manifold coincides with the strong unstable leaf. Hence, we are reduced to proving that the unstable manifolds of the periodic points of index 1 in S are dense in the ambient manifold. Now, a blender argument similar to the one in Section 7.3.3 implies that, for an open dense subset of C^1 partially hyperbolic diffeomorphisms, the closure of the unstable manifolds of any such periodic point contains the 2-dimensional unstable of some periodic orbit of index 2. Such a 2-dimensional invariant manifold is always dense, just because it cuts the local stable leaf of some point with dense forward orbit. This proves that $\mathcal{F}^u$ is minimal, whenever f is in the intersection with that open dense subset.

Let us now explain how to construct a dynamical complete section for $\mathcal{F}^u$ of $\mathcal{F}^s$. It is clear that existence of such a section is an open property, so the real issue is density. So, we may suppose that all periodic points are hyperbolic. For simplicity, we consider only the case when the central bundle is uniquely integrable and orientable and f preserves this orientation. To each periodic saddle p of the diffeomorphism f we associate the half central leaf $L(p)$ to the right of p. This contains a finite number of periodic points, since they all have the same period as p. If $L(p)$ is not a circle, we say that p is of *type* 1 (respectively *type* 2) if the last periodic point p_* in $L(p)$ has Morse index 1 (respectively 2). If $L(p)$ is a circle then p is both of type 1 and type 2, and we set $p_* = p$.

Let Σ_1 be the set of type 1 and Σ_2 be the set of type 2 periodic points. Robust transitivity implies that, generically, periodic points are dense in M. It follows that the union of the two sets is dense in M, and so one of them must be dense in some open subset. Using transitivity once more, we may assume that at least of the two sets is dense in M. Let us suppose Σ_1 is dense: in this case we get a section for $\mathcal{F}^u$, allowing us to prove that the strong unstable foliation is minimal. A similar argument applies if Σ_2 is dense just replacing the diffeomorphism by its inverse: in this case we get a section for $\mathcal{F}^s$ and conclude that the strong stable foliation is minimal. In particular, if Σ_1 and Σ_2 are simultaneously dense then both strong foliations are minimal.

For every $p \in \Sigma_1$ there is a segment I_p contained in $L(p)$, with length bigger than some uniform constant and such that $f^n(I_p) \subset \operatorname{int}(I_p)$ for some positive n: if $L(p)$ is a circle, just take $I_p = L_p$; otherwise let I_p be a segment starting close to p (to the left if p has index 1 and to the right if p has index 2) and extending long enough to the right of p_*. We construct a dynamical complete section of $\mathcal{F}^u$ by saturating finitely many (orbits of) segments I_p by local strong stable leaves. Since Σ_1 is dense, we may take sufficiently many such segments to ensure that this section does cut every strong unstable leaf.

For generic robustly transitive diffeomorphisms there is a unique homoclinic class, which coincides with the whole manifold. This is as immediate consequence of Theorem 10.20 below although, historically, it was first obtained in [57] as a consequence of Hayashi's connecting lemma [150]. As an application of Theorem 7.22, we can improve this result in dimension 3:

Proposition 7.23. *There is a C^1 open and dense subset of robustly transitive partially hyperbolic diffeomorphisms on any 3-dimensional manifold admitting some periodic point whose homoclinic class is the whole ambient manifold.*

To deduce this from Theorem 7.22, suppose for instance that the strong unstable foliation is robustly minimal. Then the stable and unstable manifolds of any hyperbolic periodic point p with Morse index 1 are both dense in M. This is immediate, by definition of minimality, for the unstable manifold. For the stable manifold it follows by observing that it intersects each sufficiently large strong unstable segment and so it is dense in the whole M; recall the arguments in Section 7.3.3. The density of the two invariant manifolds of p, together with partial hyperbolicity, imply that the homoclinic class of p is the whole manifold.

There are extensions of the previous results to higher dimensions, under the extra assumption that the central sub-bundle is robustly uniquely integrable [61]. Since it is known whether this property is dense, the following question is still open:

Problem 7.24. Let f be a robustly transitive partially hyperbolic diffeomorphism. Suppose that p is a periodic saddle of f whose continuation p_g is defined for every g in an open neighborhood $\mathcal{U}$ of f. Is $H(p_g, g) = M$ for all g in an open dense subset of $\mathcal{U}$?

A partial result is obtained in [59]: given any two hyperbolic periodic points p and q with different indices, there is an open set $\mathcal{U}$ of such that $H(p_g, g) = H(q_g, g)$ for all $g \in \mathcal{U}$.

7.4.4. *Compact central leaves.* The proof of the previous theorem leads to some natural questions about the dynamics in the central foliation. We continue to consider diffeomorphisms on 3-dimensional manifolds, partially hyperbolic with three invariant sub-bundles. Moreover, we assume that the periodic points are all hyperbolic. For simplicity, we consider only the case when the central sub-bundle is orientable and uniquely integrable, and its orientation is preserved by the diffeomorphism.

We say that a periodic central leaf L is of *stable type* if the first and the last periodic points in L have Morse index 1, and of *unstable type* if the first and the last periodic points of L have Morse index 2. The leaf L is of *saddle-node type* if the first saddle has index 1 and the last one has index 2, or vice-versa, and of *circle type* if it is compact.

The examples of robustly transitive diffeomorphisms listed in Section 7.3 follow into three classes, in what concerns their central leaves:

(1) all the central leaves are compact (perturbations of dominated products of Anosov diffeomorphism by other diffeomorphisms),
(2) there are both compact and non-compact central leaves, but those with periodic points are compact (perturbations of time-1 maps of transitive Anosov flows)
(3) all central leaves are non-compact and all are of the same (stable or unstable) type (diffeomorphisms derived from Anosov).

Problem 7.25. Are there partially hyperbolic transitive diffeomorphisms having

- simultaneously non-compact central leaves of stable and of unstable type ?
- non-compact central leaves of saddle-node type ?
- simultaneously non-compact and compact central leaves containing hyperbolic periodic points?

In a sense, all known robustly transitive diffeomorphisms are obtained by some sort of perturbation of hyperbolic systems. Examples as above, if they exist, should require essentially different constructions. For instance, we believe transitive diffeomorphisms with central leaves of both stable and unstable type can not be obtained via perturbations of Anosov diffeomorphisms, or of skew-products based on Anosov diffeomorphisms. In this direction we have

Theorem 7.26 (Bonatti, Wilkinson [68]). *Let f be a transitive partially hyperbolic on a compact 3-dimensional manifold M. Assume there exists a periodic circle γ such that the intersection $\mathcal{F}^u(\gamma) \cap \mathcal{F}^s(\gamma)$ has some compact connected component. Then, f leaves invariant a Seifert circle bundle* [9]. *Consequently, the lift of some iterate of f to some finite covering of M is a dominated skew-product over an Anosov diffeomorphism of T^2.*

Under some extra assumptions on the central foliation, Theorem 7.22 can be improved. A partially hyperbolic diffeomorphism f is *dynamically coherent* if its central foliation is uniquely integrable and it is the intersection of the center stable and center unstable foliations. We say that f is *robustly (stably) dynamically coherent* if any C^1 perturbation of it is dynamically coherent.

Theorem 7.27 (Bonatti, Díaz, Ures [61]). *Let $\mathcal{D}$ be the subset of partially hyperbolic C^1 diffeomorphisms which are robustly dynamically coherent and have a periodic compact central leaf. There is an open and dense subset of $\mathcal{D}$ for which the strong stable and the strong unstable foliations are both minimal.*

In particular, we may apply Theorem 7.27 to perturbations of time-1 maps of transitive Anosov flows, and to perturbations of dominated skew products over Anosov diffeomorphisms, to conclude that their strong stable and the strong unstable foliations are both minimal. As for the other construction in Section 7.3, partially hyperbolic diffeomorphisms derived from Anosov, including Mañé's example in Section 7.3.2, we may apply Theorem 7.22 to conclude that, for a dense open subset, the strong stable foliation is minimal.

Problem 7.28. In Mañé's example is the strong unstable foliation also minimal (for a residual, or even open and dense subset) ?

A crucial question, having in mind possible applications of these results to the study of ergodic properties:

Problem 7.29. Does robust minimality of a strong invariant foliation imply unique ergodicity of the foliation (at least for a residual subset) ?

Another major challenge is to fill the gap between the C^1 and C^2 categories: while the generic statements above are proved for the C^1 topology, control of the ergodic behavior usually requires C^2 regularity at least.

8. Stable ergodicity

A classical argument for proving ergodicity of conservative systems, going back to Hopf [160] and extended in a fundamental way by Anosov [11], may be summarized as follows. By Birkhoff's ergodic theorem, forward and time averages are well-defined and they coincide for every point on a full measure subset E. Clearly, forward time averages are constant on stable sets, and similarly for backward time averages and unstable sets. Therefore, to prove ergodicity it suffices to show that almost any pair of points in E may be joined by a path obtained by concatenation of finitely many arcs contained in stable sets and unstable sets, and *such that all endpoints are in E.*

This approach allowed Anosov [11] to prove that the geodesic flow of any compact Riemannian manifold with negative curvature is ergodic and, more generally, the

[9]Foliation of M by circles, which is trivial except for finitely many exceptional leaves.

same is true for any C^2 uniformly hyperbolic (Anosov) flow or diffeomorphism. The case of hyperbolic surfaces, as well as manifolds with constant negative curvature, had been dealt with by Hedlund and Hopf. This provided the first open sets of volume preserving systems which are ergodic: one speaks of stable ergodicity. Quite in contrast, KAM (Kolmogorov, Arnold, Moser) theory shows that *non-ergodicity* is very common among conservative systems: there often exist positive measure sets consisting of invariant tori; see for instance [347].

The material in this section is motivated by the fundamental problem of deciding which type of behavior is predominant. Recent conjectures of Pugh, Shub assert that ergodicity should prevail among partially hyperbolic systems: it should hold in an open dense subset. A detailed survey on the subject of stable ergodicity appeared recently, by Burns, Pugh, Shub, Wilkinson [80]. We only recall some main ideas and results of the theory, especially the most recent ones.

A diffeomorphism f is called *stably ergodic* if it is in the C^2 interior of the set of all volume preserving ergodic diffeomorphisms [10]. There is a similar notion for volume preserving vector fields.

8.1. **Examples of stably ergodic systems.** The ergodicity problem was treated in the pioneer work of Brin, Pesin [78] in the context of partially hyperbolic systems. They needed very restrictive regularity assumptions on invariant foliations, and so could not exhibit new stable examples. In fact, the first non-hyperbolic stably ergodic systems were found only a few years ago, by Grayson Pugh, Shub [142]. The examples currently available parallel those of robustly transitive systems, reviewed in Section 7.3. The reasons for this are not well understood.

8.1.1. *Perturbations of time-*1 *maps of geodesic flows.* Curiously, the first non-hyperbolic examples were found in a setting very close to the classical one:

Theorem 8.1 (Grayson Pugh, Shub [142]). *The time-*1 *map of the geodesic flow on the unit tangent bundle of a surface with constant negative curvature is stably ergodic.*

Theorem 8.1 was extended to surfaces with variable curvature by Wilkinson [343], and to compact manifolds with almost constant negative curvature in the work of Pugh, Shub [282]. In fact, the geodesic flow of any compact manifold with negative curvature and arbitrary dimension is stably ergodic, as we are going to see in Section 8.3.

8.1.2. *Perturbations of skew-products.* Stable ergodicity of compact Lie group extensions $F : M \times G \to M \times G$, $F_\varphi(x, g) = (f(x), \varphi(x) \cdot g)$ over Anosov diffeomorphisms $f : M \to M$ has been investigated by several authors. We just quote

Theorem 8.2 (Burns, Wilkinson [82]). *Fix $k \in (2, \infty]$. Let $f : M \to M$ be a volume preserving Anosov diffeomorphism of class $C^{k-\varepsilon}$ on a compact manifold, and let G be a compact connected Lie group. Then there exists a dense subset $\mathcal{E}$ of the space of $C^{k-\varepsilon}$ functions $\varphi : M \to G$ such that the corresponding skew-product F_φ is stably ergodic.*

[10] Sometimes a stronger notion is used: every C^2 volume preserving in a C^1 neighborhood of f is ergodic.

Theorem B in [82] also provides an explicit characterization of non-stably ergodic skew-products, in the case when M is an infranilmanifold. Recall that all known examples of Anosov diffeomorphisms live in this class of manifolds, which includes tori as a special case. In this setting, a skew-product is stably ergodic if and only if it is stably ergodic among skew-products.

8.1.3. *Stable ergodicity without partial hyperbolicity.* The previous constructions yield partially hyperbolic maps with invariant splittings $TM = E^u \oplus E^c \oplus E^s$, where all sub-bundles have positive dimensions. The first examples of robustly transitive diffeomorphisms without any invariant hyperbolic sub-bundle were given in [67]; recall Section 7.3.4. These examples contain a C^1 open subset of the space of volume preserving diffeomorphisms. Recently, Tahzibi has shown that such diffeomorphisms, and even a somewhat larger class, are stably ergodic:

Theorem 8.3 (Tahzibi [323]). *There exists a C^1 open set of volume preserving diffeomorphism of the 4-dimensional torus which admit no uniformly contracting nor uniformly expanding invariant sub-bundle and such that every $C^{1+\alpha}$, $\alpha > 0$ diffeomorphism in this open set is ergodic.*

Conjecture 8.4.
(1) Every stably ergodic diffeomorphism admits a dominated splitting and, in fact, is volume hyperbolic.
(2) Every symplectic diffeomorphism stably ergodic (within symplectic diffeomorphisms) is partially hyperbolic.

When the ambient manifold has dimension 4, the second conjecture is contained in Arnaud [17, 19]. Observe that in the symplectic context, existence of a dominated splitting implies partial hyperbolicity, see [19, 50].

Concerning the first conjecture, [58] shows that in the absence of dominated splitting a completely elliptic periodic point may be created by a smooth C^1 small conservative perturbation. It remains to prove that such a periodic point may be used to break ergodicity, and perhaps even the transitivity, after another perturbation.

8.2. **Accessibility and ergodicity.** In this section we always consider diffeomorphisms admitting a partially hyperbolic splitting $TM = E^u \oplus E^c \oplus E^s$ where all three sub-bundles have positive dimensions.

The notion of accessibility, which we recalled in Definition 7.19, embodies the idea of joining different points in the manifold by paths formed by strong stable and strong stable arcs. However, by itself, accessibility is not enough to carry out the Hopf argument: what is missing is the crucial requirement that *all concatenation points of the us-path should belong to the full measure set E where the ergodic theorem holds.* This can obtained, for instance, for the geodesic flows on manifolds with constant negative curvature, using the fact that the strong stable and strong stable foliations are smooth. Moreover, Hopf [160] proved that in the case of hyperbolic surfaces the two invariant foliations are always of class C^1, which was the crucial step in his proof of ergodicity of the corresponding geodesic flow.

A fundamental contribution of Anosov [11] was to realize that a much weaker regularity property, called absolute continuity, suffices for the Hopf argument, and that the stable and unstable foliations of geodesic flows of manifolds with negative curvature are always absolutely continuous. In fact, this is true for all C^2 uniformly hyperbolic systems. See Appendix C.

In the partially hyperbolic context, we still have absolute continuity, for the strong stable and the strong unstable foliations (Appendix B.2), but this alone is not sufficient to implement the Hopf argument: this time, stable/unstable leaves are not cross-sections for the stable/unstable foliation. Brin, Pesin [78] could bypass this difficulty, at the price of assuming Lipschitz regularity of the central foliation, which is not a robust property under small perturbations.

A breakthrough came with the recent work of Pugh, Shub, and their coauthors, see [80, 282] and references therein. Based on several partial results, they proposed

Conjecture 8.5 (Pugh, Shub [282]).
(1) Stable accessibility is dense among partially hyperbolic C^2 diffeomorphisms (volume-preserving or not).
(2) A partially hyperbolic volume preserving C^2 diffeomorphism with the essential accessibility property is ergodic.
(3) Thus, stable ergodicity is a dense property in the space of partially hyperbolic volume preserving C^2 diffeomorphisms.

Part (1) has been proved in the C^1 topology by Dolgopyat, Wilkinson, as we commented upon in Section 7.4.2. In the direction of part (2), the best current result is due to Pugh, Shub [282, 283]. This is the subject of the next section. Part (3) has been solved in the C^1 topology, when the central sub-bundle is 1-dimensional, as we shall see in Theorem 8.16 below.

8.3. The theorem of Pugh-Shub.

Theorem 8.6 (Pugh, Shub [282, 283]). *Every C^2 volume preserving diffeomorphism which is partially hyperbolic, essentially accessible, dynamically coherent, and center bunched is ergodic.*

Recall that *essential accessibility* means that any subset saturated by us-paths has either zero of full Lebesgue measure, and *dynamical coherence* means existence and uniqueness of invariant center stable and center unstable foliations; see Appendix B.2. *Center bunching* is a technical condition which, roughly, means that the derivative is (uniformly) almost conformal along each of the three invariant sub-bundles, and almost an isometry along E^c. It ensures that the sub-bundles and the holonomies of their integral foliations are θ-Hölder for some θ not far from 1.

We only say a few words about the proof of Theorem 8.6, referring the reader to [80] for a detailed discussion and references. The basic strategy is to try and implement the Hopf argument by propagating *density points* of the set E, where the ergodic theorem holds, along the strong stable and strong unstable foliations. Density points are usually defined in terms of relative measure of the set inside small balls around the point. On important difficulty is that strong stable and strong unstable holonomies need not preserve the family of balls, nor any family of sets with bounded eccentricity. So, conceivably, they may not preserve the set of density points.

Instead, Pugh, Shub consider *julienne density points*, defined using a different class of neighborhoods called juliennes such that, under the hypotheses of the theorem, their shape is not too distorted by the holonomies. For instance, center stable juliennes tend to be much longer in the center direction than in the stable direction (dually for center unstable juliennes), to compensate for poor distortion control along the central direction. Then the set of julienne density points is essentially preserved by the strong stable and the strong unstable holonomies, which is the key ingredient in setting the Hopf argument.

Based on Theorem 8.6, Burns, Pugh, Wilkinson [81] prove that the time-1 map of any volume preserving Anosov flow is stably ergodic, unless its stable and unstable sub-bundles are jointly integrable. Joint integrability can not occur for contact Anosov flows, because they always satisfy accessibility [173]. Hence, in particular, *the time-*1 *map of the geodesic flow on any compact manifold with negative curvature is stably ergodic* [11].

8.4. **Stable ergodicity of torus automorphisms.** Let $A \in \mathrm{SL}(d, \mathbb{Z})$ and f_A be the automorphism induced by A on the d-dimensional torus. It is a classical fact that f_A ie ergodic if and only if no eigenvalue of A is a root of unity [147]. If A is ergodic then f_A is an Anosov diffeomorphism, and so it is stably ergodic. The following question was asked by Hirsch, Pugh, Shub [158]:

Problem 8.7. Is every ergodic linear automorphism of the torus stably ergodic ?

A partial solution, and a very original approach to proving stable ergodicity, have been recently proposed by Rodriguez-Hertz [296]. In particular, he proves

Theorem 8.8 (Rodriguez-Hertz [296])**.** *Any ergodic linear automorphism of T^d is C^5-stably ergodic, for $d \leq 5$.*

For more precise statements we need the notion of pseudo-Anosov linear map. We say that A is *pseudo-Anosov* if no eigenvalue is a root of unity and the characteristic polynomial is irreducible over the integers and not a polynomial of a power x^k for any $k \geq 2$. This terminology is motivated by the following observation: Given a homeomorphism h on a surface S of genus g, if the action of h on the first homology group $H_1(S, \mathbb{Z}) = \mathbb{Z}^{2g}$ is pseudo-Anosov, in this sense, then h is isotopic to a pseudo-Anosov homeomorphism of S.

Theorem 8.8 is a direct consequence of the first claim in the statement that follows, together with the fact that for $d = 4$ any ergodic linear automorphism must be pseudo-Anosov, and for $d = 3$ or 5 it must even be Anosov. As a matter of fact, for odd dimensions pseudo-Anosov automorphisms are necessarily Anosov.

Theorem 8.9 (Rodriguez-Hertz [296])**.** *Let T^d denote the d-dimensional torus.*

1. *Any pseudo-Anosov linear automorphism of T^4 is C^5 stably ergodic.*
2. *Any pseudo-Anosov linear automorphism of T^d with* 2*-dimensional central sub-bundle is C^{22} stably ergodic, for $d \geq 6$.*

The central sub-bundle corresponds to the eigenvalues of norm 1. We mention that Shub, Wilkinson [314] proved that f_A is *approximated* by stably ergodic diffeomorphisms, if A acts as an isometry on the central sub-bundle.

Let us comment on the proof of Theorem 8.9. The starting point is the criterium given by Theorem 8.6. Observe that the automorphism f_A is partially hyperbolic, dynamically coherent, and center bunched. Partial hyperbolicity and center bunching are stable properties under C^1 perturbations. So is dynamical coherence in this situation, because the unperturbed map has a C^1 central foliation; see [158, Proposition 2.3]. Therefore, the whole issue is accessibility or, more precisely, essential accessibility: notice that the automorphism f_A itself is essentially accessible, but not accessible.

The main novelty in Rodriguez-Hertz's approach is that he is able to give a surprisingly strong description of the topology of accessibility classes. Let f be a

[11]We learned this from Amie Wilkinson.

small perturbation of f_A. The *accessibility class* of a point in T^d is the set of all points which may be joined to it by a us-path of f. It is useful to consider the lift $F : \mathbb{R}^d \to \mathbb{R}^d$ of f to the universal covering, which is also a partially hyperbolic map. Let us denote $C(x)$ the accessibility class of a point $x \in \mathbb{R}^d$ for the lift F. One key step is

Proposition 8.10. *Either $C(0) = \mathbb{R}^d$, in which case f has the accessibility property, or else $\#\big(C(x) \cap W^c(0)\big) = 1$ for all $x \in \mathbb{R}^d$.*

In fact, this is a consequence of the following, more detailed, statement:

Proposition 8.11. *For any $x \in \mathbb{R}^d$ one of the following holds:*
(1) *either $C(x)$ is open*
(2) *or $C(x) \cap W^c(0)$ is homeomorphic to the circle or to an interval*
(3) *or $C(x) \cap W^c(0)$ contains exactly 1 point.*

Moreover, alternative (2) can not occur for the origin $x = 0$, and the set of points x for which (3) occurs is either empty or the whole $\mathbb{R}^d$.

To prove Theorem 8.9 from Proposition 8.10, there are two cases. If $C(0) = \mathbb{R}^d$, the map f is accessible and thus ergodic. In the case where $C(x) \cap W^c(0)$ reduces to a single point, for all $x \in \mathbb{R}^d$, Rodriguez-Hertz ingeniously appeals to KAM arguments to prove that f is conjugate to f_A. It follows that f is essentially accessible and so, once more, ergodic.

8.5. **Stable ergodicity and robust transitivity.** The theories of stable ergodicity and robust transitivity had very parallel developments in these last few years and, as we mentioned before, all new classes of examples of robustly transitive systems may be adapted to get similar stably ergodic examples. It is then natural to ask

Problem 8.12. Does robust transitivity of conservative systems imply stable ergodicity?

Another fundamental question concerns the relations between accessibility and transitivity, namely in the dissipative context. The following very nice result of Brin gives a partial answer:

Theorem 8.13 (Brin [74]). *Let f be a partially hyperbolic diffeomorphism with the accessibility property and whose non-wandering set $\Omega(f) = M$. Then f is transitive.*

The hypothesis concerning the wandering set is automatic if the diffeomorphism is conservative. In that setting the hypotheses of Theorem 8.13 are satisfied on a C^1 open and dense subset, by a recent result of Dolgopyat, Wilkinson [122] quoted in Theorem 7.20. Thus, *robust transitivity is a dense property among partially hyperbolic C^1 diffeomorphisms.* Even more recently, Bonatti, Crovisier used their closing lemma for pseudo orbits, Theorem A.7, to prove that *generic volume preserving C^1 diffeomorphisms are transitive.*

Theorem 8.13 is false outside the conservative world: Niţică, Török [247] construct an open set of non-transitive accessible diffeomorphisms. This is an application of the following result:

Theorem 8.14 (Niţică, Török [247]). *Stably accessible maps are C^2 dense in a neighborhood of the product of an Anosov volume preserving transitive diffeomorphism by the identity on the circle.*

Consider $f = h \times \phi$, where $h : M \to M$ is a transitive volume preserving Anosov diffeomorphism and ϕ is a circle diffeomorphism close to the identity and exhibiting some periodic attractor. Then there exists some open set $U \subset S^1$ such that $g(M \times \bar{U}) \subset M \times U$ for $g = f$ and, thus, for any g in a C^0 neighborhood. It follows that no such g is transitive, and we may take a C^2 open subset which are accessible.

Nevertheless, Niţică, Török [247] also show that stably ergodic diffeomorphisms are C^2 dense in a neighborhood of the product of a transitive volume preserving Anosov diffeomorphism by the identity on the circle, as well as in a neighborhood of the time-1 map of any transitive volume preserving Anosov flow, *inside volume preserving diffeomorphisms.*

8.6. **Lyapunov exponents and stable ergodicity.** Burns, Dolgopyat, Pesin [79] address Conjecture 8.5 within non-uniformly hyperbolic conservative systems, that is, such that all Lyapunov exponents are different from zero almost everywhere. The following theorem summarizes their main results:

Theorem 8.15 (Burns, Dolgopyat, Pesin [79]). *Let $f : M \to M$ be a partially hyperbolic volume preserving C^2 diffeomorphism on a compact manifold. Assume that the central Lyapunov exponents are negative on an invariant positive Lebesgue measure set A. Then*

(1) *Every ergodic component of $f \mid A$ is an open set* mod 0*. If f is transitive then A is dense and $f \mid A$ is ergodic.*
(2) *If f is essentially accessible then it is ergodic. In particular, the central Lyapunov exponents are negative almost everywhere. If f is accessible then it is stably ergodic.*

The statement in (1) means that each ergodic component (more precisely, the set of points whose time averages are given by it) coincides with an open set up to zero Lebesgue measure. Let us sketch some main points in the proof of this theorem.

Using ideas from Alves, Bonatti, Viana [7, 67], together with the hypothesis on the Lyapunov exponents, one shows that almost every point in A is in the basin of some SRB measure ν which is an ergodic Gibbs u-state with negative central exponents. See Theorem 11.35 below. Then ν is the ergodic component of the Lebesgue measure that contains the point. For ν-almost every point, its (local) Pesin stable manifold is contained mod 0 in the basin of ν. Considering strong unstable leaves with uniform size through the points in that stable manifold, we get an open set which is mod 0 contained in the basin of μ. This gives the first half of (1). The second one is a direct consequence.

Next, essential accessibility implies that, given any ergodic component, almost every point may be connected to it by a us-path whose corner points are all regular, as in the Hopf argument. This gives ergodicity, as claimed in the first part of (2). The second one follows immediately. To get the last part, one begins by obtaining a uniform lower bound for the sizes of ergodic components of the perturbations of f, using the idea of hyperbolic times from [7]. Then one observes that if f is accessible then any perturbation is ε-accessible (the us-path saturation of any ε-ball is the whole manifold) which, in view of the previous observation, suffices for ergodicity.

In this context, Fayad [129] announces that essential accessibility alone does not ensure stable ergodicity.

Dolgopyat, Wilkinson [122] prove that stable accessibility is dense in the space of partially hyperbolic C^1 diffeomorphisms, as we have seen in Theorem 7.20. In another very recent paper, Baraviera, Bonatti [36] show how to approximate any volume preserving partially hyperbolic diffeomorphism by another such that the integrated sum of its central Lyapunov exponents is different from zero. This will be discussed in Section 12.6. In particular, in the case when the central sub-bundle is 1-dimensional the central Lyapunov exponent is non-zero for an open and dense subset of diffeomorphisms.

Combining these results, one gets a proof of Conjecture 8.5(3) in the C^1 topology, when the central sub-bundle is 1-dimensional:

Theorem 8.16 (Bonatti, Matheus, Viana, Wilkinson [64]). *Any partially hyperbolic volume preserving C^2 diffeomorphism on a compact manifold M having 1-dimensional central bundle is C^1 approximated by another which is stably ergodic.*

9. Robust attractors of flows

While continuous time systems contain the main features of their discrete time counterparts, there is a very important phenomenon specific to the continuous time case: robust accumulation of equilibrium points by regular orbits of the flow.

The first example of this phenomenon originated from the famous strange system of differential equations

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= xy - \beta x, \end{aligned} \tag{27}$$

proposed by Lorenz [195] as a simple model for sensitive dependence on initial conditions, loosely related to fluid convection and weather prediction. A detailed analysis of these equations may be found in the book of Sparrow [320].

The observations of Lorenz were explained by Afraimovich, Bykov, Shil'nikov [3] and Guckenheimer, Williams [143, 344] in terms of existence of a robust attractor containing a singularity (equilibrium point) together with regular orbits. Their constructions of so-called *geometric Lorenz attractors* showed that flows with such robust attractors exist. However, it took yet about two decades before Tucker proved that they do occur in the original system (27):

Theorem 9.1 (Tucker [327, 328]). *For the classical parameters, $\sigma = 10$, $\rho = 28$, $\beta = 8/3$, the Lorenz equations support a robust strange (sensitive) attractor.*

By then, starting from the mid-nineties, a theory was being developed to explain the phenomenon of robust coexistence of singular and regular orbits in the same transitive set, and to describe the properties of such sets. A major progress was the work of Morales, Pacifico, Pujals [229, 227] stating that *robust singular sets of 3-dimensional flows are singular hyperbolic attractors* (possibly after time-reversal), in fact, they share all the main properties of the geometric Lorenz attractors. This opened the way for a rather complete theory of generic flows in dimension 3, unifying uniformly hyperbolic flows and Lorenz-like flows.

9.1. **Singular invariant sets.**

In this section we describe two main models of singular hyperbolic sets.

9.1.1. *Geometric Lorenz attractors.* An *attractor* A for a flow X^t is a compact invariant set containing a dense orbit and admitting an open neighborhood U such that

$$X^t(\overline{U}) \subset U \quad \text{for all} \quad t > 0 \quad \text{and} \quad A = \bigcap_{t \geq 0} X_t(U).$$

The attractor is called *robust* if $A_Y = \bigcap_{t\geq 0} Y_t(U)$ is also an attractor for every vector field Y close to X in the C^1 topology. The attractor is *singular* if it contains equilibrium points of the vector field. Finally, we say that the attractor is *strange* if a dense orbit may be found exhibiting some positive Lyapunov exponent.

Historically, the first examples of robust singular strange attractors were the geometric Lorenz attractors, proposed in [3, 143, 344] as models for the behavior of the Lorenz equations (27) at parameter values near $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. These are smooth flows in dimension 3, with the following main features (see Figure 15):

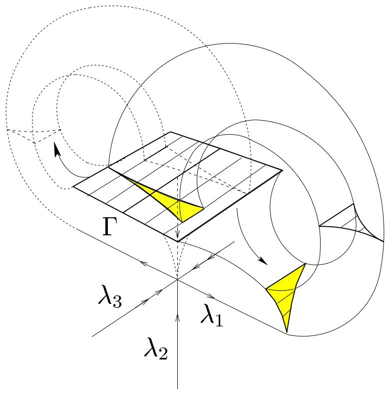

FIGURE 15. Geometric Lorenz attractor

(1) (existence of singularity) The vector field X has a singularity θ such that $DX(\theta)$ has real eigenvalues $\lambda_1 > 0 > \lambda_2 > \lambda_3$ with $\lambda_1 + \lambda_2 > 0$.
(2) (existence of cross-section) There is a cross-section Σ intersecting the stable manifold of θ along a curve Γ that separates Σ into 2 connected components, and a smooth first return map $F : \Sigma \backslash \Gamma \to \operatorname{int}(\Sigma)$ defined on the complement of Γ.
(3) (invariant stable foliation) There is a smooth foliation of Σ into curves, having Γ as one of the leaves, invariant and uniformly contracted by forward iterates of the return map F. The quotient space of Σ by this foliation is a compact interval I.
(4) (one-dimensional reduction) The quotient map f defined by F on the leaves space I is piecewise smooth and uniformly expanding on $I \setminus \{c\}$, with a discontinuity and infinite right and left derivatives at the point $c \in I$ that represents the leaf Γ. See Figure 16. Moreover, f is topologically mixing (thus, transitive) in a strong sense: every interval is mapped, eventually, onto the whole I.

The corresponding attractor Λ is characterized as follows. First, note that the restrictions of f to both $\{x < c\}$ and $\{x > c\}$ admit continuous extensions to the point $x = c$. Hence, f extends to a map $\tilde{f}$ that is 2-valued at $x = c$ and continuous

FIGURE 16. Poincaré transformation and associated Lorenz map

on both $\{x \le c\}$ and $\{x \ge c\}$. Correspondingly, the restriction of the Poincaré F to each of the connected components of $\Sigma \setminus \Gamma$ admits a continuous extension to the closure, each one collapsing the curve Γ to a single point. Thus, F may be extended to a 2-valued transformation $\tilde{F} : \Sigma \to \Sigma$ defined on the whole cross-section and continuous on the closure of each of the connected components. Let

$$\Lambda_F = \bigcap_{n \ge 0} \tilde{F}^n(\Sigma) \subset \Sigma$$

and Λ be the saturated of Λ_F by the flow, that is, the union of the orbits of its points. Thus orbits in Λ intersect the cross-section infinitely often, both forward and backward.

??

The attractor has a very complicated structure that was nicely described by Nitecki as a Cantor book with uncountably many pages joined along a spine corresponding to the two branches of the unstable manifold of the singularity and the ways in which the two branches of its unstable manifold return to transverse section of the local stable manifold of the singularity.

The theory of symbolic dynamics gives two parameters characterizing the map f up to topological equivalence, such parameters are obtained as follows.

Let $f : I \to I$ be a Lorenz interval map, and $\tilde{f}$ be the two-valued extension introduced before. The itinerary of a point $x \in I$ is the sequence $(x_n)_n$ defined, for each $n \ge 0$ by $x_n = 0$ if $\tilde{f}^n(x) \le c$ and $x_n = 1$ if $\tilde{f}^n(x) \ge c$. Note that the itinerary is multi-valued if the orbit of x contains the discontinuity c.

The *kneading sequences* of f are the itineraries $\alpha = (a_i)_{i \ge 0}$ and $\beta = (b_i)_{i \ge 0}$ of the endpoints $a = \tilde{f}(0^+)$ and $b = \tilde{f}(0^-)$.

Consider the space $\Sigma_2(\mathbb{N}) = \{0,1\}^{\mathbb{N}}$ and the shift defined on $\Sigma_2(\mathbb{N})$, $\sigma\colon \Sigma_2(\mathbb{N}) \to \Sigma_2(\mathbb{N})$, $\sigma((x_i)) = (y_i)$, where $y_i = x_{i+1}$. In the space $\Sigma_2(\mathbb{N})$ there are defined the distance d and the ordering $<$ as follows, given a pair of sequences $x = (x_i)_{i \ge 0}$ and $y = (y_i)_{i \ge 0}$, one lets $d(x,y) = \sum_{i=0}^{\infty} |x_i - y_i| \, 2^{-i}$ and write $x < y$ if there is n such that $x_i = y_i$ for all $i < n$ and $x_n = 0$ and $y_n = 1$.

Consider the σ-invariant subset Γ of $\Sigma_2(\mathbb{N})$ of sequences $x = (x_i)_i$ such that $\alpha \le \sigma^i(x) \le \beta$ for all $i \ge 0$. Let $I_0 = [-1, 0]$ and $I_1 = [0, 1]$. Given $x_0, x_1, \dots, x_n$, $x_i \in \{0, 1\}$, let $I(x_0, x_1, \dots, x_n) = \cap_{i=0}^{n} f^{-i}(I_{x_i})$. Using the expansivity of f it follows that the map $\varphi\colon \Gamma \to [0,1]$, $\varphi((x_i)) = \cap_{i=0}^{\infty} f^{-i}(I_{x_i})$, is well defined. In fact, this map is onto (observe that $\varphi(\alpha) = -1$ and $\varphi(\beta) = 1$). To get the surjectivity of φ, for a given $x \in (-1, 1)$ consider the sequence $(x_i) \in \Sigma_2(\mathbb{N})$ defined by $x_i = j$ if $f^i(x^+) \in I_j$ (i. e., there is an interval (x, δ) such that $f^i(x, \delta)$ is contained in the interior of I_j). By construction, $\varphi((x_i)) = x$. So it remains to see that (x_i) belongs to Γ. For instance, to see that $\alpha \le (x_i)$ observe that $f^{i+1}(-1) < f^{i+1}(x)$ whenever $f^k(-1)$ and $f^k(x)$ are in the same side of 0 for all $k = 0, 1 \dots, i$. Let n be the biggest i with such a property. Thus $(a_0, a_1, \dots, a_n) = (x_0, x_1, \dots, x_n)$ and $a_{n+1} = 0 < 1 = x_{n+1}$, which implies the claim. Similarly one gets $(x_i) \le \beta$ (thus $(x_i) \in \Gamma$).

Finally, the map φ is injective except for the sequences mapped into $f^{-i}(0)$ by φ, where the map φ is $2-1$. Observe that

$$\varphi(x_1,\ldots,x_n,0,1,\alpha)=\varphi(x_i,\ldots,x_n,1,0,\beta).$$

Let $\tilde{\Gamma}$ be the quotient space of Γ obtained by identifying every pair of sequences of Γ with $(x_1,\ldots,x_n,0,1,\alpha)$ and $(x_i,\ldots,x_n,1,0,\beta)$. The map f is topologically equivalent to the shift $\tilde{\sigma}\colon\tilde{\Gamma}\to\tilde{\Gamma}$ induced by σ on $\tilde{\Gamma}$. Observe that the shift $\tilde{\sigma}$ has a discontinuity at the point corresponding to $(0,1,\alpha)$ or $(1,0,\beta)$.

The topological conjugacy of two (expanding) Lorenz maps f_1 and f_2 is now stated in terms of the kneading sequences (α_1,β_1) and (α_2,β_2) of f_1 and f_2: the maps f_1 and f_2 are conjugate if one of the following possibilities hold: $\alpha_1=\alpha_2$ and $\beta_1=\beta_2$, or $\alpha_1=\beta_2'$ and $\beta_1=\alpha_2'$, where (x_i') is the *dual sequence* of (x_i) defined as $x_i'=1$ if $x_i=0$ and $x_i'=0$ if $x_i=1$ (the first possibility corresponds to the case where the conjugacy preserves the orientation and the second one to a reversing orientation conjugacy). This and the fact that φ is almost a bijection give the two parameters describing the equivalence classes of Lorenz maps.

The maximal invariant set Λ of F in R is the ω-limit set of almost all points of R (for instance, think of the points which are mapped into the line ℓ in the local stable manifold of the singularity). Flowing Λ and considering its closure we get the Lorenz attractor of the flow. The restriction of F to Λ is the set *pinched inverse limit set* for f (here there is an additional problem when considering inverse limits due to the fact that the map f is bi-valuated at the discontinuity 0). The *inverse limit set* for a well defined map g defined on a metric space X is constructed as follows. Consider the space $\hat{X}$ of all the sequences (x_i) contained in X such that $g(x_{i+1})=x_i$ and define $\hat{g}\colon\hat{X}\to\hat{X}$ by $\hat{g}((x_i))=(y_i)$ where $y_i=g(x_i)$. The pair $(\hat{g},\hat{X})$ is the inverse limit for g on X. In our setting the extra difficulty of the existence of discontinuities appears when considering inverse limits. This difficulty is solved as follows.

Consider the square R as a tubular neighborhood of the horizontal interval with vertical fibers and projection map π. By definition, F preserves and contracts the fibers (vertical segments) and projects onto f (i.e., $\pi\circ F=f\circ\pi$). The restriction of F to Λ is topologically equivalent to a certain quotient space of f. Define $\hat{I}$ as above and consider the equivalence relation $\sim$ defined on $\hat{I}$ by identifying every pair of sequences (x_i) and (y_i) of $\hat{I}$ with $x_i=y_i=-1$ or $x_i=y_i=1$ for some i. This corresponds to identify some points in the pre-orbit of the discontinuity 0 of f.

Let $\tilde{I}=\hat{I}/\sim$ and consider the map $\psi\colon\Lambda\to\tilde{I}$ defined by $\psi(x)=(x_i)$ where $x_i=\pi(F^{-i}(x))$. Observe that for points $x\in\Lambda$ that do not project onto the orbit of the discontinuity there is no possible ambiguity on the definition of ψ and this map is well defined. On the other hand, for points $x\in\Lambda$ projecting onto a point of the pre-orbit of 0 there is a ambiguity that disappears after considering the quotient of $\hat{I}$ by $\sim$. By construction one has $\hat{f}\circ\psi=\psi\circ F$. To see that ψ is a topological equivalence it remains to prove that ψ is a homeomorphism. The injectivity of ψ follows from the fact that F expands horizontal segments. The surjectivity of ψ essentially follows as the one of the map $\varphi\colon\Gamma\to I$ above. This completes the study of the geometric Lorenz attractor.

9.1.2. *Singular horseshoes.* The primary motivation for this model of singular set, introduced by Labarca, Pacifico in [187], was to exhibit vector fields on compact manifolds *with boundary* having a singularity accumulated by regular closed orbits and which are structurally stable. Note that geometric Lorenz attractors, although robust, are not stable: the dynamics on the attractor usually changes when the system is perturbed.

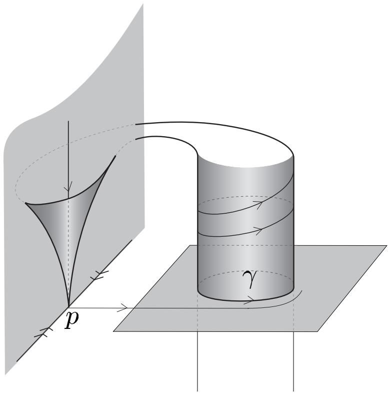

FIGURE 17. Singular cycle giving rise to a singular horseshoe

Consider a vector field X on a 3-dimensional manifold and having a singularity θ of saddle type and a regular periodic orbit γ of saddle type such that

(1) $DX(\theta)$ has three real eigenvalues with $\lambda_1 < \lambda_2 < 0 < \lambda_3$ and $\lambda_3 + \lambda_2 > 0$,
(2) there is a branch of $W^u(\theta)$ contained in $W^s(\gamma)$,
(3) there are two orbits of transverse intersection between the two dimensional manifolds $W^u(\gamma)$ and $W^s(\theta)$.

This geometric configuration is a codimension one bifurcation corresponding to a singular cycle formed by θ and γ. See Figure 17.

To study the dynamics of this vector field one argues similarly as in the case of the geometric Lorenz attractor. Consider a transverse section $S \simeq [0,1] \times [0,1]$ to the flow at some point of p of the closed regular orbit γ and the first return map F of the flow to S.

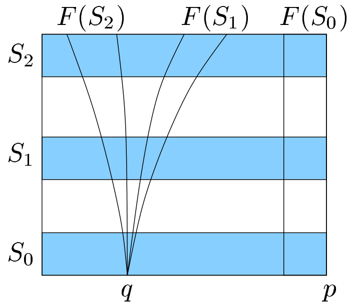

FIGURE 18. Return map for singular horseshoe

Define q as the first intersection of the branch of the unstable manifold of the singularity contained in $W^s(\gamma)$ and S. The Poincaré return map F is a modification of the horseshoe of Smale of three symbols called *singular horseshoe*. The set of points of S having a first forward return to S consists of three horizontal strips (sub-rectangles), say $S_0=[0,1]\times[0,a]$, $S_1=[0,1]\times[b,c]$ and $S_2=[0,1]\times[d,1]$, where $0<a<b<c<d<1$. The strip S_0 contains the point $p\in\gamma$ (thus p is a hyperbolic fixed point of F), and the rectangles S_1 and S_2 are such that their bottom sides $[0,1]\times\{b\}$ and $[0,1]\times\{d\}$ are contained in the stable manifold of the singularity θ and thus the first return is not defined. These strips are mapped hyperbolically by F (roughly speaking, contracting the horizontal direction and expanding the vertical one), the image of S_0 by F is the sub-rectangle $[e,1]\times[0,1]$ of S (thus $p=(0,1)$).

The main difference with respect to the usual horseshoe comes from the fact that the bottom parts of S_1 and S_2 do not return to S: consequently, the images of these two strip by F are cuspidal triangles joined at the common cuspidal vertex q and going across S from top to bottom. This return map is depicted in Figure 18. The point q can be thought as the image of the bottom parts of S_1 and S_1.

The analysis of the dynamics of the singular horseshoe follows now as in the Lorenz attractor. After proving the existence of a strong stable foliation one reduces the dynamics of the return map to an one-dimensional one. For studying this one-dimensional dynamics one considers a shift of three symbols (corresponding to the strips S_0, S_1 and S_2) and identify some sequences (essentially corresponding to the discontinuities). The one-dimensional dynamics is conjugate to such a shift.

While the geometric Lorenz attractors leads to the category of expanding Lorenz maps the singular horseshoes leads us to consider maps $f\colon I\to[0,1]$ where $I=[0,1]\setminus(a',b')\cup(c',d')$, $[a',b']$ and $[c',d']$ are disjoint and contained in the interior of $[0,1]$, such that

(HL1) $f(0)=0$, $f(a')=1=f(b')=1$, $f(c')=f(d')=0$, and $f(1)=1$,
(HL2) there is $k>1$ such that $|f'(x)|>k$ for all $x\in I$.

Let Λ_f be the forward invariant set of f in I. Write $I_0=[0,a']$, $I_1=[b',c']$ and $I_2=[d',1]$. To each point $x\in\Lambda$ we associate the element $(i_j(x))$ of $\Sigma_3(\mathbb{N})=\{0,1,2\}^{\mathbb{N}}$ defined by $i_j(x)=k$ if and only if $f^j(x)\in I_k$. It is easy to check that the restriction of f to Λ_f is conjugate to the shift defined on $\Sigma_3(\mathbb{N})$.

Consider now a map G having a linear horseshoe in S conjugate to the complete shift of three symbols such that G is affine on each strip S_0, S_1 and S_2, and $G(S_0)=[a',b']\times[0,1]$, $G(S_1)=[c'd']\times[0,1]$, and $G(S_2)=[e'1]\times[0,1]$. Write $I=[a',b']$, $J=[c'd']$. Let Λ_G the maximal invariant set of G in S and h a homeomorphism $h\colon\Lambda_\Gamma\to\Sigma_3(\mathbb{Z})$ conjugating G and the shift σ defined on $\Sigma_3(\mathbb{Z})$.

The only difference between F and G is that F is not injective on the bottom parts of S_0 and S_1, and collapse these intervals to a single point (the point q). In particular F fails to be injective. Essentially, F is obtaining from G by squeezing the intervals I and J getting a unique image. Consider the invariant set Λ of the F in S. Roughly speaking, the restriction of F to Λ is conjugate to a part of the shift of three symbols. Dynamically the pinching procedure carries some identifications of points of $\Sigma_3(\mathbb{Z})$. At the first stage, the points of I and J are identified. But this identification carries new identifications when considering the backward orbits of I and J. Let us explain all that in details.

9.2. **Singular cycles.** A *cycle* of a vector field X is a compact invariant subset Γ of the ambient manifold f the form $\Gamma=\Gamma_c\cup\Gamma_r$, where $\Gamma_c=\{\theta_0,\theta_1,\dots,\theta_k\}$ is formed by finitely many critical elements (singularities or regular periodic orbits), and Γ_r consists of non-periodic regular orbits whose ω and α-limit are contained in Γ_c and such that for each $i \in \{0, \dots k\}$ there is an orbit γ of Γ_r such that $\alpha(\gamma) = \theta_i$ and $\omega(\gamma) = \theta_{i+1}$ (respectively, $\omega(\gamma) = \theta_0$ if $i = k$). The cycle Γ is called *singular* if it contains at least a singularity, and *hyperbolic* if all the critical elements involved in it are hyperbolic. The cycle in Figure 17 involves exactly one singularity and one regular periodic orbit.

The notion of singular cycle was introduced by Labarca, Pacifico [187], and a systematic study was initiated by Bamón, Labarca, Mañé, Pacifico [35], in the setting of vector fields on 3-dimensional manifolds. Singular cycles are a main mechanism for producing singular hyperbolic sets, in a sense they play a role similar to that of heterodimensional cycles for diffeomorphisms.

9.2.1. *Explosions of singular cycles.* Bamón, Labarca, Mañé, Pacifico [35] considered 1-parameter families of vector fields $(X_\mu)_{\mu\in[-1,1]}$ on 3-dimensional manifolds, starting at a simple vector field X_{-1}, meaning one whose chain recurrent set [12] consists of finitely many hyperbolic critical elements, and leaving the region of simple systems at a parameter $\mu = 0$ for which the vector field $X = X_0$ exhibits a singular cycle Γ. Their main assumptions are

(a) The singular cycle Γ is hyperbolic and contains exactly one singularity θ.
(b) The eigenvalues of $DX(\theta)$ are real and satisfy $\lambda_1 < \lambda_2 < 0 < \lambda_3$ (up to time reversal).
(c) The intersections between the invariant manifolds of the elements of Γ_c are all transverse, except for the intersection between the unstable manifold of the singularity and the stable manifold of a periodic orbit $\gamma_1 \in \Gamma_c$.
(d) The latter consists of a unique orbit γ, which is tangent at infinity to the eigenspace of $DX(\theta)$ associated to the eigenvalue λ_2.
(e) The cycle has an *isolating block*, that is, an open neighborhood U such that Γ is the maximal invariant set of X in U.

The property that the eigenvalues be real, in (b), is a consequence of the assumption that the cycle is a first bifurcations. Concerning condition (c), a simple dimension counting argument shows that in a hyperbolic cycle not all intersections can be transverse.

Then one wants to describe the dynamics on the maximal invariant set

$$\Lambda(X_\mu, U) = \bigcap_{t\in\mathbb{R}} X_\mu^t(U)$$

of X_μ inside U, for small positive values of the parameter μ. For technical reasons, mostly having to do with the use of linearization results [321], one restricts to vector fields of class C^r, with $r \geq 3$:

Theorem 9.2 (Bamón *et al* [35])**.** *There is a neighborhood $\mathcal{U}$ of $X = X_0$ and a codimension-*1 *submanifold $\mathcal{N}$ of $\mathcal{X}^r(M)$ containing X such that every $Y \in \mathcal{N} \cap \mathcal{U}$ has a singular cycle, which is the continuation of that of X.*

---

[12]The chain recurrent set is formed by the points $p \in M$ such that there exist $(x_i)_{i=0}^m$ in M and positive numbers $(t_i)_{i=1}^{m-1}$ such that $x_0 = p = x_m$, and $X^{t_i}(x_i)$ is at distance less than ε from x_{i+1} for all $i = 0, \dots, m-1$.

Moreover, $\mathcal{U} \setminus \mathcal{N}$ has two connected components $\mathcal{U}^-$ and $\mathcal{U}^+$, such that every vector field Y in $\mathcal{N}^-$ is simple: the chain recurrent set of Y in the isolating block U consists of the continuations of the critical elements of X inside U.

One also wants to describe the behavior of vector fields in the component $\mathcal{U}^+$ corresponding to (a priori) non-simple vector fields. According to the next theorem, for the majority of such vector fields the dynamics related to the cycle is actually hyperbolic. Let $\mathcal{U}_h^+$ be the set of vector fields $Y \in \mathcal{U}^+$ such that the maximal invariant set $\Lambda(Y, U)$ is a (disjoint) union $\{\theta(Y)\} \cup \Sigma(Y)$, where $\theta(Y)$ is the continuation for Y of the singularity θ, and $\Sigma(Y)$ is a hyperbolic transitive set.

Theorem 9.3 (Bamón *et al* [35]). *There is $\delta > 0$ such that $B(t) = \{\mu \in [0,t] : X_\mu \notin \mathcal{U}_h^+\}$ has*

$$\lim_{t\to 0} \frac{1}{t^\delta} \mathrm{Leb}\big(B(t)\big) = 0,$$

for every parametrized family (X_μ) of vector fields starting in $\mathcal{U}^-$ and crossing $\mathcal{N}$ transversely at $\mu = 0$.

9.2.2. *Expanding cycles and contracting cycles.* More precise information on the way $\mathcal{U}_h^+$ fills $\mathcal{U}^+$ depends on the eigenvalues at the singularity. The singular cycle is called *expanding* if $\lambda_3 + \lambda_2 > 0$, and *contracting* if $\lambda_3 + \lambda_2 < 0$. Let us explain why this distinction is important.

Theorem 9.4 (Bamón *et al* [35]). *Theorem 3.*

Let us explain what distinguishes the two cases. For simplicity, we focus the case when the cycle involves only one regular periodic orbit.

The dynamics in the unfolding of a singular cycle as above may be reduced to that of a 1-parameter family of maps of the interval. First, one may reduce the flow to the Poincaré return map to some cross-section to the regular orbit of the cycle **??**. The assumptions imply that this return map admits an invariant strong stable foliation, for small values of the parameter $\mu > 0$. Next, taking the quotient with respect to this foliation, one reduces the return map to a map f_μ of the interval of the form

$$f_\mu : [0, \rho_\mu^{-1}] \cup [a_\mu, 1] \to [0,1] \quad \begin{cases} f_\mu(x) = \rho_\mu x, & \text{if } x \in [0, \rho_\mu^{-1}], \\ f_\mu(x) = \mu - K_\mu(x)(1-x)^{\alpha_\mu}, & \text{if } x \in [a_\mu, 1], \end{cases}$$

where $a_\mu \to 1^-$ as $\mu \to 0^+$, $\rho_0 > 1$ is the expanding eigenvalue of the return to F_0 at its fixed point p, α is the quotient $|\lambda_2|/|\lambda_3|$, and $K_m u(1) > 1$ for all $x \in [a_\mu, 1]$. The constants ρ_μ and α_μ and the map K_μ depend continuously on the parameter μ.

In the expanding case one has $\alpha > 1$, thus the map f_μ is uniformly expanding in its domain of definition (the derivative at 1 goes to infinite as μ goes to zero). This explains why in the expanding case there are no points with derivative equal to one (thus there are no saddle-node of flip bifurcations) In fact, when $\alpha > 1$ the bifurcations of periodic orbits occur when the point 1 (corresponding to the separatrix of the singularity in the cycle) becomes periodic (giving the collapse of a periodic orbit and a cycle).

In the contracting case, $\alpha < 1$, the derivative at 1 is zero, thus there are points with derivative greater than or less that one. This explains the appearance of saddle-node bifurcations. The occurrence of flip bifurcation depends on the global geometry of the cycle (in fact, in some cases the map f_μ has a decreasing branch).

Theorem 9.5 (Bamón *et al* [35]). *Theorem 4.*

In the expanding case the bifurcation set $B(t)$ is very thin: the limit capacity goes to zero as t goes to zero. Moreover, $\mathcal{U}^+ \setminus \mathcal{U}_h^+$ is laminated by codimension-1 submanifolds which correspond to topological equivalence classes: for any two vector fields X, Y in the same lamina, there is a homeomorphism between the chain recurrent sets $\Gamma(X,U)$ and $\Gamma(Y,U)$, sending orbits to orbits and preserving the direction of time. **??** In addition, given any two families $(X_\mu)_\mu$ and $(Y_\mu)_\mu$ crossing $\mathcal{N}$ transversely at $\mu = 0$ this homeomorphism depends continuously on the parameter. See [35, Theorem 5].

The unfolding of contracting singular cycles involves other bifurcations, such as saddle-node or period-doubling, which are not present in the expanding case. A first step in the study of such cycles was made in [35], where it was shown that the set of parameters μ for which X_μ has an attracting periodic orbit whose stable manifold intersects the unstable manifold of the singularity accumulates at the bifurcation parameter $\mu = 0$. This has been extended by Pacifico, Rovella [255], San Martín [306], Labarca [185], as we shall comment upon later.

In [255] Pacifico and Rovella considers the case of a cycle consisting of a singularity and a regular closed orbit where the singularity is *strongly contracting,* that is, its eigenvalues $\lambda_1 < \lambda_2 < 0 < \lambda_3$ verify $\lambda_1 > \lambda_2 + 2\lambda_3$. In this case the map f_μ is such that $f'_\mu(x) > 0$ for all $x \neq 1$ and $f''_\mu(x) < 0$ in $[a_\mu, 1]$. The hypothesis on the strong contractivity of the cycle is a technical condition guaranteeing the C^3-differentiability of the stable foliation of a first return map and of the map f_μ. High differentiability is crucial and required because the Schwarzian derivative of this map is heavily used as we see. Moreover, in this case the Schwarzian derivative Sf_μ of f_μ is negative in $[a_\mu, 1]$.

Theorem 9.6 (Pacifico, Rovella [255]). *Theorem.*

Let $\mathcal{U}_{sh}^+$ be the subset of $\mathcal{U}^+$ of vector fields X whose chain recurrent set of (in U) consists of the continuation of the equilibrium θ, a hyperbolic transitive set and a unique attracting regular closed orbit. For $t > 0$ let $H(t)$ the set of parameters $\mu \in [0,t]$ such that X^μ belongs to $\mathcal{U}_h^+ \cup \mathcal{U}_{sh}^+$. In the strongly contracting case there is $t > 0$ such that $H(t)$ has full measure in $[0,t]$, see [255, Theorem A].

The use of the condition $Sf_\mu < 0$ is rather standard. Let Λ_μ be the maximal (forward) invariant set of f_μ in $[0,1]$ and write $G_\mu = (\rho_\mu^{-1}, a_\mu)$ (the gap of the domain of definition of f_μ). First, if the orbit of the critical point meets the gap G_μ the set f_μ is hyperbolic (the set Λ_μ contains finitely many attracting hyperbolic orbits whose complement in Λ_μ is uniformly expansive) for details see [Theorem?]MS93. Denote by S the set of parameters such that the orbit of 1 meets G_μ (i.e., $1 \notin \Lambda_\mu$), by E the set of parameters such that Λ_μ contains and attracting orbit, and by H the set of hyperbolic parameters. We claim that $E \cup S \subset H$ (we know that $S \subset H$).

To get $E \subset H$ consider μ such that Λ_μ contains as attracting periodic orbit. We claim that f_μ is hyperbolic. To prove this assertion suppose that Λ_μ contains an attracting periodic orbit p of period n. Since the orbit of p necessarily meets $[a_\mu, 1]$ (where the Schwarzian derivative of f is negative) and f_μ is linear in $[0, \rho_\mu^{-1}]$ one has that Sf^n is negative in the basin of attraction of p. Singer's Theorem in [317] implies that the critical point of f_μ (i.e., the point 1) is in this basin. This immediately implies that Λ_μ contains at most one periodic attracting orbit.

To prove the claim it suffices to see that the complement of the basin of attraction of p in Λ_μ is a hyperbolic set. But this follows since $Sf_\mu < 0$ in $[a_\mu, 1]$ and linear in $[0, \rho_\mu^{-1}]$.

The final step is to see that the parameters μ such that either the orbit of 1 meets G_μ or Λ_μ has a sink have full measure in $[0, \mu_0]$ for μ_0 small enough.

This result of Pacifico and Rovella was generalized first by San Martín in [306] for singular cycles with two critical elements and a contracting singularity (instead of a strongly contracting one). Latter, in [185] Labarca extends this results for cycles with a unique contracting singularity and an arbitrary number of regular closed orbits.

In fact, in [185] Labarca extended the results in [35] for expanding singular cycles for contracting ones having a unique equilibrium (with real eigenvalues) and an arbitrary number of closed orbits, where, under certain regularity assumptions, a detailed and complete description of a neighborhood of a vector field X with a contracting singular cycle X is given.

Previous results give a rather complete description of the bifurcation of singular cycles (the singularity with real eigenvalues) in three manifolds corresponding to bifurcations from simple systems. In [34] Bamón gave a complete characterization of hyperbolic singular cycles of vector fields defined on a manifold of dimension $n \geq 2$ (any n) containing only one singularity which are in the boundary of the Morse-Smale systems. This characterization involves two properties: the weakest contracting/expanding eigenvalue of the singularity in the cycle (such eigenvalue is necessarily real, as it occurs in the cases studied above), and the type of non-transverse intersection between the invariant manifolds of the critical elements in the cycle (this is rigorously stated using the notion of *signature.*)

9.2.3. *Saddle-focus connections.* An important class of singular cycles (which do not correspond to first bifurcations) are those associated to the so called *Shil'nikov loops*: a saddle-focus equilibrium θ having eigenvalues λ and $-\rho \pm i\omega$ where $0 < \lambda < \rho$ such that a branch of the unstable manifold of θ is contained in the stable manifold of θ. In [310] Shil'nikov proved that in any neighborhood of the homoclinic orbit (the branch of the unstable manifold of the equilibrium contained in its stable manifold) there are a infinitely many attracting periodic orbits. This result was generalized by Tresser in [326] proving the existence of infinitely many linked horseshoes in any neighborhood of the homoclinic orbit. The geometrical configuration in this type of cycle also leads Arneodo-Coullet-Tresser to conjecture in [20] the existence of spiral attractors in the unfolding of such a homoclinic orbit. The existence of such attractors was proved by Pacifico-Rovella-Viana in [257]. Before, Pumariño and Rodriguez proved in [288] the existence and persistence of (Hénon-like) strange attractors in the unfolding of such a homoclinic loop (for arcs of vector fields unfolding the loop the parameters corresponding to strange attractors have positive measure). In fact, motivated by the results of Colli in [94] on the existence of diffeomorphisms with infinitely many Hénon-like attractors arbitrarily close to one with a dissipative homoclinic tangency, they also proved in [289] an analogous coexistence result for vector fields exhibiting a Shil'nikov loop.

The Shil'nikov loop is an important motivation for the cycles considered in [307] and [186] by San Martín and Labarca. They considered codimension two singular cycles on three manifolds consisting of two singularities θ and σ, the singularity θ

with three real eigenvalues $\lambda_1 < \lambda_2 < 0 < \lambda_3$ and σ being a saddle-focus contracting singularity (i.e., with eigenvalues λ_1, $\bar{\lambda_1}$ with negative real part and non-zero imaginary part, and $\lambda_3 > 0$), such that $W^u(\theta) \subset W^s(\sigma)$ and $W^u(\sigma) \subset W^s(\theta)$. They consider two-parameter families of vector fields unfolding this type loops and prove result about the prevalence of hyperbolic dynamics in the unfolding of the loop (similar results to those in [35]).

Finally, other important class of singular cycles are the so called *inclination flip* bifurcations. These type of homoclinic loops was first considered by Yanagida in [346], although in this paper it is not used this terminology, introduced by Deng in [104]. Given an equilibrium θ with eigenvalues $\lambda_1 < \lambda_2 < 0 < \lambda_3$ a homoclinic orbit γ of θ is an *inclination-flip orbit* if the unstable and the center unstable (tangent to the eigendirections associated to λ_2 and λ_3) manifolds of θ are tangent along γ. We will see latter that (in some situations) the unfolding of such cycles leads to geometric Lorenz attractors. In [105] Deng gave a mechanism through which a perturbation of a vector field with an inclination-flip homoclinic orbit would exhibit a Smale horseshoe. This result was extended in [159], where it is also analyzed the annihilations (and its associated bifurcations) of such cusp horseshoes. Latter, in [239] Naudot proved that the existence of Hénon-like attractors in the unfolding of inclination-flip bifurcations. Finally, in [225] Morales-Pacifico studied these cycles from the point of view of prevalence of hyperbolicity, obtaining results similar to those in [35].

9.2.4. *Singular attractors through singular cycles.* Many authors have tried to describe bifurcation mechanisms leading to robust singular attractors.

In [305] Rychlik gave an example of a vector field having a singularity with two inclination-flip homoclinic orbits (the two branches of the unstable manifold of an equilibrium) such that the generic unfolding of this codimension two bifurcation leads to Lorenz attractors different [13] from the ones in the geometric model. Let us observe that in [330] Ushiki, Oka and Kokubu claimed that if the original Lorenz equation has an attractor (as nowadays we know that it occurs) then bifurcations like the ones in [305] lead to attractors. The Rychlik's model was modified by Robinson in [294] to obtain cubic differential equation with Lorenz-type attractors, see also the results by Dumortier, Kokubu and Oka in [125] about the generation of Lorenz attractors in the unfolding of certain degenerate singularities.

On the other hand, Morales, Pacifico and Pujals recently present different types of singular cycles whose unfolding lead to new geometrical models of singular attractors which are not Lorenz-like. Their results describe how robust singular strange attractors arise out from simple bifurcations of structurally stable systems. Roughly speaking, one of the reasons why the bifurcation in the singular horseshoe does not lead to attractors is that *only* one separatrix of the singularity is involved in the cycle (compare with the singularity in the geometric Lorenz model). Thus, this gives a restriction on the geometrical configuration of the singular cycles leading to attractors. The new resulting singular attractors we present now will depend on whether the hyperbolic flow is Morse-Smale with periodic orbits, Morse-Smale without periodic orbits, or another type hyperbolic flow.

[13] Meaning non-equivalent. Two attractors Λ and Σ of the vector field X and Y are *equivalent* if there are neighborhoods U and V of Λ and Σ and a homeomorphism $h\colon U \to V$ which maps the orbits of X into the orbits of Y. We say that an attractor is *Lorenz-like* if it is equivalent to a geometric Lorenz attractor.

The first result, [232], states the existence of codimension one submanifolds $\mathcal{N}$ of $\mathcal{X}^r(M)$, $r \geq 3$, locally separating the set of Morse-Smale vector fields and the set of flows exhibiting singular strange attractors. Given an open set $\mathcal{U}$ of $\mathcal{X}^r(M)$ and $\mathcal{N}$ as before let $\mathcal{U} \setminus \mathcal{N} = \mathcal{U}_{ms} \cup \mathcal{U}_{sa}$, where $\mathcal{U}_{ms}$ is the connected component consisting of Morse-Smale systems and $\mathcal{U}_{sa}$ the one of vector fields with singular attractors. The submanifold $\mathcal{N}$ can be chosen in such a way the resulting attractors fail to be Lorenz-like. Thus new class of singular attractors is obtained (roughly speaking) by replacing the hyperbolic regular orbit in the singular cycle in the construction of the singular horseshoe by a saddle-node orbit.

They also considers in [231] this type of cycles in a much more complicated dynamical setting: the saddle node belongs to a nontrivial transitive set. Such bifurcations lead to new classes of singular strange attractors. In this construction the initial vector field X has a local transverse section S homeomorphic to an annulus and a persistent stable foliation. The section S is transverse to the stable manifold of the singularity p in the cycle. Thus the first return map is only defined $S \setminus \ell$, where ℓ is a curve contained in $W^s(p)$ (as in the previous examples, the existence of ℓ carries discontinuities). Moreover, the non-wandering set of X is the union of the singularity p and a transitive set which is equal to the homoclinic class of a regular saddle-node orbit q. Such a point is obtained bifurcating a hyperbolic saddle of a Plykin attractor. An important novelty of this examples is that the one dimensional dynamics induced by the return map and the stable foliation is defined on a branched one-dimensional manifold.

Finally, the third class of singular strange attractors (see [284, 228]) is obtained bifurcating Morse-Smale vector fields via singular cycles without closed regular orbits. This construction involves three singularities, a non-hyperbolic a saddle-node focus singularity σ_0 (having a one dimensional unstable manifold), a two hyperbolic singularities, a saddle-focus σ_1 (with a two dimensional unstable direction) and saddle σ_2 with real eigenvalues and a stable manifold of dimension 2. The singularities σ_0 and σ_1 have a saddle connection (i.e., $W^u(\sigma_0)$ is a branch of $W^s(\sigma_1)$) and

Finally, very surprising is the result in [224, 228] proving that an iterate of the suspension of an expanding attractor of a diffeomorphism of the two disk can be deformed into a robustly singular strange attractor through a one dimensional parameter family of vector field with only one bifurcating parameter. The resulting singular strange attractors are distinct from those described above.

9.3. **Robust transitivity and singular hyperbolicity.** In the previous sections, we saw examples illustrating the fact that for vector fields lack of hyperbolicity may also comes from the accumulation of regular closed orbits on equilibria (besides the coexistence of closed orbits of different orbits in the same transitive set). Nevertheless, as in the case of the geometric Lorenz attractors, these systems behave (in many aspects) as hyperbolic ones. In this section, we will give a precise meaning to this vague notion by introducing the concept of *singular hyperbolicity.* By the moment, this notion is only applicable and works for vector fields of three manifolds. A first challenge is to develop this theory in higher dimensions, the first step being to find an appropriate notion for the role of singular hyperbolicity reflecting the coexistence of closed regular orbits and equilibria in the same transitive set (probably, something like *singular dominated decomposition*).

Consider a closed manifold M of dimension 3 and denote by $\mathcal{X}^r(M)$ the set of C^r vector fields defined over M, $r \geq 1$. Given a vector field X denote by X_t the flow of X at time t. An invariant set Λ is *isolated* if there is a neighborhood U of Λ such that $\Lambda = \bigcap_{t\in\mathbb{R}} X_t(U)$. The set U is an *isolating block* of Λ. The set Λ is *transitive* if it is the ω-limit set of some point of it. Finally, an isolated compact transitive set Λ of a vector field X is *C^r-robustly transitive* if there are neighborhoods U of Λ in the ambient manifold M and $\mathcal{U}$ of X in $\mathcal{X}^r(M)$ such that for every Y in $\mathcal{U}$ the set $\Lambda_Y = \bigcap_{t\in\mathbb{R}} Y_t(U)$ is transitive. These notions are similar to the ones introduced in Section 7 for diffeomorphisms.

The notions of hyperbolicity, partial hyperbolicity and existence of a dominated splitting of a compact invariant set Λ of a vector field X are defined analogously as in the case of diffeomorphisms (see Appendix B), that is, there is $t_0 > 0$ such that the splitting is hyperbolic, partially hyperbolic or dominated for the map X_{t_0}. In the case of vector fields of three manifolds, there are two possibilities for a partially hyperbolic splitting over an invariant set Λ, either $T_\Lambda M = E^s \oplus E^c$, where E^s is one-dimensional and uniformly contracting and E^c is two-dimensional, or $T_\Lambda M = E^c \oplus E^u$, where E^u is one-dimensional and uniformly expanding and E^c is two-dimensional. In both cases, the flow direction is contained in the central bundle E^c. Clearly, if the set Λ does not contain singularities of X the two possibilities may occur simultaneously (for example, consider the suspension of a horseshoe), but this coexistence is not possible if the transitive set Λ simultaneously contains equilibria and closed orbits. In fact, the prototypical examples of partially hyperbolic sets which are non-hyperbolic are geometric Lorenz attractors, in this case near the singularity the angles between the flow and the unstable directions are arbitrarily small (in this case the splitting is of the form $E^s \oplus E^c$, E^s containing the strong stable direction of the singularity). Finally, the notion of volume hyperbolicity (volume expanding and volume contracting) is also defined similarly. We are now ready to state the notion of singular hyperbolicity.

Definition 9.7. A compact invariant set Λ of a vector field X defined on a closed manifold of dimension 3 is *singular hyperbolic* if every singularity of Λ is hyperbolic and there is a partially hyperbolic splitting over Λ having a central direction which is volume hyperbolic (volume contracting if the one-dimensional hyperbolic direction is unstable and volume expanding if otherwise). A *singular basic set* is a singular hyperbolic set which is transitive and isolated.

Let us make some comments about this definition. Clearly, a singular hyperbolic set which does not contain singularities is hyperbolic (that is, an usual basic set). Models of non-hyperbolic singular basic sets are the expanding Lorenz attractors and the singular horseshoes. In these examples, the singular basic set contains exactly one singularity.

Observe that a singular hyperbolic set only contains finitely many equilibria (this follows from the hyperbolicity of the singularities).

Finally, the flow direction is necessarily contained in the central bundle E^c.

9.3.1. *Robust attractors of flows.* In this section we study robustly transitive sets and attractors. As in the case of diffeomorphisms we start with the simplest case of vector fields which are robustly transitive: the transitive set is the whole ambient manifold.

Theorem 9.8 (Doering [116]). *Let X be a C^1-robustly transitive vector field defined on a closed manifold of dimension 3. Then X is Anosov. In particular, X has no singularities.*

A few words about this result. The key point of the proof is to show that there are no singularities. Observe that the closed regular orbits of X are persistently hyperbolic (otherwise, since the Poincaré return map is two dimensional, after an arbitrarily small perturbation one gets a contracting or repelling closed orbit, contradicting the robust transitivity of X in M). In this way, there is a natural splitting over the closed orbits (the one given by the stable and unstable hyperbolic directions). Assuming that the hyperbolic closed orbits are dense in the whole manifold such a dominated splitting can be extended to the whole manifold in a dominated way, see the properties of dominated splittings in Appendix B.

We now explain how the dominated splitting prevents the existence of equilibria. In fact, to prove Theorem 9.8 the main problem is to understand the interaction of equilibria and regular orbits. Next result immediately implies that a robustly transitive vector field of a three manifold does not support equilibria:

Proposition 9.9 (Doering [116]). *Let Λ be a subset of closed manifold M of dimension 3 consisting of regular orbits of a vector field X and σ a singularity of X of saddle-type Suppose that Λ admits a dominated splitting $E^s \oplus E^u$. Then the singularity σ does not belong to the interior of $\Lambda \cup \{\sigma\}$.*

The idea of the proof of this proposition is the when Λ is the complement of the singularities in M (as in our case) is following. Suppose, for instance, that $W^s(\sigma)$ has dimension two and consider a cylinder C transverse to the local stable manifold of the equilibrium σ and orthogonal to the flow direction. Let $C \cap W^s_{loc}(\sigma) = S$ and C^+ and C^- be the connected components of $C \setminus S$. Consider a point $y \in W^u_{loc}(\sigma)$ and a disk D containing y and orthogonal to the flow. After shrinking D we can assume that there is t_0 such that (for instance) the cylinder C^+ is the union of circles $S_t = X_{-t}(D)$, $t \geq t_0$.

Since D does not contain singularities, the fiber E^u_x is well defined for all $x \in D$ (i.e., E^u_x is a continuous bundle over D). Moreover, the cylinder C^+ can be chosen in such a way E^u_x is tangent to C and transverse to S_t for all t. Let D_t the image of S_t by the flow. Then $D \setminus \{y\}$ is foliated by the circles D_t. Finally, for a given $x \in D_t$ the image by the flow of E^u_x is $E^u_{X_t(x)}$ and transverse to D_t. This necessarily implies that E^u has a singularity at y, which is a contradiction.

Clearly Theorem 9.8 can not be generalized to higher dimensions: it is enough to consider the suspension of a non-hyperbolic robustly transitive diffeomorphism on a 3-manifold. There are many questions arising naturally when trying to generalize the previous result to higher dimensions. The first one is that if robustly transitive vector fields defined on a manifold of dimension four can support singularities.

The previous result also leads to consider robustly transitive sets properly contained in the ambient manifold and/or containing singularities. For these sets containing singularities a key point is to understand the dynamical role of the singularity. The main result of the theory of singular hyperbolicity is the following:

Theorem 9.10 (Morales, Pacifico, Pujals [229, 227]). *Let M be a closed compact manifold of dimension 3 and Λ be a C^1 robust singular transitive set of a vector field X of $\mathcal{X}^1(M)$ in an open set U. Then Λ is singular hyperbolic and the singularities of X in Λ have the same index and are Lorenz-like Moreover, Λ is an attractor either for X or for $-X$.*

As a consequence from the proof of this theorem (in dimension three) every robustly transitive set without singularities is uniformly hyperbolic.

It is worthwhile comparing this result with the ones for robust transitive diffeomorphisms in Section 7 (see Theorem 7.2 and the examples in Section 7.3). The difficulty in the case of diffeomorphisms is the coexistence of points of different indices in the same transitive set. Here the main problem is the co-existence of singularities and periodic points in the same transitive set (clearly, since a transitive set does not contains neither sinks nor sources, in dimension three all the regular closed orbits are hyperbolic of saddle-type and thus with the same index). In fact, the first easy step of the proof of Theorem 9.10 is to see that the periodic orbits are hyperbolic of saddle-type (otherwise, after an arbitrarily small perturbation one gets either an attracting or repelling periodic orbit in U, contradicting the robust transitivity of Λ).

Let us explain some ideas of the proof of Theorem 9.10. The first step is the local analysis of the singularities, proving that the singularities have the same index and are Lorenz-like. Consider a singularity $\sigma \in \Lambda$ and let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of $DX(\sigma)$. Clearly, the singularities in Λ are of saddle-type. The first step is to see that the eigenvalues are real. Suppose that there is a singularity $\sigma \in \Lambda$ with a pair of non-real eigenvalues (say of modulus less than one). Using the Connecting Lemma (see Theorem A.5) one gets a singular cycle associated to σ corresponding to the so called *Shil'nikov loop.* The unfolding of such a cycle leads to the creation of sinks or sources (according to the sum of the eigenvalues), for details see, for instance, [288, page 30].

Thus all the singularities of Λ have real eigenvalues with multiplicity one. We need to prove that every singularity σ of Λ is Lorenz-like. Suppose that $\lambda_1 < \lambda_2 < \lambda_3$. Clearly, $\lambda_1 < 0 < \lambda_3$. Suppose that $\lambda_2 < 0$, we now prove that then $\lambda_2 + \lambda_3 > 0$, thus the singularity is Lorenz-like. Again using the connecting lemma one gets a singular cycle associated to σ. We can assume (after a small perturbation) that $W^u(\sigma)$ is a center stable manifold of σ (transverse to the strong stable direction). Let γ be the branch of $W^u(\sigma)$ involved in the cycle. This loop is normally hyperbolic (with normal direction corresponding to the strong stable direction of the singularity). Thus we can think of the cycle as unfolded in a surface N transverse to the strong stable direction. If $\lambda_2 + \lambda_3 < 0$ the unfolding of the singular cycle creates a sink in N, which (due to the normal contraction) gives a sink in the ambient manifold M, which is a contradiction. Thus $\lambda_2 + \lambda_3 > 0$ (the case $\lambda_2 + \lambda_3 = 0$ also leads to sink after new perturbations) and the singularity is Lorenz-like.

The next step is to prove the following:

Proposition 9.11. *All the singularities of the set Λ have the same index.*

This proposition implies that all the singularities of Λ are hyperbolic: otherwise bifurcating a non-hyperbolic singularity of Λ one gets either two singularities of saddle-type having different indices or an attracting or repelling singularity, in both cases we get a contradiction.

To prove this proposition the authors state (sufficient) conditions for an isolated set of a C^1-vector field to be an attractor. The first step is the following proposition (which holds in any C^r-topology) motivated by the properties of hyperbolic attractors):

Proposition 9.12. *Let Σ be a transitive set of X having an isolating block U such that every critical element of X in Σ is hyperbolic. Suppose that there is a critical* *element $\sigma \in \Sigma$ such that for every vector field Y close to X the unstable manifold of the continuation of σ for Y is contained in U. Then Σ is an attractor.*

From this proposition one deduces the following:

Proposition 9.13. *Let Λ be a robustly transitive non trivial set of X in an open set U defined on a manifold of dimension $n \geq 3$. Suppose that Λ contains a singularity σ whose unstable manifold is one dimensional. Then Λ is an attractor.*

In view of Proposition 9.12 this proposition is a consequence of the next lemma.

Lemma 9.14. *Let Λ be a robustly transitive set of a vector field X in U. Suppose that Λ has a singularity σ having a unstable manifold of dimension 1. Then $W^u(\sigma)$ is contained in Λ.*

To prove the lemma let $W^u(+)$ and $W^u(-)$ be the two branches of the unstable manifold of σ. The Connecting Lemma implies that at least one branch of $W^u(\sigma)$, say $W^u(+)$, is contained in Λ and thus in U. Moreover, after a new application of the Connecting Lemma, one can assume that the branch $W^u(+)$ is contained in $W^s(\sigma)$, so σ has a homoclinic loop.

Suppose that $W^u(-)$ is not contained in U. Then there are $q \in W^u(-)$, a neighborhood V of q and $t > 0$ such that $Y_t(V)$ is in the complement of U for all Y close to X. Unfolding now the homoclinic loop of σ one can make the branch $W^u(+)$ passes through V for some Y arbitrarily close to X. Thus the vector field Y is such that the two branches of σ are not contained in U, which is a contradiction.

We are now ready to prove Proposition 9.11. Suppose that Λ contains two singularities with different indices, say σ and θ such that $W^u(\sigma)$ and $W^s(\theta)$ have both dimension one. Since Λ is robust transitive for X and $-X$, applying Lemma 9.14 to σ and X one gets that $W^u(\sigma)$ is contained in Λ, thus by Proposition 9.13 Λ is an attractor. Repeating the argument with θ we get that Λ is an attractor for $-X$, thus a repeller of X. Thus Λ simultaneously is an attractor and a repeller. The final step is to see that if Λ simultaneously is attracting and repelling then it is the whole manifold. This step uses the results in [162] to prove that Λ simultaneously is open and closed. Thus $\Lambda = M$ and X is robustly transitive in the whole manifold. But Theorem 9.8 implies that X has no singularities, which is a contradiction.

In view of the previous results one can assume that every singularity σ of Λ is Lorenz-like and has (for instance) two negative eigenvalues. Consider the splitting with three directions defined over the set of regular closed orbits of the vector field, denoted by $E^{ss} \oplus E^X \oplus E^{uu}$, where E^{ss} and E^{uu} are the stable and unstable directions and E^X is the flow direction. For simplicity let us assume that the set Λ is the closure of the periodic orbits of X. The objective is to extend this splitting to the closure of the regular closed orbits to get a dominated splitting. In general this is not possible, due to the presence of the equilibria, think for instance in the case of the geometric Lorenz attractor, where the angle between the flow direction and the unstable direction is arbitrarily small close to the singularity. Thus it is not possible to extend this splitting in a dominated way to the whole Λ.

The idea now is to consider the splitting $E^{ss} \oplus E^c$, where $E^c = E^{uu} \oplus E^X$, defined over the periodic orbits and extend it to its closure (that is the whole set Λ). For that it is enough to prove that such a splitting is dominated (recall the *extension to the closure property* of the dominated splittings in Appendix B). To obtain the domination one proves that given any small neighborhood V of the finite set of singularities of Λ and any point p in a regular closed orbit that does not meet this vicinity it holds that the angle between $E^{ss}(p)$ and $E^c(p)$ is uniformly bounded from below. Otherwise, after a perturbation and by using the connecting lemma, one gets a homoclinic tangency associated to a regular closed orbit. The unfolding of such a homoclinic tangency leads to the creation of either a sink or a source in U, contradicting the robust transitivity of Λ in U. The fact that the angles between E^{ss} and E^c are persistently bounded away from zero is the key to prove the domination of the splitting $E^{ss} \oplus E^c$ (recall the 2-dimensional argument in Section 7.1.1).

On the other hand, the analysis near a singularity σ gives the following. Since all the singularities have one dimensional unstable manifolds and are Lorenz-like, there is defined a splitting $E^{ss}(\sigma) \oplus E^s(\sigma) \oplus E^{uu}(\sigma)$. Then it holds that $E^{ss}(p) \to E^{ss}(\sigma)$ and $E^c(p) \to E^s(\sigma) \oplus E^{uu}(\sigma)$ as p approaches the singularity σ.

To prove that the splitting is partially hyperbolic (i.e., the one dimensional bundle E^{ss} is uniformly contracting) follows by using the Mañé's ergodic closing lemma as it is illustrated in Appendix A.2. The contradiction being that if there are points in regular orbits of Λ whose averages of contraction in E^{ss} go to 1^- then, after a perturbation, one gets a regular closed orbit which is expanding in the direction E^{ss}, the domination of $E^{ss} \oplus E^c$ implies that such an orbit is expanding, which is forbidden. The proof of the fact that E^c is volume expanding follows along the same ideas.

Finally, it leaves to prove that the set Λ is an attractor. Otherwise there is a sequence of points x_n converging to some point $x \in \Lambda$ and of numbers t_n such that $X_{t_n}(x_n) \in \bar{U}$ ($t_n \to \infty$ as $n \to \infty$). Let us get a contradiction when the point x is a singularity $\sigma \in \Lambda$ (the other cases follows arguing similarly and will be omitted here). By Lemma 9.14 the two branches of σ are (robustly) contained in Λ. Proposition 9.12 implies that Λ is an attractor. This completes our sketch of the proof of Theorem 9.10.

We finish this section with some comments about robustly transitive sets. In [72] Bowen proved that basic sets of Axiom A vector fields without cycles of generic vector fields are mixing. Recall that an invariant set Λ is *mixing* for X if for every pair of open sets U and V of the ambient manifold intersecting Λ there is $t = t(U, V)$ such that $X_s(U \cap \Lambda) \cap V \neq \emptyset$ for all $s > t$. Clearly, mixing implies transitive. The converse is false: the (constant time) suspension of a transitive diffeomorphism is a transitive vector field which fails to be mixing. The Bowen's result means that the property of being a constant time suspension can be easily destroyed by small perturbations.

Vector fields corresponding to suspensions have global cross sections (the converse is false). In [72] Bowen also studied Anosov flows having global cross sections and gave a topological condition on this section guaranteeing the mixing property. In fact, in [69] characterizes the fact of a basic set Λ of an Axiom A vector field without cycles, being not mixing is equivalent to the following: there is $\tau > 0$ such that $t = n\,\tau$, $n \in \mathbb{Z}$ whenever $X_t(x) = x$ for some $x \in \Lambda$.

Using the Bowen's characterization of mixing and Theorem 9.10, Morales and Pacifico showed in [226] that, for every 3-dimensional closed manifold M, there is a residual subset $\mathcal{R}(M)$ of $\mathcal{G}^1(M)$ (the C^1-interior of the vector field with hyperbolic critical elements) of vector fields whose non-trivial attractors are mixing. To prove this result the authors analyzes the dynamics inside a singular attractor and prove that generically a singular attractor Λ is the closure of the one dimensional unstable manifold of any singularity $\sigma \in \Lambda$ (see [226]). In fact, in [226] it is also proved that non-trivial hyperbolic attractors of generic vector fields are mixing (any dimension $n \geq 3$ and any C^r-topology).

9.4. **Consequences of singular hyperbolicity.** Let us also briefly discuss the relationship between robust transitivity, singular hyperbolicity and singular basic sets. On the one hand, every (non-singular) basic set is robustly transitive. On the other hand, a basic singular set may be neither robust transitive nor the closure of its critical elements. This can be obtained by using a Cherry flow, see [91] and [261, pages 181-187], as follows. Recall that a Cherry flow is defined by a vector field in T^2 having exactly two hyperbolic singularities (a sink S and a saddle P) and admitting a global transverse section Σ (a circle being a parallel of the torus) such that every regular orbit intersects Σ. Let Λ be the complement of the stable manifold of the sink S in T^2. Consider the first return map of the flow to the section Σ obtaining a circle map. In fact, one can construct an one-parameter family of vector fields $X^\mu(\cdot)$ such that Σ is a global transverse section for each $X^\mu(\cdot)$. For each μ there is defined a circle map $F_\mu \colon \Sigma \to \Sigma$ with rotation number $\rho(\mu)$. The construction can be done in such a way $\rho(\cdot)$ is not constant at say $\mu = 0$. Thus there are infinitely many parameters μ_1 and μ_2 accumulating on 0 such that $\rho(\mu_1)$ is rational and $\rho(\mu_2)$ is irrational. Observe that for μ_1 we can suppose that the set $\Lambda(\mu_1)$ (defined in the obvious way as the complement of the stable manifold of the sink) is the union of the saddle singularity P and finitely many regular hyperbolic closed orbits. For μ_2 the set $\Lambda(\mu_2)$ is a transitive set without regular closed orbits. We now embed the Cherry flows $X^\mu(\cdot)$ in a manifold of dimension 3 in such a way the initial torus T^2 is a normally (contracting) hyperbolic submanifold for the resulting one parameter family of vector fields denoted by $Y^\mu(\cdot)$. For $\mu = \mu_2$ (i.e., $\rho(\mu_2)$ is irrational) the set $\tilde\Lambda(\mu_2)$ (the invariant set corresponding and homeomorphic to $\Lambda(\mu_2)$) does not contain periodic orbits, thus it is a singular basic set which is not the closure of their critical elements. Moreover, the set $\tilde\Lambda(\mu_2)$ is transitive, but it fails to be robustly transitive: it is enough to perturb the vector field in order to obtain μ_1 such that $\rho(\mu_1)$ is rational (in this case the set $\tilde\Lambda(\mu_1)$ only consists of finitely many regular closed orbits and the saddle singularity).

Other pertinent comment concerns the *implosion and explosion* of singular basic sets. As we have seen in Section 9.2, in many cases the vector fields having a singular hyperbolic set form a codimension one submanifold in the space of vector fields which locally separates the space into two components, a component consisting of vector fields having a very simple dynamics (for instance, Morse-Smale), this component corresponding to the implosion of the singular basic set, and other component, corresponding to an explosion, containing non-hyperbolic vector fields displaying nontrivial transitive singular hyperbolic sets. It is worthy to compare these results with the ones for usual homoclinic classes (which in many cases are basic sets) that only can explode (recall that homoclinic classes are defined as the closure of the transverse intersections between invariant manifolds, which are persistent, thus their implosion is forbidden).

A final comment is the following, if Λ is a singular hyperbolic set of X then there is a neighborhood U of Λ such that for every vector field Y close enough to X any invariant subset of Y in U is hyperbolic (if the set does not contain the singularity) or singular hyperbolic (if otherwise). This follows from the existence of a dominated splitting in the set U, see Properties of Dominated splittings in Appendix B, and the continuation of the hyperbolic singularities.

The examples in previous sections lead us to consider singularities attached to invariant sets having infinitely many regular orbits. Let σ be a singularity of a vector field X having an invariant compact set Λ. The singularity σ is *attached* to Λ if it is accumulated by regular periodic orbits of Λ (in particular, the singularity is contained in Λ). In the examples above the invariant set contains a unique singularity, but it is possible to construct examples of singular basic sets containing a (finite) arbitrarily big number of singularities (see **complete**). But we will see in these cases all the singularities have the same number of expanding eigenvalues counted with multiplicity. Moreover, the stable manifolds of the singularities have the same dimension and their eigenvalues λ_1, λ_2 and λ_3 are all real and verify $\lambda_3 < \lambda_2 < 0 < -\lambda_2 < \lambda_1$ (or $\lambda_3 < -\lambda_2 < 0 < \lambda_2 < \lambda_1$, according to the case) as in the case of the geometric Lorenz attractor (roughly speaking, this condition means normal hyperbolicity). This motivates the following definition, a singularity σ of a vector field X is *Lorenz-like for the vector field X* if $DX(\sigma)$ has three different real eigenvalues $\lambda_1, \lambda_2, \lambda_3$ satisfying $\lambda_3 < \lambda_2 < 0 < -\lambda_2 < \lambda_1$. A singularity is *Lorenz-like* if it is Lorenz-like either for X of for $-X$. Observe that, by definition, the strong stable or strong unstable manifold of a Lorenz-like singularity is defined (according to the different cases).

In view of the previous examples of (non-hyperbolic) singular basic sets, an important problem is to determine whether (transitive) singular hyperbolic sets may have attached singularities with different indices and the type of these singularities, as well as the way in which the equilibria are inserted in the singular set:

Theorem 9.15 (Morales, Pacifico, Pujals [230]). *Consider a vector field X defined on a 3-dimensional closed manifold M. Let Λ be a singular hyperbolic set of X which is not hyperbolic. Then the set Λ contains at least one (attached) singularity and every singularity of Λ is Lorenz-like. Finally, if $W^s(\sigma)$ has dimension two then $W^{ss}(\sigma) \cap \Lambda = \{\sigma\}$ and if $W^u(\sigma)$ has dimension two then $W^{uu}(\sigma) \cap \Lambda = \{\sigma\}$.*

Motivated by the examples above of singular hyperbolic sets with attached singularities in [230] is proposed the following definition:

Definition 9.16. A vector field X is *Singular Axiom A* if the non-wandering set $\Omega(X)$ of X is the closure of the critical elements of X and $\Omega(X)$ is the disjoint union of finitely many singular basic sets. We call this union *singular spectral decomposition* of $\Omega(X)$.

The *no-cycle condition* for singular Axiom A vector fields is defined as in the hyperbolic case: let $\Lambda_1, \dots, \Lambda_m$ be singular basic sets of the spectral decomposition of $\Omega(X)$. We say that $\Lambda_1, \dots, \Lambda_m$ are involved in a *cycle* if for each $i = 1, \dots, m$ there is a regular orbit whose α-limit is contained in Λ_i and whose ω-limit is contained in Λ_{i+1}, where for $i = m$ we replace $m + 1$ by 1.

Define $\mathcal{G}^r(M)$ as the interior of the $\mathcal{C}^r$-vector fields X of M such that every critical element of X is hyperbolic. In the case of diffeomorphisms and when $r = 1$ the corresponding set, denoted by $\mathcal{F}^1(M)$, is the set of Axiom A diffeomorphisms without cycles (see the results by Aoki and Hayashi, [12] and [149]). Since the vector fields with geometrical Lorenz attractors have only hyperbolic critical elements, they belong to $\mathcal{G}^r(M)$ and this characterization can not be translated straightforwardly for flows. Clearly, the first natural step for characterizing the set $\mathcal{G}^1(M)$ is to understand the singular hyperbolic sets (or the robustly transitive sets). Next result is a first attempt into this characterization for vector fields defined on three manifolds. In higher dimensions very little is known about $\mathcal{G}^1(M)$ and there are examples of Lorenz-like attractors having any number of expanding directions and only one singularity, [65]. Other problem in this setting in higher dimensions is the possible coexistence of singularities having different index in the same singular attractor.

Theorem 9.17 (Morales, Pacifico, Pujals [230])**.** *Let M be a closed manifold of dimension* 3. *Then every C^r singular Axiom A vector field without cycles belongs to $\mathcal{G}^r(M)$ for all $r \geq 1$.*

The converse of this theorem is not true. For instance, the vector fields with expanding singular cycles in [35] are counter-examples. Such vector fields belong to $\mathcal{G}^1(M)$: their critical elements are persistently hyperbolic, recall that in such a expanding case the periodic orbits are always hyperbolic (elementary bifurcations are forbidden). But the critical elements of for the vector field X with the cycle fail to be dense in the non-wandering set: the branch of the unstable manifold of the singularity is in $\Omega(X)$. This led Morales, Pacifico and Pujals to conjecture that *singular Axiom vector fields are open and dense in $\mathcal{G}^1(M)$,* [230]. Related to this conjecture, Morales and Pacifico recently proved in [226] that generically in $\mathcal{G}^1(M)$ (M a closed manifold of dimension 3) non-trivial attractors are either hyperbolic or singular hyperbolic.

Let us say a few words about the proofs of Theorems 9.15 and 9.17 above. We begin by stating an important property of singular hyperbolic sets:

Hyperbolicity of proper subsets of singular hyperbolic sets: *Every proper compact invariant subset without singularities of a singular hyperbolic set is uniformly hyperbolic. In particular, every critical element of a singular hyperbolic set is hyperbolic.*

This property implies almost immediately that every singular hyperbolic set Λ which is not hyperbolic contains at least one singularity. Let Σ be the elements of Λ which are not singularities. Given a singularity $\sigma \in \Lambda$ there are two possibilities, either σ is in the closure of Σ, thus σ is attached to Λ, or σ does not belong to the closure of Σ. Let $\sigma_1, \dots, \sigma_k$ the singularities of Λ which are not attached to Σ and observe that $\Lambda = \{\sigma_1, \dots, \sigma_k\} \cup \overline{\Sigma}$. Since Λ is not hyperbolic the same holds for $\overline{\Sigma}$. Thus $\overline{\Sigma}$ contains at least one singularity (otherwise, by the property above $\overline{\Sigma}$ is hyperbolic).

We now prove that the singularities are Lorenz-like. We claim that if Λ has a splitting of the form $E^s \oplus E^c$ (E^s of dimension one) then the stable manifold of any singularity attached to Λ has dimension 2. Suppose that there is a singularity σ attached to Λ having only one contracting eigenvalue (i.e., $W^s(\sigma)$ is one dimensional). This implies that the regular orbits of Λ accumulate on some point q of $(W^s(\sigma) \setminus \{\sigma\})$, thus the branch of $W^s(\sigma)$ containing q is contained in Λ (i.e., Λ contains at least one branch of $W^s(\sigma)$).

Recall that E^c contains the flow direction. Consider the point q in this branch and observe that (since the singularity σ has two expanding eigenvalues) every vector which is not tangent to the flow direction at q is eventually expanded by DX_t, thus the stable direction at q is trivial which is a contradiction.

Thus we can assume that $DX(\sigma)$ has (exactly) two contracting eigenvalues. The fact that E^s has dimension one immediately implies that these eigenvalues are real. Finally, the volume expansion implies that the singularity is Lorenz-like (and thus we can define the strong stable manifold of σ).

To see that the strong stable manifold of σ does not meet Λ one argues as above. Suppose that there is $q \in W^{ss}(\sigma) \cap \Lambda$. Since the flow direction is contained in E^c the angle between $X_t(q) \in E^c$ and $E^s(X_t(q))$ goes to 0 as t goes to $+\infty$, which is a contradiction.

Finally, to prove Theorem 9.17 one writes $\Omega(X) = \Lambda_1 \cup \cdots \cup \Lambda_m$ where each Λ_i is a singular basic set. For each set Λ_i there is a neighborhood U_i where the maximal invariant set of every vector field close to X is singular hyperbolic. The proof ends by using a filtration.

Concerning the structure of the non-wandering set of C^1-vector fields recently Morales and Pacifico have announced the following result extending for vector fields of 3-manifolds the generic dichotomy between hyperbolicity and existence of infinitely many periodic attractors or repellers obtained by Mañé for C^1-surface diffeomorphisms, see [206] **theorem of mane deserves a precise reference and formulation in the text**

Theorem 9.18 (Morales, Pacifico)**.** *Let M be a closed manifold of dimension 3. There is a residual subset $\mathcal{R}$ of $\mathcal{X}^1(M)$ consisting of vector fields X such that either X is a singular Axiom A vector field with no-cycles of X has infinitely many regular attracting or repelling orbits.*

9.5. **Multidimensional Lorenz attractors.** The situation in higher dimension is much more complicated, and this theory is almost completely unexplored. Extra difficulties appear in a natural way, let us mention two of them: the occurrence of cycles containing several singularities, and the fact that the cycle may contain regular closed orbits of different Morse index, as in the case of the heterodimensional cycles in Section 6.

Let us now talk about robust attractors of flows in dimension bigger than 3, where the theory is presently much more incomplete. Flows with robust singular attractors can be constructed in any dimension, in a sort of trivial way: the attractor lies on a transversely contracting 3-dimensional submanifold. However, for some time it remained unknown whether truly high dimensional Lorenz-like attractors (the equilibrium point should have more than one expanding direction) do exist. This problem was solved, affirmatively, in [65]:

Theorem 9.19 (Bonatti, Pumariño, Viana [65])**.** *Given $k \geq 2$ and $n \geq k+3$, there exists a manifold M of dimension n, and there exists a C^∞ flow on M, exhibiting a robust attractor that contains a hyperbolic equilibrium with unstable dimension equal to k.*

Moreover, these attractors support a unique SRB measure, with the basin property. Contrary to the classical geometric Lorenz models, they have *infinite modulus of structural stability*: the set of classes of topological equivalence in a neighborhood of the flow can not be parametrized using only finitely many parameters. A variation of the arguments extends these conclusions to the case $k = 2$ and $n = 4$. For $k = 2$ an alternative construction has been announced by Shil'nikov, Turaev.

Let us comment on the ingredients of the proof of Theorem 9.19, in the most interesting case $n = k+3$. In simple terms, one tries to construct the flow as a kind

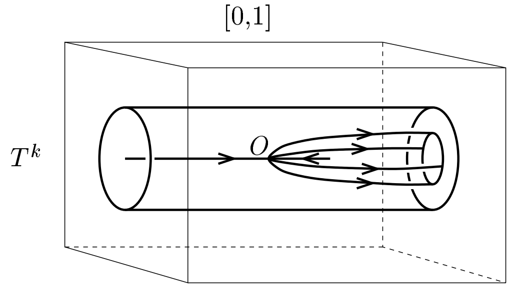

FIGURE 19. Multidimensional singular attractors via suspension

of suspension of a hyperbolic map defined on a manifold Σ of dimension $k+2$, in very much the same way as the classical geometric Lorenz models in dimension 3 may be thought of as suspensions of a piecewise smooth hyperbolic map on a rectangle. As it turns out, obstructions appear in higher dimensions that have no analog in the low-dimensional context. For instance, there are no smooth expanding maps from $D^k \setminus \{0\}$ to the unit k-dimensional disk D^k if $k \geq 2$ (clearly, there are plenty if $k = 1$). To overcome this sort of difficulties cross-sections Σ should have non-trivial topology: in our construction $\Sigma = T^k \times D^2$, where T^k is the k-dimensional torus.

The first step is to take some uniformly expanding map $g : T^k \to T^k$ on T^k, and to consider its natural extension (or inverse limit), realized as a map $G : T^k \times D^2 \to T^k \times D^2$ on the product of T^k by the 2-dimensional disk D^2. For instance, if $k = 1$ then G would be the usual solenoid embedding, see [318]. Next, we suspend G to a flow: this flow lives on the $(k+3)$-dimensional manifold $T^k \times D^2 \times [0,1]/\sim$, where $\sim$ is the equivalence relation generated by $(z,1) \sim (G(z),0)$. Then we deform the flow, so as to create an equilibrium point O with k expanding eigenvalues. This last step, resembling the classical construction of Cherry flows, is illustrated by Figure 19 (for clearness, the D^2 coordinate is not represented).

9.6. **Persistent singular attractors.** More generally than robust attractors, one wants to understand *persistent attractors*, that is, attractors that occur with *positive probability* in generic parametrized families through the initial flow. As usual, probability is understood in the sense of Lebesgue measure in parameter space.

The first examples of persistent non-robust singular attractors were constructed by Rovella [299]. He started from a variation of the classical geometric Lorenz models, assuming that the eigenvalues at the equilibrium satisfy

$$-\lambda_3 < -\lambda_2 < 0 < \lambda_1 < \lambda_2$$

instead. In his setting, an open and dense subset of the parameter space consists of flows with uniformly hyperbolic attractors, while flows with singular attractors occupy a positive Lebesgue measure subset of the complement.

Another class of persistent singular attractors was studied in recent works of Luzzatto and Viana [199], [197], and Luzzatto and Tucker [196]. These papers deal with *Lorenz-like attractors with hooks*, an extension of the classical geometric Lorenz models where one allows for the image of the Poincaré map to fold, as in the figure. This kind of behavior seems to occur in Lorenz equations (27) for slightly larger values of the Rayleigh number $r \approx 30$. It was the observation of this phenomenon

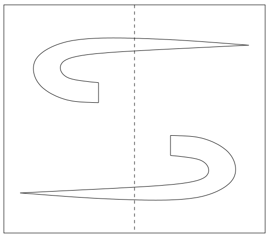

FIGURE 20. Poincaré return map with hooks

that lead Hénon [152] to introduce the map that is named after him, as a model for the dynamics of the Poincaré map near the fold. See Glendinning and Sparrow [137] for an interpretation. The models of Luzzatto, Viana, and Tucker encompass fold-type behavior with the dynamics associated to the presence of equilibria; it is shown that these flows exhibit persistent attractors even after the hooks are formed.

Exhibiting even more complex geometric features, are the *spiraling attractors* as in the figure above: the attractor contains an equilibrium point with two complex contracting eigenvalues, which causes orbits to spiral around the unstable (one-dimensional) separatrix of the equilibrium. This class of systems has been studied by several groups of mathematicians, mostly from a numerical point of view, and a great deal of information is known about their dynamics. Recently, Pacifico, Rovella, Viana [257, 256], proved, for the first time, that such attractors do exist: they occur with positive Lebesgue probability in an open class of one-parameter families of flows.

10. Generic diffeomorphisms

10.1. **Introduction.** The objective of this section is to present a collection of recent results which describe in a rather satisfactory way the dynamics of a large class of systems in the space $\mathrm{Diff}^1(M)$ of C^1-diffeomorphisms of a closed manifold M endowed with the usual C^1-topology. Here we are interested in properties verified by diffeomorphisms in a *residual subset* (i.e., a set containing a countable intersection of dense open subsets) of $\mathrm{Diff}^1(M)$. We call this approach *C^1-generic.* Sometimes our study will be local, in such a case we consider *locally generic sets* of diffeomorphisms, i.e., a set which is residual in a non-empty open set of $\mathrm{Diff}^1(M)$.

The first step for understanding the dynamics of diffeomorphisms is to localize the region of the ambient manifold containing the interesting part of the dynamics, this region depends on the different notions of recurrence one considers. Next, one tries to arrange the global dynamics using the region concentrating all recurrences.

Historically, different notions of recurrence have been considered by the dynamicists, in Section 10.2 we will briefly discuss the following ones: periodic points, recurrent points, limit points, non-wandering points, pre-periodic points (i.e., points that can be made periodic after perturbing the dynamics), and chain recurrent points. Of course, different notions of recurrence lead to different relevant dynamical sets, but we will see that, at least in the generic context, all these notions lead to the same recurrent set (Theorems 10.1 and 10.8).

In the attempt for structuring the dynamics two natural routes appear. The first one is to consider the periodic points, which are generically dense in the recurrent set (Theorem 10.1). A second less obvious way, motivated by the gradient-like dynamics and the spectral decomposition theorem of Smale for hyperbolic diffeomorphisms, is to split the recurrent set into different pieces having independent dynamics. For that one considers filtrations (recall Definition 1.6) guiding the direction of the orbits and "elementary pieces of the dynamics" (of course, to find a good notion for that is a major problem in this theory that we will discuss later).

In recent years, different candidates have been proposed for the role of elementary piece of dynamics. Before introducing these candidates, let us state some properties which are requested (or at least desired) in any attempt for defining such pieces (inspired by the archetypical model of basic piece of dynamics in the Smale's theory):

(1) *undecomposability,* the classical way (in the topological setting) to ensure that the dynamics of a compact set cannot be split into smaller pieces is to impose *transitivity,* that is, the set is the closure of the forward orbit of a point of it;
(2) *maximality* (preserving undecomposibility), the sets are not part of larger transitive sets;
(3) *isolation and independence,* they are independent and each set is isolated from the others, there are no cycles in the dynamics (orbits visiting infinitely many times and successively two different pieces). On one hand, isolation and independence are guaranteed by the existence of a filtration. This filtration guides the dynamics and its levels define neighborhoods where the pieces of dynamics are the maximal invariant sets. Conversely, the notion of isolation immediately suggests (at least in the very archetypical model) the use of filtrations, as in the case of the Smale theory, and it leads to the notion of locally maximal sets.

The concept of transitivity (or at least a some weak form of it) has been a cornerstone in any attempt for generalizing the notion of basic set. Moreover, in the hyperbolic case, transitivity and one of the other two notions above imply the third one. We will see that is not always the case in the non-hyperbolic case. Following the three main principles in the proposal of elementary piece of dynamics, one gets different candidates for this role. Some of them, with their pros and cons, are:

(1) the *maximal robustly transitive sets* (defined as a nonhyperbolic generalization of the basic sets in the Smale theory, see Section 7), these sets verify, by definition, the three prescriptions above;
(2) the *maximal transitive sets,* these sets verify the two first rules, but a priori fail to be isolated, and the *maximal weakly transitive sets,* a type of extrinsic transitivity proposed by Arnaud and by Gan,Wen [136, 18], meaning roughly that it is possible to join two points of the set by segment of orbits of the ambient manifold.
(3) Other interesting approach (in general, neglected by the differentiable dynamicists) is to consider the Conley theory (fundamental in the setting of dynamics of homeomorphisms), which gives a filtration of the dynamics (a priori, with infinitely many levels) such that the relevant dynamics in each level corresponds to a *chain recurrent class* (see Section 10.3.1). This leads

to a new candidate for the role of elementary piece of dynamics, the chain recurrence classes. The main con of this candidate is that, a priori, it only verifies a weak form of transitivity.

Another approach for studying and structuring the recurrent set comes from the very constructive point of view of the periodic orbits. When grouping periodic points in dynamical unities the notion of homoclinic class arises: the homoclinic relation naturally associates to each periodic point a canonical transitive set containing it, its *homoclinic class*. Homoclinic classes appear as a fourth candidate for the role of elementary piece of the dynamics. Transitivity is the main pro of this approach. The cons are that, in general, homoclinic classes fail to be pairwise disjoint (this problem is solved in the generic setting, see Theorems 10.10 and 10.11), and locally maximal (in general, this problem is unavoidable). Moreover, in general, homoclinic classes fail to cover the whole recurrent set (although they are generically dense in it). A psychological pro (not very crucial in our setting) is that in the Axiom A case the homoclinic classes are the basic pieces of dynamics.

Currently and finally, we guess that a good candidate for the role of elementary piece of generic dynamics, unifying all the approaches mentioned above, is available: the *maximal weakly transitive sets.* Using a perturbation connecting-like lemma (Theorem A.7) it is possible to show that, for C^1-generic diffeomorphisms, the chain recurrence classes of the Conley theory are maximal weakly transitive sets. Moreover, these sets form a partition of the non-wandering set and they are separated by a filtration directing the dynamics, see Theorem 10.8.

Concerning homoclinic classes and their importance for organizing the dynamics, using perturbation connecting-like lemmas one can prove that homoclinic classes generically are maximal transitive (also maximal weakly transitive) sets and, also, chain recurrence classes (see Theorems 10.11 and 10.12). Thus, generically, any pair of homoclinic classes either coincide or are disjoint. So one can consider the number of homoclinic classes, which is locally generically constant (Theorem 10.17). Other important property of homoclinic classes of generic diffeomorphisms is their continuous dependence. We say that a generic diffeomorphism is *tame* when it has finitely many homoclinic classes. In the tame case, the homoclinic classes are exactly the elementary pieces of dynamics above.

The previous paragraph gives a rather satisfactory glance of the topological dynamics of tame diffeomorphisms. However, there are also locally generic diffeomorphisms having simultaneously infinitely many homoclinic classes. Recall that in Section 3.2 we presented the so-called Newhouse phenomenon, consisting of locally C^2-generic surface diffeomorphisms having infinitely many sinks or sources (thus infinitely many different homoclinic classes). Although this is a C^2-phenomenon (typically associated to the unfolding of homoclinic tangencies) one can interpretate it as an indication that such a phenomenon can appear in C^1-generic dynamics. Currently, we know that this is the case in manifolds of dimension equal to or bigger than three, see Section 10.6.2. In fact, modifying the available constructions of locally generic C^1-diffeomorphisms having infinitely many sinks or sources, it is possible to construct locally generic diffeomorphisms having infinitely (and uncountably) many non-trivial elementary pieces of dynamics without periodic points, actually, these pieces are maximal transitive sets and chain recurrence classes, (see Theorem 10.26). We point out that the phenomenon of locally generic coexistence of infinitely many chain recurrence classes necessarily implies the coexistence of infinitely many homoclinic classes. In this case we speak of *wild dynamics.*

Finally, we observe that it is an open question if the Newhouse coexistence C^1-phenomenon can be displayed by C^1-generic diffeomorphisms of surfaces. This question also leads to the old key question of the genericity (in fact open and density) of the hyperbolic diffeomorphisms in surfaces.

Returning to the tame diffeomorphisms, in such a case the dynamics is nicely organized by a filtration and the maximal invariant sets in the levels of the filtration are homoclinic classes (which actually are the elementary pieces of dynamics). Moreover, tame diffeomorphisms have finitely many topological attractors (resp. repellors) whose basins contain an open and dense subset of the ambient manifold (Theorem 10.23). An important problem now is to give a good topological and ergodic description of the dynamics in each piece. At least it is known that the elementary pieces are homoclinic classes (thus transitive sets) having some hyperbolicity (volume hyperbolicity), see Theorem 10.22. Finally, in the tame case, the elementary pieces of dynamics obey the three archetypical principles above requested for the notion of elementary pieces (see Section 10.5 for a more detailed study of tame dynamics).

Finally, concerning the wild case, very few is known besides its existence. The main problem is that the existence of infinitely many basic pieces of dynamics is an important obstacle for arranging satisfactorily the dynamics. For instance, comparing with the known results in the tame case, it is an open question if wild diffeomorphisms have at least a topological attractor or repellor. Moreover, assuming that such attractors or repellors exist, it is an open question if the union of their basins contains an open and dense subset of the ambient manifold. Our ignorance concerning the dynamics of each piece of dynamics is almost complete, for example, we do not know if each piece is transitive or they are only weakly transitive. We observe that in the available examples of wild dynamics the aperiodic pieces are transitive. The dynamics of wild diffeomorphimsms will be discussed in Section 10.6.

10.2. **Notions of recurrence.** When studying the dynamics of diffeomorphisms one focus on the region of the manifold concentrating the accumulation and recurrence of orbits. In this approach, the set of periodic points appears in a very evident way and it is enough for describing the dynamics of very simple systems, as in the case of the Morse-Smale systems. For studying more complicated chaotic dynamics, other weaker notions of recurrences have been successively introduced. In this section, we will see that, for C^1-generic diffeomorphisms, all these notions of recurrence define the same (compact) region of the manifold. Let us summarize briefly the most common of these notions:

(1) the periodic points of f;
(2) the *ω and α-recurrent points,* a point is ω or α-reccurent if its forward or backward orbit returns arbitrarily close to it;
(3) the *limit points,* that is, points being accumulated by the forward or backward orbit of some point;
(4) the *non-wandering points,* that is, points x having points arbitrarily close to it whose forward orbits return close to x, in other words, for any neighborhood U_x of x there is $n > 0$ such that $f^n(U_x) \cap U_x \neq \emptyset$;

(5) the *pre-periodic points,* that is points that can be made periodic by an arbitrarily small C^1-perturbation of the dynamics[14];
(6) the *chain recurrent points,* a point x is chain recurrent if it is possible to close its orbit by doing arbitrarily small jumps (at each time) along its orbit, more precisely, for each $\varepsilon > 0$ there is an ε-pseudo-orbit x_i, $i \in \{0, \dots, n\}$, starting and ending at x, $x_0 = x_n = x$, and such that $d(f(x_i), x_{i+1}) < \varepsilon$ for all i.

By definition, every periodic point is recurrent, every recurrent point is a limit one, any limit point is non-wandering, and every non-wandering point is pre-periodic (this follows from the Pugh's closing lemma, Theorem A.1), and every pre-periodic point is chain recurrent.

The previous notions lead to the definitions of the sets of periodic points $\mathrm{Per}(f)$, recurrent points $\mathrm{Rec}(f)$, limit points, nonwandering points $\Omega(f)$, pre-periodic points $\mathrm{Per}_*(f)$, and chain recurrent points $\mathcal{R}(f)$. All these sets are f-invariant and, in general, different. The non-wandering, pre-periodic and the chain recurrent sets are compact. This is not the case, in general, of the sets of periodic, recurrent and limit points, so that one usually considers the closure of them. The closure of the limit points if the *limit set,* denoted by $L(f)$.

Using Pugh's closing lemma (see Theorem A.1) and a Kupka-Smale argument, one gets, by a standard argument of genericity, the following generic density theorem:

Theorem 10.1. *There is a residual subset R_0 of $\mathrm{Diff}^1(M)$ such that, for every $f \in R_0$,*

- *the non-wandering and the limit sets of f are equal and coincide with the closure of the periodic points of f,*
- *every periodic point of f is hyperbolic,*
- *invariant manifolds of periodic points meet transversely.*

In other words, for C^1-generic diffeomorphisms f, the following holds:

$$\overline{\mathrm{Per}(f)} = \overline{\mathrm{Rec}(f)} = L(f) = \Omega(f) \subset \mathrm{Per}_*(f) \subset \mathcal{R}(f).$$

A connecting lemma for pseudo-orbits (see Theorem A.7) now implies that any pre-periodic point can be turned to be periodic by an arbitrarily small C^1-perturbation of the diffeomorphism, that is, for general diffeomorphisms, it holds $\mathrm{Per}_*(f) = \mathcal{R}(f)$.

The same argument shows that, for generic diffeomorphisms f (see Theorem 10.8), one has:

$$\overline{\mathrm{Per}(f)} = \overline{\mathrm{Rec}(f)} = L(f) = \Omega(f) = \mathrm{Per}_*(f) = \mathcal{R}(f),$$

that is, all the regions of recurrence coincide in the C^1-generic setting.

The next step for describing the global dynamics of a generic diffeomorphism is to split the region of recurrence into elementary pieces of dynamics, following the description given, in the hyperbolic setting, by the Smale's Spectral Decomposition Theorem (see Theorem 1.3), generalized by Conley in the very general context of chain recurrent sets of homeomorphisms in [95]. This is the goal of next section.

10.3. Decomposition of the global dynamics into elementary pieces.

[14]The notion of pre-periodic points (with other name) appeared in the work of Liao [193] and played an important role in the study of the C^1-stability conjecture.

10.3.1. *Chain recurrence classes and filtrations.* In the last seventies, Conley generalized the construction of the filtration in the Spectral Decomposition Theorem (Theorem 1.3) in the most general setting for homeomorphisms of compact sets as follows.

Let (X, d) be a compact metric space and f a homeomorphism defined on X. Given $\varepsilon > 0$, an *ε-pseudo-orbit* of f is a sequence x_i of points of X such that the *jumps* $d(f(x_i), x_{i+1})$ are less than ε for all i. The chain recurrent set $\mathcal{R}(f)$ of f is the set of points x admitting closed pseudo-orbits starting and ending at x and having arbitrarily small jumps. Just by definition, the chain recurrent set is an f-invariant compact set containing the non-wandering set. In general, the chain recurrent set is strictly bigger than the non-wandering set: consider, for instance, an Axiom A diffeomorphism with a cycle between its basic pieces, then the heteroclinic points corresponding to two different basic sets in a cycle are wandering and chain recurrent.

Using pseudo-orbits we define a natural binary relation between the points of X. We write $x \dashv y$ if there are (finite and non-trivial) pseudo-orbits with arbitrarily small jumps joining x to y. By definition, a point x is chain recurrent if, and only if, $x \dashv x$. This relation is clearly transitive and, in general, non symmetric. The symmetrization $\vdash\dashv$ of $\dashv$ (that is, $x \vdash\dashv y$ if and only if $x \dashv y$ and $y \dashv x$) induces an equivalence relation on $\mathcal{R}(f)$ whose classes are the *chain recurrence classes*.

Notice that two points in the same chain recurrence class cannot be separated by a filtration (recall Definition 1.6), that is, they always belong to the same level of a filtration. In fact, Conley proved that the converse is also true: points in distinct chain recurrence classes can be separated by a filtration. More precisely:

Theorem 10.2 (Conley [95])**.** *Let (X, d) be a compact metric space and f a homeomorphism defined on X. There is a continuous map $\psi\colon X \to \mathbb{R}$ verifying the following properties:*

- *ψ is decreasing along the orbits of f, that is, $\psi(f(x)) \le \psi(x)$ for all $x \in X$;*
- *$\psi(f(x)) = \psi(x)$ if, and only if, $x \in \mathcal{R}(f)$;*
- *for any pair (x, y) of chain recurrent points, $\psi(x)$ is equal to $\psi(y)$ if, and only if, x and y belong to the same chain recurrence class;*
- *the set $\psi(\mathcal{R}(f))$ has empty interior in $\mathbb{R}$ (i.e. this set is a totally disconnected compact set of $\mathbb{R}$).*

Given real numbers $\alpha_1 > \cdots > \alpha_n$ in $(\mathbb{R} \setminus (\psi(\mathcal{R}(f))))$ one gets a filtration associated to f by considering the sets $X_i = \psi^{-1}([-\infty, \alpha_i])$. This allows us to separate two fixed distinct classes by a filtration. In fact, in this way, by considering a (may be infinite) family of real numbers α_i such that each connected component of $(\mathbb{R} \setminus (\psi(\mathcal{R}(f))))$ just contains one α_i, one gets a filtration X_i of f (a priori this filtration may have infinitely many levels) separating any two distinct chain recurrence classes of f. We observe that a homeomorphism may have infinitely (even uncountably) many chain recurrence classes.

The previous theorem shows that, if we want to consider pieces of dynamics which are separable by a filtration, the chain recurrence classes are such pieces and cannot be subdivided. However, in many cases this decomposition is very rough as in the following very simple example:

Consider an irrational rotation f defined on the sphere S^2. In this case, the non-wandering set $\Omega(f)$ coincides with $\mathcal{R}(f)$ and is the entire sphere. Moreover, there is a unique chain recurrence class which is the whole manifold. On the other hand, there is a natural finer decomposition of the sphere into pairwise disjoint f-invariant compact sets: the closures of the orbits, which are the parallel circles. In Section 10.3.2 we will introduce another decomposition of the recurrent set, that time of $\Omega(f)$, whose classes, in this simple example, are exactly the parallel circles.

10.3.2. *Maximal weakly transitive sets.* Consider a compact manifold M and a diffeomorphism f defined on M. In [14], motivated by the definition of non-wandering points, Arnaud introduced the following binary relation between points of M: $x \prec y$ if for every pair of neighborhoods U_x of x and U_y of y there is $n > 0$ with $f^n(U_x) \cap U_y \neq \emptyset$[15]. Notice that $x \prec x$ if and only if $x \in \Omega(f)$. Also, $x \prec y$ implies that $x \dashv y$.

In the example of the irrational rotation of S^2 in Section 10.3.1, the relation $\prec$ is an equivalence relation whose classes are exactly the parallel circles. However, in general, the relation $\prec$ is neither transitive nor symmetric.

Theorem 10.3 (Arnaud [18])**.** *There is a residual subset of* $\mathrm{Diff}^1(M)$ *consisting of diffeomorphisms for which* $\prec$ *is a transitive relation.*

This result implies that, for generic diffeomorphisms, the symmetrization $\prec\!\succ$ of $\prec$ induces an equivalence relation on the non-wandering set. The equivalence classes of this relation are the *maximal weakly transitive sets.* A compact f-invariant set Λ is *weakly transitive* if for every pair of points $x, y \in \Lambda$ and every pair of neighborhoods U_x of x and U_y of y, there is $n > 0$ with $f^n(U_x) \cap U_y \neq \emptyset$. Weakly transitive sets were studied by Arnaud in [18] (see also the paper by Gan,Wen [136]). Two very important examples of weakly transitive sets (besides the transitive sets) are the ω-limit $\omega(x)$ of a point x and the Hausdorff limit of a sequence of periodic orbits.

An important difference between the notions of transitivity and weakly transitivity is that the second one is not intrinsic: the definition of weakly transitive set Λ involves orbits of points which do not belong to the set, thus the restriction of a diffeomorphism f to a weakly transitive set Λ of it may fail to be weakly transitive. An easy illustration of this situation (and also and example of a weakly transitive set which is not transitive) is the following, consider a diffeomorphism f defined on the annulus $S^1 \times [0,1]$ such that every circle $S^1 \times \{a\}$ is f-invariant and f has a unique fixed point P, say in the circle $S^1 \times \{1/2\}$, which is a saddle-node, thus the backward orbit of any point of $S^1 \times \{1/2\}$ goes to P and the same for the forward one. It is now immediate to to check that the entire circle $S^1 \times \{1/2\}$ is a weakly transitive set of f defined on $S^1 \times [0,1]$. Clearly, this set is not transitive. Moreover, for the restriction F of f to the circle $S^1 \times \{1/2\}$ one has that the circle is not a weakly transitive of F (although it is a chain recurrent set of F).

Notice that, by definition, the closure of the union of an increasing sequence of weakly transitive sets is a weakly transitive set, so Zorn's Lemma implies that any weakly transitive set is contained in a maximal one. We observe that, in general, maximal weakly transitive sets may fail to be disjoint. Theorem 10.3 asserts that, for generic diffeomorphisms, maximal weakly transitive sets are pairwise disjoint and induce a partition of the non-wandering set.

When the relation $\prec$ is transitive, the partition of $\Omega(f)$ by the maximal weakly transitive sets refines the partition obtained by intersecting the chain recurrence

[15]This notion was also considered by Gan, Wen [136] with the name of *attainability relation*. The symmetrization of this relation is called *bi-attainability*.

classes with $\Omega(f)$. The next step is to understand the relation between chain recurrence classes and maximal weakly transitive sets in the generic setting.

10.3.3. *Elementary pieces of dynamics for generic diffeomorphisms.* In view of the decompositions of the dynamics in the two previous sections, we now propose definitions of elementary piece of dynamics and of diffeomorphism having a decomposition of its global dynamics into elementary pieces. We first define a *filtrating neighborhood U* of an f-invariant compact set Λ as a generalization of the level containing Λ of a filtration of f, that is, the set U is a compact neighborhood of Λ which can be written as $U = (M_1 \setminus \mathrm{int}(M_2))$, where M_1 and M_2 are compact submanifolds with boundary (of the same dimension as M) such that $f(M_i)$ is contained in the interior of M_i, $i = 1, 2$.

Remark 10.4.

(1) If U is a filtrating neighborhood of some compact f-invariant set and x belongs to $\mathcal{R}(f) \cap U$ then the chain recurrence class of x is contained in U.
(2) If Λ is the maximal invariant set in some filtrating neighborhood U, then it admits a basis of filtrating neighborhoods: it is enough to consider the intersections $\bigcap_{-n}^{n} f^i(U)$, $n \in \mathbb{N}$.
(3) Theorem 10.2 implies that any chain recurrence class admits a basis of filtrating neighborhoods.
(4) Every isolated chain recurrence class is the maximal invariant set in some filtrating neighborhood of it (this also follows from Theorem 10.2).
(5) When the diffeomorphism f has infinitely many chain recurrence classes (we will present examples of this situation later), there is a recurrence class Λ which is accumulated by a sequence of classes, thus Λ is not maximal invariant in any neighborhood of it.
(6) A transitive or weakly transitive set Λ admits a basis of filtrating neighborhoods if and only if it is a chain recurrence class: the set Λ is contained in a chain recurrence class Σ by definition, and the existence of arbitrarily small filtrating neighborhoods of Λ implies, by the first item in this remark, that Σ is equal to Λ. Moreover, in this case the set Λ is a maximal weakly transitive set.

The previous remark is the main motivation for the following definition:

Definition 10.5. An f-invariant compact set Λ is an *elementary piece of the dynamics of f* if it is simultaneously weakly transitive and a chain recurrence class.

In view of Remark 10.4, any elementary piece of dynamics of a diffeomorphism f has a basis o filtrating neighborhoods. Moreover, any two different elementary pieces of dynamics Λ and Σ are disjoint and separated by a filtration. To see the last assertion, consider a filtrating neighborhood $(M_1 \setminus \mathrm{int}(M_2))$ of Λ, (with $f(M_i) \subset \mathrm{int}(M_i)$) disjoint from Σ. For each $i = 1, 2$, the class Σ is included in M_1 or disjoint from M_i. One has now that Σ is either contained in M_2 or disjoint from M_1 and M_2. Thus (M, M_2) in the first case or (M, M_1) in the second one is the announced filtration.

Definition 10.6. The global dynamics of a diffeomorphism f admits a *decomposition into elementary pieces* if the non-wandering set $\Omega(f)$ is the union of elementary pieces of dynamics.

Lemma 10.7. *If the global dynamics of a diffeomorphism f admits a decomposition into elementary pieces then $\Omega(f) = \mathcal{R}(f)$. Moreover, in such a case the induced equivalence relations $\vdash\!\dashv$ and $\prec\!\succ$ coincide.*

Proof. It is enough to show that $\mathcal{R}(f) \subset \Omega(f)$. Let y be a point which does not belong to $\Omega(f)$ and consider its forward limit set $\omega(y)$. This set is contained in $\Omega(f)$ and weakly transitive. So $\omega(y)$ is contained in some elementary piece of dynamics Λ of f and, by hypothesis, $y \notin \Lambda$. Consider a small filtrating neighborhood $U = (M_1 \setminus \mathrm{int}(M_2))$ of Λ such that $y \notin U$, that is, $y \notin M_1$ or $y \in M_2$. In the first case, any pseudo-orbit starting at y and having small jumps will enter in M_1 and thereafter it will remain in M_1, so this pseudo-orbit never returns to y, and thus $y \notin \mathcal{R}(f)$. In the second case, the argument is similar. $\square$

In the hyperbolic case (Axiom A and no-cycles condition), the elementary pieces of dynamics of f are exactly the basic sets of the spectral decomposition of $\Omega(f)$, recall Theorem 1.3. Next result asserts that the dynamics of C^1-generic diffeomorphisms have a decomposition into elementary pieces (as we will see in next sections, these pieces may be infinitely many in number and highly complicated dynamically).

Theorem 10.8 (Bonatti-Crovisier [54]). *There is a residual subset of* $\mathrm{Diff}^1(M)$ *of diffeomorphisms f such that the relations $\prec$ and $\dashv$ coincide. In particular, $\Omega(f) = \mathcal{R}(f)$, the chain recurrence classes are the maximal weakly transitive sets, and the global dynamics of f has a decomposition into elementary pieces of dynamics.*

This result follows from a perturbation lemma (Theorem A.7) based on the Connecting Lemma (Theorem A.5): for a generic diffeomorphism f and any pair of points x and y such that $x \dashv y$, there is a diffeomorphism g arbitrarily C^1-close to f such that y belongs to the forward orbit of x.

10.4. **Periodic orbits and elementary pieces.** The homoclinic class of a hyperbolic periodic orbit (the closure of the transverse intersections of its invariant manifolds) is a compact transitive f-invariant set canonically associated to the periodic orbit. Another way of defining homoclinic classes is using the homoclinic relation introduced by Smale. Two hyperbolic periodic orbits are homoclinically related if their invariant manifolds meet transversely (the stable manifold of each orbit meets transversely the unstable manifold of the other one), thus both periodic points have the same Morse index (dimension of the unstable bundle). By the λ-lemma, to be homoclinically related defines an equivalence relation in the set of periodic orbits. The homoclinic class of a hyperbolic periodic saddle p is the closure of all periodic orbits in the equivalence class of the orbit of p.

On the one hand, in general, distinct homoclinic classes may intersect (see for instance [114]) and they do not define a partition of the closure of the periodic points. On the other hand, in the case of Axiom A diffeomorphisms, the homoclinic classes are exactly the basic pieces in the Smale's Theory and (in the terminology above) they are the elementary pieces of dynamics. Thus, it is a natural temptation to try to generalize the notion of basic piece of the Smale's Theory in the generic context by considering homoclinic classes. As we will see, this approach is partially succesfull: the homoclinic class of a periodic orbit of a generic diffeomorphism is the elementary piece of the dynamics containing it. However, as we will see in

Section 10.6, there are locally generic diffeomorphisms having elementary pieces of dynamics without any periodic orbit, thus which are not homoclinic classes.

10.4.1. *Homoclinic classes and maximal transitive sets.* A *maximal transitive set* of a diffeomorphism f is a transitive set which is a maximal element of the family of transitive sets of f ordered by inclusion. Next result lays the foundation of this concept, implying that every transitive set is contained in a maximal one.

Lemma 10.9. *The closure of an increasing family of transitive sets is transitive.*

Proof. Let $(K_i)_i$ be an increasing family of transitive sets and K its closure. To verify that K is transitive it is enough to see that for any pair U, V of non-empty open sets of K there is $n > 0$ with $f^n(U) \cap V \neq \emptyset$. By hypothesis, there is j such that $U \cap K_j \neq \emptyset$ and $V \cap K_j \neq \emptyset$. Thus, since K_j is transitive, $f^n(U) \cap V \neq \emptyset$ for some $n > 0$. □

This lemma and the Zorn lemma imply that any transitive set is contained in a maximal one.

Let us observe that there are examples of different maximal transitive sets with non-empty intersection (see for instance [114]). This is an obstruction for considering the maximal transitive sets as elementary pieces of dynamics. This leads to the notion of *saturated* transitive set[16]: a transitive set is saturated if it contains any transitive set intersecting it. In other words, a saturated transitive set is a transitive class of the equivalence relation generated by the relation *to be in the same transitive set,* that is, two points x and y are related if there are finite families of points $(x_i)_{i=0}^m$ and of transitive sets $(K_i)_{i=0}^{m-1}$ such that $x_0 = x$ and $x_m = y$ and each set K_i contains x_i and x_{i+1}. By definition, every saturated transitive set is maximal. However, there are maximal transitive sets with are not contained in any saturated one (as the maximal transitive sets with non-empty intersection in [114]).

In [57] it was conjectured that, generically in $\mathrm{Diff}^1(M)$, the notions of homoclinic class and maximal transitive set coincide. Currently, we know that this conjecture is partially false. One the one hand, generically, every homoclinic class is a maximal transitive set. On the other hand, there are locally generic diffeomorphisms having maximal transitive sets (in fact, saturated ones and chain recurrence classes) without periodic points. In particular, such maximal transitive sets are not homoclinic classes. We now discuss the first of these results, the second one will be analyzed in Section 10.6.

First, [57] uses the connecting lemma (Theorem A.4) to prove that, generically, two periodic points belong to the same transitive set if and only if their homoclinic classes coincide. Later, using an extension of the connecting lemma in Theorem A.5, Arnaud proved that for generic diffeomorphisms f, the homoclinic class $H(p, f)$ of any saddle p of f is the intersection of the closures of its invariant manifolds $W^s(p, f)$ and $W^u(q, f)$. Note that the inclusion $H(p, f) \subset \overline{W^s(p, f)} \cap \overline{W^u(q, f)}$ is obvious. This acute generic characterization of homoclinic classes is an important step for proving the following:

Theorem 10.10 (Arnaud [18])**.** *Generically in* $\mathrm{Diff}^1(M)$*, every transitive set containing a periodic orbit* p *is contained in the homoclinic class of* p*. In particular, generically, every homoclinic classes is a maximal transitive set.*

[16]This notion is called *maximal transitivity* in [89].

The previous result was generalized by considering saturated transitive sets:

Theorem 10.11 (Carballo, Morales, Pacifico [89]). *Generically in* $\mathrm{Diff}^1(M)$*, every homoclinic class is a saturated transitive set. In particular, two homoclinic classes are either equal or disjoint.*

This theorem is very interesting for our purposes, it shows that homoclinic classes of generic diffeomorphisms f define a partition of a part of the nonwandering set of f.

In fact, [54, Remarque 1.10] gives a slightly stronger version of the previous result which very convenient here:

Theorem 10.12 (Bonatti, Crovisier). *Generically in* $\mathrm{Diff}^1(M)$*, every homoclinic class is a chain recurrence class. Equivalently, every chain recurrence class containing a periodic point* p *is the homoclinic class of* p*.*

For generic diffeomorphisms, the (hyperbolic) periodic orbits are dense in the chain recurrent set. In fact, Crovisier recently proved in [100] a strong version of this density: the periodic orbits accumulate to the chain recurrence classes with some topological regularity.

Theorem 10.13 (Crovisier). *For* C^1*-generic diffeomorphisms, any chain recurrence class is the Hausdorff limit of a sequence of periodic orbits.*

In view of the previous results, one naively could guess that every elementary piece of dynamics should contain a period point and thus (by the previous result) it should be a homoclinic class. In fact, this naive argument holds when there are only finitely many chain recurrent classes or finitely many homoclinic classes, but, as we will see in Section 10.6, this is not the general generic case. We have the following consequences from the generic density of the periodic points in the chain recurrent set:

Corollary 10.14. C^1*-generically, the following holds:*

- *every isolated chain recurrence class is a homoclinic class,*
- *every chain recurrence class with non-empty interior is a homoclinic class,*
- *the set of chain recurrence classes is infinite if and only if the set of homoclinic classes also is infinite.*

Since the set of hyperbolic periodic points is countable, C^1-generically, the number of homoclinic classes is always countable, but a priori, the number of chain recurrence classes may be uncountable (see Section 10.6 for an example of this situation). Thus, a priori, in the last item of the corollary one may have countable many homoclinic classes and uncountable many chain recurrence classes.

Notice that an isolated chain recurrent class of a diffeomorphism is the maximal invariant set in a filtrating neighborhood. For generic diffeomorphisms, an isolated class is a transitive (as being a homoclinic class) and maximal invariant set. It is natural to ask whether such a class is a robustly transitive set (see Section 7), this is an open question, but we can give a partial result in this direction.

An f-invariant compact set Λ is *robustly chain recurrent* if there is an open neighborhood U of Λ and a neighborhood $\mathcal{V}$ of f in $\mathrm{Diff}^1(M)$ such that, for every $g \in \mathcal{V}$, the maximal invariant set Λ_g of g in the closure of U is a chain recurrent set contained in U and $\Lambda_f = \Lambda$.

Corollary 10.15. *For C^1-generic diffeomorphisms, every isolated homoclinic class is robustly chain recurrent.*

We will discuss the role of isolated homoclinic classes in Section 10.5.

10.4.2. *Ingredients of the proofs.* We now explain in a very few words the main ideas involved in the proofs of these results. First, the invariant manifolds of a saddle p and its homoclinic class $H(p,f)$ depend lower semi-continuously on f. Thus, for diffeomorphisms f in a residual set $\mathcal{R}$ of $\mathrm{Diff}^1(M)$, homoclinic classes depend continuously on f. Consider now a point x in the intersection of the closures of $W^s(p,f)$ and $W^u(p,f)$. Using Theorem A.5, one creates a transverse intersection between the invariant manifolds of p at x (this implication is straightforward if x is not periodic, and much more subtle if x is periodic). As the homoclinic class $H(p,f)$ depends continuously on $f \in \mathcal{R}$, the point x belongs to $H(p,f)$. This shows that, generically:

$$\overline{W^s(p,f)} \cap \overline{W^u(p,f)} = H(p,f).$$

The remainder of the proof involves the notion of Lyapunov stability. An f-invariant set K is *Lyapunov stable* if for every neighborhood U of K there is a neighborhood V of K such that $f^n(V) \subset U$ for every $n \geq 0$. The next step is to prove that,

for generic diffeomorphisms f and any saddle p of f, the closure of the unstable manifold of p is Lyapunov stable and the closure of the stable manifold of p is Lyapunov stable for f^{-1}.

The proof of this claim goes as follows. Consider a diffeomorphism f in the residual set $\mathcal{R}$ above, a saddle p of f, and a neighborhood U of the closure of $W^u(p,f)$. Arguing by contradiction, assume that $\overline{W^u(p,f)}$ is not Lyapunov stable. Thus there is a sequence (x_n) of points converging to some $x \in \overline{W^u(p,f)}$ whose forward orbits escape from U. Considering the orbits of the points x_n one can get a new sequence converging to some point $q \notin U$. Theorem A.5 now implies (once more, directly if x is not periodic) that, after a perturbation of f, the point q is in the closure of $W^u(p,f)$. But the continuity of $W^u(p,f)$ for $f \in \mathcal{R}$ prevents such a possibility.

Finally, consider any transitive (or weakly transitive) set K intersecting $H(p,f)$, $f \in \mathcal{R}$. The Lyapunov stability of $\overline{W^u(p,f)}$ implies that K is contained in any neighborhood of $\overline{W^u(p,f)}$. Hence, $K \subset \overline{W^u(p,f)}$. Using the same argument for the stable manifold one has

$$K \subset \overline{W^u(p,f)} \cap \overline{W^u(p,f)} = H(p,f).$$

This ends our discussion of Theorems 10.10 and 10.11.

Finally, to prove of Theorem 10.12 just observe that by Theorem 10.8, for generic diffeomorphisms, the chain recurrence classes are maximal weakly transitive sets.

10.5. **Wild versus tame behavior.** In the previous section, we saw that there is a residual subset $\mathcal{R}$ of $\mathrm{Diff}^1(M)$ where the homoclinic classes are chain recurrence classes (also elementary pieces of dynamics) and depend continuously on the diffeomorphism. In particular, for diffeomorphism in $\mathcal{R}$, any pair of homoclinic classes either coincide or are disjoint. Thus the homoclinic classes define a partition of a part of the limit set (in general, as we will see in Section 10.6, the homoclinic classes do not cover the whole limit set).

In dimension 2, Mañé showed that there is a nitid C^1-generic dichotomy between diffeomorphisms having finitely and infinitely many homoclinc classes:

Theorem 10.16 (Mañé [206])**.** *Given any closed surface M there is a residual subset $\mathcal{R}$ of $\mathrm{Diff}^1(M)$ such that, for every $f \in \mathcal{R}$, there are two possibilities,*

- *either f is an Axiom A diffeomorphism without cycles,*
- *or f simultaneously has infinitely many sinks or sources.*

Concerning the coexistence of infinitely many sinks or sources, there are no avaliable examples of locally generic sets of C^1 surface diffeomorphisms displaying such a dynamics. Recall that in Section 3 we saw that, in the C^2-setting, the unfolding of homoclinic tangencies leads to the appearance of locally generic sets consisting of diffeomorphisms with infinitely many sinks or sources, the so-called (Newhouse) coexistence phenomenon.

In the C^1-topology and in dimension bigger or equal than 3, the locally generic coexistence phenomenon will appear associated not to homoclinic tangencies but to a special type of strongly non-hyperbolic homoclinic classes, see [57]. We will discuss this phenomenon (and some stronger forms of it) in Section 10.6.

In order to use the decomposition of the non-wandering set into elementary pieces of dynamics in Theorem 10.8, and keeping in mind Theorem 10.16, it is crucial to determine under which conditions the number of homoclinic classes is finite and how this number varies with the (generic) diffeomorphisms.

Recall that in the residual subset $\mathcal{R}$ of $\mathrm{Diff}^1(M)$ we are considering every homoclinic class depends continuously on f and two homoclinic classes are either disjoint or coincide. Given $f \in \mathcal{R}$, let $\#(f) \in \mathbb{N} \cup \{\infty\}$ be the cardinal of the set of homoclinic classes of f. Observing that, locally in the set $\mathcal{R}$, the number of homoclinic classes only can increase, thus the function $\#(f)$ is lower semi-continuously. So there is a new residual subset of $\mathrm{Diff}^1(M)$ (also denoted by $\mathcal{R}$ for simplicity) where $\#(f)$ is continuous, that is, locally constant. We have the following result.

Theorem 10.17 (Abdenur [1])**.** *There is a residual subset $\mathcal{R}$ of $\mathrm{Diff}^1(M)$ such that the map $\#(\cdot)$ is well defined and locally constant in $\mathcal{R}$.*

This result means that the property of having k homoclinic classes, $k \in \mathbb{N} \cup \{\infty\}$, is stable in $\mathcal{R}$. This justifies and motivates the introduction of two types of generic dynamics:

- *Wild dynamics,* diffeomorphisms $f \in \mathcal{R}$ having infinitely many homoclinic classes and such that the same holds for every $g \in \mathcal{R}$ close to f. We denote this subset of $\mathcal{R}$ by $\mathcal{W}$.
- *Tame dynamics,* diffeomorphisms $f \in \mathcal{R}$ with finitely many homoclinic classes, say k, such that every $g \in \mathcal{R}$ close to f also has k homoclinic classes. We denote this subset of $\mathcal{R}$ by $\mathcal{T}$.

By Theorem 10.17, $\mathcal{R} = \mathcal{T} \cup \mathcal{W}$. Examples of tame diffeomorphisms include the Axiom A diffeomorphisms with the no-cycles property as well as the robustly transitive diffeomorphisms (see Section 7). One can also easily build non-hyperbolic tame diffeomorphisms, admitting a filtration such that the maximal invariant set in each level of the filtration is a robustly transitive set: these kind of examples can be obtained through generic bifurcation of Morse-Smale diffeomorphisms *via* heterodimensional cycles (see [107, 112]). Examples of wild diffeomorphisms are, for instance, those displaying the Newhouse coexistence phenomenon, see also Section 10.6.

Recall that Theorem 7.3 states that every robustly transitive set has a dominated splitting and is volume hyperbolic. The definition of robustly transitive set Λ of a diffeomorphism f involves *all* the diffeomorphisms in a vicinity of f. Since here we are adopting the generic (or locally generic) approach, it is convenient to consider only diffeomorphisms in a relative neighborhood of f in the residual subset $\mathcal{R}$. This leads to the concept of *generically transitive sets*, much more suited for our purposes, defined as follows: an f-invariant transitive set Λ_f is generically transitive if there are neighborhoods U of Λ_f in the ambient manifold and $\mathcal{U}$ of f in $\mathrm{Diff}^1(M)$ and a residual subset $\mathcal{R}$ of $\mathcal{U}$ containing f such that, for every diffeomorphism $g \in \mathcal{R}$, the maximal invariant set of g in the closure of U is a transitive set contained in U and $\Lambda_f = \cap_{\mathbb{Z}} f^i(\overline{U}) \subset U$.

As there are not known examples of generic transitive sets which are not robustly transitive ones, it is natural to pose the following question:

Problem 10.18. Is any generically transitive set a robustly transitive one?

Theorem 7.3 is generalized in [1] for this context: *every generically transitive set has a dominated splitting which is volume hyperbolic.* In fact, next result gives a convenient reformulation of Theorem 7.2 to the generic context.

Theorem 10.19 (Abdenur [1]). *Consider a diffeomorphism f and a saddle p of f. Suppose that $H(p, f)$ is generically transitive. Then $H(p, f)$ has a dominated splitting which is volume hyperbolic.*

Currently we know homoclinic classes of generic diffeomorphisms are chain recurrence classes. Thus the Conley theory (Theorem 10.2) gives a filtration separating these classes. Moreover, when a class is isolated it is the maximal invariant set in the corresponding level of a filtration (see Remark 10.4). This leads to the next result, previously obtained by Abdenur by using different methods.

Theorem 10.20 (Abdenur [1]). *Consider a tame diffeomorphism $f \in \mathcal{T}$. Then there are a filtration $\mathcal{M} = (M_i)_{i=1}^k$ of f, finitely many hyperbolic periodic points $P_f^1, \dots, P_f^k$ of f, and a neighborhood $\mathcal{U}$ of f in $\mathrm{Diff}^1(M)$ such that, for every $g \in \mathcal{U} \cap \mathcal{T}$,*

- *the continuations $P_g^1, \dots, P_g^k$ of $P_f^1, \dots, P_f^k$ are defined,*
- *for every $i = 1, \dots, k$, the maximal invariant set of g in the level $L_i = (M_i \setminus M_{i+1})$ of the filtration is the homoclinic class of P_g^i, and*
- *every homoclinic class of g is volume hyperbolic.*

The filtration in Theorem 10.20 gives directly:

Corollary 10.21. *Given any tame diffeomorphism $f \in \mathcal{T}$ there are no cycles*[17] *between any two homoclinic classes of f.*

Putting all the previous results together one gets the following dichotomy theorem:

Theorem 10.22. *There is a residual subset $\mathcal{R} = \mathcal{T} \cup \mathcal{W}$ of $\mathrm{Diff}^1(M)$, where $\mathcal{T}$ and $\mathcal{W}$ are disjoint open subsets of $\mathcal{R}$, such that, for every f in $\mathcal{R}$, the number of homoclinic classes is locally constant in $\mathcal{R}$ and*

[17] Let $\Lambda_1, \dots, \Lambda_m$ be a finite family of different homoclinic classes of a diffeomorphism f. These classes form a *cycle* if $W^u(\Lambda_i) \cap W^s(\Lambda_{i+1}) \neq \emptyset$ for all $i = 1, \dots, (m-1)$ and $W^u(\Lambda_m) \cap W^s(\Lambda_1) \neq \emptyset$.

- *if $f \in \mathcal{W}$ then f has infinitely many homoclinic classes,*
- *if $f \in \mathcal{T}$ then the limit set of f is the union of finitely many pairwise disjoint homoclinic classes such that*
  - *there is a filtration such that every homoclinic class is a generically transitive set in a level of the filtration,*
  - *every class is volume hyperbolic,*
  - *there are no-cycles between any two homoclinic classes.*

This result is the generalization in higher dimensions of Theorem 10.16 for generic surface diffeomorphisms.

In the context of tame diffeomorphisms, it is possible to go further in the description of the topological dynamics. An important type of homoclinic classes are those being *topological attractors* (or *repellers)*, that is, the stable (or unstable) manifold of the class is a neighborhood of it. The existence of the filtration and the finiteness of the number of homoclinic classes directly imply the existence of at least one attractor (and one repellor): the last level of the filtration is a strictly invariant open set contained in the basin of the corresponding homoclinic class, which is an attractor. Thus it is natural to ask whether the union of the basins of the attractors are dense in the ambient manifold.

Theorem 10.23 (Carballo, Morales [88]). *For any diffeomorphism $f \in \mathcal{T}$, the union of the basins of the attractors is a dense open subset of the whole manifold.*

In the generic general case (i.e., for wild diffeomorphisms), it is not known if there is an elementary piece of dynamics being a topological attractor or repellor. In [162], it is conjectured that for C^r-generic diffeomorphisms, $r \geq 1$, the union of the basins of all the chain recurrence classes being *quasi-attractors* is a residual subset of the ambient manifold. Recall that a qausi-attractor is an intersection of topological (not necessarily transitive) attractors. This conjecture was proved by Hurley in [162] for homeomorphisms (for results on generic properties of attractors of homeomorphisms see [164, 163]). In fact, the previous conjecture was proved for C^1 diffeomorphisms in [54].

We close this section by reformulating Problem 10.18 (every generic transitive set is robustly transitive) as a natural conjecture arising from the results in this section:

Conjecture 10.24. Every homoclinic class of a tame diffeomorphism is a robustly transitive set.

For tame diffeomorphisms, homoclinic classes are isolated chain recurrence classes (or equivalently elementary pieces of dynamics). Thus, by Corollary 10.15, each class is robustly chain recurrent. As a consequence we have the following:

Corollary 10.25. *There is an open subset $\mathcal{O}_{\mathcal{T}} \subset \mathrm{Diff}^1(M)$ containing $\mathcal{T}$ such that any $f \in \mathcal{O}_{\mathcal{T}}$ has finitely many chain recurrence classes. Moreover, the number of classes is locally constant and each class is volume hyperbolic, the maximal invariant set in the level of a filtration, and contains one homoclinic class.*

In the statement above we cannot ensure that the classes always remain weakly transitive, thus we do not know if they are elementary pieces of dynamics in the whole $\mathcal{O}_{\mathcal{T}}$ (of course, they are elementary pieces, generically in $\mathcal{O}_{\mathcal{T}}$).

Finally, if the conjecture above were true, each recurrence class of $f \in \mathcal{O}_{\mathcal{T}}$ should be a homoclinic class.

10.6. **A sample of wild dynamics.** Besides their existence, very little is known about the wild diffeomorphisms. We just know a few mechanisms or dynamical configurations leading to wild dynamics. Using these examples, we are learning the most elementary topological properties of such systems.

10.6.1. *Coexistence of infinitely many attractors.* The first examples of wild diffeomorphisms (in the C^2-setting and for surface diffeomorphisms) were discovered by Newhouse, [241, 242]. He proved the existence of an open set $\mathcal{U}$ in $\mathrm{Diff}^r(M)$, $r \geq 2$, and of a residual subset $\mathcal{R}$ of $\mathcal{U}$ consisting of diffeomorphisms having infinitely many sinks or sources, see Theorem 3.6. This phenomenon is associated to the unfolding of homoclinic tangencies and it was widely discussed in Section 3.2. For future applications and motivations, it is important to point out that this coexistence phenomenon is closely related to the persistence of cycles (in this context, homoclinic tangencies). Moreover, in these constructions by shrinking the sizes of the sets $\mathcal{U}$ and $\mathcal{R}$, one has that, for every $f \in \mathcal{R}$, the whole homoclinic class of the saddle involved in the tangency is contained in the closure of the (infinite) set of sinks or sources.

In higher dimensions, a trivial way to get C^2-locally generic wild diffeomorphisms is to consider a normally hyperbolic invariant surface S where the restricted dynamics is wild (having infinitely many sinks or sources). Depending on the type of normal hyperbolicity of the surface S (attracting, repelling or of saddle type), this configuration may lead to infinitely many sinks, sources or saddles with trivial homoclinic classes. A much deeper fact is that, in higher dimensions, the coexistence phenomenon (of sinks, sources or saddles with trivial homoclinic classes) does occur close to any diffeomorphism with a homoclinic tangency, see Section 3.5.1.

10.6.2. *C^1-coexistence phenomenon of sinks or sources in dimension* 3. We next discuss the C^1-coexistence phenomenon in dimension 3. This phenomenon is not explicitly related to the persistence of tangencies (although such tangencies appear in this context) but to the persistent lack of domination of a homoclinic class.

In [57], given a closed 3-manifold M, it is considered an open set $\mathcal{U}$ of $\mathrm{Diff}^1(M)$, such that every diffeomorphism $f \in \mathcal{U}$ has a periodic saddle Q_f (depending continuously on $f \in \mathcal{U}$) such that its homoclinic class $H(Q_f, f)$ does not admit (in a persistent way) any dominated splitting.

A simple way to get this property is the following. Consider the model heterodimensional cycle in Section 6.1.2. Perturbing this cycle, one has an open set $\mathcal{V}$ of C^1-diffeomorphisms such that the homoclinic class of the point Q of Morse index two in the cycle contains the point P of index 1 in the cycle. We recall that this construction only involves semi-local properties of the bifurcating diffeomorphisms, and so it is compatible with other global assumptions on the dynamics. For instance, the point P may be homoclinically related to a point P' of index 1 having a non-real (contracting) eigenvalue. Similarly, the point Q may be homoclinically related to a point Q' of index 2 with a non-real (expanding) eigenvalue. Hence, for every diffeomorphisms g in $\mathcal{V}$,

$$H(P', g) = H(P, g) \subset H(Q, g) = H(Q', g) = \Lambda_g.$$

This implies that the homoclinic class Λ_g does not admit any dominated splitting. The existence of the point P' prevents the existence of a dominated splitting $T_{\Lambda_g} = E \oplus F$ such that F is two dimensional. In such a case, the intersection of the stable bundle of P' and F should be a one-dimensional invariant direction, but this is not compatible with the existence of a non-real eigenvalue of Dg at P'. Similarly, using the point Q', one has that E cannot have dimension 2. So there is no any dominated splitting over Λ_g.

Under this kind of dynamical configuration, in [57] it is proved that there is a residual subset $\mathcal{R}$ of $\mathcal{V}$ such that, for every $g \in \mathcal{R}$, the homoclinic class $H(Q, g)$ is contained in the closure of the set of sources or sinks of g. Thus, since the homoclinic class of Q is non-trivial, every g in $\mathcal{R}$ has infinitely many sources or sinks. More precisely, if the Jacobian of f at the some of the points P, P', Q or Q' (here, for simplicity, we are assume that the points are fixed) is less than one then the homoclinic class is contained in the closure of the (infinite) set of sinks. Analogously, if the Jacobian of f at P, P', Q or Q' is greater than one then $H(Q, g)$ is contained in the closure of the (infinite) set of sources. In particular, $H(Q, g)$ can be simultaneously contained in the closures of the sets of sinks and sources.

Currently, the previous result is a direct consequence of Theorem 7.3 (generically, every homoclinic class has a dominated splitting of it is contained in the closure of an infinite set of sinks or sources). In fact, the robust non-dominated homoclinic classes are the corner stones in the construction of wild dynamics, we will discuss this point in the next section.

10.6.3. *C^1-generic existence of aperiodic pieces in dimension* 3. Consider now diffeomorphisms $f \in \mathrm{Diff}^1(M)$ having a *wild homoclinic class* $H(Q_f, f)$, that is, a homoclinic class which robustly does not admit any dominated splitting. In [53] dynamical properties of of generic diffeomorphisms f having wild homoclinic classes are obtained under the following additional hypothesis: assume that there is a neighborhood $\mathcal{U}$ of f such that, for every $g \in \mathcal{U}$, the set of periodic saddles homoclinically related to Q_g (in particular, with the same index than Q_g) simultaneously contains saddles with Jacobian greater and less than 1. In this section, for simplicity, the terminology *wild homoclinic class* means a wild homoclinic class verifying this extra hypothesis on the jacobians.

Let us recall that an f-invariant set Σ is *minimal* if the orbit of any point of Σ is dense in the whole Σ. An $f$invariant set Σ is *uniquely ergodic* if it supports only an f-invariant measure (which is necessarily ergodic).

We have the following result for wild homoclinic classes:

Theorem 10.26 (Bonatti, Díaz [53]). *Let M be a closed 3-manifold, g a diffeomorphism defined on M and Q_g a saddle of g. Suppose that there is an open set $\mathcal{U}$ of $\mathrm{Diff}^1(M)$ such that, for every $f \in \mathcal{U}$, the continuation Q_f of Q_g is defined and its homoclinic class $H(Q_f, f)$ is wild. Then there is a residual subset $\mathcal{R}$ of $\mathcal{U}$ such that, for every $f \in \mathcal{R}$, the homoclinic class $H(Q_f, f)$ is contained in the closure of infinitely (uncountable) many disjoint compact f-invariant sets $\mathcal{M}_i$ such that:*

- *Every $\mathcal{M}_i$ is a maximal transitive set without periodic points. Moreover, the dynamics of f in $\mathcal{M}_i$ is minimal and uniquely ergodic.*
- *Every $\mathcal{M}_i$ simultaneously is Lyapunov stable for f and f^{-1}. As a consequence, the sets $\mathcal{M}_i$ are saturated transitive sets*[18].

[18]The Lyapunov stability for f and f^{-1} implies that these sets are, in fact, chain recurrence classes, and so elementary pieces of dynamics, according to the terminology introduced in Section 10.3.3

This result implies that there are (locally generic) saturated transitive sets which are not homoclinic classes (recall the discussion of the notion of elementary pieces of dynamics in Section 10.3).

The main step for proving Theorem 10.26 is to construct a residual subset $\mathcal{R}$ of $\mathcal{U}$ consisting of diffeomorphisms f with *universal dynamics* at the wild homoclinic class $H(Q_f, f)$. Let us define this notion. Consider the compact unitary ball D^3 of $\mathbb{R}^3$ and denote by $\mathrm{Diff}_0^+(D^3)$ the set of C^1-diffeomorphisms $\phi\colon D^3 \to \mathrm{int}(D^3)$ preserving the orientation. A diffeomorphism f has *universal dynamics* if for any open set $\mathcal{O} \subset \mathrm{Diff}_0^+(D^3)$ there is a periodic disk $D \subset M$ (that is, there is n such that $f^n(D) \subset \mathrm{int}(D)$ and the disks $f^i(D)$, $0 \le i < n$, are pairwise disjoint) such that $f^n\colon D \to D$ is differentially conjugate to some element $g \in \mathcal{O}$. Moreover, there is another disk satisfying the same properties for f^{-1}.

The proof of the existence of a residual subset of diffeomorphisms of $\mathcal{U}$ displaying the universal dynamics uses heavily the constructions in [58], specially the results involving transitions discussed in Section 7.1.2. Since the homoclinic class $H(Q_f, f)$ is robustly non-dominated, the constructions in Section 7.1.2 imply that (after an arbitrarily small perturbation) we can assume that f has a periodic point e_f such that $Df^{n(e_f)}(e_f)$ is a homothety (here $n(e_f)$ is the period of e_f). Here, strongly using the fact that the homoclinic class $H(Q_f, f)$ contains points a_f and b_f with Jacobian bigger are less than one, one gets that $Df^{n(e_f)}(e_f)$ is the identity. For this one selects orbits having an appropriate number of iterates close to the points a_f and b_f, such that after *multiplying* the derivatives of the orbit (exactly as in Section 7.1.2) the result is an orbit whose derivative is a homothety with Jacobian close to 1 and, so, close to the identity.

Performing a new C^1-arbitrarily small perturbation, we can assume that there is a small disk D containing e_f such that $D, f(D), \dots, f^{n(e_f)-1}(D)$ are pairwise disjoint disks and the restriction of $f^{n(e_f)}$ to D is the identity. Perturbing the identity in D, we now get a new small periodic disk D' of arbitrarily large period, such that the first return map of f to D' is a prescribed diffeomorphism $g \in \mathcal{O}$. It is now clear how to prove that f has universal dynamics at the homoclinic class $H(Q_f, f)$.

To prove Theorem 10.26 we consider the open set $\mathcal{O} \subset \mathrm{Diff}_0^+(D^3)$ of diffeomorphisms having wild homoclinic classes.

Thus every locally generic diffeomorphism f having a wild homoclinic class admits a periodic disk D_0 such that the induced dynamics on D_0 is conjugate to some diffeomorphism in $\mathcal{O}$. In particular, the orbit of D_0 contains a wild homoclinic class and so contains a periodic sub-disk $D_1 \subset D_0$ on which the induced dynamics is conjugate to some element of $\mathcal{O}$. Arguing now inductively, one gets sequences of nested disks D_k, whose diameters go to zero as k to ∞, and of natural numbers n_k, $n_k \to \infty$, such that

- for every k, the disks $D_k, \dots, f^{n_k-1}(D_k)$ are pairwise disjoint,
- $f^{n_k}(D_k)$ is contained in the interior of D_k,
- the restriction of f^{n_k} to D_k is conjugated to some element of $\mathcal{O}$, and
- the diameters of the disks of the forward orbit of D_k are bounded by some d_k, where $d_k \to 0$ as $k \to \infty$.

Consider now the set

$$\mathcal{M}_f = \bigcap_k (\cup_{i=0}^{n_k-1} f^i(D_k)).$$

By construction, the set $\mathcal{M}_f$ is an f-invariant Cantor set without periodic orbits, every forward orbit of it is dense in $\mathcal{M}_f$, and it is Lyapunov stable. To get that the set $\mathcal{M}_f$ is Lyapunov stable for f^{-1} one needs to consider disks D_k such that, alternately, one has either $f^{n_k}(D_k) \subset \operatorname{int}(D_k)$ or $f^{n_k}(D_k) \supset \operatorname{int}(D_k)$ (in fact, this is the reason because in the definition of universal dynamics involves f^{-1}).

Observe that the transitive set $\mathcal{M}_f$ are obtained as infinite renormalizable sets or as an adding machine sets.

We have the following important observation which follows from the definition of universal dynamics and the procedure above:

Remark 10.27. Generic diffeomorphisms with a wild homoclinic class display infinitely many times any robust feature of diffeomorphisms of the disk D^3.

The previous remark gives the following strong form of the Newhouse coexistence phenomenon for C^1-diffeomorphisms of three manifolds:

Theorem 10.28. *Let $\mathcal{W}_h$ be the open subset of* $\operatorname{Diff}^1(M)$ *(where* $\dim(M) = 3$*) of diffeomorphisms having a wild homoclinic class. There is a residual subset $\mathcal{R}$ of $\mathcal{W}_h$ such that every f in $\mathcal{R}$ simultaneously has infinitely many*

- *sinks and sources,*
- *independent saddles (saddles with trivial homoclinic classes),*
- *non-trivial uniformly hyperbolic attractors and repellers,*
- *non-trivial partially hyperbolic attractors or repellers,*
- *wild homoclinic classes, and*
- *maximal saturated transitive Cantor sets with minimal dynamics, in particular, without periodic orbits.*

The same kind of locally generic diffeomorphisms can be constructed in higher dimensions in two ways: either, trivially, by multiplying the 3-dimensional example by a transverse uniformly hyperbolic dynamics, or by considering homoclinic classes robustly without any dominated splitting and with the extra hypothesis on the jacobians.

We finish this section posing some question concerning the wild dynamics, the basin of the attractors, the geometrical distribution of the independent saddles, and the Smale's ordering. The *Smale's ordering* in the set of saddles is defined as follows, given two saddles s_1 and s_2 one writes $s_1 \prec s_2$ if the unstable manifold of s_1 meets the stable one of s_2.

Problem 10.29.

(1) Let f be a locally generic diffeomorphism having infinitely many independent saddles but having only finitely many non-trivial homoclinic classes. Are there finitely many (local?) hyper-surfaces containing all the independent saddles?
(2) Is it possible to have (generically) an infinite sequence of independent saddles $(s_i)_i$ with $s_1 \prec s_2 \prec \cdots s_i \prec \cdots$.
(3) Generically, does the union of the basins of the topological attractors be a dense open subset of the ambient manifold? (compare with the result in [88] stating that, for tame dynamics, the union of the basins of the attractors is a dense open subset of the whole manifold).

(4) Do locally generic diffeomorphisms having non-transitive (aperiodic) elementary pieces of dynamics exist?

11. SRB measures and Gibbs states

We call *Sinai, Ruelle, Bowen (SRB) measure* for a diffeomorphism $f : M \to M$, any invariant probability measure μ whose *basin* $B(\mu)$, the set of points $x \in M$ such that

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi(f^j(x)) = \int \phi d\mu$$

for any continuous function ϕ, has positive Lebesgue measure. Every uniformly hyperbolic C^2 diffeomorphism admits a finite number of SRB measures, and the union of their basins contains Lebesgue almost every point in the ambient manifold [73, 300, 316]. We are going to discuss some recent extensions of this fundamental result beyond the hyperbolic world, especially for dominated, or even partially hyperbolic, diffeomorphisms.

11.1. **SRB measures for certain non-hyperbolic maps.**

We begin by recalling certain constructions of SRB measures for non-hyperbolic skew-products and their perturbations.

11.1.1. *Intermingled basins of attraction.* In the uniformly hyperbolic context, each attractor Λ supports exactly one SRB measure μ, and the basin $B(\mu)$ is a full measure subset of the (open) stable set $B(\Lambda)$ of the attractor. The construction we describe here, due to Kan [168], shows that the topology of the basins of SRB measures may be much more complicated, in general: in this example both basins are dense in the ambient space.

For simplicity, we describe the example as an endomorphism of the 2-dimensional cylinder preserving both boundary components. An interesting point is that in this setting the construction is C^2 robust. At the end we mention a version for diffeomorphisms of the 3-torus.

Consider $F : S^1 \times [0,1] \to S^1 \times [0,1]$ of the form $F(x,t) = (kx \bmod \mathbb{Z}, f_x(t))$ where k is some integer number with $|k| \geq 3$. Then $x \mapsto kx \bmod \mathbb{Z}$ has at least two fixed points, denoted p and q. Assume F is C^2 and

(1) F preserves the boundary components: $f_x(0) = 0$ and $f_x(1) = 1$;
(2) f_p and f_q have exactly two fixed points each, a source at $t = 1$ (respectively $t = 0$) and a sink at $t = 0$ (respectively $t = 1$);
(3) $|f_x'(t)| < k$ at every point $(x,t) \in S^1 \times [0,1]$;
(4) $\int \log f_x'(0) dx < 0$ and $\int \log f_x'(1) dx < 0$.

Proposition 11.1. *Every map satisfying (1)–(4) admits exactly two SRB measures. Both basins are dense, and their union has total Lebesgue measure in the cylinder $S^1 \times [0,1]$.*

Proof. Let μ_0 and μ_1 be the normalized arc-length measures along the two boundary components of the circle.These are ergodic F-invariant probabilities, and hypothesis (4) means that their Lyapunov exponents along the t-direction are negative. Consequently, both μ_0 and μ_1 are SRB measures for F: Pesin local stable manifolds of typical points in $t = 0$ form a positive area subset B_0 of $B(\mu_0)$, and similarly for $t = 1$.

To prove the remaining statements, consider any smooth segment γ in the open cylinder, transverse to the t-direction. We are going to show both basins intersect γ on positive Lebesgue measure subsets whose union has full measure in γ. Up to replacing γ by some forward iterate, we may suppose that it intersects $W^s(p,0) = \{p\} \times [0,1)$ and $W^s(q,0) = \{q\} \times (0,1]$. Then the orbit of γ accumulates on both boundary components. It follows (see [66, Section 2]) that some forward iterate $f^n(\gamma)$ intersects both $B_0 \subset B(\mu_0)$ and $B_1 \subset B(\mu_1)$ on positive measure subsets. Then $\gamma \cap f^{-n}(B_i)$ is contained in $B(\mu_i)$ and, in fact, its points have Pesin local stable manifolds (short segments in the t-direction) whose union has positive area and is contained in the basin of μ_i. As a consequence, every segment sufficiently C^1 close to γ intersects $B(\mu_i)$ on a subset whose measure is at least half that of $\gamma \cap f^{-n}(B_i)$.

One can easily deduce that γ is essentially contained in $B(\mu_0) \cup B(\mu_1)$. Indeed, suppose there was a positive measure subset in the complement of the union. By considering a density point and using the fact that forward iterates of F do not distort Lebesgue measure much, we find a sequence $\gamma_n \subset F^n(\gamma)$ of segments with constant length and such that the relative measure of the union of the basins inside γ_n goes to zero when n goes to infinity. Restricting to a subsequence if necessary, we may suppose that the γ_n converge to some segment β contained in the closed cylinder. This contradicts the conclusion of the previous paragraph, for β in the place of γ. □

Along the same lines we may construct diffeomorphisms of the 3-dimensional torus with intermingled basins. Roughly, one replaces the expanding map of S^1 by an Anosov diffeomorphism on T^2, and then glues along the boundary two diffeomorphisms of $T^2 \times [0,1]$ obtained in this way. These examples are still some robust, but only restricted to maps that preserve the two invariant tori obtained from the boundary components.

11.1.2. *A transitive map with two SRB measures.* Here we refine the previous construction, to exhibit a dynamical system where the topological and the ergodic notions of indecomposability are strikingly different: the map is transitive, yet it admits two different SRB measures. These examples are robust as endomorphisms of the compact cylinder preserving the boundary.

Consider a map $F : S^1 \times [0,1] \to S^1 \times [0,1]$ satisfying, in addition to properties (1)–(4), that there exist two compact sectors $I_j \times [0,1]$, $j = 1,2$, such that

(5) F maps each $I_j \times [0,1]$ onto the whole cylinder, the image of the boundary being disjoint from $I_1 \cup I_2$.
(6) There exists a vector field X which vanishes exactly at $t = 0$ and $t = 1$, and a negative irrational number α such that the transformation f_x is given by the time-1 map of X on the first sector, and by the time-α map of X on the second sector.

Proposition 11.2. *Every transformation F satisfying (1)–(6) is transitive.*

Proof. Any open subset U has an iterate which crosses the sector $I_1 \times [0,1]$. Consider a small square inside this iterate and intersecting the stable manifold of a fixed point in the sector. Using that F is given by a product on $I_1 \times [0,1]$, we get that some further iterate $F^m(U)$ contains a rectangle $I_1 \times [t_1, t_2]$. Now, using that α is irrational, we may find sub-sectors $J_\pm \times [0,1]$ of $I_1 \times [0,1]$, and iterates $F^{n_\pm}$

which map the corresponding sub-sector bijectively onto $I_1 \times [0,1]$ and are given by the time-$s_\pm$ maps of X, with $s_- < 0 < s_+$ arbitrarily close to zero. Then the orbit of $F^m(U)$ under $F^{n_\pm}$ contains the sector $I_1 \times (0,1)$. It follows that the complete forward orbit of U contains the whole open cylinder. Clearly, this implies transitivity. □

As in the previous section, this construction extends to diffeomorphisms of the 3-dimensional torus. To ensure transitivity, just ask that the halves of the 3-torus bounded by the invariant 2-tori be mapped to one another. Such examples *do not* fill an open subset of the space of all diffeomorphisms on T^3. In fact, no robust examples are known of transitive diffeomorphisms on boundariless manifolds with more than one SRB measure.

11.1.3. *Robust multidimensional attractors.* We describe a construction, by Viana and Alves, of open sets of non-uniformly hyperbolic maps with multidimensional expanding direction. It applies also to diffeomorphisms, but we restrict ourselves to a non-invertible situation.

Let $F : S^1 \times [-2,2] \to S^1 \times [-2,2]$ be a C^3 map of the form

$$F(\theta, x) = (k\theta \bmod \mathbb{Z}, x^2 - a_0 + a(\theta)),$$

where $a : S^1 \to (1,2)$ is a Morse function, for instance $a(\theta) = \alpha \sin(2\pi\theta)$, and a_0 is a Misiurewicz parameter (the critical point $x = 0$ is non-recurrent and there is no periodic attractor) for the quadratic transformation $x \mapsto x^2 - a_0$.

Theorem 11.3. *Assume k is large enough.*

(1) (Viana [335]) *For every $\alpha > 0$ sufficiently small there exists $c_0 > 0$ and a C^3 neighborhood $\mathcal{U}$ of F in the space of endomorphisms of the cylinder, such that every $G \in \mathcal{U}$ has only positive Lyapunov exponents*

$$\liminf_{n\to+\infty} \frac{1}{n} \log \|DF^n(\theta, x)^{-1}\| \leq -c_0 < 0 \tag{28}$$

at Lebesgue almost every point $(\theta, x) \in S^1 \times [-2,2]$.

(2) (Alves [4]) *Every $G \in \mathcal{U}$ admits some ergodic invariant probability measure μ_G absolutely continuous with respect to Lebesgue measure.*

(3) (Alves, Viana [9]) *The measure μ_G is unique and depends continuously on the map $G \in \mathcal{U}$.*

In the sequel we are going to outline the proof of these statements. Notice that [335] considered $k \geq 16$, but Buzzi, Sester, Tsujii [84] show that $k \geq 2$ suffices.

Beforehand, let us mention that the ergodic properties of these systems have been studied by several authors very recently. Especially, Alves, Luzzatto, Pinheiro [8] prove that the correlations of Hölder continuous observables decay superpolynomially i.e. faster than $Cn^{-\gamma}$ for any $\gamma > 0$. See [30, 336] for definitions and background information. Their result should not be optimal: Baladi, Gouezel [32] have just announced that the decay of correlations is stretched-exponential i.e. faster than $Ce^{-c\sqrt{n}}$. No lower bounds are known and, in particular, the following question remains open:

Problem 11.4. Does (G, μ_G) have exponential decay of correlations in the space of Hölder continuous functions ? Same question for $G = F$.

For proving part 1 of the theorem, one writes the Lyapunov exponent along the vertical direction as

$$\liminf_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \log(2|x_j|)$$

where $(\theta_j, x_j) = F^j(\theta, x)$ for $j \geq 0$. It suffices to prove (28) for almost every point on any admissible (nearly horizontal) curve. The assumption on k ensures that the family of such curves is invariant under forward iteration by F. The assumption that $a(\cdot)$ is a Morse function, together with a delicate analysis of iterates of admissible curves, shows that at each iterate j the set of points (θ, x) in any admissible curve for which

$$\log|x_j| \in [-r-1, -r]$$

decays exponentially with r. Combining this with the fact that the horizontal dynamics $x \mapsto kx$ is rapidly mixing, via a large deviations argument, one finds $c_0 > 0$ such that

$$\frac{1}{n}\sum_{j=0}^{n-1} \log(2|x_j|) > c_0 n,$$

but for a subset of the admissible curve whose Lebesgue measure decreases as $e^{-c\sqrt{n}}$ with n. using a Borel-Cantelli argument, one obtains the conclusion of part 1 for $F = G$.

To obtain the same conclusion for general G in a neighborhood, one uses the fact that the vertical fibration is normally expanding for F. Consequently, by a version of the stability Theorem B.10 for non-invertible maps, it gives rise to an invariant central foliation, for every nearby G. Then the argument is carried out as before, with the tangent to this central foliation in the role of vertical direction. The condition of C^3-closeness ensures that the critical set of G, where the derivative along the central direction vanishes, is a circle C^2 close to the original $\{x = 0\}$.

The proof of part 2 is based on the idea of *hyperbolic time*. The precise definition will appear later, here we just mention a consequence: if n is a hyperbolic time for a point $z = (\theta, x)$ then $DG^k(G^{n-k}(z))$ is an expansion

$$\|DG^k(G^{n-k}(z))^{-1}\| \leq e^{-ck}, \quad c > 0$$

for every $1 \leq k \leq n$. From the proof of (28) one gets that Lebesgue almost every point has infinitely many hyperbolic times, in fact, the set of hyperbolic times has positive density at infinity. This is based on the following remarkably useful (see e.g. [7] for a proof)

Lemma 11.5 (Pliss [275]). *Given $A \geq c_0 > c$ there exists $\theta = (c_0 - c)/(A - c)$ such that, given any real numbers $a_1, \ldots, a_N$ with*

$$\sum_{i=1}^{N} a_i \geq c_0 N \quad \textit{and} \quad a_i \leq A \textit{ for every } i,$$

there exist $l > \theta N$ and $1 \leq n_1 < \cdots < n_l \leq N$ such that

$$\sum_{i=n+1}^{n_j} a_i \geq c(n_j - n) \quad \textit{for all } 0 \leq n < n_j \textit{ and } j = 1, \ldots, l.$$

From the construction of hyperbolic times, one shows that every map $G \in \mathcal{U}$ induces a map $\tilde{G}$, of the form $\tilde{G}(z) = G^{i(z)}(z)$ which is uniformly expanding (though not a Markov map). One proves that the Perron-Frobenius operator of $\tilde{G}$ has a non-negative fixed point h_G in the space of bounded variation functions. Then h_G is the density of a $\tilde{G}$-invariant probability measure $\tilde{\mu}_G$. Now

$$\mu = \sum_{j=0}^{\infty} G_*^j(\tilde{\mu}_G | \{z : i(z) > j\})$$

defines an absolutely continuous G-invariant measure μ_G. The fact that μ is finite follows from $e^{-c\sqrt{n}}$ estimates as above. By normalization and ergodic decomposition of μ one easily gets an ergodic G-invariant probability, absolutely continuous with respect to Lebesgue measure.

Concerning part 3 of Theorem 11.3, we just mention that continuity follows from analyzing the variation of the fixed point h_G of the Perron-Frobenius operator as a function of the dynamics.

11.1.4. *Open sets of non-uniformly hyperbolic maps.* We describe a construction of robustly non-hyperbolic maps, proposed in [7, Appendix], which includes the examples in Sections 7.3.2 and 7.3.4. We start with the version for endomorphisms.

Let $f : M \to M$ be a C^1 map on a manifold with finite volume, such that there exist $p, q \in \mathbb{N}$, $\delta_1 > 0$, $\sigma_1 > q$, and a covering of M by measurable subsets $B_1, \ldots, B_p, B_{p+1}, \ldots, B_{p+q}$ satisfying

(a) f is injective on B_j for every $1 \le j \le p+q$

(b) $|\det(Df)| \ge \delta_1$ at every point, and $|\det(Df)| \ge \sigma_1$ on $B_{p+1}, \ldots, B_{p+q}$.

Assume, in addition, that there exist $\sigma_2 < 1$ and $\delta_2 > 0$ such that

(c) $\|Df^{-1}\| \le e^{\delta_2}$ at every point, and $\|Df^{-1}| \le \sigma_2$ on $B_1, \ldots, B_p$.

Proposition 11.6. *Given $p, q, \delta_1, \sigma_1, \sigma_2$ there exists $\bar{\delta}_2 > 0$ and $c_0 > 0$ such that any map satisfying conditions (a), (b), (c) for some $\delta_2 \le \bar{\delta}_2$ has*

$$\limsup_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \|Df(f^j(x))^{-1}\| \le -c_0 < 0$$

at Lebesgue almost every point $x \in M$.

First, one shows that under hypotheses (a) and (b) almost every point spends at least a fraction $\varepsilon_0 = \varepsilon_0(p, q, \delta_1, \sigma_1) > 0$ of time inside $B_1 \cup \cdots \cup B_p$: for every $n \ge 1$ the measure of the set of points with less than $\varepsilon_0 n$ visits to $B_1 \cup \cdots \cup B_p$ during the first n iterates is bounded by

$$\mathrm{Leb}(M) \sigma_1^{-(1-\varepsilon_0)n} \delta_1^{-\varepsilon_0 n} e^{\gamma_0 n} p^{\varepsilon_0 n} q^n$$

where $\gamma_0 = \gamma_0(\varepsilon_0)$ goes to zero when ε_0 goes to zero. Since $q < \sigma_1$, this expression decreases exponentially with n, if ε_0 is chosen small enough. For all other points, using (c) and taking δ_2 small,

$$\sum_{j=0}^{n-1} \log \|Df(f^j(x))^{-1}\| \le (\varepsilon_0 n) \log \sigma_2 + (1 - \varepsilon_0) n \delta_2 \le -c_0 n$$

with $c_0 = (\varepsilon_0/2) |\log \sigma_2|$. The proposition follows, using the Borel-Cantelli lemma.

Here is a typical application of Proposition 11.6. Let $f_0 : M \to M$ be a uniformly expanding map on the d-dimensional torus M:

$$\|Df_0^{-1}\| \le \sigma_2 < 1.$$

Let V be some small domain in M, and f_1 be any volume expanding map in M coinciding with f_0 outside V and only mildly contracting inside V: $\|Df_1^{-1}\| \le e^{\delta_2}$ for some small δ_2. Then *every map f in a C^1 neighborhood of f_1 is non-uniformly expanding*, in the sense that it satisfies the conclusion of Proposition 11.6. Just take $q = 1$ and $B_{p+1} = V$.

Along the same lines, one may give robust examples of C^1 diffeomorphisms admitting a dominated splitting $TM = E^{cu} \oplus E^{cs}$ with E^{cu} non-uniformly expanding and E^{cs} non-uniformly contracting:

$$\begin{gathered}
\limsup_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \left\| \left(Df \mid E^{cu}_{f^j(x)}\right)^{-1} \right\| \le -c_0 < 0 \\
\text{and} \\
\limsup_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \left\| Df \mid E^{cs}_{f^j(x)} \right\| \le -c_0 < 0
\end{gathered} \tag{29}$$

at Lebesgue almost every point x. In fact, one may construct these examples so that they have simultaneous hyperbolic times, in the sense of Section 11.3.3. Thus the methods in there prove existence and finiteness of SRB measures for these diffeomorphisms (assuming C^2 regularity).

Start with an Anosov diffeomorphism $f_0 : M \to M$ on the d-dimensional torus $M = T^d$, with hyperbolic splitting $TM = E^u \oplus E^s$. Let $V \subset M$ be a small domain: both V and $f_0(V)$ are contained in the projections of unit cubes of $\mathbb{R}^d$. Let f be a diffeomorphism such that

(1) f admits thin invariant cone fields $\mathcal{C}^{cu}$ and $\mathcal{C}^{cs}$ containing E^u and E^s, respectively;
(2) f contracts volume along any disk D^{cs} tangent to $\mathcal{C}^{cs}$ and expands volume along any disk D^{cu} tangent to $\mathcal{C}^{cu}$:
$$|\det(Df \mid TD^{cs})| < \sigma_1^{-1} < 1 \quad \text{and} \quad |\det(Df \mid TD^{cu})| > \sigma_1 > 1;$$
(3) f is C^1 close to f_0 outside V, so that $f(V)$ is still contained in the projection of a unit cube, and
$$\|Df \mid TD^{cs}\| < \sigma_2 < 1 \quad \text{and} \quad \|(Df \mid TD^{cu})^{-1}\| < \sigma_2 < 1$$
for any disks D^{cu} and D^{cs} as before;
(4) inside V the behavior of the two sub-bundles may reversed, but only mildly:
$$\|Df \mid TD^{cs}| < e^{\delta_2} \quad \text{and} \quad \|(Df \mid TD^{cu})^{-1}| < e^{\delta_2}$$
for any disks D^{cu} and D^{cs} as before.

Notice that 1–4 are C^1 open conditions. Arguing as in the proof of Proposition 11.6 one shows that they imply (29), if δ_2 is small enough.

11.2. **Gibbs u-states for $E^u \oplus E^{cs}$ systems.** .

Let Λ be a compact invariant subset for a C^r diffeomorphism $f : M \to M$ with a dominated splitting $T_\Lambda M = E^u \oplus E^{cs}$ where the sub-bundle E^u is uniformly expanding and has positive dimension.

Through every point $x \in \Lambda$ there exists a unique C^r injectively immersed strong unstable manifold $\mathcal{F}^u(x)$ tangent to E^u_x and which is uniformly contracted by the negative iterates of f. The strong unstable foliation $\mathcal{F}^u = \{\mathcal{F}^u(x) : x \in \Lambda\}$ is invariant under the diffeomorphism. In all that follows we assume that the set Λ consists of entire strong unstable leaves. That is the case, for instance, if Λ is a topological attractor, that is, the maximal invariant set in a positively invariant neighborhood.

11.2.1. *Existence of Gibbs u-states.* We are going to see that if f is of class C^2 then it admits a special class of invariant probability measures characterized by having conditional measures along (local) strong unstable leaves which are absolutely continuous with respect to the corresponding Lebesgue measure. We shall refer to these measures as Gibbs u-states, by analogy with the uniformly hyperbolic case, where they correspond to the Gibbs states associated to a potential defined by the Jacobian along the unstable direction. There is a dual notion of Gibbs s-state, when Λ admits a uniformly contracting direction E^s and consists of entire strong stable leaves.

We begin by introducing the notion of *foliated box* for the strong unstable foliation $\mathcal{F}^u$. Let X be any small unstable disk, that is, contained in some strong unstable leaf, and let Σ be a small cross section to the foliation through the center p of X. Denote by Y the intersection of Σ with the invariant set Λ. Then there exists $\phi : X \times Y \to \Lambda$ which is a homeomorphism onto its image, such that ϕ maps each horizontal $X \times \{y\}$ diffeomorphically to an unstable domain through y. We may choose ϕ such that $\phi(p, y) = y$ for all $y \in Y$ and $\phi(x, p) = x$ for all $x \in X$. In what follows we identify $X \times Y$, and each $X \times \{y\}$, with their images under such a chart ϕ.

Definition 11.7. An invariant probability measure μ is a *Gibbs u-state* if, for every foliated box $X \times Y$ such that $\mu(X \times Y) > 0$, the conditional measures of $\mu \mid X \times Y$ with respect to the partition into unstable plaques $\{X \times \{y\} : y \in Y\}$ are absolutely continuous with respect to Lebesgue measure along the corresponding plaque.

See Appendix C for a definition of conditional measures. For any domain D inside a strong unstable leaf, we shall denote by m_D the Lebesgue measure induced on D by some Riemannian structure on the manifold.

Theorem 11.8 (Pesin, Sinai [272])**.** *For any unstable domain $D \subset \Lambda$, every accumulation point of the sequence of probability measures*

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j\left(\frac{m_D}{m_D(D)}\right)$$

is a Gibbs u-state with densities $\rho(\cdot)$ with respect to Lebesgue measure along unstable plaques satisfying

$$\frac{\rho(z_1)}{\rho(z_2)} = \prod_{j=0}^{\infty} \frac{\det(Df^{-1} \mid E^u_{f^{-j}(z_1)})}{\det(Df^{-1} \mid E^u_{f^{-j}(z_2)})} \tag{30}$$

for any points z_1 and z_2 in the same unstable plaque.

The proof of Theorem 11.8 is based on the fact that the distortion of Lebesgue measure along strong unstable leaves by the iterates of f is bounded. More precisely,

Lemma 11.9. *Given any $L > 0$ there exists $C_0 > 0$ such that if $x, y \in D$, and $N \geq 1$ are such that* $\operatorname{dist}(f^N(x), f^N(y)) \leq L$ *then*

$$\frac{\det(Df^N|E^u_x)}{\det(Df^N|E^u_y)} \leq C_0 .$$

Here $\operatorname{dist}(\cdot,\cdot)$ means distance along the corresponding strong unstable leaf. The proof is a classical argument (see e.g. [67, Lemma 3.3]). One ingredient is the uniform contraction of unstable leaves by backward iterates of f. Another is uniform Lipschitz (or Hölder) variation of the unstable Jacobian $\det(Df \mid E^u)$ along strong unstable leaves (with bounded diameter). The latter is a consequence of the fact that strong unstable manifolds have uniformly bounded curvature, when f is C^2. See [78, 158].

Using the distortion lemma 11.9 one proves a uniform bound for the measure induced by every iterate $f^n_*(m_D/m_D(D))$ on each foliated box $X \times Y$:

$$f^n_*\Big(\frac{m_D}{m_D(D)}\Big)(A \times Y) \leq C\frac{m_X(A)}{m_X(X)} + \varepsilon_n \tag{31}$$

for every measurable subset A of X, where $C > 0$ and $\varepsilon_n > 0$ depend only on f, D, X, and ε_n goes to zero when n goes to infinity. The last term bounds the total weight of connected components of $f^n(D) \cap (X \times Y)$ that do not cross the foliated box completely; convergence to zero is exponential. Property (31) passes to the limit: every accumulation measure μ satisfies

$$\mu(A \times Y) \leq C\frac{m_X(A)}{m_X(X)} \tag{32}$$

for every measurable $A \subset X$. This suffices to prove that μ is a Gibbs u-state: the densities of the conditional measures with respect to Lebesgue measure along the plaques are bounded by $C/\mu(X \times Y)$, on every foliated box which has positive weight for μ. This ends our outline of the proof of Theorem 11.8.

Remark 11.10. If Λ is a topological attractor, the center stable sub-bundle E^{cs} admits a unique continuous invariant extension to the basin $B(\Lambda)$. This extension, which we also denote E^{cs} may be found as the limit of backward iterates of any continuous extension of the center stable sub-bundle to a small neighborhood of Λ: the limit exists by a contraction argument. The subspaces E^{cs}_y are characterized by the fact that their forward iterates grow slower than any vector in the complement.

Remark 11.11. When Λ is an attractor, including the case $\Lambda = M$, the conclusion of Theorem 11.8 extends to any C^2 disk contained in the basin of attraction of Λ and transverse to the center stable direction E^{cs}.

11.2.2. *Structure of Gibbs u-states.* Our notion of Gibbs u-state seems more general than the one of Pesin, Sinai, who require (30) as part of the definition. Observe that this implies that the densities are uniformly bounded from zero and infinity, with bounds independent of the Gibbs u-state. In particular, the support consists of entire strong stable manifolds. In this section we show that the two definitions do coincide.

The first step is to show that Theorem 11.8 remains valid if we replace D by a positive Lebesgue measure subset E:

Lemma 11.12. *Given any unstable domain D and any positive Lebesgue measure set $E \subset D$, every accumulation point of*

$$\mu_{n,E} = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \Big(\frac{m_E}{m_D(E)} \Big)$$

is a Gibbs u-state with densities satisfying (30).

Proof. Given any $\delta > 0$ we may find pairwise disjoint domains $D_1, \ldots, D_s$ in D such that the relative Lebesgue measure of E inside each D_i is larger than $1 - \delta$ and the total measure of E outside the union of the D_i is less than $\delta m_D(E)$. This follows from a Lebesgue density argument: just decompose D into small "cubes", and pick as D_i those which are entirely contained in D (no intersection with the boundary) and have a large fraction occupied by E. Then, for any $j \geq 1$, we may write

$$f_*^j \Big(\frac{m_E}{m_D(E)} \Big) = \sum_{i=1}^{s} \frac{m_D(D_i)}{m_D(E)} f_*^j \Big(\frac{m_{D_i}}{m_D(D_i)} \Big) + \frac{1}{m_D(E)} f_*^j m_{(E \setminus \cup_{i=1}^s D_i)} - \frac{1}{m_D(E)} \sum_{i=1}^{s} f_*^j m_{D_i \setminus E} \,.$$

The total masses of both the second and the third term do not depend on j, and are less than δ. Therefore, every accumulation point of $\mu_{n,E}$ differs from an accumulation point of

$$\sum_{i=1}^{s} \frac{m_D(D_i)}{m_D(E)} \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \Big(\frac{m_{D_i}}{m_D(D_i)} \Big)$$

by a measure whose total mass is less than δ. By Theorem 11.8 applied to each domain D_i, every accumulation point of this last sequence is a Gibbs u-state, with densities satisfying (30); in particular, they are uniformly bounded. Since this is a closed property, making δ go to zero we get that every accumulation point of $\mu_{n,E}$ is a Gibbs u-state with densities satisfying (30). $\square$

Lemma 11.13. *Ergodic components of any Gibbs u-state μ are Gibbs u-states whose densities satisfy* (30).

Proof. Let $B(f)$ be the full μ-measure set of points where forward and backward time averages converge and coincide. Since μ is absolutely continuous along strong unstable manifolds, μ-almost every point x is contained in an unstable disk which intersects $B(f)$ on a positive Lebesgue measure subset D_x containing x. The limit time averages of all points $y \in D_x$ coincide with μ_x . Thus,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \Big(\frac{m_{D_x}}{m_D(D_x)} \Big) = \lim_{n\to\infty} \frac{1}{m_D(D_x)} \int_{D_x} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(y)} \, dm_D(y) = \mu_x \,.$$

On the other hand, by Lemma 11.12 the limit on the left is a Gibbs u-state with densities satisfying (30). $\square$

Now we can prove that our Gibbs u-state satisfy all the conditions in [272]:

Corollary 11.14. *If μ is any Gibbs u-state then*

(1) *It has densities with respect to Lebesgue measure along unstable plaques satisfying* (30).
(2) *The densities are uniformly bounded from zero and infinity, and the support of μ consists of entire strong unstable leaves.*

Proof. For part 1 just note that each strong unstable leaf intersects at most one ergodic component, so that the density is given by that of the ergodic component. The first statement in 2 is a direct consequence of part 1. The densities being bounded from zero implies that the support consists of entire leaves. □

Remark 11.15. It follows that the set of Gibbs u-states is closed, hence compact, relative to the weak* topology in the space $\mathcal{M}$ of probabilities. If Λ is a topological attractor (maximal invariant set in a forward invariant neighborhood), we even have that the set of pairs (μ, g) such that μ is a Gibbs u-stated supported in the continuation of Λ for a nearby diffeomorphism g is closed in the product of $\mathcal{M}$ by a neighborhood of f.

11.2.3. *Every SRB measure is a Gibbs u-state.* When Λ is a topological attractor with partially hyperbolic splitting $T_\Lambda M = E^u \oplus E^{cs}$, it is easy to see that every SRB measure μ supported in Λ must be a Gibbs u-state. Indeed, consider any disk D inside $B(\Lambda)$ transverse to the center stable sub-bundle and intersecting the basin of μ on a positive Lebesgue measure subset D_0. On the one hand,

$$\frac{1}{n}\sum_{j=0}^{n-1} f_*^j\left(\frac{m_{D_0}}{m_D(D_0)}\right) = \frac{1}{m_D(D_0)}\int_{D_0} \frac{1}{n}\sum_{j=0}^{n-1} \delta_{f^j(x)}\, dm_D(x)$$

converges to μ when $n \to \infty$. On the other hand, combining Remark 11.11 and Lemma 11.12, the limit must be a Gibbs u-state.

Here we are going to prove a stronger fact: *for Lebesgue almost every point on any unstable disk, all accumulation points of time averages are Gibbs u-states.* We only assume that Λ admits a partially hyperbolic splitting $T_\Lambda M = E^u \oplus E^{cs}$ and consists of entire strong unstable leaves.

Theorem 11.16. *There exists $E \subset \Lambda$ intersecting every unstable disk on a full Lebesgue measure subset, such that for any $x \in E$ every accumulation point ν of*

$$\nu_{n,x} = \frac{1}{n}\sum_{j=0}^{n-1} \delta_{f^j(x)}$$

is a Gibbs u-state.

In the remainder of this section we prove this theorem. Let D be any unstable domain, and $X \times Y$ be a foliated box. Our goal is to obtain a bound like (32) for every measurable subset A of X. It is no restriction to take X with radius less than 1/10. and to suppose that A itself is a disk, contained in the unstable disk with the same center as X and radius 10 times smaller. We say that A is far from the boundary of X.

Fix $C_* > 0$ such that the Lebesgue measure of any unstable disk is at most C_* times larger than the Lebesgue measure of any unstable disk with half the radius. C_* may be chosen depending only on the dimension of E^u and a bound on the curvature of strong unstable leaves. Define $K_0 = C_0 C_*^2$ where C_0 is the distortion bound given by Lemma 11.9.

Proposition 11.17. *Under the previous hypotheses,there exists $c > 0$ such that, given any sub-disk A far from the boundary of X, there exists $n_0 \geq 1$ such that, for all $n \geq n_0$,*

$$m_D\left(\{x \in D : \nu_{n,x}(A \times Y) \geq 10K_0 \frac{m_X(A)}{m_X(X)}\}\right) \leq e^{-cn}.$$

The first step in the proof of this proposition is a Markov type construction. Replacing f by some iterate f^ℓ, if necessary, we may assume that the least expansion $\sigma = \|Df^{-1} \mid E^u\|^{-1}$ along the unstable direction satisfies

$$\sigma \operatorname{diam}(A) > 1.$$

Observe that we have the right to do this, since the limits of time averages $\nu_{n,x}$ of f are given by the limits of time averages $\nu_{n,f^i(x),\ell}$ of f^ℓ, averaged over $0 \leq i < \ell$.

Lemma 11.18. *There exist families $\{\hat{X}_y : y \in Y\}$ and $\{\hat{A}_y : y \in Y\}$ of unstable domains such that*

(1) *$X \times \{y\} \subset \hat{X}_y$ and $\hat{X}_y$ is contained in the unstable disk with the same center as $X \times \{y\}$ and radius twice as large;*
(2) *if $f^n(X_y)$, $n \geq 1$ intersects some $X \times \{z\}$ then $f^n(\hat{X}_y)$ contains $\hat{X}_z$*
(3) *$A \times \{y\} \subset \hat{A}_y \subset X \times \{y\}$ and $\hat{A}_y$ is contained in the unstable disk with the same center as $A \times \{y\}$ and radius twice as large;*
(4) *if $f^n(A \times \{y\})$, $n \geq 1$ intersects some $X \times \{z\}$ then $f^n(\hat{A}_y)$ contains $\hat{X}_z$.*

Proof. First we take $\hat{X}_y$ to be the (increasing) union of the sequence $(\hat{X}_y^n)_n$ defined by $\hat{X}_y^0 = X \times \{y\}$ and $\hat{X}_y^n =$ the union of $\hat{X}_y^{n-1}$ with all pre-images $f^{-n}(\hat{X}_z^0)$, $z \in Y$ that intersect $\hat{X}_y^{n-1}$. The diameter of each $f^{-n}(\hat{X}_z^0)$ is bounded by $\sigma^{-n} \operatorname{diam}(X)$. As $\sigma > (\operatorname{diam} A)^{-1}$ is much larger than 1, it follows that $\hat{X}_y^n$ is contained in the neighborhood of radius

$$\sum_{i=1}^{n} \sigma^{-i} \operatorname{diam}(X) < \frac{1}{2} \operatorname{diam}(X)$$

around $X \times \{y\}$. Claim 1 follows immediately.

By definition, if $f^{-n}(\hat{X}_z^0)$ intersects $\hat{X}_y^0$ then it is contained in $\hat{X}_y^n$. Now let $i \geq 1$ and assume we have shown that $f^{-n}(\hat{X}_z^{i-1})$ is contained in $\hat{X}_y^{n-1+i}$. Consider any w such that $f^{-i}(\hat{X}_w^0)$ intersects X_z^{i-1} and, thus, is part of $\hat{X}_z^i$. Then $f^{-n-i}(\hat{X}_w^0)$ intersects $f^{-n}(\hat{X}_z^{i-1})$ and, consequently, also $\hat{X}_y^{n+i-1}$. So, by definition, $f^{-n-i}(\hat{X}_w^0)$ is part of $\hat{X}_y^{n+i}$. This proves that $f^{-n}(\hat{X}_z^i)$ is contained in $\hat{X}_y^{n+i}$. By induction, we have established this fact for every $i \geq 1$. This implies that $f^{-n}(\hat{X}_z) \subset \hat{X}_y$, and so claim 2 is proved.

Next, define $\hat{A}_y$ as the union of $A \times \{y\}$ with all the pre-images $f^{-n}(\hat{X}_z)$, $n \geq 1$ that intersect $A \times \{y\}$. The diameter of any of these pre-images is bounded by $\sigma^{-1} \operatorname{diam}(\hat{X}_z) < \operatorname{diam}(A) \operatorname{diam}(\hat{X}_z) \leq \frac{1}{2} \operatorname{diam}(A)$. Hence, $\hat{A}_y$ is contained in a disk with the same center as A and radius twice as large. In particular, it is contained in $X \times \{y\}$, as claimed in part 3.

Finally, if $f^{-n}(\hat{X}_z^0)$ intersects $A \times \{y\}$, so does $f^{-n}(\hat{X}_z)$. Then, by definition, $f^{-n}(\hat{X}_z)$ is contained in $\hat{A}_y$. This proves claim 4. $\square$

Lemma 11.19. *Let $1 \le k \le n$ and $0 \le n_1 < \cdots < n_k < n$ be fixed. Then the m_D-measure of the set of points $x \in D$ which hit $A \times Y$ at times $n_1, \ldots, n_k$ is bounded by $(K_0 m_X(A)/m_X(X))^k$.*

Proof. Let Δ be the union of D with all the pre-images $f^{-n_1}(\hat{X}_y)$ that intersect it. The diameter of each pre-image is bounded by $\sigma^{-n_1} 2 \operatorname{diam}(X) \ll 1$, and so Δ is only slightly larger than D. Since $D \subset \Delta$, it suffices to prove that the m_Δ measure of the set of points $x \in \Delta$ which hit A at times $n_1, \ldots, n_k$ is bounded by $(K_0 m_X(A)/m_X(X))^k$.

By construction, if $f^{n_1}(\Delta)$ intersects some $X \times \{y_1\}$ then it contains it. Let $\Delta_X(y_1) = \Delta \cap f^{-n_1}(X \times \{y_1\})$ and $\Delta_A(y_1) = \Delta \cap f^{-n_1}(A \times \{y_1\})$. The union Δ_A^1 of the $\Delta_A(y_1)$ over all these y_1 contains the set of points $x \in \Delta$ which hit $A \times Y$ at time n_1. Given any $1 \le i < k$, suppose we have subsets $\Delta_X(y_1, \ldots, y_i) \subset \Delta_A(y_1, \ldots, y_i)$ of Δ such that

(a) f^{n_i} maps $\Delta_X(y_1, \ldots, y_i)$ and $\Delta_A(y_1, \ldots, y_i)$ diffeomorphically onto $\hat{X}_{y_i}$ and $\hat{A}_{y_i}$, respectively;
(b) the union Δ_A^i of the $\Delta_A(y_1, \ldots, y_i)$ over all $y_1, \ldots, y_i$ for which it is defined contains the set of points $x \in \Delta$ which hit $A \times Y$ at times $n_1, \ldots, n_i$.

For each such subset, consider all $y_{i+1} \in Y$ such that $f^{n_{i+1}-n_i}(A \times \{y_i\})$ intersects $A \times \{y_{i+1}\}$. By part 4 of Lemma 11.18,

$$f^{n_{i+1}}(\Delta_A(y_1, \ldots, y_i)) = f^{n_{i+1}-n_i}(\hat{A}_{y_i}) \supset \hat{X}_{y_{i+1}}.$$

Define $D_X(y_1, \ldots, y_i, y_{i+1})$ and $D_X(y_1, \ldots, y_i, y_{i+1})$ to be the pre-images of $\hat{X}_{y_{i+1}}$ and $\hat{A}_{y_{i+1}}$ under $f^{n_{i+1}}$. Observe that the union Δ_A^i of the $\Delta_A(y_1, \ldots, y_i, y_{i+1})$ over the all $y_1, \ldots, y_i, y_{i+1}$ for which it is defined contains the set of points $x \in \Delta$ which hit $A \times Y$ at times $n_1, \ldots, n_i, n_{i+1}$. This follows from the inductive hypotheses, together with the observation that if $f^{n_{i+1}-n_i}(A \times \{y_i\})$ is disjoint from $A \times \{y\}$ then there are no points in $\Delta_A(y_1, \ldots, y_i)$ hitting $A \times Y$ at times $n_1, \ldots, n_i$ and $A \times \{y\}$ at time n_{i+1}. Thus, we have recovered at time n_{i+1} the inductive properties (a) and (b) above.

After k steps we obtain a family of subsets $\Delta_X(y_1, \ldots, y_k) \subset \Delta_A(y_1, \ldots, y_k)$ of Δ satisfying (a) and (b) for $i = k$. Using Lemma 11.9 and parts 1 and 3 of Lemma 11.18,

$$\frac{m_\Delta(\Delta_A(y_1, \ldots, y_i))}{m_\Delta(\Delta_X(y_1, \ldots, y_i))} \le C_0 \frac{m_X(\hat{A}_{y_i})}{m_X(\hat{X}_{y_i})} \le K_0 \frac{m_X(A)}{m_X(X)}$$

for every $1 \le i \le k$. Recall the definition of $K_0 = C\, C_*^2$ preceding the statement of Proposition 11.17. Denoting Δ_X^i the union of all $\Delta_X(y_1, \ldots, y_i)$, we deduce that

$$\frac{m_\Delta(\Delta_A^i)}{m_\Delta(\Delta_A^{i-1})} \le \frac{m_\Delta(\Delta_A^i)}{m_\Delta(\Delta_X^i)} \le K_0 \frac{m_X(A)}{m_X(X)}$$

for every $1 < i \le k$. This is also true for $i = 1$, if one interprets $\Delta_A^0 = \Delta$. Multiplying these inequalities over all i, we find that

$$m_\Delta(\Delta_A^i) \le \left(K_0 \frac{m_X(A)}{m_X(X)}\right)^k$$

and this implies the claim. ∎

We are ready to give the

Proof of Proposition 11.17. By definition, $n\nu_{n,x}(A \times Y)$ is the number of times a point x hits $A \times Y$ during the first n iterates. Thus, in view of Lemma 11.19, the m_D-measure of the set of points with $\nu_{n,x}(A \times Y) > 10K_0 m_X(A)/m_X(X)$ is bounded by

$$\sum_{k>10nK_0 m_X(A)/m_X(X)} \binom{n}{k} \left(K_0 \frac{m_X(A)}{m_X(X)}\right)^k.$$

The first factor counts the sequences $0 \le n_1 < \cdots < n_k < n$ of hitting times. To conclude the proof, use Lemma 11.20 below with $\alpha = K_0 m_X(A)/m_X(X)$ and $\beta = 10\alpha$. □

Lemma 11.20. *Given $0 < 10\alpha \le \beta \le 1/2$ there exist constants $C > 1 > c > 0$ such that*

$$\sum_{k>\beta n} \binom{n}{k} \alpha^k \le Ce^{-cn} \quad \textit{for all } n \ge 1.$$

Proof. Stirling's formula implies, for every $0 \le j \le k$,

$$\binom{n}{k} \le B \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k}$$

where B is a universal constant. We claim that

$$\left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} \alpha^k \le e^{-\beta k} \quad \text{for all } k > \beta n. \tag{33}$$

This implies

$$\sum_{k>\beta n} \binom{n}{k} \alpha^k \le \sum_{k>\beta n} e^{-\beta k} \le Ce^{-\beta^2 n}$$

for some constant $C > 1$ which depends only on β. Taking $c = \beta^2$ we obtain the conclusion of the lemma. Hence, we are left to prove the claim (33).

For this purpose, we write the inequality as

$$x \log \frac{1}{x} + (1-x)\log\frac{1}{1-x} - x\log\frac{1}{\alpha} \le -\beta x \quad \text{for } x > \beta,$$

where $x = k/n$ and we have taken n:th roots on both sides. Since the sum of the first two terms is a concave function of x, the statement holds if only if it holds at the point $x = \beta$:

$$\beta \log\frac{1}{\beta} + (1-\beta)\log\frac{1}{1-\beta} - \beta\log\frac{1}{\alpha} \le -\beta^2 .$$

Since β is not far from zero, $\log 1/(1-\beta) \le 2\beta$. Hence, for the previous inequality it suffices that

$$\beta\log\frac{1}{\beta} + (1-\beta)2\beta - \beta\log\frac{1}{\alpha} \le -\beta^2 \quad \Leftrightarrow \quad \log\frac{\alpha}{\beta} + 2 \le \beta .$$

The assumption $10\alpha \le \beta$ implies that this last inequality is indeed true. Hence, (33) is also true, and that completes the proof of the lemma. □

Finally, we give the

Proof of Theorem 11.16. Let D be an unstable domain. Consider any foliated box $X \times Y$ as before. Proposition 11.17, together with a Borel-Cantelli argument, proves that, given any disk $A \subset X$ far from the boundary of X, there exists a full m_D-measure subset of points $x \in D$ such that

$$\nu_{n,x}(A \times Y) \leq 10K_0 \frac{m_X(A)}{m_X(X)}$$

for all but, at most, finitely many $n \geq 1$. Consequently, the same upper bound holds for $\nu(A \times Y)$, for any accumulation measure ν of the sequence $\nu_{n,x}$. Applying this conclusion to a countable generating family of disks far from the boundary of X, and then taking the intersection of the corresponding full measure subsets of D, we get a full m_D-measure subset of points $x \in D$ for which every accumulation measure ν satisfies

$$\nu(A \times Y) \leq 10K_0 \frac{m_X(A)}{m_X(X)}.$$

for any measurable set A far from the boundary of X.

To complete the proof, cover the manifold by foliated boxes $X \times Y$ such that the radius of X is fairly small, and the foliated boxes obtained replacing each X by the disk $\bar{X}$ with the same center and half the radius still cover M. If ν does not give weight to $\bar{X} \times Y$, there is nothing to prove. Otherwise, by the previous paragraph,

$$\frac{\nu(A \times Y)}{\nu(\bar{X} \times Y)} \leq \frac{10K_0}{\nu(\bar{X} \times Y)} \frac{m_X(A)}{m_X(X)} \leq \frac{10K_0 C_*}{\nu(\bar{X} \times Y)} \frac{m_X(A)}{m_X(\bar{X})},$$

for every measurable set $A \subset \bar{X}$ (the constant C_* was defined prior to the statement of Proposition 11.17). This implies that ν is a Gibbs u-state, as claimed. $\square$

Remark 11.21. When $\Lambda = M$ the set E has full Lebesgue measure in M. This follows immediately by absolute continuity of the strong unstable foliation.

Theorem 11.16 has the following immediate consequence for conservative diffeomorphisms. Simply note that, by the ergodic theorem, forward and backward time averages exist and coincide at almost every point.

Corollary 11.22. *Suppose $f : M \to M$ is a volume preserving C^2 diffeomorphism admitting a partially hyperbolic splitting $TM = E^u \oplus E^c \oplus E^S$ where E^u and E^s have positive dimensions. For Lebesgue almost every point $x \in M$, the time average*

$$\lim_{n \to \pm\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)}$$

is, simultaneously, a Gibbs u-state and a Gibbs s-state.

This motivates the following

Conjecture 11.23. If f is a volume preserving partially hyperbolic C^2 diffeomorphism having the accessibility property, then it is ergodic with respect to Lebesgue measure.

11.2.4. *Mostly contracting central direction.* In general, Gibbs u-states are not SRB measures, even if they are ergodic. However,

Proposition 11.24. *Any ergodic Gibbs u-state μ whose Lyapunov exponents along the center stable sub-bundle E^{cs} are all negative is an SRB-measure.*

This is a simple consequence of Pesin theory (see Appendix C): by absolute continuity of the stable foliation and the definition of Gibbs u-states, the union of the local stable manifolds through the points $x \in \Lambda \cap B(\mu)$ form a positive Lebesgue measure subset of M contained in the basin of μ.

In this section we pursue this idea further on, leading to certain general criteria for existence and finiteness of SRB measures, and control of their basins. We consider Λ a topological attractor of a C^2 diffeomorphism $f : M \to M$, admitting a dominated splitting $T_\Lambda M = E^u \oplus E^{cs}$ such that E^u is uniformly expanding and has positive dimension.

Recall that the center stable sub-bundle admits a unique continuous invariant extension to the basin of Λ. For any $x \in B(\Lambda)$ we define

$$\lambda^c(x) = \limsup_{n\to+\infty} \frac{1}{n} \log \|Df^n \mid E^{cs}_x\| .$$

The main result is

Theorem 11.25 (Bonatti, Viana [67]). *Assume that $\lambda^c(x) < 0$ for a positive Lebesgue measure subset of every unstable disk D. Then there exist finitely many SRB measures $\mu_1, \dots, \mu_N$ supported in Λ, which are the ergodic Gibbs u-states of f. Moreover, the union of their basins is a full Lebesgue measure subset of the topological basin of attraction $B(\Lambda)$.*

The proof goes as follows. By Theorem 11.8 and Lemma 11.13, there exist ergodic Gibbs u-states μ_α. The hypothesis $\lambda^c < 0$ implies that the central Lyapunov exponents of every μ_α are all negative. Hence, by Proposition 11.24, every ergodic Gibbs u-state is an SRB measure. If there were infinitely many μ_α then we could find a sequence μ_{α_n} accumulating some ergodic Gibbs state μ; recall Remark 11.15. Using that all the supports are saturated by strong unstable manifolds, by Corollary 11.14, we would find that the stable manifolds of points in $B(\mu)$ would capture a positive measure subset of points in $B(\mu_{\alpha_n})$ for all large n, which is a contradiction. Thus, there is only a finite number of ergodic Gibbs u-states $\mu_1, \dots, \mu_N$. Finally, suppose there was a positive Lebesgue measure subset $A \subset B(\Lambda)$ disjoint from the basins of all these SRB measures μ_i. Consider a disk transverse to the center stable sub-bundle and intersecting A on a positive Lebesgue measure subset. Using Lemmas 11.12 and 11.13, we would construct a new ergodic Gibbs u-state, contradicting the definition of N.

Kan's construction in Section 11.1.1 shows that one can not expect the SRB measure to be unique, even if the attractor is transitive. Indeed, his examples are partially hyperbolic with splitting $E^u \oplus E^{cs}$ and the SRB measures are ergodic Gibbs u-states with negative central Lyapunov exponents (in the endomorphism version one should consider the natural extension).

On the other hand, as a fairly direct consequence of Theorem 11.25, one gets

Theorem 11.26 (Bonatti, Viana [67]). *Assume the strong unstable foliation of Λ is minimal, and there exists some unstable disk $\hat{D}$ such that $\lambda^c(x) < 0$ for a positive Lebesgue measure subset of D. Then Λ supports a unique SRB measure μ, and $B(\mu) = B(\Lambda)$ up to zero Lebesgue measure.*

The minimality assumption gives that every unstable disk D has an iterate close to D_0. This permits to recover the assumptions of Theorem 11.25 by considering the intersection of the iterate with stable manifolds of points in D_0. Thus, there exist finitely many SRB measures, which are ergodic Gibbs u-states with negative central exponents.Using minimality once more, one concludes that the SRB measure is unique.

For the conclusion of Theorem 11.25 it is not enough to suppose that $\lambda^c < 0$ Lebesgue almost everywhere in $B(\Lambda)$. For instance, [37, 120] gives such an example with $\Lambda = M$ and an infinite number of SRB measures. The problem is that absolute continuity of the strong unstable foliation ensures $\lambda^c < 0$ at m_D-almost every point only inside *almost every* (not every) unstable disk D. On the other hand, the much stronger methods we are going to discuss in Section 11.3 permit to prove

Theorem 11.27. *Assume the center stable sub-bundle E^{cs} has a dominated splitting $E^{cs} = E^c \oplus E^s$ where E^c is one-dimensional and E^s is uniformly contracting. Let $c > 0$ and $H \subset B(\Lambda)$ be a positive Lebesgue measure set such that $\lambda^c(x) < -c$ for almost every $x \in H$. Then there are finitely many SRB measures, which are ergodic Gibbs u-states, such that H is contained* mod 0 *in the union of their basins.*

The stable sub-bundle E^s may have dimension zero. The theorem will be part of a more general result, Theorem 11.35, which needs no assumption on the dimension of the central sub-bundle.

11.2.5. *Differentiability of Gibbs u-states.* Another important problem concerns the variation of SRB measures as a function of the dynamical system. In this direction, Dolgopyat [118] proves a differentiability result at certain partially hyperbolic maps f admitting a unique SRB measure, an ergodic Gibbs u-state, whose basin has full Lebesgue measure in the ambient. We refer the reader to [118] for the precise statement.

Here we just mention a striking application of Dolgopyat's main result to the time-1 map f of the geodesic flow in the unit tangent bundle $M = T^1 S$ of a surface S with negative curvature. It is well-known that f is volume preserving and partially hyperbolic, with splitting $TM = E^u \oplus E^c \oplus E^s$. The central bundle E^c corresponds to the direction of the flow. The Lebesgue measure is ergodic and, hence, the unique SRB measure for f. Furthermore, [142] prove that f is stably ergodic: the Lebesgue measure is ergodic for every volume preserving map in a C^2 neighborhood of f.

Let us consider smooth one-parameter families of diffeomorphisms $(f_t)_t$ through $f_0 = f$, volume preserving or not. Of course, f_t is partially hyperbolic for every small t. For any f_t-invariant probability ν, define the integrated central Lyapunov exponent

$$\lambda^c(f_t\,, \nu) = \int \log |Df_t \mid E^c| \, d\nu \, .$$

Theorem 11.28 (Dolgopyat [118])**.** *There exists a quadratic form $c(\cdot)$ in the space of vector fields on M such that*

$$\lambda^c(f_t, \nu_t) = c\big(\frac{\partial f_t}{\partial t}\big)\, t^2 + o(t^2)$$

for every Gibbs u-state ν_t of f_t. Moreover, $c(\cdot)$ is not identically zero, even restricted to divergence free vector fields.

As usually, $o(t^2)$ represents an expression going to zero faster than t^2 when $t \to 0$. Here are some consequences of this theorem:

(1) For generic families of volume preserving diffeomorphisms through f, and for all small t, either f_t or f_t^{-1} has negative central Lyapunov exponent.
(2) Moreover, (f_t, m) has exponential decay of correlations and satisfies the central limit theorem, for any small t, and so does (g, m) for any volume preserving diffeomorphism g in a neighborhood of f_t.
(3) For generic families of diffeomorphisms through f, volume preserving or not, and for all small t, the diffeomorphism f_t has a unique SRB measure.

Related to (1), Baraviera, Bonatti [36] prove that one can always perturb a volume preserving partially hyperbolic diffeomorphism, in the C^1 topology, to render the sum of the integrated central Lyapunov exponents different from zero.

Related to (2), Castro [90] and Dolgopyat [119, 117] had proven exponential decay of correlations and limit theorems for other classes of partially hyperbolic diffeomorphisms with splitting $E^u \oplus E^{cs}$, especially when the central sub-bundle is mostly contracting.

11.3. **SRB measures for dominated dynamics.** In this section we drop the requirement of existence of a uniformly expanding sub-bundle: the only standing assumptions are existence of a dominated splitting $T_K M = E^{cu} \oplus E^{cs}$ over a compact forward invariant set K, and C^2 regularity of the diffeomorphism.

This makes the situation much more subtle since, in the absence of uniform expansion, distortion control is not guaranteed. Recall that expansion was a key ingredient for Lemma 11.9. On the other hand, domination does provide the other key ingredient, namely Hölder control of the Jacobian:

Proposition 11.29. *There exists $\gamma > 0$ and, given $L > 0$ and any C^2 disk $D \subset K$ transverse to the center stable direction E^{cs}, there exists $C > 0$ such that*

$$x \mapsto \log |\det(Df \mid T_y f^n(D))|$$

is (C, γ)-Hölder on every domain of diameter less than L inside any $f^n(D)$, $n \geq 1$.

The crucial idea to overcome the lack of uniform expansion is the concept of *hyperbolic times*, introduced by Alves [4] and which we already mentioned in Section 11.1.3.

11.3.1. *Non-uniformly expanding maps.* To explain how the use of hyperbolic times permits to handle distortion, without having to deal with additional geometric difficulties, we begin by exploring this idea in a context of non-invertible dynamics.

A local diffeomorphism $g : M \to M$ on a compact manifold is *non-uniformly expanding* on a set H (not necessarily invariant), if there exists $c_0 > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \sum_{j=0}^{n-1} \|Dg(g^j(z))^{-1}\| < -c_0 \tag{34}$$

at Lebesgue almost every $z \in H$. Note that $Dg(\zeta)^{-1}$ represents the inverse of the derivative $Dg(\zeta)$ at a point $\zeta \in M$. If $\dim M = 1$, this condition just means that the (lower) Lyapunov exponent is bounded from zero on H:

$$\liminf_{n \to +\infty} \frac{1}{n} \log |Dg^n(z)| > c_0.$$

Theorem 11.30 (Alves, Bonatti, Viana [7]). *Suppose $g : M \to M$ is a C^2 local diffeomorphism non-uniformly expanding on some positive Lebesgue measure set*

H. Then H is contained mod 0 *in the union of the basins of finitely many SRB measures, which are ergodic measures absolutely continuous with respect to Lebesgue measure.*

In particular, *if the limit in* (34) *is negative Lebesgue almost everywhere in M then Lebesgue almost every point is in the basin of some SRB measure.* Observe that this implies that time averages converge Lebesgue almost everywhere. In addition, if the limit is less than some $-c_0 < 0$, the SRB measures are finitely many.

Let us comment on the proof of Theorem 11.30. Fix some $c \in (0, c_0)$. An iterate n is a *hyperbolic time* for a point z if

$$\prod_{j=n-k}^{n-1} \|Dg(g^j(z))^{-1}\| \leq e^{-ck} \quad \text{for every } 1 \leq k \leq n.$$

On the one hand, using Lemma 11.5 one proves that for every point z as in (34) the set of hyperbolic times has positive density at infinity: a lower bound on the density depends only on c_0, c, and a bound on the C^1 norm of f. On the other hand, at hyperbolic times we have uniform control of the distortion:

Lemma 11.31. *There exists $\rho > 0$ and $C_0 > 0$ such that, given any pair (z, n) such that n is a hyperbolic time for z, and given any $1 \leq k \leq n$, the inverse branch g_k of g^k mapping $g^n(z)$ to $g^{n-k}(z)$ satisfies*

(1) *g_k is well defined and is a $e^{-ck/2}$-contraction on the ρ-ball around $g^n(z)$;*
(2) *$\det(Dg_k(x)) \leq C \det(Dg_k(y))$ for any x and y in the ρ-ball around $g^n(z)$.*

The first claim is proved by recurrence on k, using uniform continuity of the derivative and its inverse. The second one is a consequence, using also that the logarithm of the Jacobian is Lipschitz continuous (Hölder suffices).

Now, existence of absolutely continuous invariant probabilities can be proved as follows. Let m be Lebesgue measure on M, and

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} g_*^j m.$$

For each j let H_j is the set of points $z \in H$ for which j is a hyperbolic time. Then we decompose $\mu_n = \xi_n + \eta_n$ where

$$\xi_n = \frac{1}{n} \sum_{j=0}^{n-1} g_*^j(m \mid H_j) \quad \text{and} \quad \eta_n = \frac{1}{n} \sum_{j=0}^{n-1} g_*^j(m \mid H_j^c).$$

Using part 2 of Lemma 11.31 one sees that every ξ_n is absolutely continuous with respect to Lebesgue measure, with uniformly bounded density. Using that the set of hyperbolic times has positive density at infinity, for almost every point in H, we find $\kappa > 0$ such that $\xi_n(M) \geq \kappa$ for all large n. Taking the limit along some convenient subsequence we find $\mu = \xi + \eta$ where μ is an g-invariant probability, and ξ is an absolutely continuous measure with total mass $\xi(M) \geq \kappa$. Let $\mu = \mu_a + \mu_s$ be the decomposition of μ as the sum of an absolutely continuous measure and a totally singular measure, relative to Lebesgue measure. Then both μ_a and μ_s are g-invariant and $\mu_a(M) \geq \xi(M) \geq \kappa$. Thus, normalizing μ_a we obtain an absolutely continuous g-invariant probability.

The proof proceeds by ergodic decomposition. Using Lemma 11.31 and a density point argument, we see that any g-invariant set with positive ξ-measure must contain a full Lebesgue subset of some ρ-ball. Thus, up to zero measure subsets, there are only finitely many of these invariant sets. Taking the normalized restriction of μ_a to such an invariant set with minimal measure, we obtain an ergodic absolutely continuous probability measure μ_1, and we also check that the basin $B(\mu_1)$ intersects H on a positive Lebesgue measure subset. Repeating the whole argument with $H \setminus B(\mu_1)$ in the place of H, we exhaust H by basins of ergodic absolutely continuous probabilities, after a finite number of steps.

Remark 11.32. A version of Theorem 11.30 for piecewise smooth maps, possibly with critical points, is also given in [7]. One requires g to be a local diffeomorphism only on the complement of some singular set $\mathcal{S} \subset M$, with polynomial C^2 bounds close to $\mathcal{S}$: the norms of $Dg(z)$, $Dg^{-1}(z)$, $D^2g(z)$ are less than $C \operatorname{dist}(z, \mathcal{S})^{-c}$. In this setting, to control distortion it is necessary to keep a hold on visits to the neighborhood of $\mathcal{S}$. In addition to (34), one assumes that given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\liminf_{n \to +\infty} \frac{1}{n} \sum_j \log \operatorname{dist}(g^j(z), \mathcal{S}) \geq -\varepsilon \tag{35}$$

for Lebesgue almost every $z \in H$, where the sum is over all $0 \leq j \leq n-1$ for which the distance is less than δ. Then the full statement of Theorem 11.30 remains valid for these maps.

An example are the cylinder maps treated in Section 11.1.3, where the singular set is a circle. Let us also point out that Alves, Araújo [5] have recently proven stochastic stability for a large class of transformations as in Remark 11.32.

11.3.2. *Existence of Gibbs cu-states.* Now we apply these ideas to C^2 diffeomorphisms having a dominated splitting $E^{cu} \oplus E^{cs}$ over some positively invariant set K.

An invariant probability measure μ is called a *Gibbs cu-state* if the $\dim E^{cu}$ largest exponents are positive at μ-almost every point and μ is absolutely continuous along the corresponding Pesin strong unstable manifolds. See Appendix C. We say that E^{cu} is *non-uniformly expanding* on a set $H \subset K$ if there is $c_0 > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \left\| \left(Df \mid E^{cu}_{f^j(x)}\right)^{-1} \right\| \leq -c_0 < 0 \tag{36}$$

for Lebesgue almost every $z \in H$.

Theorem 11.33 (Alves, Bonatti, Viana [7]). *Suppose E^{cu} is non-uniformly expanding on a positive Lebesgue measure set $H \subset K$. Then f has some ergodic Gibbs cu-state supported inside K.*

For the proof one considers iterates $\mu_n = n^{-1} \sum_{j=0}^{n-1} f_*^j m_D$ of Lebesgue measure along some disk $D \subset K$ transverse to the center stable direction E^{cs}. Proposition 11.29 provides control of the geometry of the iterates of D. Otherwise, the argument goes along the same lines as that of Theorem 11.30. As before, the key idea is a notion of hyperbolic time: given $c > 0$ we say that n is a *cu-hyperbolic time* for a point $x \in K$ if for every $1 \leq k \leq n$

$$\prod_{j=n-k}^{n-1} \|(Df \mid E^{cu}_{f^j(x)})^{-1}\| \leq e^{-ck}.$$

The hypothesis (36) implies that x has infinitely many cu-hyperbolic times, with positive density at infinity. Decomposing $\mu_n = \xi_n + \eta_n$ according to cu-hyperbolic times, and considering convenient accumulation points, one constructs some Gibbs cu-state for f. The last step in the proof is an ergodic decomposition argument.

Just as in Proposition 11.24, we have that every ergodic Gibbs cu-state such that the remaining $\dim E^{cs}$ Lyapunov exponents are negative is an SRB measure. In the next section we discuss some criteria that ensure this property.

11.3.3. *Simultaneous hyperbolic times.* We say that E^{cs} is *non-uniformly contracting* on a set H if the condition

$$\limsup_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \left\| \left(Df \mid E^{cs}_{f^j(x)}\right) \right\| \le -c_0 < 0 \tag{37}$$

holds almost everywhere on H. Observe that, just as in (36), we consider $n \to +\infty$. As before, if x satisfies (37) then the set of *cs-hyperbolic times* n

$$\prod_{j=n-k}^{n-1} \|Df \mid E^{cs}_{f^j(x)}\| \le e^{-ck} \quad \text{for all } 1 \le k \le n$$

has positive density at infinity. We say that n is a *simultaneous hyperbolic time* for x if it is both a cu- and a cs-hyperbolic time.

??

Theorem 11.34. *(Alves, Bonatti, Viana* [7]*) Let $H \subset K$ be a positive Lebesgue measure set such that almost every point in H has positive density at infinity of simultaneous hyperbolic times. Then H is covered* mod 0 *by the basins of finitely many SRB measures, which are ergodic Gibbs cu-states with* $\dim E^{cs}$ *negative Lyapunov exponents.*

In the remainder of this section we highlight applications cases of this theorem, corresponding to the cases when either of the invariant sub-bundles is uniform.

Clearly, if E^{cu} is uniformly expanding then every $n \ge 1$ is a hyperbolic time. So, in this case, to ensure the hypotheses of Theorem 11.34 we only need E^{cs} to be non-uniformly contracting:

Theorem 11.35. *Suppose the sub-bundle E^{cu} is uniformly expanding on K, and E^{cs} is non-uniformly contracting on a positive Lebesgue measure set $H \subset K$. Then H is covered* mod 0 *by the basins of finitely many SRB measures, which are ergodic Gibbs cu-states with* $\dim E^{cs}$ *negative Lyapunov exponents.*

Notice that this contains Theorem 11.27. Indeed, if E^{cs} splits as $E^c \oplus E^s$ with the $\dim E^c = 1$ and E^s uniformly contracting, then condition (37) reduces to

$$\limsup_{n\to\infty} \frac{1}{n} \log |Df^n \mid E^c_x| < -c_0 \,.$$

Moreover, in this context the notions of Gibbs cu-state and u-state coincide.

Similarly, if E^{cs} is uniformly contracting then any ergodic Gibbs cu-state is an SRB measure. Hence,

Theorem 11.36. *Suppose the sub-bundle E^{cs} is uniformly contracting on K, and E^{cu} is non-uniformly expanding on a positive Lebesgue measure set $H \subset K$. Then*

H is covered mod 0 *by the basins of finitely many SRB measures, which are ergodic Gibbs cu-states with* $\dim E^{cs}$ *negative Lyapunov exponents.*

Remark 11.37. When E^{cu} splits as $E^u \oplus E^c$ with E^u uniformly expanding, we may speak of both Gibbs cu-states and u-states. In fact, every cu-state is also a u-state, as a consequence of the absolute continuity of the strong unstable foliation. In particular, the support of a Gibbs cu-state always consists of complete strong unstable manifolds. In general, the support is not saturated by center unstable leaves, even assuming a center unstable foliation does exist.

Example 11.38. We describe a volume preserving diffeomorphism of the 3-torus with partially hyperbolic splitting $TM = E^u \oplus E^c \oplus E^s$, admitting several SRB measures which are ergodic Gibbs cu-states. Start with $f_0 = h \times \mathrm{id}$ where h is a linear Anosov diffeomorphism of the 2-dimensional torus and id is the identity map on the circle. Using the techniques in [36, 313] one may obtain a small C^1 perturbation f, coinciding with f_0 outside the union of pairwise disjoint horizontal strips $T^2 \times I_j$ and such that the average of the central Lyapunov exponent on each strip is positive. By the previous theorem, each $K = T^2 \times I_j$ contains the support of some ergodic Gibbs cu-state with 2 positive and 1 negative Lyapunov exponent.

In a very nice recent work, Vásquez [331] extends several results from Section 11.2 to Gibbs cu-states of topological attractors with a dominated splitting $E^{cu} \oplus E^{cs}$. He proves that the densities along Pesin strong unstable manifolds satisfy (30), at almost every point, and he deduces that the support consists of entire strong unstable manifolds. Assuming the center unstable bundle E^{cu} is non-uniformly expanding, in the sense of (36), Theorem 11.33 gives a constructive proof of the existence of Gibbs cu-states. Vásquez shows that, in fact, this construction produces *all* Gibbs cu-states. Most interesting, assuming the constant c_0 remains uniform in a neighborhood of the diffeomorphism, he proves continuity of cu-states with the dynamics in the C^2 topology: if $g_n \to g$ and μ_n is a Gibbs cu-state for g_n for each n, then any weak limit μ is a Gibbs cu-state for g. He also extends the connection between Gibbs states and SRB measures in Section 11.2.3 to this generality: SRB measures must be Gibbs cu-states. These conclusions allow him to prove that the SRB measure varies continuously with the diffeomorphism, for an open class of systems with dominated decomposition containing the construction in Section 7.3.4. Finally, Vásquez proves that the conclusion of Theorem 11.34 remains valid if the simultaneous hyperbolic times condition is replaced by

$$\limsup_{n\to+\infty} \frac{1}{n} \log \|Df^n \mid E^{cs}_x\| < 0$$

on a full Lebesgue measure subset of every disk transverse to the center stable bundle E^{cs} (compare Theorem 11.25).

11.4. **Generic existence of SRB measures. ?? ??**

11.5. **Further related results.**

11.5.1. *Zero-noise limit and the entropy formula.* The development of a theory of non-uniformly hyperbolic systems (stable manifold theorem, absolute continuity of foliations) allowed Pesin [274] to prove that if a $C^{1+\alpha}$ diffeomorphism $f : M \to M$

preserves a probability μ absolutely continuous with respect to Lebesgue measure, then the entropy of (f, μ) is given by

$$h_\mu(f) = \lambda_\mu^+(f), \tag{38}$$

where $\lambda_\mu^+(f)$ is the integral, with respect to μ, of the sum of all positive Lyapunov exponents at each point (with multiplicity).

This remarkable formula holds, more generally, if μ has absolutely continuous conditional measures along unstable manifolds; see [190]. Furthermore, this sufficient condition is also necessary for the entropy formula to hold, as shown by Ledrappier, Young [188, 191]. Combining this with the absolute continuity theorem C.4, one concludes that ergodic probabilities that satisfy the entropy formula are SRB measures, if the Lyapunov exponents are all different from zero: local stable manifolds through a typical local unstable manifold form a positive Lebesgue measure set contained in the basin.

Cowieson, Young [99] recently proposed to construct invariant probabilities satisfying the entropy formula as zero-noise levels of stationary measures of random perturbations of the system. They consider C^∞ diffeomorphisms with some attractor Λ, and perturbations of such diffeomorphisms by random maps, determined by probability distributions ν_ε supported in the ε-neighborhood of f; see Appendix D.2. Given a stationary probability μ_ε for the corresponding Markov chain, define $h_{\mu_\varepsilon} = \mathbf{??}$ and $\lambda^+_{\mu_\varepsilon} = \mathbf{??}$.

Theorem 11.39 (Cowieson, Young [99]). *Assume the diffeomorphism $f: M \to M$ is C^∞. Consider a perturbation scheme by random maps, as above, such that the corresponding transition probabilities $p_\varepsilon(\cdot, | \, x)$ are absolutely continuous with respect to Lebesgue measure. Let μ_ε be stationary measures for the Markov chain, and μ_0 be an accumulation point of μ_ε as $\varepsilon \to 0$.*

If $\lambda^+_{\mu_\varepsilon} \to \lambda^+_{\mu_0}$ as $\varepsilon \to 0$ then μ_0 satisfies the entropy formula: $h_{\mu_0} = \lambda^+_{\mu_0}$.

The proof combines three main ingredients. The first one is a random version of the entropy formula stating that $h_{\mu_\varepsilon} = \lambda^+_{\mu_\varepsilon}$. This uses the hypothesis that the noise is absolutely continuous. A second ingredient is the fact that h_{μ_ε} depends lower semi-continuously on ε, and hence $h_{\mu_0} \geq \limsup_{\varepsilon \to 0} h_{\mu_\varepsilon}$. This follows from a random version of work of Yomdin [349] and Newhouse [243] and depends on considering the C^∞ topology. The final main ingredient is Ruelle's inequality [301], according to which $h_{\mu_0} \leq \lambda_0^+$.

As an application, they consider the case when the attractor Λ admits a dominated splitting $T_\Lambda M = E^+ \oplus E^-$ such that for every point $x \in \Lambda$ and every vector $v \in E^+$ (respectively, $v \in E^-$) such that

$$\lim_{|n| \to \infty} \frac{1}{n} \log \|Df^n(x) v\|$$

exists the limit is non-negative (respectively, non-positive). This assumption implies that

$$\lambda^+_{\mu_0} = \int \log |\det Df \mid E^+| \, d\mu = \int \log |\det Df| \, d\mu - \int \log |\det Df \mid E^-| \, d\mu.$$

Moreover, just because the splitting is dominated,

$$\lambda^+_{\mu_\varepsilon} \geq \inf \int \log |\det Dg| \, d\mu_\varepsilon \, d\nu_\varepsilon(g) - \int \log |\det Df \mid E^-| \, d\mu_\varepsilon \, d\nu_\varepsilon(g).$$

Taking $\mu_\varepsilon \to \mu_0$, along some sequence of values of ε, we get that $\liminf \lambda_\varepsilon^+ \geq \lambda_0^+$, and this suffices for the argument in Theorem 11.39.

11.5.2. *Equilibrium states for non-hyperbolic maps.* Let $f : M \to M$ be a continuous transformation on a compact space, and $\phi : M \to \mathbb{R}$ be a continuous function. An f-invariant probability measure is an *equilibrium state* of f for ϕ if it maximizes the functional

$$F_{f,\phi}(\nu) = h_\nu(f) + \int \phi \, d\nu$$

among all f-invariant probability measures. By classical results of Sinai, Ruelle, Bowen [70, 303, 316], equilibrium states always exist if f is a uniformly expanding transformation or a uniformly hyperbolic diffeomorphism. Moreover, the equilibrium state is unique if the *potential* ϕ is Hölder continuous. In this setting the SRB measure coincides with the equilibrium state for the potential $\phi = -\log|\det Df \mid E^u|$.

Problem 11.40. Investigate the existence and uniqueness of equilibrium states for non-uniformly expanding maps and non-uniformly hyperbolic diffeomorphisms and flows.

??

One difficulty is to formulate the assumption of non-uniform hyperbolicity in a suitable fashion. Assuming (34) at *Lebesgue* almost every point permits to prove existence of SRB measures, as we have seen in Section 11.3.1. But that is a less useful condition for the construction of other equilibrium states, as most equilibrium states are likely to be singular with respect to Lebesgue measure.

It is tempting to require (34) to hold for almost every point, relative to any invariant probability, but that is much too strong: it implies that f is uniformly expanding, as observed by Alves, Araújo, Saussol [6]. This has been improved by Cao [86]: if every f-invariant probability has only positive Lyapunov exponents then f is uniformly expanding. His argument goes as follows.

Let ν be any f-invariant probability. By assumption, there exists $\delta > 0$ such that

$$\frac{1}{m}\int \log \|Df^m(x)^{-1}\|^{-1} \, d\nu(x) > \delta \tag{39}$$

for all $m \geq 1$. It is clear that we may fix m and δ such that the same is true for every probability in a neighborhood of μ, in the weak topology. So, since the space of invariant probabilities is weak-compact, we may fix m and δ so that (34) is true for every f-invariant probability. Replacing f by f^m, we may suppose that $m = 1$. Now, if f is not uniformly expanding, given any $k \geq 1$ there exists $x_k \in M$ such that

$$\prod_{j=0}^{k-1} \|Df(f^j(x_k))^{-1}\|^{-1} \leq \|Df^k(x_k)^{-1}\|^{-1} \leq 2.$$

Let ν_k be the average of Dirac measures over the first k iterates of x_k and ν be any accumulation point of the sequence ν_k. Then ν is an f-invariant probability and

$$\int \log \|Df(x)^{-1}\|^{-1}\, d\nu = \lim_k \int \log \|Df(x)^{-1}\|^{-1}\, d\nu_k$$
$$= \lim_k \frac{1}{k}\sum_{j=0}^{k-1} \log \|Df(f^j(x_k))^{-1}\|^{-1} \le \lim_k \frac{2}{k} = 0.$$

This contradicts (39), and so f must be uniformly expanding.

Nevertheless, a substantial contribution to the solution of Problem 11.40 was recently provided by Oliveira:

Theorem 11.41 (Oliveira [251]). *There exists $\tau > 0$ and an open set of local C^2 diffeomorphisms f of the d-dimensional torus, $d \ge 2$ which are not uniformly expanding (they exhibit periodic saddles) and which have equilibrium states for every continuous potential ϕ satisfying $\max\phi - \min\phi \le \tau h_{top}(f)$.*

Actually, Oliveira imposes a weaker condition $P_f(\phi) \ge \max\phi + (1-\tau)h_{top}(f)$ where $P_f(\phi)$ is the topological pressure [19] of ϕ. The way he overcomes the difficulty discussed above is by identifying a subset $\mathcal{K}$ of invariant probabilities which contains suitable candidates for the equilibrium state, and whose elements are, *a priori*, known to be non-uniformly expanding, in the sense that condition (34) holds almost everywhere. Indeed, $\mathcal{K}$ satisfies

- f is ν-expansive (uniformly) for every $\nu \in \mathcal{K}$: there exists a partition $\mathcal{P}$ of M such that ν-almost every point $x \in M$ is determined by its itinerary relative to $\mathcal{P}$, for every $\nu \in \mathcal{K}$;
- the supremum of $F_{f,\phi}$ on $\mathcal{K}$ is attained and coincides with the supremum on the whole set of f-invariant probabilities.

The first property, which uses (34), ensures that the supremum of $F_{f,\phi}$ is attained, because the functional is upper semi-continuous, whereas the second one implies that it is a global maximum.

In a more recent work [252], Oliveira proposes a different approach which also proves uniqueness of equilibrium states, under similar hypotheses. In both cases, the systems he considers are of the type described in Section 11.1.4. The hypothesis that the potential is not far from constant is important to ensure that the elements of $\mathcal{K}$ are indeed non-uniformly expanding: they give small weight to the regions where the map is not expanding. Currently, there is some evidence that, for potentials with large variation, equilibrium states may exist but fail to be non-uniformly expanding.

12. Lyapunov exponents

An invariant probability μ is *hyperbolic* for a diffeomorphism $f : M \to M$ if the Lyapunov exponents of f are all different from zero at μ-almost every point: this means that the tangent space splits as the sum of two subspaces which are exponentially contracted or expanded by all *large enough* iterates of the derivative. Pesin theory [273] provides detailed geometric information on the dynamics of such *non-uniformly hyperbolic* systems (f, μ), building on which several deep results have

[19] The pressure $P_f(\phi)$ is the supremum of $h_\nu(f) + \int \phi\, d\nu$ over all invariant probabilities ν. Notice that $P_f(\phi) \ge h_{top}(f) + \min\phi$.

been proved, such as the entropy formula for smooth measures [274], the existence of uniformly hyperbolic sets carrying large entropy [170], and the converse to the entropy formula [188, 191].

We are going to discuss a number of recent results in the direction of the following two fundamental problems: *How frequently are Lyapunov exponents different from zero ? How do Lyapunov exponents vary with the dynamics ?* These results establish a surprising connection between measurable objects such as Lyapunov exponents and Oseledets subspaces, and geometric and topological features of the dynamics such as existence of (Hölder) continuous invariant subbundles, opening the way for the use of geometric and topological tools in this intrinsically ergodic setting.

A survey on Lyapunov exponents has appeared recently, by Bochi, Viana [50], and the discussion that follows has some intersection with their own.

Let M be a compact manifold with dimension $d \geq 1$ and $f : M \to M$ be a C^r transformation, $r \geq 1$. The *largest Lyapunov exponent* of f at a point $x \in M$ is the exponential rate of growth

$$\lambda(f, x) = \lim_{n\to\infty} \frac{1}{n} \log \|Df^n(x)\|$$

of the iterates of the derivative over the orbit of x. It follows from the sub-additive ergodic theorem [180] that the limit exists almost everywhere, with respect to any f-invariant probability measure μ. It is clear that it does not depend on the choice of the Riemannian structure $\|\cdot\|$ on M. We denote

$$\lambda(f, \mu) = \int \lambda(f, x)\, d\mu.$$

A more detailed statement is provided by Oseledets theorem (see Appendix C): Assuming f is a diffeomorphism, for μ-almost all $x \in M$ there exist $k = k(f, x) \in \mathbb{N}$, real numbers $\hat\lambda_1(f, x) > \cdots > \hat\lambda_k(f, x)$, and a splitting

$$T_x M = E_x^1 \oplus \cdots \oplus E_x^k \tag{40}$$

of the tangent space, such that the splitting is invariant under Df, and

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df^n(x) v_i\| = \hat\lambda_i(f, x) \quad \text{for every non-zero } v_i \in E_x^i\,.$$

When f is non-invertible, instead of a splitting one gets a filtration into vector subspaces $T_x M = E_x^1 > \cdots > E_x^k > 0 = E_x^{k+1}$ such that

$$\lim_{n\to+\infty} \frac{1}{n} \log \|Df^n(x) v_i\| = \hat\lambda_i(f, x) \quad \text{for every } v_i \in E_x^i \setminus E_x^{i+1}\,.$$

In either case, the *Lyapunov exponents* $\hat\lambda_i(f, x)$ and the *Oseledets subspaces* E_x^i are unique μ-almost everywhere, and they vary measurably with the point x. Moreover, the Lyapunov exponents are constant on orbits, and so they are constant μ-almost everywhere if μ is ergodic. We also have $\lambda(f, x) = \hat\lambda_1(f, x)$ wherever they are defined.

12.1. **Continuity of exponents and projective hyperbolicity.** We begin by introducing a proper set-up for discussing dependence of Lyapunov exponents on the dynamical system. Let

$$\lambda_1(f, x) \geq \lambda_2(f, x) \geq \cdots \geq \lambda_d(f, x), \quad d = \dim M,$$

be the Lyapunov exponents of f at x *counted with multiplicity* (the dimension of the corresponding Oseledets subspace), and define

$$\lambda_i(f,\mu) = \int \lambda_i(f,x)\, d\mu(x) \quad \text{for } 1 \le i \le d.$$

We think of the $\lambda_i(f,\mu)$ as functions of f in the space $\mathrm{Diff}^1_\mu(M)$ of C^1 diffeomorphisms preserving a given probability measure μ, endowed with the C^1 topology. We also define

$$\Lambda_i(f,\mu) = \lambda_1(f,\mu) + \cdots + \lambda_i(f,\mu) \quad \text{and} \quad \mathrm{V}_i(f,\mu) = \lambda_i(f,\mu) + \cdots + \lambda_d(f,\mu)$$

Proposition 12.1. *Let μ be any Borel measure on M. For every $1 \le i \le d$ the functions*

$$g \mapsto \Lambda_i(g,\mu) \quad \textit{and} \quad g \mapsto \mathrm{V}_i(g,\mu)$$

are, respectively, upper semi-continuous and lower semi-continuous on $\mathrm{Diff}^1_\mu(M)$.

Upper semi-continuity is proved as follows (see [189]). For $i = 1$ one uses the relation

$$\Lambda_1(g,\mu) = \lambda_1(g,\mu) = \inf_{n\ge 1} \frac{1}{n} \int \log \|Dg^n\|\, d\mu\,. \tag{41}$$

The infimum of continuous functions being upper semi-continuous, the conclusion follows. For $i \ge 2$ one considers the action $Dg^{\wedge i} : W^i \to W^i$ induced by Dg on the vector bundle W^i over M whose fiber W^i_x is the space of i-forms on the co-tangent space $(T_xM)^*$. A natural choice of a norm is such that $\|(Dg^{\wedge i}(x))^n\|$ is the supremum of the i-volume of $Df^n(x)(P)$ over i-dimensional parallelepipeds $P \subset T_xM$ of unit i-dimensional volume. The largest Lyapunov exponent of $Dg^{\wedge i}$ is precisely $\Lambda_i(g,\mu)$. Thus we have the relation corresponding to (41) for $Dg^{\wedge i}$:

$$\Lambda_i(g,\mu) = \inf_{n\ge 1} \frac{1}{n} \int \log \|(Dg^{\wedge i})^n\|\, d\mu\,, \tag{42}$$

and semi-continuity follows as before. A dual argument gives the lower semi-continuity statement in the proposition.

The following theorem characterizes continuity of Lyapunov exponents for volume preserving C^1 diffeomorphisms:

Theorem 12.2 (Bochi, Viana [49, 52, 51])**.** *Let μ be normalized Lebesgue measure on a compact manifold M. If $f \in \mathrm{Diff}^1_\mu(M)$ is a continuity point for the map*

$$\mathrm{Diff}^1_\mu(M) \to \mathbb{R}^d, \quad g \mapsto (\lambda_1(g,\mu), \cdots, \lambda_d(g,\mu))$$

then, for almost every point $x \in M$,

(a) *either all Lyapunov exponents $\lambda_i(f,x) = 0$ for $1 \le i \le d$,*
(b) *or the Oseledets splitting of f is dominated on the orbit of x.*

The second alternative means that there are at least two distinct Oseledets subspaces and there exists $m \in \mathbb{N}$ such that

$$\frac{\|Df^n(y)u\|}{\|u\|} \ge 2\, \frac{\|Df^n(y)v\|}{\|v\|} \quad \text{for all} \quad n \ge m \tag{43}$$

and all non-zero $u \in E^{1,i}_y = E^1_y \oplus \cdots \oplus E^i_y$ and $v \in E^{i+1,k}_y = E^{i+1}_y \oplus \cdots \oplus E^k_y$, for any $1 \le i < k$, and every iterate $y = f^l(x)$. In other words, the fact that Df^n will eventually expand $E^{1,i}$ more than $E^{i+1,k}$ can be observed in a finite number m of iterates, uniform over the orbit of x. This implies that *the Oseledets splitting extends continuously to the closure of the orbit* and, in particular, the angles between the Oseledets subspaces are bounded from zero over the orbit. Compare Remark C.2 in Appendix C.

Problem 12.3. Does the theorem remain valid if the conclusion is replaced by

(a') either $\lambda_j(f,x) = 0$ for all $1 \le j \le d$ and μ-almost every point $x \in M$,

(b') or the Oseledets splitting of f extends continuously to a dominated splitting on the whole ambient space M.

This more global version of the conclusion does hold when $d = 2$: Bochi [49] proves that continuity points either have vanishing Lyapunov exponents almost everywhere or else they are uniformly hyperbolic (Anosov). Notice also that the two versions (a)-(b) and (a')-(b') coincide if f is ergodic: the number and dimensions of the Oseledets subspaces are constant almost everywhere, and so either (a) or (b) must hold on a full measure subset.

Remark 12.4. Conversely, if f satisfies either (a') or (b') then it is a continuity point for all Lyapunov exponents. In case (a') this is a direct consequence of the fact that the largest Lyapunov exponent is an upper semi-continuous function of the dynamics, by Proposition 12.1. In case (b') one applies a similar argument, and the dual fact that the smallest exponent is lower semi-continuous, to each subbundle of the dominated splitting: by domination, every nearby diffeomorphism has an invariant splitting to the same number of subbundles, which are uniformly close to the initial ones.

Let us say a few words about the proof of Theorem 12.2. Clearly, $(\lambda_1, \ldots, \lambda_d)$ is continuous at some f if and only $(\Lambda_1, \ldots, \Lambda_d)$ is. Suppose the Oseledets splitting of f is neither trivial nor dominated, over a positive measure set of orbits: for some i and for arbitrarily large m there exist iterates y for which (43) does not hold. The strategy is to take advantage of this fact to, by a small perturbation of the map, cause a vector originally in $E_y^{1,i}$ to move to $E_z^{i+1,k}$, $z = f^m(y)$ thus "blending" different expansion rates.

More precisely, given a perturbation size $\epsilon > 0$ one takes m sufficiently large with respect to ϵ. Then for $n \gg m$ one chooses a convenient $y = f^l(x)$, with $l \approx n/2$, where domination fails. By composing Df with small rotations near the first m iterates of y, one causes the orbit of some $Df^l(x)v \in E_y^{1,i}$ to move to $E_{f^m(y)}^{i+1,k}$. Thus, one constructs a perturbation g preserving the orbit segment $\{x, \ldots, f^n(x)\}$ such that $Dg^n(x)v \in E_{f^n(x)}^{i+1,k}$. As a result, it is possible to show that

$$\|(Dg^{\wedge j}(x))^n\| \lessapprox \exp\left[n\left(\Lambda_{j-1} + \frac{\lambda_j + \lambda_{j+1}}{2}\right)\right] \ll \exp\left(n\Lambda_j\right),$$

where $j = \dim E^{1,i}$ and $\lambda_j = \lambda_j(f,x)$, $\Lambda_j = \Lambda_j(f,x)$.

This local procedure is then repeated for several points $x \in M$. Using (42) and a Kakutani tower argument, one proves that Λ_j drops under such arbitrarily small perturbations, contradicting continuity.

Problem 12.5. (1) Characterize the continuity points of Lyapunov exponents relative to more general Borel measures μ on M.

(2) Is there a non-conservative version of Theorem 12.2, in the space $\mathrm{Diff}^1(M)$ of all C^1 diffeomorphisms ?

Even more important, one would like to have a corresponding understanding for higher topologies (see also Section 12.4):

Problem 12.6. Let μ be Lebesgue measure on M. What are the continuity points of $\mathrm{Diff}^r_\mu(M) \ni g \mapsto (\lambda_1(g,\mu), \cdots, \lambda_d(g,\mu))$ when $r > 1$?

Now we state a version of Theorem 12.2 for symplectic diffeomorphisms. Let (M, ω) be a compact symplectic manifold of dimension $\dim M = 2q$. The space $\mathrm{Symp}^\nu_\omega(M)$ of all C^ν symplectic diffeomorphisms is a subspace of $\mathrm{Diff}^\nu_\mu(M)$, where μ is the volume measure associated to the volume form $\omega^q = \omega \wedge \cdots \wedge \omega$. The Lyapunov exponents of symplectic diffeomorphisms satisfy

$$\lambda_i(f,x) + \lambda_{2q-i+1}(f,x) = 0 \quad \text{for all } 1 \le i \le q.$$

In particular, $\lambda_q(x) \geq 0$ and $\Lambda_q(f,\mu)$ is the integral of the sum of all non-negative exponents. Moreover, let E^+_x, E^0_x, and E^-_x be the sums of all Oseledets spaces associated to positive, zero, and negative Lyapunov exponents, respectively. Then $\dim E^+_x = \dim E^-_x$ and $\dim E^0_x$ is even.

Theorem 12.7 (Bochi, Viana [51])**.** *Suppose $f \in \mathrm{Symp}^1_\omega(M)$ is a continuity point for the function*

$$\mathrm{Symp}^1_\omega(M) \ni g \mapsto \Lambda_q(g,\mu).$$

Then for μ-almost every $x \in M$ either $\dim E^0_x \geq 2$ *or the splitting $T_xM = E^+_x \oplus E^-_x$ is uniformly hyperbolic along the orbit of x.*

In the second alternative, one just needs to prove that the splitting is dominated, because for symplectic diffeomorphisms dominated splittings into two subspaces of the same dimension are uniformly hyperbolic. See [50, 51].

12.2. **A dichotomy for conservative systems.** From Theorem 12.2 one gets the following dichotomy for generic volume preserving diffeomorphisms:

Theorem 12.8 (Bochi, Viana [49, 52, 51])**.** *Let μ be normalized Lebesgue measure on a compact manifold M. There exists a residual subset $\mathcal{R}$ of* $\mathrm{Diff}^1_\mu(M)$ *such that, for every $f \in \mathcal{R}$ and μ-almost every point x,*

(a) *either all Lyapunov exponents $\lambda_i(f,x) = 0$ for $1 \le i \le d$,*
(b) *or the Oseledets splitting of f is dominated on the orbit of x.*

This result was first proved for $d = 2$ by Bochi [49], partially based on an outline of proof by Mañé [212]. The statement had been announced by Mañé in the early eighties [207]: *for a residual subset of C^1 area preserving diffeomorphisms on any surface, either the Lyapunov exponents vanish almost everywhere or the diffeomorphism is Anosov* and the surface is the torus.

Theorem 12.8 is a direct consequence of Proposition 12.1 and Theorem 12.2. The proposition gives that every $g \mapsto \Lambda_i(g,\mu)$ is upper semi-continuous. Since the set of continuity points of a semi-continuous function on a Baire space always contains a residual subset, we find that the elements of a residual subset of $\mathrm{Diff}^1_\mu(M)$ are continuity points for $g \mapsto (\Lambda_1, \ldots, \Lambda_d)$ and, hence, also for $g \mapsto (\lambda_1, \ldots, \lambda_d)$. By Theorem 12.2 they satisfy the conclusion of Theorem 12.8.

A similar argument, starting from Theorem 12.7, gives the following generic dichotomy for symplectic diffeomorphisms, which implies that generic symplectic diffeomorphisms on most manifolds (not supporting Anosov diffeomorphisms) are *not* hyperbolic!

Theorem 12.9 (Bochi, Viana [51]). *There exists a residual set $\mathcal{R} \subset \mathrm{Symp}^1_\omega(M)$ such that for every $f \in \mathcal{R}$ either the diffeomorphism f is Anosov or Lebesgue almost every point has zero as Lyapunov exponent, with multiplicity ≥ 2.*

A very useful ingredient in deducing Theorem 12.9 is the fact that generic C^1 diffeomorphisms (including all C^2 diffeomorphisms) either are Anosov or all their hyperbolic sets have zero Lebesgue measure. See [50] for a proof. This is true also in the volume-preserving category. On the other hand, no analogue is known for invariant sets with a dominated splitting. In fact, it is easy to find examples of C^∞ diffeomorphisms for which both alternatives (a) and (b) in Theorem 12.8 occur simultaneously:

Example 12.10. Let $\{f_t : t \in S^1\}$ be a family of volume preserving diffeomorphisms of $\mathbb{T}^3$ such that $f_t = \mathrm{id}$ for t in some interval and f_t is partially hyperbolic for t in another interval. One way such a family can be obtained is using the construction of partially hyperbolic maps isotopic to the identity given in [56]. Then $F : \mathbb{T}^4 \to \mathbb{T}^4$, $F(t,x) = (t, f_t(x))$ satisfies both (a) and (b) in Theorem 12.8 on positive measure sets.

It is unclear whether such examples can be made generic. In this direction we ask the following strong form of Problem 12.3:

Problem 12.11. Is there a residual subset of C^r volume preserving or symplectic diffeomorphisms, $r \geq 1$, for which the compact invariant sets with a dominated splitting either have zero volume or coincide with the ambient space ? Same in the space of all diffeomorphisms.

Remark 12.12. For r sufficiently large, KAM theory yields C^r open sets of symplectic maps which are not hyperbolic, due to the presence of invariant Lagrangian tori restricted to which the map is conjugate to rotations and whose union has positive volume. Moreover, Herman has constructed C^∞ open sets of volume preserving diffeomorphisms having invariant sets with positive volume, which are unions of codimension-1 invariant tori. See [347, § 4.6]. In all these examples where hyperbolicity fails, *all* Lyapunov exponents vanish. Question 1 in [313] asks whether for generic volume preserving diffeomorphisms either all Lyapunov exponents are zero or none is zero, almost everywhere.

The next almost-example describes a possible surprising consequence of Theorem 12.8: Lyapunov exponents with multiplicity persistently larger than 1. The argument would be complete if we knew the answer to

Problem 12.13. Are C^2 maps C^1 dense among volume preserving C^1 diffeomorphisms ?

Zehnder [354] gives a complete solution in the symplectic category: C^∞ maps are C^k dense among C^k symplectic diffeomorphisms, for any $k \geq 1$. He also proves that C^∞ maps are C^k dense among volume preserving diffeomorphisms of class $C^{k,\gamma}$ (the derivative of order k is γ-Hölder), for any $k \geq 1$ and $0 < \gamma \leq 1$.

Example 12.14 (Bonatti,Viana [67], Tahzibi [323]). For $M = \mathbb{T}^4$ and $\mu =$ Lebesgue measure on M, there exists an open subset $\mathcal{U}$ of $\mathrm{Diff}^1_\mu(M)$ such that, for every $f \in \mathcal{U}$,

(1) f admits a dominated splitting $TM = E \oplus F$ with $\dim E = \dim F = 2$;
(2) there are no other continuous invariant subbundles;

(3) f is transitive and if f is C^2 then it is even ergodic.

(4) f is not hyperbolic: $Df \mid E$ is not expanding and $Df \mid F$ is not contracting.

Assume C^2 diffeomorphisms are dense in $\mathcal{U}$ (this could be true even if the global density statement fails). The set of ergodic maps is always a G_δ (countable intersection of open subsets) in $\mathrm{Diff}^r_\mu(M)$, $r \geq 0$, by [254, § 8]. Thus property (3) implies that the set of ergodic diffeomorphisms contains a residual (*dense* G_δ) subset $\mathcal{E}$ of $\mathcal{U}$. Take $\mathcal{S} = \mathcal{R} \cap \mathcal{E}$, where $\mathcal{R}$ is the residual set in Theorem 12.8. Because of (1), the Lyapunov exponents of f can not be all zero. So, using Theorem 12.8 and ergodicity, the Oseledets splitting must extend to a dominated splitting on the whole M. By (2), this implies that the Oseledets splitting has the form $E^1_x \oplus E^2_x$ with $\dim E^i_x = 2$. In other words, there exists a residual subset $\mathcal{S}$ of $\mathcal{U}$ such that every $f \in \mathcal{S}$ is ergodic, and its Lyapunov exponents satisfy

$$\lambda_1(f,\mu) = \lambda_2(f,\mu) > \lambda_3(f,\mu) = \lambda_4(f,\mu).$$

By Remark 12.4, every $f \in \mathcal{S}$ is a continuity point for all Lyapunov exponents.

A classical theorem of Oxtoby, Ulam [254] states that ergodic maps are generic among all volume preserving *homeomorphisms*. The following question, posed to us by A. Katok, asks for a C^1 version of this result:

Problem 12.15 (Katok)**.** Do ergodic diffeomorphisms contain a residual subset of all area preserving C^1 diffeomorphisms on any surface ? Same question in higher dimensions.

12.3. **Deterministic products of matrices.** Lyapunov exponents have an important role in several other areas, such as the theory of random matrices and the spectral theory of Schrödinger operators. Such applications are usually formulated in the more general setting of linear cocycles over a dynamical system. The questions we have been discussing extend in a natural way to this setting.

Let $f : M \to M$ be a continuous transformation on a compact metric space M. A *linear cocycle* over f is a vector bundle automorphism $F : \mathcal{E} \to \mathcal{E}$ covering f, where $\pi : \mathcal{E} \to M$ is a finite-dimensional (real or complex) vector bundle over M. This means that

$$\pi \circ F = f \circ \pi$$

and F acts as a linear isomorphism on every fiber. The prototype is the derivative $F = Df$ of a diffeomorphism on a manifold (*dynamical cocycle*). A linear cocycle over a flow $f^t : M \to M$, $t \in \mathbb{R}$, is a flow $F^t : \mathcal{E} \to \mathcal{E}$, $t \in \mathbb{R}$, such that F^t is a linear cocycle over f^t for every $t \in \mathbb{R}$.

For simplicity, we focus on the case when the vector bundle is trivial $\mathcal{E} = M \times \mathbb{R}^d$ (the complex case $\mathcal{E} = M \times \mathbb{C}^d$ is entirely similar). Then the cocycle has the form $F(x,v) = (f(x), A(x)v)$ with $A : M \to \mathrm{GL}(d,\mathbb{R})$. The iterates are given by

$$F^n(x,v) = (f^n(x), A^n(x)v) \quad \text{with } A^n(x) = A(f^{n-1}(x)) \cdots A(f(x))\, A(x).$$

When f is invertible, $A^{-n}(x)$ is the inverse of $A^n(f^{-n}(x))$. In the continuous time case $F^t(x,v) = (f^t(x), A^t(x)v)$ for $t \in \mathbb{R}$.

Given any measurable $A : M \to \mathrm{GL}(d,\mathbb{R})$ such that $\log^+ \|A^{\pm 1}\|$ are integrable relative to some invariant probability μ, the largest Lyapunov exponent

$$\lambda(A,x) = \lim_{n\to\infty} \frac{1}{n} \log \|A^n(x)\|$$

exists μ-almost everywhere. In fact, the theorem of Oseledets holds in this setting of linear cocycles, see Appendix C. The symbols $\hat{\lambda}_j(A,x)$, $\lambda_i(A,x)$, $\lambda_i(A,\mu)$, $\Lambda_i(A,\mu)$, $\mathrm{V}_i(A,\mu)$ will have the same meaning as in the dynamical case. Except where stated otherwise, we assume cocycles to be continuous. Then the integrability condition is automatic, by compactness.

Example 12.16. (Products of iid random matrices) Let G be a matrix group, $M = G^{\mathbb{N}}$ be the space of sequences $(\alpha_j)_{j\geq 0}$ in G, and $f : M \to M$ be the shift map. Let $A : M \to G$ be the projection onto the zero:th coordinate: $A((\alpha_j)_{j\geq 0}) = \alpha_0$. Then

$$A^n((\alpha_j)_{j\geq 0}) = \alpha_{n-1}\cdots\alpha_1\,\alpha_0\,. \tag{44}$$

Let ν be a probability in G and $\mu = \nu^{\mathbb{N}}$. The cocycle defined by A over (f,μ) models the product of random matrices chosen in G independently according to the distribution ν.

A classical theory initiated by Furstenberg and Kesten [133, 135] states, in its simplest form, that *for the majority of choices of the probability* ν *in* $G = \mathrm{SL}(d,\mathbb{R})$ *the product* (44) *grows exponentially fast.* More precisely, as long as there is no probability on the projective space $\mathbb{RP}^{d-1}$ invariant under the action of *all* the matrices in $\operatorname{supp}\nu$, the largest Lyapunov exponent $\lambda(A,\mu)$ of the corresponding cocycle is positive. Also in great generality, one even has that the Lyapunov spectrum is simple: all Oseledets subspaces have dimension 1. See Guivarch, Raugi [144] and Gold'sheid, Margulis [138].

Quite in contrast, the following version the dichotomy theorem shows that vanishing Lyapunov exponents are quite common for linear cocycles. We give the statement in the ergodic invertible case. Recall that a group action is *transitive* if for any pair of points on the manifold there is a group element sending one point to the other.

Theorem 12.17 (Bochi, Viana [49, 52, 51])**.** *Assume* $f : (M,\mu) \to (M,\mu)$ *is ergodic and invertible. Let* $G \subset \mathrm{SL}(d,\mathbb{R})$ *be any closed subgroup acting transitively on the projective space* $\mathbb{RP}^{d-1}$*. Then there exists a residual subset* $\mathcal{R}$ *of continuous maps* $A : M \to G$ *for which either the Lyapunov exponents* $\lambda_i(A,\mu)$ *are all zero or the Oseledets splitting of* A *extends to a dominated splitting over the support of* μ*.*

Proposition 12.1 extends to linear cocycles: the sums $\Lambda_i(A,\mu)$ and $\mathrm{V}_a(A,\mu)$ of the largest and the smallest Lyapunov exponents are, respectively, upper semi-continuous and lower semi-continuous functions of A in the space $C^0(M,G)$ of continuous maps from M to G endowed with the C^0 topology. Then Theorem 12.17 follows from a corresponding version of Theorem 12.2: if A is a continuity point for the map

$$C^0(M,G) \ni B \mapsto (\lambda_1(B,\mu),\ldots,\lambda_d(B,\mu)),$$

then either the Lyapunov exponents are all zero or the Oseledets splitting extends to a dominated splitting over the whole support. The converse of this latter result is also true, and follows from the arguments in Remark 12.4.

An important point to be noted in Theorem 12.17 is that existence of a dominated splitting is a strong property, which can often be excluded *a priori* for topological or geometric reasons. The next couple of examples illustrate this point. In such cases we are left with the first alternative in the dichotomy: vanishing Lyapunov exponents for generic cocycles.

Example 12.18. Let $f : M \to M$ and $A : M \to \mathrm{SL}(d,\mathbb{R})$ be such that for every $1 \le i < d$ there exists a periodic point p_i in the support of μ, with period q_i, such that the eigenvalues $\{\beta_j^i : 1 \le j \le d\}$ of $A^{q_i}(p_i)$ satisfy

$$|\beta_1^i| \ge \cdots \ge |\beta_{i-1}^i| > |\beta_i^i| = |\beta_i^{i+1}| > |\beta_{i+2}^i| \ge \cdots \ge |\beta_d^i| \tag{45}$$

and β_i^i, β_{i+1}^i are complex conjugate (not real). Such an A may be found, for instance, starting with a constant cocycle and deforming it on disjoint neighborhoods of the periodic orbits. Property (45) remains valid for every B in a C^0 neighborhood $\mathcal{U}$ of A. It implies that no B admits an invariant dominated splitting over the support of μ: if such a splitting $E \oplus F$ existed then, at every periodic point, the $\dim E$ largest eigenvalues would be strictly larger than the other eigenvalues, which is incompatible with (45). It follows, by Theorem 12.17, that every cocycle in a residual subset $\mathcal{U} \cap \mathcal{R}$ of the neighborhood has all the Lyapunov exponents equal to zero.

Example 12.19. Let $f : S^1 \to S^1$ be a homeomorphism and μ be any invariant ergodic measure with $\operatorname{supp}\mu = S^1$. Let $\mathcal{N}$ be the set of all continuous $A : S^1 \to \mathrm{SL}(2,\mathbb{R})$ non-homotopic to a constant. For a residual subset of $\mathcal{N}$, the Lyapunov exponents of the corresponding cocycle over (f,μ) are zero. That is a direct consequence of Theorem 12.17, together with the fact that the cocycle has no invariant continuous subbundle. The latter can be shown as follows.

A continuous subbundle defines a continuous map from S^1 to $\mathbb{RP}^1$ and, hence, an element of $\pi_1(\mathbb{RP}^1) = \mathbb{Z}$. The space $\mathrm{SL}(2,\mathbb{R})$ has the homotopy type of the circle, and the map A defines an element $\deg A$ of its fundamental group $\mathbb{Z}$. The cocycle associated to A induces an action in the space of continuous sub-bundles. This action is defined at the level of the fundamental group $\pi_1(\mathbb{RP}^1)$, and is given by a translation: $i \mapsto i + \deg A$. An invariant subbundle yields a fixed point for this action, and so it can only exist if $\deg A = 0$ or, in other words, if A is homotopic to a constant.

Theorem 12.17 also carries over to the space $L^\infty(M, \mathrm{SL}(d,\mathbb{R}))$ of measurable bounded cocycles, still with the uniform topology. Let us point out that in weaker topologies cocycles having a dominated splitting may cease to constitute an open set. In fact,

Theorem 12.20 (Arnold, Cong [22], Arbieto, Bochi [15])**.** *For every $1 \le p < \infty$, a residual subset of all L^p linear cocycles over any ergodic transformation have one-point Lyapunov spectra, that is, all the Lyapunov exponents coincide.*

First, Arnold, Cong [22] proved the conclusion on a dense subset. Then Arbieto, Bochi [15] proved the rather surprising fact that semi-continuity of Lyapunov exponents (Proposition 12.1) is true in any L^p topology, and they deduce the full statement of the theorem.

12.4. **Abundance of non-zero exponents.** We are going to see that the conclusions of the previous section change radically if one considers cocycles which are better than just continuous: assuming the base dynamics is hyperbolic (uniformly or not), almost all Hölder continuous or even differentiable cocycles have non-zero Lyapunov exponents.

Let G be any closed subgroup of $\mathrm{SL}(d,\mathbb{R})$. For $0 < \nu \le \infty$ denote by $C^\nu(M,G)$ the space of C^ν maps from M to G endowed with the C^ν topology. For $\nu \ge 1$ it is implicitly assumed that M has a smooth structure. When ν is not an integer, C^ν means the map is $[\nu]$ differentiable and the derivative of order $[\nu]$ is $(\nu-[\nu])$-Hölder continuous. In the integer case, C^ν means either that f is ν times differentiable with continuous derivative of order ν, or that it is $\nu-1$ times differentiable with Lipschitz continuous derivative of order $\nu-1$: all statements are meant for both interpretations.

Definition 12.21. A matrix Lie group G *acts independently* on the projective space if the map

$$G \to (\mathbb{RP}^{d-1})^d, \quad B \mapsto (B\xi_1, \dots, B\xi_d)$$

is a submersion, for any linearly independent set $\{\xi_1, \dots, \xi_d\} \subset \mathbb{RP}^{d-1}$.

This holds, for instance, for the special linear group $\mathrm{SL}(d,\mathbb{R})$ and for the symplectic group.

Let μ be a hyperbolic measure for a diffeomorphism $f : M \to M$. We assume the derivative Df to be Hölder continuous, so that Pesin theory applies (Appendix C). The notion of measure with local product structure is recalled in Section C.5. Most interesting hyperbolic measures have local product structure, e.g. Lebesgue measure, measures absolutely continuous along stable or along unstable manifolds, and the equilibrium states of Axiom diffeomorphisms relative to Hölder potentials [70].

Theorem 12.22 (Viana [332])**.** *Assume (f,μ) is non-uniformly hyperbolic with local product structure, and let $G \subset \mathrm{SL}(d,\mathbb{R})$ be a Lie subgroup acting independently on the projective space. Assume also that (f,μ) is ergodic. Then, for every $\nu > 0$ the set of cocycles A with largest Lyapunov exponent $\lambda(A,x) > 0$ at μ-almost every point contains an open dense subset $\mathcal{A}$ of $C^\nu(M,G)$. Moreover, its complement has ∞-codimension.*

The last property means that the set of cocycles with vanishing exponents is locally contained inside finite unions of closed submanifolds of $C^\nu(M,G)$ with arbitrary codimension. Thus, generic parametrized families of cocycles do not intersect this exceptional set at all!

Remark 12.23. We may drop the ergodicity assumption in Theorem 12.24, at the price of replacing open and dense by residual in the statement and allowing for countable unions in the definition of ∞-codimension. This is because any measure with local product structure has countably many ergodic components, which also have local product structure.

One gets a stronger conclusion when $f : M \to M$ is uniformly hyperbolic, for instance, a two-sided shift of finite type or an Axiom A diffeomorphism restricted to some hyperbolic basic set. In this situation every invariant measure is hyperbolic. The main improvement is that the set $\mathcal{A}$ may be taken the same for all invariant measures with local product structure, ergodic or not.

Theorem 12.24 (Bonatti, Gomez-Mont, Viana [63, 332])**.** *Assume $f : M \to M$ is uniformly hyperbolic, and let $G \subset \mathrm{SL}(d,\mathbb{R})$ be a Lie subgroup acting independently on the projective space. Then, for every $\nu > 0$, the set of cocycles A having largest Lyapunov exponent $\lambda(A,x) > 0$ at μ-almost every point and for every invariant measure with local product structure contains an open dense subset $\mathcal{A}$ of $C^\nu(M,G)$. Moreover, its complement has ∞-codimension.*

Theorem 12.24 was first proved in [63] under an additional hypothesis called domination: the cocycle induced on the projective bundle $M \times \mathbb{RP}^{d-1}$ is partially hyperbolic, with the projective fibers as central leaves. Under this extra hypothesis [66] gets a stronger conclusion: for the majority of cocycles all Lyapunov exponents have multiplicity 1, in other words, the Oseledets subspaces E^i are one-dimensional. We expect this to be true in full generality:

Conjecture 12.25. Theorems 12.22 and 12.24 should remain true if one replaces $\lambda(A, x) > 0$ by all Lyapunov exponents $\lambda_i(A, x)$ having multiplicity 1.

Theorems 12.22 and 12.24 extend to cocycles over non-invertible transformations, respectively, local diffeomorphisms equipped with invariant non-uniformly expanding probabilities (all Lyapunov exponents positive), and uniformly expanding continuous maps, like one-sided shifts of finite type, or smooth expanding maps.

Problem 12.26. What are the continuity points of Lyapunov exponents as functions of the cocycle in $C^\nu(M, \mathrm{SL}(d, \mathbb{R}))$ when $\nu > 0$? Analogously, assuming the base system (f, μ) is hyperbolic.

Lyapunov exponents may vary discontinuously in the C^ν topology if ν is small. The known examples are not dominated in the sense of [63, 66].

In the remainder of this section we outline some main ingredients in the proofs of Theorems 12.22 and 12.24 that are of independent interest. For clearness, we focus on a particular situation: we suppose f is uniformly expanding, $G = \mathrm{SL}(2, \mathbb{R})$, and μ is ergodic with $\operatorname{supp} \mu = M$. It is no restriction to suppose $\nu \geq 1$, because the cases $0 < \nu < 1$ can be transformed to $\nu = 1$ by replacing the distance $d(x, y)$ in M by $d(x, y)^\nu$.

12.4.1. *Bundle-free cocycles.* We call a cocycle bundle-free if it admits no Lipschitz continuous invariant sub-bundle, not even up to finite covering. More precisely:

Definition 12.27. We say that $A : M \to \mathrm{SL}(2, \mathbb{R})$ is *bundle-free* if, given any $\eta \geq 1$, there exists *no* Lipschitz continuous map $\psi : x \mapsto \{v_1(x), \ldots, v_\eta(x)\}$ assigning to each $x \in M$ a subset of $\mathbb{RP}^1$ with exactly η elements, invariant under the cocycle in the sense that

$$A(x)\big(\{v_1(x), \ldots, v_\eta(x)\}\big) = \{v_1(f(x)), \ldots, v_\eta(f(x))\} \quad \text{for all } x \in M.$$

Notice that the definition is independent of the invariant measure μ. We also point out that regularity is crucial here: existence of Lipschitz invariant subbundles is rare (most cocycles are bundle-free) whereas existence of Hölder invariant subbundles with poor Hölder constants is often robust. The following construction illustrates the issues involved, in a simplified situation:

Example 12.28. Let $F : S^1 \times \mathbb{R} \to S^1 \times \mathbb{R}$, $F(\theta, x) = (f(\theta), g(\theta, x))$ be a smooth map with

$$\sigma_1 \geq |f'| \geq \sigma_2 > \sigma_3 > |\partial_x g| > \sigma_4 > 1.$$

Let θ_0 be a fixed point of f and x_0 be the fixed point of $g(\theta_0, \cdot)$. Then

(1) The set of points whose forward orbit is bounded is the graph of a continuous function $u : S^1 \to \mathbb{R}$ with $u(\theta_0) = x_0$. This function is ν-Hölder for any $\nu < \log \sigma_4 / \log \sigma_1$. Typically it is not Lipschitz:

(2) The fixed point $p_0 = (\theta_0, x_0)$ has a strong-unstable set $W^{uu}(p_0)$ invariant under F and which is locally a Lipschitz graph over S^1. If u is Lipschitz then its graph must coincide with $W^{uu}(p_0)$.
(3) However, for an open dense subset of choices of g the strong-unstable set is not globally a graph: it intersects vertical lines at infinitely many points.

Theorem 12.29. *Suppose $A \in C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$ has $\lambda(A,\mu) = 0$. Then A is approximated in $C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$ by stably bundle-free maps.*

Here is a sketch of the proof, which is inspired in the previous example. Using the hypothesis one finds periodic points p of f, such that the matrix $A^q(p)$, $q = \mathrm{per}(p)$ is hyperbolic and the normalized eigenvalues are close to 1. Then the eigenspaces, seen as periodic points of the cocycle acting in the projective space, have strong-unstable sets that are *locally* Lipschitz graphs over M. The graph of any Lipschitz map ψ as in Definition 12.27 has to be contained in the strong-unstable sets. But a simple transversality argument shows that *globally* the strong-unstable sets are not graphs, not even up to finite covering, if certain configurations with positive codimension are avoided.

12.4.2. *A geometric criterium for non-zero exponents.* Another main step is

Theorem 12.30. *Suppose $A \in C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$ is bundle-free and there exists some periodic point $p \in M$ of f such that A is hyperbolic over the orbit of p. Then $\lambda(A,\mu) > 0$ for every invariant probability μ with local product structure.*

The second condition, on the existence of some periodic point over which the cocycle is hyperbolic, is satisfied by an open and dense subset of $C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$, that we denote HP. See the last section of [66] for a proof, and note that the same statement for $\mathrm{SL}(2,\mathbb{C})$ is easy. We also denote by BF the subset of bundle-free maps. The next example helps understand the significance of Theorem 12.30.

Example 12.31. Let $M = S^1$, $f : M \to M$ be given by $f(x) = kx \bmod \mathbb{Z}$, for some $k \geq 2$, and μ be Lebesgue measure on M. Let

$$A : M \to \mathrm{SL}(2,\mathbb{R}), \qquad A(x) = \begin{pmatrix} \beta(x) & 0 \\ 0 & 1/\beta(x) \end{pmatrix}$$

for some smooth function β such that $\int \log \beta \, d\mu = 0$. It is easy to ensure that the set $\beta^{-1}(1)$ is finite and does not contain $x = 0$. Then $A \in$ HP and indeed the matrix A "looks hyperbolic" at most points. Nevertheless, the Lyapunov exponent $\lambda(A,\mu) = \int \log \beta \, d\mu = 0$. Notice that A is not bundle-free.

This suggests the following heuristic principle: assuming there is some source of hyperbolicity (here the fact that $A \in$ HP), the only way Lyapunov exponents may happen to vanish is by having expanding directions mapped *exactly* onto contracting directions, thus causing hyperbolic behavior to be "wasted way".

12.4.3. *Conclusion and an application.* It is now easy to finish the argument using Theorems 12.29 and 12.30. Let ZE be the subset of $A \in C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$ such that $\lambda(A,\mu) = 0$. Theorem 12.29 says that any $A \in$ ZE is approximated by the interior of BF. Since HP is open and dense, A is also approximated by the interior of BF $\cap$ HP. By Theorem 12.30, the latter is contained in the complement of ZE. This proves that the interior of $C^\nu \setminus$ ZE is dense in ZE, and so it is dense in the whole $C^\nu(M, \mathrm{SL}(2,\mathbb{R}))$, as claimed. To get the ∞-codimension statement observe that it suffices to avoid the positive codimension configurations mentioned at the end of Section 12.4.1, for *some* of infinitely many periodic points of f. This completes our outline of the proof of Theorems 12.22 and 12.24, in a particular case.

We conclude with an application that puts Theorems 12.17 and 12.30 together to give a very complete description of Lyapunov exponents for a whole open set of continuous cocycles. It also shows that these results are, in some sense, optimal.

Example 12.32. Let $M = S^1$, $f : M \to M$ be a uniformly expanding map, and μ be any ergodic measure invariant supported on the whole M. Given any continuous map $\alpha : S^1 \to S^1$, define

$$A : M \to \mathrm{SL}(2, \mathbb{R}), \qquad A(x) = R_{\alpha(x)} \cdot A_0$$

where R_α denotes the rotation of angle α and A_0 is some constant matrix.

Corollary 12.33. *Assume* $2\deg(\alpha)$ *is not a multiple of* $\deg(f) - 1$ *and the matrix* $A(p)$ *is hyperbolic at some fixed point* p. *Then there exists a* C^0 *neighborhood* $\mathcal{U}$ *of* A *such that*

(1) *for* B *in a residual subset* $\mathcal{R} \cap \mathcal{U}$ *we have* $\lambda(B, \mu) = 0$;
(2) *for every* $B \in C^\nu(M, \mathrm{SL}(2, \mathbb{R}))$, $\nu > 0$ *we have* $\lambda(B, \mu) > 0$.

Proof. Start by taking $\mathcal{U}$ to be the isotopy class of A in the space of continuous maps from M to SL. We claim that, given any $B \in \mathcal{U}$, there is no *continuous* B-invariant map

$$\psi : M \ni x \mapsto \{\psi_1(x), \ldots, \psi_\eta(x)\}$$

assigning a constant number $\eta \geq 1$ of elements of $\mathbb{RP}^1$ to each point $x \in M$. The proof is by contradiction. Suppose there exists such a map and the graph

$$G = \{(x, \psi_i(x)) \in S^1 \times \mathbb{RP}^1 : x \in S^1 \text{ and } 1 \leq i \leq \eta\}$$

is connected. Then G represents some element (η, ζ) of the fundamental group $\pi_1(S^1 \times \mathbb{RP}^1) = \mathbb{Z} \oplus \mathbb{Z}$. Because B is isotopic to A, the image of G under the cocycle must represent $(\eta \deg(f), \zeta + 2\deg(\alpha)) \in \pi_1(S^1 \times \mathbb{RP}^1)$; here the factor 2 comes from the fact that S^1 is the 2-fold covering of $\mathbb{RP}^1$. By the invariance of ψ we get

$$\zeta + 2\deg(\alpha) = \deg(f)\zeta$$

which contradicts the hypothesis that $\deg(f) - 1$ does not divide $2\deg(\alpha)$. If the graph G is not connected, consider the connected components instead. Since connected components are pairwise disjoint, they all represent elements with the same direction in the fundamental group. Then the same type of argument as before proves the claim in full generality.

Now let $\mathcal{R}$ be the residual subset in Theorem 12.17. The previous observation implies that no $B \in \mathcal{R} \cap \mathcal{U}$ may have an invariant dominated splitting. Then B must have all Lyapunov exponents equal to zero as claimed in (1). Similarly, that observation ensures that every $B \in \mathcal{U} \cap C^\nu$ is bundle-free. Our assumptions ensure that A is in HP, and so is any map C^0 close to it. Thus, reducing $\mathcal{U}$ if necessary, we may apply Theorem 12.30 to conclude that $\lambda(B, \mu) > 0$. This proves (2). $\square$

12.5. **Lower bounds for Lyapunov exponents. ?? ??**

12.6. **Non-uniformly hyperbolic systems.** We refer the reader to the supplementary chapter by Katok, Mendoza in [171], and to the book of Barreira, Pesin [38] for broad presentations of non-uniform hyperbolicity theory. Here we just discuss a few important recent developments.

12.6.1. *Exact dimensional measures.* Given a Borel measure μ on a complete metric space M, the *pointwise dimension* of μ at $x \in M$ is defined by

$$d(\mu, x) = \lim_{r \to 0} \frac{\log \mu(B_r(x))}{\log r}. \tag{46}$$

In other words, if the pointwise dimension does exist then small balls around x have

$$\mu(B_r(x)) \approx r^{d(\mu,x)}.$$

In general, one speaks of upper and lower dimensions, $\overline{d}(\mu, x)$ and $\underline{d}(\mu, x)$, defined by the lim sup and the lim inf in (46). If μ is invariant under a diffeomorphism $f : M \to M$ then $\overline{d}$ and $\underline{d}$ are constant on trajectories; hence, they are constant almost everywhere if μ is ergodic.

Eckmann, Ruelle [126] highlighted the importance of this concept for invariant measures of smooth systems, and thus raised the question whether the pointwise dimension always exists at almost every point if μ is a hyperbolic invariant measure. Young [350] proved that if $d(\mu, x) = d$ at μ-almost every point then several natural definitions of dimension of the measure μ coincide with d. Then μ is said to be *exact dimensional.* She also proved that ergodic hyperbolic measures of a *surface* $C^{1+\alpha}$ diffeomorphism f are exact dimensional, in fact,

$$d(\mu, x) = h_\mu(f)\big(\lambda_1^{-1} - \lambda_2^{-1}\big)$$

where $h_\mu(f)$ is the entropy of (f, μ) and $\lambda_1 > 0 > \lambda_2$ are the Lyapunov exponents. Ledrappier, Young [191, 192] introduced stable and unstable pointwise dimensions of an invariant probability, denoted $d^s(\mu, x)$ and $d^u(\mu, x)$. Avoiding technicalities, let us just say that these are the pointwise dimensions of μ conditioned (see Appendix C) to stable manifolds and to unstable manifolds, respectively. They proved that stable and unstable pointwise dimensions always exist almost everywhere, and

$$\overline{d}(\mu, x) \le d^s(\mu, x) + d^u(\mu, x),$$

if μ is hyperbolic. Finally, the matter was settled by the following remarkable

Theorem 12.34 (Barreira, Pesin, Schmeling [39])**.** *Let μ be a hyperbolic invariant measure of $C^{1+\alpha}$ diffeomorphism $f : M \to M$ on any compact manifold M. Then $d(\mu, x)$ exists and is equal to $d^s(\mu, x) + d^u(\mu, x)$ at almost every point. In particular, if μ is ergodic then it is exact dimensional and $d(\mu) = d^s(\mu) + d^u(\mu)$.*

The main ingredient in the proof is a property of *local quasi product structure* for arbitrary invariant probability measures: for any δ there exists a subset Λ with $\mu(M \setminus \Lambda) < \delta$ and a constant $\kappa > 1$ such that

$$r^\delta \mu_x^s(B_r(\kappa^{-1}x))\mu_x^u(B_r(\kappa^{-1}x)) \le \mu(B_r(x)) \le r^{-\delta}\mu_x^s(B_r(\kappa x))\mu_x^u(B_r(\kappa x))$$

for every $x \in \Lambda$ and $r < r(x)$, where μ_x^s and μ_x^u are conditional measures of μ along the stable manifold and the unstable manifold of x. Compare Definition C.6.

12.6.2. *Removing zero Lyapunov exponents.* Since existence of an Anosov diffeomorphism imposes very strong restrictions on the manifold, it is natural to ask whether all manifolds support non-uniformly hyperbolic diffeomorphisms. This problem has a fairly long history. Katok [169] constructed non-uniformly hyperbolic Bernoulli diffeomorphisms on any surface. Then Brin, Feldman, Katok [77] showed that every compact surface supports Bernoulli diffeomorphisms. Their examples had 2 Lyapunov exponents equal to zero, but Brin [75] could bring that number down to 1. Very recently, Dolgopyat, Pesin used a perturbation argument going back to [313] to get rid of this last vanishing exponent, thus settling the matter completely:

Theorem 12.35 (Dolgopyat, Pesin [121]). *Every compact manifold supports volume preserving C^∞ diffeomorphisms which are Bernoulli and non-uniformly hyperbolic relative to Lebesgue measure.*

Even more recently, the perturbation argument of [313] was given a local form by Baraviera, Bonatti [36], and this allowed them to prove

Theorem 12.36 (Baraviera, Bonatti [36]). *Every volume preserving partially hyperbolic diffeomorphism $f : M \to M$ on a compact manifold may be approximated, in the C^1 topology, by another for which the integrated sum of all central Lyapunov exponents is different from zero.*

The integrated sum of all central Lyapunov exponents is given by

$$\lambda^c(f) = \int_M \log |\det(Df \mid E^c)| \, dm$$

where E^c is the central sub-bundle of f. Observe that this number varies continuously with f, within the set of partially hyperbolic C^1 diffeomorphisms, because the central sub-bundle does.

Related to all this, we recall

Conjecture 12.37 (Viana [337]). Let a diffeomorphism $f : M \to M$ be non-uniformly hyperbolic at Lebesgue almost every point: there exists a measurable splitting of the tangent space $T_xM = E^u_x \oplus E^s_x$, defined Lebesgue almost everywhere, and such that

$$\liminf_{n\to\pm\infty} \frac{1}{n} \log \|Df^n(x)v_u\| > 0 > \limsup_{n\to\pm\infty} \frac{1}{n} \log \|Df^n(x)v_s\|$$

for every $v_u \in E^u_x$ and $v_s \in E^s_x$. Then f admits some SRB measure.

The results of [7, 67] contain special cases of this conjecture, recall Theorem 11.34. Another special case is Keller's theorem [175] (a simpler proof is given in [103]): *for unimodal maps with negative Schwarzian derivative, there exists an absolutely continuous invariant measure if and only if there exists a positive Lebesgue measure set of points with positive Lyapunov exponent.* If such a measure exists, it is ergodic and, hence, an SRB measure.

Appendix A. Perturbation lemmas

Most genericity results mentioned in the text are based on certain general perturbation statements that we review in this appendix.

A.1. **Closing lemmas.** A typical question is: *Given a recurrent point of a system, is there a C^r nearby system for which that point is periodic ?* The difficulty of the question depends, dramatically, on the topology and, to some extent, on the dimension of the ambient space and the class of systems one considers.

For $r = 0$ the answer is an easy Yes. For $r = 1$ it is provided by Pugh's closing lemma:

Theorem A.1 (Pugh [278, 279]). *Given any C^1 flow $\mathbf{f} = \{f^t : t \in \mathbb{R}\}$ on a compact manifold M, and a point $x \in M$ such that $f^{t_j}(x) \to x$ for some sequence $t_j \to \infty$, there exists a flow $\mathbf{g} = \{g^t : t \in \mathbb{R}\}$ arbitrarily close to $\mathbf{f}$ in the C^1 topology, for which x is a periodic point.*

Partial results had been obtained by Peixoto [270], for flows on orientable surfaces, and Anosov [11], for the class of systems named after him. Actually, in the special case of hyperbolic systems this problem has a complete answer, known as the shadowing lemma: a closing orbit exists in the system itself, no perturbation is needed.

Theorem A.1 was extended by Pugh, Robinson [280] to several other classes of systems, including C^1 diffeomorphisms, Hamiltonian flows, symplectic diffeomorphisms, and volume preserving diffeomorphisms and flows. See also [16] for recent proofs based on a different approach. Let us mention that the C^1 closing lemma remains open in the setting of geodesic flows (C^2 perturbation of the Riemannian metric).

For $r > 1$ the question of the closing lemma remains mostly open. Gutierrez [145] shows that local perturbations as used by Pugh in the C^1 case do not suffice if $r \geq 2$. Herman [156, 157] constructed counterexamples to the C^∞ closing lemma in the setting of conservative systems (symplectic diffeomorphisms or Hamiltonian vector fields) in sufficiently high dimensions. The result of Peixoto mentioned above holds for any $0 \leq r \leq \infty$. It suggests that a perhaps more realistic goal could be

Problem A.2. Given a C^r system with a recurrent point x, is there an arbitrarily C^r close system for which either x is periodic or else there are no recurrent points in a neighborhood of x ?

On the other hand, the closing lemma has been refined in two important directions, always in the C^1 topology:

A.2. **The ergodic closing lemma.** In many situations one would like to know: *Does the periodic orbit obtained after perturbation remain close, up to the period, to a segment of the original orbit ?* Mañé's ergodic closing lemma asserts that this is true for a *full probability* set of points, that is, having full measure with respect to any invariant probability of the unperturbed map. For the precise statement, let $\Sigma(f)$ be the set of points $x \in M$ such that given any C^1 neighborhood $\mathcal{U}$ of f and any $\varepsilon > 0$ there exist $g \in \mathcal{U}$, coinciding with f outside the ε-neighborhood of the f-orbit of x, and there exists a g-periodic point $y \in M$ such that

$$d(f^i(x), g^i(y)) \leq \varepsilon \quad \text{for all } 0 \leq i \leq \operatorname{per}(y).$$

Theorem A.3 (Mañé [206]). *For any C^1 diffeomorphism f of a compact manifold, the set $\Sigma(f)$ has full probability for f.*

A major application of this statement is to express in terms of periodic orbits dynamical behaviors that are detected by invariant measures. The following argument, which is part of Mañé's proof of the 2-dimensional version of Theorem 7.2 in [206], is a good illustration:

Let f be a surface diffeomorphism with a dominated splitting $E \oplus F$ and suppose the sub-bundle F is *not* uniformly expanding. Then there are arbitrarily long orbit segments over which expansion along F is arbitrarily weak. Any weak accumulation point of Dirac averages along such orbit segments is an invariant measure for which the sub-bundle F is not expanding. Pugh's closing lemma permits to close a typical orbit of such a measure. However, not knowing whether the periodic orbit remains close to the original one, one loses control on the dilation along F. On the contrary, the ergodic closing lemma ensures that the derivative of the perturbed map g is, at best, weakly expanding along the continuation F_g of F for g over the orbit of the periodic point y. By another small perturbation we can now turn any slight expansion into a contraction, which means that y becomes a sink.

A.3. **Connecting lemmas.** The connecting lemma is motivated by the following question. Assume that the unstable manifold of a periodic point accumulates on the stable manifold of another (possibly the same) periodic point. *Is there a small perturbation of the dynamical system for which the two invariant manifolds corresponding to the continuations of the two periodic points actually intersect each other?*

While this is closely related to the problem of the closing lemma, there is one subtle difference. In the context of the closing lemma it suffices to create a periodic orbit passing close to the initial recurrent point x: once that is done, a change of coordinates close to the identity allows us to suppose that the periodic point is x itself. In other words, points near x are indistinguishable from each other. This is clearly no longer true in the setting of the connecting lemma, because now one wants to create orbits connecting invariant (stable and unstable) manifolds that are given a priori.

Even in the simplest case, flows on orientable surfaces, this problem remained open for more than three decades after the corresponding closing problem had been settled. This was also a major difficulty in the proof of the stability conjecture for diffeomorphisms [211], where a main step consists in showing that if the map is not hyperbolic then homoclinic connections may be created by small C^1 perturbations. In [210], Mañé bypassed the lack, at the time, of a connecting lemma by proving several connecting results with additional assumptions: in a few words, he supposed that orbits accumulating the two invariant manifolds spent a positive fraction of time near the associated periodic point.

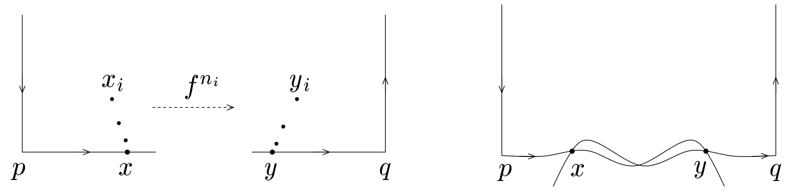

FIGURE 21. Connecting invariant manifolds

Eventually, the problem was solved by Hayashi [150] for C^1 diffeomorphisms and flows. In fact he proves a stronger statement, cf. Figure 21:

Theorem A.4 (Hayashi [150]). *Given two periodic points p and q, if there are orbits visiting the neighborhoods of some point in the unstable manifold of p and some point in the stable manifold of q then, by an arbitrarily C^1-small perturbation, one can make the two invariant manifolds intersect each other.*

Using this result, Hayashi could simplify Mañé's proof of the stability conjecture substantially, as well as extend it to C^1 flows. Moreover, these results were extended to the conservative setting by Xia, Wen [342, 345].

It is fair to say that the C^1 connecting lemma was a decisive ingredient in the recent progress towards understanding C^1 generic dynamical systems. On the other hand, many of these applications require variations of the original statement. In particular, the following question posed itself quite naturally: Assume the unstable manifold of p and the stable manifold of q accumulate on some point x. *Is there a C^1 perturbation of the system for which the two invariant manifolds intersect each other ?* This has been answered recently, assuming x is not periodic itself, independently by Arnaud [18], Hayashi [151], and Wen, Xia [342]. Their statements are slightly different, here we quote [18]:

Theorem A.5. *Let f be a diffeomorphism on a compact manifold, $x \in M$ be a non-periodic point, and $\mathcal{U}$ be a C^1 neighborhood of f. Then there exist $N \geq 1$ such that for any neighborhood V of x there exists another neighborhood $W \subset V$ such that given any points y and z outside $\cup_{0 \leq n < N} f^n(V)$ and such that W contains some forward iterate of y and some backward iterate of z, there exists $g \in \mathcal{U}$ coinciding with f outside $\cup_{0 \leq n < N} f^n(V)$ and for which z is in the forward orbit of y, passing through V.*

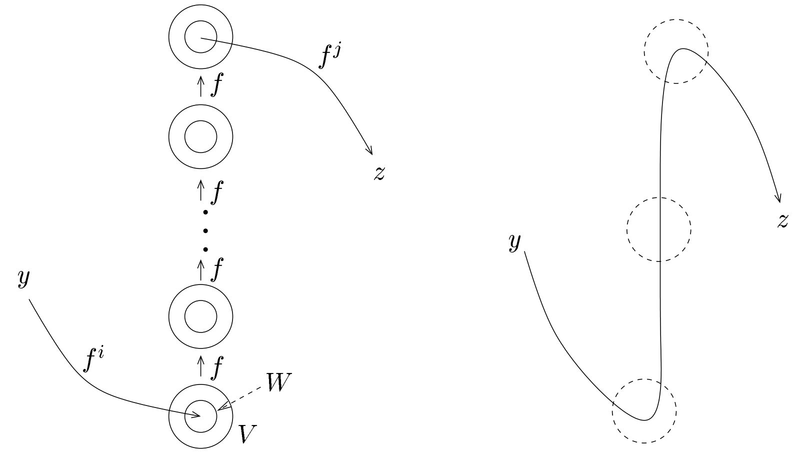

FIGURE 22. Connecting orbits

See Figure 22. The statement remains true for non-compact manifolds, as long as the orbit of x has some accumulation point in M.

A.4. **Sketch of the proof of the closing and connecting lemmas.** Consider a diffeomorphism f on a compact manifold M, and a point $x \in M$ having a positive iterate $f^n(x)$ very close to x. The naive strategy for "closing" the orbit of x is to compose f with a diffeomorphism h close to the identity, such that $h(f^n(x)) = x$. If *the support of* h, that is, the region where h differs from the identity, does not contain any intermediate iterate $f^i(x)$, $i \in \{1, \ldots, n-1\}$ then the diffeomorphism $g = h \circ f$ satisfies $g^n(x) = x$. This simple idea actually gives the proof of the C^0 closing lemma: if $\dim M \geq 2$ and the distance $d(x, f^n(x))$ is less than some number δ then there is a topological disk Δ of diameter δ containing x and $f^n(x)$ and disjoint from $f^i(x)$, $i \in \{1, \ldots, n-1\}$. Now there is a diffeomorphism h supported in Δ, and so δ-close to identity in the C^0 topology, such that $h(f^n(x)) = x$.

This idea is no longer sufficient in the C^1 topology. To explain this, consider a diffeomorphism h such that $h(y) = x$ for some point y with $d(x, y) \leq \delta$. In order that the C^1 distance from h to the identity be bounded by ε we need some "safety space": the support of h must contain the geometric (round) disk of radius $d(x, y)/\varepsilon$ around y. Taking $y = f^n(x)$, there is no reason why this disk of prescribed radius would be disjoint from the orbit segment $f^i(x), i \in \{1, \ldots, n-1\}$. So, while trying to close the orbit by sending $f^n(x)$ to x one might be opening it at some intermediate place.

Pugh's idea for solving this difficulty was to spread the perturbation in time along some segment of the orbit of x. Since we deal with local perturbations, we may think of the dynamics as being essentially linear, so that the key ingredient for reducing the safety zone is the following lemma from linear algebra:

Lemma A.6 ([280]). *Let $\mathbb{R}^d$ be endowed with some Euclidean metric $d(\cdot, \cdot)$, and let M_i be any sequence of linear isomorphisms of $\mathbb{R}^d$. Given $\eta > 0$ and $\varepsilon > 0$ there exists an integer $N \geq 1$ and some basis of $\mathbb{R}^d$ such that the following holds; given any pair of points z, y in the cube $[-1, 1]^d$ of radius 1 relative to this basis, there exist points z_i in the cube of radius $1 + \eta$ such that $z_0 = z$, $z_N = y$, and*

$$d\big(M_i(z_{i-1}), M_i(z_i)\big) \leq \varepsilon d\big(M_i(z_{i-1}), \partial M_i([-(1+2\eta), (1+2\eta)]^d)\big)$$

for all $i = 1, \ldots, N$.

Notice that the case $d = 1$ is easy: it suffices to take for z_i an equally spaced sequence between z and y, and to use the simple fact that one-dimensional linear maps preserve ratios of distances. Thus, in this case by spreading the perturbation along N iterates one reduces the safety zone by a factor N.

To finish the proof of the closing lemma, one considers those iterates of x in the time interval $\{1, \ldots, n\}$ that return to some local chart around x. Among them, one chooses a pair $f^j(x)$ and $f^k(x)$, with $j > k$, which are closest in the following sense. If C represents the smallest cube, in local coordinates, containing $f^j(x)$ and $f^k(x)$ then there is no other iterate $f^l(x)$, $1 \leq l \leq n$ in the cube C_η with the same center as C and radius $(1 + 2\eta)$ as large. Taking the local chart sufficiently small, we may consider the N first iterates of f to be roughly linear restricted to it, so that the previous lemma may be applied with $z = f^j(x)$, $y = f^k(x)$, and $M_i = Df^i(x)$. The factor ε permits to construct a perturbation h supported inside slightly smaller domains: h is the identity outside the union of $f^i(C_\eta)$ over $1 \leq i \leq N$, and satisfies $h(f^i(z_{i-1})) = f^i(z_i)$, where z_i is the sequence given by the lemma. Taking $g = h \circ f$ one has that $g^N(f^j(x)) = f^{k+N}(x)$, and $f^j(x)$ is a periodic orbit for g. Notice that $f^j(x)$ is close to x.

For proving the connecting lemma, Hayashi renounced trying to exhibit a perturbation that would leave the intermediate iterates unchanged. Given two orbit segments $f^i(y)$, $0 \le i \le m$ and $f^{-i}(z)$, $0 \le i \le n$, such that $f^m(y)$ is close to $f^{-n}(z)$, we want to perturb the orbit segments in order that y and z belong in the same orbit for the perturbed map. The strategy is to make perturbations at several places in order to shorten the two segments every time one of them comes close to itself or to the other: at each visit close to some point of a segment, instead of following the original trajectory one jumps directly to last visit close to that point, discarding all intermediate iterates.

Should "closeness" be an equivalence relation, which of course it is not, this strategy could be neatly implemented. As it is, one must replace this relation by more technical versions which, on the other hand, are transitive. The first step is to consider a small cube contained in a local chart around $f^m(y)$ and $f^{-n}(z)$ as in the closing lemma (using Lemma A.6). Next one tiles this cube into sub-cubes such that the size of each tile is comparable, up to a bounded factor, to the size of every one of its neighbors, and so the number of neighbors is uniformly bounded. Moreover, $f^m(y)$ and $f^{-n}(z)$ belong to the same tile. We consider two points as being "close" if they belong to the same tile, and then we carry out the shortening procedure as explained before: orbit segments in between the first and the last visit to each tile are discarded.

We still have to deal with the possibility of having points in neighboring tiles which are actually very close to each other. This is where having a bounded number of neighbors, all with comparable sizes, becomes crucial. We consider another closeness relation for such points, where we say that they are close if the distance is much smaller than the size of the tile and all its neighbors. As the number of neighbors is uniformly bounded, this is an equivalence relation, and the equivalence classes have small diameter. Then, we carry out a second shortening of the orbit segments as before.

At this point local perturbations as in the proof of the closing lemma, with disjoint supports, allow us to close all the jumps introduced by the various shortenings performed in the previous steps, and to complete the proof.

Notice that while the connecting lemma aims at closing a pseudo-orbit with a single jump, the proof creates and then closes pseudo-orbits with any number of jumps. This suggests that one should look for a more general *closing lemma for arbitrary pseudo-orbits.* Such a statement has indeed been obtained, very recently:

Theorem A.7 (Bonatti, Crovisier [54, 55]). *Let f be a C^1 diffeomorphism whose periodic points are all hyperbolic, and y, z be any two points which may be joined by pseudo-orbits with arbitrarily small jumps. Then there exists a C^1 nearby diffeomorphism g for which y and z are in the same orbit.*

This statement remains true within volume-preserving and within symplectic maps, and in the symplectic case one may relax the hypothesis to include elliptic points.

A.5. **Dynamical realizations of perturbations of the derivative.** The following result permits to perform dynamically perturbations of the linear co-cycle defined by the derivative of a diffeomorphism f. It may also be stated uniformly for any g in a neighborhood of f.

Theorem A.8 (Franks [131]). *Let f be a C^1-diffeomorphism defined on a closed manifold M and consider any $\delta > 0$. Then there is $\varepsilon > 0$ such that, given any finite set Σ, any neighborhood U of Σ, and any linear maps*

$$A_x \colon T_x M \to T_{f(x)} M\,, \quad x \in \Sigma$$

such that every A_x is ε-close to the derivative of f at x, there is a diffeomorphism g which is δ-close to f in the C^1 topology, coinciding with f on $M \setminus U$ and on Σ, and such that the derivative of g at any point $x \in \Sigma$ is A_x.

The key point in the proof is that the C^1 norm of a map is not affected by conjugacy under a homothetic transformation. This allows one to perform perturbations of the derivative in arbitrarily small neighborhoods of the point without changing the C^1 size of the perturbation.

This argument clearly fails in any C^r topology for any $r \geq 2$, and so does the conclusion. Since one often uses Franks' lemma to perturb the derivative along periodic orbits, let us give a C^2 counterexample in that context. Consider a linear Anosov map f of the 2-torus and fix some $\delta > 0$. Then for any $\varepsilon > 0$ there is a periodic orbit Σ which is "sufficiently dense" in the torus such that there are ε-perturbations of the derivative at the points of Σ that cannot be realized as δ-perturbations of f in the C^2 topology: they can be realized as C^1 perturbations but the variation of the derivative must be very fast, and so the second derivative is necessarily large.

The previous counter-example is rather artificial, but Pujals, Sambarino [285] give a natural dynamical situation where Franks' lemma fails in the C^2 category. They consider a diffeomorphism f having a non-trivial homoclinic class Λ which is locally maximal, admits a nontrivial dominated splitting, and contains a non-hyperbolic point p (for instance, a saddle-node). A class of examples are the saddle-node horseshoes obtained by collapsing a sink into a hyperbolic horseshoe, as considered in [324].

Clearly, there are periodic points $p_n \in \Lambda$ or arbitrarily large period having a Lyapunov exponent which is arbitrarily close to one: it suffices to consider points passing close enough to p and having only finitely many iterates outside a neighborhood of p. Then, by small perturbations of the derivatives along these orbits one destroys their hyperbolicity. However, the results in [285] ensure that the periods of the non-hyperbolic periodic points of diffeomorphisms C^2 close to f are uniformly bounded.

Appendix B. Normal hyperbolicity and foliations

There are several notions extending uniform hyperbolicity in the following spirit: one still requires the existence of an invariant splitting of the tangent bundle, but one allows one or more of the subbundles involved to be such that the derivative exhibits mixed contracting/neutral/expanding behavior along them. Typically, these central (non-hyperbolic) subbundles concentrate the main dynamical difficulties: often one can project the dynamics along the remaining hyperbolic directions thus reducing the study to the central part.

Brin, Pesin [78] and Hirsch, Pugh, Shub [158] considered *partially hyperbolic systems*, that is, admitting an invariant splitting of the tangent bundle as the sum $E^s \oplus E^c \oplus E^u$ of three subbundles, where E^s is uniformly contracting, E^u is uniformly expanding, and the central direction E^c is "in between" the other two.

Examples include the time-1 map of an Anosov flow and the product of an Anosov diffeomorphism by the identity.

In the context of the stability conjecture, Liao [194], Mañé [206], Pliss [275] were led to the more general notion of *dominated splitting* into two subbundles: one of them is definitely more contracted (or less expanded) than the other, after a uniform number of iterates. We begin by giving the precise definition of dominated splitting (or dominated decomposition) with any number of subbundles, and discussing some basic properties.

B.1. **Dominated splittings.**

Definition B.1. Let $f : M \to M$ be a diffeomorphism on a closed manifold M and K be any f-invariant set. A Df-invariant splitting $T_xM = E_1(x) \oplus \cdots \oplus E_k(x)$, $x \in K$ of the tangent bundle over K is *dominated* if there exists $\ell \in \mathbb{N}$ such that for every $i < j$, every $x \in K$, and every pair of unitary vectors $u \in E_i(x)$ and $v \in E_j(x)$, one has

$$\frac{\|Df^\ell(x)u\|}{\|Df^\ell(x)v\|} < \frac{1}{2} \tag{47}$$

and the dimension of $E_i(x)$ is independent of $x \in K$ for every $i \in \{1, \ldots, k\}$.

We write $E_i \prec E_j$ or, equivalently, $E_j \succ E_i$. In some contexts we need to specify the strength of the domination, measured by the number ℓ of iterates needed to get the inequality in the definition. Then we write $E_i \prec_\ell E_j$, or $E_j \succ_\ell E_i$, and we say that the splitting is *ℓ-dominated*.

Let us list some useful elementary properties of dominated splittings, referring the reader to [58] for proofs. An important ingredient in the proofs of the properties *extension to a neighborhood* and *persistence* is a characterization of existence of a dominated splitting in terms of invariant cone-fields.

Uniqueness: The dominated splitting is unique if one fixes the dimensions of the subbundles: Let d be the dimension of the ambient manifold and let

$$d = n_1 + n_2 + \cdots + n_k\,.$$

Then there is at most one dominated splitting $T_KM = E_1 \oplus E_2 \oplus \cdots \oplus E_k$ over K with dimension of E_i equal to n_i for all i. It is clear that without the dimension requirement, there may be more than one dominated splitting. For instance, if $E^s \oplus E^c \oplus E^u$ is dominated then so are $E \oplus E^u$ and $E^s \oplus F$, where $E = E^s \oplus E^c$ and $F = E^c \oplus E^u$.

Transversality: The angles between any two subbundles of a dominated splitting are uniformly bounded from zero. In fact, the angle between $\oplus_{i\in I} E_i$ and $\oplus_{j\in J} E_j$ is uniformly bounded from zero, for any disjoint sets $I, J \subset \{1, \ldots, k\}$.

Continuity: Every dominated splitting is continuous, meaning that the subbundles $E_i(x)$ depend continuously on the point x.

Extension to the closure: Every ℓ-dominated splitting over a set K may be extended to an ℓ-dominated splitting over the closure of K. Even more, if g_n is a sequence of diffeomorphisms converging to f and K_n is a g_n-invariant set with an ℓ-dominated splitting, with dimensions independent of n, then the f-invariant set

$$K = \limsup_n K_n = \bigcap_N \overline{\bigcup_{n\geq N} K_n}$$

admits an ℓ-dominated splitting.

Extension to a neighborhood: Let K be an f-invariant set with a dominated splitting. Then such a splitting can be extended in a dominated way to the maximal invariant set of f in a neighborhood of K.

Persistence: Every dominated splitting persists under C^1-perturbations. More precisely, consider an f-invariant set K with an ℓ-dominated splitting. Then, for every $\varepsilon > 0$ there is a neighborhood U of K and a C^1-neighborhood $\mathcal{U}$ of f such that, for every $g \in \mathcal{U}$ the maximal g-invariant set in the closure of U admits an $(\ell - \varepsilon)$-dominated splitting, having the same dimensions of the initial dominated splitting over K.

Clustering: Suppose $E_1 \oplus \cdots \oplus E_k$ is a dominated splitting and let $1 \leq i < j \leq k$, with either $i > 1$ or $j < k$. Then

$$E_1 \oplus \cdots \oplus E_{i-1} \oplus F \oplus E_{j+1} \oplus \cdots \oplus E_k\,,$$

where $F = E_i \oplus \cdots \oplus E_j$, is a dominated splitting.

Sub-splittings: Let $E_1 \oplus \cdots \oplus E_k$ be a dominated splitting and suppose that $E_j = F_1 \oplus \cdots \oplus F_m$ is a dominated splitting. Then

$$E_1 \oplus \cdots \oplus E_{j-1} \oplus F_1 \oplus \cdots \oplus F_m \oplus E_{j+1} \oplus \cdots \oplus E_k$$

also is a dominated splitting.

These properties lead to and justify the following notion introduced in [58]:

Proposition B.2. *Let K be an f-invariant set admitting a dominated splitting. Then there is a unique* finest dominated splitting $T_K M = E_1 \oplus \cdots \oplus E_k$ *on K such that no E_i admits a dominated splitting. Any other invariant splitting is obtained by considering a partition of $\{1, \ldots, k\}$ into sub-intervals, and clustering the corresponding subbundles.*

Curiously, the following basic question, which was raised already in [158, page 5], remains open in general:

Problem B.3. Given a dominated splitting $T_K M = E \oplus F$ over an invariant set K, is there a continuous Riemannian metric over K which is adapted to the splitting in the sense that for this norm condition (47) holds with $\ell = 1$?

The following notions were used in the characterization of robustly transitive dynamics given in Section 7:

Definition B.4. Let f be a diffeomorphism and K be an f-invariant set having a dominated splitting $T_K M = E_1 \oplus \cdots \oplus E_k$. We say that the splitting and the set K are

- *partially hyperbolic* if there is $n \geq 1$ such that the restriction of Df^n either contracts uniformly E_1 or expands uniformly E_k.
- *volume hyperbolic* if there is $n \geq 1$ such that the Jacobian of $Df^n \mid E_1$ is strictly less than 1 (uniform volume contraction) and the Jacobian of $Df^\ell \mid E_k$ is strictly greater than 1 (uniform volume expansion).

We say that the diffeomorphism f is partially/volume hyperbolic if the ambient space M is a partially/volume hyperbolic set for f.

Remark B.5. An invariant set is volume hyperbolic if and only if its finest dominated splitting is volume hyperbolic. If a volume hyperbolic splitting has at least one of the extremal bundles E_1 or E_k of dimension 1 then it is partially hyperbolic.

Remark B.6. The notion of partial hyperbolicity is sometimes used in a much more restrictive sense. Many authors (e.g. Pugh, Shub [282]) require both extremal subbundles to be hyperbolic. Moreover, when defining domination one may ask that the maximal expansion of $Df^\ell \mid E_i$ at any point x of K be less than the minimal expansion of $Df^\ell \mid E_{i+1}$ at any point y of K. This corresponds to what Hirsch, Pugh, Shub [158] call *absolute* normal hyperbolicity, whereas the more general notion we use here corresponds to their *relative* normal hyperbolicity.

Just to give some flavor of the dynamical consequences of these properties, let us quote a result on bifurcations of invariant sets with a dominated splitting, from [59]: *Let f be a diffeomorphism having a dominated splitting into one-dimensional subbundles, defined on the whole ambient manifold. Then no diffeomorphism close to f exhibits homoclinic tangencies.* Section 7.2 and [59] explore other consequences.

B.2. **Invariant foliations.** Let f be a C^r diffeomorphism on a closed manifold M admitting a dominated splitting $TM = E^u \oplus E^c \oplus E^s$ where Df is uniformly expanding along E^u and uniformly contracting along E^s.

Theorem B.7 (Brin, Pesin [78], Hirsch, Pugh, Shub [158])**.** *Assuming* $\dim E^u > 0$, *there exists a unique family $\mathcal{F}^u$ of injectively C^r immersed submanifolds $\{\mathcal{F}^u(x) : x \in M\}$ such that $x \in \mathcal{F}^u(x)$ and $\mathcal{F}^u(x)$ is tangent to E^u_x at every $x \in M$. This family is invariant, in the sense that $f(\mathcal{F}^u(x)) = \mathcal{F}(f(x))$ for all $x \in K$, and the leaves $\mathcal{F}^u(x)$ are uniformly contracted by some negative iterate of f. Moreover, $\mathcal{F}^u$ is a continuous lamination of M, meaning that every point has a continuous local chart which trivializes the leaves and is uniformly C^r along each leaf.*

$\mathcal{F}^u$ is the *strong unstable foliation* and its leaves $\mathcal{F}^u(x)$ are called strong unstable manifolds. Local strong unstable manifolds are neighborhoods of the point inside the corresponding leaf. By reversing time one gets strong stable foliations and (local) strong stable manifolds, assuming $\dim E^s > 0$.

More generally, strong stable and strong unstable foliations exist for any invariant subset $K \subset M$ admitting a dominated splitting as above. The family $\mathcal{F}^u$ is indexed by $x \in K$. These strong invariant foliations are continuous laminations, in the same sense as before.

Our knowledge of the properties of the central bundle E^c is quite different. Some of the most basic questions remain open, including the existence in general of a tangent foliation. On the other hand, when they exist such central foliations usually exhibit strong rigidity. Let us discuss the problems and the results in some detail.

As we just mentioned, it is not known in general whether there is a foliation tangent to the central bundle E^c. When E^c has dimension 1, there always exists curves tangent to E^c passing through every point. That is because continuous vector fields always have solutions. But these curves need not be unique, because the central bundle is usually not Lipschitz continuous.

Problem B.8. Let f be a partially hyperbolic diffeomorphism on a compact manifold such that the central bundle E^c has dimension equal to 1. Is there some foliation tangent to E^c ? Is this foliation unique ? Is there some invariant foliation tangent to E^c ? (clearly, invariance follows from uniqueness)

When the central bundle has dimension greater than 1, integrability of E^c in general seems more unlikely:

Problem B.9. (1) Build a robustly transitive partially hyperbolic diffeomorphism on a compact 3-dimensional manifold having open sets where no local surface is tangent to E^c.
(2) Build a robustly transitive diffeomorphism on a compact manifold with dimension larger than 3 having no invariant foliation.

On the other hand, unlike the strong stable and strong unstable foliations, which in general are not robust, most central foliations are structurally stable: actually, that is the case for all known examples. By structural stability, one means that given any nearby C^1 diffeomorphism f there exists a globally defined homeomorphism h_g sending leaves of the central foliation of f to leaves of the central foliation of g and such that $h_g \circ f \circ h_g^{-1}$ is isotopic to g along the leaves.

The following general structural stability result assumes that the initial central foliation is *plaque expansive*: given ε there exists δ such that if x_n and y_n are δ-pseudo orbits with all jumps along the central leaves, and they remain within distance ε for all times, then x_n and y_n are in the same local central leaf for all n.

Theorem B.10 (Hirsch, Pugh, Shub [158])**.** *Suppose f admits a plaque expansive central foliation $\mathcal{F}^c$. Then $(f, \mathcal{F}^c)$ is structurally stable.*

It is shown in [158] that every central foliation of class C^1 is plaque expansive. It is not clear whether differentiability is really necessary:

Problem B.11. Is every central foliation of a transitive partially hyperbolic diffeomorphism plaque expansive ?

The same questions may be asked about *center stable* and *center unstable* foliations, tangent to the sums $E^c \oplus E^s$ and $E^c \oplus E^u$, respectively. When both invariant foliations $\mathcal{F}^{cs}$ and $\mathcal{F}^{cu}$ do exist, the intersections of leaves define an invariant central foliation $\mathcal{F}^c$, and one says that the system is *dynamically coherent*. There are indications that the following question should have an affirmative answer:

Problem B.12. For partially hyperbolic diffeomorphisms on 3-dimensional manifolds with 1-dimensional invariant subbundles, does existence of a unique central foliation imply dynamical coherence ?

The ergodic properties of orbits propagate along stable and unstable sets, since forward time averages are constant on stable sets and backward time averages are constant on unstable sets. In order to use this idea, often called the Hopf argument, one needs the holonomies of foliations to preserve zero measure sets. This *absolute continuity* property does hold for the strong stable and strong unstable foliations of partially hyperbolic C^2 diffeomorphisms [78]. We discuss this issue in much more generality in Appendix C.3.

In contrast, even when there exists a unique central foliation, it needs not be absolutely continuous. Open sets of partially hyperbolic diffeomorphisms with such pathological central foliations were first exhibited by Shub, Wilkinson [313]. See Section 7.4.1. Their construction also shows that neither $\mathcal{F}^{cu}$ nor $\mathcal{F}^{cs}$ are absolutely continuous, in general.

Appendix C. Non-uniformly hyperbolic theory

We review some fundamental results from Pesin theory that are used in the text. For more information, see Mañé [209, Chapter 4], Pollicott [277], or the supplementary chapter by Katok, Mendoza in [171].

C.1. **The linear theory.** Let $f : M \to M$ be an invertible measurable transformation, and μ be an f-invariant probability on M. Let $\pi : \mathcal{E} \to M$ be a finite-dimensional vector bundle over M, endowed with a measurable Riemannian metric $\|\cdot\|$. Let $F : \mathcal{E} \to \mathcal{E}$ be a measurable vector bundle automorphism over f: this means that $\pi \circ F = f \circ \pi$ and the action $F_x : \mathcal{E}_x \to \mathcal{E}_{f(x)}$ of F on each fiber $\mathcal{E}_x = \pi^{-1}(x)$ is a linear isomorphism.

Oseledets ergodic theorem states that, under an integrability condition, almost every fiber splits as a direct sum of subspaces such that the iterates $F_x^n : \mathcal{E}_x \to \mathcal{E}_{f^n(x)}$ have well-defined rates of exponential growth, in norm, restricted to each subspace:

Theorem C.1 (Oseledets [253]). *Assume $\log^+ \|F_x\|$ and $\log^+ \|F_x^{-1}\|$ are integrable for μ. For μ-almost all $x \in M$, there exist $k(x) \in \mathbb{N}$, real numbers $\hat\lambda_1(x) > \cdots > \hat\lambda_{k(x)}$, and a splitting $\mathcal{E}_x = E_x^1 \oplus \cdots \oplus E_x^{k(x)}$, such that $F_x(E_x^i) = E_{f(x)}^i$ for $1 \le i \le k(x)$, and*

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|F_x^n(v_i)\| = \hat\lambda_i(x) \quad \textit{for every non-zero } v_i \in E_x^i\,.$$

The Lyapunov exponents $\hat\lambda_i(x)$, and the Lyapunov subspaces E_x^i are unique μ-almost everywhere, and they depend measurably on x.

In most situations we are interested in, M is a compact space and the transformation f, the vector bundle $\pi : \mathcal{E} \to M$, the automorphism F, and the norm $\|\cdot\|_x$ are all continuous. Then the integrability hypothesis is automatically satisfied, and the Lyapunov exponents and Lyapunov subspaces do not depend on the actual choice of the norm. An important particular case: $f : M \to M$ is a C^1 diffeomorphism on a compact Riemannian manifold M, $\mathcal{E} = TM$ is the tangent bundle, and $F = Df$ is the derivative of f.

Theorem C.1 extends to the case when f is non-invertible: instead of a splitting one gets a filtration $\mathcal{E}_x > E_x^1 > \cdots > E_x^{k(x)} > 0 = E_x^{k(x)+1}$ of the fiber into vector subspaces such that

$$\lim_{n\to+\infty} \frac{1}{n} \log \|F_x^n(v_i)\| = \hat\lambda_i(x) \quad \text{for every } v_i \in E_x^i \setminus E_x^{i+1}\,.$$

There is also a continuous time version of the theorem, where one considers a flow (or semi-flow) $(F^t)_{t\in\mathbb{R}} : \mathcal{E} \to \mathcal{E}$ of vector bundle automorphisms over a flow (or semi-flow) $(f^t)_{t\in\mathbb{R}} : M \to M$.

An outline of Mañé's proof [209] of Theorem C.1 follows. It is no restriction to suppose that μ is ergodic. The largest Lyapunov exponent is the growth rate of the norm over the whole fiber:

$$\hat\lambda_1(x) = \limsup_{n\to+\infty} \frac{1}{n} \log \|F_x^n\|$$

and E_x^1 is the subset of vectors $v \in \mathcal{E}_x$ realizing this growth rate, in the sense that

$$\liminf_{n\to-\infty} \frac{1}{n} \log \|F_x^n(v)\| \ge \hat\lambda_1(x). \tag{48}$$

Then E_x^1 is a vector subspace, the limit actually exists for both $n \to \pm\infty$, and it coincides with $\hat\lambda_1(x)$. One key step in the proof is to show that the dimension of E_x^1 is positive at μ-almost every point. This is done as follows.

Fix $\varepsilon > 0$. Given $N \geq 1$, consider the set $Y_{\varepsilon,N}$ of points x such that

$$\log \|F_x^n(u)\| \leq n(\hat{\lambda}_1(x) - \varepsilon)$$

for some non-zero vector u and for all $0 > n > -N$. Using the definition of $\hat{\lambda}_1(x)$ and Pliss Lemma 11.5, one finds that $\mu(Y_{\varepsilon,N}) \geq \delta$ where $\delta > 0$ is independent of N. Consequently, $\mu(\cap_N Y_{\varepsilon,N}) \geq \delta$. Thus, there is a positive measure set of points such that

$$\liminf_{n\to-\infty} \frac{1}{n} \log \|F_x^n(u)\| \geq \hat{\lambda}_1(x) - \varepsilon$$

for some non-zero vector u. By ergodicity, this (invariant) set has full measure. Making $\varepsilon \to 0$, it follows that $\dim E_x^1 > 0$ at μ-almost every point.

Another key step in the proof of Theorem C.1 is a graph-transform construction of a measurable subbundle $\mathcal{E}'$, invariant under F and such that $\mathcal{E} = \mathcal{E}' \oplus E^1$. The proof proceeds by induction on the dimension of the vector bundle, considering the restriction of F to $\mathcal{E}'$.

Remark C.2. In general, the angles between Lyapunov subspaces are not bounded from zero, not even over individual orbits. However, angles may decrease at most sub-exponentially fast: for μ-almost x and any disjoint subsets I, J of $\{1, \ldots, k(x)\}$,

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \operatorname{angle}\big(\oplus_{i\in I} E^i_{f^n(x)}, \oplus_{j\in J} E^j_{f^n(x)}\big) = 0.$$

One consequence is the following useful relation:

$$\sum_{j=1}^{k(x)} \hat{\lambda}_i(x) \dim E_x^i = \lim_{n\to\pm\infty} \frac{1}{n} \log |\det(F_x^n)| \,.$$

C.2. **Stable manifold theorem.** Let $f: M \to M$ be a C^1 diffeomorphism on a compact manifold, and let $\mathcal{E} = TM$ and $F = Df$. The *stable set* of a point $x \in M$ is defined by

$$W^s(x) = \{y \in M : d(f^n(x), f^n(y)) \to 0 \text{ as } n \to +\infty\}$$

The unstable set $W^u(x)$ is defined analogously, with f^{-1} in the place of f. For $x \in M$ as in Theorem C.1, define

$$E_x^s = \oplus_{\hat{\lambda}_j(x)<0} E_x^j \quad \text{and} \quad E_x^u = \oplus_{\hat{\lambda}_j(x)>0} E_x^j.$$

Moreover, let $\tau_x > 0$ be smaller than the norm of every non-zero Lyapunov exponent at x. Pesin's stable manifold theorem states that x has a *local stable manifold* which is a smooth disk tangent to E_x^s at x and whose points approach the orbit of x exponentially fast:

Theorem C.3 ([128, 273, 281, 302]). *Assume the derivative of f is Hölder continuous. Then there exists a C^r embedded disk $W^s_{loc}(x)$ through x and there exists $C_x > 0$ such that*

(1) *$W^s_{loc}(x)$ is tangent to E^s_x at x;*
(2) *$d(f^n(x_1), f^n(x_2)) \leq C_x e^{-n\tau_x} \operatorname{dist}(x_1, x_2)$ for all $x_1, x_2 \in W^s_{loc}(x)$;*
(3) *$f\big(W^s_{loc}(x)\big) \subset W^u_{loc}(f(x))$;*
(4) *if all Lyapunov exponents are non-zero, $W^s(x) = \bigcup_{n=0}^\infty f^{-n}\big(W^u_{loc}(f^n(x)\big)$.*

Analogously, x has a *local unstable manifold* $W^u_{loc}(x)$ satisfying corresponding properties with f^{-1} in the place of f.

In addition, the local stable manifold and local unstable manifold depend measurably on x (as C^1 embedded disks), and the constants C_x, τ_x may also be chosen depending measurably on the point. One way to understand this statement is that one may find compact *hyperbolic blocks* $\mathcal{H}(\kappa)$, $\kappa \in \mathbb{N}$, such that

- $C_x < \kappa$ and $\tau_x > 1/\kappa$ for every $x \in \mathcal{H}(\kappa)$;
- $\mathcal{H}(\kappa) \subset \mathcal{H}(\kappa+1)$ for all $\kappa \in \mathbb{N}$ and $\mu(\mathcal{H}(\kappa)) \to 1$ as $\kappa \to \infty$.
- the C^1 disks $W^s_{loc}(x)$ and $W^u_{loc}(x)$ vary continuously with x in $\mathcal{H}(\kappa)$.

In particular, the sizes of $W^s_{loc}(x)$ and $W^u_{loc}(x)$ are uniformly bounded from zero on each $x \in \mathcal{H}(\kappa)$, and so is the angle between the two disks.

C.3. **Absolute continuity of foliations.** Most important for the applications, the holonomy maps associated to Pesin's stable lamination $\mathcal{W}^s = \{W^s_{loc}(x)\}$ preserve zero Lebesgue measure sets. More precisely, let x be in some hyperbolic block $\mathcal{H}(\kappa)$ and x_1, x_2 be points of $W^s_{loc}(x)$ close to x. Let Σ_1 and Σ_2 be small smooth disks transverse to $W^s_{loc}(x)$ at x_1 and x_2:

$$T_{x_i}\Sigma_i \oplus T_{x_i}W^s_{loc}(x) = T_{x_i}M \quad \text{for } i = 1, 2.$$

Given any point $y \in \mathcal{H}(\kappa)$ close to x, the local stable manifold $W^s_{loc}(y)$ intersects each cross section Σ_i in exactly one point y_i. This defines a holonomy map

$$\pi^s : y_1 \to y_2$$

which is a homeomorphism between subsets of Σ_1 and Σ_2.

Theorem C.4. [273, 281] *Assume the derivative of f is Hölder continuous. Then Pesin's stable lamination $\mathcal{W}^s$ is absolutely continuous: every holonomy map π^s as before maps zero Lebesgue measure subsets of Σ_1 to zero Lebesgue measure subsets of Σ_2. A dual statement holds for the unstable lamination $\mathcal{W}^u$.*

This fundamental property was first observed by Anosov [11], for the class of systems named after him. Even in that uniformly hyperbolic setting the statement is not true, in general, if one drops the assumption of Hölder continuity of the derivative [295].

C.4. **Conditional measures along invariant foliations.** Let κ and $x \in \mathcal{H}(\kappa)$ be fixed. For each small $\delta > 0$, let $\mathcal{B}^s_\delta(\kappa, x)$ be the union of the local stable manifolds through points $y \in \mathcal{H}(\kappa)$ in the δ-neighborhood of x. Let $\nu = \nu(\kappa, x, \delta)$ be the restriction of μ to $\mathcal{B}^s_\delta(\kappa, x)$.

Let $\mathcal{P}$ be the coarsest partition of $\mathcal{B}^s_\delta(\kappa, x)$ such that every local stable manifold $W^s_{loc}(y)$ is entirely contained in some atom of $\mathcal{P}$. Then $\mathcal{P}$ is a *measurable partition* in the sense of Rokhlin [297]: it may be written as the product

$$\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n \tag{49}$$

of an increasing sequence of *finite* partitions $\mathcal{P}_n$ of $\mathcal{B}^s_\delta(\kappa, x)$.

Let $\hat{\nu}$ be the projection of ν to $\mathcal{P}$, that is, the push-forward of ν under the map $x \mapsto \mathcal{P}(x)$ assigning to each point x the atom of $\mathcal{P}$ that contains it. By Rokhlin's disintegration theorem, there exists a family $\{\nu_P : P \in \mathcal{P}\}$ of *conditional probability measures* satisfying

(1) $\nu(P) = 1$ for $\hat{\nu}$-almost every $P \in \mathcal{P}$;
(2) $\nu(E) = \int \nu_P(E)\, d\hat{\nu}(P)$ for every measurable subset E,

It is not difficult to see that such a family is essentially unique: given any other choice $\{\nu'_P : P \in \mathcal{P}\}$, we have $\nu_P = \nu'_P$ for $\hat{\nu}$-almost every P.

One way to construct the family of conditional measures is as follows. For each $P \in \mathcal{P}$ and continuous function φ define

$$E_n(\varphi, P) = \frac{1}{\nu(P_n)} \int_{P_n} \varphi\, d\nu$$

where P_n is the element of the partition $\mathcal{P}_n$ that contains P. We may restrict to the case when the denominator is non-zero, for the union of all $P \in \mathcal{P}$ such that $\nu(P_n) = 0$ for some n has zero measure. The main step is to show that for each function φ the sequence $E_n(\varphi, P)$ converges at $\hat{\nu}$-almost every P. Let $E(\varphi, P)$ be the limit. Using the fact that the space of continuous function has a countable dense subset, there is a full ν-measure subset of $P \in \mathcal{P}$ such that $E(\varphi, P)$ is defined for all continuous functions φ. By the Riesz-Markov theorem, each $E(\cdot, P)$ corresponds to a probability ν_P on M, and one checks that these probabilities satisfy (1) and (2) above.

Definition C.5. An invariant measure μ is *absolutely continuous along stable manifolds* if, given any $\mathcal{B}^s_\delta(\kappa, x)$ as above, the conditional measure ν_P are absolutely continuous with respect to Lebesgue measure in P for $\hat{\mu}$-almost every $P \in \mathcal{P}$.

This makes sense because the atoms $P \in \mathcal{P}$ are embedded submanifolds. One defines absolute continuity along unstable manifolds in the same way.

Lebesgue measure is always absolutely continuous along stable manifolds and along unstable manifolds. This is a consequence of Theorem C.4.

C.5. **Local product structure.** Here we assume that μ is a *hyperbolic measure* for f, that is, an invariant probability such that all Lyapunov exponents $\hat{\lambda}_j(x)$ are non-zero at μ-almost every point x. We also assume that μ is non-atomic.

Let $x \in \mathcal{H}(\kappa)$ and y be any point of $\mathcal{H}(\kappa)$ in a closed δ-neighborhood $B_\delta(x)$ of x. If δ is small enough, $W^s_{loc}(y)$ intersects $W^u_{loc}(x)$ exactly in one point, and analogously for $W^u_{loc}(y)$ and $W^s_{loc}(x)$. Let

$$\mathcal{H}^u_\delta(\kappa, x) \subset W^u_{loc}(x) \quad \text{and} \quad \mathcal{H}^s_\delta(\kappa, x) \subset W^s_{loc}(x)$$

be the (compact) intersection sets obtained in this way. Reducing δ if necessary, $W^s_{loc}(\xi) \cap W^u_{loc}(\eta)$ consists of exactly one point $[\xi, \eta]$, for every $\xi \in \mathcal{H}^u_\delta(\kappa, x)$ and $\eta \in \mathcal{H}^s_\delta(\kappa, x)$. Let $\mathcal{H}_x(\kappa, \delta)$ be the set of such points $[\xi, \eta]$. By construction, $\mathcal{H}_x(\kappa, \delta)$ contains $\mathcal{H}(\kappa) \cap B_\delta(x)$ and its diameter goes to zero when $\delta \to 0$. Moreover, $\mathcal{H}_x(\kappa, \delta)$ is homeomorphic to $\mathcal{H}^u_\delta(\kappa, x) \times \mathcal{H}^s_\delta(\kappa, x)$ via

$$(\xi, \eta) \mapsto [\xi, \eta]\,.$$

Definition C.6. A hyperbolic measure μ has *local product structure* if, for μ-almost every point x and every small $\delta > 0$ as before, the restriction $\nu = \mu \mid \mathcal{H}_\delta(\kappa, x)$ is equivalent to the product measure $\nu^u \times \nu^s$, where ν^u and ν^s are the projections of ν to $\mathcal{H}^u_\delta(\kappa, x)$ and $\mathcal{H}^s_\delta(\kappa, x)$, respectively.

The measure μ has local product structure if and only if stable holonomy maps (respectively unstable holonomy maps) are absolutely continuous relative to the conditional measures along leaves of the unstable lamination (respectively the stable lamination). In particular, Lebesgue measure always has local product structure if it is invariant and hyperbolic. More generally, using Theorem C.4, every hyperbolic invariant measure which is absolutely continuous along stable manifolds or along unstable manifolds has local product structure.

Appendix D. Random perturbations

Here we discuss the notion of stochastic stability more precisely, first in the case of discrete-time systems $f : M \to M$. Let U be an open subset of M such that $f(U)$ is relatively compact in U. The noise level $\varepsilon > 0$ will always be smaller than $\operatorname{dist}(f(U), M \setminus U)$, so that random orbits can not escape from U.

General presentations of random dynamical systems can be found in the books of Arnold [21] and Kifer [178, 179].

D.1. **Markov chain model.** This is defined by a family $\{p_\varepsilon(\cdot \mid z) : z \in U,\ \varepsilon > 0\}$ of Borel probability measures, such that every $p_\varepsilon(\cdot \mid z)$ is supported inside the ε-neighborhood of $f(z)$. The random orbits are the sequences $\{z_j\}$ where each z_{j+1} is a random variable with probability distribution $p_\varepsilon(\cdot \mid z_j)$.

A probability measure μ_ε is *stationary* for the Markov chain if

$$\mu_\varepsilon(E) = \int p_\varepsilon(E \mid z)\, d\mu_\varepsilon(z) \tag{50}$$

for every Borel set $E \subset U$. Equivalently, the skew-product measure $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$

$$d(\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}})(z_0, z_1, \dots, z_n, \dots) = \mu_\varepsilon(dz_0)\, p_\varepsilon(dz_1 \mid z_0) \cdots p_\varepsilon(dz_n \mid z_{n-1}) \cdots$$

is invariant under the shift map $\mathcal{F} : U \times U^{\mathbb{N}} \to U \times U^{\mathbb{N}}$ in the space of random orbits $\{z_j : j \geq 0\}$. By the ergodic theorem, the time average of every continuous function $\varphi : U \to \mathbb{R}$

$$\tilde{\varphi}(\mathbf{z}) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(z_j) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ \pi_0)(\mathcal{F}^j(\mathbf{z}))$$

exists over a full $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-measure subset of random orbits $\mathbf{z} = \{z_j\}$. Stationary measures always exist, if the transition probabilities $p_\varepsilon(\cdot \mid z)$ depend continuously on the point z. Consider the operator $\mathcal{T}_\varepsilon$ acting on the space of Borel measures by

$$\mathcal{T}_\varepsilon \eta(E) = \int p_\varepsilon(E \mid z)\, d\eta(z).$$

Lemma D.1. *Suppose $z \mapsto p_\varepsilon(\cdot \mid z)$ is continuous, relative to the weak* topology in the space of Borel probabilities. Then every weak* accumulation point μ_ε of the sequence $n^{-1} \sum_{j=0}^{n-1} \mathcal{T}_\varepsilon^j \eta$ is a stationary measure, for any probability measure η supported in U.*

Proof. The space of probability measures on the closure of $f(U)$ is weak* compact, and so accumulation points do exist. The assumption on the transition probabilities ensures that the operator $\mathcal{T}_\varepsilon$ is continuous, relative to the weak* topology. It follows that every accumulation point is a fixed point for $\mathcal{T}_\varepsilon$. This is equivalent to being a stationary measure. □

A function $\phi : U \to \mathbb{R}$ is *invariant* for the Markov chain if $\phi(x) = \int \phi(y) p_\varepsilon(x, dy)$ for μ_ε-almost every x. A stationary measure μ_ε is *ergodic* if every invariant function is constant μ_ε-almost everywhere.

Lemma D.2. *If μ_ε is ergodic then $\tilde{\varphi}(\mathbf{z}) = \int \varphi \, d\mu_\varepsilon$ for every continuous function $\varphi : U \to \mathbb{R}$ and $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-almost every random orbit $\mathbf{z}$.*

Proof. For each $k \geq 0$, define

$$\tilde{\varphi}_k(z_0, \dots, z_k) = \int \tilde{\varphi}(\mathbf{z}) \, p_\varepsilon(dz_{k+1} \mid z_k) \cdots p_\varepsilon(dz_{n+1} \mid z_n) \cdots$$

Using $\tilde{\varphi} = \tilde{\varphi} \circ \mathcal{F}$, we get that $\tilde{\varphi}_0$ is an invariant function, and so $\tilde{\varphi}_0$ is constant $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-almost everywhere. Moreover, $\tilde{\varphi}_k = \tilde{\varphi}_{k-1} \circ \mathcal{F}$ for every $k \geq 1$, and so every $\tilde{\varphi}_k$ is constant almost everywhere. Using $\tilde{\varphi}(\mathbf{z}) = \lim \tilde{\varphi}_k(z_0, \dots, z_k)$, we conclude that the same is true for $\tilde{\varphi}$, and that implies the lemma. □

By Kifer [178, Proposition 2.1], every stationary measure is a convex combination of ergodic ones. Araújo [13] proves that, in great generality, the set of all stationary measures is a finite-dimensional simplex, whose extreme points are the ergodic measures. Moreover, the time average $\tilde{\varphi}(\mathbf{z})$ exists for every initial point z_0 and almost every choice of z_j, $j \geq 1$, and it is given by one of the ergodic measures.

D.2. **Iterations of random maps.** In this model one considers sequences obtained by iteration $z_j = g_j \circ \cdots \circ g_1(z_0)$ of maps g_j chosen at random ε-close to the original f. One is given a family $\{\nu_\varepsilon : \varepsilon > 0\}$ of probabilities in the space of C^r maps, some $r \geq 0$, such that each $\operatorname{supp} \nu_\varepsilon$ is contained in the ε-neighborhood of f. The random orbit associated to a $(z_0, \mathbf{g})$ is the sequence $z_j = g_j \cdots g_1(z_0)$, where the g_j are independent random variables with probability distribution ν_ε.

There is an associated Markov chain, defined by

$$p_\varepsilon(E \mid z) = \nu_\varepsilon\big(\{g : g(z) \in E\}\big). \tag{51}$$

Observe that, for any Borel subsets $A_0, A_1, \dots, A_m$ of U,

$$\big(\mu_\varepsilon \times \nu_\varepsilon^{\mathbb{N}}\big)\big(\{(z_0, \mathbf{g}) : z_0 \in A_0, g_1(z_0) \in A_1, \dots, g_m \cdots g_1(z_0) \in A_m\}\big) = \\ = (\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}})(A_0 \times A_1 \times \cdots \times A_m).$$

This means that the statistics of the random orbits obtained from randomly perturbing the dynamical system are faithfully reproduced by the Markov chain. A probability measure μ_ε is stationary, in the sense of (50), if and only if

$$\mu_\varepsilon(E) = \int \mu_\varepsilon(g^{-1}(E)) \, d\nu_\varepsilon(g) \tag{52}$$

for every Borel set $E \subset U$. Most of the time we deal with continuous maps. Then the transition probabilities $p_\varepsilon(\cdot \mid z)$ given by (51) vary continuously with z, and so stationary measures do exist.

Example D.3. (Additive noise) Let $f : M \to M$ be a diffeomorphism on some Lie group M, e.g., a d-dimensional torus. Let $\{\theta_\varepsilon : \varepsilon > 0\}$ be a family of probability measures on M such that every $\operatorname{supp} \theta_\varepsilon$ is contained in the ε-neighborhood of unity $e \in M$. Let ν_ε be the measure induced by θ_ε in the space of diffeomorphisms, via left multiplication $M \ni \xi \mapsto \xi \cdot f$. The corresponding transition probabilities are

$$p_\varepsilon(E \mid z) = \theta_\varepsilon(\{\xi \in M : \xi \cdot f(z) \in E\}) = \theta_\varepsilon(E \cdot f(z)^{-1}),$$

where $f(z)^{-1}$ represents the algebraic inverse of $f(z)$.

Example D.4. (Noise in parameter space) Let $F : \mathbb{R}^k \times U \to U$ be a C^r map such that every $F(0,\cdot) = f$ and $f_t(U)$ is relatively compact in U for every $f_t = F(t,\cdot)$. Let θ_ε be the normalized Lebesgue measure on the ε-ball around 0, and ν_ε be the probability induced by θ_ε in the space of diffeomorphisms via the parametrization $t \mapsto f_t$. Clearly, $\operatorname{supp} \nu_\varepsilon \to f$ in the C^r topology as $\varepsilon \to 0$.

D.3. **Stochastic stability.** Let μ be an invariant probability measure for $f : U \to U$. The main case we have in mind is when μ is an SRB measure supported in an attractor Λ with the basin property (6), and U is contained in the basin of attraction of Λ. In that case,

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z)) \to \int \varphi\, d\mu$$

for Lebesgue almost every $z \in U$ and every continuous φ.

First, suppose there is a unique stationary probability measure μ_ε, for every small $\varepsilon > 0$. Then

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(z_j) \to \int \varphi\, d\mu_\varepsilon$$

for almost every random orbit $\{z_j\}$ and every continuous φ. Uniqueness holds, for instance, if the attractor $\cap_{n\geq 1} f^n(U)$ is transitive, assuming $p_\varepsilon(\cdot \mid z)$ is absolutely continuous and the support contains a $\rho(\varepsilon)$-neighborhood of $f(z)$. See Araújo [14, Proposition 2.1].

Definition D.5. The system (f,μ) is *stochastically stable* under the perturbation scheme $\{p_\varepsilon(\cdot \mid z) : z \in U,\ \varepsilon > 0\}$ (or $\{\nu_\varepsilon : \varepsilon > 0\}$) if μ_ε converges to μ in the weak* sense:

$$\lim_{\varepsilon\to 0} \int \varphi\, d\mu_\varepsilon = \int \varphi\, d\mu \quad \text{for every continuous } \varphi : U \to \mathbb{R}.$$

The definition extends naturally to the case where the stationary measure is not unique: the whole simplex should converge to μ when $\varepsilon \to 0$.

Remark D.6. It is useful to extend the definition to domains containing several relevant invariant measures, including the whole manifold M. For instance, an attractor may support more than one SRB measure (although it is not yet clear how general this phenomenon is). Also, if the basin of attraction is just a positive Lebesgue measure set, or if one considers random noise which is not supported on small neighborhoods, then random orbits may escape from the basin of attraction. In such cases, a more global notion of stochastic stability can be applied: the simplex of stationary measures should be weak* close to the simplex generated by the SRB measures of the unperturbed system, when ϵ is small. This is substantiated by the following important result of Araújo [14] for absolutely continuous random noise: if there is stochastic stability restricted to each basin of attraction, and the basins contain Lebesgue almost every point, then the system is stochastically stable in the whole ambient space $U = M$, in the previous sense.

In existing results of stochastic stability, the assumptions on the random noise often depend on the class of systems. Even in the simplest uniformly hyperbolic setting, one can not expect stability under *every* Markov chain perturbation:

Example D.7. (Keller [174], Froyland [132]) Let $f : S^1 \to S^1$, $f(z) = 2z \bmod \mathbb{Z}$. Lebesgue measure on $S^1 = \mathbb{R}/\mathbb{Z}$ is the unique SRB measure of f. Consider the Markov chain defined by

$$p_\varepsilon(\cdot \mid z) = \begin{cases} \text{normalized Lebesgue on } (-\varepsilon, \varepsilon) & \text{if } z \in (-\varepsilon, \varepsilon) \\ \text{normalized Lebesgue on } (f(z) - \varepsilon, f(z) + \varepsilon) & \text{if } z \notin (-\varepsilon, \varepsilon) \end{cases}$$

The Markov chain has a unique stationary measure, μ_ε = normalized Lebesgue measure on $(-\varepsilon, \varepsilon)$, but μ_ε does not converge to the SRB measure when $\varepsilon \to 0$.

In the uniformly hyperbolic case, Kifer [179] assumes absolutely continuity with respect to Lebesgue measure m

$$p_\varepsilon(\cdot \mid z) = \rho_{\varepsilon,z}\, m$$

for some integrable function $\rho_{\varepsilon,z}$, and a kind of Lipschitz dependence of the transition probabilities on the point z; see [179, Chapter 2]. On the other hand, C^2 uniformly hyperbolic diffeomorphisms are stochastically stable under any random maps scheme $\{\nu_\varepsilon : \varepsilon > 0\}$: see Young [351] and [336, Section 4.5], where it is shown that her absolute continuity condition is superfluous.

Problem D.8. Are uniformly expanding or uniformly hyperbolic maps stochastically stable under any Markov chain perturbation with transition probabilities varying continuously/Lipschitz with the point ?

Stochastic stability under every random maps scheme $\{\nu_\varepsilon : \varepsilon > 0\}$ includes stability under deterministic perturbations (let ν_ε = Dirac measure): the SRB measure must vary continuously with the dynamical system. Alves, Viana [9] call the latter property *statistical stability*. This is a fairly restrictive property: An example of Keller [174, § 6] shows that even uniformly expanding piecewise smooth maps in dimension 1 are not always statistically stable. However, [9] exhibits open sets of statistical stability outside the uniformly hyperbolic domain.

Problem D.9. Are piecewise smooth uniformly expanding maps always stable under additive absolutely continuous random noise ?

Restricting to Markov chains that come from a random iteration process via (51) enriches the set of tools available for proving stochastic stability. But we lack good examples to decide whether it is strictly necessary for the conclusion. Whether a given Markov chain does derive from the iteration of random maps is largely a matter of smoothness of the noise:

D.4. **Realizing Markov chains by random maps.** This problem is discussed by Kifer [178, Section 1.1] and Benedicks, Viana [42]. Kifer proves that under a mild condition on the ambient space, and assuming $z \mapsto p_\varepsilon(z \mid E)$ is measurable for every Borel set E, such a realization is possible in the space of measurable maps. When the transition probabilities have positive densities we can say more [42]: the Markov chain is represented by parametrized families of maps at least as regular as the densities themselves.

For simplicity, we suppose $U \subset \mathbb{R}^d$, but the arguments extend immediately to manifolds. And the support may be any homeomorphic image of $f(z) + [-\varepsilon, \varepsilon]^d$, under a homeomorphism that depends continuously on the point z.

Proposition D.10. *Suppose every $p_\varepsilon(\cdot \mid z)$, $z \in U$, is absolutely continuous with respect to Lebesgue measure m, with density $\rho_{\varepsilon,z} > 0$ on the support $f(z) + [-\varepsilon, \varepsilon]^d$.*

Suppose $r_\varepsilon : (\xi, z) \mapsto \rho_{\varepsilon,z}(f(z)+\xi)$ is continuous on $[-\varepsilon,\varepsilon]^d \times U$. Then there exist continuous maps $F_\varepsilon : [0,1]^d \times U \to U$ such that

$$p_\varepsilon(E \mid z) = m(\{\omega : F_\varepsilon(\omega, z) \in E\}) \quad \textit{for every Borel set } E.$$

Proof. For $z \in U$ and $\xi = (\xi_1, \dots, \xi_d) \in [-\varepsilon,\varepsilon]^d$, let

$$\omega_i = \frac{p_\varepsilon(f(z) + [-\varepsilon,\xi_1] \times \cdots \times [-\varepsilon,\xi_i] \times [-\varepsilon,\varepsilon]^{d-i} \mid z)}{p_\varepsilon(f(z) + [-\varepsilon,\xi_1] \times \cdots \times [-\varepsilon,\xi_{i-1}] \times [-\varepsilon,\varepsilon]^{d-i+1} \mid z)} \tag{53}$$

for $i = 1, \dots, d$ (for $i = 1$ the denominator is 1). Since the density is positive, this is well-defined if $\xi_1, \dots, \xi_{i-1}$ are all larger than $-\varepsilon$. Given any $1 \le s \le j$, the probability $p_\varepsilon(f(z) + [-\varepsilon,\xi_1] \times \cdots \times [-\varepsilon,\xi_j] \times [-\varepsilon,\varepsilon]^{d-j} \mid z)$ is comparable to

$$(\xi_s + \varepsilon) \int_{-\varepsilon}^{\xi_1} \cdots \int_{-\varepsilon}^{\xi_{s-1}} \int_{-\varepsilon}^{\xi_{s+1}} \cdots \int_{-\varepsilon}^{\xi_j} \int_{-\varepsilon}^{\varepsilon} \cdots \int_{-\varepsilon}^{\varepsilon} (r_\varepsilon \mid \{\xi_s = -\varepsilon\})\, dm,$$

in the sense that the quotient goes to 1 when $\xi_s + \varepsilon$ goes to zero. There is a similar estimate when several $\xi_s + \varepsilon$ go to zero simultaneously. It follows that (53) extends continuously to the points where the denominator is zero.

For fixed $\xi_1, \dots, \xi_{i-1}$ the map $\xi_i \mapsto \omega_i$ is a homeomorphism from $[-\varepsilon,\varepsilon]$ onto $[0,1]$. So, $\psi_{\varepsilon,z} : (\xi_1, \dots, \xi_d) \mapsto (\omega_1, \dots, \omega_d)$ is a homeomorphism onto $[0,1]^d$. Moreover, $(\xi, z) \mapsto \psi_{\varepsilon,z}(\xi)$ is continuous, because the densities depend continuously on both variables. Let $\phi_{\varepsilon,z}$ be the inverse of $\psi_{\varepsilon,z}$. Then $(\omega, z) \mapsto \phi_{\varepsilon,z}(\omega)$ and $F_\varepsilon(\omega, z) = f(z) + \phi_{\varepsilon,z}(\omega)$ are continuous. Finally, the definition (53) gives

$$p_\varepsilon(E \mid z) = \omega_1 \cdots \omega_d = m(\{\omega : \phi_{\varepsilon,z}(\omega) \in E - f(z)\}) = m(\{\omega : F_\varepsilon(\omega, z) \in E\})$$

for every $E = f(z) + [-\varepsilon,\xi_1] \times \cdots \times [-\varepsilon,\xi_d]$. Since these rectangles generate the σ-algebra on the support, the relation must be true for every Borel set. □

Thus, the Markov chain $\{p_\varepsilon(\cdot \mid z) : z \in U\}$ is realized by the probability measure induced by Lebesgue measure in the space of continuous maps through $\omega \mapsto F_\varepsilon(\omega, \cdot)$.

If $(\xi, z) \mapsto r_\varepsilon(\xi, z)$ is C^r, for some $r \ge 1$, then the map $(\xi, z) \mapsto \psi_{\varepsilon,z}(\xi)$ constructed in the proposition is C^r. Moreover, $\psi_{\varepsilon,z}$ is a diffeomorphism, because

$$\frac{\partial \omega_i}{\partial \xi_i} > 0 \quad \text{and} \quad \frac{\partial \omega_i}{\partial \xi_j} = 0 \quad \text{for all } 1 \le i < j \le d.$$

It follows that F_ε is C^r, and so ν_ε lives in the space of C^r maps.

For $\{\nu_\varepsilon : \varepsilon > 0\}$ to be a small perturbation of f in the space of C^r maps, we need $\operatorname{supp} \nu_\varepsilon \to f$ in the C^r topology, as $\varepsilon \to 0$. For $r = 0$ this comes for free, because $\sup \|F_\varepsilon(\omega, z) - f(z)\| \le \varepsilon$. For $r \ge 1$, the following additional conditions suffice (but they are not optimal):

$$c\varepsilon^{-d} \le r_\varepsilon \le C\varepsilon^{-d} \quad \text{and} \quad \|\partial^{p+q} r_\varepsilon / \partial\xi^p \partial z^q\| \le C\varepsilon^{-d-p} \quad \text{for} \quad 1 \le p + q \le r.$$

Let us check this, in the case $r = 1$.

Step 1: $\|\dfrac{\partial \psi_{\varepsilon,z}}{\partial z}\| \le C$.

Taking derivative with respect to z_j in the expression of ω_i, and then using $c\varepsilon^{-d} \le r_\varepsilon \le C\varepsilon^{-d}$ and $\|\partial r_\varepsilon / \partial z_j\| \le C\varepsilon^{-d}$, we find

$$|\frac{\partial \omega_i}{\partial z_j}| \le \frac{2C^2(\xi_i + \varepsilon)}{2c^2\varepsilon} \le \frac{2C^2}{c^2}$$

for every $1 \le j \le i \le d$. The claim follows immediately.

Step 2: $\|(\frac{\partial\psi_{\varepsilon,z}}{\partial\xi})^{-1}\| \le C$.

Since the derivative is given by a (lower) triangular matrix, to get this bound on the inverse it suffices to prove that

$$|(\frac{\partial\omega_i}{\partial\xi_i})^{-1}| \le C\varepsilon \quad\text{and}\quad |\frac{\partial\omega_i}{\partial\xi_j}| \le C\varepsilon^{-1} \tag{54}$$

for every $1 \le j < i \le d$. By direct derivation of ω_i, we get

$$\frac{\partial\omega_i}{\partial\xi_i} \mid \{\xi_i = \xi\} = \frac{\int_{-\varepsilon}^{\xi_1}\cdots\int_{-\varepsilon}^{\xi_{i-1}}\int_{-\varepsilon}^{\varepsilon}\cdots\int_{-\varepsilon}^{\varepsilon} r_\varepsilon \mid \{\xi_i = \xi\}}{\int_{-\varepsilon}^{\xi_1}\cdots\int_{-\varepsilon}^{\xi_{i-1}}\int_{-\varepsilon}^{\varepsilon}\cdots\int_{-\varepsilon}^{\varepsilon} r_\varepsilon}$$

where the integral in the denominator is with respect to $d\xi_1\cdots d\xi_{i-1}d\xi_{i+1}\cdots d\xi_d$, and the one on the denominator is with respect to $d\xi_i$ too. Using $c\varepsilon^{-d} \le r_\varepsilon \le C\varepsilon^{-d}$ we get that

$$\frac{\partial\omega_i}{\partial\xi_i} \ge \frac{c}{C(2\varepsilon)}$$

which proves the first half of (54). To prove the second half, take the derivative of ω_i with respect to ξ_j and then use the mean value theorem, to conclude that

$$|\frac{\partial\omega_i}{\partial\xi_j} \mid \{\xi_j = \xi\}| \le \frac{\int_{-\varepsilon}^{\xi_1}\cdots\int_{-\varepsilon}^{\xi_{j-1}}\int_{-\varepsilon}^{\xi_{j+1}}\cdots\int_{-\varepsilon}^{\xi_i}\int_{-\varepsilon}^{\varepsilon}\cdots\int_{-\varepsilon}^{\varepsilon} |\partial r_\varepsilon/\partial\xi_j| \mid \{\xi_j = \eta\}}{\int_{-\varepsilon}^{\xi_1}\cdots\int_{-\varepsilon}^{\xi_{j-1}}\int_{-\varepsilon}^{\xi_{j+1}}\cdots\int_{-\varepsilon}^{\xi_{i-1}}\int_{-\varepsilon}^{\varepsilon}\cdots\int_{-\varepsilon}^{\varepsilon} r_\varepsilon / \mid \{\xi_j = \eta\}}$$

for some $\eta \in (-\varepsilon, \xi)$. Using $r_\varepsilon \ge c\varepsilon^{-d}$ and $|\partial r_\varepsilon/\partial\xi_j| \le C\varepsilon^{d-1}$, we deduce

$$|\frac{\partial\omega_i}{\partial\xi_j}| \le \frac{(\xi+\varepsilon)C\varepsilon^{-d-1}}{(2\varepsilon)c\varepsilon^{-d}} \le \frac{C}{c\varepsilon}.$$

This completes the proof of (54) and step 2.

Step 3: Conclusion $\|\frac{\partial F_\varepsilon(\omega,\cdot)}{\partial z} - \frac{\partial f}{\partial z}\| \le C\varepsilon$.

Taking derivatives with respect to z in the equality $\psi_{\varepsilon,z} \circ \phi_{\varepsilon,z}(\omega) \equiv \omega$, and using the conclusions of the previous two steps, we obtain

$$\|\frac{\partial F_\varepsilon(\omega,\cdot)}{\partial z} - \frac{\partial f}{\partial z}\| = \|\frac{\partial\phi_{\varepsilon,z}}{\partial z}\| = \|\frac{\partial\phi_{\varepsilon,z}}{\partial z}(\frac{\partial\phi_{\varepsilon,z}}{\partial z})^{-1}\| \le C\varepsilon.$$

This completes the argument.

Problem D.11. Give examples of systems that are stochastically stable in the C^r but not in the C^s category, $s < r$.

D.5. **On shadowing versus stochastic stability.** The shadowing property of hyperbolic systems means that given every pseudo-orbit $(x_n)_n$, that is, such that $f(x_n)$ is close to x_{n+1} for every n, is uniformly approximated (shadowed) by a true orbit. See [70]. An important reformulation of this result for non-uniformly hyperbolic maps was due to Katok [170].

Shadowing is somewhat related to stochastic stability, since random orbits are pseudo-orbits, and proving stability involves showing that the behavior of (most) random orbits mimics that of (typical) true orbits. However, in the strong form stated above, the shadowing property is very rare outside the uniformly hyperbolic world. For instance, it fails already for piecewise smooth uniformly expanding maps of the interval [98]. Smooth counterexamples should occur among quadratic maps of the interval, and Hénon maps, and have been explicitly exhibited in the context of heterodimensional cycles [60, 353]. Namely, [60] constructs a C^1 open set of partially hyperbolic diffeomorphisms such that there is $\delta > 0$ such that, for any $\varepsilon > 0$ there are $\varepsilon-$pseudo orbits which are not δ-shadowed by any true orbit.

On the other hand, as mentioned in the text, many of these systems have been shown to be stochastically stable, under very general random perturbations. So uniform shadowing is by no means necessary, and a much weaker shadowing property might suffice for stochastic stability:

Problem D.12. Formulate a general criterium for stochastic stability in terms of a shadowing property.

Piecewise smooth uniformly expanding maps of the interval are a natural context to start investigating this problem.

D.6. **Random perturbations of flows.** The main model is *diffusion*: the random orbits are the solutions of stochastic differential equations (for simplicity, pretend $M = \mathbb{R}^d$)

$$dz_t = X(z_t)\,dt + \epsilon A(z_t)\,dw_t \tag{55}$$

where X is the deterministic vector field, $A(\cdot)$ is a matrix-valued function, and dw_t is the standard Brownian motion. See Kifer [179], and also Oksendal [250] or Kunita [184] for general information about stochastic differential equations. A solution of (55) with initial condition $z \in M$ is a stochastic process

$$\mathbb{R} \times \Omega \to M, \quad (t, \omega) \mapsto f_\omega^t(z)$$

on an (abstract) probability space Ω, satisfying

$$f_\omega^t(z) = z + \int_0^t X(f_\omega^s(z))\,ds + \varepsilon \int_0^t A(f_\omega^s(z))\,dw_s$$

where the last term is a stochastic integral in the sense of Ito. Under reasonable conditions on $X(\cdot)$ and $A(\cdot)$, there exists exactly one solution *with continuous paths,* that is, such that

$$[0, +\infty) \ni t \mapsto f_\omega^t(z) \tag{56}$$

is continuous for almost $\omega \in \Omega$. See e.g. Theorem 5.5 in [250].

Stochastic stability means that, given any continuous function φ with compact support and any $\delta > 0$, if ε is small enough then

$$\left| \lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi(f_\omega^s(z))\,ds - \int \varphi\,d\mu_{SRB} \right| < \delta$$

for Lebesgue almost every z in the relevant domain U (see the next paragraph) and almost every ω. Here μ_{SRB} is the SRB measure of the deterministic vector field X inside U.

An important difference with respect to the previous models, where noise was supported in small balls, is that diffusion tends to mingle all the attractors together, because the solutions (56) tend to spread over the whole ambient M for all $t > 0$. For instance, if A is everywhere nonsingular then the stochastic flow has a unique stationary measure μ_ε, that is, such that

$$\mu_\varepsilon(E) = \int \mu_\varepsilon\big((f_\omega^t)^{-1}(E)\big)\,d\omega$$

for all $t > 0$ and every measurable set E. Thus, one should use a global notion of stability as in Remark D.6. However, it makes sense to speak of stochastic perturbations restricted to some invariant set U of X, if one considers $A(\cdot)$ supported inside U.

Kifer [179] proved that hyperbolic attractors of flows are stochastically stable, and so are the geometric Lorenz attractors. This was extended by Metzger [218] for the contracting Lorenz-like attractors introduced by Rovella [299].

References

[1] F. Abdenur. Generic robustness of spectral decompositions. *Annales Scient. E.N.S.*

[2] R. Abraham and S. Smale. Nongenericity of ω-stability. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*. Amer. Math. Soc., 1970.

[3] V. S. Afraimovich, V. V. Bykov, and L. P. Shil'nikov. On the appearence and structure of the Lorenz attractor. *Dokl. Acad. Sci. USSR*, 234:336–339, 1977.

[4] J. F. Alves. SRB measures for non-hyperbolic systems with multidimensional expansion. *Ann. Sci. École Norm. Sup.*, 33:1–32, 2000.

[5] J. F. Alves and V. Araújo. Stochastic stability for robust classes of non-uniformly expanding maps. *Astérisque*. To appear.

[6] J. F. Alves, V. Araújo, and B. Saussol. On the uniform hyperbolicity of some nonuniformly hyperbolic systems. *Proc. Amer. Math. Soc.*, 131:1303–1309 (electronic), 2003.

[7] J. F. Alves, C. Bonatti, and M. Viana. SRB measures for partially hyperbolic systems whose central direction is mostly expanding. *Invent. Math.*, 140:351–398, 2000.

[8] J. F. Alves, S. Luzzatto, and V. Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. Preprint University of Porto, 2002.

[9] J. F. Alves and M. Viana. Statistical stability for a robust class of multidimensional non-hyperbolic attractors. *Ergod. Th. & Dynam. Sys.*, 21, 2001.

[10] A. Andronov and L. Pontryagin. Systèmes grossiers. *Dokl. Akad. Nauk. USSR*, 14:247–251, 1937.

[11] D. V. Anosov. Geodesic flows on closed Riemannian manifolds of negative curvature. *Proc. Steklov Math. Inst.*, 90:1–235, 1967.

[12] N. Aoki. The set of axiom a diffeomorphisms with no cycles. *Bol. Soc. Bras. Mat., Nova Sér.*, 23 (1–2):21–25, 1992.

[13] V. Araújo. Attractors and time averages for random maps. *Annales de l'Inst. Henri Poincaré - Analyse Non-linéaire*, 17:307–369, 2000.

[14] V. Araújo. Infinitely many stochastically stable attractors. *Nonlinearity*, 14:583–596, 2001.

[15] A. Arbieto and J. Bochi. L^p-generic cocycles have one-point Lyapunov spectrum. *Stochastics and Dynamics*. To appear. Preprint www.preprint.impa.br 2002.

[16] M.-C. Arnaud. Le "closing lemma" en topologie C^1. *Mém. Soc. Math. Fr.*, 74:vi+120, 1998.

[17] M.-C. Arnaud. Difféomorphismes symplectiques de classe C^1 en dimension 4. *C. R. Acad. Sci. Paris Sér. I Math.*, 331:1001–1004, 2000.

[18] M.-C. Arnaud. Création de connexions en topologie C^1. *Ergod. Th. & Dynam. Systems*, 21:339–381, 2001.

[19] M.-C. Arnaud. The generic C^1 symplectic diffeomorphisms of symplectic 4-dimensional manifolds are hyperbolic, partially hyperbolic, or have a completely elliptic point. *Ergod. Th. & Dynam. Sys.*, 22:1621–1639, 2002.

[20] A. Arnéodo, P. Coullet, and C. Tresser. Possible new strange attractors with spiral structure. *Comm. Math. Phys.*, 79:573–579, 1981.

[21] L. Arnold. *Random dynamical systems*. Springer-Verlag, 1998.

[22] L. Arnold and N. D. Cong. On the simplicity of the Lyapunov spectrum of products of random matrices. *Ergod. Th. & Dynam. Sys.*, 17(5):1005–1025, 1997.

[23] V. I. Arnold. Small denominators I: On the mapping of a circle to itself. *Izv. Akad. Nauk. Math. Series*, 25:21–86, 1961. Transl. A.M.S. 46 (1965).

[24] A. Arroyo and F. Rodriguez-Hertz. Homoclinic bifurcations and uniform hyperbolicity for three-dimensional flows. Preprint IMPA, 2002.

[25] A. Avila. *Bifurcations of unimodal maps: the topological and metric picture*. PhD thesis, IMPA, 2001.

[26] A. Avila, M. Lyubich, and W. de Melo. Regular or stochastic dynamics in real analytic families of unimodal maps. In preparation.
[27] A. Avila and C. G. Moreira. Statistical properties of unimodal maps: periodic orbits and pathological foliations. In preparation.
[28] A. Avila and C. G. Moreira. Statistical properties of unimodal maps: smooth families with negative Schwarzian derivative. *Astérisque.*
[29] A. Avila and C. G. Moreira. Statistical properties of unimodal maps: the quadratic family. *Annals of Math.*
[30] V. Baladi. *Positive transfer operators and decay of correlations.* World Scientific Publishing Co. Inc., 2000.
[31] V. Baladi, C. Bonatti, and B. Schmitt. Abnormal escape rates from nonuniformly hyperbolic sets. *Ergod. Th. & Dynam. Sys.*, 19:1111–1125, 1999.
[32] V. Baladi and S. Gouezel. Seminar (by Baladi) at Paris-Jussieu, November 2002.
[33] V. Baladi and M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. *Ann. Sci. École Norm. Sup.*, 29:483–517, 1996.
[34] R. Bamón. Singular cycles of vector field on regular parts of the boundary of Morse-Smale systems. *Bull. Braz. Mat. Soc.*, 30 (2):223–246, 1999.
[35] R. Bamón, R. Labarca, R. Mañé, and M.J. Pacifico. The explosion of singular cycles. *Publ. Math. IHES*, 78:207–232, 1993.
[36] A. Baraviera and C. Bonatti. Removing zero central Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*
[37] L. Barreira and Ya. Pesin. Lectures on Lyapunov exponents and smooth ergodic theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 3–106. Amer. Math. Soc., 2001. Appendix A by M. Brin and Appendix B by D. Dolgopyat, H. Hu and Ya. Pesin.
[38] L. Barreira and Ya. Pesin. *Lyapunov exponents and smooth ergodic theory*, volume 23 of *Univ. Lecture Series.* Amer. Math. Soc., 2002.
[39] L. Barreira, Ya. Pesin, and J. Schmeling. Dimension and product structure of hyperbolic measures. *Ann. of Math.*, 149:755–783, 1999.
[40] M. Benedicks and L. Carleson. On iterations of $1-ax^2$ on $(-1,1)$. *Annals of Math.*, 122:1–25, 1985.
[41] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Annals of Math.*, 133:73–169, 1991.
[42] M. Benedicks and M. Viana. Random perturbations and statistical properties of Hénon-like maps. Preprint www.preprint.impa.br 2002.
[43] M. Benedicks and M. Viana. Solution of the basin problem for Hénon-like attractors. *Invent. Math.*, 143:375–434, 2001.
[44] M. Benedicks and L.-S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergod. Th. & Dynam. Sys.*, 12:13–37, 1992.
[45] M. Benedicks and L.-S. Young. SBR-measures for certain Hénon maps. *Invent. Math.*, 112:541–576, 1993.
[46] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Astérisque*, 261:13–56, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).
[47] D. Bennequin. Seminar at Jussieu-Paris, April 2002.
[48] A. M. Blokh and M. Yu. Lyubich. Ergodicity of transitive maps of the interval. *Ukrainian Math. J.*, 41:985–988, 1989.
[49] J. Bochi. Genericity of zero Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*, 22:1667–1696, 2002.
[50] J. Bochi and M. Viana. Lyapunov exponents: How often are dynamical systems hyperbolic ? Preprint www.preprint.impa.br 2002.
[51] J. Bochi and M. Viana. The Lyapunov exponents of generic volume preserving and symplectic systems. Preprint www.preprint.impa.br 2002.
[52] J. Bochi and M. Viana. Uniform (projective) hyperbolicity or no hyperbolicity: a dichotomy for generic conservative maps. *Annales de l'Inst. Henri Poincaré - Analyse Non-linéaire*, 19:113–123, 2002.
[53] C. Bonatti and L. J. Díaz. On maximal transitive sets of generic diffeomorphisms. *Publ. Math. IHES.*

[54] C. Bonatti and S. Crovisier. Recurrence and genericity. *Comptes Rendus Acad. Sci. Paris.* To appear.

[55] C. Bonatti and S. Crovisier. Recurrence et généricité. Preprint Dijon, 2003.

[56] C. Bonatti and L. J. Díaz. Nonhyperbolic transitive diffeomorphisms. *Annals of Math.*, 143:357–396, 1996.

[57] C. Bonatti and L. J. Díaz. Connexions heterocliniques et genericité d'une infinité de puits ou de sources. *Ann. Sci. École Norm. Sup.*, 32:135–150, 1999.

[58] C. Bonatti, L. J. Díaz, and E. Pujals. A C^1-generic dichotomy for diffeomorphisms: weak forms of hyperbolicity or infinitely many sinks or sources. *Annals of Math.* To appear.

[59] C. Bonatti, L. J. Díaz, E. Pujals, and J. Rocha. Robust transitivity and heterodimensional cycles. *Astérisque.* To appear.

[60] C. Bonatti, L. J. Díaz, and G. Turcat. Pas de shadowing lemma pour les dynamiques partiellement hyperboliques. *C. R. Acad. Sci. Paris Sér. I Math.*, 330:587–592, 2000.

[61] C. Bonatti, L. J. Díaz, and R. Ures. Minimality of strong stable and unstable foliations for partially hyperbolic diffeomorphisms. *Bull. Inst. Math. Jussieu.*

[62] C. Bonatti, L. J. Díaz, and M. Viana. Discontinuity of the Hausdorff dimension of hyperbolic sets. *C. R. Acad. Sci. Paris Sér. I Math.*, 320:713–718, 1995.

[63] C. Bonatti, X. Gomez-Mont, and M. Viana. Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices. *Annales de l'Inst. Henri Poincaré - Analyse Non-linéaire.* To appear. Preprint www.preprint.impa.br 2001.

[64] C. Bonatti, C. Matheus, M. Viana, and A. Wilkinson. Abundance of stable ergodicity. Preprint www.preprint.impa.br 2002.

[65] C. Bonatti, A. Pumariño, and M. Viana. Lorenz attractors with arbitrary expanding dimension. *C. R. Acad. Sci. Paris Sér. I Math.*, 325:883–888, 1997.

[66] C. Bonatti and M. Viana. Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergod. Th. & Dynam. Sys.*

[67] C. Bonatti and M. Viana. SRB measures for partially hyperbolic systems whose central direction is mostly contracting. *Israel J. Math.*, 115:157–193, 2000.

[68] C. Bonatti and A. Wilkinson. Transitive partially hyperbolic diffeomorphisms on 3-manifolds. In preparation.

[69] R. Bowen. Periodic orbits for hyperbolic flows. *Amer. J. Math.*, 94:1–30, 1972.

[70] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[71] R. Bowen. A horseshoe with positive measure. *Invent. Math.*, 29:203–204, 1975.

[72] R. Bowen. Mixing Anosov flows. *Topology*, 15:77–79, 1977.

[73] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29:181–202, 1975.

[74] M. Brin. Topological transitivity of a certain class of dynamical systems, and flows of frames on manifolds of negative curvature. *Funkcional. Anal. i Priložen.*, 9:9–19, 1975.

[75] M. Brin. Bernoulli diffeomorphisms with $n-1$ nonzero exponents. *Ergod. Th. & Dynam. Sys.*, 1:1–7, 1981.

[76] M. Brin, D. Burago, and S. Ivanov. Lecture (by Burago) at the John Mather Conference, Princeton, October 2002.

[77] M. Brin, J. Feldman, and A. Katok. Bernoulli diffeomorphisms and group extensions of dynamical systems with nonzero characteristic exponents. *Annals of Math.*, 113:159–179, 1981.

[78] M. Brin and Ya. Pesin. Partially hyperbolic dynamical systems. *Izv. Acad. Nauk. SSSR*, 1:177–212, 1974.

[79] K. Burns, D. Dolgopyat, and Ya. Pesin. Partial hyperbolicity, Lyapunov exponents, and stable ergodicity. Preprint Northwestern University, 2002.

[80] K. Burns, C. Pugh, M. Shub, and A. Wilkinson. Recent results about stable ergodicity. In *Smooth ergodic theory and its applications (Seattle WA, 1999)*, volume 69 of *Procs. Symp. Pure Math.*, pages 327–366. Amer. Math. Soc., 2001.

[81] K. Burns, C. Pugh, and A. Wilkinson. Stable ergodicity and Anosov flows. *Topology*, 39:149–159, 2000.

[82] K. Burns and A. Wilkinson. Stable ergodicity of skew products. *Ann. Sci. École Norm. Sup. (4)*, 32:859–889, 1999.

[83] G. Buzzard. Infinitely many periodic attractors for holomorphic maps. *Annals of Math.*, 145:389–417, 1997.

[84] J. Buzzi, O. Sester, and M. Tsujii. Weakly expanding skew-products of quadratic maps. *Ergod. Th. & Dynam. Sys.*

[85] C. Camacho and A. Lins Neto. *Geometric theory of foliations.* Birkhäuser Boston Inc., 1985. Translated from the Portuguese by Sue E. Goodman.

[86] Y. Cao. Lyapunov exponents and uniform hyperbolicity. Preprint 2002.

[87] Y. Cao and J. M. Mao. The non-wandering set of some Hénon maps. *Chaos, Solitons & Fractals*, 11:2045–2053, 2000.

[88] C. Carballo and C. Morales. Homoclinic classes and finitude of attractors for vector fields on n-manifolds. Preprint 2001.

[89] C. Carballo, C. Morales, and M. J. Pacifico. Homoclinic classes for generic C^1 vector fields. *Ergod. Th. & Dynam. Sys.* To appear.

[90] A. A. Castro. Backward inducing and exponential decay of correlations for partially hyperbolic attractors with mostly contracting central direction. *Israel J. Math.* To appear.

[91] T. Cherry. Analytic quasi-periodic curves of discontinuous type in a torus. *Proc. London Math. Soc.*, 44, 1938.

[92] P. Collet and J. P. Eckmann. On the abundance of aperiodic behaviour for maps of the interval. *Comm. Math. Phys.*, 73:115–160, 1980.

[93] P. Collet and J. P. Eckmann. Positive Lyapunov exponents and absolute continuity for maps of the interval. *Ergod. Th. & Dynam. Sys.*, 3:13–46, 1983.

[94] E. Colli. Infinitely many coexisting strange attractors. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15:539–579, 1998.

[95] C. Conley. *Isolated invariant sets and the Morse index*, volume 38 of *CBMS Regional Conference Series in Mathematics.* American Mathematical Society, 1978.

[96] M. J. Costa. Saddle-node horseshoes giving rise to global hénon-like attractors. *Anais Acad. Bras. Ciências*, 70:393–400, 1998.

[97] P. Coullet and C. Tresser. Itérations d'endomorphims et groupe de renormalization. *C. R. Acad. Sci. Paris*, 287, Série I:577–580, 1978.

[98] E. Coven, I. Kan, and J. Yorke. Pseudo-orbit shadowing in the family of tent maps. *Trans. Amer. Math. Soc.*, 308:227–241, 1988.

[99] B. Cowieson and L.-S. Young. Lecture (by Young) at the Charles Pugh Conference, Evanston, May 2001.

[100] S. Crovisier. Saddle-node bifurcations for hyperbolic sets. Preprint Orsay 2000.

[101] S. Crovisier. Une remarque sur les ensembles hyperboliques localement maximaux. *C. R. Acad. Sci. Paris*, 334:401–404, 2002.

[102] W. de Melo. Structural stability of diffeomorphisms on two-manifolds. *Invent. Math.*, 21:233–246, 1973.

[103] W. de Melo and S. van Strien. *One-dimensional dynamics.* Springer Verlag, 1993.

[104] B. Deng. The Silnikov problem, exponential expansion, strong λ-lemma, C^1-linearizations and homoclinic orbits. *J. Dynam. Diff. Eq.*, 3:177–244, 1990.

[105] B. Deng. Homoclinic twisting bifurcations and cusp horseshoe maps. *J. Dynam. Diff. Eq.*, 5:417–467, 1993.

[106] L. J. Díaz. Persistence of cycles and nonhyperbolic dynamics at heteroclinic bifurcations. *Nonlinearity*, 8:693–713, 1995.

[107] L. J. Díaz. Robust nonhyperbolic dynamics and heterodimensional cycles. *Ergod. Th. & Dynam. Sys.*, 15:291–315, 1995.

[108] L. J. Díaz, E. Pujals, and R. Ures. Partial hyperbolicity and robust transitivity. *Acta Math.*, 183:1–43, 1999.

[109] L. J. Díaz and J. Rocha. Heterodimensional cycles, partial hyperbolicity, and limit dynamics. Preprint 2000.

[110] L. J. Díaz and J. Rocha. Nonconnected heteroclinic cycles: bifurcation and stability. *Nonlinearity*, 5:1315–1341, 1992.

[111] L. J. Díaz and J. Rocha. Largue measure of hyperbolic dynamics when unfolding heterodimensional cycles. *Nonlinearity*, 10:857–884, 1997.

[112] L. J. Díaz and J. Rocha. Partial hyperbolicity and transitive dynamics generated by heteroclinic cycles. *Ergod. Th. & Dynam. Sys.*, 21:25–76, 2001.

[113] L. J. Díaz, J. Rocha, and M. Viana. Strange attractors in saddle-node cycles: prevalence and globality. *Invent. Math.*, 125:37–74, 1996.

[114] L. J. Díaz and B. Santoro. Collision, explosion, and collapse of homoclinic classes. Master Thesis PUC-Rio 2001 and in preparation.

[115] L. J. Díaz and R. Ures. Persistent homoclinic tangencies at the unfolding of cycles. *Ann. Inst. H. Poincaré, Anal. Non-linéaire*, 11:643–659, 1996.

[116] C. I. Doering. Persistently transitive vector fields on three-dimensional manifolds. In *Procs. on Dynamical Systems and Bifurcation Theory*, volume 160, pages 59–89. Pitman, 1987.

[117] D. Dolgopyat. Limit theorems for partially hyperbolic diffeomorphisms. Preprint 2000.

[118] D. Dolgopyat. On differentiability of SRB states. Preprint 2000.

[119] D. Dolgopyat. On dynamics of mostly contracting diffeomorphisms. *Comm. Math. Phys*, 213:181–201, 2000.

[120] D. Dolgopyat, H.Hu, and Ya. Pesin. An example of a smooth hyperbolic measure with countably many ergodic components. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., 2001.

[121] D. Dolgopyat and Ya. Pesin. Every manifold supports a completely hyperbolic diffeomorphism. *Ergod. Th. & Dynam. Sys.*, 22:409–437, 2002.

[122] D. Dolgopyat and A. Wilkinson. Stable accessibility is C^1 dense. *Astérisque.* To appear.

[123] P. Duarte. Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non. Linéaire*, 11:359–409, 1994.

[124] P. Duarte. Abundance of elliptic isles at conservative bifurcations. *Dynam. Stability Systems*, 14:339–356, 1999.

[125] F. Dumortier, H. Kokubu, and H. Oka. A degenerate singularity generating geometric Lorenz attractors. *Ergodic Theory Dyn. Syst*, 15(5):833–856, 1995.

[126] J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57:617–656, 1985.

[127] K. Falconer. *Fractal geometry.* John Wiley & Sons Ltd., 1990. Mathematical foundations and applications.

[128] A. Fathi, M. Herman, and J.-C. Yoccoz. A proof of Pesin's stable manifold theorem. In *Geometric dynamics (Rio de Janeiro 1981)*, volume 1007 of *Lect. Notes in Math.*, pages 177–215. Springer Verlag, 1983.

[129] B. Fayad. Lecture at the Colloque Herman, IHP-Paris, November 2002.

[130] M. Feigenbaum. Qualitative universality for a class of nonlinear transformations. *J. Stat. Phys.*, 19:25–52, 1978.

[131] J. Franks. Necessary conditions for the stability of diffeomorphisms. *Trans. A.M.S.*, 158:301–308, 1971.

[132] G. Froyland. *Estimating physical invariant measures and space averages of dynamical systems indicators.* PhD thesis, Univ. Western Australia.

[133] H. Furstenberg. Non-commuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963.

[134] H. Furstenberg. Boundary theory and stochastic processes on homogeneous spaces. In *Harmonic analysis in homogeneous spaces*, volume XXVI of *Proc. Sympos. Pure Math. (Williamstown MA, 1972)*, pages 193–229. Amer. Math. Soc., 1973.

[135] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.

[136] S. Gan and L. Wen. Heteroclinic cycles and homoclinic closures for generic diffeomorphisms. Preprint Peking University 2001.

[137] P. Glendinning and C. Sparrow. T-points: A codimension two heteroclinic bifurcation. *Jour. Stat. Phys.*, 43:479–488, 1986.

[138] I. Ya. Gold'sheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Uspekhi Mat. Nauk.*, 44:13–60, 1989.

[139] S. V. Gonchenko, L. P. Shil'nikov, and D. V. Turaev. Dynamical phenomena in systems with structurally unstable Poincaré homoclinic orbits. *Chaos*, 6:15–31, 1996.

[140] J. Graczyk and G. Swiatek. Generic hyperbolicity in the logistic family. *Annals of Math.*, 146:1–52, 1997.

[141] J. Graczyk and G. Swiatek. Smooth unimodal maps in the 1990s. *Ergod. Th. & Dynam. Sys.*, 19:263–287, 1999.

[142] M. Grayson, C. Pugh, and M. Shub. Stably ergodic diffeomorphisms. *Annals of Math.*, 140:295–329, 1994.
[143] J. Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Publ. Math. IHES*, 50:59–72, 1979.
[144] Y. Guivarc'h and A. Raugi. Products of random matrices : convergence theorems. *Contemp. Math.*, 50:31–54, 1986.
[145] C. Gutierrez. A counter-example to a C^2 closing lemma. *Ergod. Th. & Dynam. Sys.*, 7:509–530, 1987.
[146] M. Hall. On the sums and products of continued fractions. *Annals of Math.*, 48:966–993, 1947.
[147] P. Halmos. On automorphisms of compact groups. *Bulletin Amer. Math. Soc.*, 49:619–624, 1943.
[148] S. Hayashi. Hyperbolicity and homoclinic bifurcations generating nonhyperbolic dynamics. Preprint 2002.
[149] S. Hayashi. Diffeomorphisms in $\mathcal{F}^1(m)$ satisfy Axiom A. *Ergod. Th. & Dynam. Sys.*, 12 (2):233–253, 1992.
[150] S. Hayashi. Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjectures for flows. *Annals of Math.*, 145:81–137, 1997.
[151] S. Hayashi. A C^1 make or break lemma. *Bol. Soc. Brasil. Mat.*, 31:337–350, 2000.
[152] M. Hénon. A two dimensional mapping with a strange attractor. *Comm. Math. Phys.*, 50:69–77, 1976.
[153] M. Hénon and Y. Pomeau. Two strange attractors with a simple structure. In *Turbulence and Navier-Stokes equations*, volume 565, pages 29–68. Springer Verlag, 1976.
[154] M. Herman. Lecture at the International Conference on Dynamical Systems, Edinburgh, July 2000.
[155] M. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle a des rotations. *Publ. Math. IHES*, 103–104, 1979.
[156] M. Herman. Différentiabilité optimale et contre-exemples à la fermeture en topologie C^∞ des orbites récurrentes de flots hamiltoniens. *C. R. Acad. Sci. Paris Sér. I Math.*, 313:49–51, 1991.
[157] M. Herman. Exemples de flots hamiltoniens dont aucune perturbation en topologie C^∞ n'a d'orbites périodiques sur un ouvert de surfaces d'énergies. *C. R. Acad. Sci. Paris Sér. I Math.*, 312:989–994, 1991.
[158] M. Hirsch, C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lect. Notes in Math.* Springer Verlag, 1977.
[159] A. J. Homburg, H. Kokubu, and M. Krupa. The cusp horseshoe and its bifurcations in the unfolding of an inclination-flip homoclinic orbit. *Ergod. Th. & Dynam. Sys.*, 14(4):667–693, 1994.
[160] E. Hopf. Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. *Ber. Verh. Sächs. Akad. Wiss. Leipzig*, 91:261–304, 1939.
[161] B. Hunt and V. Kaloshin. A stretched exponential bound on the rate of growth of the number of periodic points for prevalent diffeomorphisms I-II. *Electron. Res. Announc. A. M. S.*, 7:17–27 and 28–36, 2001.
[162] M. Hurley. Attractors: persistence and density of their basins. *Trans. AMS*, 269:247–271, 1982.
[163] M. Hurley. Properties of attractors of generic homeomorphisms. *Ergod. Th. & Dynam. Systems*, 16:1297–1310, 1996.
[164] Mike Hurley. Generic homeomorphisms have no smallest attractors. *Proc. Amer. Math. Soc.*, 123:1277–1280, 1995.
[165] M. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.*, 81:39–88, 1981.
[166] J. Rocha J.M. Gambaudo. Maps of the sphere at the boundary of chaos. *Nonlinearity*, 7:1251–1259, 1999.
[167] V. Kaloshin. Generic diffeomorphisms with superexponential growth of the number of periodic orbits. *Comm. Math. Phys.*, 211:253–271, 2000.
[168] I. Kan. Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin. *Bull. Amer. Math. Soc.*, 31:68–74, 1994.
[169] A. Katok. Bernoulli diffeomorphisms. *Annals of Math.*, 110:529–547, 1979.

[170] A. Katok. Lyapunov exponents, entropy and periodic points of diffeomorphisms. *Publ. Math. IHES*, 51:137–173, 1980.

[171] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Cambridge University Press, 1995.

[172] A. Katok and Yu. Kifer. Random perturbations of transformations of an interval. *J. Analyse Math.*, 47:193–237, 1986.

[173] A. Katok and A. Kononenko. Cocycle's stability for partially hyperbolic systems. *Math. Res. Lett.*, 3:191–210, 1996.

[174] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94:313–333, 1982.

[175] G. Keller. Markov extensions, zeta functions, and Fredholm determinants for piecewise invertible dynamical systems. *Trans. Amer. Math. Soc.*, 314:433–497, 1989.

[176] G. Keller and T. Nowicki. Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. *Comm. Math. Phys.*, 149:31–69, 1992.

[177] Yu. Kifer. On small random perturbations of some smooth dynamical systems. *Math. USSR Izv.*, 8:1083–1107, 1974.

[178] Yu. Kifer. *Ergodic theory of random perturbations.* Birkhäuser, 1986.

[179] Yu. Kifer. *Random perturbations of dynamical systems.* Birkhäuser, 1988.

[180] J. Kingman. The ergodic theorem of subadditive stochastic processes. *J. Royal Statist. Soc.*, 30:499–510, 1968.

[181] O. Kozlovski. *Structural stability in one-dimensional dynamics.* PhD thesis, Univ. Amsterdam, 1998.

[182] O. Kozlovski. Getting rid of the negative Schwarzian derivative condition. *Ann. of Math.*, 152:743–762, 2000.

[183] O. Kozlovski. Axiom A maps are dense in the space of unimodal maps in the C^k topology. *Ann. of Math.*, 157:1–43, 2003.

[184] H. Kunita. *Stochastic flows and stochastic differential equations.* Cambridge University Press, 1990.

[185] R. Labarca. Bifurcation of contracting singular cycles. *Ann. Sci. Éc. Norm. Supér.*, 28:705–745, 1995.

[186] R. Labarca and B. San Martín. Prevalence of hyperbolicity for complex singular cycles. *Bull. Braz. Mat. Soc.*, 28:343–362, 1997.

[187] R. Labarca and M.J. Pacifico. Stability of singular horseshoes. *Topology*, 25:337–352, 1986.

[188] F. Ledrappier. Propriétés ergodiques des mesures de Sinaï. *Publ. Math. I.H.E.S.*, 59:163–188, 1984.

[189] F. Ledrappier. Quelques propriétés des exposants caractéristiques. *Lect. Notes in Math.*, 1097:305–396, 1984.

[190] F. Ledrappier and J.-M. Strelcyn. A proof of the estimation from below in pesin's entropy formula. *Ergod. Th & Dynam. Sys*, 2:203–219, 1982.

[191] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. *Ann. of Math.*, 122:509–539, 1985.

[192] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math.*, 122:540–574, 1985.

[193] S.-T. Liao. Obstruction sets. II. *Beijing Daxue Xuebao*, 2:1–36, 1981.

[194] S.T. Liao. On the stability conjecture. *Chinese Annals of Math*, 1:9–30, 1980.

[195] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmosph. Sci.*, 20:130–141, 1963.

[196] S. Luzzatto and W. Tucker. Non-uniformly expanding dynamics in maps with singularities and criticalities. *Inst. Hautes Études Sci. Publ. Math.*, 89:179–226 (2000), 1999.

[197] S. Luzzatto and M. Viana. Lorenz-like attractors without invariant foliations. In preparation.

[198] S. Luzzatto and M. Viana. Parameter exclusions in Hénon-like systems. Preprint www.preprint.impa.br 2003.

[199] S. Luzzatto and M. Viana. Positive Lyapunov exponents for Lorenz-like families with criticalities. *Astérisque*, 261:201–237, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[200] M. Lyubich. Almost every real quadratic map is either regular or stochastic. To appear Annals of Math.

[201] M. Lyubich. Dynamics of quadratic maps I-II. *Acta Math.*, 178:185–247, 248–297, 1997.

[202] M. Lyubich. Feigenbaum-Coullet-Tresser, universality, and Milnor's hairiness conjecture. *Annals of Math.*, 149:319–420, 1999.

[203] M. Lyubich. The quadratic family as a qualitatively solvable model of chaos. *Notices Amer. Math. Soc.*, 47:1042–1052, 2000.

[204] M. Lyubich. Dynamics of quadratic maps III: Parapuzzle and SRB measure. *Astérisque*, 261:173–200, 2001. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[205] R. Mañé. Contributions to the stability conjecture. *Topology*, 17:383–396, 1978.

[206] R. Mañé. An ergodic closing lemma. *Annals of Math.*, 116:503–540, 1982.

[207] R. Mañé. Oseledec's theorem from the generic viewpoint. In *Procs. International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1269–1276, Warsaw, 1984. PWN Publ.

[208] R. Mañé. Hyperbolicity, sinks and measure in one-dimensional dynamics. *Comm. Math. Phys.*, 100:495–524, 1985.

[209] R. Mañé. *Ergodic theory and differentiable dynamics.* Springer Verlag, 1987.

[210] R. Mañé. On the creation of homoclinic points. *Publ. Math. I.H.E.S.*, 66:139–159, 1988.

[211] R. Mañé. A proof of the C^1 stability conjecture. *Publ. Math. I.H.E.S.*, 66:161–210, 1988.

[212] R. Mañé. The Lyapunov exponents of generic area preserving diffeomorphisms. In *International Conference on Dynamical Systems (Montevideo, 1995)*, pages 110–119. Longman, 1996.

[213] A. Manning and H. McCluskey. Hausdorff dimension of horseshoes. *Ergod. Th. & Dynam. Sys.*, 3:251–260, 1983.

[214] J. M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc.*, 4:257–302, 1954.

[215] M. Martens and T. Nowicki. Invariant measures for lebesgue typical quadratic maps. *Astérisque*, 261:239–252, 2001. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[216] J. Mather. Commutators of diffeomorphisms. *Comm. Math. Helv.*, 48:195–233, 1973.

[217] C. McMullen. *Renormalization and 3-manifolds which fiber over the circle*, volume 142 of *Annals of Math. Studies.* Princeton Univ. Press, 1996.

[218] R. Metzger. Stochastic stability for contracting Lorenz maps and flows. *Comm. Math. Phys.*, 212:277–296, 2000.

[219] J. Milnor and W. Thurston. On iterated maps of the interval: I, II. In *Dynamical Systems - Maryland 1986-87*, volume 1342, pages 465–563. Springer Verlag, 1988.

[220] M. Misiurewicz. Absolutely continuous invariant measures for certain maps of the interval. *Publ. Math. IHES*, 53:17–51, 1981.

[221] Y. Mitsumatsu. Foliations and contact structures on 3-manifolds. In *Geometry and dynamics (Warsaw 2000)*, pages 75–125. World Scientific Publ., 2002.

[222] L. Mora and C. Martin-Ribas. A complement to Hayashi's connecting lemma. Preprint 2001.

[223] L. Mora and M. Viana. Abundance of strange attractors. *Acta Math.*, 171:1–71, 1993.

[224] C. Morales and M. J. Pacifico. New singular strange attractors arising from hyperbolic flows. Preprint.

[225] C. Morales and M. J. Pacifico. Degenerated singular cycles of inclination-flip type. *Ann. Sci. École Norm. Sup.*, 31:1–16, 1998.

[226] C. Morales and M. J. Pacifico. Mixing attractors for 3-flows. *Nonlinearity*, 14:359–378, 2001.

[227] C. Morales, M. J. Pacifico, and E. Pujals. Robust transitive singular sets for 3-flows are partially hyperbolic attractors and repellers. *Annals of Math.*

[228] C. Morales, M. J. Pacifico, and E. Pujals. Global attractors from the explosion of singular cycles. *C. R. Acad. Sci. Paris Sér. I Math.*, 325:1317–1322, 1997.

[229] C. Morales, M. J. Pacifico, and E. Pujals. On C^1 robust singular transitive sets for three-dimensional flows. *C. R. Acad. Sci. Paris Sér. I Math.*, 326:81–86, 1998.

[230] C. Morales, M. J. Pacifico, and E. Pujals. Singular hyperbolic systems. *Proc. Amer. Math. Soc.*, 127:3393–3401, 1999.

[231] C. Morales, M. J. Pacifico, and E. Pujals. Strange attractors across the boundary of hyperbolic systems. *Comm. Math. Phys.*, 211:527–558, 2000.

[232] C. Morales and E. Pujals. Singular strange attractors on the boundary of Morse-Smale systems. *Ann. Sci. École Norm. Sup.*, 30:693–717, 1997.

[233] C. G. Moreira. Geometric properties of the Markov and Lagrange spectra. Preprint www.preprint.impa.br, 2002.

[234] C. G. Moreira. Stable intersections of Cantor sets and homoclinic bifurcations. *Ann. Inst. H. Poincaré Anal. Non. Linéaire*, 13:741–781, 1996.

[235] C. G. Moreira, J. Palis, and M. Viana. Homoclinic tangencies and fractal invariants in arbitrary dimension. *C. R. Acad. Sci. Paris*, 333:475–480, 2001.

[236] C. G. Moreira and J.-C. Yoccoz. Tangences homocliniques stables pour les ensembles hyperboliques de grande dimension fractale. In preparation.

[237] C. G. Moreira and J.-C. Yoccoz. Stable intersections of regular Cantor sets with large Hausdorff dimensions. *Ann. of Math.*, 154:45–96, 2001. Corrigendum in Annals of Math. 154:527, 2001.

[238] N. Muniz. *Hénon-like families in higher dimensions: SRB measures and basin problem.* PhD thesis, IMPA, 2000.

[239] V. Naudot. Strange attractors in the unfolding of an inclination-flip homoclinic orbit. *Ergod. Th. & Dynam. Syst*, 16(5):1071–1086, 1996.

[240] S. Newhouse. Nondensity of Axiom A(a) on S^2. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*, pages 191–202. Amer. Math. Soc., 1970.

[241] S. Newhouse. Diffeomorphisms with infinitely many sinks. *Topology*, 13:9–18, 1974.

[242] S. Newhouse. The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms. *Publ. Math. I.H.E.S.*, 50:101–151, 1979.

[243] S. Newhouse. Continuity properties of entropy. *Annals of Math.*, 129:215–235, 1990. Errata in Annals of Math. 131:409–410,1990.

[244] S. Newhouse and J. Palis. Bifurcations of Morse-Smale dynamical systems. In *Dynamical systems (Proc. Sympos. Univ. Bahia, Salvador, 1971)*, pages 303–366. Academic Press, 1973.

[245] S. Newhouse and J. Palis. Cycles and bifurcation theory. *Astérisque*, 31:44–140, 1976.

[246] S. Newhouse, J. Palis, and F. Takens. Bifurcations and stability of families of diffeomorphisms. *Publ. Math. I.H.E.S.*, 57:5–71, 1983.

[247] V. Niţică and A. Török. An open dense set of stably ergodic diffeomorphisms in a neighborhood of a non-ergodic one. *Topology*, 40:259–278, 2001.

[248] T. Nowicki. A positive Lyapunov exponent for the critical value of an S-unimodal mapping implies uniform hyperbolicity. *Ergod. Th. & Dynam. Sys.*, 8:425–435, 1988.

[249] T. Nowicki and S. van Strien. Invariant measures exist under a summability condition for unimodal maps. *Invent. Math.*, 105:123–136, 1991.

[250] B. Oksendal. *Stochastic differential equations. An introduction with applications.* Universitext. Springer Verlag, 1995.

[251] K. Oliveira. Equilibrium states for certain non-uniformly hyperbolic systems. Preprint www.preprint.impa.br, 2002.

[252] K. Oliveira. *Thermodynamical formalism for a robust class of non-uniformly hyperbolic maps.* PhD thesis, 2003. Preprint www.preprint.impa.br.

[253] V. I. Oseledets. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.

[254] J. C. Oxtoby and S. M. Ulam. Measure-preserving homeomorphisms and metrical transitivity. *Ann. of Math.*, 42:874–920, 1941.

[255] M. J. Pacifico and A. Rovella. Unfolding contracting singular cycles. *Annales Sci. École Norm. Sup.*, 26:691–700, 1993.

[256] M. J. Pacifico, A. Rovella, and M. Viana. Persistence of global spiraling attractors. In preparation.

[257] M. J. Pacifico, A. Rovella, and M. Viana. Infinite-modal maps with global chaotic behavior. *Ann. of Math.*, 148:441–484, 1998. Corrigendum in Annals of Math. 149, page 705, 1999.

[258] J. Palis. A note on Ω-stability. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*, pages 221–222. Amer. Math. Soc., 1970.

[259] J. Palis. On the C^1 Ω-stability conjecture. *Publ. Math. I.H.E.S.*, 66:211–215, 1988.

[260] J. Palis. A global view of Dynamics and a conjecture on the denseness of finitude of attractors. *Astérisque*, 261:335–347, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[261] J. Palis and W. de Melo. *Geometric Theory of Dynamical Systems. An introduction.* Springer Verlag, 1982.

[262] J. Palis and S. Smale. Structural stability theorems. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*, pages 223–232. Amer. Math. Soc., 1970.

[263] J. Palis and F. Takens. Hyperbolicity and the creation of homoclinic orbits. *Annals of Math.*, 125:337–374, 1987.

[264] J. Palis and F. Takens. *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations.* Cambridge University Press, 1993.

[265] J. Palis and M. Viana. On the continuity of Hausdorff dimension and limit capacity for horseshoes. In *Dynamical systems, Valparaiso 1986*, pages 150–160. Springer Verlag, 1988.

[266] J. Palis and M. Viana. High dimension diffeomorphisms displaying infinitely many periodic attractors. *Ann. of Math.*, 140:207–250, 1994.

[267] J. Palis and J.-C. Yoccoz. Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension. *Acta Math.*, 172:91–136, 1994.

[268] J. Palis and J.-C. Yoccoz. Fers à cheval non uniformément hyperboliques engendrés par une bifurcation homocline et densité nulle des attracteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 333:867–871, 2001.

[269] J. Palis and J.-C. Yoccoz. Implicit formalism for affine-like maps and parabolic composition. In *Global analysis of dynamical systems*, pages 67–87. IOP Publ., 2001.

[270] M. Peixoto. Structural stability on two-dimensional manifolds. *Topology*, 1:101–120, 1962.

[271] Ya. Pesin. *Dimension theory in dynamical systems.* University of Chicago Press, 1997. Contemporary views and applications.

[272] Ya. Pesin and Ya. Sinai. Gibbs measures for partially hyperbolic attractors. *Ergod. Th. & Dynam. Sys.*, 2:417–438, 1982.

[273] Ya. B. Pesin. Families of invariant manifolds corresponding to non-zero characteristic exponents. *Math. USSR. Izv.*, 10:1261–1302, 1976.

[274] Ya. B. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Math. Surveys*, 324:55–114, 1977.

[275] V. Pliss. On a conjecture due to Smale. *Diff. Uravnenija*, 8:262–268, 1972.

[276] R.V. Plykin. Sources and currents of A-diffeomorphisms of surfaces. *Math USSR Sbornik*, 94:243–264, 1974.

[277] M. Pollicott. *Lectures on ergodic theory and Pesin theory on compact manifolds.* Cambridge University Press, 1993.

[278] C. Pugh. The closing lemma. *Amer. J. of Math.*, 89:956–1009, 1967.

[279] C. Pugh. An improved closing lemma and a general density theorem. *Amer. J. of Math.*, 89:1010–1021, 1967.

[280] C. Pugh and C. Robinson. The C^1 closing lemma, including Hamiltonians. *Ergod. Th. & Dynam. Sys.*, 3:261–313, 1983.

[281] C. Pugh and M. Shub. Ergodic attractors. *Trans. Amer. Math. Soc.*, 312:1–54, 1989.

[282] C. Pugh and M. Shub. Stably ergodic dynamical systems and partial hyperbolicity. *J. Complexity*, 13:125–179, 1997.

[283] C. Pugh and M. Shub. Stably ergodicity and julienne quasi-conformality. *J. Europ. Math. Soc.*, 2:1–52, 2000.

[284] E. Pujals. *On saddle-node focus singular cycles.* PhD thesis, IMPA, 1995.

[285] E. Pujals and M. Sambarino. The dynamics of dominated splitting. Preprint 2001.

[286] E. Pujals and M. Sambarino. Homoclinic tangencies and hyperbolicity for surface diffeomorphisms. *Annals of Math.*, 151:961–1023, 2000.

[287] E. Pujals and M. Sambarino. On homoclinic tangencies, hyperbolicity, creation of homoclinic orbits, and variation of entropy. *Nonlinearity*, 13:921–926, 2000.

[288] A. Pumariño and A. Rodriguez. *Persistence and coexistence of strange attractors in homoclinic saddle-focus connections*, volume 1658 of *Lect. Notes in Math.* Springer Verlag, 1997.

[289] A. Pumariño and J. A. Rodriguez. Coexistence and persistence of infinitely many strange attractors. *Ergod. Th. & Dynam. Sys.*, 21:1511–1523, 2001.

[290] J. Robbin. A structural stability theorem. *Annals of Math.*, 94:447–493, 1971.

[291] C. Robinson. C^r structural stability implies Kupka-Smale. In *Dynamical Systems (Proc. Sympos. Univ. Bahia, Salvador, 1971)*, pages 443–449. Academic Press, 1971.

[292] C. Robinson. Structural stability of vector fields. *Annals of Math.*, 99:154–175, 1974. Errata in Annals of Math. 101:368, 1975.

[293] C. Robinson. Structural stability of C^1 diffeomorphisms. *J. Diff. Equ.*, 22:28–73, 1976.

[294] C. Robinson. Homoclinic bifurcation to a transitive attractor of Lorenz type. *Nonlinearity*, 2:495–518, 1989.

[295] C. Robinson and L. S. Young. Nonabsolutely continuous foliations for an Anosov diffeomorphism. *Invent. Math.*, 61:159–176, 1980.

[296] F. Rodriguez-Hertz. *Stable ergodicity of certain linear automorphisms of the torus.* PhD thesis, IMPA, 2001. Preprint premat.fing.edu.uy 2001.
[297] V.A. Rokhlin. Lectures on the entropy theory of measure-preserving transformations. *Russ. Math. Surveys*, 22 -5:1–52, 1967. Transl. from Uspekhi Mat. Nauk. 22 - 5 (1967), 3–56.
[298] N. Romero. Persistence of homoclinic tangencies in higher dimensions. *Ergod. Th. & Dynam. Sys.*, 15, 1995.
[299] A. Rovella. The dynamics of perturbations of the contracting Lorenz attractor. *Bull. Braz. Math. Soc.*, 24:233–259, 1993.
[300] D. Ruelle. A measure associated with Axiom A attractors. *Amer. J. Math.*, 98:619–654, 1976.
[301] D. Ruelle. An inequality for the entropy of differentiable maps. *Bull. Braz. Math. Soc.*, 9:83–87, 1978.
[302] D. Ruelle. Ergodic theory of diferentiable dynamical systems. *Publ. Math. IHES*, 50:27–58, 1981.
[303] D. Ruelle. The thermodynamical formalism for expanding maps. *Comm. Math. Phys.*, 125:239–262, 1989.
[304] D. Ruelle and A. Wilkinson. Absolutely singular dynamical foliations. *Comm. Math. Phys.*, 219:481–487, 2001.
[305] M. Rychlik. Lorenz attractors through Shil'nikov-type bifurcation. Part 1. *Erg. Th. & Dynam. Syst.*, 10:793–821, 1989.
[306] B. san Martín. Contracting singular cycles. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 15:651–659, 1998.
[307] B. san Martín. Saddle-focus singular cycles and prevalence of hyperbolicity. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 15:623–649, 1998.
[308] A. Sannami. An example of a regular Cantor whose arithmetic difference is a Cantor set with positive measure. *Hokkaido Math. Journal*, 21:7–23, 1992.
[309] W. Shen. *The metric property of multimodal interval maps and density of Axiom A.* PhD thesis, Univ. Hiroshima, 2000.
[310] L. P. Shil'nikov. A case of the existence of a denumerable set of periodic motions. *Sov. Math. Dokl.*, 6:163–166, 1965.
[311] M. Shub. Endomorphisms of compact differentiable manifolds. *Amer. Journal of Math.*, 91:129–155, 1969.
[312] M. Shub. Topologically transitive diffeomorphisms on T^4. In *Dynamical Systems*, volume 206 of *Lect. Notes in Math.*, page 39. Springer Verlag, 1971.
[313] M. Shub and A. Wilkinson. Pathological foliations and removable zero exponents. *Invent. Math.*, 139:495–508, 2000.
[314] M. Shub and A. Wilkinson. Stably ergodic approximation: two examples. *Ergod. Th. & Dynam. Sys.*, 20:875–893, 2000.
[315] R. Simon. A 3-dimensional Abraham-Smale example. *Procs. Amer. Math. Soc.*, 34:629–630, 1972.
[316] Ya. Sinai. Gibbs measures in ergodic theory. *Russian Math. Surveys*, 27:21–69, 1972.
[317] D. Singer. Stable orbits and bifurcations of maps of the interval. *SIAM J. Appl. Math.*, 35:260–267, 1978.
[318] S. Smale. Diffeomorphisms with many periodic points. *Bull. Am. Math. Soc.*, 73:747–817, 1967.
[319] J. Sotomayor. Generic bifurcations of dynamical systems. In *Dynamical Systems (Procs. Sympos., Univ. Bahia, Salvador, 1971)*, pages 561–582. Academic Press, 1973.
[320] C. Sparrow. *The Lorenz equations: bifurcations, chaos and strange attractors*, volume 41 of *Applied Mathematical Sciences.* Springer Verlag, 1982.
[321] E. Sternberg. On the structure of local homeomorphisms of Euclidean n-space - II. *Amer. J. Math.*, 80:623–631, 1958.
[322] D. Sullivan. *Bounds, quadratic differentials, and renormalization conjectures*, volume II of *A.M.S. Centennial Publications.* Amer. Math. Soc., 1992.
[323] A. Tahzibi. *Stably ergodic diffeomorphisms which are not partially hyperbolic.* PhD thesis, IMPA, 2002.
[324] F. Takens. Intermittancy: Global aspects. In *Dynamical systems, Valparaiso 1986*, volume 1331 of *Lecture Notes in Mathematics*, pages 213–239. Springer Verlag, 1988.

[325] L. Tedeschini-Lalli and J. Yorke. How often do simple dynamical systems have infinitely many coexisting sinks ? *Comm. Math. Phys.*, 106:635–657, 1986.

[326] C. Tresser. About some theorems by L. P. Shil'nikov. *Ann. Inst. Henri Poincaré*, 40:441–461, 1984.

[327] W. Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris Sér. I Math.*, 328:1197–1202, 1999.

[328] W. Tucker. A rigorous ODE solver and Smale's 14th problem. *Found. Comput. Math.*, 2:53–117, 2002.

[329] R. Ures. On the approximation of Hénon-like attractors by homoclinic tangencies. *Ergod. Th. & Dynam. Sys.*, 15, 1995.

[330] S. Ushiki, H. Oka, and H. Kokubu. Existence of strange attractors in the unfolding of a degenerate singularity of a vector field in invariant by translation. *C.R. Acad. Sci. Paris., Sér I*, 298:39–42, 1984.

[331] C. Vásquez. *Statistical stability for diffeomorphisms with a dominated splitting.* PhD thesis, IMPA, 2003. Preprint www.preprint.impa.br.

[332] M. Viana. Almost all cocycles over any hyperbolic system have positive Lyapunov exponents. Preprint www.preprint.impa.br 2003.

[333] M. Viana. Strange attractors in higher dimensions. *Bull. Braz. Math. Soc.*, 24:13–62, 1993.

[334] M. Viana. Global attractors and bifurcations. In *Nonlinear dynamical systems and chaos (Groningen, 1995)*, pages 299–324. Birkhäuser, 1996.

[335] M. Viana. Multidimensional nonhyperbolic attractors. *Publ. Math. I.H.E.S.*, 85:63–96, 1997.

[336] M. Viana. *Stochastic dynamics of deterministic systems.* Lecture Notes XXI Braz. Math. Colloq. IMPA, Rio de Janeiro, 1997.

[337] M. Viana. Dynamics: a probabilistic and geometric perspective. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, pages 557–578, 1998.

[338] Q. Wang and L.-S. Young. Strange attractors with one direction of instability in higher dimensions. Preprint 2002.

[339] Q. Wang and L.-S. Young. Strange attractors with one direction of instability. *Comm. Math. Phys.*, 218:1–97, 2001.

[340] Q. Wang and L.-S Young. From invariant curves to strange attractors. *Comm. Math. Phys.*, 225:275–304, 2002.

[341] L. Wen. Homoclinic tangencies and dominated splittings. *Nonlinearity.*

[342] L. Wen and Z. Xia. C^1 connecting lemmas. *Trans. Amer. Math. Soc.*, 352:5213–5230, 2000.

[343] A. Wilkinson. Stable ergodicity of the time-one map of a geodesic flow. *Ergod. Th. & Dynam. Sys.*, 18:1545–1587, 1998.

[344] R.F. Williams. The structure of the Lorenz attractor. *Publ. Math. IHES*, 50:73–99, 1979.

[345] Z. Xia. Homoclinic points in symplectic and volume-preserving diffeomorphisms. *Comm. Math. Phys.*, 177(2):435–449, 1996.

[346] E. Yanagida. Branching of double pulse solutions from single pulse solutions in nerve axon equations. *J. Diff. Eq.*, 66:243–262, 1987.

[347] J.-C. Yoccoz. Travaux de Herman sur les tores invariants. *Astérisque*, 206:Exp. No. 754, 4, 311–344, 1992. Séminaire Bourbaki, Vol. 1991/92.

[348] Y. Yomdin. Volume growth and entropy. *Israel J. of Math.*, 57:285–300, 1987.

[349] Y. Yomdin. Volume growth and entropy. *Israel J. Math.*, 57:285–300, 1987.

[350] L.-S. Young. Dimension, entropy and Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*, 2:109–124, 1982.

[351] L.-S. Young. Stochastic stability of hyperbolic attractors. *Ergod. Th. & Dynam. Sys.*, 6:311–319, 1986.

[352] L.-S. Young. Decay of correlations for certain quadratic maps. *Comm. Math. Phys.*, 146:123–138, 1992.

[353] G.-C. Yuan and J. Yorke. An open set of maps for which every point is absolutely nonshadowable. *Procs. Amer. Math. Soc.*, 128:909–918, 2000.

[354] E. Zehnder. Note on smoothing symplectic and volume preserving diffeomorphisms. *Lect. Notes in Math.*, 597:828–854, 1977.

Laboratoire de Topologie, UMR 5584 du CNRS, B.P. 47 870, 21078 Dijon Cedex, France

E-mail address: bonatti@satie.u-bourgogne.fr

Departamento de Matemática, PUC-Rio, R. Marquês de S. Vicente 225, 22453-900 Rio de Janeiro, Brazil
E-mail address: lodiaz@mat.puc-rio.br

IMPA, Estrada D. Castorina 110, 22460-320 Rio de Janeiro, Brazil
E-mail address: viana@impa.br