Lecture Notes on Attractors and Physical Measures

Marcelo Viana [1]

IMPA - Rio de Janeiro

[1]Partially supported by PRONEX - Dynamical Systems, Brazil

Contents

1 Definitions and Examples

Let $f : M \to M$ be a measurable map on some measurable space M. Let μ be a probability measure defined on the σ-algebra of M, invariant under f:

$$\mu(f^{-1}(B)) = \mu(B) \quad \text{for every measurable set } B \subset M.$$

We use δ_p to represent the Dirac measure supported on a point p in M.

Definition 1.1. The *basin* of μ is the set $B(\mu)$ of points $z \in M$ such that,

$$\frac{1}{n}\sum_{j=0}^{n-1} \delta_{f^j(z)} \to \mu \text{ in the weak* sense,} \quad \text{as } n \to \infty. \tag{1}$$

In other words, $z \in B(\mu)$ if and only if the time average

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z))$$

exists and coincides with the space average $\int \varphi\, d\mu$, for every continuous function $\varphi : M \to \mathbb{R}$. Then one also says that z is a *generic point* of μ.

The main concept in this section is that of observable invariant measure: if an initial state $z \in M$ is chosen at random then $z \in B(\mu)$ with positive probability, in a *physically relevant* sense. In which case, the measure μ can be computed as the time average along the orbit of z. To turn this into a definition, one has to express the idea of physical chance precisely: as usual, we postulate that this corresponds to Lebesgue measure in the space M.

For this last notion to be defined M should have some smooth structure, e.g., a finite-dimensional manifold, or branched manifold, possibly with boundary. More generally, M could be a subdomain, or even a tower over a subdomain of such a manifold. Then we call *Lebesgue measure* to any measure m on the Borel σ-algebra of M that is generated by a volume form. More precisely, for every $p \in M$ there exists a volume form ω_p on a neighbourhood V_p of p, so that

$$m(B) = \int_B d\omega_p \quad \text{for every measurable set } B \subset V_p.$$

Of course, this is not uniquely defined. However, since different Lebesgue measures in M are all equivalent, in the sense that they all have the same zero measure sets, it is often irrelevant to which of them we are referring. So, except where otherwise specified, we use "Lebesgue measure" to mean any measure in the Lebesgue class.

Definition 1.2. An f-invariant Borel probability measure μ is a *physical*, or *Sinai-Ruelle-Bowen (SRB) measure* for f if its basin $B(\mu)$ has positive Lebesgue measure.

The previous definitions extend naturally to continuous-time dynamical systems, i.e. systems described by flows or by semi-flows. Let X^t, $t \geq 0$ be a semi-flow on M: X^0 is the identity, and each X^t, $t \geq 0$, is a measurable transformation on M, with $X^{t+s} = X^t \circ X^s$ for every t, s. A probability measure μ is invariant under the semi-flow if it is invariant under every map X^t, $t \geq 0$.

Definition 1.3. The *basin* of μ is the set $B(\mu)$ of points $z \in M$ such that, given any continuous function $\varphi : M \to \mathbb{R}$,

$$\frac{1}{T} \int_0^T \varphi(X^t(z))\, dt \to \int \varphi \, d\mu \quad \text{as} \quad T \to +\infty. \tag{2}$$

We say that μ is a *physical*, or *Sinai-Ruelle-Bowen (SRB) measure* for X^t if its basin has positive Lebesgue measure.

Example 1.4. Suppose the map $f : M \to M$ admits an invariant probability measure μ that is absolutely continuous with respect to Lebesgue measure, and ergodic. Then μ is a physical measure for f, as a simple consequence of Birkhoff's ergodic theorem. Indeed, the theorem states that $B(\mu)$ has non-zero (even full) measure for μ, and so it must have non-zero Lebesgue measure.

For the same reasons, ergodic absolutely continuous invariant measures of a semi-flow are physical measures for the semi-flow.

Example 1.5. Let $f : [0,1] \to [0,1]$ be a piecewise C^2 map with a neutral fixed point at the origin, as in Figure 1. That is, we suppose that $f(0) = 0$ and $f'(0) = 1$, but the second derivative $f''(0)$ is non-zero. On the other hand, $|f'(z)| > 1$ for every $z \neq 0$, including $z = c^{\pm}$. It can be seen that the orbit of Lebesgue almost every point $z \in [0,1]$ spends almost all the time in an arbitrarily small neighbourhood of the origin: given any $\delta > 0$,

$$\frac{1}{n} \#\big\{ j \in \{0, 1, \ldots, n-1\} : |f^j(z)| < \delta \big\} \to 1, \qquad \text{as } n \to \infty.$$

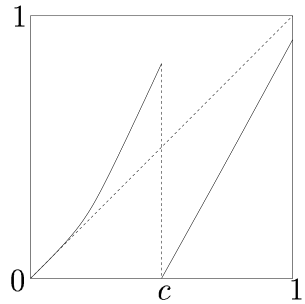

Figure 1: A map with a neutral fixed point

It follows that, given any continuous function $\varphi : [0,1] \to \mathbb{R}$ and any $\varepsilon > 0$, we have

$$\Big|\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z)) - \varphi(0)\Big| < \epsilon$$

for every large n. So, for every continuous function φ and Lebesgue almost every point z,

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z)) = \varphi(0) = \int \varphi \, d\delta_0 \, .$$

This means that the Dirac measure at zero is the unique SRB measure of f.

Now we describe an example, due to Bowen, of a flow in the plane for which time averages fail to converge for a whole open set of points. In particular, there is no physical measure whose basin intersects this open set. Similar arguments apply to the time-1 map $f = X^1$ of this flow.

Example 1.6. The flow is described in Figure 2. A main feature is the existence of a double saddle-connection between saddle-points A and B. We denote by L the region bounded by the separatrices that form this connection. Let $-\lambda_A < 0 < \sigma_A$ and $-\lambda_B < 0 < \sigma_B$ be the eigenvalues of the flow at A and B, respectively. We assume that

$$\frac{\lambda_A \lambda_B}{\sigma_A \sigma_B} > 1,$$

to ensure that the boundary of L attracts the orbits of all points $z \in L$ that are close enough to it. Then those orbits must visit, alternately, the

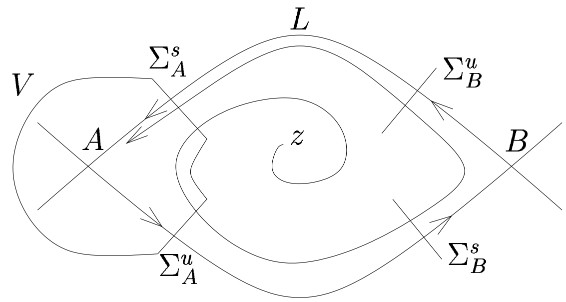

Figure 2: A flow without physical measure

vicinities of A and B. Fix cross sections Σ_A^s, Σ_A^u close to A and intersecting its stable and unstable separatrices, respectively. Similarly, let Σ_B^s, Σ_B^u be cross sections intersecting the stable and unstable separatrices of B. Fix z and let

$$\cdots < T_A^s(j) < T_A^u(j) < T_B^s(j) < T_B^u(j) < T_A^s(j+1) < \cdots$$

be the successive times at which the orbit of z intersects these cross-sections. Then $\tau_A(j) = T_A^u(j) - T_A^s(j)$ and $\tau_B(j) = T_B^u(j) - T_B^s(j)$ correspond to the successive times spent by the orbit near each of the saddles. It is an elementary exercise to check that

- $\tau_A(j) \approx \tau_B(j)$ and $\tau_B(j) \approx \tau_A(j+1)$, where $\approx$ means that the quotients of the two expressions are bounded by some constant independent of j;
- both sequences $\tau_A(j)$ and $\tau_B(j)$ increase exponentially fast with j, at the rate $(\lambda_A \lambda_B)/(\sigma_A \sigma_B)$;
- the transition times $T_B^s(j) - T_A^u(j)$ and $T_A^s(j+1) - T_B^u(j)$ are bounded by some constant independent of j.

As a consequence, each visit time is comparable to the total time elapsed thus far: there exists $c > 0$ such that

$$\tau_A(j) \geq c T_A^s(j) \quad \text{and} \quad \tau_B(j) \geq c T_B^s(j),$$

for every j. Now we may easily conclude that the time averages $T^{-1} \int \delta_{X^t(z)}\, dt$ do not converge as $T \to +\infty$, for any point $z \in L$ close to the boundary of

L. Indeed, suppose otherwise, and let μ be the limit. Let V be some neighbourhood of A as in Figure 2. Up to slightly modifying V we may suppose that its boundary has zero μ-measure. Then $\mu(V) = \lim_{T\to+\infty} \tau_V(T)$, where

$$\tau_V(T) = \frac{1}{T}\int \mathcal{X}_V(X^t(z))\,dt$$

is the fraction of the time interval $[0,T]$ spent by z in V. Now,

$$\tau_V(T_A^s(j)) \geq \frac{\tau_A^s(j)}{T_A^s(j)} \geq c \qquad \text{but} \qquad \tau_V(T_A^u(j+1)) \leq \frac{1}{1+c}\,\tau_V(T_A^s(j))$$

for every j. This implies that $\tau_V(T)$ has no limit as $T \to +\infty$, and so we have reached a contradiction.

Such examples show that *SRB measures need not exist* for *all* systems. Existence results are usually difficult, and are known only for certain classes of systems. In particular, it is unknown in which generality do the basins of physical measures cover at least a full Lebesgue measure subset of the phase-space M. We will return to this fundamental problem later.

Right now, let us suppose that a map $f : M \to M$ does have some SRB measure μ. Let m denote Lebesgue measure restricted to the basin of μ, and normalized so as to be a probability. By definition,

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

for every $z \in B(\mu)$, and every continuous function $\varphi : M \to \mathbb{R}$. Suppose, for the sake of simplicity, that M is compact. Then the sequence on the left is bounded in norm by $\sup|\varphi|$. As a direct consequence of the dominated convergence theorem,

$$\int \frac{1}{n}\sum_{j=0}^{n-1}\varphi\circ f^j\,dm \to \int\int \varphi\,d\mu\,dm = \int \varphi\,d\mu.$$

The expression on the left is precisely the integral of φ with respect to the measure $n^{-1}\sum_{j=0}^{n-1} f_*^j m$. In other words, we have proved that

$$\frac{1}{n}\sum_{j=0}^{n-1} f_*^j m \;\to\; \mu \quad \text{in the weak}^* \text{ topology.} \tag{3}$$

This simple observation suggests that SRB measures might be found as limits or, at least, accumulation points of the averages of forward iterates $f_*^j m$ of Lebesgue measure, possibly restricted to some subset of the phase-space M and normalized.

It is a well-known consequence of the Banach-Alaoglu theorem that the space of probability measures on a compact metric space is compact with respect to the weak* topology. See [33, Section I.8] for instance. Therefore, accumulation points

$$\mu = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j m \tag{4}$$

always exist. Moreover, assuming that the map f is continuous on M, the push-forward operator $f_* : \eta \mapsto f_*\eta$ is also continuous, relative to the weak* topology in the space of Borel measures in M.

Using this fact, one concludes readily that any such accumulation point μ is an invariant measure for f. In fact,

$$f_*\mu = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^{j+1} m = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j m + \lim_{k\to\infty} \frac{1}{n_k} \left(f_*^{n_k} m - m\right).$$

The first limit on the right is μ, by assumption, and the second one is identically zero: given any continuous function φ,

$$\frac{1}{n_k} \int (\varphi \circ f^{n_k} - \varphi)\, dm \to 0 \quad \text{as } n_k \to \infty.$$

This shows that $f_*\mu = \mu$, which is just the same as saying that μ is invariant.

On the other hand, there is no a priori reason for a measure μ as in (4) to have particularly interesting properties: recall for instance Example 1.6. Indeed, to be able to conclude that such a μ is an SRB measure one must keep a fair control of the sequence

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j m, \tag{5}$$

which is where most of the difficulty lies.

In the next three sections we present some important cases where such a control is possible: the measures μ_n are all absolutely continuous with respect to Lebesgue measure, with uniform bounds on the *densities* (Radon-Nikodym derivatives). More general situations will appear subsequently.

2 Uniformly Expanding Maps

In this section we prove that any uniformly expanding map on a compact (connected) manifold with Hölder continuous Jacobian admits a unique physical measure μ. Moreover, the basin of μ is a full Lebesgue measure subset of the manifold. See Theorem 2.1 for the precise statement. This is a central result from Ruelle's theory [**?**] of equilibrium states for expanding maps.

2.1 Definitions and Basic Properties

Let M be a compact manifold and $f : M \to M$ be a C^1 map.

Definition 2.1. We say that f is *(uniformly) expanding* if there exist constants $C > 0$ and $\sigma > 1$ such that

$$\|Df^n(x)v\| \geq C\sigma^n \|v\| \quad \text{for every } x \in M,\ v \in T_xM, \text{ and } n \geq 1. \tag{6}$$

Here $\|\cdot\|$ denotes an arbitrary Riemannian norm on the manifold M: since all norms are equivalent, (6) holds for $\|\cdot\|$ if and only if it holds for any other norm, apart from the fact that the constants may vary. As a matter of fact, up to choosing a convenient norm, we may always suppose that $C = 1$.

Indeed, let (6) hold for some norm $\|\cdot\|$, and constants C, σ. Given any $1 < \sigma_* < \sigma$, fix $N \geq 1$ large enough so that $C(\sigma/\sigma_*)^N \geq 1$. Then, consider the Riemannian norm $\|\cdot\|_*$ defined by

$$\|v\|_*^2 = \sum_{j=0}^{N-1} \sigma_*^{-2j} \|Df^j(x)v\|^2,$$

for each $x \in M$ and $v \in T_xM$. Direct substitution gives

$$\|Df(x)v\|_* \geq \sigma_* \|v\|_* \quad \text{for every } x \in M \text{ and } v \in T_xM. \tag{7}$$

A Riemannian norm as in (7) is said to be *adapted* to f. It also follows from these remarks that the set of expanding maps is open in the C^1 topology: if f satisfies (7) then, up to slightly reducing $\sigma_* > 1$, so does any map g in some C^1 neighbourhood.

Example 2.2. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map such that $F(\mathbb{Z}^n) \subset \mathbb{Z}^n$. Then there exists a unique map f on the n-dimensional torus $M = \mathbb{R}^n/\mathbb{Z}^n$ such that $f \circ \pi = \pi \circ F$, where $\pi : \mathbb{R}^n \to M$ is the canonical projection. If all the eigenvalues $\lambda_1, \ldots, \lambda_n$ of F have norm larger than 1 then this map f is expanding: any $1 < \sigma < \inf_i |\lambda_i|$ will do in (6).

According to [55], every expanding map on the n-torus is topologically conjugate to a linear model f as in Example 2.2. More generally, cf. [21], a manifold admits expanding maps if and only if it is an infranilmanifold, and then any such map is topologically conjugate to an algebraic expanding endomorphism. In general, the conjugacy is a *singular map*: it does not preserve the class of sets with zero Lebesgue measure.

Definition 2.3. Let X, Y be metric spaces and $f : X \to Y$ be a continuous map. Given $C > 0$ and $0 < \nu \leq 1$, we say that f is (C, ν)-*Hölder* if

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2)^\nu \quad \text{for every } x_1, x_2 \text{ in } X.$$

When $\nu = 1$ we also say that f is C-*Lipschitz*. In general, f is ν-*Hölder* if it is (C, ν)-Hölder for some $C > 0$.

The following theorem summarizes results of existence and uniqueness of SRB measures that we prove in Subsections 2.2 and 2.3.

Theorem 2.1. *Let $f : M \to M$ be a uniformly expanding map on a compact manifold M. Assume that there exists $0 < \nu_0 \leq 1$ such that the logarithm $M \ni x \mapsto \log|\det Df(x)|$ of the Jacobian of f is ν_0-Hölder.*

Then f admits a unique invariant measure μ which is absolutely continuous with respect to Lebesgue measure. Moreover, μ is ergodic, its support coincides with M, and its basin $B(\mu)$ is a full Lebesgue measure subset of M. In particular, μ is the unique SRB-measure of f.

The assumption of expansiveness is used in the proof of this theorem through the consequences provided by the following proposition. We fix, once and for all, a Riemannian norm $\|\cdot\|$ adapted to f, and denote $d(\cdot\,,\cdot)$ the corresponding distance on the manifold.

Proposition 2.2. *Let M be a compact manifold and $f : M \to M$ be a C^1 expanding map. Then there exists $k \geq 1$ such that every point $y \in M$ has exactly k pre-images under f. Moreover, there exists $\rho_0 > 0$ such that given any pre-image x of a point $y \in M$ there exists a C^1 map $h : B(y, \rho_0) \to M$ with $f \circ h = \mathrm{id}$, $h(y) = x$, and*

$$d(h(y_1), h(y_2)) \leq \sigma^{-1}\, d(y_1, y_2) \quad \textit{for every } y_1, y_2 \in B(y, \rho_0).$$

Proof. We only sketch the arguments, as they are quite standard. Clearly, (6) implies that the derivative Df is an isomorphism at every point. So, given any $x \in M$ there exists $\rho_0 > 0$ such that f maps some neighbourhood $V(x)$ of x diffeomorphically onto the ball of radius ρ_0 around $y = f(x)$. By compactness, ρ_0 may be chosen independent of x. Then the number of pre-images of any $y \in M$ must be finite and even bounded. It also follows that the set of points with exactly n pre-images is open, for every $n \geq 0$. So, by connectedness, the number of pre-images must be the same for every $y \in M$. Finally, let us denote $h = (f \mid V(x))^{-1}$. Since the norm is adapted to f,

$$\|Dh(z)\| = \|Df(h(z))^{-1}\| \leq \sigma^{-1}$$

for every z in the domain of h, and so h contracts distances by a factor σ^{-1}, as stated. □

Maps h as in the statement are called *(local) inverse branches* of f. More generally, we can define inverse branches h^n of f^n, $n \geq 1$, as follows. Given $y \in M$ and $x \in f^{-n}(y)$, let $h_1, \ldots, h_n$ be inverse branches of f with

$$h_j(f^{n-j+1}(x)) = f^{n-j}(x)$$

for every $1 \leq j \leq n$. Since each h_j is a contraction, its image is contained in a ball of radius less than ρ_0 around $f^{n-j}(x)$. Then $h^n = h_n \circ \cdots \circ h_1$ is well-defined on the ball of radius ρ_0 around y. Clearly, $f^n \circ h^n = \mathrm{id}$ and $h^n(y) = x$.

2.2 Upper Bounds on the Densities

It is easy to see that the pre-image of a zero Lebesgue measure set under an expanding map f also has zero Lebesgue measure. It follows that if a probability measure ν is absolutely continuous with respect to Lebesgue measure, then the same is true for its push-forward $f_*\nu$ under f. Let m be Lebesgue measure on M, normalized so that $m(M) = 1$. Then, in particular, $f_*^n m$ is absolutely continuous with respect to m for every $n \geq 1$.

We prove in Proposition 2.4 that if f is an expanding map with Hölder continuous Jacobian, as in the statement of Theorem 2.1, then the densities $d(f_*^n m)/dm$ are bounded by some constant independent of $n \geq 1$. From this we deduce that any accumulation point of the sequence (5) is absolutely continuous with respect to Lebesgue measure, with density bounded by that same constant. In particular, f has some invariant measure μ that is absolutely continuous with respect to Lebesgue measure.

The main step is the following result of bounded distortion, which is also the only place where the assumption of Hölder continuity is needed in the proof.

Lemma 2.3. *There exists $C_1 > 0$ such that given any $n \geq 1$, any $y \in M$, and any inverse branch $h^n : B(y, \rho_0) \to M$ of f^n,*

$$\frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} \leq \exp(C_1 d(y_1, y_2)^{\nu_0}) \leq \exp(C_1 (2\rho_0)^{\nu_0})$$

for every $y_1, y_2 \in B(y, \rho_0)$.

Proof. Let us write h^n as a composition $h^n = h_n \circ \cdots \circ h_1$ of inverse branches of f. We also denote $h^i = h_i \circ \cdots \circ h_1$ for $1 \leq i < n$, and $h^0 = \mathrm{id}$. Then

$$\log \frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} = \sum_{i=1}^{n} \log |\det Dh_i(h^{i-1}(y_1))| - \log |\det Dh_i(h^{i-1}(y_2))| .$$

Note that $\log |\det Dh_i| = -\log |\det Df| \circ h_i$ and, by assumption, $\log |\det Df|$ is (C_0, ν_0)-Hölder for some $C_0 > 0$. Moreover, cf. Proposition 2.2, each h_j is a σ^{-1}-contraction. Then,

$$\log \frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} \leq \sum_{i=1}^{n} C_0 \, d(h^i(y_1), h^i(y_2))^{\nu_0} \leq \sum_{i=1}^{n} C_0 \, \sigma^{-i\nu_0} d(y_1, y_2)^{\nu_0}.$$

So, to prove the lemma it is enough to take $C_1 = C_0 \sum_{i=1}^{\infty} \sigma^{-i\nu_0}$. □

Proposition 2.4. *There exists $C_2 > 0$ such that $(f_*^n m)(B) \leq C_2 m(B)$ for every measurable set $B \subset M$ and every $n \geq 1$.*

Proof. It is no restriction to take B contained in some ball $B_0 = B(z, \rho_0)$ of radius ρ_0 around a point $z \in M$. Lemma 2.3 implies that

$$\frac{m(h^n(B))}{m(h^n(B_0))} = \frac{\int_B |\det Dh^n| \, dm}{\int_{B_0} |\det Dh^n| \, dm} \leq \exp(C_1 (2\rho_0)^{\nu_0}) \frac{m(B)}{m(B_0)} ,$$

for each inverse branch h^n of f^n at z. Moreover, $(f_*^n m)(B) = m(f^{-n}(B))$ is the sum of $m(h^n(B))$ over all inverse branches, and analogously for B_0. So, we get that

$$\frac{(f_*^n m)(B)}{(f_*^n m)(B_0)} \leq \exp(C_1 (2\rho_0)^{\nu_0}) \frac{m(B)}{m(B_0)} .$$

Of course, $(f_*^n m)(B_0) \le (f_*^n m)(M) = 1$. Moreover, the Lebesgue measure of the balls of fixed radius ρ_0 is bounded from zero by some $\alpha_0 > 0$ that depends only on ρ_0. Now it suffices to take $C_2 = \exp(C_1(2\rho_0)^{\nu_0})/\alpha_0$. □

Lemma 2.5. *Let ν be a probability measure on a compact metric space X, and $\varphi : X \to [0, +\infty)$ be integrable with respect to ν. Let μ_i, $i \ge 1$, be a sequence of probability measures on X converging to some μ, in the weak* sense. If $\mu_i \le \varphi\nu$ for every $i \ge 1$ then $\mu \le \varphi\nu$.*

Proof. Let B be any measurable set. For each $\varepsilon > 0$, let K_ε be a compact subset of B such that $\mu(B \setminus K_\varepsilon)$ and $(\varphi\nu)(B \setminus K_\varepsilon)$ are both less than ε. Then let A_ε be the open neighbourhood of K_ε defined by $A_\varepsilon = \{z : d(z, K_\varepsilon) < r\}$, where $r > 0$ is small enough so that the measure of $A_\varepsilon \setminus K_\varepsilon$ is less than ε for both μ and $\varphi\nu$. Changing r if necessary, we may suppose that the boundary of A_ε has zero μ-measure (there are at most countably many exceptional values of r). Then $\mu = \lim \mu_i$ implies $\mu(A_\varepsilon) = \lim \mu_i(A_\varepsilon) \le (\varphi\nu)(A_\varepsilon)$. Making $\varepsilon \to 0$ we get $\mu(B) \le (\varphi\nu)(B)$. □

Now we apply this lemma to our situation, with $\varphi \equiv C_2$, $\nu = m$, and $\mu_i = n_i^{-1} \sum_{j=0}^{n_i-1} f_*^j m$ for any subsequence $(n_i)_i$ such that $(\mu_i)_i$ converges to some measure μ. We immediately get

Corollary 2.6. *Every accumulation point μ of the sequence $n^{-1} \sum_{j=0}^{n-1} f_*^j m$ is an f-invariant measure absolutely continuous with respect to Lebesgue measure.*

2.3 Ergodicity and Uniqueness

Now we show that the probability measure μ constructed above is ergodic and so, recall Example 1.4, is a physical measure for f. We also get that the basin of μ is a full Lebesgue measure subset of M. In particular, the physical measure is unique.

We begin by fixing some partition $\mathcal{P}_0 = \{U_1, \ldots, U_s\}$ of M into regions with non-empty interior and diameter less than ρ_0. Then, for each $n \ge 1$, we let $\mathcal{P}_n$ be the partition of M consisting of the images of each of the U_i, $1 \le i \le s$, under corresponding inverse branches of f^n. The diameter of $\mathcal{P}_n$, defined as the supremum of the diameters of its elements, is less than $\rho_0 \sigma^{-n}$.

Lemma 2.7. *Let $\mathcal{P}_n$, $n \ge 1$, be a sequence of partitions in a compact metric space with diameters converging to zero as $n \to \infty$. Let ν be a probability* *measure in that space, and B be any measurable subset such that $\nu(B) > 0$. Then there are $V_n \in \mathcal{P}_n$, for $n \geq 1$, so that*

$$\nu(V_n) > 0 \quad \textit{and} \quad \frac{\nu(B \cap V_n)}{\nu(V_n)} \to 1 \quad \textit{as } n \to \infty.$$

Proof. Given any $0 < \varepsilon < \nu(B)$, let K_ε be some compact subset of B with $\nu(B \setminus K_\varepsilon) < \varepsilon$. As the diameter of the partitions converges to zero, the measure of the union $A_{\varepsilon,n}$ of all the elements of $\mathcal{P}_n$ that intersect K_ε satisfies $\nu(A_{\varepsilon,n} \setminus K_\varepsilon) < \varepsilon$ for every large enough n. If we had

$$\nu(K_\varepsilon \cap V_n) \leq \frac{\nu(B) - \varepsilon}{\nu(B) + \varepsilon}\, \nu(V_n)$$

for every $V_n \in \mathcal{P}_n$ that intersects K_ε, it would follow that

$$\nu(K_\varepsilon) \leq \frac{\nu(B) - \varepsilon}{\nu(B) + \varepsilon}\, \nu(A_{\varepsilon,n}) \leq \frac{\nu(B) - \varepsilon}{\nu(B) + \varepsilon}\, (\nu(K_\varepsilon) + \varepsilon) \leq \nu(B) - \varepsilon,$$

a contradiction. So, there must be some $V_n \in \mathcal{P}_n$ with

$$\nu(B \cap V_n) \geq \nu(K_\varepsilon \cap V_n) > \frac{\nu(B) - \varepsilon}{\nu(B) + \varepsilon}\, \nu(V_n)$$

and this also implies $\nu(V_n) > 0$. The statement follows, taking $\varepsilon \to 0$. □

We say that $A \subset M$ is an *invariant set* of $f : M \to M$ if $f^{-1}(A) = A$.

Lemma 2.8. *Let $A \subset M$ be an invariant set of a $C^{1+\nu_0}$ expanding map f such that $m(A) > 0$. Then A has full Lebesgue measure in some $U_i \in \mathcal{P}_0$, that is, there exists $1 \leq i \leq s$ so that $m(U_i \setminus A) = 0$.*

Proof. By Lemma 2.7, there exist $V_n \in \mathcal{P}_n$ such that $m(V_n \setminus A)/m(V_n)$ converges to zero as $n \to \infty$. Let $U_{i(n)} = f^n(V_n)$. Applying Lemma 2.3 to the inverse branch of f^n mapping $U_{i(n)}$ to V_n, we conclude that

$$\frac{m(U_{i(n)} \setminus A)}{m(U_{i(n)})} = \frac{m(f^n(V_n) \setminus A)}{m(f^n(V_n))} \leq \exp\big(C_1 (2\rho_0)^{\nu_0}\big) \frac{m(V_n \setminus A)}{m(V_n)}$$

also converges to zero. Since $\mathcal{P}_0$ is finite, there must exist $1 \leq i \leq s$ such that $i(n) = i$ for infinitely many values of n. Then $m(U_i \setminus A) = 0$. □

Corollary 2.9. *Any $C^{1+\nu_0}$ expanding map $f : M \to M$ has some ergodic absolutely continuous invariant measure.*

Proof. As a consequence of the lemma, there exist at most $\#\mathcal{P}_0$ two-by-two disjoint invariant sets with positive Lebesgue measure. It follows that M can be partitioned into finitely many minimal positive Lebesgue measure invariant sets $A_1, \ldots, A_s$, $s \le \#\mathcal{P}_0$: minimality means there are no invariant subsets $B_i \subset A_i$ with $0 < m(B_i) < m(A_i)$. Given any f-invariant absolutely continuous measure μ, there is some i such that $\mu(A_i) > 0$. Then the normalized restriction μ_i of μ to A_i,

$$\mu_i(B) = \frac{\mu(B \cap A_i)}{\mu(A_i)}$$

is invariant, absolutely continuous, and ergodic (because A_i is minimal). □

This argument also gives that there exist only finitely many measures as in the statement. The last step in the proof of Theorem 2.1 is to show that, in fact, such a measure is unique. This requires the following topological result.

Lemma 2.10. *Given any non-empty open set $U \subset M$, there exists $N \ge 1$ such that $f^N(U) = M$.*

Proof. Let $x \in U$ and $r > 0$ be such that the ball of radius r around x is contained in U. Given any $n \ge 1$, suppose that $f^n(U)$ does not cover the whole manifold. Then, there is some curve γ connecting $f^n(x)$ to a point $y \in M \setminus f^n(U)$, and γ may be taken with length less than $\operatorname{diam} M + 1$. By lifting γ through the local diffeomorphism f^n we get a curve γ_n connecting x to a point $y_n \in M \setminus U$. Then $r \le \operatorname{length}(\gamma_n) \le \sigma^{-n}(\operatorname{diam} M + 1)$, which gives an upper bound on n. Thus, $f^n(U) = M$ for every large enough n, as stated. □

Corollary 2.11. *If $A \subset M$ is an f-invariant set with positive Lebesgue measure, then A has full Lebesgue measure in the whole manifold M.*

Proof. Let U be the interior of a set U_i as given by Lemma 2.8, and let $N \ge 1$ be such that $f^N(U) = M$. Then $m(U \setminus A) = 0$, and so $M \setminus A = f^N(U) \setminus f^N(A) \subset f^N(U \setminus A)$ also has zero Lebesgue measure. □

The following statement completes the proof of Theorem 2.1.

Corollary 2.12. *Let μ be an absolutely continuous invariant measure of f. Then μ is ergodic and its basin $B(\mu)$ has full Lebesgue measure in M. Moreover, the support of μ is the whole manifold M.*

Proof. If A is an f-invariant subset of M then, by the previous corollary, either A or A^c have zero Lebesgue measure. So either $\mu(A) = 0$ or $\mu(A^c) = 0$. This proves ergodicity. Then $B(\mu)$ is an invariant set with positive Lebesgue measure, and so it must have full Lebesgue measure. Similarly, as the support of μ is a compact invariant subset with positive Lebesgue measure, it must coincide with M. □

In particular, the map f has a unique absolutely continuous invariant measure μ. It can be shown, from our previous arguments, that the density $d\mu/dm$ may be taken Hölder continuous and bounded away from zero on M.

3 Piecewise Expanding Maps

In quick terms, we call a transformation $f : M \to M$ piecewise expanding if the ambient manifold M or, at least, a full Lebesgue measure subset of it, can be partitioned into countably many domains restricted to which the transformation is expanding and sufficiently differentiable. Precise definitions will appear later. Figure 3 describes some simple examples we have in mind.

Example 3.1. We say that $f : [0,1] \to [0,1]$ is a *tent map* if it is continuous and there exists $c \in (0,1)$ such that

- the derivative Df is constant and larger than one in norm, in each of the intervals $[0,c)$ and $(0,1]$.

See Figure 3. More generally, one may consider maps that are affine and expanding on each interval (c_{i-1}, c_i), $1 \le i \le N$, for some finite sequence of points $0 = c_0 < c_1 < \cdots < c_N = 1$.

The next class of examples play a key role in the theory of Lorenz-like attractors of flows, see [2], [22].

Example 3.2. We call $f : [0,1] \to [0,1]$ a *Lorenz-like map* if there there exist $c \in (0,1)$ and $\sigma > 1$ such that

- f is C^2 on each of intervals $[0,c)$ and $(c,1]$, with a discontinuity at c;

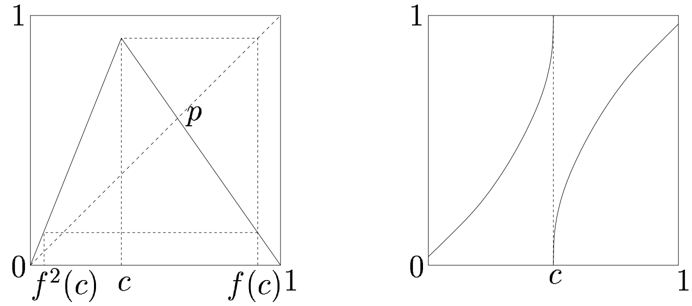

Figure 3: Tent map and Lorenz-like map

- $|Df(x)| > \sigma$ for every $x \in [0,1] \setminus \{c\}$
- the left and the right derivatives at c are both infinite, and $1/|Df|$ extends to c as a Lipschitz function on $[0,1]$.

See Figure 3. This can also be generalized, to include maps with any number of singular points (any number of regularity intervals).

Besides the intrinsic interest of this class of systems, there are several reasons for studying piecewise expanding maps. For one thing, they provide a fairly simple setting for dealing with difficulties that are common to much more complicated systems. This point will be illustrated in a little while. At least as important, piecewise expanding maps are often found in the course of studying other dynamical systems, such as Lorenz-like attractors of flows, unimodal maps of the circle or the interval, or more general non-uniformly hyperbolic transformations on manifolds: properties of the system can be understood through constructing and analyzing certain piecewise expanding maps that are associated to it.

In this section we only deal with one-dimensional maps, that is, we always take M to be either the circle S^1 or the compact interval $I = [0,1]$. Throughout, m denotes some normalized Lebesgue measure in M. The higher dimensional case is discussed in Section 4.

3.1 Definitions and Statements

Let $f : M \to M$ be so that there exist $C > 0$, $\sigma > 1$, and a family $\mathcal{P}^1$ of two-by-two disjoint intervals covering a full Lebesgue measure subset of M, such that f is C^2 restricted to each $\xi \in \mathcal{P}^1$, and

$$|Df^n| \geq C\sigma^n \quad \text{for } n \geq 1, \tag{8}$$

at any point where the derivative exists. For each $n \geq 1$, let $\mathcal{P}^n$ be the family of *regularity intervals* of f^n. That is, the elements of $\mathcal{P}^n$ are the maximal intervals η such that $f^j(\eta)$ is contained in some atom of $\mathcal{P}^1$, for every $0 \leq j \leq n-1$.

Suppose that $\log |Df \mid \xi|$ is Lipschitz continuous for every $\xi \in \mathcal{P}^1$, with uniform Lipschitz constant. This holds, for instance, if $\mathcal{P}^1$ is finite and every $f \mid \xi$ admits a C^2 extension to the boundary. Then, the same arguments as in Lemma 2.3 give that the inverse branches $h^n_\eta = (f^n \mid \eta)^{-1}$ of f^n have uniformly bounded distortion: there exists $K > 0$ such that

$$\sup \Big\{ \frac{|Dh^n_\eta(y_1)|}{|Dh^n_\eta(y_2)|} : y_1, y_2 \in f^n(\eta) \Big\} \leq K \tag{9}$$

for every $\eta \in \mathcal{P}^n$ and $n \geq 1$. In particular, each measure $f^n_*(m \mid \eta)$ is absolutely continuous with respect to Lebesgue measure on $f^n(\eta)$, with density bounded by some uniform constant.

Since $f^n_* m$ is the sum of the $f^n_*(m \mid \eta)$ over all $\eta \in \mathcal{P}^n$, one may hope to show that the measures $f^n_* m$, $n \geq 1$, have uniformly bounded densities, which would imply that f has some invariant measure absolutely continuous with respect to Lebesgue measure. This would follow from the proof of Proposition 2.4, if one knew that the Lebesgue measure of the intervals $f^n(\eta)$, $\eta \in \mathcal{P}^n$, $n \geq 1$, is uniformly bounded away from zero. However, this is generally *not* the case, even if f has finitely many regularity intervals.

This fact is a main source of difficulties, and we shall return to it in a while. Before that, let us briefly discuss a special class of maps for which a lower bound for the measure of the $f^n(\eta)$ does exist.

Markov Expanding Maps Suppose the map $f : M \to M$ satisfies

(M1) the image $f(\xi)$ of every $\xi \in \mathcal{P}^1$ coincides with some union of elements of $\mathcal{P}^1$, up to a zero Lebesgue measure set;

(M2) there exists $\delta > 0$ such that $m(f(\xi)) \geq \delta$ for any $\xi \in \mathcal{P}^1$.

The first condition implies that $f^n(\eta)$ contains the image $f(\xi)$ of some $\xi \in \mathcal{P}^1$, for every $\eta \in \mathcal{P}^n$. Then (M2) gives $m(f^n(\eta)) \geq \delta$ for every $n \geq 1$. So, the

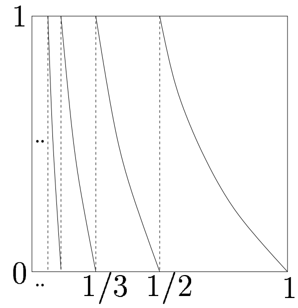

Figure 4: The Gauss map

proof of Proposition 2.4 carries on to this case. It follows, as in Corollary 2.6, that f has absolutely continuous invariant measures.

Moreover, we can combine this condition (M2) with the distortion bound (9) to conclude, as in Subsection 2.3, that any f-invariant set A with positive Lebesgue measure must have $m(A) \geq \delta$. As a consequence, f has finitely many ergodic absolutely continuous invariant probabilities, and any absolutely continuous invariant measure is a linear combination of the ergodic ones. Finally, if f is transitive (in the sense that for any $\xi, \eta \in \mathcal{P}^1$, there exists n such that $\xi \cap f^n(\eta)$ has positive Lebesgue measure) then it has a unique absolutely continuous invariant probability.

These facts are often referred to as the Folklore theorem. See e.g. [11] for references and a discussion of the origin of this statement.

Observe that the proof still works if we just assume $\log |D(f \mid \xi)|$ to be Hölder continuous, with uniform constants. In fact, it also extends to higher dimensions, assuming $\log |\det Df|$ is Hölder continuous on each regularity domain, always with uniform constants.

Example 3.3. The *Gauss map* $G : [0,1] \to [0,1]$ is defined by

$$G(x) = 1/x - [1/x] \quad \text{for } x \neq 0, \quad \text{and} \quad G(0) = 0.$$

The family of regularity intervals is $\mathcal{P}^1 = \{\big(1/n+1, 1/n\big) : n \geq 1\}$. Note that $|DG(x)| = 1/x^2 \geq 1$ wherever the derivative is defined. Moreover,

$$|DG(x)| > 2 \ \text{ if } \ x \leq 2/3 \quad \text{and} \quad |DG^2(x)| \geq 4 \ \text{ if } \ x > 2/3.$$

This implies (8), with $\sigma = 2$ and $C = 1/2$. Finally, given any $1/(n+1) < x < y < 1/n$,

$$\big| \log |DG(x)| - \log |DG(y)| \big| \leq 2(n+1)(y-x) \leq 4(y-x)^{1/2}.$$

So $\log Df$ is $(4, 1/2)$-Hölder on each element of $\mathcal{P}^1$. Properties (M1) and (M2) are clear, with $\delta = 1$.

Now we go back to general, possibly non-Markov, piecewise expanding maps in dimension one. As we already mentioned, the "chopping" that takes place at the boundary of the regularity intervals may cause the length of the iterates $f^n(\eta)$, $\eta \in \mathcal{P}^n$, to be arbitrarily small. Then, conceivably, one might have several small intervals $f^n(\eta)$ piling-up in a same region, thus causing the density of $f_*^n m$ to grow unbounded as $n \to \infty$.

That this does not actually occur was first proved by [29], for maps with finitely many regularity intervals: the densities

$$\phi_n = \frac{d(f_*^n m)}{dm}$$

have *uniformly bounded variation* and, hence, they are uniformly bounded. Existence of some absolutely continuous invariant measure is a direct consequence, cf. Corollary 2.9.

Building on this approach, [30] showed that f admits some ergodic absolutely continuous invariant measure, and the number of such measures is finite. It was observed by [60] that similar arguments apply under a weaker regularity assumption, see (E2) below, that allows for maps with unbounded derivative as in Example 3.2. Then [**?**] extended these results to maps with infinitely many regularity intervals.

Definition 3.4. We call $f : M \to M$ a *piecewise expanding map*, on either $M = S^1$ or $M = [0, 1]$, if there exists a countable family $\mathcal{P}^1$ of two-by-two disjoint intervals covering a full Lebesgue measure subset of M, such that

(E1) the restriction of f to each $\xi \in \mathcal{P}^1$ is a C^1 monotonic map, and the function $\xi \ni x \mapsto 1/|Df(x)|$ has bounded variation;

(E2) there exist constants $C > 0$ and $\sigma > 1$ such that $|Df^n(x)| \geq C\sigma^n$ for every $n \geq 1$, and every $x \in M$ for which the derivative is defined.

We call $\mathcal{P}^1$ a *partition into regularity intervals* of f. We say that f has finitely many regularity intervals if the partition $\mathcal{P}^1$ may be chosen finite. The boundary points of the regularity intervals $\xi \in \mathcal{P}^1$ that are not on the boundary of M are called *singular points* of f.

The following theorem is proved in Subsections 3.3 and 3.4.

Theorem 3.1. *Let f be a piecewise expanding map of the circle or the interval, with finitely many regularity intervals.*

Then f has some ergodic invariant probability measure absolutely continuous with respect to Lebesgue measure, and the number of such measures is bounded by the number of singular points of f. The union of their basins is a full Lebesgue measure subset of M. Moreover, any absolutely continuous invariant measure μ can be written $\mu = \varphi m$ where φ has bounded variation.

It is not difficult to see that this result can not hold in the general infinite case. The following counterexample is due to [**?**], another had been given by [29].

Example 3.5. Let $f : [0,1] \to [0,1]$ be given by $f(0) = 0$ and

$$f(x) = 2x - 2^{-j+1} \quad \text{for} \quad x \in (2^{-j}, 2^{-j+1}] \text{ and each } j \geq 1.$$

Then $f(x) \leq x$ for every $x \in [0,1]$, and the equality holds if and only if x is in $\mathrm{Fix}(f) = \{2^{-k} : k \geq 0\} \cup \{0\}$. Then, by Poincaré's recurrence theorem, every f-invariant measure is supported in $\mathrm{Fix}(f)$, and so it is a linear combination of Dirac measures on fixed points of the map. In particular, f has no absolutely continuous invariant measure. We leave it to the reader to check that Lebesgue almost every orbit of f converges to 0, and so the Dirac measure at zero is the unique SRB measure.

On the other hand, [**?**] also provides a natural condition under which most of the conclusion of Theorem 3.1 does extend to piecewise expanding maps with infinitely many regularity intervals. For each $\xi \in \mathcal{P}^1$, let $\hat{g}_\xi : M \to \mathbb{R}$ be the function defined by

$$\hat{g}_\xi(x) = \frac{1}{|Df(x)|} \quad \text{if } x \in \xi \qquad \text{and} \qquad \hat{g}_\xi(x) = 0 \quad \text{otherwise.} \tag{10}$$

Conditions (E1) and (E2) in the definition imply that each $\hat{g}_\xi$ has bounded variation. So, the next result generalizes Theorem 3.1.

Theorem 3.2. *Let f be a piecewise expanding map of the circle or the interval. Assume that, for some choice of a partition $\mathcal{P}^1$,*

$$\sum_{\xi \in \mathcal{P}^1} \mathrm{var}\, \hat{g}_\xi < \infty.$$

Then f has some ergodic invariant probability measure absolutely continuous with respect to Lebesgue measure, and there are finitely many such measures. The union of their basins covers a full Lebesgue measure subset of M. Moreover, if μ is any absolutely continuous invariant measure then $\mu = \varphi m$ where φ has bounded variation.

The proof of this theorem is given in Subsection 3.5, where we also describe a few examples and applications.

3.2 Bounded Variation Functions

Let us begin by recalling the definition and some elementary properties of the notion of variation of real functions defined on the circle or the interval. See, for instance, [?] for more information.

Definition 3.6. Let $\varphi : M \to \mathbb{R}$ and $\eta = [a, b]$ be a compact interval in M. The *variation* of φ on η is

$$\operatorname*{var}_{\eta} \varphi = \sup \sum_{i=1}^{n} |\varphi(x_{i-1}) - \varphi(x_i)|$$

where the supremum is over all finite sequences $a = x_0 < x_1 < \cdots < x_n = b$, $n \geq 1$, with $<$ representing an arbitrary orientation on η.

The variation $\operatorname{var}_\eta \varphi$ of φ on an arbitrary connected subset η of M (including $\eta = M = S^1$) is the supremum of its variations over all compact intervals contained in η. We represent the variation $\operatorname{var}_M \varphi$ over the whole ambient manifold M simply as $\operatorname{var} \varphi$.

Definition 3.7. A function $\varphi : M \to \mathbb{R}$ has *bounded variation* if $\operatorname{var} \varphi$ is finite. Given any connected subset η of M, φ has *bounded variation on* η if $\operatorname{var}_\eta \varphi < \infty$.

The following properties follow directly from the definition.

Lemma 3.3. *Let $\varphi_1, \varphi_2 : M \to \mathbb{R}$ and η be a connected subset of M.*

1. $\operatorname*{var}_{\eta}(\varphi_1 + \varphi_2) \leq \operatorname*{var}_{\eta} \varphi_1 + \operatorname*{var}_{\eta} \varphi_2$;
2. $\operatorname*{var}_{\eta}(\varphi_1 \varphi_2) \leq \operatorname*{var}_{\eta} \varphi_1 \sup_{\eta} |\varphi_2| + \sup_{\eta} |\varphi_1| \operatorname*{var}_{\eta} \varphi_2$;

3. $\sup_\eta \varphi_1 \le \operatorname{var}_\eta \varphi_1 + \inf_\eta \varphi_1 \le \operatorname{var}_\eta \varphi_1 + \dfrac{1}{m(\eta)} \displaystyle\int_\eta \varphi_1 \, dm$;

4. $\operatorname{var}_\eta |\varphi_1| \le \operatorname{var}_\eta \varphi_1$;

5. $\operatorname{var}_\eta (\varphi_1 \circ h) = \operatorname{var}_{h(\eta)} \varphi_1$ *if* $h : \eta \to h(\eta)$ *is a homeomorphism.*

The claims in the next lemma are also simple consequences of the definition. See for instance [**?**] for proofs.

Lemma 3.4. *Suppose* $\varphi : M \to \mathbb{R}$ *has bounded variation on some interval* $\eta \subset M$*. Then*

1. *the restriction of* φ *to* η *can be written as the difference* $\varphi_1 - \varphi_2$ *of two non-decreasing functions;*

2. φ *has at most countably many discontinuity points;*

3. *the lateral limits* $\lim_{x \to z^\pm} \varphi(x)$ *exist at every point* $z \in \eta$ *(for points on the boundary consider only the limit from the inside of* η*).*

Now we prove Helly's theorem: sets of functions which are uniformly bounded and have uniformly bounded variation are relatively compact in $L^1(m)$.

Lemma 3.5. *Let* $K_1, K_2 > 0$ *and* $\psi_n : M \to \mathbb{R}$*,* $n \ge 1$*, be a sequence of functions on* M *such that* $\sup \psi_n \le K_1$ *and* $\operatorname{var} \psi_n \le K_2$ *for every* $n \ge 1$*.*

Then there exists a subsequence $(\psi_{n_k})_k$ *and a function* $\psi_0 : M \to \mathbb{R}$ *with* $\sup |\psi_0| \le K_1$ *and* $\operatorname{var} \psi_0 \le K_2$ *such that* ψ_{n_k} *converges to* ψ_0 *as* $k \to \infty$*,* m*-almost everywhere and in* $L^1(m)$*.*

Proof. We consider $M = [0,1]$, the case of the circle is analogous. Write

$$\psi_n^+(x) = \operatorname{var}_{[0,x]} \psi_n \quad \text{and} \quad \psi_n^- = \psi_n^+ - \psi_n.$$

Then $(\psi_n^-)_n$ and $(\psi_n^+)_n$ are uniformly bounded sequences of non-decreasing functions. Choose $(n_k)_k$ so that $\psi_{n_k}^\pm(q)$ converges to some real number $\phi_0^\pm(q)$ as $k \to \infty$, for every rational $q \in [0,1]$. Clearly, $\phi_0^\pm(q_1) \le \phi_0^\pm(q_2)$ whenever $q_1 \le q_2$. Then, extend $\phi_0^\pm$ to non-decreasing functions in the whole $[0,1]$ by setting

$$\phi_0^\pm(x) = \inf\{\phi_0^\pm(q) : q \in [x,1] \cap \mathbb{Q}\}.$$

We claim that $\psi_{n_k}^{\pm}(x)$ converges to $\phi_0^{\pm}(x)$ as $k \to \infty$, for every continuity point x of $\phi_0^{\pm}$ (a co-countable set). Indeed, given any such x and any $\delta > 0$, we may fix rational numbers $q_1 \le x \le q_2$ such that

$$\phi_0^{\pm}(x) - \delta \le \phi_0^{\pm}(q_1) \le \phi_0^{\pm}(x) \le \phi_0^{\pm}(q_2) \le \phi_0^{\pm}(x) + \delta.$$

Then, for every sufficiently large k,

$$\phi_0^{\pm}(x) - 2\delta \le \phi_0^{\pm}(q_1) - \delta \le \psi_{n_k}^{\pm}(q_1) \le \psi_{n_k}^{\pm}(x)$$

and, analogously, $\psi_{n_k}^{\pm}(x) \le \phi_0^{\pm}(x) + 2\delta$. This proves the claim.

Next, let $\psi_0^{\pm}$ be right-continuous functions coinciding with $\phi_0^{\pm}$ at every point of continuity of $\phi_0^{\pm}$, and define $\psi_0 = \psi_0^{+} - \psi_0^{-}$. It follows that ψ_{n_k} converges to ψ_0 except, possibly, on a countable set of points E. In particular, $\psi_{n_k} \to \psi_0$ m-almost everywhere and in $L^1(m)$. Finally,

$$|\psi_0(x)| = \lim_k |\psi_{n_k}(x)| \le \sup_k \sup \psi_{n_k} \qquad \text{and}$$

$$\sum_{j=1}^{s} |\psi_0(x_j) - \psi_0(x_{j-1})| = \lim_k \sum_{j=1}^{s} |\psi_{n_k}(x_j) - \psi_{n_k}(x_{j-1})| \le \sup_k \operatorname{var} \psi_{n_k},$$

for every x and $x_0 \le x_1 \le \cdots \le x_s$ in $[0,1] \setminus E$. Since ψ_0 is right-continuous, this implies that $\sup |\psi_0| \le K_1$ and $\operatorname{var} \psi_0 \le K_2$. □

It is useful to extend the notion of bounded variation to elements of the space $L^1(m)$, in the following way.

Definition 3.8. The *variation* $\operatorname{var}_\eta[\varphi]$ of an element $[\varphi] \in L^1(m)$ on an interval $\eta \subset M$ is the infimum of the variations $\operatorname{var}_\eta \varphi$ on η taken over all representatives of $[\varphi]$. We say that $[\varphi]$ has *bounded variation on* η if $\operatorname{var}_\eta[\varphi]$ is finite.

Remark 3.9. It is not difficult to check that a representative φ realizes the infimum in the definition of $\operatorname{var}_\eta[\varphi]$ if and only if φ is continuous at the boundary points of η, and $\varphi(z)$ is between $\lim_{x\to z^-} \varphi(x)$ and $\lim_{x\to z^+} \varphi(z)$ for every point z in the interior of η.

Moreover, the variation of such a function φ over any $[a,b] \subset \eta$ coincides with the supremum of $\sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})|$ taken only over the sequences $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ such that φ is continuous at every x_i.

Lemma 3.6. *For any $[\varphi] \in L^1(m)$ with bounded variation and any interval $\eta \in M$,*

$$\operatorname*{var}_{\eta}[\varphi] = \sup \left\{ \Big| \int_\eta \varphi \, D\omega \, dm \Big| : \omega \in C_0^1(\eta) \quad \textit{with} \quad \sup|\omega| \le 1 \right\}.$$

where $C_0^1(\eta)$ is the space of all C^1 functions on M that are zero in the complement of η.

Proof. Let φ be a representative of $[\varphi]$ such that $\operatorname{var}_\eta \varphi = \operatorname{var}_\eta[\varphi] < \infty$. According to Lemma 3.4, we can always extend φ continuously to the closure of η. This extension is also a representative of $[\varphi]$, and its variation on $\operatorname{clos}(\eta)$ is equal to $\operatorname{var}_\eta \varphi$. Therefore, it is no restriction to suppose that η is compact $\eta = [a, b]$.

The lemma is a consequence of claims (1) and (2) below.

(1) Given any sequence $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, such that φ is continuous at x_i for every $0 \le i \le n$, and given any $\varepsilon > 0$, there exists $\omega \in C_0^1(\eta)$ with $\sup|\omega| \le 1$, satisfying

$$\Big| \int_\eta \varphi \, D\omega \, dm \Big| \ge \sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})| - \varepsilon.$$

Fix $\delta > 0$ such that $|\varphi(x) - \varphi(x_i)| \le \varepsilon/2n$ for every x such that $|x - x_i| \le \delta$, and every $0 \le i \le n$. Let ω be any C^1 function on η such that

- $\omega(x_i) = 0$ for $0 \le i \le n$, and ω is monotone (either non-increasing or non-decreasing) on $[x_{i-1}, x_{i-1} + \delta]$ and on $[x_i - \delta, x_i]$, for all $1 \le i \le n$;
- $\omega | [x_{i-1} + \delta, x_i - \delta] \equiv \operatorname{sgn}(\varphi(x_{i-1}) - \varphi(x_i))$ for every $1 \le i \le n$.

Here sgn denotes the usual sign function $\operatorname{sgn}(z) = z/|z|$, with $\operatorname{sgn}(0) = 0$. Let $1 \le i \le n$ be fixed. Then

$$\int_{x_{i-1}}^{x_i} \varphi \, D\omega \, dm = \int_{x_{i-1}}^{x_{i-1}+\delta} \varphi \, D\omega \, dm + \int_{x_i - \delta}^{x_i} \varphi \, D\omega \, dm.$$

If $\operatorname{sgn}(\varphi(x_{i-1}) - \varphi(x_i)) = 1$ then $D\omega \ge 0$ in the first integral, and $D\omega \le 0$ in the second one. It follows that the integral of $\varphi \, D\omega$ on $[x_{i-1}, x_i]$ is bounded from below by

$$\big(\varphi(x_{i-1}) - \frac{\varepsilon}{2n}\big) \int_{x_{i-1}}^{x_{i-1}+\delta} D\omega \, dm + \big(\varphi(x_i) + \frac{\varepsilon}{2n}\big) \int_{x_i-\delta}^{x_i} D\omega \, dm$$
$$= \big(\varphi(x_{i-1}) - \frac{\varepsilon}{2n}\big) - \big(\varphi(x_i) + \frac{\varepsilon}{2n}\big) = |\varphi(x_i) - \varphi(x_{i-1})| - \frac{\varepsilon}{n}.$$

Analogously, we get the same conclusion also when the sign is -1 or zero. Then, adding over all $1 \le i \le n$,

$$\int_a^b \varphi \, D\omega \, dm \ge \sum_{i=1}^n |\varphi(x_i) - \varphi(x_{i-1})| - \varepsilon.$$

This proves (1), which implies the inequality $\le$ in the statement.

(2) Given any function $\omega \in C_0^1(\eta)$ with $\sup|\omega| \le 1$, and given any $\varepsilon > 0$, there exists a sequence $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, satisfying

$$\sum_{i=1}^n |\varphi(x_i) - \varphi(x_{i-1})| \ge \Big|\int_\eta \varphi \, D\omega \, dm\Big| - \varepsilon.$$

Let $n \ge 1$ and $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ be arbitrary. We use $\mathcal{E}$ to represent various expressions converging to zero as $\sup_i |x_i - x_{i-1}|$ converges to zero. On the one hand,

$$\int \varphi \, D\omega \, dm = \sum_{i=1}^n \varphi(x_i) D\omega(x_i)(x_i - x_{i-1}) + \mathcal{E}.$$

Since ω is C^1 and $\omega(x_0) = \omega(x_n) = 0$, the term on the right can also be written as

$$\sum_{i=1}^n \varphi(x_i)\big(\omega(x_i) - \omega(x_{i-1})\big) + \mathcal{E} = \sum_{i=2}^{n-1} \big(\varphi(x_{i-1}) - \varphi(x_i)\big)\omega(x_{i-1}) + \mathcal{E}.$$

As $\sup|\omega| \le 1$, it follows that

$$\Big|\int \varphi \, D\omega \, dm\Big| \le \sum_{i=2}^{n-1} |\varphi(x_i) - \varphi(x_{i-1})| + \mathcal{E} \le \sum_{i=1}^n |\varphi(x_i) - \varphi(x_{i-1})| + \mathcal{E}.$$

This means that any sequence with sufficiently small $\sup_i |x_i - x_{i-1}|$ satisfies the inequality in (2). $\square$

3.3 Absolutely Continuous Invariant Measures

Let f be an expanding map of the circle or the interval, and $\mathcal{P}^1$ be a corresponding partition into regularity intervals. Here, as well as in the next subsection, we assume that $\mathcal{P}^1$ is finite. However, *finiteness is used only in Proposition 3.9 and in Corollary 3.14*.

For every $n \geq 1$, let $\mathcal{P}^n$ be the partition into regularity intervals of f^n, defined by $\mathcal{P}^n(x) = \mathcal{P}^n(y)$ if and only if $\mathcal{P}^1(f^j(x)) = \mathcal{P}^1(f^j(y))$ for all $0 \leq j < n$. For each $n \geq 1$ and $\eta \in \mathcal{P}^n$, we define $g_\eta^n : M \to \mathbb{R}$ by

$$g_\eta^n(y) = \frac{1}{|Df^n|} \circ (f^n \mid \eta)^{-1}(y) \quad \text{if } y \in f^n(\eta),$$

and $g_\eta^n(y) = 0$ otherwise. For simplicity, we also write $g_\eta = g_\eta^1$ for $\eta \in \mathcal{P}^1$.

Observe that, cf. part 5 of Lemma 3.3,

$$\sup g_\eta = \sup \hat{g}_\eta \quad \text{and} \quad \operatorname{var} g_\eta \leq \operatorname{var} \hat{g}_\eta\,, \tag{11}$$

where $\hat{g}_\eta$ was defined in (10). Actually, the two variations coincide except, possibly, if $f(\eta)$ is the whole M.

In general, given any $\varphi : M \to \mathbb{R}$ we consider $\varphi \circ (f^n \mid \eta)^{-1}$ as a function on M, identically zero in the complement of $f^n(\eta)$.

Lemma 3.7. *For every integrable function $\varphi : M \to \mathbb{R}$, and every $n \geq 1$, the iterate of Lebesgue measure under f^n can be written $f_*^n(\varphi m) = \varphi_n m$, with*

$$\varphi_n = \sum_\eta g_\eta^n \cdot \left(\varphi \circ (f^n \mid \eta)^{-1}\right)$$

where the sum is over all $\eta \in \mathcal{P}^n$ such that $m(\eta) > 0$.

Proof. Let $B \subset M$ be an arbitrary measurable set. By definition,

$$f_*^n(\varphi m \mid \eta)(B) = \varphi m(f^{-n}(B) \cap \eta) = \int_{f^{-n}(B)\cap\eta} \varphi(x)\, dm(x).$$

Changing to the variable $y = (f^n \mid \eta)(x)$, we find

$$f_*^n(\varphi m \mid \eta)(B) = \int_{B\cap f^n(\eta)} \frac{\varphi}{|Df^n|} \circ (f^n \mid \eta)^{-1}(y)\, dm(y).$$

The term on the right can be rewritten as $\int_B g^n_\eta \cdot (\varphi \circ (f^n \mid \eta)^{-1})\, dm$, since the functions in this last integral are identically zero outside $f^n(\eta)$. This proves that the density of each $f^n_*(\varphi m \mid \eta)$ is given by $g^n_\eta \cdot (\varphi \circ (f^n \mid \eta)^{-1})$, for every $\eta \in \mathcal{P}^n$. On the other hand,

$$f^n_*(\varphi m) = \sum_{\eta \in \mathcal{P}^n} f^n_*(\varphi m \mid \eta),$$

and the intervals η with $m(\eta) = 0$ play no role in this sum, since the corresponding term $f^n_*(\varphi m)$ is identically zero. The conclusion of the lemma follows by summing over all the $\eta \in \mathcal{P}^n$ having non-zero Lebesgue measure. □

Remark 3.10. By definition, for every $n \geq 1$ we have

$$\int \varphi_n \, dm = \int 1 \, d(f^n_*(\varphi m)) = \int 1 \, d(\varphi m) = \int \varphi \, dm.$$

Let $|\varphi|_n$ be the sequence one obtains as in the previous lemma, when φ is replaced by $|\varphi|$. Then $-|\varphi|_n \leq \varphi_n \leq |\varphi|_n$, since $-|\varphi| \leq \varphi \leq |\varphi|$. In particular,

$$\int |\varphi_n| \, dm \leq \int |\varphi|_n \, dm = \int |\varphi| \, dm \quad \text{for every } n \geq 1. \tag{12}$$

Lemma 3.8. *There exist $C_1 > 0$ and $0 < \lambda_1 < 1$ such that*

$$\sup g^n_\eta \leq C_1 \lambda_1^n \quad \textit{and} \quad \operatorname{var} g^n_\eta \leq C_1 \lambda_1^n \quad \textit{for every } n \geq 1 \textit{ and } \eta \in \mathcal{P}^n.$$

Proof. The first claim is a direct consequence of the definition of g^n_η and the expansivity condition (E2): it is enough to take $C_1 > 1/C$ and $\lambda_1 > 1/\sigma$.

Next, given $\eta \in \mathcal{P}^n$ and $0 \leq j < n$, let $\xi_j \in \mathcal{P}^j$, $\eta_j \in \mathcal{P}^1$, $\zeta_j \in \mathcal{P}^{n-j-1}$ be defined by

$$\eta \subset \xi_j, \qquad f^j(\eta) \subset \eta_j, \qquad f^{j+1}(\eta) \subset \zeta_j.$$

By definition, there exists some constant $C_2 > 0$ such that, for every $\xi \in \mathcal{P}^1$,

$$\begin{aligned}\operatorname{var} g_\xi &\leq \operatorname*{var}_{f(\xi)} \frac{1}{|Df|} \circ (f \mid \xi)^{-1} + 2 \sup_{f(\xi)} \frac{1}{|Df|} \circ (f \mid \xi)^{-1} \\ &\leq \operatorname*{var}_{\xi} \frac{1}{|Df|} + 2 \sup_{\xi} \frac{1}{|Df|} \leq C_2,\end{aligned}$$

(the supremum term bounds the variation of g_ξ at the boundary points of $f(\xi)$). Since $g_\eta^n = g_{\xi_j}^j \, g_{\eta_j} \, g_{\zeta_j}^{n-j-1}$ for every $0 \le j < n$, using Lemma 3.3 we obtain

$$\operatorname{var} g_\eta^n \le \sum_{j=0}^{n-1} \sup g_{\xi_j}^j \ \operatorname{var} g_{\eta_j} \ \sup g_{\zeta_j}^{n-j-1} \le \sum_{j=0}^{n-1} \frac{C_2}{C^2 \sigma^{n-1}} = (C_2 C^{-2} \sigma) n \sigma^n .$$

Fixing $\lambda_1 > 1/\sigma$ and choosing $C_1 > \sup\{C_2 C^{-2} \sigma \, n \, (\lambda_1 \sigma)^{-n} : n \ge 1\}$, we get the second claim. □

The crucial ingredient for most results in this section is the following result of [29], stating that the variations of the densities φ_n tend to decrease as n grows.

Proposition 3.9. *There are $C_0 > 0$ and $0 < \lambda_0 < 1$ such that, given any bounded variation function $\varphi : M \to \mathbb{R}$,*

$$\operatorname{var} \varphi_n \le C_0 \, \lambda_0^n \ \operatorname{var} \varphi + C_0 \int |\varphi| dm$$

for every $n \ge 1$, where φ_n is as in Lemma 3.7.

Proof. Combining Lemma 3.8 with properties in Lemma 3.3 we find

$$\begin{aligned} \operatorname{var} \varphi_n &\le \sum_\eta \operatorname{var} g_\eta^n \sup_\eta |\varphi| + \sup g_\eta^n \big(\operatorname*{var}_\eta \varphi + 2 \sup_\eta |\varphi| \big) \\ &\le \sum_\eta 3 C_1 \lambda_1^n \sup_\eta |\varphi| + C_1 \lambda_1^n \operatorname*{var}_\eta \varphi . \end{aligned} \tag{13}$$

Note that the expression in parentheses is an upper bound for the variation of $\varphi \circ (f^n \mid \eta)^{-1}$, taking into account the jumps at the boundary of $f^n(\eta)$. Using the third property in Lemma 3.3 we get

$$\begin{aligned} \operatorname{var} \varphi_n &\le \sum_\eta 4 C_1 \lambda_1^n \operatorname*{var}_\eta \varphi + 3 C_1 \lambda_1^n \frac{1}{m(\eta)} \int_\eta |\varphi| \, dm \\ &\le 4 C_1 \lambda_1^n \operatorname{var} \varphi + K(n) \int |\varphi| \, dm , \end{aligned} \tag{14}$$

where $K(n) = 3 C_1 \lambda_1^n \sup\{1/m(\eta) : \eta \in \mathcal{P}^n\}$. Recall that we only have to deal with intervals $\eta \in \mathcal{P}^n$ with positive Lebesgue measure.

This is close to what we want, but we still have to explain why $K(n)$ can be replaced by some constant independent of n. For that, we fix $N \geq 1$ such that $4C_1\lambda_1^N \leq 1/2$, and we denote $K_0 = \max\{K(n) : 1 \leq n \leq N\}$. Then, given any $n \geq 1$, we write $n = qN + r$ with $q \geq 0$ and $1 \leq r \leq N$. Using the previous bound with φ replaced by φ_{n-N}, ..., φ_{n-qN}, and φ, respectively, and recalling (12),

$$\begin{aligned}
\operatorname{var} \varphi_n &\leq K_0 \int |\varphi_{n-N}|\, dm + \frac{1}{2} \operatorname{var} \varphi_{n-N} \\
&\leq \left(1 + \cdots + 2^{-q+1}\right) K_0 \int |\varphi| dm + \frac{1}{2^q} \operatorname{var} \varphi_q \\
&\leq \left(1 + \cdots + 2^{-q}\right) K_0 \int |\varphi| dm + \frac{1}{2^q} 4C_1 \lambda_1^r \operatorname{var} \varphi.
\end{aligned}$$

To finish the proof of the proposition, choose $C_0 \geq \max\{2K_0, 4C_1\}$ and $\lambda_0 \geq \max\{2^{-1/N}, \lambda_1\}$. □

Remark 3.11. Let ν and μ_n, $n \geq 1$, be finite measures on a compact metric space, such that μ_n is absolutely continuous with respect to ν for every $n \geq 1$. If the Radon-Nikodym derivatives $d\mu_n/d\nu$ converge in $L^1(\nu)$ to some function ψ then μ_n converges to $\mu = \psi\nu$ in the weak* topology. Indeed, given any continuous function $\varphi : M \to \mathbb{R}$,

$$\left| \int \varphi \, d\mu_n - \int \varphi \, d\mu \right| \leq \int |\varphi| \left| \frac{d\mu_n}{d\nu} - \psi \right| d\nu \leq \sup |\varphi| \left\| \frac{d\mu_n}{d\nu} - \psi \right\|_1,$$

and the last term converges to zero as $n \to \infty$.

In particular, the densities ϕ_n of the iterates $f_*^n m$ of Lebesgue measure are uniformly bounded: by Proposition 3.9 and Remark 3.10,

$$\sup \phi_n \leq \operatorname{var} \phi_n + \int \phi_n \, dm \leq (C_0 + 1) \int 1 \, dm = C_0 + 1, \tag{15}$$

for every $n \geq 1$. Hence, as in Corollary 2.9, we may conclude that f has some absolutely continuous invariant measure. In fact, we can prove more:

Corollary 3.10. *The map f has some absolutely continuous invariant probability measure μ whose density $d\mu/dm$ has bounded variation.*

Proof. Let ϕ_n be as above, and define $\psi_n = n^{-1}\sum_{j=0}^{n-1}\phi_j$ for each $n \geq 1$. Then $\mu_n = n^{-1}\sum_{j=0}^{n-1} f_*^j m$ can be written as $\mu_n = \psi_n m$. Proposition 3.9 implies that $\operatorname{var}\phi_j \leq C_0$ for every j, and so $\operatorname{var}\psi_n \leq C_0$ for every $n \geq 1$. Moreover, by (15), we have $\sup|\psi_n| \leq C_0 + 1$ for every $n \geq 1$. This means that we may apply Lemma 3.5 to conclude that there exists a subsequence $(\psi_{n_k})_k$ converging in $L^1(m)$ to some function φ_0 with $\operatorname{var}\varphi_0 \leq C_0$. In particular, cf. Remark 3.11, $(\mu_{n_k})_k$ converges to $\mu = \varphi_0 m$ in the weak* topology. This ensures that μ is invariant under f. □

3.4 Bounding the Number of Physical Measures

Now we show that a piecewise expanding map has finitely ergodic absolutely continuous invariant measures. This was proved under a C^2 assumption by [30], see also [28]. Then [60] observed that the bounded variation condition (E1) in Definition 3.4 suffices for the argument. The main step is the following consequence of Proposition 3.9.

Proposition 3.11. *Any absolutely continuous probability measure ν of a piecewise expanding map can be written $\nu = \theta m$ where θ has bounded variation.*

Proof. By assumption, one may write $\nu = \psi m$ for some integrable function ψ with L^1-norm $\|\psi\|_1 = \int |\psi|\, dm$ equal to 1. We want to prove that ψ coincides Lebesgue almost everywhere with some bounded variation function θ.

Let $(\xi_l)_l$ be some sequence of bounded variation functions converging to ψ in $L^1(m)$. It is no restriction to suppose that $\|\xi_l\|_1 \leq 2$ for all l. For each $n \geq 1$ and $l \geq 1$, let

$$f_*^n(\psi m) = \psi_n m \quad \text{and} \quad f_*^n(\xi_l m) = \xi_{l,n} m$$

where ψ_n and $\xi_{l,n}$ are obtained from ψ and ξ_l, respectively, as in Lemma 3.7. Note that, since $\nu = \psi m$ was taken invariant, $f_*^n(\psi\, m) = \psi\, m$ and so $\psi_n = \psi$ Lebesgue almost everywhere. Applying Proposition 3.9 to ξ_l we get that

$$\operatorname{var}\xi_{l,n} \leq C_0 \lambda_0^n \operatorname{var}\xi_l + C_0 \int |\xi_l|\, dm \leq 3\, C_0$$

for every large enough n. Moreover, by Lemma 3.3 and Remark 3.10,

$$\sup|\xi_{l,n}| \leq \operatorname{var}\xi_{l,n} + \int |\xi_{l,n}|\, dm \leq 3\, C_0 + \int |\xi_l|\, dm \leq 3\, C_0 + 2.$$

Hence, the sequence $(\xi_{l,n})_n$ satisfies the assumptions of Lemma 3.5, for each fixed l. As a consequence, there exists some function θ_l with $\sup|\theta_l| \le 3C_0+2$ and $\operatorname{var}\theta_l \le 3C_0$, and there exists a subsequence $(n_k)_k$ such that $(\xi_{l,n_k})_k$ converges to θ_l in $L^1(m)$.

On the one hand, using (12) for the function $\varphi = \xi_l - \psi$ we get that

$$\|\theta_l - \psi\|_1 = \lim_{k\to\infty} \|\xi_{l,n_k} - \psi\|_1 = \lim_{k\to\infty} \|\xi_{l,n_k} - \psi_{n_k}\|_1 \le \|\xi_l - \psi\|_1 \, ,$$

for every l. This implies that θ_l converges to ψ in $L^1(m)$, since ξ_l does. On the other hand, we may apply Lemma 3.5 to the sequence $(\theta_l)_l$, to conclude that some subsequence $(\theta_{l_j})_j$ converges in $L^1(m)$ to a function θ with $\operatorname{var}\theta \le 3C_0$. This implies that $\psi = \theta$ Lebesgue almost everywhere. Therefore, $\nu = \psi\, m = \theta\, m$. □

Remark 3.12. We even obtained a uniform bound $3C_0$ for the variation of any density ψ of an absolutely continuous invariant measure.

Lemma 3.12. *Given any f-invariant set $A \subset M$ with positive Lebesgue measure, there exists some absolutely continuous f-invariant probability measure ν_A such that $\nu_A(A) = 1$.*

Proof. Let $\hat{A}$ be the set of points $x \in A$ such that every neighbourhood of x intersects A in a positive Lebesgue measure subset. By Lebesgue's differentiation theorem, $\hat{A}$ contains a full Lebesgue measure subset of A, and so $m(\hat{A}) = m(A) > 0$. Moreover, $\hat{A}$ is close to being invariant, in the following sense. Let $x \in \hat{A}$ and suppose that x is not a singular point. Then f is a local diffeomorphism near x. In particular, since every neighbourhood of x intersects A in a positive Lebesgue measure subset, the same is true for $f(x)$. In other words, $f(x)$ is also in $\hat{A}$. Thus, we have shown that $\hat{A}$ is contained in the union of $f^{-1}(\hat{A})$ with the (finite) set of singular points. Then, $\hat{A}$ is also contained in the union of $f^{-n}(\hat{A})$ with some finite set, for every $n \ge 1$.

Let $(m \mid \hat{A})$ represent the normalized restriction of Lebesgue measure to $\hat{A}$, and consider the sequence of probability measures

$$\mu_{A,n} = \frac{1}{n}\sum_{j=0}^{n-1} f_*^j(m \mid \hat{A}).$$

By definition, $(m \mid \hat{A}) = \varphi m$ where $\varphi = \mathcal{X}_{\hat{A}}/m(\hat{A})$. Let φ_n be the corresponding density of $f_*^n(m \mid \hat{A})$, as given in Lemma 3.7. Then, cf. (15),

$$\varphi_j \leq \frac{1}{m(\hat{A})}\,\phi_j \leq \frac{C_0+1}{m(\hat{A})} \quad \text{for every } j \geq 0.$$

This implies that $\mu_{A,n}$ admits a density bounded by $(C_0+1)/m(\hat{A})$ for every $n \geq 1$. It follows from Lemma 2.5 that every accumulation point of the sequence $\mu_{A,n}$ is an absolutely continuous invariant measure. Let μ_A be any of these invariant measures.

The property of almost invariance of $\hat{A}$ we proved before implies that $f_*^j(m \mid \hat{A})(\hat{A}) \geq (m \mid \hat{A})(\hat{A}) = 1$ for every $j \geq 1$. This gives $\mu_{A,n}(\hat{A}) = 1$ for every $n \geq 1$, and so $\mu_A(\mathrm{clos}(\hat{A})) = 1$. On the other hand, by Proposition 3.11, we have $\mu_A = \theta_A m$ for some function θ_A with bounded variation. Since bounded variation functions have at most countably many discontinuity points, there exist some open interval $J \subset M$ and some $\delta > 0$ such that $\theta_A(x) > \delta$ for every $x \in J$. This implies that the restrictions of μ_A and m to J are equivalent measures. As the closure of $\hat{A}$ has full measure for μ, it must intersect J in a full Lebesgue measure subset. In particular, $\hat{A}$ has some point in J, and so $A \cap J$ has positive Lebesgue measure. Then, $\mu_A(A) \geq \delta m(A \cap J) > 0$.

Finally, let ν_A be the normalized restriction of μ_A to A:

$$\nu_A(B) = \frac{\mu_A(B \cap A)}{\mu_A(A)} \quad \text{for any measurable set } B.$$

Then $\nu_A(A) = 1$ and ν_A is absolutely continuous with respect to Lebesgue measure, since it is absolutely continuous with respect to μ_A. Moreover, ν_A is invariant under f, because A and μ_A are f-invariants. □

Corollary 3.13. *Every f-invariant set $A \subset M$ with positive Lebesgue measure has full Lebesgue measure in the neighbourhood of some singular point: there are $\varepsilon > 0$ and a singular point c of f such that $m([c-\varepsilon, c+\varepsilon] \setminus A) = 0$.*

Proof. Let ν be any absolutely continuous invariant measure so that $\nu(A) = 1$. By Proposition 3.11, $\nu = \theta m$ for some function θ with bounded variation. In particular, there exists an open interval $J \subset M$ such that the infimum of θ on J is strictly positive. Then, $\nu(J \setminus A) = 0$ implies $m(J \setminus A) = 0$.

Now we consider the iterates $f^n(J)$, $n \geq 1$, of the interval J. The expansivity condition (E2) in Definition 3.4 implies that

$$\text{length}(f^n(J)) \geq C\sigma^n \,\text{length}(J)$$

as long as $f^j(J)$ does not intersect the singular set of f, for any $0 \leq j \leq n-1$. So, since $\sigma > 1$ whereas the term on the left is bounded by the diameter of M, there must be a first time $N \geq 1$ such that $f^N(J)$ contains some singular point c. In particular, $f^N \mid J$ is a diffeomorphism onto its image. Then, $f^N(J)$ is an open interval and $m(f^N(J) \setminus A) = m(f^N(J \setminus A)) = 0$. □

The last step in the proof of Theorem 3.1 is

Corollary 3.14. *The map f has some ergodic absolutely continuous invariant probability measure, and the number of such measures does not exceed the number s of singular points of f. Moreover, their basins cover a full Lebesgue measure subset of M.*

Proof. It follows from Corollary 3.13 that there are at most s two-by-two disjoint f-invariant sets with positive Lebesgue measure. As a consequence, the manifold M can be partitioned into $r \leq s$ invariant sets $A_1, \ldots, A_r$, such that $m(A_i) > 0$ for every $1 \leq i \leq r$, and which are minimal: there is no invariant set $B_i \subset A_i$ with $0 < m(B_i) < m(A_i)$. As we have seen, for each $1 \leq i \leq r$ there exists an absolutely continuous invariant measure ν_i such that $\nu_i(A_i) = 1$. The fact that A_i is minimal implies that ν_i is ergodic.

Moreover, given any absolutely continuous invariant measure μ we may write $\mu = \sum_i \mu(A_i)\mu_i$, where the sum is over all the values of i such that $\mu(A_i) > 0$, and μ_i denotes the normalized restriction of μ to A_i. Since ν_i and μ_i are both ergodic, either they coincide or they are mutually singular. The second possibility would contradict the assumption that A_i is minimal, and so we must have $\mu_i = \nu_i$. This proves that $\nu_1, \ldots, \nu_r$ are precisely the ergodic absolutely continuous invariant measures of f.

Finally, let E be the complement of $B(\nu_1) \cup \cdots \cup B(\nu_r)$ in M. Then E is f invariant, and $\nu_i(E) = 0$ for every $1 \leq i \leq r$. As a consequence, E has zero measure with respect to any absolutely continuous invariant measure of f. By Lemma 3.12 the set E must have zero Lebesgue measure. □

Example 3.13. In particular, maps with a unique singular point have a unique absolutely continuous invariant probability measure, and it is is ergodic. This includes the tent maps and the Lorenz-like maps as in Figure 3.

On the other hand, piecewise expanding maps with more than one singular point may have several physical measures:

Example 3.14. Let $f : [0,1] \to [0,1]$ be given by

$$f(x) = \begin{cases} (1/2) - |1/2 - 2x| & \text{for } x \in [0,1/2] \\ 1 - |3/2 - 2x| & \text{for } x \in (1/2,1]. \end{cases}$$

Both $2\mathcal{X}_{[0,1/2]}m$ and $2\mathcal{X}_{[1/2,1]}m$ are invariant, ergodic, and absolutely continuous.

This shows that to ensure uniqueness one needs some assumption of indivisibility of the dynamics.

Definition 3.15. We say that a piecewise expanding map $f : M \to M$ is *transitive* if there exists some compact $I_* \subset M$ such that

(1) $f(I_*) \subset I_*$ and the orbit $f^n(x)$, $n \geq 0$, of Lebesgue almost every $x \in M$ intersects the interior of I_*;

(2) given any pair of intervals V_1 and V_2 contained in I_*, there exists $n \geq 0$ such that $f^n(V_1) \cap V_2$ has positive Lebesgue measure.

Corollary 3.15. *If f is transitive then it has a unique absolutely continuous invariant probability measure μ. Moreover, the support of μ coincides with I_*.*

Proof. As part of the proof of Corollary 3.14, we showed that any absolutely continuous invariant measure can be written as a linear combination of finitely many ergodic ones $\nu_1, \ldots, \nu_s$. Therefore, to obtain the first claim we only have to prove that there exists at most one of these ergodic measures.

Let $1 \leq i \leq s$. We also know, from Corollary 3.13, that there exists some interval U_i such that $m(U_i \setminus B(\nu_i)) = 0$. Transitivity implies that for Lebesgue almost any point $x \in U_i \cap B(\nu_i)$ there exists $n \geq 1$ such that $f^k(x)$ is in the interior of I_*. Take x such that neither iterate $f^j(x)$, $j \geq 0$, is a singular point. Then f^k is a local diffeomorphism near x. Hence, there exists some open neighbourhood $V_i \subset U_i$ of x, such that $f^k(V_i) \subset I_*$. Moreover,

$$m(f^k(V_i) \setminus B(\nu_i)) = m(f^k(V_i \setminus B(\nu_i))) = 0.$$

Consequently, $m(W_i \setminus B(\mu_i)) = 0$ where $W_i = \cup_{n=k}^{\infty} f^n(V_i)$. On the other hand, the second condition in Definition 3.15 implies that $W_i \cap W_j$ has positive Lebesgue measure, for any $1 \leq j \leq s$. In particular, $B(\nu_i) \cap B(\nu_j)$

must be non-empty, which implies that $\nu_i = \nu_j$. This proves uniqueness and ergodicity.

The proof of the second claim is similar. Since the density of μ has bounded variation, there exists an interval U such that $d\mu/dm > 0$ on U. Then, U is contained in the support of μ. As before, we can find $k \geq 1$ and an open interval $V \subset U$ such that $f^k(V) \subset I_*$. On the one hand, $W = \cup_{n=k}^{\infty} f^n(V)$ is contained in the support of μ. On the other, the properties in Definition 3.15 imply that W is a dense subset of I_*. Thus, I_* is contained in the support. Finally, since μ is ergodic and I_* is an invariant set with non-zero μ-measure, $\mu(M \setminus I_*) = 0$. As I_* is compact, this implies that no point outside I_* is in the support of μ. □

Let $f : I \to I$ be a tent map as in Figure 3, with c denoting the singular point. We already know that f has a unique absolutely continuous invariant measure μ. As an application of the previous corollary, we show that if the derivative Df is large enough then the support of μ coincides with the interval $I_* = [f^2(c), f(c)]$. Moreover, in that case the map is transitive. Neither of these conclusions is true, in general, if the derivative is only larger than 1.

Lemma 3.16. *If f is a tent map with $|Df(x)| \geq \sigma > \sqrt{2}$ then f is transitive, with $I_* = [f^2(c), f(c)]$ and a strong form of property (2): for any interval $J \subset I_*$ there exists $N \geq 1$ such that $f^N(J) = I_*$.*

Proof. The first condition in Definition 3.15 is easy: since $|Df| > 1$, no orbit $f^n(x)$ with $x \neq 0, 1$ can remain forever in $[0, f^2(c)] \cup [f(c), 1]$. We are left to prove the last statement in the lemma.

We claim that $f^n(J)$ must eventually contain the fixed point $p > c$ of f. Indeed, otherwise one would be able to construct a sequence of intervals

$$J = J_0 \supset J_1 \supset \cdots \supset J_n \supset \cdots$$

in the following way.

- If $f^{n-1}(J_{n-1})$ does not contain c, take $J_n = J_{n-1}$. Observe that in this case $m(f^n(J_n)) = \sigma\, m(f^{n-1}(J_{n-1}))$.

- If $f^{n-1}(J_{n-1})$ does contain c, then take $J_n \subset J_{n-1}$ such that $f^{n-1}(J_n)$ coincides with the largest of the two intervals

$$f^{n-1}(J_{n-1}) \cap [f^2(c), c] \quad \text{or} \quad f^{n-1}(J_{n-1}) \cap [c, f(c)].$$

  Then $m(f^n(J_n)) \geq (\sigma/2) m(f^{n-1}(J_{n-1}))$.

Moreover, in this last case $f^n(J_n)$ can not contain c, since we suppose that $p \notin f^{n+1}(J_n)$. This means that for the construction of J_{n+1} we fall in the first case: $J_{n+1} = J_n$. In particular,

$$m(f^{n+1}(J_{n+1})) \geq \sigma m(f^n(J_n)) \geq \frac{\sigma^2}{2} m(f^{n-1}(J_{n-1})).$$

As we suppose $\sigma^2 > 2$, it follows that the sequence $m(f^n(J_n))$ is unbounded, a contradiction. This proves that $p \in f^{n_1}(J_{n_1})$ for some $n_1 \geq 0$, as we claimed.

Then, $p \in f^n(J)$ for every $n \geq n_1$ and, by expansivity, $[c,p] \subset f^{n_2}(J)$ for some $n_2 \geq n_1$. It follows that $[f^3(c), f(c)] \subset f^{n_2+3}(J)$. On the other hand, it is easy to check that $f^3(c) < p$ for all $\sigma > \sqrt{2}$. Now there are two cases to consider. If $f^3(c) \leq c$, we get $[f^2(c), f(c)] \subset f^{n_2+4}(J)$, as we wanted to prove. Otherwise, there must be some odd number $k > 3$ such that $f^k(c) < c$ and $[f^k(c), f(c)] \subset f^{n_2+k}(J)$. Then $[f^2(c), f(c)] \subset f^{n_2+k+1}(J)$ and this also proves that I_* is contained in some iterate of J. □

Example 3.16. A similar argument applies if f is a Lorenz-like map with $|Df(x)| \geq \sigma > \sqrt{2}$ for all $x \neq c$. See [22]. Assuming $f(c^-) = 1$, $f(c^+) = 0$, as in Figure 3, then for any subinterval J of $I_* = [0,1]$ there exists $N \geq 1$ such that $f^N(J) = I_*$. Therefore, the map is transitive. Furthermore, the support of the absolutely continuous invariant measure is the whole interval $[0,1]$.

3.5 Maps With Infinitely Many Branches

In this subsection we prove Theorem 3.2, and describe a few useful applications. Throughout, f is a piecewise expanding map as in the theorem: there exists a partition $\mathcal{P}^1$ into regularity intervals of f, such that

$$\sum_{\eta \in \mathcal{P}^1} \operatorname{var} \hat{g}_\eta < \infty.$$

As was pointed out before, finiteness of regularity intervals intervened only in two occasions while we were proving Theorem 3.1: in Proposition 3.9 and in Corollary 3.14. Our first step is to extend the proposition to the present situation. For this we need the following lemma.

Let $\hat{g}_\xi$ be the weight functions introduced in (10), for each $\xi \in \mathcal{P}^1$. More generally, given any $i \geq 1$ and any interval J contained in an element η of

$\mathcal{P}^i$, we define $\hat{g}^i_J : M \to \mathbb{R}$ by

$$\hat{g}^i_J(x) = \frac{1}{|Df^i(x)|} \quad \text{if } x \in J \quad \text{and} \quad \hat{g}^i_J(x) = 0 \quad \text{otherwise.}$$

It is clear from the definitions that the variation of these functions depends monotonically on the domain: if $J \subset L$ then $\operatorname{var} \hat{g}^i_J \le \operatorname{var} \hat{g}^i_L$.

Lemma 3.17. *For every $n \ge 1$,*

$$\sum_{\eta \in \mathcal{P}^n} \operatorname{var} g^n_\eta \le \sum_{\eta \in \mathcal{P}^n} \operatorname{var} \hat{g}^n_\eta < \infty.$$

Proof. Using part 5 of Lemma 3.3, one checks easily that $\operatorname{var} g^n_\eta \le \operatorname{var} \hat{g}^n_\eta$ for every $\eta \in \mathcal{P}^n$. In fact, these variations can be different only if f^n maps η onto the whole M. Compare (11). The first inequality is a direct consequence.

Next, given $n \ge 1$ and any $\eta \in \mathcal{P}^{n+1}$,

$$\hat{g}^{n+1}_\eta = \big(\hat{g}_{f^n(\eta)} \circ f^n\big) \hat{g}^n_\eta.$$

So, using Lemma 3.3 once more,

$$\operatorname{var} \hat{g}^{n+1}_\eta \le \operatorname{var} \hat{g}_{f^n(\eta)} \, \sup \hat{g}^n_\eta + \sup \hat{g}_{f^n(\eta)} \, \operatorname{var} \hat{g}^n_\eta \,.$$

Observe that $\sup \hat{g}^n_\eta \le \operatorname{var} \hat{g}^n_\eta$, because $\inf \hat{g}^n_\eta = 0$, and analogously for $\hat{g}_{f^n(\eta)}$. Let $\xi \in \mathcal{P}^n$ and $\zeta \in \mathcal{P}^1$ be defined by $\eta \subset \xi$ and $f^n(\eta) \subset \zeta$. Then, the previous inequality can be replaced by

$$\operatorname{var} \hat{g}^{n+1}_\eta \le 2 \operatorname{var} \hat{g}_{f^n(\eta)} \, \operatorname{var} \hat{g}^n_\eta \le 2 \operatorname{var} \hat{g}_\zeta \, \operatorname{var} \hat{g}^n_\xi \,.$$

Since each pair $(\xi, \zeta) \in \mathcal{P}^n \times \mathcal{P}^1$ determines $\eta \in \mathcal{P}^{n+1}$ uniquely, we obtain

$$\sum_{\eta \in \mathcal{P}^{n+1}} \operatorname{var} \hat{g}^{n+1}_\eta \le 2 \sum_{\zeta \in \mathcal{P}^1} \operatorname{var} \hat{g}_\zeta \sum_{\xi \in \mathcal{P}^n} \operatorname{var} \hat{g}^n_\xi \,.$$

Now the claim in the lemma follows immediately, by induction on n. □

Now we can prove, following [**?**], that Proposition 3.9 remains valid in the generality of Theorem 3.2.

Proposition 3.18. *There are $C_0 > 0$ and $0 < \lambda_0 < 1$ such that, given any bounded variation function $\varphi : M \to \mathbb{R}$,*

$$\operatorname{var} \varphi_n \leq C_0 \lambda_0^n \operatorname{var} \varphi + C_0 \int |\varphi| \, dm$$

for any $n \geq 1$, where φ_n is as in Lemma 3.7.

Proof. Using Lemmas 3.7 and 3.3 we get, as in (13),

$$\operatorname{var} \varphi_n \leq \sum_\eta \operatorname{var} g_\eta^n \sup_\eta |\varphi| + \sup g_\eta^n (\operatorname{var}_\eta \varphi + 2 \sup |\varphi|),$$

recall that the sum is over the intervals $\eta \in \mathcal{P}^n$ that have positive Lebesgue measure. Moreover, $\sup g_\eta^n \leq \operatorname{var} g_\eta^n$ because $\inf g_\eta = 0$. So, the previous inequality may be replaced by

$$\operatorname{var} \varphi_n \leq \sum_\eta \operatorname{var} g_\eta^n (3 \sup_\eta |\varphi| + \operatorname{var}_\eta \varphi). \tag{16}$$

The terms involving variation pose no problem: using Lemma 3.8,

$$\sum_\eta \operatorname{var} g_\eta^n \operatorname{var}_\eta \varphi \leq C_1 \lambda_1^n \sum_\eta \operatorname{var}_\eta \varphi \leq C_1 \lambda_1^n \operatorname{var} \varphi$$

One would like to replace supremum by variation and integral in the remaining terms, using part 3 of Lemma 3.3, as we did before in (14). The problem is that the measure $m(\eta)$ of these intervals is no longer bounded away from zero.

To bypass this, we split the sum into two parts. Given any finite subset $\mathcal{Q}^n$ of $\mathcal{P}^n$, we may estimate the sum corresponding to the intervals $\eta \in \mathcal{Q}^n$ in the same way as in the finite case. On the other hand, using the summability in Lemma 3.17, we may choose $\mathcal{Q}^n$ in such a way that the total contribution of the remaining terms is much smaller than $\sup |\varphi| \leq \operatorname{var} \varphi + \int |\varphi| \, dm$.

More precisely, we begin by fixing a finite subset $\mathcal{Q}^n$ of $\mathcal{P}^n$ such that

$$\sum_{\eta \notin \mathcal{Q}^n} \operatorname{var} g_\eta^n \leq C_1 \lambda_1^n.$$

Then,

$$\begin{aligned}\sum_{\eta \notin \mathcal{Q}^n} 3 \operatorname{var} g_\eta^n \sup_\eta |\varphi| &\leq 3 \sup |\varphi| \sum_{\eta \notin \mathcal{Q}^n} \operatorname{var} g_\eta^n \leq 3 C_1 \lambda_1^n \sup |\varphi| \\ &\leq 3 C_1 \lambda_1^n \operatorname{var} \varphi + 3 C_1 \lambda_1^n \int |\varphi| \, dm.\end{aligned}$$

On the other hand, compare (14)

$$\sum_{\eta\in\mathcal{Q}^n} 3\operatorname{var} g_\eta^n \sup_\eta |\varphi| \le \sum_{\eta\in\mathcal{Q}^n} 3C_1\lambda_1^n\big(\operatorname{var}_\eta \varphi + \frac{1}{m(\eta)}\int_\eta |\varphi|\,dm\big)$$
$$\le 3C_1\lambda_1^n \operatorname{var}\varphi + K(n)\int |\varphi|\,dm,$$

with $K(n) = 3C_1\lambda_1^n \sup\{1/m(\eta) : \eta \in \mathcal{Q}^n\}$. So, 16 leads to

$$\operatorname{var}\varphi_n \le 7C_1\lambda_1^n \operatorname{var}\varphi + 2K(n)\int |\varphi|\,dm.$$

Now we only have to remove the dependence on n of the integral term, and this can be done in the same way as in the proof of Proposition 3.9. Fixing a large enough integer $N \ge 1$ so that $7C_1\lambda_1^N \le 1/2$, we get that it is enough to take $\lambda_0 \ge \max\{2^{-1/N}, \lambda_1\}$ and $C_0 \ge \max\{4K_0, 7C_1\}$, where $K_0 = \max\{K(n) : 1 \le n \le N\}$. □

Combining this proposition with Lemma 3.5 we immediately get the analog of Corollary 3.10:

Corollary 3.19. *The map f has some absolutely continuous invariant measure whose density has bounded variation.*

It would not be difficult to obtain the remaining claims in Theorem 3.2 at this point, but we do not do it right away. Instead, we postpone the proof to Subsection 4.2 where analogous facts are obtained in much more generality.

In the rest of the present section we indicate some applications of the previous results.

Definition 3.17. A piecewise expanding map $f : M \to M$ has *long branches* if there exist $\delta > 0$, $K > 0$, and a partition $\mathcal{P}^1$ into regularity intervals such that

(a) the restriction $f \mid \eta$ of f to each $\eta \in \mathcal{P}^1$ is C^2, and

$$\frac{|D^2(f \mid \eta)|}{|D(f \mid \eta)|^2} \le K \quad \text{for every } \eta \in \mathcal{P}^1.$$

(b) the image of every interval $\eta \in \mathcal{P}^1$ has Lebesgue measure $m(\eta) \ge \delta$.

The next proposition shows that if f has long branches then it satisfies the assumptions of Theorem 3.2.

Proposition 3.20. *If $\mathcal{P}^1$ satisfies (a) and (b) above, for some $\delta, K > 0$, then $\sum_{\eta \in \mathcal{P}^1} \operatorname{var} \hat{g}_\eta < \infty$. Therefore, f has some ergodic absolutely continuous invariant measure, and there are finitely many such measures.*

Proof. Let η be any element of $\mathcal{P}^1$. Condition (a) can be rewritten as

$$\left| D\left(\frac{1}{D(f \mid \eta)}\right)\right| \le K$$

which implies

$$\operatorname*{var}_\eta \frac{1}{|D(f \mid \eta)|} \le \operatorname*{var}_\eta \frac{1}{D(f \mid \eta)} = \int_\eta \left| D\left(\frac{1}{D(f \mid \eta)}\right)\right| \, dm \le K m(\eta).$$

On the other hand, by the mean value theorem there exists some $x_\eta \in \eta$ such that

$$\frac{1}{|Df(x_\eta)|} = \frac{m(\eta)}{m(f(\eta))} \le \frac{1}{\delta} m(\eta).$$

In particular,

$$\sup_\eta \frac{1}{|D(f \mid \eta)|} \le \operatorname*{var}_\eta \frac{1}{|D(f \mid \eta)|} + \inf_\eta \frac{1}{|D(f \mid \eta)|} \le \big(K + \frac{1}{\delta}\big) m(\eta).$$

Then,

$$\operatorname{var} \hat{g}_\eta \le \operatorname*{var}_\eta \frac{1}{|D(f \mid \eta)|} + 2 \sup_\eta \frac{1}{|D(f \mid \eta)|} \le \big(3K + \frac{2}{\delta}\big) m(\eta),$$

and so $\sum_{\eta \in \mathcal{P}^1} \operatorname{var} \hat{g}_\eta \le (3K + 2/\delta)$.

The last part of the statement is now a consequence of Theorem 3.2 □

Example 3.18. Let $f : [0,1] \to [0,1]$ be a map with a neutral fixed point as in Example 1.5. From f we construct a new map $\hat{f} : [0,1] \to (0,1]$, as follows. See Figure 5. Let $h_0 : (0, f(c)) \to (0, c)$ be the inverse of $f \mid (0, c)$. Then let $c_1 = c$ and $c_{j+1} = h_0(c_j)$, for each $j \ge 1$. Finally, define $\hat{f}$ by

$$\hat{f} \mid (c_1, 1] = f \mid (c_1, 1] \quad \text{and} \quad \hat{f} \mid (c_{j+1}, c_j) = f^j \mid (c_{j+1}, c_j) \ \text{ for each } j \ge 1.$$

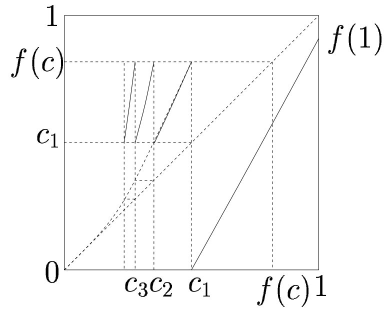

Figure 5: Inducing near a neutral fixed point

For completeness, we also set $\hat{f}(x) = f(c)$ for any $x \in \{0, \ldots, c_3, c_2, c_1\}$, although this is quite arbitrary. Clearly, $\hat{f}$ is C^2 in the interior of each element of

$$\mathcal{P}^1 = \big\{(c_{j+1}, c_j] : j \geq 1\big\} \cup \big\{(c_1, 1]\big\}.$$

Let $\sigma = \inf\{|Df(x)| : x \in (c_2, c_1) \cup (c_1, 1]\}$. Then $\sigma > 1$, and $|D\hat{f}(x)| \geq \sigma$ at every point x where the derivative is defined. Moreover,

$$\hat{f}\big((c_{j+1}, c_j]\big) = (c, f(c)] \quad \text{for} \quad j \geq 1 \quad \text{and} \quad \hat{f}\big((c_1, 1]\big) = (0, f(1)].$$

So, $\hat{f}$ satisfies condition (2) in Definition 3.17, with $\delta = \min\{|f(c) - c|, |f(1)|\}$. It is not difficult to check that $\hat{f}$ also satisfies condition (1), and we leave this as an exercise to the reader. Altogether, this shows that $\hat{f}$ is a piecewise expanding map with long branches, and so Theorem 3.2 applies to it.

4 Piecewise Expanding Maps in Higher Dimensions

The ergodic theory of piecewise expanding maps in higher dimensions is presently much less satisfactory than in the one-dimensional case, despite the progress attained over the last two decades.

The first existence results for absolutely continuous invariant measures (apart from the Markov case) appeared in [**?**], [27]. Other constructions were proposed e.g. in [7] and [20]. See this last paper for an account of results obtained in the eighties. More recently, the scope of these results was considerably extended in works such as [1], [**?**], [**?**], [**?**], [59].

In essentially all the cases, the authors consider some notion of variation for functions in higher dimensional domains, and prove a Lasota-Yorke type of inequality, as in Proposition 3.9. Ergodic and spectral properties of the system can then be deduced along the lines of the one-dimensional case.

We discuss this approach in the next subsection, sketching an application to a class of multidimensional piecewise expanding maps with long branches.

4.1 The Bounded Variation Approach

The theory of bounded variation functions on higher dimensional domains is presented in [19] and [16]. Here quote some main notions and facts that are more directly relevant for our purposes.

First, we give a definition of variation for functions on domains of $\mathbb{R}^d$, any $d \geq 1$. Instead of Definition 3.6, that depends strongly on the order structure of 1-dimensional manifolds, our starting point is Lemma 3.6.

Let U be an open subset in some Euclidean space $\mathbb{R}^d$, and $\varphi : U \to \mathbb{R}$ be a Lebesgue integrable function. Given a C^1 vector field $\omega : U \to \mathbb{R}^d$, we denote by $\operatorname{div}\omega$ its *divergent*. That is

$$\operatorname{div}\omega = \frac{\partial\omega_1}{\partial x_1} + \cdots + \frac{\partial\omega_d}{\partial x_d} \qquad \text{if} \qquad \omega = (\omega_1, \ldots, \omega_d).$$

Definition 4.1. The *variation* of $\varphi : U \to \mathbb{R}$ on U is

$$\operatorname*{var}_U \varphi = \sup \left\{ \Big| \int_U (\varphi \operatorname{div}\omega)\, dm \Big| : \omega \in C_0^1(U) \;\; \text{with} \;\; \sup \|\omega\| \leq 1 \right\}$$

where $C_0^1(U)$ is the space of C^1 vector fields $\omega : U \to \mathbb{R}^d$ whose support is a compact subset of U. A function φ has *bounded variation on* U if $\operatorname{var}_U \varphi < \infty$.

Clearly, the variation of a function depends only on its L^1 class, that is, functions that coincide Lebesgue almost everywhere have the same variation. The space of L^1 classes with bounded variation on U is denoted $\mathrm{BV}(U)$.

The next proposition provides a useful criterium for deciding whether a function $\varphi : U \to \mathbb{R}$ has bounded variation by, basically, reducing the problem to dimension 1. Assume that φ has compact support. Then it may be extended to the whole $\mathbb{R}^d$, with the same support and $\operatorname{var}_{\mathbb{R}^d} \varphi = \operatorname{var}_U \varphi$. Hence we may just as well take U to coincide with $\mathbb{R}^d$. Given $1 \leq i \leq d$ and $\hat{x} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$, let

$$\varphi_{i,\hat{x}} : \mathbb{R} \to \mathbb{R} \quad \text{with} \quad \varphi_{i,\hat{x}}(x) = \varphi(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_d).$$

We represent by $\operatorname{var}_{[a,b]}[\varphi_{i,\hat{x}}]$ the variation of the L^1 class of $\varphi_{i,\hat{x}}$ over any compact interval $[a,b]$, recall Definition 3.8. Moreover, m_{d-1} denotes Lebesgue measure in $\mathbb{R}^{d-1}$.

Proposition 4.1. *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be an integrable function with compact support. Then $\hat{x} \mapsto \operatorname{var}_{[a,b]}[\varphi_{i,\hat{x}}]$ is measurable, for every $1 \le i \le d$ and any real numbers $a < b$. Moreover, φ has bounded variation in $\mathbb{R}^d$ if and only if*

$$\int_K \operatorname*{var}_{[a,b]}[\varphi_{i,\hat{x}}]\, dm_{d-1}(\hat{x}) < \infty$$

for every $1 \le i \le s$, every $a < b$, and any compact subset K of $\mathbb{R}^{d-1}$.

For a proof see Lemma 1 and Theorem 2 in [16, Section 5.10].

Example 4.2. Let $\alpha > 0$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ be defined by

$$\varphi(x) = \|x\|^{-\alpha} \quad \text{if} \quad 0 < \|x\| \le 1 \quad \text{and} \quad \varphi(x) = 0 \quad \text{otherwise.}$$

Then φ has bounded variation in $\mathbb{R}^d$ if and only if $\alpha < d - 1$. Indeed, for any $1 \le i \le d$ and $\hat{x} \ne 0$, the function $\varphi_{i,\hat{x}}$ is monotone increasing on $(-\infty, 0]$ and monotone decreasing on $[0, +\infty)$. So, given any $a < -1 < 1 < b$,

$$\operatorname*{var}_{[a,b]}[\varphi_{i,\hat{x}}] = 2\,\varphi_{i,\hat{x}}(0) = 2\,\|\hat{x}\|^{-\alpha}.$$

Now it suffices to note that, if K is the closed unit ball in $\mathbb{R}^{d-1}$,

$$\int_K 2\,\|\hat{x}\|^{-\alpha}\, dm_{d-1}(\hat{x}) < \infty \quad \text{if and only if} \quad \alpha < d - 1.$$

This example shows that bounded variation functions in dimension higher than 1 need not be bounded. On the other hand, the next proposition (Sobolev's inequality, see Theorem 1.28 of [19]) ensures that functions with bounded variation on a d-dimensional domain are in an L^p space, where p is determined by the dimension d. It also follows from Example 4.2 that the expression of p in the proposition can not be improved.

Proposition 4.2. *There exists $C(d) > 0$, depending only on d, such that, for any bounded variation function $\varphi : U \to \mathbb{R}$ with compact support, we have*

$$\Big(\int_U |\varphi|^p\, dm\Big)^{1/p} \le C(d)\, \operatorname*{var}_U \varphi \qquad \textit{where} \quad p = \frac{d}{d-1}.$$

In particular, $\mathrm{BV}(U) \subset L^p(U, m)$.

Proposition 4.3. *If $\varphi_n : U \to \mathbb{R}$, $n \geq 1$, are bounded variation functions converging to $\varphi : U \to \mathbb{R}$ in $L^1(U, m)$ then*

$$\operatorname*{var}_U \varphi \leq \liminf_n \operatorname*{var}_U \varphi_n$$

In other words, the variation is lower semi-continuous, with respect to the L^1 norm. Proposition 4.3 is contained in Theorem 1.9 of [19]. It is not difficult to deduce, cf. Remark 1.12 in [19], that the expression

$$\|\varphi\|_{\mathrm{BV}} = \operatorname*{var}_U \varphi + \|\varphi\|_1 \tag{17}$$

defines a complete norm in the space $\mathrm{BV}(U)$.

Also related to Proposition 4.3, we have the following important extension of Lemma 3.5, for domains with Lipschitz boundary in any dimension: sets of functions with uniformly bounded variation are relatively compact with respect to the L^1 norm.

Proposition 4.4. *Suppose the domain U is bounded, and its boundary is Lipschitz continuous. Let $K_1, K_2 > 0$ and φ_n, $n \geq 1$, be a sequence in $L^1(U, m)$ such that $\|\varphi_n\|_1 \leq K_1$ and $\operatorname{var}_U \varphi_n \leq K_2$ for every $n \geq 1$. Then there exists a subsequence $(n_k)_k$ such that $(\varphi_{n_k})_k$ converges in $L^1(U, m)$ to a function φ_0 with $\|\varphi_0\|_1 \leq K_1$ and $\operatorname{var}_U \varphi_0 \leq K_2$.*

A proof of this last fact can be found in Theorem 1.19 of [19]. The bounds on $\|\varphi_0\|_1$ and $\operatorname{var}_U \varphi_0$ follow from the L^1 convergence and Proposition 4.3.

These results show that some of the main tools from the one-dimensional case remain valid for bounded variation functions in any dimension. On the other hand, this notion of variation is very sensitive to the geometry of the domain. For instance, cf. Example 1.4 in [19], if $E \subset U$ is a compact domain bounded by a C^2 hypersurface, then

$$\operatorname*{var}_U \mathcal{X}_E = m_{d-1}(\partial E), \tag{18}$$

where $\mathcal{X}_E$ is the characteristic function $\mathcal{X}_E$ of E and $m_{d-1}(\partial E)$ denotes the $(d-1)$-dimensional Hausdorff measure of the boundary of E. So, even characteristic functions of open sets may have unbounded variation.

This observation is at the origin of serious difficulties one encounters in higher dimensions. Not surprisingly, the cases of high dimensional piecewise expanding maps one has been able to treat depend on restrictive conditions on the geometry of the domains of smoothness of the map, e.g. their boundaries should be fairly regular. At least in some cases, see the discussion in [48], such conditions are an artifact of this method, and not really necessary for the existence of absolutely continuous invariant measures. Alternative notions of variation have been proposed, but they also require technical restrictions on the shapes of the smoothness domains.

Maps with Long Branches Here we outline an application of the previous ideas to a class of piecewise expanding maps in any dimension that generalizes the one-dimensional maps with long branches treated in Proposition 3.20. This is due to [20], for maps with finitely many regularity domains, and [**?**] in the general case.

Let R be a bounded region in $\mathbb{R}^d$. We say that $f : R \to R$ is a C^2 *piecewise expanding map* if there is a partition $\mathcal{P}^1$ of R into domains η such that

(E1) the boundary of η is piecewise C^2 and has finite $(d-1)$-dimensional Hausdorff measure

(E2) the restriction of f to the interior of η is a C^2 diffeomorphism onto its image, and it admits a C^2 extension to the closure of η;

(E3) there is $\sigma > 1$ such that $\|Df(x)^{-1}\| \le \sigma^{-1}$ for every point x where the derivative is defined.

We say that a C^2 piecewise expanding map $f : R \to R$ has *long branches* if it satisfies two additional properties, (D) and (G), resemblant of parts (a) and (b) of Definition 3.17. The first one is a condition of bounded distortion:

(D) There is some $K > 0$ such that

$$\frac{\|D(Jf_\eta^{-1})\|}{|Jf_\eta^{-1}|} \le K, \quad \text{for every } \eta \in \mathcal{P}^1,$$

where $Jf_\eta^{-1} = \det D(f \mid \eta)^{-1}$ is the Jacobian of the inverse of $(f \mid \eta)$.

(G) is a geometric requirement on the images $f(\eta)$ of the regularity domains: they should not be too small (sizes uniformly bounded away from zero), and the angles at the border corners should also be bounded from below. More precisely, we suppose that

(G) There are $\alpha > 0$, $\delta > 0$, and for each $\eta \in \mathcal{P}^1$ there is a C^1 unitary vector field H_η on the boundary of $f(\eta)$, such that

1. $|\sin \operatorname{angle}(v, H_\eta(x))| \geq \alpha$ for every $x \in \partial f(\eta)$ and any vector v tangent to $\partial f(\eta)$ at x;
2. the segments $[x, x + \delta H_\eta(x)]$, $x \in \partial f(\eta)$, are two-by-two disjoint and their union is a neighbourhood of the boundary in $f(\eta)$.

By convention, a vector field is C^1 on the boundary of $f(\eta)$ if its restriction to each smooth component of the boundary is C^1 on that component. Moreover, the tangent space of $f(\eta)$ at a corner point is the union of the tangent spaces of all the smooth components that contain that point.

Theorem 4.5. *Let $f : R \to R$ be a C^2 piecewise expanding map with long branches, i.e., f satisfies (E1)–(E3), (D), (G). Assume that*

$$\sigma > 1 + \alpha^{-1}.$$

Then f has some invariant probability measure absolutely continuous with respect to Lebesgue measure in R.

Let m be the d-dimensional Lebesgue measure on $\mathbb{R}^d$, normalized so that $m(R) = 1$. As in Lemma 3.7, given any integrable function $\varphi : R \to \mathbb{R}$ and any $n \geq 1$, there exists $\varphi_n : R \to \mathbb{R}$ such that $f_*^n(\varphi m) = \varphi_n m$. In fact, we may take

$$\varphi_n = \sum_\eta g_\eta^n \cdot \left(\varphi \circ (f^n \mid \eta)^{-1}\right) \tag{19}$$

where the sum is over all the regularity domains η of f^n with $m(\eta) > 0$, and

$$g_\eta^n(y) = \frac{1}{|\det Df^n|} \circ (f^n \mid \eta)^{-1}(y) = Jf_\eta^{-n}(y) \qquad \text{if } y \in f^n(\eta),$$

with $g_\eta^n(y) = 0$ otherwise. Moreover, as in (12),

$$\int_R |\varphi_n| \, dm \leq \int_R |\varphi| \, dm \quad \text{for every } n \geq 1. \tag{20}$$

The main step in the proof of Theorem 4.5 is the following version of Proposition 3.9:

Proposition 4.6. *Suppose f is a C^2 piecewise expanding map with long branches. Then there exists $C_0 > 0$ such that*

$$\operatorname*{var}_R \varphi_1 \le \lambda \operatorname*{var}_R \varphi + C_0 \int_R |\varphi| \, dm, \qquad \lambda = \sigma^{-1}(1 + \alpha^{-1}),$$

for every $\varphi \in \mathrm{BV}(R)$.

Then, an absolutely continuous invariant measure for f can be found as follows. Fix $\varphi \equiv 1$. By Proposition 4.6 and (20),

$$\operatorname*{var}_R \varphi_n \le \lambda \operatorname*{var}_R \varphi_{n-1} + C_0 \int_R |\varphi_{n-1}| \, dm \le \lambda \operatorname*{var}_R \varphi_{n-1} + C_0$$

for every $n \ge 1$. Recall that we are assuming $\lambda < 1$. It follows that the sequence $(\varphi_n)_n$ has uniformly bounded variation: for every $n \ge 1$,

$$\operatorname*{var}_R \varphi_n \le \lambda^n \operatorname*{var}_R \varphi + C_0\big(1 + \cdots + \lambda^{n-1}\big) \le \frac{C_0}{1-\lambda}$$

Therefore, the variation of the sequence $\psi_n = n^{-1} \sum_{j=0}^{n-1} \varphi_j$ is also uniformly bounded, by the same constant. Using Proposition 4.4, we conclude that $(\psi_n)_n$ has some subsequence $(\psi_{n_k})_k$ converging in $L^1(R, m)$ to a a bounded variation function ψ_0. In particular the sequence of measures

$$\frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j m = \psi_{n_k} m$$

converges to $\mu = \psi_0 \, m$, an invariant absolutely continuous measure. This completes our sketch of the proof of Theorem 4.5.

4.2 Finiteness of Physical Measures

In this subsection we prove that, quite in general, a Lasota-Yorke type of inequality suffices to ensure that there are finitely many ergodic absolutely continuous invariant measures. We also need the map to be piecewise regular, but not necessarily piecewise expanding. The result applies to maps in any dimension and with any number of regularity domains, including the situations in Subsection 3.5 and in Theorem 4.5 as special cases. The main idea in the proof is the approximation argument used by [30] to obtain Proposition 3.11.

Let M be a compact domain in $\mathbb{R}^d$, whose boundary is contained in a finite union of C^2 hypersurfaces. We assume that $f : M \to M$ satisfies

(H1) there exists an open set $U \subset M$ such that $M \setminus U$ has zero Lebesgue measure and f is a local C^1 diffeomorphism at every point of U;

(H2) there exist constants $C_0 > 0$ and $0 < \lambda_0 < 1$ such that for any bounded variation function $\varphi : M \to \mathbb{R}$ we have $f_*^n(\varphi m) = \varphi_n m$ for some function $\varphi_n : M \to \mathbb{R}$ with

$$\operatorname*{var}_M \varphi_n \le C_0 \lambda_0^n \operatorname*{var}_M \varphi + C_0 \int |\varphi| \, dm.$$

Theorem 4.7. *If f satisfies (H1) and (H2), then it admits ergodic absolutely continuous invariant probability measures $\nu_1, \ldots, \nu_s$ such that*

1. *the union of the basins of $\nu_1, \ldots, \nu_s$ has full Lebesgue measure in M;*
2. *every absolutely continuous invariant measure μ is a linear combination of $\nu_1, \ldots, \nu_s$;*

Moreover, the density $d\mu/dm$ of any such measure μ has bounded variation.

The main step is the following proposition.

Proposition 4.8. *Given any function $\psi : M \to \mathbb{R}$ with $\int |\psi| \, dm = 1$, there exists a subsequence $(n_k)_k$ and a function $\theta : M \to \mathbb{R}$ with $\operatorname{var}_M \theta \le 4C_0$, such that*

$$\frac{d}{dm}\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j(\psi m)\right) \to \theta \quad \textit{in} \quad L^1(M, m).$$

Proof. Let $(\xi_l)_l$ be some sequence of bounded variation functions converging to ψ in $L^1(M, m)$. We may suppose that every ξ_l has L^1-norm less than 2. For each $n \ge 1$ and $l \ge 1$, let

$$f_*^n(\psi m) = \psi_n m \quad \text{and} \quad f_*^n(\xi_l m) = \xi_{l,n} m$$

where ψ_n and $\xi_{l,n}$ are obtained from ψ and ξ_l, respectively, as in (19). Let us fix l for a while. Assumption (H2) implies that

$$\operatorname*{var}_M \xi_{l,n} \le C_0 \lambda_0^n \operatorname*{var}_M \xi_l + C_0 \int |\xi_l| \, dm \le 3\, C_0$$

for every large enough n. So, increasing n if necessary,

$$\operatorname*{var}_M\left(\frac{1}{n}\sum_{j=0}^{n-1}\xi_{l,j}\right) \le \frac{1}{n}\sum_{j=0}^{n-1}\operatorname*{var}_M \xi_{l,j} \le 4C_0.$$

It follows from Proposition 4.4 that there exists a function θ_l and a sequence $(m(l,i))_i \to \infty$ such that

$$\frac{1}{m(l,i)}\sum_{j=0}^{m(l,i)-1}\xi_{l,j} \to \theta_l$$

as $i \to \infty$. Moreover, by Proposition 4.3, we have $\operatorname{var}\theta_l \le 4C_0$.

Then, using Proposition 4.4 for the sequence θ_l, we conclude that there exists a subsequence $(l_k)_k$ such that θ_{l_k} converges in $L^1(M,m)$ to some function θ with $\operatorname{var}_M\theta \le 4C_0$. It follows, by a triangular inequality argument, that there exists a subsequence $n_k = m(l_k, i_k)$, $k \ge 1$, such that

$$\frac{1}{n_k}\sum_{j=0}^{n_k-1}\xi_{l_k,j} \to \theta$$

in $L^1(M,m)$ as $k \to \infty$. On the other hand, as $\|\xi_{l,j} - \psi_j\|_1 \le \|\xi_l - \psi\|_1$ for every j, l,

$$\Big\|\frac{1}{n_k}\sum_{j=0}^{n_k-1}(\xi_{l_k,j} - \psi_j)\Big\|_1 \le \frac{1}{n_k}\sum_{j=0}^{n_k-1}\|\xi_{l_k} - \psi\|_1 = \|\xi_{l_k} - \psi\|_1\,,$$

and the last term goes to zero as $k \to \infty$. This implies that

$$\frac{1}{n_k}\sum_{j=0}^{n_k-1}\psi_j \to \theta \quad \text{in} \quad L^1(M,m),$$

as claimed. □

Corollary 4.9. *Any absolutely continuous probability measure μ of a piecewise expanding map can be written $\mu = \theta\, m$ where θ has $\operatorname{var}_M\theta \le 4C_0$.*

Proof. By assumption $\mu = \psi m$ for some $\psi \in L^1(M, m)$ with $\psi \geq 0$ and $\int \psi \, dm = 1$. Proposition 4.8 states that a subsequence

$$\frac{1}{n_k} \sum_{j=0}^{n_k - 1} \psi_j$$

converges in $L^1(M, m)$ to some function θ whose variation is bounded by $4C_0$. Now, $\psi_n = \psi$ for every n because μ is invariant. This implies that $\psi = \theta$. □

Lemma 4.10. *Given any f-invariant set $A \subset M$ with positive Lebesgue measure, there exists some absolutely continuous f-invariant probability measure ν_A such that $\nu_A(A) = 1$.*

Proof. Let $(m \mid A)$ represent the normalized restriction of Lebesgue measure to A. In other words, $(m \mid A) = \psi m$ where $\psi = \mathcal{X}_A / m(A)$. Let us consider the sequence of probability measures

$$\mu_{A,n} = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j (m \mid A).$$

By Proposition 4.8, there exists a function θ_A with $\operatorname{var}_M \theta_A \leq 4C_0$ such that some subsequence

$$\frac{d\mu_{A,n_k}}{dm} \to \theta_A \quad \text{in} \quad L^1(m),$$

as $k \to \infty$. By Remark 3.11, the sequence μ_{A,n_k} converges to $\nu_A = \theta_A \, m$ in the weak* sense. Moreover, ν_A is an absolutely continuous invariant measure for f. Since A is assumed to be invariant, $f_*^j(m \mid A)(A) = (m \mid A)(A) = 1$ for every $j \geq 1$. This gives $\mu_{A,n}(A) = 1$ for every $n \geq 1$. Finally, L^1 convergence of the densities implies that

$$\nu_A(A) = \int_A \theta_A \, dm = \lim_k \int_A \frac{d\mu_{A,n_k}}{dm} \, dm = \lim_k \mu_{A,n_k}(A) = 1.$$

So, ν_A does satisfy the conclusion of the lemma. □

It is worth pointing out that the argument in Lemma 3.12 does not carry on to higher dimensions. This is because the support of functions with bounded variation may have empty interior, see the following example.

Example 4.3. Let U be an open subset of $\mathbb{R}^d$ containing the closed unit ball $B_1(0)$ around the origin. Let $\{q_n : n \in \mathbb{N}\}$ be a countable dense subset of $B_1(0)$, and B_n be the open ball of radius 2^{-n-1} around each q_n. Define $E_n = B_1(0) \setminus \cup_{j=1}^n B_j$ and $E = B_1(0) \setminus \cup_{j=1}^\infty B_j$. Note that E is non-empty, in fact it has positive d-dimensional Lebesgue measure:

$$m_d(E) \geq \sigma(d) - \sum_{j=1}^{\infty} 2^{-(j+1)d} \sigma(d) \geq \frac{\sigma(d)}{2},$$

where $\sigma(d)$ is the Lebesgue measure of the unit ball in $\mathbb{R}^d$. Clearly, E, E_n are compact, and E is nowhere dense. According to (18)

$$\operatorname{var} \mathcal{X}_{E_n} \leq \omega(d) + \sum_{j=1}^{n} 2^{-jd} \omega(d) \leq 2\omega(d).$$

for every $n \geq 1$, with $\omega(d)$ denoting the $(d-1)$-dimensional measure of the unit sphere in $\mathbb{R}^d$. Then, $\operatorname{var} \mathcal{X}_E \leq \liminf_n \operatorname{var} \mathcal{X}_{E_n} \leq 2\omega(d)$, as a consequence of Proposition 4.3.

Corollary 4.11. *If $A \subset M$ is an f-invariant set with positive Lebesgue measure then $m(A) \geq (4C_0C(d))^{-1/d}$.*

Proof. Let ν_A be an absolutely continuous invariant measure giving full weight to A, as in Lemma 4.10, and $\theta_A = d\nu_A/dm$. Let $p = d/(d-1)$ and $q = 1/d$, with $p = \infty$ in the case $d = 1$. By Sobolev's inequality Propositionp.vsobolev,

$$\|\theta_A\|_p \leq C(d) \operatorname*{var}_M \theta_A \leq 4C_0C(d).$$

Combining this with Hölder's inequality we get

$$1 = \|\theta_A\|_1 \leq \|\theta_A\|_p \, m(A)^{1/q} \leq 4C_0C(d)\, m(A)^d,$$

as we claimed. □

Corollary 4.11 implies that there are finitely many two-by-two disjoint f-invariant sets with positive Lebesgue measure. As an immediate consequence, M can be partitioned into finitely many minimal f-invariant sets:

Corollary 4.12. *There exist f-invariant sets $A_1, \ldots, A_s$ such that*

1. *$m(A_i) > 0$ for every $1 \leq i \leq s$, and $M = A_1 \cup \cdots \cup A_s$;*

2. there are no f-invariant sets $B_i \subset A_i$ with $0 < m(B_i) < m(A_i)$.

Finally, let $\nu_1, \ldots, \nu_s$ be absolutely continuous invariant measures with $\nu_i(A_i) = 1$ for $i = 1, \ldots, s$, as in Lemma 4.10. The fact that A_i is minimal implies that ν_i is ergodic and $B(\nu_i)$ has full measure in A_i. Moreover, any absolutely continuous invariant measure μ can be written as $\mu = \sum_i \mu(A_i)\nu_i$. To see this, write $\mu = \sum_i \mu(A_i)\mu_i$, where the sum is over the values of i such that $\mu(A_i) > 0$ and μ_i is the normalized restriction of μ to A_i. Each μ_i is also an ergodic measure, because A_i is minimal. Consequently, either $\mu_i = \nu_i$ or there exists some invariant subset $B \subset A_i$ with $\mu_i(B) = 0$ and $\nu_i(B) = 1$. The last case would imply $0 < m(B) < m(A_i)$, contradicting the minimality. So, we must have $\mu_i = \nu_i$ for every i.

This shows that these measures ν_i satisfy all the conclusions of Theorem 4.7. We have finished the proof of the theorem.

5 Hyperbolic Sets

Here we recall the definitions and some basic facts about hyperbolic sets of diffeomorphisms, specially attractors. Proofs may be found in [41], [33], [56], [?].

5.1 Definitions and Examples

The next definition involves the notion of *splitting* $E^1 \oplus E^2$ of the tangent space $T_\Lambda M$ of the manifold M over a subset Λ. By that we understand a map $x \mapsto (E^1_x, E^2_x)$ associating to each point $x \in \Lambda$ two complementary subspaces of the tangent space $T_x M$. We always assume that the subspaces E^1_x have constant dimension at every point $x \in \Lambda$, then the same is true for the E^2_x. We call the splitting *continuous* if given any $p \in \Lambda$ there exist continuous vector fields $X_1, \ldots, X_u, Y_1, \ldots, Y_s$ on a neighbourhood $U_p \subset \Lambda$ of p, linearly independent at every point and such that E^1_x is the subspace generated by $X_1(x), \ldots, X_u(x)$ and E^2_x is the subspace generated by $Y_1(x), \ldots, Y_s(x)$, for every $x \in U_p$.

Definition 5.1. Let $f : M \to M$ be a C^1 diffeomorphism and Λ be a compact subset of M that is invariant under f, that is, $f(\Lambda) = \Lambda$. We say that $\Lambda \subset M$ is a *(uniformly) hyperbolic set* for f if there exists a continuous splitting $T_\Lambda M = E^u \oplus E^s$ of the tangent space M over Λ such that

1. the splitting is *invariant* under the derivative Df: for every $x \in \Lambda$

$$Df(x)^{-1} \cdot E^u_x = E^u_{f^{-1}(x)} \quad \text{and} \quad Df(x) \cdot E^s_x = E^s_{f(x)};$$

2. the subbundle E^u is *expanding* and the subbundle E^s is *contracting* for Df: there are constants $C > 0$ and $\lambda < 1$ so that

$$\|Df^{-n}(x)|E^u_x\| \le C\lambda^n \quad \text{and} \quad \|Df^n(x)|E^s_x\| \le C\lambda^n$$

   for every $x \in \Lambda$ and every $n \ge 1$.

Clearly, this last condition is independent of the choice of the Riemannian norm $\|\cdot\|$ on M, that affects only the value of the constant $C > 0$. According to the next lemma, we can always find a Riemannian norm on the manifold M for which $C = 1$. Such a norm is said to be *adapted* to f on Λ.

Proposition 5.1. *Let Λ be a hyperbolic set for a diffeomorphism f. Then, given any $\lambda_* \in (\lambda, 1)$, there exists a Riemannian norm $\|\cdot\|_*$ on M such that*

$$\|Df(x)^{-1}v^u\|_* \le \lambda_*\|v^u\|_* \quad \textit{and} \quad \|Df(x)v^s\|_* \le \lambda_*\|v^s\|_*$$

for every $v^u \in E^u_x$, $v^s \in E^s_x$, and $x \in \Lambda$.

Proof. Fix $\lambda < \lambda_+ < \lambda_*$ and $N \ge 1$ large enough so that $C(\lambda/\lambda_+)^N < 1$. Given any vector $v = v^u + v^s$ in $E^u_x \oplus E^s_x$, define

$$\|v\|_+^2 = \|v^u\|_+^2 + \|v^s\|_+^2 \,,$$

with

$$\|v^u\|_+^2 = \sum_{j=0}^{N-1} \lambda_+^{-2j}\|Df^{-j}(x)v^u\|^2 \quad \text{and} \quad \|v^s\|_+^2 = \sum_{j=0}^{N-1} \lambda_+^{-2j}\|Df^{j}(x)v^s\|^2.$$

It is easy to see that this defines a continuous norm on Λ, and

$$\|Df(x)^{-1}v^u\|_+ \le \lambda_+\|v^u\|_+ \quad \text{and} \quad \|Df(x)v^s\|_+ \le \lambda_+\|v^s\|_+ \tag{21}$$

for any $v^u \in E^u_x$, $v^s \in E^s_x$, and $x \in \Lambda$. In general, the subbundles E^s and E^u do not admit smooth extensions to a neighbourhood of Λ, and so $\|\cdot\|_+$ may fail to extend to a smooth norm on M. However, this can be easily solved. Let $\|\cdot\|_*$ be any C^∞ Riemannian norm on M whose restriction to Λ is uniformly close to $\|\cdot\|_+$: if the two norms are close enough then (21) remains valid with $\|\cdot\|_*$ in the place of $\|\cdot\|_+$, and λ_* in the place of λ_+. $\square$

Example 5.2. (Linear Anosov maps) Let $A \in \mathrm{Sl}(d, \mathbb{Z})$, that is, A is a linear isomorphism of $\mathbb{R}^d$, $d \geq 2$, with determinant equal to ± 1 and whose matrix relative to the canonical basis of $\mathbb{R}^d$ has integer coefficients. Then A preserves the lattice $\mathbb{Z}^d$, and so there exists a unique smooth map f from the d-dimensional torus $M = \mathbb{R}^d/\mathbb{Z}^d$ to itself satisfying $\pi \circ A = f \circ \pi$, where $\pi : \mathbb{R}^d \to M$ is the canonical projection. Besides, f is a diffeomorphism: its inverse may be obtained through the same construction, with A^{-1} in the place of A.

Now, suppose the isomorphism A is *hyperbolic*: all its eigenvalues have norm different from 1. For example, this is the case for the 2-dimensional isomorphism

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Let $\hat{E}^u$ and $\hat{E}^s$ be the direct sum of the (generalized) eigenspaces of A corresponding to the eigenvalues with norm larger than 1 and smaller than 1, respectively. Given any point $w \in M$, choose $z \in \mathbb{R}^d$ such that $\pi(z) = w$, and then let

$$E^u_w = D\pi(z) \cdot \hat{E}^u \quad \text{and} \quad E^s_w = D\pi(z) \cdot \hat{E}^s.$$

These objects do not depend on the choice of z, and so this defines subbundles $E^u = (E^u_w)_{w \in M}$ and $E^s = (E^s_w)_{w \in M}$ of the tangent space of M. Moreover, $E^u \oplus E^s$ is a hyperbolic splitting for f: the derivative Df leaves both E^u and E^s invariant, while expanding the vectors in E^u and contracting the vectors in E^s.

Indeed, let $0 < \lambda < 1$ be fixed close enough to 1 so that no eigenvalue of A has norm in the interval $[\lambda, \lambda^{-1}]$. Let $\| \cdot \|_e$ be any norm in $\mathbb{R}^d$, and endow T^d with the Riemannian metric $\| \cdot \|$ defined by $\|D\pi(z)v\| = \|v\|_e$ for every $z \in \mathbb{R}^d$ and $v \in \mathbb{R}^d$. Then

$$\|Df^{-n} \mid E^u_w\| = \|A^{-n} \mid \hat{E}^u\|_e \quad \text{and} \quad \|Df^n(w) \mid E^s_w\| = \|A^n \mid \hat{E}^s\|_e$$

are less than $C\lambda^n$ for all $w \in M$ and $n \geq 1$, as long as the constant C is fixed sufficiently large. This proves that the ambient manifold $\Lambda = M$ is a hyperbolic set for f.

A fundamental property of hyperbolic sets is that they are a *robust* feature of the system: if a diffeomorphism f has a hyperbolic set Λ then any other diffeomorphism g in a C^1 neighbourhood has a hyperbolic set Λ_g, close to Λ.

Furthermore, the dynamics of g on Λ_g is topologically equivalent (conjugate) to that of f on Λ. That is the content of the next theorem.

Theorem 5.2. *Let Λ be a hyperbolic set for a diffeomorphism $f : M \to M$. Then there is a neighborhood $\mathcal{N}$ of f in $\mathrm{Diff}^1(M)$, and there is a continuous map $\phi : \mathcal{N} \longrightarrow \mathrm{Emb}(\Lambda, M)$ such that $\phi(f)$ is the inclusion map of Λ in M and $\Lambda_g = \phi(g)(\Lambda)$ is a hyperbolic set for every $g \in \mathcal{N}$. Moreover,*

$$\phi(g) \circ (f \mid \Lambda) = (g \mid \Lambda_g) \circ \phi(g).$$

$\mathrm{Emb}(\Lambda, M)$ denotes the space of continuous one-to-one maps from Λ to M, endowed with the topology of uniform convergence. For $r \geq 1$, $\mathrm{Diff}^r(M)$ is the space of C^r diffeomorphisms on M, with the C^r topology. We call Λ_g the *hyperbolic continuation* of Λ for g. Similarly, we call $\tilde{x} = \phi(g)(x)$ the *hyperbolic continuation* of $x \in \Lambda$ for g.

Example 5.2 is somewhat special in that we were able to exhibit the hyperbolic splitting explicitly, which is hardly ever possible. Fortunately, in order to prove that an invariant set is hyperbolic it suffices to have some reasonable approximation of the invariant subbundles E^u and E^s. The precise formulation uses the notion of stable and unstable cone fields.

Let $E \oplus F = T_K M$ be a splitting of the tangent space $T_\Lambda M$ over some subset $K \subset M$. Given $a > 0$, the *cone field of width a around E* is the family $C_a(E) = \big(C_a(E, x)\big)_{x \in K}$ defined by

$$C_a(E, x) = \{v_1 + v_2 \in E \oplus F : \|v_2\| \leq a \leq \|v_1\|\}.$$

Definition 5.3. $C_a(E)$ is an *unstable cone field* for f on K if

1. $C_a(E)$ is *forward invariant*: there exists $\theta < 1$ such that
$$Df(x) \cdot C_a(E, x) \subset C_{\theta a}(E, f(x)) \tag{22}$$
for every $x \in K \cap f^{-1}(K)$.

2. there exist $C > 0$ and $\sigma > 1$ such that,
$$\|Df^n(x)v\| \geq C\sigma^n \|v\| \tag{23}$$
for every $v \in C_a(E, x)$, $n \geq 1$, and $x \in K \cap f^{-1}(K) \cap \cdots \cap f^{-n+1}(K)$.

Observe that we do not require the splitting $E \oplus F$ to be continuous. Note also that $C_a(E)$ depends on the choice of the Riemannian norm $\| \cdot \|$ in M, besides the subbundles E and F. A cone field is *stable*, respectively *backward invariant*, for f if it is unstable, respectively forward invariant for f^{-1}.

Proposition 5.3. *Let Λ be a compact invariant set of a diffeomorphism f. Suppose there is a splitting $T_\Lambda M = E \oplus F$ and there are constants $a > 0$ and $b > 0$ such that $C_a(E)$ is an unstable cone field and $C_b(F)$ is a stable cone field for f on Λ. Then Λ is a hyperbolic set for f.*

The converse is also true: if Λ is a hyperbolic set with splitting $E^u \oplus E^s$ then any cone field with sufficiently small width (relative to an adapted norm) around E^u is an unstable cone field for f and, analogously, any cone field with small width around E^s is a stable cone field for f on Λ.

It is important to observe that the conditions in Definition 5.3 are *open*: if $C_a(E)$ is an unstable cone field for f, then it is also an unstable cone field for any other diffeomorphism C^1 near it. This leads to

Corollary 5.4. *Given a hyperbolic set Λ of a diffeomorphism $f : M \to M$, there exists a neighbourhood U of Λ in M, and a neighbourhood $\mathcal{N}$ of f in $\mathrm{Diff}^1(M)$, such that if $g \in \mathcal{N}$ and Γ is any compact subset of U that is invariant under g, then Γ is a hyperbolic set for g.*

Example 5.4. We say that $f : M \to M$ is an *Anosov diffeomorphism* if the whole manifold M is a hyperbolic set for f. See [4]. A special case are the hyperbolic automorphisms of the d-torus T^d constructed in Example 5.2. As an application of the previous corollary, the class of Anosov diffeomorphisms is open in the C^1 topology.

Anosov diffeomorphisms on tori are always topologically conjugate to a hyperbolic automorphism as in Example 5.2, cf. [18], [35]. More generally, Anosov diffeomorphisms may be constructed on infranilmanifolds [58], and then they are topologically conjugate to algebraic models. It is not known whether these diffeomorphisms may exist on other manifolds. On the other hand, by [17], [39] Anosov diffeomorphisms such that either E^u or E^s have dimension 1 exist only on topological tori.

Example 5.5. (Solenoids) Let $S^1 = \mathbb{R}/\mathbb{Z}$, D^2 be the closed unit disk in the complex plane, and let Q be the solid torus $Q = S^1 \times D^2$. Given constants

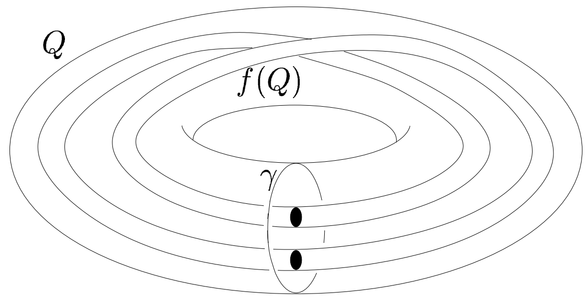

Figure 6: A solenoid

$0 < \lambda < \rho < 1/(2\pi)$, let $f : Q \to Q$ be the map given by

$$f(\theta, z) = (2\theta \bmod \mathbb{Z}, \rho e^{2\pi i\theta} + \lambda z).$$

Geometrically, f acts on the solid torus by contracting along the D^2 direction, and stretching and wrapping the image twice around the S^1 direction. See Figure 6. The assumptions on ρ and λ ensure that f is an embedding of Q strictly into itself. Then the set

$$\Lambda = \bigcap_{n \geq 0} f^n(Q)$$

of those points whose orbit is defined for all times (both positive and negative) is a hyperbolic set for f. Indeed, as the reader may check,

$$C^u_a(p) = \{(\dot{\theta}, \dot{z}) \in T_p(S^1 \times D^2) : |\dot{z}| \leq a|\dot{\theta}|\}$$

$$C^s_b(p) = \{(\dot{\theta}, \dot{z}) \in T_p(S^1 \times D^2) : |\dot{\theta}| \leq b|\dot{z}|\}$$

are, respectively, unstable cone field and stable cone field for f on Q, if $a = 1$ and b is sufficiently small.

An invariant set Λ of a diffeomorphism $f : M \to M$ is *isolated* if there exists a neighbourhood U of it such that Λ is the set of points whose orbits never leave U, neither in the future nor in the past:

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

Theorem 5.5. *If Λ is an isolated hyperbolic set for a diffeomorphism f, then its hyperbolic continuation Λ_g for a nearby map g is an isolated hyperbolic set for g, indeed $\Lambda_g = \cap_{n \in \mathbb{Z}} g^n(U)$.*

We say that an invariant set Λ of a transformation f is *transitive* if there exists some $x \in \Lambda$ whose forward orbit $\{f^n(x) : n \geq 0\}$ is dense in Λ.

Definition 5.6. A transitive hyperbolic set Λ is a *hyperbolic attractor* for $f : M \to M$ if there exists an open neighbourhood U of Λ such that

$$\operatorname{clos}\big(f(U)\big) \subset U \qquad \text{and} \qquad \Lambda = \bigcap_{n \geq 0} f^n(U).$$

A set Λ is a *hyperbolic repeller* for f if it is a hyperbolic attractor for f^{-1}.

If f is a transitive Anosov map then $\Lambda = M$ is a hyperbolic attractor (and a hyperbolic repeller) for f. The solenoids in Example 5.5 are transitive sets and, thus, they are hyperbolic attractors for the corresponding maps.

Clearly, hyperbolic attractors are isolated hyperbolic sets. In particular, the hyperbolic continuation of a hyperbolic attractor is again a hyperbolic attractor. Similar facts hold for repellers.

Definition 5.7. The *basin* of a hyperbolic attractor Λ is the set $B(\Lambda)$ of points $z \in M$ such that

$$d(f^n(z), \Lambda) \to 0 \quad \text{as} \quad n \to +\infty \tag{24}$$

It is easy to see that $B(\Lambda)$ coincides with the union of all backward iterates $f^{-n}(U)$, for any neighbourhood U as in Definition 5.6.

5.2 Stable and Unstable Manifolds

Among the most important geometric properties of hyperbolic sets is the existence of invariant foliations (or laminations) that are dynamically defined. This is the subject of the present subsection. For the time being Λ denotes any hyperbolic set of a C^r diffeomorphism $f : M \to M$, $r \geq 1$. Near the end of the subsection we focus on the case when Λ is an attractor.

Definition 5.8. The *stable manifold* $W^s(x)$ of $x \in M$ is the set of points $y \in M$ whose forward orbit is asymptotic to that of x:

$$\lim_{n \to +\infty} d(f^n(x), f^n(y)) = 0.$$

Given $\varepsilon > 0$, the *local stable manifold* of size ε of a point $x \in M$ is the set $W^s_\varepsilon(x)$ of points $y \in M$ such that

$$\lim_{n\to+\infty} d(f^n(x), f^n(y)) = 0 \quad \text{and} \quad d(f^n(x), f^n(y)) \leq \varepsilon \ \text{ for all } n \geq 0.$$

It follows immediately from the definitions that $y \in W^s(x)$ if and only if $f^n(y)$ is in the local stable manifold of $f^n(x)$ for some $n \geq 0$. That is,

$$W^s(x) = \bigcup_{n\geq 0} f^{-n}\big(W^s_\varepsilon(f^n(x))\big).$$

In general these sets are not submanifolds of M, in fact they may have a very complicated geometric structure. However, according to the next theorem, if x belongs to some hyperbolic set then $W^s_\varepsilon(x)$ is a disk C^r embedded in M. Then, $W^s(x)$ is a C^r immersed submanifold.

We represent by $\mathrm{Emb}^s(N, M)$ the space of C^s embeddings of a manifold N in M, for any integer $s \geq 1$.

Theorem 5.6. *Let Λ be a hyperbolic set for a diffeomorphism f. Provided $\varepsilon > 0$ is small enough, every local stable manifold $W^s_\varepsilon(x)$, $x \in \Lambda$, is a disk C^r embedded in M, with $T_x W^s_\varepsilon(x) = E^s_x$.*

Moreover, $W^s_\varepsilon(x)$ varies continuously with the point $x \in \Lambda$: given any $p \in \Lambda$ there exists a neighbourhood V_p of p in Λ, and a continuous map

$$\Phi_p : V_p \to \mathrm{Emb}^r(W^s_\varepsilon(p), M),$$

such that $\Phi_p(p)$ is the inclusion of $W^s_\varepsilon(p)$ in M, and every $W^s_\varepsilon(x)$, $x \in V_p$, is given by the image of $W^s_\varepsilon(p)$ under $\Phi_p(x)$.

The *local unstable manifold* of size ε, denoted $W^u_\varepsilon(x)$, and the *unstable manifold*, $W^u(x)$, of a point $x \in M$ are defined in the same way as the local stable manifold and the stable manifold, respectively, replacing f by its inverse. By Theorem 5.6 applied to f^{-1}, local unstable manifolds of points in a hyperbolic set are C^r embedded disks, and the unstable manifolds are C^r immersed.

Definition 5.9. Let Λ be a hyperbolic set for a diffeomorphism f in M. The *stable set* $W^s(\Lambda)$ is the set of points $x \in M$ such that

$$W^s(\Lambda) = \{x \in M : \lim_{n\to+\infty} d(f^n(x), \Lambda) = 0\}.$$

The *unstable set* $W^u(\Lambda)$ of Λ is defined similarly, taking $n \to -\infty$.

Clearly, $W^s(\Lambda)$ contains the union of the stable manifolds $W^s(x)$ of all the points $x \in \Lambda$. In general, this inclusion may be strict but, according to the next theorem, this never happens if the hyperbolic set is isolated: any orbit that approaches an isolated hyperbolic set Λ is asymptotic to some orbit inside Λ. A dual statement for unstable manifolds follows immediately, by considering the inverse map f^{-1}.

Theorem 5.7. *If Λ is an isolated hyperbolic set, $W^s(\Lambda) = \cup_{x\in\Lambda} W^s(x)$. In fact, given any $\varepsilon > 0$ there exists a neighbourhood U_ε of Λ such that*

$$\{x \in M : f^n(x) \in U_\varepsilon \text{ for all } n \geq 0\} \subset \bigcup_{x\in\Lambda} W^s_\varepsilon(x).$$

In particular, this theorem applies when Λ is a hyperbolic attractor. By definition, in that case $W^s(\Lambda) = B(\Lambda)$ contains a neighbourhood of Λ. In fact, this last property characterizes attractors among the transitive hyperbolic sets:

Theorem 5.8. *Let Λ be a transitive hyperbolic set for $f : M \to M$. The following conditions are equivalent:*

1. *Λ is a hyperbolic attractor for f;*
2. *$W^s(\Lambda)$ contains a neighbourhood of Λ in M;*
3. *$W^u(p) \subset \Lambda$ for every $p \in \Lambda$.*

Foliated Charts Let Λ be a hyperbolic attractor and $\mathcal{F}^u$ be its unstable foliation, i.e., the family of unstable manifolds of the points in the attractor. Let $\varepsilon > 0$ be fixed, small enough so that the conclusion of Theorem 5.6 holds for Λ.

Given any $p \in \Lambda$, let V_p be a neighbourhood of p in Λ, and

$$\Phi_p : V_p \to \mathrm{Emb}^r(W^u_\varepsilon(p), M)$$

be a continuous map as in Theorem 5.6: $W^u_\varepsilon(z) = \Phi_p(z)(W^u_\varepsilon(p))$ for every $z \in V_p \cap \Lambda$. Moreover, let Σ_p be some smooth disk transverse to the unstable foliation at p, in the sense that $T_p\Sigma_p \oplus E^u_p = T_pM$. We take Σ_p small enough so that it intersects each local unstable manifold of size ε in not more than

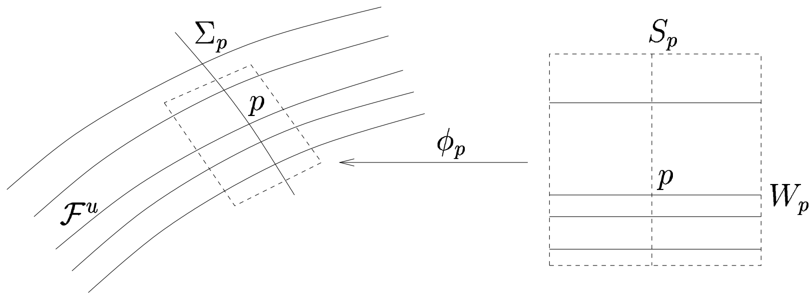

Figure 7: A foliated chart

one point. Note that, by transversality, Σ_p does intersect the local unstable manifold of any point of Λ close enough to p. We denote, cf. Figure 7,

$$W_p = W^u_\varepsilon(p) \cap V_p \quad \text{and} \quad S_p = \Sigma_p \cap V_p \cap \Lambda.$$

Up to replacing V_p by some smaller compact neighbourhood, we may suppose that W_p is a compact disk around p and S_p is also a compact set. Then we define

$$\phi_p : W_p \times S_p \to \Lambda, \qquad \phi_p(x, y) = \Phi_p(y)(x).$$

Observe that ϕ_p does take values in Λ, by the last property in Theorem 5.8. Moreover, as a consequence of Theorem 5.6,

(A1) ϕ_p is a homeomorphism onto a neighbourhood Z_p of p in Λ, with $\phi_p(x, p) = x$ for every $x \in W_p$;

(A2) every $\phi_{p,y} = \phi_p \mid (W_p \times \{y\})$ is a C^r diffeomorphism onto a neighbourhood of y inside $W^u(y)$;

(A3) the diffeomorphisms $\phi_{p,y}$ vary continuously with the point $y \in S_p$, in the C^r topology.

We call $\phi_p : W_p \times S_p \to Z_p$ a *foliated chart* for the unstable foliation $\mathcal{F}^u$ at the point p.

Observe that the size of W_p, S_p, Z_p is essentially determined by $\varepsilon > 0$. In particular, foliated charts can be constructed at any point p of Λ in such a way that Z_p contains a neighbourhood with fixed radius around the point. Of course, decreasing ε we can also make Z_p arbitrarily small.

6 Partial Hyperbolicity

Here we introduce an important extension of the notions discussed in the previous section: partially hyperbolic sets and attractors. As we shall see, partial hyperbolicity is intimately related to *robustness* of the dynamics, e.g., attractors that can not be destroyed by any small perturbation of the system.

6.1 Definitions and Examples

Definition 6.1. Let Λ be a compact invariant set for a C^1 diffeomorphism $f : M \to M$, and $T_\Lambda M = E^1 \oplus E^2$ be a continuous Df-invariant splitting of the tangent space over Λ. We say that the splitting is *dominated* if there are constants $C > 0$ and $0 < \lambda < 1$ such that

$$\|Df^{-n}(f^n(x)) \mid E^1_{f^n(x)}\| \, \|Df^n(x) \mid E^2_x\| \leq C\lambda^n$$

for every $x \in \Lambda$ and $n \geq 1$.

Recall that whenever we speak of a splitting of the tangent space it is implicit that the subspaces E^1_x and E^2_x have constant dimensions over the invariant set. The condition in the definition may be rewritten

$$\frac{\|Df^n(x)v_2\|}{\|Df^n(x)v_1\|} \leq C\lambda^n$$

for every norm 1 vectors $v_1 \in E^x_1$ and $v_2 \in E^2_x$, and $x \in \Lambda$. So, roughly speaking, the splitting is dominated if E^1 is more expanding/less contracting than E^2.

Definition 6.2. Let Λ be a compact invariant set for $f : M \to M$. We say that Λ is *partially hyperbolic* if there is a dominated splitting $T_\Lambda M = E^1 \oplus E^2$ such that

1. either E^1 is expanding: $\|Df^{-n}(x) \mid E^1_x\| \leq C\lambda^n$ for every $x \in \Lambda$ and $n \geq 1$,
2. or E^2 is contracting: $\|Df^n(x) \mid E^2_x\| \leq C\lambda^n$ for every $x \in \Lambda$ and $n \geq 1$.

In the first case we write $E^1 = E^u$, $E^2 = E^{cs}$, and say that the partially hyperbolic set Λ is of *type* $E^u \oplus E^{cs}$. In the second one we write $E^1 = E^{cu}$, $E^2 = E^s$, and say that Λ is of *type* $E^{cu} \oplus E^s$.

Definition 6.3. A diffeomorphism $f : M \to M$ is *partially hyperbolic* (of type $E^u \oplus E^{cs}$, respectively $E^{cu} \oplus E^s$) if the manifold M is a partially hyperbolic set for f (of type $E^u \oplus E^{cs}$, respectively $E^{cu} \oplus E^s$).

Λ is partially hyperbolic of type $E^u \oplus E^{cs}$ for f if and only if it is of type $E^{cu} \oplus E^s$ for the inverse f^{-1}. Of course, hyperbolic sets are also partially hyperbolic, of both types. In the sequel we describe a few other examples. The first one, a generalization of Example 5.2, illustrates the fact that, unlike in the hyperbolic case, the splitting in Definition 6.2 is usually not unique.

Example 6.4. Let $d \geq 3$ and $A \in \mathrm{Sl}(d, \mathbb{Z})$ be such that the spectrum of A splits into three nonempty subsets, contained in

$$\{\lambda : b < |\lambda|\}, \quad \{\lambda : a < |\lambda| < b\}, \quad \{\lambda : |\lambda| < a\}.$$

for some $0 < a < 1 < b$. Let $\mathbb{R}^d = \hat{F}^u \oplus \hat{F}^c \oplus \hat{F}^s$ be the corresponding splitting of $\mathbb{R}^d$ into invariant subspaces: A preserves the three subspaces, the eigenvalues of $A \mid \hat{F}^u$ have norm larger than b, those of $A \mid \hat{F}^s$ have norm smaller than a, and the eigenvalues of $A \mid F^c$ are all in (a, b). Let $M = T^d$ be the d-dimensional torus, $\pi : \mathbb{R}^d \to M$ be the canonical projection, and f be the diffeomorphism induced in M by A, cf. Example 5.2. Then $F^u = D\pi \cdot \hat{F}^u$, $F^c = D\pi \cdot \hat{F}^c$, $F^s = D\pi \cdot \hat{F}^s$ are continuous subbundles of the tangent space of M. It is easy to see that both

$$E^1 = F^u, \quad E^2 = F^c \oplus F^s, \quad \text{and} \quad E^1 = F^u \oplus F^c, \quad E^2 = F^s,$$

define splittings of TM as in Definition 6.2, respectively of type $E^u \oplus E^{cs}$ and $E^{cu} \oplus E^s$. Note that this construction is valid even if A is hyperbolic (in which case f is Anosov).

Remark 6.5. Partial hyperbolicity is sometimes defined through a stronger condition: existence of a dominated splitting into three subbundles, one of which is expanding whereas another is contracting. Here such systems are called strongly partially hyperbolic. That is, a compact invariant set Λ is *strongly partially hyperbolic* for f if there exists a continuous splitting

$$T_\Lambda M = E^u \oplus E^c \oplus E^s$$

of the tangent bundle into three Df-invariant subbundles, such that E^u is uniformly expanding, E^s is uniformly contracting, and both splittings

$$E^1 = E^u, \ E^2 = E^c \oplus E^s \qquad \text{and} \qquad E^1 = E^u \oplus E^c, \ E^2 = E^s$$

are dominated.

Example 6.6. Let $(X^t)_{t\in\mathbb{R}}$ be an Anosov flow on a compact manifold M: $(X^t)_t$ is a C^1 flow without singularities such that there exists a continuous splitting $TM = E^s \oplus E^0 \oplus E^u$ of the tangent bundle of M into three subbundles that are invariant under DX^t, for every $t \in \mathbb{R}$; moreover, E^u is expanding and E^s is contracting: there are $C > 0$ and $0 < \lambda < 1$ such that

$$\|DX^{-t} \mid E^u\| \leq C\lambda^t, \quad \|DX^t \mid E^s\| \leq C\lambda^t, \quad \text{for } t > 0,$$

and E^0 is the one-dimensional subbundle generated by the vector field. Let $f = X^1$ be the time-1 map of the flow. Then M is a strongly partially hyperbolic set for f, with splitting

$$E^u \oplus E^0 \oplus E^s.$$

More generally, a hyperbolic set of a flow is strongly partially hyperbolic for the corresponding time 1 diffeomorphism.

Now we state a generalization of Proposition 5.3 for partially hyperbolic systems. The terminology is as in Definition 5.3.

Proposition 6.1. *Let Λ be a compact invariant set for $f : M \to M$.*

1. *Λ admits a dominated splitting if and only if there exists a continuous splitting $T_\Lambda M = E \oplus F$ and there exists $a > 0$ such that the cone field $C_a(E)$ is forward invariant.*

2. *Λ is partially hyperbolic of type $E^u \oplus E^{cs}$ if and only if $C_a(E)$ may be taken to be an unstable cone field for f.*

As a consequence, one also gets an analog of Corollary 5.4: if Λ is a partially hyperbolic set for f, there exists a neighbourhood U of Λ in M and a neighbourhood $\mathcal{N}$ of f in $\mathrm{Diff}^1(M)$, such that if $g \in \mathcal{N}$ and Γ is any compact subset of U that is invariant under g, then Γ is partially hyperbolic set for g.

Example 6.7. Let $f : M \to M$ be a diffeomorphism with some hyperbolic attractor Λ, e.g. an Anosov diffeomorphism. Let $T_\Lambda M = E^u \oplus E^s$ be the corresponding splitting of the tangent space over Λ. Fix $\lambda < 1$ and some Riemannian metric on M so that

$$\|Df^{-1} \mid E^u\| < \lambda \quad \text{and} \quad \|Df \mid E^s\| < \lambda.$$

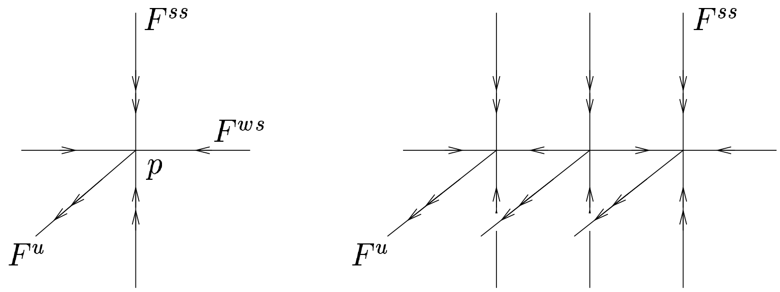

Figure 8: Deforming an Anosov diffeomorphism

Now let $g : N \to N$ be any diffeomorphism on a compact manifold N such that $\|Dg\| < \lambda$ and $\|Dg^{-1}\| < \lambda$. Note that this is satisfied by any map close enough to the identity. Then $\Lambda \times N$ is a strongly partially hyperbolic set for the map $f \times g : M \times N \to M \times N$, with splitting

$$\big(E^u \times \{0\}\big) \oplus \big(\{0\} \times TN\big) \oplus \big(E^s \times \{0\}\big).$$

Example 6.8. Partially hyperbolic sets can also be obtained by deformation of hyperbolic ones, as in the following construction of [31]. Start with an Anosov diffeomorphism f_0 in T^3 that admits an invariant splitting into three subbundles

$$F^u \oplus F^{ws} \oplus F^{ss}.$$

Here F^u is expanding, and F^{ws} and F^{ss} are contracting, with F^{ss} dominating F^{ws}: there are $\lambda < 1$ and a Riemannian metric on M such that

$$\|Df_0 \mid F^{ss}_x\| \, \|Df_0^{-1} \mid F^{ws}_{f(x)}\| \leq \lambda$$

for every x. Let p be a fixed point of f_0. Deform f_0 by isotopy in a small neighbourhood of p, as described in Figure 8: keep the diffeomorphism essentially unchanged in the directions of F^u and F^{ss}, while modifying it in the direction of F^{ws} so that the fixed point goes through a pitchfork bifurcation, that gives rise to two new fixed points. As shown in [31], this can be done in such a way that the whole $M = T^3$ is a partially hyperbolic transitive set for the resulting diffeomorphism f, as well as for any other one in a C^1 neighbourhood of it. Since f has periodic saddles with either 1 or 2 contracting eigenvalues, it can not be an Anosov diffeomorphism.

6.2 Invariant Foliations

Partially hyperbolic sets share some of the nice geometric properties of hyperbolic sets. Crucial among these is the existence of invariant foliations tangent to expanding or contracting subbundles, stated in the next theorem.

By a *foliation* $\mathcal{F}$ on a set $\Lambda \subset M$ we mean a family of two-by-two disjoint C^r immersed submanifolds, $1 \leq r \leq \infty$, called the *leaves* of $\mathcal{F}$, such that every one of them intersects Λ, and every point $p \in \Lambda$ is contained in some of the leaves. The foliation is *f-invariant* if $f(\mathcal{F}(p)) = \mathcal{F}(f(p))$ for every $p \in \Lambda$, where $\mathcal{F}(p)$ denotes the leaf that contains p.

All the foliations we deal with here are *continuous*, in the sense of Theorem 5.6: given any $p \in \Lambda$ there exists a neighbourhood V_p of p in Λ, a disk W_p around p inside $\mathcal{F}(p)$, and a continuous map

$$\Phi_p : V_p \to \text{Emb}^r(W_p, M),$$

such that $\Phi_p(p)$ is the inclusion of W_p in M, and the image $\Phi_p(z)(W_p)$ of the embedding $\Phi_p(z)$ is a neighbourhood of z inside $\mathcal{F}(z)$. When the value of r is relevant, we say that $\mathcal{F}$ is a *continuous foliation with C^r leaves*.

Theorem 6.2. *Let Λ be a partially hyperbolic set of type $E^u \oplus E^{cs}$ for a C^r diffeomorphism f, any $r \geq 1$. Then there exists a unique f-invariant foliation $\mathcal{F}^u$ on Λ such that $T_p\mathcal{F}^u(p) = E^u_p$ at every point p of Λ. This foliation is continuous, its leaves are C^r submanifolds, and they are exponentially contracted by backward iterates: there is $\lambda < 1$ and, for each pair of points z_1, z_2 in a same leaf of $\mathcal{F}^u$, there exists $C > 0$ such that*

$$d(f^{-n}(z_1), f^{-n}(z_2)) \leq C\lambda^n, \quad \textit{for every } n \geq 1.$$

See [13, Section 2], [25], and [56, Appendix IV]. We call $\mathcal{F}^u$ *strong-unstable foliation* of Λ. Dually, given a partially hyperbolic set of type $E^{cu} \oplus E^s$ there exists a unique f-invariant *strong-stable foliation* $\mathcal{F}^s$ tangent to E^s. Furthermore, it is continuous, its leaves are at least as smooth as the diffeomorphism f, and they are exponentially contracted by forward iterates. This follows from applying Theorem 6.2 to the inverse map f^{-1}.

Definition 6.9. A transitive partially hyperbolic set Λ is a *partially hyperbolic attractor* for $f : M \to M$ if there exists an open neighbourhood U of Λ such that

$$\text{clos}\big(f(U)\big) \subset U \qquad \text{and} \qquad \Lambda = \bigcap_{n \geq 0} f^n(U).$$

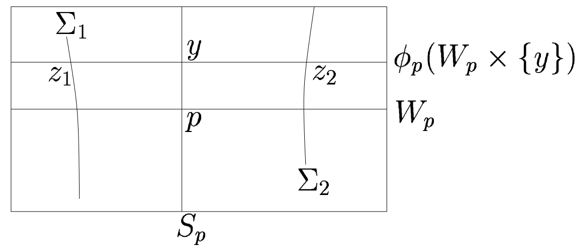

Figure 9: A local holonomy map for $\mathcal{F}^u$

The following extension of Theorem 5.8 is part of [46, Proposition 1].

Proposition 6.3. *If Λ is a partially hyperbolic attractor then it is a union of entire unstable leaves: $\mathcal{F}^u(p) \subset \Lambda$ for every $p \in \Lambda$.*

For the remainder of this section we take Λ to be a partially hyperbolic attractor of type $E^u \oplus E^{cs}$. At some points we suppose that Λ is hyperbolic, and then it is always implicit that the subbundle E^{cs} is contracting, that is, the splitting is the one in Definition 5.1. In this case we also call $\mathcal{F}^u$ the *unstable foliation*, and $\mathcal{F}^s$ the *stable foliation* of Λ.

The continuity of $\mathcal{F}^u$ allows us to define foliated charts $\phi_p : W_p \times S_p \to Z_p$ for the strong-unstable foliation $\mathcal{F}^u$ at each point p in Λ, just as we did for hyperbolic sets at the end of Subsection 5.2. Let $p \in M$ and $\phi_p : W_p \times S_p \to Z_p$ be a foliated chart for $\mathcal{F}^u$ at p. Let Σ_1 and Σ_2 be C^1 submanifolds embedded in Z_p, with $\dim \Sigma_i = \dim E^{cs}$ and transverse to the strong-unstable foliation: Σ_i intersects each strong-unstable disk $\phi_p(W_p \times \{y\})$ in at most one point, and this intersection is transverse, for $i = 1, 2$.

Define $\tilde{\Sigma}_1$ to be the set of points $z_1 \in \Sigma_1$ such that the unstable disk $\phi_p(W_p \times \{y\})$ passing through z_1 intersects Σ_2 in some point $z_2 = \pi(z_1)$. See Figure 9. The *local holonomy map* of $\mathcal{F}^u$ from Σ_1 to Σ_2 is

$$\pi : \tilde{\Sigma}_1 \to \Sigma_2, \quad z_1 \mapsto \pi(z_1) = z_2 .$$

Since ϕ_p is a homeomorphism, π is a homeomorphism onto its image $\tilde{\Sigma}_2$.

According to Theorem 6.2, the leaves of the unstable foliation are at least as smooth as the diffeomorphism f itself. In general, one can not expect the local holonomy maps to be very regular: holonomy maps of real analytic Anosov diffeomorphisms may fail to be Lipschitz continuous. However, one always has at least Hölder regularity, as long as the diffeomorphism is twice differentiable:

Theorem 6.4. *If Λ is a partially hyperbolic attractor of a C^2 diffeomorphism f, every local holonomy map of the strong-unstable foliation is Hölder continuous, with uniform Hölder constants.*

Somewhat related to this, we also have that the invariant subbundles E^u and E^{cs} are Hölder continuous if f is C^2. Recall that a splitting is continuous if the subbundles are locally generated by continuous linearly independent vector fields. Correspondingly, we say that the splitting $T_\Lambda M = E^u \oplus E^{cs}$ is Hölder continuous if those vector fields can be taken Hölder, with uniform Hölder constants.

Theorem 6.5. *If Λ is a partially hyperbolic attractor of a C^2 diffeomorphism, the corresponding splitting $T_\Lambda M = E^u \oplus E^{cs}$ is Hölder continuous.*

See [13, Corollary 2.1] and [24, Theorem 6.4]. If Λ is strongly partially hyperbolic, as in Remark 6.5, then the splitting $T_\Lambda M = E^u \oplus E^c \oplus E^s$ is Hölder continuous (the definition extends immediately to splittings into any number of subbundles).

Although, as mentioned before, holonomy maps of the strong-unstable foliation need not be smooth in general, they do have a weaker regularity property called *absolute continuity*: zero Lebesgue measure sets are mapped into zero Lebesgue sets, at least if the diffeomorphism is twice differentiable. That is the content of the next theorem, where m_{Σ_1} and m_{Σ_2} represent the Riemannian volumes induced by the Riemannian metric on Σ_1 and Σ_2, respectively.

Theorem 6.6. *If Λ is a partially hyperbolic attractor of a C^2 diffeomorphism f, there exists a constant $C_0 > 1$, depending only on f, such that any local holonomy map π of the strong-unstable foliation of Λ satisfies*

$$\frac{1}{C_0} m_{\Sigma_1}(B) \leq m_{\Sigma_2}(\pi(B)) \leq C_0 m_{\Sigma_1}(B)$$

for any measurable set $B \subset \Sigma_1$, and any transverse sections Σ_1, Σ_2 as above.

This crucial fact was first established by [4], as a main step in the proof that C^2 Anosov diffeomorphisms and flows that preserve Lebesgue measure are ergodic. See also [5]. The extension to the partially hyperbolic context was due to [13].

7 Measures Absolutely Continuous Along $\mathcal{F}^u$

In this section we take $f : M \to M$ to be a C^2 diffeomorphism. We prove that every partially hyperbolic attractor Λ of type $E^u \oplus E^{cs}$, supports some invariant probability measure μ that is *absolutely continuous along the strong-unstable foliation* $\mathcal{F}^u$. The precise statement will be given in Theorem 7.1, after we have defined this last notion. We shall show that, in many situations, such probabilities are physical measures for f on Λ. In particular, this is the case if the subbundle E^{cs} is contracting, corresponding to Λ being hyperbolic.

We begin by defining absolute continuity of a measure along a foliation, in an abstract setting. Let X, Y be compact metric spaces, and $\mathcal{F}$ denote the partition of $X \times Y$ into "horizontal lines"

$$\mathcal{F} = \big\{X \times \{y\} : y \in Y\big\}.$$

Let ν be some probability measure in X. We also denote $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ the canonical projections. Given any measure μ on $X \times Y$, we let $\hat{\mu} = \pi_{Y*}(\mu)$. That is, $\hat{\mu}$ is the measure defined on Y by

$$\hat{\mu}(\xi) = \mu(\pi_Y^{-1}(\xi)) = \mu(X \times \xi) \quad \text{for every measurable set} \quad \xi \subset Y.$$

Definition 7.1. A measure μ on $X \times Y$ is *absolutely continuous with respect to ν along the horizontal foliation $\mathcal{F}$* if there exists a measurable function $\rho : X \times Y \to [0, +\infty)$ such that

$$\mu(B) = \int_B \rho(x, y)\, d\nu(x)\, d\hat{\mu}(y)$$

for every measurable set $B \subset X \times Y$.

In other words, μ is absolutely continuous with respect to ν along $\mathcal{F}$ if and only if μ is absolutely continuous with respect to the product measure $\nu \times \hat{\mu}$. Then, ρ is the Radon-Nikodym derivative of μ relative to $\nu \times \hat{\mu}$. We call $\{\rho(\cdot, y)\nu : y \in Y\}$ *conditional measures* of μ relative to $\mathcal{F}$.

Going back to the dynamical context, let Λ be a partially hyperbolic attractor of type $E^u \oplus E^{cs}$ for a C^2 diffeomorphism $f : M \to M$, and let $\mathcal{F}^u$ be the corresponding strong-unstable foliation. The definition of absolute continuity of a measure along $\mathcal{F}^u$ uses the notion of foliated chart introduced in Subsections 5.2 and 6.2.

Definition 7.2. A measure μ supported in Λ is *absolutely continuous along* $\mathcal{F}^u$ if for every $p \in \Lambda$ there exists a foliated chart $\phi_p : W_p \times S_p \to Z_p$ for $\mathcal{F}^u$ at p such that the pull-back $\phi_p^* \mu$ of μ under ϕ_p is absolutely continuous with respect to Lebesgue measure along $\mathcal{F}$.

The *pull-back* is the measure on $W_p \times S_p$ defined by $\phi_p^* \mu(B) = \mu(\phi_p(B))$, for every measurable subset B.

Let U be any compact disk contained in some leaf of $\mathcal{F}^u$, and m_U be a Riemannian volume on U. That is, m_U is the Lebesgue measure induced on U by the volume element associated to some Riemannian metric of M. In this section we prove the following result of [46].

Theorem 7.1. *Any accumulation point of the sequence*

$$\frac{1}{n} \sum_{j=0}^{n-1} f_*^j m_U$$

is an invariant measure for f, absolutely continuous along $\mathcal{F}^u$.

7.1 Conditional Measures

As a first step in the proof of Theorem 7.1, we obtain a few abstract results about measures absolutely continuous along a foliation. The setting is the same as in Definition 7.1: we suppose that μ is a measure on the product $X \times Y$ of two compact metric spaces, and ν is a probability measure in X.

Lemma 7.2. *Suppose there exists a measurable function $\psi : X \to [0, +\infty)$ with $\int \psi \, d\nu < \infty$, and a family $\mathcal{R}$ of rectangles $A \times \xi \subset X \times Y$ generating the σ-algebra of all measurable subsets of $X \times Y$, so that*

$$\mu(A \times \xi) \le \hat{\mu}(\xi) \int_A \psi \, d\nu \quad \textit{for every } A \times \xi \in \mathcal{R}.$$

Then μ is absolutely continuous with respect to ν along $\mathcal{F}$, with $\rho(x, y) \le \psi(x)$ at $(\nu \times \hat{\mu})$-almost every point $(x, y) \in X \times Y$.

Proof. It is easy to check that the family of measurable subsets $B \subset X \times Y$ for which

$$\mu(B) \le \int_B \psi(x) \, d\nu(x) \, d\hat{\mu}(y) \tag{25}$$

is a σ-algebra. By assumption, this family contains $\mathcal{R}$. Since we also take $\mathcal{R}$ to be generating, (25) must hold for every measurable subset B. This implies that μ is absolutely continuous with respect to $(\nu \times \hat{\mu})$, with Radon-Nikodym derivative $\rho = d\mu/d(\nu \times \hat{\mu})$ satisfying $\rho(x, y) \leq \psi(x)$ at $(\nu \times \hat{\mu})$-almost every point. □

Remark 7.3. In the same setting, let $\phi : X \to [0, +\infty)$ be any measurable function such that

$$\mu(A \times \xi) \geq \hat{\mu}(\xi) \int_A \phi \, d\nu$$

for every $A \times \xi$ in some generating family of rectangles. Then, by similar arguments, the Radon-Nikodym derivative also satisfies $\rho(x, y) \geq \phi(x)$ at $(\nu \times \hat{\mu})$-almost every point.

Proposition 7.3. *Let μ_k, $k \geq 1$, be a sequence of measures on $X \times Y$ converging to some measure μ in the weak* topology. Suppose there exists a measurable function $\psi : X \to [0, +\infty)$ with $\int \psi \, d\nu < \infty$, and a family $\mathcal{R}$ of rectangles $A \times \xi \subset X \times Y$ generating the σ-algebra of measurable subsets of $X \times Y$, so that*

$$\mu_k(A \times \xi) \leq \hat{\mu}_k(\xi) \int_A \psi \, d\nu \quad \textit{for every } A \times \xi \in \mathcal{R} \textit{ and } k \geq 1.$$

Then μ is absolutely continuous with respect to ν along $\mathcal{F}$, with density $\rho \leq \psi$ at $(\nu \times \hat{\mu})$-almost every point in $X \times Y$.

Proof. Let $\hat{\mu}_k = \pi_{Y*}(\mu_k)$. By the previous lemma, for each $k \geq 1$ there exists $\rho_k : X \times Y \to [0, +\infty)$ such that $\mu_k = \rho_k(\nu \times \hat{\mu}_k)$ and $\rho_k \leq \psi$ at $(\nu \times \hat{\mu}_k)$-almost every point. In particular,

$$\mu_k(A \times \xi) \leq \hat{\mu}_k(\xi) \int_A \psi \, d\nu$$

for any measurable rectangle $A \times \xi$ in $X \times Y$. Since π_Y is a continuous map, the assumption $\mu_k \to \mu$ implies that $\hat{\mu}_k \to \hat{\mu}$. Then, assuming the boundary of ξ has zero $\hat{\mu}$-measure, $\hat{\mu}_k(\xi)$ converges to $\hat{\mu}(\xi)$ as $k \to \infty$. Let us suppose, furthermore, that $A \subset X$ and $\xi \subset Y$ are open subsets. Then

$$\mu(A \times \xi) \leq \liminf_{k \to \infty} \mu_k(A \times \xi) \leq \liminf_{k \to \infty} \hat{\mu}_k(\xi) \int_A \psi \, d\nu = \hat{\mu}(\xi) \int_A \psi \, d\nu.$$

Now the proposition follows from Lemma 7.2, together with the observation that the family of open rectangles $A \times \xi$ such that $\hat{\mu}(\partial \xi) = 0$ generates the σ-algebra of $X \times Y$. □

Proposition 7.4. *In the setting of Proposition 7.3, let $\phi : X \to [0, +\infty)$ be any measurable function such that*

$$\mu_k(A \times \xi) \geq \hat{\mu}_k(\xi) \int_A \phi \, d\nu \quad \textit{for any } A \times \xi \in \mathcal{R} \textit{ and } k \geq 1.$$

Then the derivative ρ satisfies $\rho \geq \phi$ at $(\nu \times \hat{\mu})$-almost every point.

Proof. This is similar to the previous proposition. According to Remark 7.3, we have $\rho_k \geq \phi$ at $(\nu \times \hat{\mu}_k)$-almost every point. In particular,

$$\mu_k(A \times \xi) \geq \hat{\mu}_k(\xi) \int_A \phi \, d\nu$$

for any measurable rectangle $A \times \xi$ in $X \times Y$. Taking $A \subset X$ and $\xi \subset Y$ to be closed, and the boundary of ξ to have zero measure for $\hat{\mu}$, we conclude that

$$\mu(A \times \xi) \geq \limsup_k \mu_k(A \times \xi) \geq \limsup_k \hat{\mu}_k(\xi) \int_A \phi \, d\nu = \hat{\mu}(\xi) \int_A \phi \, d\nu.$$

Since these closed rectangles generate the σ-algebra of $X \times Y$, the proposition follows from Remark 7.3. □

7.2 Distortion Along Strong-Unstable Leaves

In what follows we suppose that some Riemannian metric has been chosen on M: determinants and lengths of curves are always meant with respect to this metric. This also determines a Riemannian volume on each of the leaves of the strong-unstable foliation, that we denote m_u. Given any measurable subset B of some strong-unstable leaf, we let m_B be the restriction of m_u to B.

For $j \in \mathbb{Z}$ and $y \in \Lambda$, we let $J^u f^j$ be the norm of the Jacobian of f restricted to the unstable subspace E^u_y:

$$J^u f^j(y) = |\det(Df^j(y) \mid E^u_y)|.$$

Let U be a compact disk contained in some leaf of $\mathcal{F}^u$.

Lemma 7.5. *For any $n \geq 1$, we have $f^n_* m_U = (J^u f^{-n}) \, m_{f^n(U)}$.*

Proof. By definition, given any measurable subset B of $f^n(U)$,

$$f_*^n m_U(B) = \int_{f^{-n}(B)} 1 \, dm_U \, .$$

Changing variables $y = (f^n \mid U)(x)$ in the integral, we get

$$f_*^n m_U(B) = \int_B (J^u f^{-n}) \, dm_{f^n(U)}$$

for any measurable subset B, as we claimed. □

Definition 7.4. Given points x and y in a same strong-unstable leaf F,

$$d_u(x,y) = \{\text{length}(\alpha) : \alpha \text{ is a piecewise } C^1 \text{ curve in } F \text{ connecting } x \text{ to } y\}.$$

It is easy to check that $d_u(\cdot\,,\cdot)$ defines a distance on each leaf F of the foliation $\mathcal{F}^u$. Since the derivative Df is uniformly expanding along the tangent bundle $E^u = T\mathcal{F}^u$, this distance is uniformly contracted by negative iterates of f. Indeed, given any piecewise C^1 curve α connecting x to y inside the leaf F, the length of $f^{-j}(\alpha)$ is less than $C\lambda^j \,\text{length}(\alpha)$, for every $j \geq 1$. Therefore,

$$d_u(f^{-j}(x), f^{-j}(y)) \leq C\lambda^j d_u(x,y) \quad \text{for every } j \geq 1. \tag{26}$$

For completeness, we also set $d_u(x,y) = \infty$ when x, y are in different strong-unstable leaves.

Proposition 7.6. *Given $L > 0$ there exists $K > 0$ such that*

$$\frac{J^u f^{-n}(y_1)}{J^u f^{-n}(y_2)} \leq K$$

for every $n \geq 1$ and any $y_1, y_2 \in \Lambda$ such that $d_u(y_1, y_2) \leq L$.

Proof. Since we suppose the diffeomorphism f to be C^2, the tangent bundle $E^u = T\mathcal{F}^u$ is ν_0-Hölder for some $0 < \nu_0 \leq 1$. Recall Theorem 6.5. As a consequence, the map

$$\varphi^u : \Lambda \to \mathbb{R}, \quad \varphi^u(x) = \log|\det(Df(x) \mid E^u_x)|\,,$$

is (C_0, ν_0)-Hölder for some constant $C_0 > 0$. In particular, its restriction to each strong-unstable leaf is (C_0, ν_0)-Hölder with respect to the d_u-distance on the leaf. Using the chain rule,

$$\log \frac{J^u f^{-n}(y_1)}{J^u f^{-n}(y_2)} = \sum_{j=1}^{n} \varphi^u(f^{-j}(y_2)) - \varphi^u(f^{-j}(y_1)).$$

By Hölder continuity and (26), this is less than

$$\sum_{j=1}^{n} C_0 d_u(f^{-j}(y_1), f^{-j}(y_2))^{\nu_0} \le \sum_{j=1}^{n} C_0 \big(C\lambda^j d_u(y_1, y_2)\big)^{\nu_0}.$$

So, since we suppose that $d_u(y_1, y_2) \le L$,

$$\log \frac{J^u f^{-n}(y_1)}{J^u f^{-n}(y_2)} \le C_0 \big(CL\big)^{\nu_0} \sum_{j=1}^{n} \lambda^{j\nu_0}.$$

Thus, we may take $K = C_0 \big(CL\big)^{\nu_0} \sum_{j=1}^{\infty} \lambda^{j\nu_0}$. □

Corollary 7.7. *Given any $L > 0$ there exists $K > 0$ so that, for $n \ge 1$ and any domain D in $f^n(U)$ with d_u-diameter less than L, we have*

$$\frac{1}{K} \frac{m_{f^n(U)}(B)}{m_{f^n(U)}(D)} \le \frac{f_*^n m_U(B)}{f_*^n m_U(D)} \le K \frac{m_{f^n(U)}(B)}{m_{f^n(U)}(D)}$$

for any measurable subset B of D.

Proof. This is a direct consequence of Lemma 7.5 and Proposition 7.6, with the same constant K as in the proposition:

$$\frac{f_*^n m_U(B)}{f_*^n m_U(D)} = \frac{\int_B (J^u f^{-n})\, dm_{f^n(U)}}{\int_D (J^u f^{-n})\, dm_{f^n(U)}} \le K \frac{m_{f^n(U)}(B)}{m_{f^n(U)}(D)},$$

and the lower inequality is obtained in the same way. □

As another consequence of Proposition 7.6, we obtain the following result about positive Lebesgue measure subsets of strong-unstable leaves: forward iterates of the set fill-in an arbitrarily large fraction of some d_u-ball with given radius. The d_u-ball of radius $r > 0$ around a point $q \in \Lambda$ is denoted $B_r^u(q)$.

Proposition 7.8. *Let $r > 0$ be fixed. Let F be a leaf of the foliation $\mathcal{F}^u$, and A be a subset of F such that $m_u(A) > 0$. Then there exist $p_n \in f^n(A)$, $n \geq 1$, such that*

$$\frac{m_u\big(f^n(A) \cap B_r^u(p_n)\big)}{m_u\big(B_r^u(p_n)\big)} \to 1 \quad \text{as } n \to \infty.$$

Proof. It is no restriction to suppose that A is compact. For each $n \geq 1$, let Q_n be a maximal finite subset of $f^n(A)$ such that the open d_u-balls $B_r^u(q)$ of radius r around the points of Q_n are two-by-two disjoint. Then the d_u-balls of radius $2r$ around the points $q \in Q_n$ cover $f^n(A)$. Since $\mathcal{F}^u$ is a continuous foliation with C^2 leaves, cf. Theorem 6.2, its leaves have uniformly bounded curvature. As a consequence, there exists $\kappa > 0$, depending only on r, such that

$$1 \leq \frac{m_u(B_{2r}^u(q))}{m_u(B_r^u(q))} \leq \kappa$$

for any point $q \in \Lambda$. Then, taking $L = 2r$ in Corollary 7.7,

$$1 \leq \frac{m_u\big(f^{-n}(B_{2r}^u(q))\big)}{m_u\big(f^{-n}(B_r^u(q))\big)} \leq K\kappa \tag{27}$$

for any $n \geq 1$ and $q \in Q_n$. Let us show that, for n large, there exists $q \in Q_n$ such that A fills-in a large fraction of $f^{-n}(B_r^u(q))$. More precisely,

Claim:

$$\min_{q \in Q_n} \frac{m_u\big(f^{-n}(B_r^u(q)) \setminus A\big)}{m_u\big(f^{-n}(B_r^u(q))\big)} \to 0 \quad \text{as } n \to \infty.$$

Proof. We prove the claim by contradiction. Suppose there exists $\delta > 0$ such that, for every $q \in Q_n$

$$m_u\big(f^{-n}(B_r^u(q)) \setminus A\big) \geq \delta m_u\big(f^{-n}(B_r^u(q))\big).$$

Since the $B_r^u(q)$, $q \in Q_n$ are two-by-two disjoint, adding over q we get

$$m_u\Big(\bigcup_{q \in Q_n} f^{-n}(B_r^u(q)) \setminus A\Big) \geq \delta m_u\Big(\bigcup_{q \in Q_n} f^{-n}(B_r^u(q))\Big). \tag{28}$$

On the one hand, the d_u-diameter of $f^{-n}(B_r^u(q))$ is less than $2rC\lambda^n$, which converges to zero as $n \to \infty$. Since $f^{-n}(q) \in A$ for all $q \in Q_n$, it follows that the union of the $f^{-n}(B_r^u(q))$ is contained in a small neighbourhood of A if

n is large. Thus, the left hand side of (28) goes to zero as $n \to \infty$. On the other hand, by (27), the right hand side is bounded from below by

$$\frac{\delta}{K\kappa} m_u\Big(\bigcup_{q \in Q_n} f^{-n}(B^u_{2r}(q))\Big) \geq \frac{\delta}{K\kappa} m_u(A) > 0.$$

We have reached a contradiction, so the claim is proved. □

For each $n \geq 1$ we pick p_n to be a point in Q_n where the minimum of the expression in the statement of the claim is attained. Then, using Proposition 7.6 once more,

$$\frac{m_u\big(B_r(p_n) \setminus f^n(A)\big)}{m_u\big(B_r(p_n)\big)} \leq K \frac{m_u\big(f^{-n}(B_r(p_n)) \setminus A\big)}{m_u\big(f^{-n}(B_r(p_n))\big)} \to 0 \quad \text{as } n \to \infty.$$

This proves the proposition. □

7.3 Proof of the Existence Theorem

Now, we are in a position to prove Theorem 7.1. Let μ_n be as in the statement. We want to show that, given any accumulation point μ of μ_n and any point p in Λ, there exists a foliated chart ϕ_p at p such that the pull-back $\phi_p^*\mu$ is absolutely continuous along the horizontal foliation $\mathcal{F}$.

We fix $\mu = \lim_k \mu_{n_k}$ and the point p in all that follows. The choice of a foliated chart $\phi_p : W_p \times S_p \to Z_p$ is rather arbitrary: we only require that the boundary of Z_p have zero μ-measure:

$$\mu(\partial Z_p) = 0, \tag{29}$$

which can always be obtained, replacing W_p and S_p by slightly smaller sets if necessary. We show that, for any such chart, $\phi^*\mu$ is indeed absolutely continuous along the horizontal foliation.

Definition 7.5. We say that a connected component γ of $f^j(U) \cap Z_p$, *crosses* Z_p if $\phi_p^{-1}(\gamma)$ is a graph over W_p, that is, $\phi_p^{-1}(\gamma)$ projects homeomorphically onto W_p under the canonical projection $\pi_1 : W_p \times S_p \to W_p$.

Each μ_n, $n \geq 1$, is supported in the union of the iterates $f^j(U)$ over all $0 \leq j \leq n-1$. We denote Γ_j^c the union of the connected components of

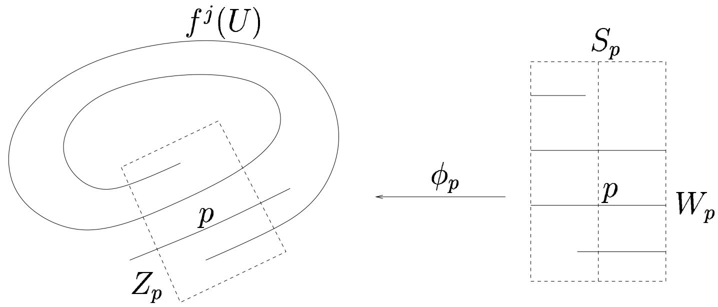

Figure 10: Crossing and non-crossing components

$f^j(U) \cap Z_p$ that cross Z_p, and Γ_j^{nc} the union of all the other components. Then we write the restriction of μ_n to Z_p as

$$(\mu_n \mid Z_p) = \mu_n^c + \mu_n^{nc},$$

where μ_n^c is the part of the measure μ_n that sits on components crossing Z_p, and μ_n^{nc} is the part of μ_n sitting on non-crossing components:

$$\mu_n^c = \frac{1}{n}\sum_{j=0}(f_*^j m_U) \mid \Gamma_j^c \quad \text{and} \quad \mu_n^{nc} = \frac{1}{n}\sum_{j=0}(f_*^j m_U) \mid \Gamma_j^{nc}.$$

Firstly, we prove that the total mass of μ_n^{nc} goes to zero as n goes to infinity.

Lemma 7.9. *We have $\mu_n^{nc}(M) = \mu_n^{nc}(Z_p) \to 0$ as $n \to \infty$.*

Proof. Let $j \geq 0$ and z be any point in Γ_j^{nc}. Recall that $\Gamma_j^{nc} \subset Z_p \cap f^j(U)$. Since Z_p can be written as a disjoint union

$$Z_p = \bigcup_{y \in S_p} \phi_p(W_p \times \{y\}),$$

there exists a unique $y \in S_p$ such that $z \in \phi_p(W_p \times \{y\})$. Then the connected component of $Z_p \cap f^j(U)$ which contains z is a subset of the strong-unstable disk $\phi_p(W_p \times \{y\})$. Since this component does not cross Z_p, the disk $f^j(U)$ can not contain $\phi_p(W_p \times \{y\})$. Therefore, there exists some $z_0 \in \phi_p(W_p \times \{y\})$ that is on the boundary of $f^j(U)$. In particular, $d_u(z, z_0) \leq \delta_0$, where $\delta_0 > 0$ is any upper bound for the d_u-diameter of the $\phi_p(W_p \times \{y\})$ over all $y \in S_p$

and $p \in \Lambda$. In other words, we have proved that Γ_n^{nc} is contained in the d_u-neighbourhood of radius δ_0 of the boundary of $f^j(U)$ inside the corresponding strong-unstable leaf.

Since Df expands distances uniformly along strong-unstable leaves, cf. (26), we may conclude that $f^{-j}(\Gamma_j^{nc})$ is contained in the d_u-neighbourhood $N(\partial U, C\lambda^j\delta_0)$ with radius $C\lambda^j\delta_0$ of the boundary of U. Therefore,

$$f_*^j m_U(\Gamma_j^{nc}) = m_U(f^{-j}(\Gamma_j^{nc})) \leq m_U(N(\partial U, C\lambda^j\delta_0))$$

for every $j \geq 0$. The last term $m_U(N(\partial U, C\lambda^j\delta_0))$ converges to $m_U(\partial U) = 0$ as $j \to \infty$, because $\lambda < 1$. So $f_*^j m_U(\Gamma_j^{nc})$ converges to zero as well, as $j \to \infty$. As a consequence,

$$\mu_n^{nc}(M) = \frac{1}{n}\sum_{j=0}^{n-1} f_*^j m_U(\Gamma_j^{nc}) \to 0 \quad \text{as } j \to \infty,$$

as claimed in the lemma. □

It follows that the fraction of the μ_n that is supported on non-crossing components has no effect on the limit measure μ:

Corollary 7.10. *We have* $\phi_p^*\mu = \lim_k \phi_p^*\mu_{n_k}^c$.

Proof. Let B be any measurable subset of $W_p \times S_p$ whose boundary ∂B has zero $\phi_p^*\mu$-measure. In other words, $\mu(\phi_p(\partial B)) = 0$. Since the boundary of $\phi_p(B)$ is contained in $\partial Z_p \cup \phi_p(\partial B)$, assumption (29) implies $\mu(\partial\phi_p(B)) = 0$. As a consequence,

$$\mu(\phi_p(B)) = \lim_{k\to\infty} \mu_{n_k}(\phi_p(B)) = \lim_{k\to\infty} \mu_{n_k}^c(\phi_p(B)) + \mu_{n_k}^{nc}(\phi_p(B)).$$

By Lemma 7.9, the last term on the right converges to zero. So,

$$\mu(\phi_p(B)) = \lim_{k\to\infty} \mu_{n_k}^c(\phi_p(B))$$

for any subset B as above. This is equivalent to the claim in the corollary. □

From now on we focus our attention on crossing components. For notational simplicity, we write $\eta = \phi_p^*\mu$ and $\eta_n = \phi_p^*\mu_n^c$, and we let $\hat\eta$ and $\hat\eta_n$ represent the quotient measures on S_p:

$$\hat\eta(\xi) = \phi_p^*\mu(W_p \times \xi) \quad \text{and} \quad \hat\eta_n(\xi) = \phi_p^*\mu_n^c(W_p \times \xi).$$

Lemma 7.11. *There exists $C_1 > 1$, depending only on the diffeomorphism f, such that*

$$\frac{1}{C_1}\,\frac{m_u(A)}{m_u(W_p)}\,\hat{\eta}_n(\xi) \le \phi_p^* \mu_n^c(A \times \xi) \le C_1\,\frac{m_u(A)}{m_u(W_p)}\,\hat{\eta}_n(\xi)$$

for any measurable sets $A \subset W_p$ and $\xi \subset S_p$.

Proof. We explain how to obtain the upper inequality, the lower one is analogous. The main step is the following

Claim: There exists $C_1 > 1$, depending only on f, such that

$$\frac{f_*^j m_U\big(\phi_p(A \times \xi) \cap \gamma\big)}{f_*^j m_U(\gamma)} \le C_1\,\frac{m_u(A)}{m_u(W_p)}$$

for every $0 \le j \le n-1$, and every connected component γ of $Z_p \cap f^j(U)$ crossing Z_p and intersecting $\phi_p(A \times \xi)$.

Proof. Since γ crosses Z_p, there exists $y \in S_p$ such that $\gamma = \phi_p(W_p \times \{y\})$. Note that γ intersects $\phi_p(A \times \xi)$ if and only if $y \in \xi$ and, in that case,

$$(A \times \xi) \cap (W_p \times \{y\}) = (A \times \{y\}).$$

Now we use bounded distortion. Let δ_0 be an upper bound for the d_u-diameter of the strong-unstable disks $\phi_p(W_p \times \{y\})$, over all $y \in S_p$ and $p \in \Lambda$. By Corollary 7.7 there exists a constant $C_2 = K(\delta_0) > 0$ such that

$$\frac{f_*^j m_U\big(\phi_p(A \times \{y\})\big)}{f_*^j m_U\big(\phi_p(W_p \times \{y\})\big)} \le C_2 \frac{m_{f^j(U)}\big(\phi_p(A \times \{y\})\big)}{m_{f^j(U)}\big(\phi_p(W_p \times \{y\})\big)}\,. \tag{30}$$

Recall that, according to property (A2) stated at the end of Subsection 5.2, $\phi_{p,y} = (\phi_p \mid W_p \times \{y\})$ is a diffeomorphism of $W_p \times \{y\}$ onto γ. Recall also that $m_{f^j(U)}$ is just the restriction of the Riemannian volume m_u to $f^j(U)$. By the mean value theorem,

$$\frac{m_{f^j(U)}\big(\phi_p(A \times \{y\})\big)}{m_{f^j(U)}\big(\phi_p(W_p \times \{y\})\big)} = \frac{|\det D\phi_{p,y}(x_1, y)|}{|\det D\phi_{p,y}(x_2, y)|}\,\frac{m_u(A)}{m_u(W_p)} \tag{31}$$

for some $x_1, x_2 \in W_p$. The quotient of the Jacobians is uniformly bounded:

$$\frac{|\det D\phi_{p,y}(z_1, y)|}{|\det D\phi_{p,y}(z_2, y)|} \le C_3 \tag{32}$$

for some $C_3 > 1$ and every z_1, z_2 in W_p. Moreover, by properties (A3) and (A1) of foliated charts, the diffeomorphisms $\phi_{p,y}$ vary continuously with the point y, and $\phi_{p,p}$ is, essentially, the identity on W_p. Since Λ is compact, these facts ensure that the constant C_3 may be chosen depending only on f (neither on y nor on p). Taking $C_1 = C_2 C_3$, the claim follows from (30), (31), (32). □

Going back to proving the lemma, we write

$$\mu_n^c\big(\phi_p(A \times \xi)\big) = \frac{1}{n} \sum_{j=0}^{n-1} \sum_{\gamma} f_*^j m_U(\phi_p(A \times \xi) \cap \gamma), \tag{33}$$

where the last sum is over the components γ of $Z_p \cap f^j(U)$ crossing Z_p that intersect $\phi_p(A \times \xi)$. As noted in the proof of the Claim, any connected component crossing Z_p may be written as $\gamma = \phi_p(W_p \times \{y\})$. Moreover, it intersects $\phi_p(A \times \xi)$ if and only if $y \in \xi$. Clearly, the last condition does not depend on the set A. So, replacing A by W_p in (33) gives

$$\mu_n^c\big(\phi_p(W_p \times \xi)\big) = \frac{1}{n} \sum_{j=0}^{n-1} \sum_{\gamma} f_*^j m_U(\gamma) \tag{34}$$

where the last sum runs over the same subset of connected components γ of $Z_p \cap f^j(U)$ as in (33). Now, using the Claim to compare the sums in (33) and (34) term by term, we find

$$\frac{\phi_p^* \mu_n^c(A \times \xi)}{\hat{\eta}_n(\xi)} = \frac{\mu_n^c\big(\phi_p(A \times \xi)\big)}{\mu_n^c\big(\phi_p(W_p \times \xi)\big)} \leq C_1 \frac{m_u(A)}{m_u(W_p)},$$

which is what we wanted to prove. □

Corollary 7.12. *If $\mu = \lim_k \mu_{n_k}$, $p \in \Lambda$, and $\phi_p : W_p \times S_p \to Z_p$ are as as above, then $\phi_p^* \mu$ is absolutely continuous with respect to normalized Lebesgue measure $m_u / m_u(W_p)$ on W_p along the horizontal foliation. In addition, the density ρ is uniformly bounded away from zero and infinity: $1/C_1 \leq \rho \leq C_1$.*

Proof. This follows from Propositions 7.3 and 7.4, with $X = W_p$, $Y = S_p$, $\phi_p^* \mu_{n_k}^c$ in the role of μ_k, $\eta = \phi_p^* \mu$ in the role of μ, $\nu = m$, $\psi = C_1$, $\phi = 1/C_1$, and $\mathcal{R}$ being the family of all measurable rectangles $A \times \xi \subset W_p \times S_p$. Lemma 7.11 states that the assumptions of the two propositions are satisfied. Recall also Corollary 7.10.

The propositions state that $1/C_1 \le \rho \le C_1$ on a subset with full $m \times \hat{\eta}$-measure. Then, modifying the values of ρ on the complement if necessary (since the complement has measure zero, the new function is again a density for $\phi_p^*\mu$), we may suppose that $1/C_1 \le \rho \le C_1$ everywhere. $\square$

Theorem 7.1 is contained in the following proposition, that summarizes the main facts we proved in this section.

Proposition 7.13. *For any accumulation point μ of $\mu_n = n^{-1}\sum_{j=0}^{n-1} f_j^* m_U$, and any point $p \in \Lambda$, there exist foliated charts ϕ_p for the strong-unstable foliation $\mathcal{F}^u$ at p, so that $\phi_p^*\mu$ is absolutely continuous with respect to normalized Lebesgue measure along the horizontal foliation, with density ρ uniformly bounded away from zero and infinity by constants that depend only on the diffeomorphism f.*

Remark 7.6. The conclusion of Theorem 7.1 is also valid for the accumulation points of

$$\frac{1}{n}\sum_{j=0}^{n-1} f_*^j(\varphi\,\mathrm{Leb}),$$

where φ is any non-negative function supported in a sufficiently small neighbourhood of Λ, and Leb is Lebesgue measure in the ambient manifold M. See Theorem 3 of [46].

8 Sinai-Ruelle-Bowen Measures

The following fundamental result was first proved by Sinai [57] for Anosov diffeomorphisms, and by Ruelle and Bowen [12], [**?**], [9] for general hyperbolic attractors of diffeomorphisms or flows.

Theorem 8.1. *Suppose that Λ is a hyperbolic attractor for a C^2 diffeomorphism $f : M \to M$. Then there exists a unique invariant probability measure μ supported in Λ that is absolutely continuous along $\mathcal{F}^u$. This measure is ergodic and its support coincides with Λ. Moreover, μ is a physical measure for f, in fact, $B(\mu)$ is a full Lebesgue measure subset of the basin $B(\Lambda)$ of Λ.*

Let us begin by giving a sketch of the proof, the details are carried out in Subsection 8.1. A similar approach can be applied to certain partially hyperbolic (non-hyperbolic) attractors, to prove existence and finiteness of SRB measures, see the last subsection.

First, we prove that any invariant measure absolutely continuous along the unstable foliation splits into finitely many ergodic components, that are also absolutely continuous along $\mathcal{F}^u$. For this purpose we introduce an equivalence relation, the *accessibility relation*, such that time averages are constant on each equivalence class. This is defined in the (full measure) set of regular points, which are the points at which forward and backward time averages exist and coincide, for any continuous function. Accessibility is the smallest equivalence relation such that regular points in a same stable or unstable manifold, or in a same orbit, are in a same accessibility class.

There are only finitely many accessibility classes having positive weight for some invariant measure μ absolutely continuous along $\mathcal{F}^u$. The ergodic components of any such μ are its normalized restrictions to those accessibility classes. As part of the proof we also get that ergodic measures absolutely continuous along $\mathcal{F}^u$ have the SRB property. This is based on the fact that the stable foliation $\mathcal{F}^s$ is absolutely continuous, cf. Theorem 6.6. Using, for the first time, the assumption that f is transitive on Λ, we show that the accessibility class is, actually, unique. In this way we conclude that μ is ergodic and unique. Moreover, its basin fills-in a full Lebesgue measure subset of the whole basin of the attractor.

8.1 Hyperbolic attractors

Definition 8.1. A point $x \in M$ is *regular* if the forward and backward time averages

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^{-j}(x))$$

exist and coincide, for every continuous function $\varphi : M \to \mathbb{R}$.

Note that, according to the ergodic theorem of Birkhoff, the set $\mathcal{R}$ of regular points has full measure with respect to any probability measure that is invariant under f.

Definition 8.2. Given $x, y \in \mathcal{R}$, we set $x \approx y$ if there exist $N \geq 1$, regular points $x = z_0, z_1, \ldots, z_{N-1}, z_N = y$, and integers $k_1, \ldots, k_N$, such that

$$z_i \in W^u(f^{k_i}(z_{i-1})) \cup W^s(f^{k_i}(z_{i-1}))$$

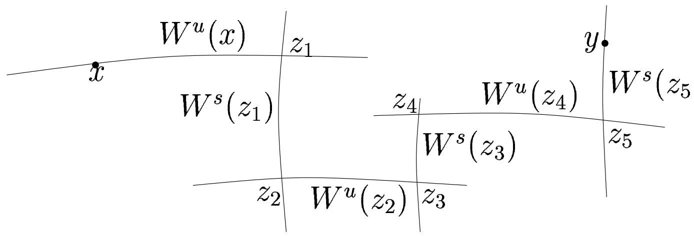

Figure 11: The accessibility relation

for every $i = 1, \ldots, N$.

See Figure 11, corresponding to a situation with $k_1 = \cdots = k_N = 0$. It is easy to check that $\approx$ is an equivalence relation. We refer to the equivalence classes as *accessibility classes*. Note that, by definition, they are invariant sets for f.

The usefulness of this notion stems from the following simple observation.

Lemma 8.2. *The time averages of any continuous function $\varphi : M \to \mathbb{R}$ are constant on each accessibility class.*

Proof. Suppose x and y are in a same stable set, that is, if $d(f^j(x), f^j(y)) \to 0$ as $j \to +\infty$. Then $|\varphi(f^j(x)) - \varphi(f^j(y))| \to 0$ as $j \to +\infty$, and so the two points have the same forward time average:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(y)).$$

Similarly, points in a same unstable set have the same backward time averages. It is also clear that x and y are iterates of each other then they have the same time averages, both in the future and in the past. Now suppose $x \approx y$, and let $x = z_0, z_1, \ldots, z_N = y$ be as in Definition 8.2. Since all these points are assumed to be regular, the previous remarks show that they all have the same time averages for φ. $\square$

Corollary 8.3. *If an invariant measure μ gives positive weight to an accessibility class $\mathcal{A}$ then its normalized restriction to $\mathcal{A}$ is an ergodic measure, and it does not depend on μ.*

Proof. $\mathcal{A}$ has full measure for this normalized restriction $\mu_{\mathcal{A}}$, and we have shown in the lemma that time averages are constant on it. So

$$\mu_{\mathcal{A}} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(z)} \quad \text{for any } z \in \mathcal{A}.$$

This proves that μ_A is ergodic and independent of μ. □

Everything we said so far in this subsection holds in great generality, e.g, for homeomorphisms in metric spaces. Now we focus again on the context of C^2 diffeomorphisms admitting a hyperbolic attractor Λ.

We use $\phi'_p : S'_p \times W'_p \to Z'_p$ as a generic notation for a foliated chart of the stable foliation $\mathcal{F}^s$ of Λ at the point p. Recall that, by definition, W'_p is contained in the local stable manifold of p, whereas S'_p is the intersection of Λ with some cross section Σ'_p of $\mathcal{F}^s$ at p. For simplicity, we always assume that Σ'_p is contained in the local unstable manifold of p. In that case, S'_p coincides with Σ'_p, cf. Theorem 5.8, and they are contained in Λ.

Lemma 8.4. *There exists $r > 0$ such that, given any accessibility class $\mathcal{A}$ in Λ that intersects some unstable leaf in a positive m_u-measure subset, there exists $p_{\mathcal{A}} \in \Lambda$ so that m_u-almost every point z the d_u-ball of radius r around $p_{\mathcal{A}}$ is in the stable manifold of some point of $\mathcal{A}$.*

Proof. We fix $r > 0$ so that the d_u-ball of radius r around any point $p \in \Lambda$ is contained in the interior of Z'_p, where Z'_p is the image of some foliated chart for the stable foliation $\mathcal{F}^s$ at p. Let F be any leaf of $\mathcal{F}^u$ for which $A = \mathcal{A} \cap F$ has positive m_u-measure. Note that $f^n(A) \subset \mathcal{A}$ for every $n \geq 1$, because $\mathcal{A}$ is invariant. So, according to Proposition 7.8, there exist points $p_n \in \mathcal{A}$ such that

$$\frac{m_u\big(B^u_r(p_n) \setminus \mathcal{A}\big)}{m_u\big(B^u_r(p_n)\big)} \to 0 \quad \text{as } n \to \infty. \tag{35}$$

By compactness, we may suppose that the sequence $(p_n)_n$ converges to some point $p_{\mathcal{A}} \in \Lambda$. In particular, the $B^u_r(p_n)$ converge to the d_u-ball of radius r around $p_{\mathcal{A}}$. In view of our choice of r, $B^u_r(p_{\mathcal{A}})$ is contained in Z'_p, and so is $B^u_r(p_n)$ for every large n. It follows from (35) and the absolute continuity property of $\mathcal{F}^s$, Theorem 6.6, that the local stable manifolds through points in $\mathcal{A} \cap B^u_r(p_n)$ intersect $B^u_r(p)$ in a subset G_n such that

$$\frac{m_u(B^u_r(p_{\mathcal{A}}) \setminus G_n)}{m_u(B^u_r(p_{\mathcal{A}}))} \leq C_0 \frac{m_u(B^u_r(p_n) \setminus \mathcal{A})}{m_u(B^u_r(p_n))} \to 0$$

as $n \to \infty$. Let $G = \cup_n G_n$. Then G has full m_u-measure in $B^u_r(p_{\mathcal{A}})$, and any $z \in G$ is in the stable manifold of some point in $\mathcal{A}$. □

Lemma 8.5. *Given $r > 0$ there exists $s > 0$ so that the following holds. Let $\mathcal{A}$ be an accessibility class and $p \in \Lambda$ be such that m_u-almost every point in the d_u-ball of radius r around p is in the stable manifold of some point of $\mathcal{A}$. Then the same is true for*

1. *Lebesgue almost every point in the s-neighbourhood $B_s(p)$ of p in the ambient manifold M,*
2. *and m_u-almost every point in $F \cap B_s(p)$, for any unstable leaf F.*

Proof. We choose $s > 0$ small enough so that, given any point $p \in \Lambda$, there exists a foliated chart $\phi'_p : S'_p \times W'_p \to Z'_p$ for the stable foliation at p such that Z_p contains $B_s(p)$. We reduce s if necessary, so that $\pi_1(B_s(p)) \subset B^u_r(p)$ for every $p \in \Lambda$, where π_1 is the projection onto S'_p along the stable leaves inside Z'_p. In other words, $\pi_1 = \phi_p \circ \tilde{\pi}_1 \circ \phi_p^{-1}$, where $\tilde{\pi}_1 : S'_p \times W'_p \to S'_p$ is the canonical projection.

Claim: Let Σ be a C^1 embedded disk in the $B_s(p)$ and transverse to $\mathcal{F}^s$. Then m_Σ-almost every point in Σ is in the stable manifold of some point of $\mathcal{A}$.

Proof. Let $\pi : \Sigma \to S'_p$ be the local holonomy map of $\mathcal{F}^s$ from Σ to S'_p. Our choice of s ensures that the image $\tilde{\Sigma}$ of π is contained in $B^u_r(p)$. By Theorem 6.6, both π and its inverse π^{-1} map zero Lebesgue measure sets into zero Lebesgue measure sets. So, the assumption implies that $\pi(z)$ is in the stable manifold of a point in $\mathcal{A}$, for m_Σ-almost every point. Of course, z and $\pi(z)$ are in the same stable manifold, so the conclusion follows. □

Part 2 of the lemma is a particular case of the Claim. To prove part 1 it suffices to consider any C^1 foliation $\mathcal{G}$ of $B_s(p)$ by disks transverse to $\mathcal{F}^s$. The Claim applies to each of the leaves of $\mathcal{G}$ and so, by Fubini's theorem, Lebesgue almost every point in that s-neighbourhood is in the stable manifold of some point of the accessibility class $\mathcal{A}$. □

Corollary 8.6. *1. There exist only finitely many accessibility classes $\mathcal{A}_1, \ldots, \mathcal{A}_N$ as in Lemma 8.4.*

2. *The ergodic decomposition of any invariant measure μ absolutely continuous along $\mathcal{F}^u$ is given by $\mu = \sum_i \mu(\mathcal{A}_i)\mu_i$ where the sum is over the values of i such that $\mu(\mathcal{A}_i) > 0$, and each μ_i is the normalized restriction of μ to $\mathcal{A}_i$.*

Proof. Suppose there are infinitely many distinct accessibility classes $\mathcal{A}_i$, $i \geq 1$. Then, combining Lemmas 8.4 and 8.5, there is $s > 0$ and there are points $p_i \in \Lambda$, $i \geq 1$, such that the union of the stable manifolds of the points in $\mathcal{A}_i$ contains a full Lebesgue measure subset of the s-neighbourhood of p_i, for every $i \geq 1$. By compactness, we may suppose that the p_i converge to some $p \in \Lambda$. Then, for every large i and j, the neighbourhoods $B_s(p_i)$ and $B_s(p_j)$ intersect each other in a set with positive Lebesgue measure. In particular, there exists $q \in B_s(p_i) \cap B_s(p_j)$ that is in the stable manifolds of points $q_i \in \mathcal{A}_i$ and $q_j \in \mathcal{A}_j$. This implies that q_i and q_j are in a same accessibility class, contradicting the assumption that $\mathcal{A}_i$ and $\mathcal{A}_j$ are different. Part 1 of the corollary is proved.

Now we prove part 2. In view of Corollary 8.3, we only have to show that the union of all the accessibility classes as in Lemma 8.4 has full measure for μ. In equivalent terms, it suffices to prove that any measurable subset B of Λ with $\mu(B) > 0$ intersects at least one accessibility class $\mathcal{A}$ that contains a positive m_u-measure subset of some unstable leaf.

Indeed, given any measurable set $B \subset \Lambda$ with $\mu(B) > 0$, there exist $p \in \Lambda$ and a foliated chart $\phi_p : W_p \times S_p \to Z_p$ for $\mathcal{F}^u$ at p such that $B \cap Z_p$ has positive μ-measure (consider some finite covering of Λ by images of foliated charts). As the set $\mathcal{R}$ of regular points has full measure, we also have $\mu(B \cap Z_p \cap \mathcal{R}) > 0$. This can be rewritten as

$$\phi_p^*\mu\big(\phi_p^{-1}(B \cap Z_p \cap \mathcal{R})\big) > 0.$$

Then, since $\phi_p^*\mu$ is absolutely continuous along the horizontal foliation of $W_p \times S_p$, there exists $y \in S_p$ such that $\phi_p^{-1}(B \cap Z_p \cap \mathcal{R}) \cap (W_p \times \{y\})$ has positive m_u-measure. Taking the image under the embedding $\phi_p \mid (W_p \times \{y\})$,

$$m_u\big((B \cap Z_p \cap \mathcal{R}) \cap \phi_p(W_p \times \{y\})\big) > 0.$$

The set $(B \cap Z_p \cap \mathcal{R}) \cap \phi_p(W_p \times \{y\})$ is contained in the accessibility class of any of its points, since it is contained in an unstable disk $\phi_p(W_p \times \{y\})$. The last inequality shows that this accessibility class intersects the unstable disk in a subset with positive m_u-measure. $\square$

Corollary 8.7. *Any ergodic measure μ absolutely continuous along the unstable foliation is a physical measure for f.*

Proof. By Corollary 8.6, μ is ergodic if and only if there is an accessibility class $\mathcal{A}_i$ so that $\mu(\mathcal{A}_i) = 1$. Moreover, cf. Corollary 8.3, in that case μ is given by the limit of $n^{-1} \sum_{j=0}^{n-1} \delta_{f^j(z)}$ as $n \to \infty$, for any point $z \in \mathcal{A}_i$. By Lemma 8.5, there exists $p_i \in \Lambda$ such that Lebesgue almost every point in the s-neighbourhood of p is in the stable manifold of some point of $\mathcal{A}_i$. Since points in a same stable manifold have the same forward time averages, it follows that

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(w)}$$

for a full Lebesgue measure subset of $B_s(p_i)$. This means that the basin of μ contains that subset of $B_s(p_i)$, and so it has positive Lebesgue measure. □

Observe that Theorem 7.1, together with the second part of Corollary 8.6, imply that there does exist at least one accessibility class as in Lemma 8.4. Now we use the assumption that f is transitive on Λ to show that, in fact, such an accessibility class is unique. Thus, f has a unique physical measure supported in Λ.

Lemma 8.8. *Let $p \in \Lambda$ and $\mathcal{A}$ be any accessibility class as in Lemma 8.4. Then m_u-almost every point in $B^u_r(p)$ is in the stable manifold of some point of $\mathcal{A}$.*

Proof. Let $\phi'_p : S'_p \times W'_p \to Z'_p$ be a foliated chart for the stable foliation at p, such that Z'_p contains the r-neighbourhood of p (recall the choice of r in Lemma 8.4). By the second part of Lemma 8.5, there is $p_{\mathcal{A}} \in \Lambda$ such that m_u-almost every point in $F \cap B_s(p_{\mathcal{A}})$ is in the stable manifold of some point of $\mathcal{A}$, for any unstable leaf F intersecting the s-neighbourhood $B_s(p_{\mathcal{A}})$. The fact that f is transitive on Λ implies that there exist points $q_k \to p_{\mathcal{A}}$ and times $n_k \to \infty$ such that $f^{n_k}(q_k) \to p$. In particular, q_k is in the $s/2$-neighbourhood of $p_{\mathcal{A}}$ and $B_r(f^{n_k}(q_k))$ is contained in Z'_p, for every large k. Then, m_u-almost every point in the d_u-ball of radius $s/2$ around q_k is in the stable manifold of some point in $\mathcal{A}$. By the invariance of $\mathcal{A}$, the same is true for m_u-almost every point in the f^{n_k}-image of this ball. Increasing n_k if necessary, this image contains the d_u-ball $B^u_r(f^{n_k}(q_k))$. So, we have shown that p is accumulated by points $f^{n_k}(p_k)$ such that m_u-almost every point in their d_u-balls of radius r are in the stable manifold of some point in $\mathcal{A}$. The same argument as in Lemma 8.4 shows that this remains true for the d_u-ball of radius r around p. □

Corollary 8.9. *There exists exactly one accessibility class that intersects some unstable manifold in a positive m_u-measure set.*

Proof. In view of the remarks preceding Lemma 8.8, we only have to prove uniqueness. Let $p \in \Lambda$, and $\mathcal{A}$, $\mathcal{B}$, be any accessibility classes as in the hypothesis. By the previous lemma, the stable sets of $\mathcal{A}$ and $\mathcal{B}$ fill-in full measure subsets of $B_r^u(p)$. In particular, they must intersect each other, and this implies that $\mathcal{A} = \mathcal{B}$. □

Corollary 8.10. *There exists a unique invariant probability measure μ absolutely continuous along $\mathcal{F}^u$, and μ is ergodic. Moreover, the basin of μ contains a full Lebesgue measure subset of the basin of Λ. Consequently, μ is also the unique SRB measure of f in Λ.*

Proof. Uniqueness and ergodicity follow from Corollaries 8.6 and 8.9:

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(z)} \tag{36}$$

for any z in the unique accessibility class $\mathcal{A}$.

Lemmas 8.5 and 8.8 imply that Lebesgue almost every point in the s-neighbourhood of any $p \in \Lambda$ is in the stable manifold of some $z \in \mathcal{A}$. So, (36) is true also for Lebesgue almost every point in the s-neighbourhood of the attractor Λ. Since the basin $B(\Lambda)$ is the union of the pre-images of this neighbourhood, we have (36) for Lebesgue almost every point in $B(\Lambda)$. The last statement is an immediate consequence. □

8.2 Attractors of Type $E^u \oplus E^{cs}$

It is easy to see that a (transitive) partially hyperbolic attractor may support infinitely many ergodic probability measures absolutely continuous along the strong-unstable foliation. Moreover, in general they are *not* SRB measures for the map.

Example 8.3. Let $f_1 : M_1 \to M_1$ and $f_2 : M_2 \to M_2$ be Anosov diffeomorphisms, with splittings $TM_i = E_i^u \oplus E_i^s$ for $i = 1, 2$. Then $f_1 \times f_2$ is an

Anosov diffeomorphism on $M_1 \times M_2$. Taking f_1 and f_2 to be transitive, i.e. so that there is some point whose forward orbit is dense in the corresponding manifold, then $f_1 \times f_2$ is also transitive. See e.g. Remark 9.3. Assume that the contraction of f_1 along E_1^s is stronger than the contraction of f_2 along E_2^s, and the expansion of f_1 along E_1^u is also stronger than the expansion of f_2 along E_2^u. More precisely, there exists $\lambda < 1$ such that

$$\|Df_1 \mid E_1^s\| \, \|(Df_2 \mid E_2^s)^{-1}\| \leq \lambda \quad \text{and} \quad \|Df_2 \mid E_2^u\| \, \|(Df_1 \mid E_1^u)^{-1}\| \leq \lambda.$$

Then we may also think of $M_1 \times M_2$ as a strongly partially hyperbolic attractor for $f_1 \times f_2$, with splitting

$$T(M_1 \times M_2) = \big(E_1^u \times \{0\}\big) \oplus TM_2 \oplus \big(E_1^s \times \{0\}\big).$$

The foliation $\mathcal{F}^u$ tangent to the strong-unstable bundle $E_1^u \times \{0\}$ is given by $\mathcal{F}^u(x_1, x_2) = \mathcal{F}_1^u(x_1) \times \{x_2\}$, where $\mathcal{F}_1^u$ is the unstable foliation of f_1. Let μ_i be the SRB measure of f_i for $i = 1, 2$, and ν be an f_2-invariant measure supported on a periodic orbit of f_2. That is, ν is the average of the Dirac measures supported on the points of the periodic orbit. Then $\mu_1 \times \nu$ is an invariant ergodic measure for $f_1 \times f_2$, absolutely continuous along $\mathcal{F}^u$. But the Anosov diffeomorphism $f_1 \times f_2$ has a unique SRB measure $\mu_1 \times \mu_2$.

There are, however, relevant situations in which SRB measures can be constructed via Theorem 7.1. The following sufficient condition was proposed by [**?**], extending [**?**]. It holds for a C^1 open set of diffeomorphisms with partially hyperbolic, possibly non-hyperbolic, attractors.

Let Λ be a partially hyperbolic attractor of type $E^u \oplus E^{cs}$ for a diffeomorphism $f : M \to M$. We say that E^{cs} is *mostly contracting* if $Df^n \mid E^{cs}$ is asymptotically contracting, as $n \to \infty$, over a large set of points: given any domain U inside a strong-unstable leaf, there exists a positive Lebesgue measure subset U_0 of U, such that

$$\limsup_{n \to +\infty} \frac{1}{n} \log \|Df^n \mid E_x^{cs}\| < 0 \tag{37}$$

for every $x \in U_0$.

Theorem 8.11. *Let Λ be a partially hyperbolic attractor of type $E^u \oplus E^{cs}$ for a C^2 diffeomorphism, such that E^{cs} is mostly contracting. Then any ergodic measure absolutely continuous along the strong-unstable foliation $\mathcal{F}^u$*

is an SRB measure. There are finitely many such measures and their basins contain a full Lebesgue measure subset of the basin of the attractor Λ. *If the orbit of any strong-unstable leaf is dense in the attractor then the SRB measure is unique.*

In the situation treated by [?], the partially hyperbolic attractor is obtained from an initial Anosov diffeomorphism $f_0 : M \to M$, through a deformation by isotopy. In the simplest case, M is the 3-dimensional torus and f_0 has hyperbolic splitting $TM = E_0^u \oplus E_0^s$ with $\dim E_0^s = 2$. One deforms f_0 close to a fixed point p_0, weakening the contraction along the stable direction while keeping the unstable direction essentially unchanged. At some stage of the deformation, the continuation of the fixed point p_0 goes through a Hopf bifurcation, after which it becomes a repeller (source). The complement Λ of the unstable set of this repeller is a partially hyperbolic attractor for the modified diffeomorphism f. It has a splitting $TM = E^u \oplus E^{cs}$ such that $\dim E^{cs} = 2$, and the orbit of any strong-unstable leaf is dense in Λ.

In Theorem 8.11 it is not necessary to suppose that Λ is an attractor, nor that it is transitive: under condition (37), the conclusion holds for any partially hyperbolic set consisting of entire strong-unstable leaves. Compare Proposition 6.3.

On the other hand, even if Λ is a transitive attractor, there may be more than one SRB measure supported on it. Indeed, [26] constructed diffeomorphisms of $T^2 \times [0,1]$ having two SRB measures (one supported on each boundary component), whose basins are both dense, their union covering a full Lebesgue measure subset. For these maps, $\Lambda = T^2 \times [0,1]$ is transitive and strongly partially hyperbolic, with $E^{cs} = E^c \oplus E^s$ mostly contracting. Similar examples can be constructed in T^3, gluing two maps on $T^2 \times [0,1]$ as above along the boundary. Most interesting, in $T^2 \times [0,1]$, the construction of [26] is robust: any nearby diffeomorphism is transitive and has two SRB measures. This relies on the fact that diffeomorphisms of $T^2 \times [0,1]$ have to preserve the boundary of the manifold.

For manifolds without boundary, it is a very interesting open problem whether there are robust examples of coexistence of several SRB measures on a same transitive attractor. For instance, let $\Lambda = \cap_{n \geq 0} f^n(U)$ be a partially hyperbolic attractor of a C^2 diffeomorphism f. Here U is an open neighbourhood of Λ such that the closure of $f(U)$ is contained in U. Assume that for every g in a neighbourhood of f, the maximal invariant set $\Lambda_g = \cap_{n \geq 0} g^n(U)$ is transitive and partially hyperbolic of type $E^u \oplus E^{cs}$, with

E^{cs} mostly contracting. Is it true that *for the generic diffeomorphism* g *close to* f *the attractor* Λ_g *supports a unique SRB measure ?*

9 Comments on Global Dynamics

We include in this subsection a few comments on global aspects of Dynamical Systems, in order to place the previous results into their broad context.

The notion of hyperbolicity was introduced in Dynamics by Smale in the early sixties, see [58], with the aim of characterizing structural stability and, hopefully, proving that most systems are structurally stable. This last expectation turned out to be unfunded, as many important systems do not fall into this class, including partially hyperbolic maps and flows. However, hyperbolicity did provide a powerful framework for analyzing *robustly complicated dynamics*: hyperbolic systems usually exhibit an infinite number of periodic orbits, as well as stochastic behaviour of typical trajectories, and this is now rather well understood in the hyperbolic setting. Moreover, hyperbolicity indeed proved to be the crucial ingredient for structural stability.

A central goal ever since has been to extend as much as possible of the conclusions of the hyperbolic theory to "most" dynamical systems. In this regard, crucial input came from the study of experimental systems, as well as of models for their behaviour, often carried out numerically. In recent years a new point of view emerged, with a distinctly more probabilistic flavour. One focus on aspects of the dynamical behaviour that persist for many systems, in terms of probability in parameter space. Also, important notions of stability of the dynamics are also formulated in probabilistic terms.

In this spirit, a comprehensive program towards a global theory of Dynamics was proposed a few years ago by Palis, see [**?**] and also [45]. First of all, he conjectured that

- every system can be approximated by another having only finitely many attractors, whose basins cover a full Lebesgue measure subset of the ambient manifold.

Here an *attractor* is a compact invariant set Λ whose stable set, or *basin*, has positive Lebesgue measure (in the relevant known cases, the basin is a neighbourhood of the attractor). Moreover,

- these attractors should have good statistical properties, including existence of physical measures whose basins cover a full measure subset of the attractor's stable set;

- the dynamics on the basin of each attractor should be stable, in a statistical sense: time averages along pseudo-orbits, obtained by randomly perturbing the system at each iterate, are close to the time averages along the orbits of the original system, if the random perturbations are small. See e.g. [**?**] for a precise definition.

We call an attractor *robust* if it can not be destroyed by any small perturbation of the dynamical system. More generally, we have to deal with attractors which are only *persistent*: they exist with positive probability in parameter space, for generic parametrized families through the original system. Hénon-like attractors [6], [36], are an important model of persistent, yet non-robust attractors. Hyperbolic and partially hyperbolic attractors, as well as Lorenz-like attractors of flows, provide important examples of robust attractors. In fact, some amount of hyperbolicity (including existence of a dominated splitting) is a necessary condition for robustness, as we comment upon below.

In the remainder of this section we detail some of these topics. To start with, we recall some basic definitions and facts about Axiom A diffeomorphisms. See also [58], [41], [45].

Definition 9.1. A point p is *non-wandering* for f if for any neighbourhood V of p there exists $n \geq 1$ such that $f^n(V) \cap V \neq \emptyset$. The set of non-wandering points, or *non-wandering set*, is denoted $\Omega(f)$.

In other words, $p \in M$ is non-wandering if and only if there are points arbitrarily close to it that return arbitrarily close to p in future times (p itself may never return).

It follows directly from the definition that $\Omega(f)$ is a closed set. It is also easy to check that any point that is in the accumulation set

$$L(f, z) = \{w \in M : \text{ there exists } n_j \to \pm\infty \text{ so that } f^{n_j}(z) \to w\}$$

of some $z \in M$ is a non-wandering point. Therefore, $\Omega(f)$ always contains the *limit set* $L(f)$ of f, which is the closure of the union of all the accumulation sets for all the points $z \in M$. Note also that if z is a periodic point then $L(f, z)$ is just the orbit of z. Therefore, the closure of the set $\mathrm{Per}(f)$ of periodic points of f is always contained in $L(f)$. Summarizing,

$$\mathrm{clos}\big(\mathrm{Per}(f)\big) \subset L(f) \subset \Omega(f)$$

for any diffeomorphism (or even homeomorphism) f. The inclusions may be strict, in general.

Definition 9.2. A diffeomorphism $f : M \to M$ is *hyperbolic* (or *Axiom A*) if its non-wandering set $\Omega(f)$ is hyperbolic for f and coincides with the closure $\operatorname{clos}\big(\operatorname{Per}(f)\big)$ of the set of periodic points of f.

A fundamental property of hyperbolic systems is that the dynamics on the non-wandering set can be decomposed into finitely many hyperbolic *basic pieces*, cf. the next theorem [58]. Recall that transitivity means that the forward orbit of some point is dense.

Theorem 9.1. *If f is hyperbolic then its non-wandering set can be written as a disjoint union*

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

of isolated hyperbolic sets $\Lambda_1, \ldots, \Lambda_N$, such that the restriction of f to Λ_i is transitive for every $1 \le i \le N$.

More generally, there is a similar decomposition for the closure of the set of periodic points whenever it is a hyperbolic set for f. It should be noted that, cf. [40], if the limit set $L(f)$ is hyperbolic then $\operatorname{clos}\big(\operatorname{Per}(f)\big) = L(f)$. This is not always true for the non-wandering set.

Remark 9.3. An Anosov diffeomorphism $f : M \to M$ is transitive if and only its periodic points are dense in M. Indeed, transitivity implies $L(f)$ is the whole manifold M, and then $\operatorname{Per}(f)$ is dense in M by the result of [40] mentioned above. In the converse direction, if $\operatorname{Per}(f)$ is dense in M then f is hyperbolic and $\Omega(f)$ coincides with M. So, by Theorem 9.1, M can be split into a finite number of compact transitive sets. By connectedness, there must be exactly one such set. This means that f is transitive.

Actually, any Anosov diffeomorphism is hyperbolic [4]. Let us also mention that all known Anosov diffeomorphisms are transitive.

If $\Lambda_1 \cup \cdots \cup \Lambda_N$ is the decomposition of $\Omega(f)$, or even of $L(f)$, then every point in M is in the stable set, respectively unstable set, of some basic piece:

$$\bigcup_{i=1}^{N} W^s(\Lambda_i) = M = \bigcup_{i=1}^{N} W^u(\Lambda_i).$$

If f is C^2 then the stable set of Λ_i has positive Lebesgue measure if and only if Λ_i is a attractor, see [9]. Recall, from Theorem 5.8, that the stable set (or basin) of a hyperbolic attractor contains a neighbourhood of it. Thus, *for hyperbolic C^2 diffeomorphisms Lebesgue almost every point in M is in the basin of some attractor*. On the contrary, C^1 diffeomorphisms may exhibit transitive hyperbolic sets that are not attractors and have positive Lebesgue measure [10].

Definition 9.4. Suppose f is a hyperbolic diffeomorphism. A *cycle* in $\Omega(f)$ is a sequence of basic pieces $\Lambda_{i_0}, \ldots, \Lambda_{i_{k-1}}, \Lambda_{i_k} = \Lambda_{i_0}$ of the non-wandering set such that

$$\big(W^u(\Lambda_{i_{j-1}}) \setminus \Lambda_{i_{j-1}}\big) \cap \big(W^s(\Lambda_{i_j}) \setminus \Lambda_{i_j}\big) \neq \emptyset.$$

for $1 \le j \le k$. We say that f has *no cycles* if there are no cycles in $\Omega(f)$.

If f has no cycles then there exists a *filtration* [58] for it, that is, a sequence $\emptyset = M_0 \subset M_1 \subset \cdots \subset M_N = M$ of compact submanifolds with boundary such that f maps each M_i into its interior, and the set of points whose orbits never leave $M_i \setminus M_{i-1}$ coincides with Λ_i for every $1 \le i \le N$:

$$\Lambda_i = \bigcap_{n \in \mathbb{Z}} f^n(M_i \setminus M_{i-1}). \tag{38}$$

If the limit set of f is hyperbolic, one defines cycles in it in the same way as in Definition 9.4. As before, if there are no cycles in $L(f)$ then there is a filtration $M_1, \ldots, M_N$ such that the basic pieces of the limit set coincide with the maximal invariant sets in each $M_i \setminus M_{i-1}$. Now, existence of such a filtration forces the non-wandering set to coincide with $L(f)$. Therefore, if the limit set is hyperbolic and there are no cycles in it, then $\Omega(f) = L(f)$, and so f is a hyperbolic diffeomorphism with no cycles [40]. Clearly, the converse is also true.

Another important conclusion is that hyperbolic diffeomorphisms with no cycles are C^r *Ω-stable* for any $r \ge 1$: any g in a small C^r neighbourhood of f is topologically conjugate to f, restricted to their non-wandering sets. That is, there exists a homeomorphism $h : \Omega(f) \to \Omega(g)$ such that

$$(f \mid \Omega(f)) \circ h = h \circ (g \mid \Omega(g)).$$

Cf. [58], this follows from the existence of a filtration as in (38), combined with the local stability Theorem 5.2. On the other hand, a C^r Ω-stable hyperbolic diffeomorphism can not have cycles [42].

Inspired on these facts, as well as on the stability results for Anosov systems [4] and for Morse-Smale systems [44], Palis and Smale conjectured that hyperbolicity together with the no cycle condition completely characterize the Ω-stable systems. In view of the results mentioned before, this *Ω-stability conjecture* reduced to proving that C^r Ω-stable systems are hyperbolic, for every $r \geq 1$.

They also proposed a similar characterization for a more global notion of stability, due to [3]. One says that a diffeomorphism f is *C^r structurally stable*, $r \geq 1$, if any g in a small C^r neighbourhood of f is topologically conjugate to f: there exists a homeomorphism $h : M \to M$ that

$$f \circ h = h \circ g.$$

The *stability conjecture* in [44] claims that a system is C^r structurally stable in and only if it is hyperbolic and satisfies the *strong transversality condition*: the stable manifold and the unstable manifold of any two points in $\Omega(f)$ are transverse.

That hyperbolic systems satisfying the transversality condition are structurally stable was proved in the seventies by [49], [14], [51], [52], [53]. Moreover, [50] implied that the strong-transversality condition is indeed necessary for structural stability. In this way, the stability conjecture was also reduced to proving that hyperbolicity is necessary for structural stability.

The proof came only after another decade. By the mid-eighties, Mañé [34] proved the remarkable fact that C^1 structurally stable diffeomorphisms must be hyperbolic, thus settling the C^1 stability conjecture. Based on his methods, [43] extended this conclusion to the C^1 Ω-stable case. Other fundamental contributions to these problems had been given, specially by [47], [**?**], [31], [32], [54]. The extension of both conjectures for C^1 flows was achieved very recently by [23].

Some of the key key tools in the proofs, such as the closing lemma of [**?**] and the connecting lemma of [23] are available only in the C^1 topology. Their C^r versions, as well as of the results in the previous paragraph, remain outstanding open problems for any $r \geq 2$.

We close with a brief discussion on robust attractors.

Definition 9.5. We say that Λ is a *C^r robust attractor* for f if there exists a neighbourhood U of Λ such that

$$\operatorname{clos}(f(U)) \subset U \quad \text{and} \quad \Lambda = \bigcap_{n=0}^{\infty} f^n(U), \tag{39}$$

and $\Lambda_g = \cap_{n=0}^{\infty} g^n(U)$ is a transitive set for any diffeomorphism g in a C^r neighbourhood of f. When $r = 1$ we just call Λ a robust attractor.

Hyperbolic attractors are robust attractors, as we have seen in the previous section. The converse is true in two dimensions, according to [31]: robust attractors of surface diffeomorphisms are always hyperbolic. On the other hand, non-hyperbolic robust attractors were exhibited by [55], in dimension at least 4, and by [31], in dimension 3 or larger. Several other constructions were proposed in recent years, see for instance [8] and [**?**], and references therein.

Remarkably, robustness does imply some weak form of hyperbolicity: [**?**] proved recently that robust attractors in any dimension always admit an invariant dominated splitting. Before that [15] had shown that robust attractors of diffeomorphisms in 3-dimensional manifolds are always partially hyperbolic.

Definition 9.5 extends to flows X^t, $t \in \mathbb{R}$, with (39) replaced by

$$\mathrm{clos}(X^t(U)) \subset U \text{ for } t > 0, \quad \text{and} \quad \Lambda = \bigcap_{t \geq 0} X^t(U).$$

The result of [31] mentioned above admits a counterpart for flows in 3-dimensions: robust attractors have to be hyperbolic, *if they contain only regular orbits of the flow*. See [**?**], [23]. On the other hand, the geometric Lorenz attractors are striking examples that robust attractors of flows may contain singularities, together with regular orbits. A theory of such *singular attractors* has been under development in recent years, see [37], [38]. In particular, cf. this last paper, robust singular attractors of 3-dimensional flows are *singular hyperbolic*: partially hyperbolic with volume expansion along one of the subbundles and volume contraction along the other.

References

[1] K. Adl-Zarabi. Absolutely continuous invariant measures for piecewise expanding C^2 transformations in $\mathbb{R}^n$ on domains with cusps on their boundaries. *Ergod. Th. & Dynam. Sys.*, 16:1–18, 1996.

[2] 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.

[3] A. Andronov and L. Pontryagin. Systèmes grossiers. *Dokl. Akad. Nauk. USSR*, 14:247–251, 1937.

[4] D. V. Anosov. Geodesic flows on closed Riemannian manifolds of negative curvature. *Proc. Steklov Math. Inst.*, 90:1–235, 1967.

[5] D. V. Anosov and Ya. G. Sinai. Certain smooth ergodic systems. *Russian Math. Surveys*, 22:103–167, 1967.

[6] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Annals of Math.*, 133:73–169, 1991.

[7] M. L. Blank. Stochastic properties of deterministic dynamical systems. *Sov. Sci. Rev. C Maths/Phys*, 6:243–271, 1987.

[8] C. Bonatti and L. J. Díaz. Nonhyperbolic transitive diffeomorphisms. *Annals of Math.*, 143:357–396, 1996.

[9] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[10] R. Bowen. A horseshoe with positive measure. *Invent. Math.*, 29:203–204, 1975.

[11] R. Bowen. Invariant measures for Markov maps of the interval. *Comm. Math. Phys.*, 69:1–17, 1979.

[12] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math.*, 29:181–202, 1975.

[13] M. Brin and Ya. Pesin. Partially hyperbolic dynamical systems. *Izv. Acad. Nauk. SSSR*, 1:177–212, 1974.

[14] W. de Melo. Structural stability of diffeomorphisms on two-manifolds. *Invent. Math.*, 21:233–246, 1973.

[15] L. J. Díaz, E. Pujals, and R. Ures. Partial hyperbolicity and robust transitivity. *Acta Math.*, 183:1–43, 1999.

[16] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, 1992.

[17] J. Franks. Anosov diffeomorphisms on tori. *Trans. Amer. Math. Soc.*, 145:117–124, 1969.

[18] J. Franks. Anosov diffeomorphisms. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*, pages 61–93. Amer. Math. Soc., 1970.

[19] E. Giusti. *Minimal surfaces and functions of bounded variation.* Birkhäuser, 1984.

[20] P. Góra and A. Boyarsky. Absolutely continuous invariant measures for piecewise expanding C^2 transformations in $\mathbb{R}^n$. *Israel Jour. Math.*, 67:272–286, 1989.

[21] M. Gromov. Groups of polynomial growth and expanding maps. *Publ. Math. IHES*, 53:53–73, 1981.

[22] J. Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Publ. Math. IHES*, 50:59–72, 1979.

[23] 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.

[24] M. Hirsch and C. Pugh. Stable manifolds and hyperbolic sets. In *Global analysis*, volume XIV of *Proc. Sympos. Pure Math. (Berkeley 1968)*, pages 133–163. Amer. Math. Soc., 1970.

[25] M. Hirsch, C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lect. Notes in Math.* Springer Verlag, 1977.

[26] I. Kan. Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin. *Bull. Amer. Math. Soc.*, 31:68–74, 1994.

[27] G. Keller. *Proprietés ergodiques des endomorphismes dilatants C^2 par morceaux des régions bornés du plan.* PhD thesis, Univ. Rennes, 1979.

[28] A. Kosjakin and E. Sandler. Ergodic properties of a certain class of piecewise smooth transformations of a segment. *Izv. Vyss. Ucebn. Zaved. Matematika*, 118:32–40, 1972.

[29] A. Lasota and J.A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[30] T.-Y. Li and J. A. Yorke. Ergodic transformations from an interval to itself. *Trans. Amer. Math. Soc.*, 235:183–192, 1978.

[31] R. Mañé. Contributions to the stability conjecture. *Topology*, 17:383–396, 1978.

[32] R. Mañé. An ergodic closing lemma. *Annals of Math.*, 116:503–540, 1982.

[33] R. Mañé. *Ergodic theory and differentiable dynamics.* Springer Verlag, 1987.

[34] R. Mañé. A proof of the C^1 stability conjecture. *Publ. Math. I.H.E.S.*, 66:161–210, 1988.

[35] A. Manning. There are no new Anosov diffeomorphisms on tori. *Amer. J. Math.*, 96:422–429, 1974.

[36] L. Mora and M. Viana. Abundance of strange attractors. *Acta Math.*, 171:1–71, 1993.

[37] 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.

[38] 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.

[39] S. Newhouse. On codimension one Anosov diffeomorphisms. *Amer. J. Math.*, 92:761–770, 1970.

[40] S. Newhouse. Hyperbolic limit sets. *Trans. A.M.S.*, 167:125–150, 1972.

[41] S. Newhouse. Lectures on Dynamical Systems. In *Dynamical Systems*, Progress in Mathematics (CIME Lectures 1978), pages 1–114. Birkhäuser, 1978.

[42] 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.

[43] J. Palis. On the C^1 Ω-stability conjecture. *Publ. Math. I.H.E.S.*, 66:211–215, 1988.

[44] 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.

[45] J. Palis and F. Takens. *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations.* Cambridge University Press, 1993.

[46] Ya. Pesin and Ya. Sinai. Gibbs measures for partially hyperbolic attractors. *Ergod. Th. & Dynam. Sys.*, 2:417–438, 1982.

[47] V. Pliss. On a conjecture due to Smale. *Diff. Uravnenija*, 8:262–268, 1972.

[48] H. Proppe, P. Góra, and A. Boyarsky. Inadequacy of the bounded variation technique in the ergodic theory of higher dimensional transformations. *Nonlinearity*, 3:1081–1087, 1990.

[49] J. Robbin. A structural stability theorem. *Annals of Math.*, 94:447–493, 1971.

[50] C. Robinson. C^r structural stability implies Kupka-Smale. In *Dynamical Systems (Proc. Sympos. Univ. Bahia, Salvador, 1971)*, pages 443–449. Academic Press, 1971.

[51] C. Robinson. Structural stability of vector fields. *Annals of Math.*, 99:154–175, 1974. Errata in Annals of Math. 101:368, 1975.

[52] C. Robinson. Structural stability of C^1 vector fields. In *Dynamical Systems - Warwick 1974*, volume 468, pages 262–277. Springer Verlag, 1975.

[53] C. Robinson. Structural stability of C^1 diffeomorphisms. *J. Diff. Equ.*, 22:28–73, 1976.

[54] A. Sannami. The stability theorems for discrete dynamical systems on two-dimensional manifolds. *Nagoya Math. Journal*, 90:1–55, 1983.

[55] M. Shub. Endomorphisms of compact differentiable manifolds. *Amer. Journal of Math.*, 91:129–155, 1969.

[56] M. Shub. *Global stability of dynamical systems.* Springer Verlag, 1987.

[57] Ya. Sinai. Gibbs measures in ergodic theory. *Russian Math. Surveys*, 27:21–69, 1972.

[58] S. Smale. Differentiable dynamical systems. *Bull. Am. Math. Soc.*, 73:747–817, 1967.

[59] M. Tsujii. Physical measures for partially hyperbolic endomorphisms. Preprint 2002.

[60] S. Wong. Some metric properties of piecewise monotonic mappings of the unit interval. *Trans. Amer. Math. Soc.*, 246:493–500, 1978.

Marcelo Viana (viana@impa.br)
IMPA, Est. D. Castorina 110
22460-320 Rio de Janeiro, Brazil