RANDOM PERTURBATIONS AND STATISTICAL PROPERTIES OF HÉNON-LIKE MAPS

MICHAEL BENEDICKS AND MARCELO VIANA

Abstract. For a large class of non-uniformly hyperbolic diffeomorphisms, we prove stochastic stability under small random noise: the unique stationary probability measure of the Markov chain converges to the Sinai-Ruelle-Bowen measure of the unperturbed attractor when the noise level tends to zero.

Introduction

A basic problem in Dynamics is that of the stability of the dynamical behaviour under perturbations of the system. In simple terms, one wants to decide whether the system's evolution in the long run depends in a sensitive way upon details of the evolution law or, on the contrary, remains roughly the same when the evolution law is slightly modified. In the latter – stable – case, the mathematical formulation of the system stands a good chance of yielding accurate conclusions, even if it does not correspond exactly to the physical phenomenon it is meant to model (as it really never does).

Structural stability was proposed by Andronov-Pontryagin [2] in the thirties. It means that the whole orbit structure remains the same, up to a global continuous change of coordinates, for every C^1-nearby system (diffeomorphism or flow). It was conjectured by Palis-Smale [23] that a system is structurally stable if and only if it is uniformly hyperbolic (Axiom A plus strong transversality condition).

This conjecture was established in the mid-eighties by Mañé [20], for diffeomorphisms, and about a decade later by Hayashi [15], for flows. They showed that stable systems are hyperbolic, the converse having been proved by Robbin [27] and Robinson [28] in the seventies. The versions of Mañé's and Hayashi's theorems for the C^k topology, that were also conjectured in [23], remain outstanding challenges for all $k > 1$. See [24] for a detailed account of this and related subjects.

Date: July 26, 2002.

This work was partially supported by a STINT grant. M.B. is partially supported by the NFR and the Göran Gustafsson Foundation. M.V. is partially supported by CNPq and Faperj.

A different notion of stability, of a statistical nature, was introduced by Kolmogorov and Sinai [31]. A precise definition of this *stochastic stability* will appear in the next section. For now, let us say that a system is stochastically stable if time-averages of continuous functions, the most basic statistical data, are only slightly affected when the evolution is perturbed by small random noise.

This notion has attracted renewed interest in recent years, in the wake of much on-going progress in understanding non-uniformly hyperbolic dynamical systems. Indeed, many such systems should be stable in this stochastic sense, although so far this has been proved only in a few cases.

In the uniformly hyperbolic context the theory was carried out by Kifer [18, 19], who proved that uniformly expanding maps and Axiom A attractors, as well as geometric Lorenz attractors of flows, are stochastically stable. See also Young [34] for the case of Axiom A diffeomorphisms.

There has also been substantial progress for one-dimensional maps, including the family of quadratic real transformations $x \mapsto a - x^2$. Katok-Kifer [16] proved stochastic stability in the Misiurewicz case, i.e. when the critical point is non-recurrent. Benedicks-Young [9] and Baladi-Viana [6] extended this conclusion for sets of values of the parameter a with positive Lebesgue measure. See also the unpublished work of Collet [12]. Recently, Avila-Moreira [5] announced that quadratic maps are stochastically stable for Lebesgue almost every parameter value. In fact, these results hold for generic families of unimodal maps of the interval.

For all these and other very interesting recent developments, including Alves, Araújo [1, 3] and Metzger [21], it is fair to say that the theory of stochastic stability remains very much incomplete. In particular, little is known about higher-dimensional systems, outside the uniformly hyperbolic domain.

In the present work we prove that *Hénon-like attractors are stochastically stable under small random perturbations.* The precise statement will appear as Theorem A in Section 1.5, but we take the remainder of this Introduction for a brief explanation.

Hénon-like attractors are modeled on the Hénon family of maps

$$f_{a,b} : (x, y) \mapsto (1 + y - ax^2, bx) \qquad \begin{matrix} 0 < b < b_0 \\ 0 < a < 2. \end{matrix} \tag{1}$$

In [7], Benedicks and Carleson proved that there is a set of positive Lebesgue measure E in the parameter space such that for $(a, b) \in E$, the map $f = f_{a,b}$ has a strange attractor. More precisely, there is an attractor $\Lambda = \operatorname{clos}(W^u(P))$, where P is the fixed point in the first quadrant, containing a point z_0 with a dense orbit and such that

$$\left\|Df^j(z_0)(0,1)\right\| \geq e^{cj} \qquad \text{for all } j \geq 0.$$

Based on [7], Mora-Viana [22] and Díaz-Rocha-Viana [14] proved that attractors combining hyperbolic and "folding" behaviour occur persistently in very general bifurcations mechanisms, such as homoclinic tangencies and saddle-node cycles.

Then it was proved by Benedicks-Young [10] that all these *Hénon-like attractors* support a unique Sinai-Ruelle-Bowen (SRB) measure, that is, an invariant probability measure μ such that

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z)) \to \int \varphi\, d\mu \tag{2}$$

for every continuous function φ, and every z in some positive volume subset B of the ambient manifold. In addition, the system (f, μ) has exponential decay of correlations in the space of Hölder continuous functions [11]. Finally, in [8] we proved that these attractors have the *no-holes property:* almost every point z in the basin of the attractor Λ satisfies property (2).

By random perturbations of f we mean that we consider pseudo-trajectories $\{z_j\}_{j=0}^{\infty}$ where z_0 is an arbitrary point, and each z_j is a random variable in the ε-neighbourhood of $f(z_{j-1})$. Our conditions include the case when the probability law of x_j is uniformly distributed in that ε-neighbourhood, but they hold in more generality, as we shall see in Section 1.5.

We prove that, for every small $\varepsilon > 0$, there exists a unique probability measure μ_ε such that

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(z_j) \to \int \varphi\, d\mu_\varepsilon \tag{3}$$

for almost every choice of $\{z_j\}$. Moreover, μ_ε converges to the SRB measure μ in the weak*-topology, when ε tends to zero. This last fact encompasses the stochastic stability of these maps.

This work is organized as follows. In Section 1 we introduce the main notations, and give the full statements of our results. In Sections 2 and 3 we recall some known material that we are to use in the sequel. Section 4 contains a modified construction of the SRB measure of the unperturbed map: the main point is that we prove that Cesaro limits of iterates of arc-length on unstable manifolds exist, and they give the SRB measure in a more explicit way, which is crucial for our purposes. In Section 5 we adapt the symbolic dynamics construction in Section 3 to the stochastic setting. These are the two main preparatory ingredients for the proof of Theorem A, that we give in Section 6.

1. Notations and statement of results

Let us make precise what we mean by random perturbations of a map. For detailed accounts of random iterations see the books of Kifer [18, 19] and Arnold [4].

We consider U an open subset of some manifold M, and $f : U \to U$ such that $f(U)$ is relatively compact in U. In what follows $\varepsilon > 0$ is always assumed to be smaller than $\operatorname{dist}(f(U), M \setminus U)$.

1.1. **Markov chains.** In heuristic terms, Markov chains model the random processes obtained by iterating each z_j under the original map f, and then making a small mistake. Formally, one is given a family $\{p_\varepsilon(\cdot \mid z) : z \in U,\ \varepsilon > 0\}$ of Borel probability measures in U, such that every $p_\varepsilon(\cdot \mid z)$ is supported inside the ε-neighbourhood of $f(z)$. This defines a Markov chain in U, with the $p_\varepsilon(\cdot \mid z)$ as its transition probabilities: the *random orbits* are the sequences $\{z_j\}$ where each z_j is a random variable with probability distribution $p_\varepsilon(\cdot \mid z_{j-1})$.

Associated to the Markov chain we have an operator $\mathcal{T}_\varepsilon$ acting on the space of Borel measures of U by

$$\mathcal{T}_\varepsilon m = \int p_\varepsilon(\cdot \mid z)\, dm(z).$$

A probability measure μ_ε is *stationary* if $\mathcal{T}_\varepsilon \mu_\varepsilon = \mu_\varepsilon$, that is,

$$\mu_\varepsilon(E) = \int p_\varepsilon(E \mid z)\, d\mu_\varepsilon(z) \tag{4}$$

for every Borel set $E \subset U$. This is equivalent to the probability measure $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$ defined on the cylinder sets of $U \times U^{\mathbb{N}}$ by

$$\mu_\varepsilon(A_0 \times \cdots \times A_m) = \int_{A_0} d\mu_\varepsilon(z_0) \int_{A_1} p_\varepsilon(dz_1 \mid z_0) \cdots \int_{A_m} p_\varepsilon(dz_m \mid z_{m-1})$$

being invariant under the shift map $\mathcal{F} : U \times U^{\mathbb{N}} \to U \times U^{\mathbb{N}}$, defined by $(z_0, \{z_i\}_{i=1}^\infty) \mapsto (z_1, \{z_i\}_{i=2}^\infty)$ Then the time-average

$$\tilde{\varphi}(\mathbf{z}) = \lim \frac{1}{n} \sum_{j=0}^{n-1} \varphi(z_j)$$

exists for every continuous $\varphi : U \to \mathbb{R}$ and a full $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-measure subset of random orbits $\mathbf{z} = \{z_j\}$.

1.2. **Existence and ergodicity of stationary measures.** Consider a family $\{p_\varepsilon(\cdot \mid z) : z \in U,\ \varepsilon > 0\}$ of transition probabilities as before.

Lemma 1.1. *Let $z \mapsto p_\varepsilon(\cdot \mid z)$ be weak*-continuous, for some $\varepsilon > 0$. Then, for any Borel probability measure m with support contained in U, every weak*-accumulation point μ_ε of the sequence $n^{-1}\sum_{j=0}^{n-1}\mathcal{T}_\varepsilon^j m$ is a stationary measure for the Markov chain $\{p_\varepsilon(\cdot \mid z) : z \in U\}$.*

Proof. Since the space of probability measures supported in the closure of U is weak*-compact, accumulation points do exist. The assumption on the transition probabilities ensures that the operator $\mathcal{T}_\varepsilon$ is weak*-continuous. It follows that every accumulation point is a fixed point for $\mathcal{T}_\varepsilon$, that is, a stationary measure. □

A function $\phi : U \to \mathbb{R}$ is called *invariant* if

$$\phi(x) = \int \phi(y)\, p_\varepsilon(dy \mid x) \quad \text{for } \mu_\varepsilon\text{-almost every } x.$$

A stationary measure μ_ε is *ergodic* if every invariant function is constant μ_ε-almost everywhere. Every stationary measure can be decomposed as a convex combination of ergodic ones; see e.g. [18, Proposition 2.1] or [3]. If μ_ε is ergodic then the time-average

$$\tilde{\varphi}(\mathbf{z}) = \int \varphi\, d\mu_\varepsilon \tag{5}$$

for every continuous function φ and $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-almost every $\mathbf{z} = \{z_j\}$. To see this, consider

$$\tilde{\varphi}_k(z_0, \ldots, z_k) = \int \tilde{\varphi}(\mathbf{z})\, p_\varepsilon(dz_{k+1} \mid z_k) \cdots p_\varepsilon(dz_{n+1} \mid z_n) \cdots$$

for each $k \geq 0$. Using the fact that $\tilde{\varphi} = \tilde{\varphi} \circ \mathcal{F}$, we easily get that $\tilde{\varphi}_0$ is an invariant function and so $\tilde{\varphi}_0$ is constant $\mu_\varepsilon \times p_\varepsilon^{\mathbb{N}}$-almost everywhere. Moreover, $\tilde{\varphi}_k = \tilde{\varphi}_{k-1} \circ \mathcal{F}$ for every $k \geq 1$, and so each $\tilde{\varphi}_k$ is constant almost everywhere. Using $\tilde{\varphi}(\mathbf{z}) = \lim \tilde{\varphi}_k(z_0, \ldots, z_k)$, we conclude that the same is true for $\tilde{\varphi}$, and that implies (5).

1.3. **Random maps.** In this paper we consider random orbits obtained by iteration $z_j = g_j \circ \cdots \circ g_1(z_0)$ of maps g_j chosen at random (independently) close to the original f. In precise terms, one is given a family $\{\nu_\varepsilon : \varepsilon > 0\}$ of probabilities in the space of maps, such that each $\Omega_\varepsilon = \operatorname{supp} \nu_\varepsilon$ is contained in the ε-neighbourhood of f (e.g. with respect to some C^k topology, $k \geq 0$). A basic tool is the skew product map

$$\mathcal{F}_\varepsilon : U \times \Omega_\varepsilon^{\mathbb{N}} \to U \times \Omega_\varepsilon^{\mathbb{N}}, \quad \mathcal{F}_\varepsilon(z, \mathbf{g}) \mapsto (g_1(z), \sigma(\mathbf{g}))$$

where $\mathbf{g} = (g_1, g_2, \dots)$ and $\sigma : \Omega_\varepsilon^{\mathbb{N}} \to \Omega_\varepsilon^{\mathbb{N}}$ is the shift map. The *random orbit* associated to a $(z, \mathbf{g})$ is the sequence $z_j = g_j \cdots g_1(z),\ \ j \geq 0$.

A probability measure μ_ε is *stationary* if

$$\mu_\varepsilon(E) = \int (g_* \mu_\varepsilon)(E)\, d\nu_\varepsilon(g) = \int \mu_\varepsilon(g^{-1}(E))\, d\nu_\varepsilon(g) \tag{6}$$

for every Borel set $E \subset U$. It is not difficult to see that μ_ε is stationary if and only if the measure $(\mu_\varepsilon \times \nu_\varepsilon^{\mathbb{N}})$ is invariant for $\mathcal{F}_\varepsilon$.

There is an associated Markov chain scheme, given by the transition probabilities

$$p_\varepsilon(E \mid z) = \nu_\varepsilon\big(\{g : g(z) \in E\}\big). \tag{7}$$

A probability measure μ_ε is stationary, in the sense of (6), if and only if it is stationary for the Markov chain defined by (7). Indeed, given any Borel set E and any Borel probability measure m in U,

$$\begin{aligned}
\int (g_* m)(E)\, d\nu_\varepsilon(g) &= \iint (\mathcal{X}_E \circ g)(z)\, d\nu_\varepsilon(g)\, dm(z) \\
&= \int \nu_\varepsilon\big(\{g : g(z) \in E\}\big)\, dm(z) \\
&= \int p_\varepsilon(E \mid z)\, dm(z) = (\mathcal{T}_\varepsilon m)(E)
\end{aligned}$$

(use Fubini's theorem 8.8 in [30]). This calculation also shows that

$$\mathcal{T}_\varepsilon m = \int (g_* m)\, d\nu_\varepsilon(g) = \pi_{1*} \mathcal{F}_{\varepsilon *}(m \times \nu_\varepsilon^{\mathbb{N}}),$$

where $\pi_1 : U \times \Omega_\varepsilon^{\mathbb{N}} \to U$ is the canonical projection $\pi_1(z, \mathbf{g}) = z$, and π_{1*} and $\mathcal{F}_{\varepsilon *}$ are the forward iterations induced by π_1 and $\mathcal{F}_\varepsilon$ in the space of Borel measures.

Moreover, given any Borel subsets $A_0, A_1, \dots, A_m$ of U,

$$\begin{aligned}
&\big(\mu_\varepsilon \times \nu_\varepsilon^{\mathbb{N}}\big)\left(\{(z, \mathbf{g}) : z \in A_0, g_1(z) \in A_1, \dots, g_m \cdots g_1(z) \in A_m\}\right) \\
&= \int_{A_0} d\mu_\varepsilon(z) \int \mathcal{X}_{\{g_1 : g_1(z) \in A_1\}} d\nu_\varepsilon(g_1) \cdots \int \mathcal{X}_{\{g_m : (g_m \cdots g_1)(z) \in A_m\}} d\nu_\varepsilon(g_m) \\
&= \int_{A_0} d\mu_\varepsilon(z) \int_{A_1} p_\varepsilon(dy_1 \mid z) \cdots \int_{A_m} p_\varepsilon(dy_m \mid y_{m-1})
\end{aligned}$$

This means that the statistics of the random orbits obtained from randomly perturbing the dynamical system are faithfully reproduced by the Markov chain.

Since we only deal with continuous maps, the probabilities $p_\varepsilon(\,\cdot \mid z)$ given by (7) vary continuously with z, and so Lemma 1.1 applies.

Remark 1.2. The Markov chains one obtains via this random perturbation scheme are special in that they exhibit *spatial correlation:* as one usually deals with fairly regular maps g, transition probabilities $p_\varepsilon(\cdot \mid z)$ and $p_\varepsilon(\cdot \mid z')$ given by (7) are strongly correlated if z and z' are close-by. See Section 1.6 below and [18, Section 1.1] for a discussion of relations between these two schemes.

1.4. **Stochastic stability.** Let us suppose $f : U \to U$ has a naturally defined invariant probability measure μ. The main case we have in mind is when μ is the unique SRB measure supported in an attractor Λ with the no-holes property, and U is contained in the basin of attraction of Λ. In that case,

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

for Lebesgue almost every $z \in U$, and every continuous $\varphi : U \to \mathbb{R}$.

Let a random perturbation scheme as in Sections 1.1 or 1.3 be given. Assume there is a unique stationary probability measure μ_ε, for every small $\varepsilon > 0$. Then, cf. previous section,

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi(z_j) \to \int \varphi\, d\mu_\varepsilon$$

for almost every random orbit $\{z_j\}$ and every continuous $\varphi : U \to \mathbb{R}$.

Definition 1.3. The system (f, μ) is *stochastically stable* with respect to $\{p_\varepsilon(\cdot \mid z) : z \in U,\ \varepsilon > 0\}$ (or with respect to $\{\nu_\varepsilon : \varepsilon > 0\}$) if $\mu_\varepsilon \to \mu$ when $\varepsilon \to 0$, in the weak*-sense: $\int \varphi\, d\mu_\varepsilon \to \int \varphi\, d\mu$ for every continuous $\varphi : U \to \mathbb{R}$.

We shall see in Section 6.1 that the stationary probability is unique in the situations we are interested in. In general, stationary measures form a simplex in the space of probabilities. The definition of stochastic stability extends naturally: the whole simplex should converge to μ when $\varepsilon \to 0$. We also observe that, in great generality, this simplex has finite dimension [3].

1.5. **Statement of the main result.** The aim of the present paper is to prove stochastic stability for Hénon-like attractors, under very general random perturbations.

Fix a bounded open neighbourhood U of the Hénon-like attractor Λ, contained in the basin of attraction $B(\Lambda)$ and such that $f(U)$ is relatively compact in U. Consider random perturbations $\{\nu_\varepsilon : \varepsilon > 0\}$ where each ν_ε is supported in the ε-neighbourhood of f relative to the

C^2 topology on the closure of U. Let $p_\varepsilon(\cdot \mid z)$ be the corresponding transition probabilities, given by (7).

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

(H1) every $\Lambda_{\varepsilon,z}$ admits a lamination into nearly horizontal (slope less than 10) curves such that the union of the laminae with length less than εt has $p_\varepsilon(\cdot \mid z)$-probability $\le K t^{1+\kappa}$, for every $t > 0$;

(H2) the conditional probability of $p_\varepsilon(\cdot \mid z)$ along each lamina $\mathbf{y}$ is absolutely continuous with respect to arc-length $m_{\mathbf{y}}$, with density $\psi_{\mathbf{y}} = \psi_{\varepsilon,z,\mathbf{y}}$ bounded by $K/\operatorname{length}(\mathbf{y})$;

(H3) restricted to some ball of radius $\rho(\varepsilon)$ around $f(z)$, the probability $p_\varepsilon(\cdot \mid z)$ is absolutely continuous with respect to Lebesgue measure, with positive density.

Figure 1 describes some domains satisfying the geometric condition (H1), when $p_\varepsilon(\cdot \mid z)$ is normalized area. In the last example it is assumed that the upper and lower cuspidal vertices have finite order contact; in the flat case (H1) may not hold. More examples are given in Figure 2.

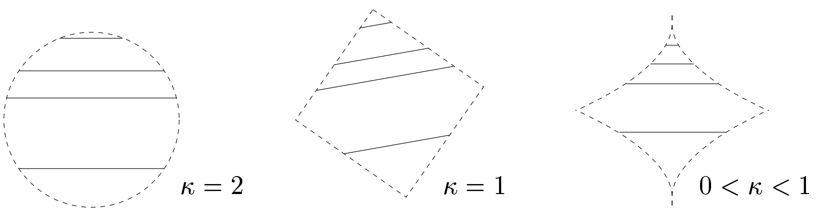

FIGURE 1.

Our main result is the following

Theorem A. *Let f be a Hénon-like map, μ be its SRB measure, and $\{\nu_\varepsilon : \varepsilon > 0\}$ fulfill conditions (H1), (H2), (H3).*

Then there is a unique stationary measure μ_ε supported in the basin $B(\Lambda)$, and this measure is ergodic. Moreover, μ_ε converges to μ in the weak-sense as $\varepsilon \to 0$.*

In the proof we use (H1)-(H2) only for points $z = (x, y)$ close to the y-axis $\{x = 0\}$. Condition (H3) is needed only for proving uniqueness, in Lemma 6.1.

Remark 1.4. The reason why we state our results for random map type perturbations is that the distortion arguments in Section 5 require Lipschitz, or at least Hölder, variation of the derivative. It is not clear how generally the conclusion will hold if one considers Markov chain perturbations, not necessarily arising from a random maps scheme. But the comments in the next section around Example 1.7 do provide an extension of Theorem A for random perturbations of Markov chain type, when the random noise satisfies a Lipschitz regularity condition.

An easier version of these methods may be applied to certain one-dimensional maps. In this way, we improve the results in [9]: since we require no lower bound on the density, we are able to treat random noise supported in general domains, not only intervals. In this regard, see the discussion around Figure 2 in the next section.

1.6. **Additional remarks.** In some cases, like in Figure 1, the space of laminae in (H1) may be parametrized by the vertical coordinate (hence our using the symbol $\mathbf{y}$ to represent laminae). However, our methods apply in more general situations, such as illustrated by Figure 2. In particular, the second and third examples in the figure emphasize the fact that $\Lambda_{\varepsilon,z}$ needs not be connected.

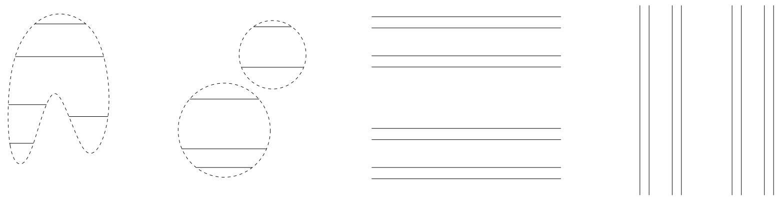

FIGURE 2.

In the fourth example we think of $p_\varepsilon(\cdot \mid z)$ as being, essentially, uniformly distributed on the product of a compact interval by a Cantor set with Lebesgue measure ε. Although the product set does not have the geometric property (H1), we can easily fit this situation into our hypotheses: it suffices to take $\Lambda_{\varepsilon,z}$ to be a rectangle containing the support.

Example 1.5. (Additive noise) Let Λ_ε be a neighbourhood of the origin in $\mathbb{R}^2$ contained in the ball of radius ε, and let θ_ε be a probability measure supported on Λ_ε. Let ν_ε be the measure induced by θ_ε in the space of C^2 diffeomorphisms via the map $t \mapsto f_t = f + t$. Here $\Lambda_{\varepsilon,z} = f(z) + \Lambda_\varepsilon$. Property (H1) translates immediately to a similar condition about Λ_ε. Conditions (H2)-(H3) hold, for instance, if θ_ε is absolutely continuous with respect to area, with density ψ_ε such that $\varepsilon^2\psi_\varepsilon$ is bounded from zero and infinity.

Example 1.6. (Noise in parameter space) Let $\phi : B \times \mathbb{R}^2 \to \mathbb{R}^2$ be a C^2 map, where $B \subset \mathbb{R}^2$ is the unit ball around the origin, such that $\phi(0, \cdot) = 0$ and $\omega \mapsto \phi(\omega, z)$ is a diffeomorphism for every z near the attractor Λ. Define $f_\omega(z) = f(z) + \phi(\omega, z)$. Let θ_ε be the normalized Lebesgue measure on the ε-ball around $\omega = 0$, and ν_ε be the probability measure induced by θ_ε in the space of diffeomorphisms via the map $\omega \mapsto f_\omega$. Then $\{\nu_\varepsilon : \varepsilon > 0\}$ satisfies the assumptions of Theorem A.

The problem of when a Markov chain can be realized by a random maps scheme is discussed by Kifer in [18, Section 1.1]. He proves that under a mild condition on the ambient space, and assuming that $x \mapsto p_\varepsilon(E \mid x)$ is measurable for every Borel set E, such a realization is possible in the space of measurable maps. When the transition probabilities have positive densities we can say more: the Markov chain is represented by a parametrized family of maps as regular as the densities themselves:

Example 1.7. Assume $p_\varepsilon(\,\cdot \mid z) = \rho^\varepsilon(\cdot, z)m$ where m denotes Lebesgue area and $\rho^\varepsilon(\cdot, z)$ is positive on its support. For simplicity we take the support to be $f(z) + [-\varepsilon, \varepsilon]^2$. For each z and $(\xi_1, \xi_2) \in [-\varepsilon, \varepsilon]^2$, define

$$\begin{aligned} \omega_1 &= p_\varepsilon(f(z) + [-\varepsilon, \xi_1] \times [-\varepsilon, \varepsilon] \mid z) \quad \text{and} \\ \omega_1\omega_2 &= p_\varepsilon(f(z) + [-\varepsilon, \xi_1] \times [-\varepsilon, \xi_2] \mid z). \end{aligned}$$

The map $\psi^\varepsilon(\cdot, z) : (\xi_1, \xi_2) \mapsto (\omega_1, \omega_2)$ is well-defined and a bijection onto $[0, 1]^2$. Let $\phi^\varepsilon(\cdot, z)$ be its inverse. The definition gives

$$m(\{\omega : \phi^\varepsilon(z, \omega) \in E\}) = p_\varepsilon(f(z) + E \mid z)$$

for every measurable set E (the definition of $\omega_1\omega_2$ gives this for a generating family of rectangles). This means that the parametrized family $f^\varepsilon_\omega(z) = f(z) + \phi^\varepsilon(\omega, z)$, endowed with Lebesgue measure on $[0, 1]^2$, realizes the Markov chain. If $(\xi, z) \mapsto \rho^\varepsilon(\xi, z)$ is C^r, for some $r \geq 0$, then $\psi^\varepsilon(z, \xi)$ is C^r in both variables, and each $\psi^\varepsilon(\cdot, z)$ is a homeo/diffeomorphism. Therefore, the parametrized family f^ε is C^r.

This type of construction shows that realizability is closely related to the regularity of the random noise. Unfortunately, one lacks good examples to decide whether stochastic stability is significantly affected by the smoothness category. However, it is clear that, even in the simplest situations, one cannot expect stability to hold for *every* Markov chain. The following example of Keller [17] was brought to our attention by Gary Froyland:

Example 1.8. Let $f : S^1 \to S^1$ be given by $f(z) = 2z \bmod \mathbb{Z}$. Lebesgue measure on $S^1 = \mathbb{R}/\mathbb{Z}$ is the unique SRB measure of f. Consider the

Markov chain defined by $p_\varepsilon(\,\cdot\mid z)$ = normalized Lebesgue measure on

$$\begin{cases} (-\varepsilon,\varepsilon) & \text{if } z\in(-\varepsilon,\varepsilon)\\ (f(z)-\varepsilon, f(z)+\varepsilon) & \text{if } z\notin(-\varepsilon,\varepsilon)\end{cases}$$

The Markov chain has a unique stationary measure, μ_ε = normalized Lebesgue measure on $(-\varepsilon,\varepsilon)$, but μ_ε does not converge to the SRB measure when $\varepsilon\to 0$.

2. Hénon-like maps

We begin by recalling certain known facts about Hénon-like attractors, from [7, 10, 22], that are needed for the sequel. This section is mostly a summary of [8, Section 2].

2.1. **Hénon-like families.** We consider parameterized families of diffeomorphisms of the plane

$$f=f_a:(x,y)\mapsto(1-ax^2,0)+R(a,x,y), \tag{8}$$

R close to zero in the C^3 norm, which we call *Hénon-like families*. More precisely, we suppose that $\|R\|_{C^3}\le J\sqrt{b}$ on $[1,2]\times[-2,2]^2$, with

$$J^{-1}b\le|\det Df|\le Jb \quad\text{and}\quad \|D\big(\log|\det Df|\big)\|_\infty\le J, \tag{9}$$

where $J>0$ is arbitrary and $b>0$ is taken sufficiently small. The Hénon model (1) is affinely conjugate to $(x,y)\mapsto(1-ax^2+\sqrt{b}y,\sqrt{b}x)$, and so fits into this framework if b is small.

Consider parameters $a\in[a_1,a_2]$ with $1\gg\delta\gg 2-a_1>2-a_2\gg b$. The parameter interval should not be too small: $(a_2-a_1)\ge(2-a_2)/10$ suffices. In this parameter range, f has a unique fixed saddle-point P such that $\Lambda=\operatorname{clos}(W^u(P))$ is compact, indeed $\Lambda\subset(-2,2)^2$. The basin of attraction $B(\Lambda)$ contains a neighbourhood of Λ, in all the cases we are dealing with. See [8, Section 5].

The present setting may be extended considerably, along lines that are now well-understood. Indeed, the properties of quadratic family that are used in this context (specially: non-flat critical point, expanding behaviour outside the critical region, and variation of the kneading invariant) are true in great generality for families of one-dimensional maps with negative Schwarzian derivative, cf. [13, 14, 32].

Thus, we can replace the quadratic maps $1-ax^2$ in the definition (8) by very general families of uni- or multi-modal maps of the circle or the interval, as described in [14, Section 5]. See also [33]. Moreover, $a=2$ may be replaced by any parameter such that all critical points are non-recurrent.

2.2. **The strange attractor.** Besides J, let $\sqrt{e} < \sigma_1 < \sigma_2 < 2$ be fixed at the very beginning. For the next theorem, one also fixes constants $1 \gg \beta \gg \alpha > 0$, and supposes $b \ll \delta \ll \alpha$. Throughout, we use $C > 1$ to represent various large constants depending only on J, σ_1, σ_2, α, or β (not on δ or b). Analogously, $c \in (0,1)$ is a generic notation for small constants depending only on J, σ_1, σ_2, α, or β.

Let $I(\delta) = \{(x,y) : |x| < \delta\}$. For $z \in W^u(P)$, let $t(z)$ be any norm 1 vector tangent to $W^u(P)$ at z (the particular choice is irrelevant). Given a non-zero vector $v = (v_1, v_2) \in \mathbb{R}^2$, $\operatorname{slope} v$ will always be taken with absolute values, i.e. $\operatorname{slope} v = |v_2|/|v_1|$.

Theorem 2.1. [7, 22] *Given any Hénon-like family, there exists a positive Lebesgue measure set E such that for every $a \in E$ the map f has a countable* critical set $\mathcal{C} \subset W^u(P) \cap I(\delta)$ *whose elements ζ satisfy*

1. *$t(\zeta)$ is almost horizontal and $t(f(\zeta))$ is almost vertical, in the sense that* $\operatorname{slope} t(\zeta) \le C\sqrt{b}$ *and* $\operatorname{slope} t(f(\zeta)) \ge c/\sqrt{b}$*;*
2. *$t(f(\zeta))$ is exponentially contracted and $w_0 = (1,0)$ is exponentially expanded under positive iterates: $\|Df^n(f(\zeta))t(f(\zeta))\| \le (C\,b)^n$ and $\|Df^n(f(\zeta))w_0\| \ge \sigma_1^n$ for all $n \ge 1$;*
3. *if $f^n(\zeta) \in I(\delta)$ then there is $\zeta_n \in \mathcal{C}$ so that* $\operatorname{dist}(f^n(\zeta), \zeta_n) \ge e^{-\alpha n}$ *and there is a C^2 curve $L = \{(x, y(x))\}$ with $|y'(x)| \le 1/10$ and $|y''(x)| \le 1/10$, tangent to $t(\zeta_n)$ at ζ_n and also containing $f^n(\zeta)$.*

In addition, there exists $\zeta \in \mathcal{C}$ such that $\{f^n(\zeta) : n \ge 0\}$ is dense in Λ.

From now on we always suppose $a \in E$. The statements that follow are part of the proof of this theorem.

2.3. **Segments of unstable manifold around critical points.**

Proposition 2.2.
1. *There exists $\zeta_0 = (x_0, y_0) \in \mathcal{C}$ with $|x_0| \le C\sqrt{b}$, so that $\mathcal{C} \cap G_0 = \{\zeta_0\}$, where G_0 denotes the segment connecting $f(\zeta_0)$ to $f^2(\zeta_0)$ in $W^u(P)$;*
2. *denoting $G_g = f^g(G_0) \setminus f^{g-1}(G_0)$, then $\mathcal{C} \cap G_g$ is finite for every $g \ge 1$, and in fact $\mathcal{C} \cap G_1$ consists of a single point ζ_1;*
3. *for every $\zeta \in \mathcal{C} \cap G_g$ and $g \ge 0$, the segment $\gamma = \gamma(\zeta)$ of radius δc^g around ζ in $W^u(P)$ may be written $\gamma = \{(x, y(x))\}$ with*
$$|y'(x)| \le C\sqrt{b} \quad \textit{and} \quad |y''(x)| \le C\sqrt{b};$$
4. *given any $\zeta \in \mathcal{C} \cap G_g$ with $g > 0$, there exist $\tilde{g} < g$ and $\tilde{\zeta} \in \mathcal{C} \cap G_{\tilde{g}}$ with* $\operatorname{dist}(\zeta, \tilde{\zeta}) \le b^{g/10}$.

The lower bound on the length of the segments $\gamma(\zeta)$ is important, so that we give a special name ρ to the constant c in the context of part 3 of the proposition. Moreover, we write K for the large constant C, and call a $C^2(b)$ *curve* any curve $\{(x, y(x))\}$ with $|y'(x)| \le K\sqrt{b}$ and $|y''(x)| \le K\sqrt{b}$.

Note that the expanding eigenvalue of $Df(P)$ is negative and so G_0 is a neighbourhood of P and ζ_0 in $W^u(P)$. It is easy to see that G_0 and G_1 contain $C^2(b)$ curves extending from $x = -9/10$ to $x = 9/10$. For $g \ge 0$, points in G_g are said to be of *generation* g.

2.4. **Contracting directions.** Since every orbit in $B(\Lambda)$ must eventually enter the square $[-2, 2]^2$, we may always assume to be dealing with orbits which never leave this square in positive time, and we do so. Given $\lambda > 0$, a point $z = (x, y)$ is called λ-*expanding* if

$$\|Df^j(z)w_0\| \ge \lambda^j \quad \text{for all } j \ge 1. \tag{10}$$

An important case is $z \in f(\mathcal{C})$, with $\lambda = \sigma_1$, cf. Theorem 2.1.2. We say that z is λ-expanding up to time n if the inequality in (10) holds for $1 \le j \le n$. We define the *contracting direction of order* $n \ge 1$ *at* z as the tangent direction $e^{(n)}(z)$ that is most contracted by $Df^n(z)$.

The next proposition summarizes a few results from [7, Section 5] and [22, Section 6]. In the statement $\lambda > 0$ and $\tau > 0$ are arbitrary constants, with τ sufficiently small (e.g. $\tau \le 10^{-20}$), and one assumes that b is much smaller than either of them.

Proposition 2.3. *Let z be λ-expanding up to time $n \ge 1$, and ξ satisfy* $\operatorname{dist}(f^j(\xi), f^j(z)) \le \tau^j$ *for every $0 \le j \le n-1$. Then, for any point η in the τ^n-neighbourhood of ξ and for every $1 \le l \le k \le n$,*

1. *$e^{(k)}(\eta)$ is uniquely defined and* $\operatorname{slope}(e^{(k)}(\eta)) \ge c/\sqrt{b}$;
2. $\operatorname{angle}(e^{(l)}(\eta), e^{(k)}(\eta)) \le (Cb)^l$ *and* $\|Df^l(\eta)e^{(k)}(\eta)\| \le (Cb)^l$;
3. $\|De^{(k)}(\eta)\| \le C\sqrt{b}$ *and* $\|D^2e^{(k)}(\eta)\| \le C\sqrt{b}$;
4. $\|D(Df^l e^{(k)})(\eta)\| \le (Cb)^l$;
5. $1/10 \le \|Df^n(\xi)w_0\|/\|Df^n(z)w_0\| \le 10$ *and* $\operatorname{angle}(Df^n(\xi)w_0, Df^n(x)w_0) \le (\sqrt{C\tau})^n$.

Parts 3 and 4 are also true for the derivatives of $e^{(k)}$ and $Df^l e^{(k)}$ with respect to the parameter a. Throughout, we write *expanding* to mean λ-expanding for some $\lambda \ge e^{-20}$.

2.5. **Long stable leaves.**

Proposition 2.4. *If z is an expanding point then its stable set $W^s(z)$ contains a segment $\Gamma = \Gamma(z) = \{(x(y), y) : |y| \le 1/10\}$ with $|x'| \le C\sqrt{b}$ and $|x''| \le C\sqrt{b}$, such that $z \in \Gamma$ and*

$$\operatorname{dist}(f^n(\xi), f^n(\eta)) \le (Cb)^n \operatorname{dist}(\xi, \eta), \quad \textit{for all } \xi, \eta \in \Gamma \textit{ and } n \ge 1.$$

Moreover, if z_1, z_2 are expanding points then

$$\operatorname{angle}(t_\Gamma(\xi_1), t_\Gamma(\xi_2)) \le C\sqrt{b}\operatorname{dist}(\xi_1,\xi_2)$$

for every $\xi_1 \in \Gamma(z_1)$ and $\xi_2 \in \Gamma(z_2)$, where $t_\Gamma(\xi_i)$ denotes any norm 1 vector tangent to $\Gamma(z_i)$ at ξ_i.

We call a *long stable leaf* any curve Γ as in this proposition, and a *stable leaf* any compact curve having some iterate contained in a long stable leaf. The first part of the proposition is proved in [7, Section 5.3], the arguments extending directly to Hénon-like maps [22, Section 7C]. The second part is an easy consequence of the construction, as explained in [8, Section 2].

2.6. **Hyperbolic behaviour away from the critical region.**

Proposition 2.5. *Given any $k \ge 1$, any $z \in [-2,2]^2$ with $f^j(z) \notin I(\delta)$ for $0 \le j < k$, and any vector v with $\|v\| = 1$ and* slope $v \le 1/5$*, then*

$$\operatorname{slope}(Df^j(z)\,v) \le (C/\delta)\sqrt{b} < 1/10 \quad \textit{and} \quad \|Df^j(z)v\| \ge c\delta\sigma_2^j$$

for $1 \le j \le k$. If either $z \in f(I(2\delta))$ or $f^k(z) \in I(2\delta)$ then

$$\|Df^k(z)\,v\| \ge \sigma_2^k,$$

and in the latter case we also have $\operatorname{slope}(Df^k(z)\,v) \le C\sqrt{b}$.

This means, in particular, that pieces of orbits outside $I(\delta)$ are (essentially) expanding. Similar statements are well-known for one-dimensional maps such as $x \mapsto 1 - ax^2$. The proposition follows using a perturbation argument, see [7, Lemmas 4.5, 4.6].

2.7. **Bound periods for critical points.** Another important notion is that of *bound period* $p(n,\zeta)$ associated to a *return* n of a critical point $\zeta \in \mathcal{C}$. These are defined through the following inductive procedure.

If $n \ge 1$ does not belong to $[\nu+1, \nu+p(\nu,\zeta)]$ for any return $1 \le \nu < n$, then n is a (*free*) return for ζ if and only if $f^n(\zeta) \in I(\delta)$. Moreover, the bound period $p = p(n,\zeta)$ is the largest integer such that

$$\operatorname{dist}(f^{n+j}(\zeta), f^j(\zeta_n)) \le e^{-\beta j} \quad \text{for all } 1 \le j \le p, \tag{11}$$

where ζ_n is the *binding point* of $f^n(\zeta)$, given by Theorem 2.1.3. If, on the contrary, n is in $[\nu+1, \nu+p(\nu,\zeta)]$ for some previous return $1 \le \nu < n$ then, by definition, n is a (*bound*) return for ζ if and only if $n - \nu$ is a return for the binding point ζ_ν, and we let $p(n,\zeta) = p(n-\nu, \zeta_\nu)$.

We may suppose that bound periods are nested, in the sense that if $n \in [\nu+1, \nu+p(\nu,\zeta)]$ then $n + p(n,\zeta) \le \nu + p(\nu,\zeta)$, that is to say, the bound period associated to n ends before the one associated to ν.

We write $d_n(\zeta) = \operatorname{dist}(f^n(\zeta), \zeta_n)$, for ζ and ζ_n as before. Moreover, $w_j(z) = Df^j(f(z))w_0$ for any point z and $j \geq 0$.

Proposition 2.6. *Let $n \geq 1$ be a free return of $\zeta \in \mathcal{C}$, and $p = p(n, \zeta)$ be the corresponding bound period. Then*

1. $(1/5)\log(1/d_n(\zeta)) \leq p \leq 5\log(1/d_n(\zeta))$;
2. $\|w_{n+p}(\zeta)\| \geq \sigma_1^{(p+1)/3}\|w_{n-1}(\zeta)\|$ *and* $\operatorname{slope} w_{n+p}(\zeta) \leq (C/\delta)\sqrt{b}$;
3. $\|w_{n+p}(\zeta)\| d_n(\zeta) \geq ce^{-\beta(p+1)}\|w_{n-1}(\zeta)\|$;
4. $\|w_j(f^n(\zeta))\| \geq \sigma_1^j$ *for* $1 \leq j \leq p$, *and* $\operatorname{slope} w_p(f^n(\zeta)) \leq (C/\delta)\sqrt{b}$.

A main ingredient here is the property in Theorem 2.1.3. We shall comment a bit more on it in a while. Actually, for free returns n, a curve L as in the theorem may be taken tangent not only to $t(\zeta_n)$ at ζ_n but also to $w_{n-1}(\zeta)$ at $f^n(\zeta)$, see [7, Section 7.3] and [22, Lemma 9.5].

2.8. **Dynamics on the unstable manifold.** The next proposition, appearing in [10], permits to extend to generic orbits in $W^u(P)$ the control given by the previous statements for orbits of critical points. This is a key step in the construction of the SRB measure of f on Λ that appeared in that paper, cf. Theorem 2.9 below.

Proposition 2.7. *Let $\tilde{z} \in W^u(P)$ be such that $f^n(\tilde{z}) \notin \mathcal{C}$ for every $n \geq 1$. Then, given any $n \geq 1$ such that $f^n(\tilde{z}) \in I(\delta)$, there exists $\zeta_n \in \mathcal{C}$ and some C^2 curve $L = \{(x, y(x))\}$ with $|y'| \leq 1/10$ and $|y''| \leq 1/10$, tangent to $t(\zeta_n)$ at ζ_n and also containing $f^n(\tilde{z})$.*

Given a point $z \in W^u(P)$, fix $k \gg 1$ so that $\tilde{z} = f^{-k}(z)$ belongs to a small neighbourhood of P in $W^u(P)$. We can now define *returns*, *binding points*, and *bound periods* for $\tilde{z}$ in the same way as we did before for critical points. That is, corresponding to a free return n of $\tilde{z}$ we choose as binding point a critical point ζ_n as in the proposition, and define the bound period $p = p(n, \tilde{z})$ of $f^n(\tilde{z})$ with respect to this ζ_n, cf. (11). As in the case of critical points, we take the bound periods nested.

We say that $z = f^k(\tilde{z})$ is a *free point* if k is outside every bound period $[\nu + 1, \nu + p(\nu, z_0)]$ of $\tilde{z}$. This is an intrinsic property of the point z: the choice of k is irrelevant, as long as it is large enough. We call a segment $\gamma \subset W^u(P)$ *free* if all its points are free. While proving Proposition 2.7, it is shown in [10] that if n is a free return for $\tilde{z}$ and $\gamma \subset W^u(P)$ is a free segment containing $f^n(\tilde{z})$, then the same binding point may be assigned to all the points in $\gamma \cap I(\delta)$. More precisely, *there is a critical point ζ_γ and a curve L as in the statement, tangent* *to* $t(\zeta_\gamma)$ *at* ζ_γ *and containing the whole* γ. In particular, L is tangent to $t(w)$ at every $w \in \gamma$. In some cases $\zeta_\gamma \in \gamma = L$, but it is not always possible to take $L \subset W^u(P)$.

Given any maximal free segment γ intersecting $I(\delta)$, we always fix L and ζ_γ as above, and set $d_{\mathcal{C}}(w) = \operatorname{dist}(w, \zeta_\gamma)$ for each $w \in L$. We extend $t(w)$ to represent a norm 1 vector tangent to the curve L at every $w \in L$, and define the bound period $p(w)$ of every $w \in L$ with respect to this ζ_γ, cf. (11).

2.9. **Bound periods following tangential returns.** The following definition is a slight extension of notions with similar denominations appearing in [7, 10, 11, 22]. Given points p, q and tangent vectors u, v, we say that p is in *tangential position* relative to (q, v) if there exists a curve $\{(x, y(x))\}$ with $|y'| \le 1/5$ and $|y''| \le 1/5$, tangent to v at q and also containing p. And we say that (p, u) is in tangential position relative to (q, v) if such a curve may be chosen tangent to u at p.

Thus, as we have seen, if z is a free point contained in the $W^u(P)$ then $(z, t(z))$ is in tangential position with respect to $(\zeta_\gamma, t(\zeta_\gamma))$ for some critical point ζ_γ. It is worth stressing that there can be no analog of this for points outside the unstable manifold. But in [8] we proved that, for points in the basin, returns are almost surely eventually tangential.

The importance of the tangential position property comes from the consequence that the diffeomorphism f behaves, essentially, as a one-dimensional quadratic map over the curve L. This is at the basis of the proof of the next result, which is similar to that of Proposition 2.6. See [7, Section 7.4] and [22, Section 10].

Proposition 2.8. *Given any curve L as before and $z \in L$,*

1. $(1/5)\log(1/d_{\mathcal{C}}(z)) \le p(z) \le 5\log(1/d_{\mathcal{C}}(z))$;
2. $\|Df^{p(z)+1}(z)t(z)\| \ge \sigma_1^{(p(z)+1)/3}$ *and* $\operatorname{slope}(Df^{p(z)+1}(z)t(z)) < (C/\delta)\sqrt{b}$;
3. $\|Df^{p(z)+1}(z)t(z)\| d_{\mathcal{C}}(z) \ge c e^{-\beta(p(z)+1)}$;
4. $\|w_j(z)\| \ge \sigma_1^j$ *for* $1 \le j \le p(z)$, *and* $\operatorname{slope} w_{p(z)}(z) < (C/\delta)\sqrt{b}$.

As noted in [8, Section 2], parts 2-4 of Proposition 2.8 remain true if one replaces $t(z)$ by any norm 1 tangent vector v such that (z, v) is in tangential position relative to $(\zeta_\gamma, t(\zeta_\gamma))$. Moreover, the arguments also allow for some freedom in the definition of bound period. For instance, suppose $z(s) \in L$ is such that (compare (11))

$$\operatorname{dist}(f^j(z(s)), f^j(\zeta_\gamma)) \begin{cases} \le 10e^{-\beta j} & \text{for } 1 \le j \le p(z) \\ \ge \frac{1}{10}\, e^{-\beta j} & \text{for } j = p(z)+1. \end{cases} \tag{12}$$

For example, this will always be the case if $z(s)$ is close enough to z. Then the same arguments as in the proof of Proposition 2.8 apply, giving conclusions 2-4 of the proposition with $z(s)$ in the place of z, and $p(z)$ unchanged. This means that one might just as well take $p(z(s)) = p(z)$ for any such s. This flexibility of the definition was used before in [8, 11].

2.10. **SRB measure and the no-holes property.** We also quote the main results of [10] and [8]:

Theorem 2.9. [10] *There exists a unique f-invariant measure μ supported in Λ, having nonzero Lyapunov exponents almost everywhere, and whose conditional measures along unstable manifolds are absolutely continuous with respect to Lebesgue measure on these manifolds. In particular, μ is an SRB measure for f.*

In addition, the support of μ coincides with Λ, and the system (f, μ) is Bernoulli.

Given any segment $\gamma \subset W^u(P)$, almost every point in γ, with respect to arc-length, satisfies (2) for every continuous φ. See [10, Section 3] and [8, Section 2].

Theorem 2.10. [8] *Through Lebesgue almost every point z in the basin of attraction $B(\Lambda)$ passes a stable leaf $W^s(\xi)$ of some point $\xi \in \Lambda$: in fact,* $\operatorname{dist}(f^n(z), f^n(\xi)) \to 0$ *exponentially fast as $n \to +\infty$.*

Moreover, for Lebesgue almost every $z \in B(\Lambda)$ and every continuous function $\varphi : \mathbb{R}^2 \to \mathbb{R}$,

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

3. Itineraries in the basin of attraction

One of the new features on this paper, as with [8], is that we have to deal with orbits that are not in tangential positions relative to the critical points. In [8] this was handled through a special sequence of dynamically defined partitions $\mathcal{P}_n$ of the basin into *rectangles*, that is, regions bounded by two segments of $W^u(P)$ and two stable leaves. Let us recall that construction here, as in Section 5 we shall need to modify it to fit the present stochastic context.

3.1. **A special family of long stable leaves.** The next proposition is the basic result allowing us to construct these partitions. It encompasses Proposition 3.3 and Remark 3.3 of [8], and some of the comments following them. The symbol $\approx$ means equality up to a factor 100. It is no restriction to take $\Delta = \log(1/\delta)$ to be a large integer, and we do so.

Proposition 3.1. *Given $\zeta \in \mathcal{C}$, let $\Gamma^s(\zeta) = \{(x^s(y), y) : |y| \leq 1/10\}$ be the long stable leaf through the critical value $f(\zeta)$.*

1. *There exist long stable leaves $\Gamma_r(\zeta) = \{(x_r(y), y) : |y| \leq 1/10\}$, for $r \geq \Delta$, accumulating Γ^s exponentially fast from the left:*
$$x^s(y) - x_r(y) \approx e^{-2r} \quad \text{for every } r \geq \Delta \text{ and } |y| \leq 1/10.$$
Moreover, the leaves $\Gamma_\Delta = \Gamma_\Delta(\zeta)$ and $\Gamma_{\Delta+1} = \Gamma_{\Delta+1}(\zeta)$ may be taken the same for all critical values.
2. *There exist long stable leaves $\Gamma_{r,l}(\zeta) = \{(x_{r,l}(y), y) : |y| \leq 1/10\}$, for $0 \leq l \leq r^2$ and $r \geq \Delta$, with $\Gamma_{r,0} = \Gamma_{r-1}$ and $\Gamma_{r,r^2} = \Gamma_r$, and such that*
$$x^s(y) - x_{r,l}(y) \approx e^{-2r} \quad \text{and} \quad x_{r,l}(y) - x_{r,l-1}(y) \approx \frac{e^{-2r}}{r^2}$$
for every $r \geq \Delta$, every $1 \leq l \leq r^2$, and every $|y| \leq 1/10$.
3. *Every $\Gamma_{r,l}(\zeta)$ intersects the unstable manifold $W^u(P)$ at some point $\eta_{r,l}$. For every free return $n \geq 1$, $d_{\mathcal{C}}(f^{n-1}(\eta_{r,l})) \geq e^{-2\beta n}$ and $(f^{n-1}(\eta_{r,l}), Df^{n-1}(\eta_{r,l})w_0)$ is in tangential position relative to $(\zeta_n, t(\zeta_n))$.*

We are going to define *itinerary* of a point z in the basin of attraction. The definition involves choosing a sequence of critical points $\tilde{\zeta}_j$ close to each iterate $f^{n_j}(z)$ that is near $x = 0$, and describing the position of $f^{n_j}(z)$ relative to $\tilde{\zeta}_j$ in terms of the long stable leaves $\Gamma_{r,l}$ in Proposition 3.1. By definition, the atoms of each partition $\mathcal{P}_n$ are the sets of points sharing the same itinerary up to time n.

More precisely, to each point $z \in B(\Lambda)$ we are going to associate sequences n_j, $i_j = (\tilde{\zeta}_j, r_j, l_j, \epsilon_j)$, $j \geq 0$, where n_j is an integer, $\tilde{\zeta}_j \in \mathcal{C}$, $\epsilon_j \in \{+, 0, -\}$, and r_j and l_j are also integers with either $(r_j, l_j) = (0, 0)$ or $r_j \geq \Delta$ and $1 \leq l_j \leq r_j^2$. Roughly speaking, n_j is the j:th free return of z, $\tilde{\zeta}_j$ is the corresponding binding point, and r_j, l_j, ϵ_j describe the position of $f^{n_j+1}(z)$ relative to the long stable leaves $\Gamma_{r,l}(\tilde{\zeta}_j)$. The precise construction follows.

3.2. **Preliminaries.** Recall that G_0, G_1 contain long $C^2(b)$ segments γ_0, γ_1, around the critical points ζ_0, ζ_1, respectively. In view of the form of our map, for each $i = 0, 1$ we may write $f(\gamma_i)$ as $\{\xi_i(x), \eta_i(x)\}$ with $\xi_i'' \approx -2a \approx 4$ and $|\eta_i|, |\eta_i'|, |\eta_i''| \leq C\sqrt{b}$. In particular, $f(\gamma_i)$ intersects each $\Gamma_{r,l}(\zeta_i)$, for $0 \leq l \leq r^2$, in exactly two points.

Let Δ_i be the region bounded by $f(\gamma_i)$ and by the long stable leaf $W^s_{loc}(P)$ passing through P, see Figure 3. Since $f(\gamma_0)$ and $f(\gamma_1)$ are disjoint, whereas Δ_0 and Δ_1 must intersect each other (e.g. extend $\{\gamma_0, \gamma_1\}$ to a foliation by nearly horizontal curves, and use that the image of each leaf intersects every vertical line in not more than two points), we have either $\Delta_1 \subset \Delta_0$ or $\Delta_0 \subset \Delta_1$.

We consider $\Delta_1 \subset \Delta_0$, as the other case is analogous. In the sequel we define $n_j(z), i_j(z)$, $j \geq 0$, for points $z \in \Delta_0$. The extension to generic points $w \in B(\Lambda)$ is, simply, by taking $n_j(w) = n + n_j(f^n(w))$ and $i_j(w) = i_j(f^n(w))$ for each $j \geq 0$, where $n \geq 0$ is the smallest integer for which $f^n(w) \in \Delta_0$. Lebesgue almost every point in the basin of Λ has some iterate contained in Δ_0, cf. [8, Section 5]. Hence, this leaves out only a zero Lebesgue measure subset of $B(\Lambda)$, which is negligible for our purposes.

Before proceeding, let us make a few simple conventions. In what follows (r, l) should be replaced by $(r-1, (r-1)^2 + l)$ if $l \leq 0$, and by $(r+1, l-r^2)$ if $l > r^2$. We say that $(r_1, l_1) > (r_2, l_2)$ if either $r_1 > r_2$ or $r_1 = r_2$ and $l_1 > l_2$. The *region in between* two long stable leaves is open on the left and closed on the right: if $\Gamma_1 = \{(x_1(y), y) : |y| \leq 1/10\}$ and $\Gamma_2 = \{(x_2(y), y) : |y| \leq 1/10\}$, with $x_1 < x_2$, then the region in between Γ_1 and Γ_2 is $\{(x, y) : x_1(y) < x \leq x_2(y),\ |y| \leq 1/10\}$.

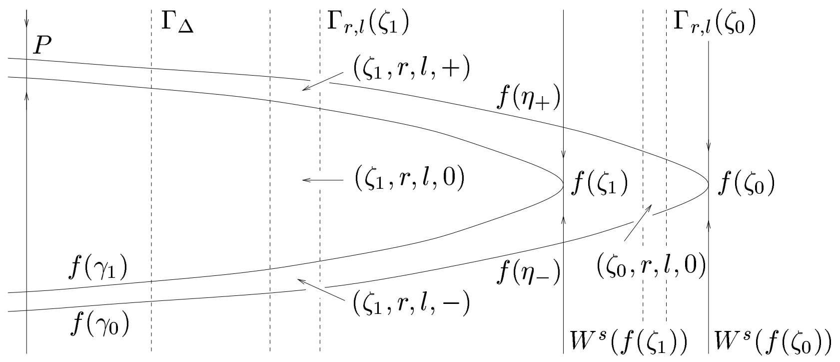

FIGURE 3.

3.3. **Itineraries: Step zero.** Let $(\hat{r}, \hat{l})$ be defined by the condition that $f(\zeta_1)$ is in the region of Δ_0 in between $\Gamma_{\hat{r},\hat{l}}(\zeta_0)$ and $\Gamma_{\hat{r},\hat{l}-1}(\zeta_0)$. For $z \in \Delta_0$ we define $n_0 = -1$ and

(a) $i_0(z) = (\zeta_0, r, l, 0)$ if z is in the region of Δ_0 in between $\Gamma_{r,l}(\zeta_0)$ and $\Gamma_{r,l-1}(\zeta_0)$, with $(r, l) > (\hat{r}, \hat{l})$;

(b) $i_0(z) = (\zeta_0, \hat{r}, \hat{l}, 0)$ if z is in the region of Δ_0 between $W^s_{loc}(f(\zeta_1))$ and $\Gamma_{\hat{r},\hat{l}}(\zeta_0)$;

(c) $i_0(z) = (\zeta_1, r, l, \pm)$ if z is in either of the two regions of $\Delta_0 \setminus \Delta_1$ in between $\Gamma_{r,l}(\zeta_1)$ and $\Gamma_{r,l-1}(\zeta_1)$, the sign $+/-$ corresponding to the upper/lower region;

(d) $i_0(z) = (\zeta_1, 0, 0, \pm)$ if z is in either of the two regions of $\Delta_0 \setminus \Delta_1$ in between Γ_Δ and $W^s_{loc}(P)$, the sign $+/-$ corresponding to the upper/lower region;
(e) $i_0(z) = (\zeta_1, r, l, 0)$ if z is in the region of Δ_1 in between $\Gamma_{r,l}(\zeta_1)$ and $\Gamma_{r,l-1}(\zeta_1)$.
(f) $i_0(z) = (\zeta_1, 0, 0, 0)$ if z is in the region of Δ_1 in between Γ_Δ and $W^s_{loc}(P)$.

We also define $R(i_0) = \{z \in \Delta_0 : i_0(z) = i_0\}$ for each $i_0 = (\tilde{\zeta}_0, r_0, l_0, \epsilon_0)$ as before. This closes the zeroth step of our definition.

3.4. **Itineraries: Step 1.** Now we explain how $n_1(z)$ and $i_1(z)$ are defined for z in $R(i_0)$, for each fixed i_0.

In cases (a), (b), (c), (e), define $p_1 = p_1(i_0) \geq 1$ to be the largest integer such that

$$\operatorname{dist}(f^j(z), f^j(\tilde{\zeta}_0)) \leq e^{-\beta j}$$

for $1 \leq j \leq p_1$ and every $z \in f^{-1}(R(i_0))$. For (d), (f) just set $p_1 = 0$. In any case, let $m_1 = n_1 > p_1$ be minimum such that $f^{n_1}(R(i_0))$ intersects $I(\delta)$. Denote γ_i^u, $i = 0, 1$, and γ_j^s, $j = 0, 1$, the four segments forming the boundary of the rectangle $R(i_0)$, with the γ_i^u contained in $W^u(P)$ and the γ_j^s contained in long stable leaves. Moreover, let $z^*_{i,j} = \gamma_i^u \cap \gamma_j^s$ be the corner points of $R(i_0)$, for $i = 0, 1$ and $j = 0, 1$.

Proposition 3.2. [8, Proposition 3.5]

1. $m_1 > p_1 \geq (4/3) r_0$;
2. *for* $i = 0, 1$, *the slope of* $f^{n_1}(\gamma_i^u)$ *is less than* $(C/\delta)\sqrt{b}$ *at every point;*
3. $\operatorname{length}(f^{n_1}(\gamma_j^s)) \leq (1/10)\, d_{\mathcal{C}}(z^*_{i,j})$ *for* $i = 0, 1$ *and* $j = 0, 1$;
4. $\operatorname{angle}(t(z^*_{0,j}), t(z^*_{1,j})) \leq (1/10)\, d_{\mathcal{C}}(z^*_{i,j})$ *for* $i = 0, 1$ *and* $j = 0, 1$.

Moreover, conclusions 2-4 of Proposition 2.8 are true at time p_1 *for any point in either of the unstable boundary segments.*

The last statement means that we may take the bound period constant equal to p_1 on the whole $f^{-1}(R(i_0))$. Recall the comments following Proposition 2.8. In particular, both segments $f^{n_1}(\gamma_i^u)$, $i = 0, 1$, are free. According to Proposition 2.7, each of these segments may be extended to a C^2 curve $K_i = \{(x, y_i(x)\}$ with $|y_i'|, |y_i''| \leq 1/10$ and tangent to $W^u(P)$ at some critical point $\eta_i \in K_i$. By definition, $d_{\mathcal{C}}(z^*_{i,j}) = \operatorname{dist}(z^*_{i,j}, \eta_i)$ for every $j = 1, 0$.

Recall that η_i may not belong to $f^{n_1}(\gamma_i^u)$. We can also not discard the possibility that $\eta_0 = \eta_1$. On the other hand, according to the next lemma, either both η_i belong to the corresponding $f^{n_1}(\gamma_i^u)$ or none does, and in the latter case we may always take the two critical points to coincide.

Lemma 3.3. [8, Lemma 3.6] *If $\eta_0 \in f^{n_1}(\gamma_0^u)$ then $\eta_1 \in f^{n_1}(\gamma_1^u)$. In the opposite case, $f^{n_1}(\gamma_1^u)$ is in tangential position relative to $(\eta_0, t(\eta_0))$: there is a C^2 curve $K_2 = \{(x, y_2(x)\}$ with $|y_2'|, |y_2''| \leq 1/5$, containing $f^{n_1}(\gamma_1^u)$ and tangent to $W^u(P)$ at η_0.*

We define $i_1(z)$ first when $\eta_i \in f^{n_1}(\gamma_i^u)$ for $i = 0, 1$. Up to interchanging subscripts, we may suppose that $f(\eta_0)$ is to the right of $f(\eta_1)$, meaning that its long stable leaf is to the right of the one passing through $f(\eta_1)$. Then $f(\eta_1)$ is contained in a region bounded by $f^{n_1+1}(\gamma_0^u)$ and some pair of long leaves $\Gamma_{\hat{r},\hat{l}-1}(\eta_0)$ and $\Gamma_{\hat{r},\hat{l}}(\eta_0)$. We let, see Figure 4,

(a1) $i_1(z) = (\eta_0, r, l, 0)$ if $f^{n_1+1}(z)$ is in the region of $f^{n_1+1}(R(i_0))$ in between $\Gamma_{r,l}(\eta_0)$ and $\Gamma_{r,l-1}(\eta_0)$, with $(r, l) > (\hat{r}, \hat{l})$;

(b1) $i_1(z) = (\eta_0, \hat{r}, \hat{l}, 0)$ if $f^{n_1+1}(z)$ is in the region of $f^{n_1+1}(R(i_0))$ in between $W^s_{loc}(f(\eta_1))$ and $\Gamma_{\hat{r},\hat{l}}(\zeta_0)$;

(c1) $i_1(z) = (\eta_1, r, l, \pm)$ if $f^{n_1+1}(z)$ is in either of the connected components of $f^{n_1+1}(R(i_0))$ in between $\Gamma_{r,l}(\eta_1)$ and $\Gamma_{r,l-1}(\eta_1)$, the sign $+/-$ corresponding to the upper/lower region.

(d1) $i_1(z) = (\eta_1, 0, 0, \pm)$ if $f^{n_1+1}(z)$ is in either of the connected components of $f^{n_1+1}(R(i_0))$ to the left of Γ_Δ, the sign $+/-$ corresponding to the upper/lower component.

The definition of $i_1(z)$ is slightly simpler in the case $\eta_i \notin f^{n_1}(\gamma_i^u)$ for $i = 0, 1$. Taking advantage of the fact that both segments $f^{n_1}(\gamma_i^u)$, $i = 0, 1$, are in tangential position relative to η_0, cf. Lemma 3.3, we define

(a2) $i_1(z) = (\eta_0, r, l, +)$ if $f^{n_1+1}(z)$ is in the region of $f^{n_1+1}(R(i_0))$ in between $\Gamma_{r,l}(\eta_0)$ and $\Gamma_{r,l-1}(\eta_0)$;

(b2) $i_1(z) = (\eta_0, 0, 0, +)$ if $f^{n_1+1}(z)$ is in the region of $f^{n_1+1}(R(i_0))$ to the left of Γ_Δ.

See Figure 4. Our choice $\epsilon_j = +$ is purely conventional: the intersection of $f^{n_1+1}(R(i_0))$ with any region in between two stable leaves is connected, and so ϵ_j has no role in this case.

This completes the definition of $i_1(z)$. For each $i_0 = (\tilde{\zeta}_0, r_0, l_0, \epsilon_0)$ and $i_1 = (\tilde{\zeta}_1, r_1, l_1, \epsilon_1)$ we set $R(i_0, i_1) = \{z \in R(i_0) : i_1(z) = i_1\}$.

3.5. **Itineraries: Conclusion.** The definition of $i_k(z)$ for a general $k \geq 1$ is very similar to the case $k = 1$. Suppose $i_j(z)$, $n_j(z)$, and $R(i_0, \ldots, i_j)$ have been defined for every $j < k$. Let $i_j = (\tilde{\zeta}_j, r_j, l_j, \epsilon_j)$, $j = 0, \ldots, k-1$, be fixed, and $z \in R(i_0, \ldots, i_{k-1})$. In cases (a1), (b1), (c1), (a2), we define $p_k = p_k(i_0, \ldots, i_{k-1}) \geq 1$ to be the largest integer

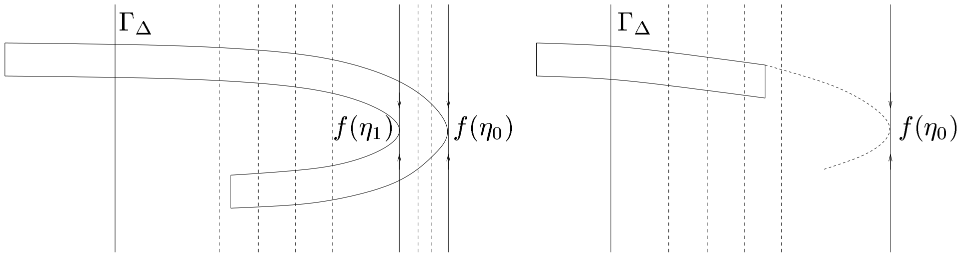

FIGURE 4.

such that

$$\operatorname{dist}(f^j(\zeta), f^j(\tilde{\zeta}_{k-1})) \leq e^{-\beta j}$$

for $1 \leq j \leq p_k$ and every $\zeta \in f^{n_{k-1}}(R(i_0, \ldots, i_{k-1}))$. For (d1), (b2) we just set $p_k = 0$. Then we let n_k be the smallest integer larger than $n_{k-1} + p_k$ such that $f^{n_k}(R(i_0, \ldots, i_{k-1}))$ intersects $I(\delta)$, and let $m_k = n_k - (n_{k-1} + 1)$. Call γ_i^u, γ_j^s the boundary segments, and $z_{i,j}^*$ the corner points of $f^{n_{k-1}+1}(R(i_0, \ldots, i_{k-1}))$, with the same conventions as before. Then,

Proposition 3.4. [8, Proposition 3.7]

1. $m_k > p_k \geq (4/3) r_{k-1}$;
2. *for* $i = 0, 1$ *the slope of* $f^{m_k}(\gamma_i^u)$ *is less than* $(C/\delta)\sqrt{b}$ *at every point;*
3. $\operatorname{length} f^{m_k}(\gamma_j^s) \leq (1/10)\, d_{\mathcal{C}}(z_{i,j}^*)$ *for* $i = 0, 1$ *and* $j = 0, 1$;
4. $\operatorname{angle}(t(z_{0,j}^*), t(z_{1,j}^*)) \leq (1/10)\, d_{\mathcal{C}}(z_{i,j}^*)$ *for* $i = 0, 1$ *and* $j = 0, 1$.

Moreover, conclusions 2-4 of Proposition 2.8 are true at time p_k *for any point in either of the unstable boundary segments.*

Thus, the bound period may be taken constant equal to p_k on the whole $f^{n_{k-1}}(R(i_0, \ldots, i_{k-1}))$. Then both $f^{m_k}(\gamma_i^u)$, $i = 0, 1$, are free segments, and the Proposition 2.7 gives us the analog of Lemma 3.3 at every return:

Lemma 3.5. [8, Lemma 3.8] *Either there are two critical points* η_0, η_1 *such that* $\eta_i \in f^{m_k}(\gamma_i^u)$ *for* $i = 0$ *and* $i = 1$, *or there is a critical point* η_0 *such that both segments* $f^{m_k}(\gamma_i^u)$, $i = 0, 1$, *are in tangential position relative to* $(\eta_0, t(\eta_0))$.

In the first case we define $\hat{r}, \hat{l}$ just as before. Then we let $i_k(z)$ be given by the rules which are obtained replacing $f^{n_1+1}(z)$ by $f^{n_k+1}(z)$, and $f^{n_1+1}(R(i_0))$ by $f^{n_k+1}(R(i_0, \ldots, i_{k-1}))$ in (a1)–(d1). In the second case we define $i_k(z)$ by the rules obtained by making the corresponding substitutions in (a2)–(b2). Finally, for each $i_0, \ldots, i_{k-1}, i_k$, we let

$$R(i_0, \ldots, i_{k-1}, i_k) = \{z \in R(i_0, \ldots, i_{k-1}) : i_k(z) = i_k\}.$$

Our definition of itinerary of a point z in the basin of Λ is complete. By construction, every $R(i_0, \ldots, i_k)$ is a rectangle. Note that the two segments of unstable manifold on its boundary are also contained in the boundary of $R(i_0, \ldots, i_{k-1})$. In the sequel, we call *unstable sides* of a rectangle the segments of unstable manifold on its boundary, and *unstable boundary* the union of the unstable sides. Stable sides and stable boundary are defined analogously.

3.6. **Abundance of long stable leaves.** The following result was proved in [8]. A related construction appeared in [11].

Proposition 3.6. *There exists a family $\mathcal{H}$ of long stable leaves and $\varepsilon_0 > 0$ such that itineraries are constant on each leaf $\Gamma \in \mathcal{H}$ and the set $H = \cup_{\Gamma \in \mathcal{H}} \Gamma$ has positive area. Moreover, H intersects every nearly horizontal C^1 curve $\gamma = \{(x, y(x))\}$, $|y'| \leq 1/5$, connecting Γ_Δ to $\Gamma_{\Delta+1}$ on a subset with arc-length measure larger than ε_0.*

It suffices to take $\mathcal{H}$ to be the family of long stable leaves forming the set $H = H(i_0)$ of [8, Section 4], for any symbol $i_0 = (\tilde{\zeta}_0, r_0, l_0, \epsilon_0)$ with $r_0 = \Delta$. The first statement in Proposition 3.6 is contained in the definition of $H(i_0)$. That H has positive area is proved in [8, Proposition 4.10]. The final statement is a consequence. Indeed, the second part of Proposition 2.4 states that the lamination $\mathcal{H}$ is Lipschitz continuous, with small Lipschitz constant $C\sqrt{b}$. It follows, by the Gronwall inequality, that the projection $\pi_{\mathcal{H}} : \gamma_1 \to \gamma_2$ along the leaves of $\mathcal{H}$ is Lipschitz continuous with Lipschitz constant smaller than 2, for any two nearly horizontal curves γ_1 and γ_2. Thus,

$$m_{\gamma_2}\big(\pi_{\mathcal{H}}(E)\big) \leq 2m_{\gamma_1}\big(E\big) \tag{13}$$

for any measurable set $E \subset \gamma_1 \cap H$. In particular, $m_\gamma(\gamma \cap H)$ is positive and bounded from zero, for every nearly horizontal curve γ, as claimed.

4. SRB measure via return maps

In this section we give an alternative construction of the SRB measure for Hénon-like attractors. It is based on constructing a kind of return map to some subset of the attractor, with good expansion, distortion, and Markov properties. A fairly standard argument shows that this return map has an SRB measure, from which we obtain the SRB measure of the original diffeomorphism f.

This modification of the original construction in [10] provides an explicit expression for the SRB measure of f, which turns out to be very important for our proof of stochastic stability.

4.1. **Itineraries and escape situations on $W^u(P)$.** Itineraries for points in the unstable manifold were implicit in [7], and an explicit construction first appeared in [10]. Here it is convenient to adopt the following definition, inherited from our construction in the basin of attraction. We use the setting and notations of Section 3.

Let Ω be the unstable side of Δ_0 contained in $f(\gamma_0)$, that is, such that P is one of its end-points. See Figure 3. For each $k \geq 0$, there is a partition $\mathcal{W}_k$ of Ω such that each $\omega_k \in \mathcal{W}_k$ is an unstable side of some rectangle $R(i_0, \ldots, i_k)$. For every $z \in \omega_k$ define $n_k(z) = n_k$ (this is determined by i_{k-1}) and $i_k(z) = i_k = (\tilde{\zeta}_k, r_k, l_k, \epsilon_k)$.

Let $\omega_{k-1} \in \mathcal{W}_{k-1}$, with itinerary i_0, i_1, $\ldots$, i_{k-1}, and let n_k be the corresponding k:th free return. We say that n_k is an *escape situation* if $f^{n_k+1}(\omega_{k-1})$ crosses the region $\mathcal{R}_\Delta$, extending to distance $\geq \delta/10$ to either side of $\mathcal{R}_\Delta$. This notion was introduced in [7], in a slightly different form.

Then let $\omega \subset \omega_{k-1}$ be the sub-segment which is mapped inside $\mathcal{R}_\Delta$ by f^{n_k+1}. We say that the points of ω escape at time n_k, and we call $f^{n_k+1}(\omega)$ an *escaping leaf*. Note that ω is the union of Δ^2 elements of the partition $\mathcal{W}_k$, corresponding to $r_k = \Delta$, and $f^{n_k+1}(\omega)$ is a nearly horizontal curve stretching across $\mathcal{R}_\Delta$: its end-points are contained in Γ_Δ and $\Gamma_{\Delta+1}$, respectively.

Given $w \in \Omega$ we define $e_0(w)$ to be the smallest $n_k \geq 0$ such that w escapes at time n_k. If n_k does not exist then, by convention, $e_0(w) = \infty$. Similarly, given $z \in f^{n_k+1}(\omega)$, let $w = f^{-n_k-1}(z)$ and $n_l = n_l(w)$ be the next escape situation of w, that is, the smallest $n_l > n_k$ such that w escapes at time n_l. We call $e(z) = n_l - n_k$ the *escaping time* of z. If n_l does not exist, the escaping time is infinite. The next lemma says that this is rather unlikely, in terms of the arc-length measure m_γ on each leaf γ:

Lemma 4.1. *There exist $C > 0$ and $c > 0$ such that, for every $n \geq 1$,*

$$m_\Omega\big(\{w \in \Omega : e_0(w) > n\}\big) \leq Ce^{-cn} \quad \textit{and}$$

$$m_\gamma\big(\{z \in \gamma : e(z) > n\}\big) \leq Ce^{-cn} \quad \textit{for every escaping leaf } \gamma.$$

This follows from the large deviations argument in [7, Section 2.2]. The definition of escape situation ensures that non-escaping segments (the connected components of the $f^{n_k+1}(\omega_{k-1}) \setminus \mathcal{R}_\Delta$) are never too small, which is important for this argument (alternatively, we could take Γ_Δ and $\Gamma_{\Delta+1}$ contained in the stable manifold of the fixed point, so that their iterates never return to $\mathcal{R}_\Delta$). See also [10, Section 3.3], where a very similar statement is used.

4.2. **Long unstable leaves.** Let $X_0 \subset W^u(P)$ be the union of all escaping leaves, over all $k \geq 1$, and X be the closure of X_0.

Lemma 4.2. *Each point z of X is located on a C^1 nearly horizontal curve $\gamma_z = \{(x, y(x))\}$, $|y'(x)| \leq 1/10$, stretching across $\mathcal{R}_\Delta$.*

Proof. Let $\{z_j\}$ be a sequence in X_0 converging to z, and $\{\gamma_j\}$ be the escaping curves with $z_j \in \gamma_j$. From the definition we have that every $f^{-1}(\gamma_j)$ is a free curve. In particular,

$$f^{-1}(\gamma_j) = \{(x, y_j(x))\} \quad \text{with } |y_j'| \leq 1/10 \text{ and } |y_j''| \leq 1/10$$

for all j. So, using Arzela-Ascoli, we can pick a subsequence of $\{\gamma_j\}$ that converges, in the C^1 topology, to some curve γ_z as in the statement. $\square$

We call *long unstable leaves* all the C^1 curves γ_z as in the lemma, including the escaping leaves as a particular case. We shall see in Lemma 4.6 that long unstable leaves are exponentially contracted by all backward iterates of f, with uniform contraction rates.

By construction, escaping leaves are two-by-two disjoint. In particular, for every pair of long unstable leaves $\gamma_1 = \{(x, y_1(x))\}$ and $\gamma_2 = \{(x, y_2(x))\}$, either $y_1 \leq y_2$ or $y_2 \leq y_1$. This provides a natural total order relation in the space of long unstable leaves. It is not clear whether general long unstable leaves are also pairwise disjoint [1]. But in this direction we can prove (see Figure 5):

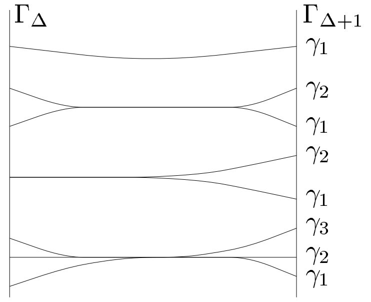

FIGURE 5.

[1] This would follow, via the Gronwall inequality, if we knew that tangent directions to escaping curves satisfy a uniform Lipschitz condition $|y_1' - y_2'| \leq C|y_1 - y_2|$. Of course, Lipschitz continuity of the unstable foliation is an interesting problem in itself. A particular case was done in [8, Section 4.1].

Lemma 4.3. 1. *If two long unstable leaves γ_1 and γ_2 intersect each other, then $\gamma_1 \cap \gamma_2$ is connected. Moreover, the two leaves are tangent at every point in the intersection.*
2. *If $\gamma_1 < \gamma_2 < \gamma_3$ are long unstable leaves such that γ_1 intersects γ_2 and γ_2 intersects γ_3, then γ_2 is an escaping leaf, whereas γ_1 and γ_3 are not.*
3. *Each connected component of X consists of not more than 3 long unstable leaves. Moreover, there are at most countably many connected components containing more than 1 leaf.*

Proof. Suppose the intersection of γ_1 and γ_2 was not connected. Then there would be some domain D bounded by pieces of the two curves. By the expansion property in Lemma 4.6 below, the lengths of the backward iterates of long unstable leaves decrease with time, even exponentially fast. Consequently, the area of $f^{-n}(D)$ would converge to zero as $n \to \infty$, contradicting the fact that f^{-1} is area-expanding. This proves the first part of claim 1.

The second part of 1, as well as claim 2, are simple consequences of the fact that escaping leaves are pairwise disjoint, and every long unstable leaf is a C^1 limit of escaping leaves: if two long unstable leaves intersect each other, there can be at most one escaping leaf between them (and, in that case, neither of the first two is an escaping leaf). Moreover, the first part of 3 is a direct consequence of claim 2: note that if a long unstable leaf is disjoint from all the others, then it coincides with its connected component.

Let us prove the last part of claim 3. By claim 1 the end-points of distinct long unstable leaves can not coincide. So, to each connected component containing two leaves $\gamma_1 < \gamma_2$ we may associate an open segment in either Γ_Δ or $\Gamma_{\Delta+1}$, bounded by end-points of γ_1 and γ_2. Clearly, different connected components are assigned disjoint intervals. Hence, there can be at most countably many of these components. $\square$

To bypass possible intersections between long unstable leaves, we introduce the extension $\tilde{X}$ of X obtained by "doubling" (or "tripling") points in the intersection of two (or three) leaves. Formally,

$$\tilde{X} = \{(z, \gamma) : z \in \gamma \text{ and } \gamma \text{ is a long unstable leaf}\}.$$

By Lemma 4.3, the canonical projection $p : \tilde{X} \to X$, $p(x, \gamma) = z$, is at most 3-to-1. We shall identify each $\{(z, \gamma) : z \in \gamma\}$ with the corresponding leaf γ.

4.3. **Itineraries and escape situations on $\tilde{X}$.** Next we need to extend the definitions of itinerary $i = \{i_k\}$ and escaping time $e(\,\cdot\,)$ to (almost all) points in the set $\tilde{X}$. This goes as follows.

Let z be a point in some long unstable leaf γ that is not an escaping leaf. We consider arbitrary sequences $\{z_j\}$ in X_0 converging to z and such that the escaping leaves $\gamma_j \ni z_j$ converge to γ in the C^1 topology. Each z_j may be written as $z_j = f^{n_{l(j)}+1}(w_j)$ for some $w_j \in \Omega$ that escapes at time $n_{l(j)} = n_{l(j)}(w_j)$. We observe that the forward itinerary of z_j converges to a limit when $j \to \infty$. By this we mean that there exist sequences $\{\nu_k\}$ and $\{\iota_k\}$ such that, for each $k \geq 1$,

$$n_{l(j)+k}(w_j) - n_{l(j)}(w_j) = \nu_k \quad \text{and} \quad i_{l(j)+k}(w_j) = \iota_k$$

for every z_j in some neighbourhood of z. Then we define the k:th free return and the k:th symbol of z by

$$n_k(z, \gamma) = \nu_k \quad \text{and} \quad i_k(z, \gamma) = \iota_k. \tag{14}$$

Moreover, the escaping time $e(z, \gamma)$ is the smallest integer for which a sequence $\{(z_j, \gamma_j)\} \to (z, \gamma)$ may be found with $e(z_j) = e(z, \gamma)$ for all large j. Using the fact that γ_j converges to γ we conclude that there exists a segment $\xi \subset \gamma$ containing z, such that $e(\,\cdot\,)$ is constant on ξ and $f^{e(z,\gamma)}(\xi)$ is a long unstable leaf. See Lemma 4.5 below.

For the statements in the previous paragraph to be fully accurate we need to be slightly more precise about the definition of itinerary on $W^u(P)$ than was necessary up to this point. Also, we must restrict the construction to a subset of $\tilde{X}$ with *full probability*, in the sense that it intersects every long unstable leaf on a subset with full arc-length measure.

One problem is that the binding points are not uniquely defined, and different choices for the various points z_j might result in their itineraries being mostly unrelated. This is resolved by setting a definite selection rule right from the start: we introduce an (arbitrary) order in the critical set $\mathcal{C}$, and always take as binding point the first eligible critical point, that is, the smallest, with respect to this order, for which the tangential position condition is satisfied. Recall Section 2.9.

For a full probability subset of $\tilde{X}$, the orbit of z does not hit the vertical lines $x = \pm\delta$, at least not before the first time ν_1 it intersects $\{(x, y) : |x| < \delta\}$. Then ν_1 is also the first free return of z_j, if j is large: $n_{l(j)+1} = n_{l(j)} + \nu_1$. Let $\zeta_j = \tilde{\zeta}_{l(j)+1}(w_j)$ be the corresponding binding point. Suppose $\operatorname{dist}(f^{\nu_1}(z_j), \zeta_j)$ is not bounded from zero. Then, by Proposition 2.8.1, the bound period $p_{l(j)+1}(w_j)$ is not bounded above. Consequently, the escaping times $e(z_j)$ are arbitrarily large. It follows from Corollary 4.4 below that this happens only for subset of each leaf with zero arc-length measure. As we are concerned with full probability subsets only, we can neglect this case: we suppose from now on that $\operatorname{dist}(f^{\nu_1}(z_j), \zeta_j)$ is bounded from zero.

Note that $t(f^{\nu_1}(z_j))$ converges to $t(z)$, the tangent direction to $f^{\nu_1}(\gamma)$ at $f^{\nu_1}(z)$, as $j \to \infty$. Together with the assumption about the distance, this implies that $(f^{\nu_1}(z_j), t(f^{\nu_1}(z_j)))$ is in tangential position relative to $(\zeta_l, t(\zeta_l))$ for all large j and l. In view of the selection rule above, this means that the ζ_j are all the same for sufficiently large j. Let η_1 be this critical point. Moreover, $(f^{\nu_1}(z), t(f^{\nu_1}(z)))$ is in tangential position relative to $(\eta_1, t(\eta_1))$.

In addition, we may assume that $f^{\nu_1+1}(z)$ does not fall in any of the long stable leaves $\Gamma_{r,l}(\eta_1)$ associated to the critical value $f(\eta_1)$: this restriction also has full probability in $\tilde{X}$. It follows, by continuity, that all symbols $(r_{l(j)}, l_{l(j)}, \epsilon_{l(j)}))$ coincide for all large j. This proves (14) in the case $k = 1$. Furthermore, it ensures that the bound periods $p_{l(j)+1}$ are all the same for large j. We define their common value p_1 to be the bound period of z.

Now the argument proceeds in the same fashion. As before for ν_1, we may suppose that the next free return $\nu_2 > \nu_1 + p_1$ is simultaneous for all z_j with large j, as well as for z. Repeating the previous reasoning, each time for z_j in a smaller neighbourhood of z, we get the convergence (14) for every $k \geq 1$.

Also important is that the exponential estimate of Lemma 4.1 remains valid for every long unstable leaf.

Corollary 4.4. *Let $C > 0$ and $c > 0$ be the constants in Lemma 4.1. Then*

$$m_\gamma(\{z \in \gamma : e(z, \gamma) > n\}) \leq Ce^{-cn}$$

for all $n \geq 1$ and every long unstable leaf γ.

Proof. Let γ be a long unstable leaf, and $\{\gamma_j\}$ be a sequence of escaping leaves C^1 converging to γ. Let $A_{n,j} = \{z_j \in \gamma_j : e(z_j) \leq n\}$, for each $n \geq 1$ and $j \geq 1$. Then let A_n be the set of points $z \in \gamma$ that are limits of sequences $\{z_j\}$ with $z_j \in A_{n,j}$ for every j. Lemma 4.1 says that $m_{\gamma_j}(A_{n,j}) \geq m_{\gamma_j}(\gamma_j) - Ce^{-cn}$. So, as m_γ is a regular measure,

$$m_\gamma(A_n) \geq \limsup m_{\gamma_j}(A_{n,j}) \geq m_\gamma(\gamma) - Ce^{-cn}$$

(m_{γ_j} converges to m_γ, in the strong sense of uniform convergence of densities projected to the horizontal direction). On the other hand, by definition, if $z \in A_n$ then there exist $k \leq n$ and some subsequence of $\{z_j\}$ such that $e(z_j) = k$ for all j. This shows that $\{z \in \gamma : e(z, \gamma) > n\}$ is disjoint from A_n, and so its m_γ-measure is less than Ce^{-cn}. $\square$

4.4. **The return map R.** In particular, $e(z, \gamma)$ is finite for m_γ-almost every $z \in \gamma$ and every long unstable leaf γ.

Lemma 4.5. *Assuming $e(z,\gamma)$ is finite, there exists a segment $\xi \subset \gamma$ containing z and a long unstable leaf γ_1 such that $f^{e(z,\gamma)}$ maps ξ onto γ_1 and $e(w,\gamma) = e(z,\gamma)$ for all $w \in \xi$.*

Proof. Let $e(z,\gamma) = k$. By definition, there exist (z_j, γ_j), with $z_j \in X_0$, arbitrarily close to (z,γ). Also by definition, there exists a segment $\xi_j \subset \gamma_j$ containing z_j such that $e(w_j) = k$ for every $w_j \in \xi_j$ and $f^k(\xi_j)$ is an escaping leaf. Up to restricting to subsequences, we may suppose that $f^k(\xi_j)$ converges to a long unstable leaf γ_1 and ξ_j converges to a segment $\xi \subset \gamma$. It is clear that $z \in \xi$ and $f^k(\xi) = \gamma$. We are left to show that $e(w,\gamma) = k$ for all $w \in \xi$.

From the definition of escaping times in $\tilde{X}$ we have that $e(w,\gamma) \le k$ for all $w \in \xi$. Suppose there was $z' \in \xi$ such that $e(z',\gamma) = l < k$. Arguing as before, with z' in the place of z, we would find $\xi' \subset \gamma$ containing z', such that $f^l(\xi')$ is a long unstable leaf and $e(w',\gamma) \le l$ for every $w' \in \xi'$. We may suppose that $f^l(z')$ is in the interior of $\mathcal{R}_\Delta$, replacing z' by some nearby point in $\xi \cap \xi'$ if necessary. Moreover, $f^l(z)$ is in the exterior of $\mathcal{R}_\Delta$, because z does not belong to ξ' nor to the f^l-pre-image of the long stable leaves Γ_Δ and $\Gamma_{\Delta+1}$ (a full probability restriction). This means that $f^l(\xi)$ would intersect both the interior and the exterior of $\mathcal{R}_\Delta$. Then the same would be true about $f^l(\xi_j)$, for large j. But that would contradict the definition of escape situation: all the points of ξ_j have the same itinerary up to time k. This proves that the escaping time is indeed constant on ξ. □

Now we are ready to define our return map $R : \tilde{X} \to \tilde{X}$: using the notations of Lemma 4.5, we set

$$R(z,\gamma) = \big(f^{e(z,\gamma)}(z), \gamma_1\big) \tag{15}$$

for every (z,γ) with finite escaping time. Thus, the domain of R is a subset of $\tilde{X}$ intersecting every long unstable leaf γ in a full m_γ-subset.

According to Lemma 4.5 this map has a Markov type property: the image of any unstable leaf is a union of complete long unstable leaves. Consequently, the same is true for every iterate R^n, $n \ge 1$.

4.5. **Expansion and distortion.** Our goal in this section is to prove that the map R is expanding, and has a bounded distortion property along long unstable leaves, cf. Proposition 4.7 below.

For every $(z,\gamma) \in \tilde{X}$ we denote by $t(z) = t(z,\gamma)$ the norm 1 vector tangent to γ at z and pointing to the right (this is independent of γ, by Lemma 4.3.1). Then we let the *unstable derivative* $R'(z)$ be the number defined by

$$Df^{e(z)}(z)\, t(z) = R'(z) t\big(R(z,\gamma)\big).$$

Lemma 4.6. *There are constants $C > 0$ and $\lambda_0 < 1$ such that, for any unstable leaf γ, and every $z, w \in \gamma$ and $n \geq 1$,*

1. $\|Df^{-n}(z)t(z)\| \leq C\lambda_0^n$ *and*
2. $\dfrac{\|Df^{-n}(z)t(z)\|}{\|Df^{-n}(w)t(w)\|} \leq C.$

Proof. Suppose first that γ is an escaping leaf $\gamma = f^{n_k+1}(\omega)$. The first claim is a consequence of the fact that n_k is a free return for the segment $\omega \subset \Omega$; see [7, Lemma 7.13]. Observe also that, by construction, points in ω have the same itinerary up to time n_k. In particular, $f^{n_i}(\omega)$ is in tangential position to some critical point, at every free return $n_i \leq n_k$. This means that [10, Proposition 2] is applicable, and we conclude the bounded distortion statement in the second claim. The constants C and λ_0 we get in this way are independent of the escaping leaf.

Now let γ be any unstable leaf. By definition, there exists a sequence $\{\gamma_j\}$ of escaping leaves converging to γ in the C^1 topology. This means that we can find $z_j \in \gamma_j$ converging to z, and $t(z_j)$ converges to $t(z)$. Then $\|Df^{-n}(z_j)t(z_j)\|$ converges to $\|Df^{-n}(z)t(z)\|$ when $j \to \infty$, for each fixed $n \geq 1$. Thus, the two properties in the lemma follow from the corresponding facts for the escaping curves γ_j, obtained in the first paragraph of the proof, and the observation that the constants did not depend on j. $\square$

Proposition 4.7. *The map R is uniformly expanding and has bounded distortion along unstable leaves:*

1. $|(R^k)'(x)|^{-1} \leq C\lambda_0^k$ *for all $(x,\gamma) \in \tilde{X}$ and $k \geq 1$, and*
2. $\dfrac{|(R^k)'(y)|}{|(R^k)'(x)|} \leq C$ *for any $k \geq 1$ and (x,γ), $(y,\gamma) \in \tilde{X}$ such that R^i is smooth on the segment $[x,y] \subset \gamma$ for all $1 \leq i \leq k$.*

Proof. Let $n = e(x,\gamma) + e\big(R(x,\gamma)\big) + \cdots + e\big(R^{i-1}(x,\gamma)\big)$. Then we have $R^k(x,\gamma) = (f^n(x), \gamma_k)$ and

$$|(R^k)'(x)| = \|Df^n(x)t(x)\| = 1/\|Df^{-n}(z)t(z)\|, \quad \text{where } z = f^n(x).$$

Thus, claim 1 is a direct consequence of the first part of Lemma 4.6, and the fact that $n > k$.

Similarly, $R^k(y,\gamma) = (f^n(y), \gamma_k)$, for the same long unstable leaf γ_k, and $|(R^k)'(y)| = 1/\|Df^n(w)t(w)\|$, with $w = f^n(y)$. So, claim 2 follows directly from the second statement in Lemma 4.6. $\square$

4.6. **Measures absolutely continuous along unstable leaves.** Fix a map $\pi_1 : \tilde{X} \to \mathbb{R}$ induced by some submersion from a neighbourhood of X to $\mathbb{R}$ sending each long unstable leaf onto the same interval $I \subset \mathbb{R}$, diffeomorphically. Let $\mathcal{U}$ be the family of long unstable leaves, endowed with the topological and measurable structure induced by the order relation. Let $\pi_2 : \tilde{X} \to \mathcal{U}$ be the canonical projection $\pi_2(z, \gamma) = \gamma$.

$\tilde{X}$ is identified with $I \times \mathcal{U}$ via the bijection $(\pi_1, \pi_2) : \tilde{X} \to I \times \mathcal{U}$. Let $\mathcal{A}$ be the σ-algebra in $\tilde{X}$ generated by the products $A \times B$ of measurable sets $A \subset I$ and $B \subset \mathcal{U}$. Given a Borel measure ν on X, let $m \times \hat{\nu}$ be the measure defined on $\mathcal{A}$ by

$$(m \times \hat{\nu})(A \times B) = m(A) \times \hat{\nu}(B),$$

where m is Lebesgue measure and $\hat{\nu} = (\pi_2)_* \nu$.

We say that ν is *absolutely continuous along unstable leaves* if it is absolutely continuous with respect to $m \times \hat{\nu}$: there exists an $\mathcal{A}$-measurable function $\rho : X \to \mathbb{R}$ such that

$$\nu = \rho\,(m \times \hat{\nu}). \tag{16}$$

Then the conditional probability measures $\{\nu_\gamma : \gamma \in \mathcal{U}\}$ of ν relative to the partition $\mathcal{U}$ (see Rokhlin [29]) are absolutely continuous with respect to m: one may take $\nu_\gamma = (\rho \mid \gamma) m$ for every $\gamma \in \mathcal{U}$.

The following simple lemma will be useful later:

Lemma 4.8. *Let $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ be Borel measures on $\tilde{X}$.*

1. *If $\lambda_1 \ll \lambda_2$ and λ_2 is absolutely continuous along unstable leaves, then λ_1 is absolutely continuous along unstable leaves.*
2. *If λ_1 and λ_2 are absolutely continuous along unstable leaves, then $\lambda_1 + \lambda_2$ is absolutely continuous along unstable leaves.*
3. *If every λ_n, $n \geq 1$, is absolutely continuous along unstable leaves and $\eta = \sum_{n=1}^{\infty} \lambda_n$ is a finite measure, then η is absolutely continuous along unstable leaves.*

Proof. The hypothesis $\lambda_1 \ll \lambda_2$ implies that $\hat{\lambda}_1 \ll \hat{\lambda}_2$. Let ϕ and $\hat{\phi}$ be the Radon-Nikodym derivatives, that is, $\lambda_1 = \phi \lambda_2$ and $\hat{\lambda}_1 = \hat{\phi} \hat{\lambda}_2$. Moreover, let $\lambda_2 = \rho_2\,(m \times \hat{\lambda}_2)$. On $I \times \{\hat{\phi} > 0\}$ we have

$$\lambda_1 = \phi \lambda_2 = \phi \rho_2\,(m \times \hat{\lambda}_2) = \phi \rho_2 \psi\,(m \times \hat{\lambda}_1),$$

where $\psi(z, \gamma) = 1/\hat{\phi}(\gamma)$. Since $I \times \{\hat{\phi} > 0\}$ has full λ_1-measure, this proves that λ_1 is absolutely continuous along unstable leaves, as claimed in 1.

To prove part 2, we begin by noting that $\lambda_i \ll \lambda_1 + \lambda_2$, and so $\hat{\lambda}_i \ll \hat{\lambda}_1 + \hat{\lambda}_2$, for $i = 1, 2$. So, let us write $\hat{\lambda}_i = \phi_i\,(\hat{\lambda}_1 + \hat{\lambda}_2)$ and $\lambda_i = \rho_i\,(m \times \hat{\lambda}_i)$. Then $\lambda_1 + \lambda_2 = (\rho_1 \psi_1 + \rho_2 \psi_2) m \times (\hat{\lambda}_1 + \hat{\lambda}_2)$ where $\psi_i(z, \gamma) = \phi_i(\gamma)$.

Finally, let $\{\lambda_n\}$ be as in 3. By part 2, every $\eta_n = \sum_{i=1}^n \lambda_i$ is absolutely continuous along unstable leaves. Let $E \subset \tilde{X}$ be any measurable subset such that $\eta(E) > 0$. Then $\eta_n(E) > 0$, and so $(m \times \hat{\eta}_n)(E) > 0$, for every large n. On the other hand, $\eta_n \nearrow \eta$ implies $\hat{\eta}_n \nearrow \hat{\eta}$, and so $(m \times \hat{\eta}_n) \nearrow (m \times \hat{\eta})$. In particular, $(m \times \hat{\eta})(E) \geq (m \times \hat{\eta}_n)(E) > 0$. This proves that $\eta \ll m \times \hat{\eta}$, as claimed in part 3. $\square$

4.7. **SRB measure for the return map.** We are going to prove that R has exactly one invariant probability measure absolutely continuous along unstable leaves. The main step is

Lemma 4.9. *There exists $K > 0$ such that, given any long unstable leaf γ, the sequence $\{\lambda_n = R_*^n m_\gamma\}$ satisfies*

$$\lambda_n(A \times B) \leq K m(A)\, \hat{\lambda}_n(B)$$

for every $n \geq 1$ and $A \times B \in \mathcal{A}$.

Proof. By the Markov property in Lemma 4.5, $R^k(\gamma)$ is a union of long unstable leaves: there exist segments ξ_i such that $\gamma = \cup_i \xi_i$, up to a zero m_γ-measure set, and each $\gamma_i = R^k(\xi_i)$ is a long unstable leaf. With our notations, a leaf γ_i intersects $A \times B$ if and only $\gamma_i \in B$. In that case there exists a segment $\eta_i \subset \xi_i$ that is mapped diffeomorphically to $A \approx A \times \{\gamma_i\}$ by R^k. Then

$$\lambda_n(A \times B) = \sum_{\gamma_i \in B} m_\gamma(\eta_i) \quad \text{and} \quad \hat{\lambda}_n(B) = \lambda_n(I \times B) = \sum_{\gamma_i \in B} m_\gamma(\xi_i).$$

By part 2 of Proposition 4.7, together with the mean value theorem, there exist positive constants C_1 and C_2 such that

$$\frac{m_\gamma(\eta_i)}{m_\gamma(\xi_i)} \leq C_1 \frac{m_{\gamma_i}(A)}{m_{\gamma_i}(\gamma_i)} \leq C_2 \frac{m(A)}{m(I)}$$

for every i. The second inequality uses the fact that π_1 is a diffeomorphism on each leaf, and so the measures m_{γ_i} are uniformly equivalent to m.

Putting these relations together we obtain $\lambda_n(A \times B) \leq K\, m(A) \hat{\lambda}(B)$, with $K = C_2/m(I)$. $\square$

Since the measurable sets $A \times B$ generate the σ-algebra $\mathcal{A}$, Lemma 4.9 implies that every λ_n is absolutely continuous along unstable leaves, with Radon-Nikodym density ρ_n bounded by K. Moreover, the same is true for the sequence

$$\nu_n = \frac{1}{n} \sum_{j=0}^{n-1} \lambda_j = \frac{1}{n} \sum_{j=0}^{n-1} R_*^j m_\gamma. \tag{17}$$

That is because $\lambda_j(A \times B) \le Km(A)\hat{\lambda}_j(B)$ for all j implies

$$\nu_n(A \times B) = \sum_{j=0}^{n-1} \lambda_j(A \times B) \le Km(A) \sum_{j=0}^{n-1} \hat{\lambda}_j(B) = Km(A)\hat{\nu}_n(B).$$

Corollary 4.10. *Given any long unstable leaf γ, every weak*-accumulation point ν of the sequence $\{\nu_n\}$ is an R-invariant probability measure absolutely continuous along unstable leaves, with density bounded by the constant K.*

Proof. Invariance $R_*\nu = \nu$ follows from $R_*\nu_n - \nu_n = n^{-1}(\lambda_n - \lambda_0)$. To see that ν is absolutely continuous along unstable leaves, consider any measurable sets $A \subset I$ and $B \subset \mathcal{U}$ such that $\nu(\partial A \times \mathcal{U}) = 0$ and $\nu(I \times \partial B) = 0$. Then $\nu(A \times B) = \lim \nu_n(A \times B)$ and $\hat{\nu}(B) = \nu(I \times B) = \lim \nu_n(I \times B) = \lim \hat{\nu}_n(B)$, because the boundaries of $A \times B$ and $I \times B$ have zero ν-measure. Using Lemma 4.9, we conclude that $\nu(A \times B) \le Km(A)\,\hat{\nu}(B)$ for any such A and B. Since the family of these sets $A \times B$ generates the σ-algebra $\mathcal{A}$, up to zero ν-measure subsets, this proves that ν is absolutely continuous with respect to $m \times \hat{\nu}$, with density bounded by K almost everywhere. □

For the next result we need some information about the dynamics transverse to the long unstable leaves. This is provided by the family $\mathcal{H}$ of long stable leaves in Proposition 3.6. Let $\tilde{\mathcal{H}}$ be the family of pre-images $\tilde{\Gamma} = p^{-1}(\Gamma)$ of the leaves $\Gamma \in \mathcal{H}$, and $\tilde{H} = p^{-1}(H)$ be the union of all such $\tilde{\Gamma}$. Since itineraries are constant on each $\Gamma \in \mathcal{H}$, every R^j corresponds to an iterate f^{m_j} with m_j constant on $\tilde{\Gamma}$. Thus, the R-orbits of any two points in each $\tilde{\Gamma}$ are forward asymptotic, because the f-orbits of points in the same long stable leaf are. We say that $\tilde{\Gamma}$ is a *stable set* for R.

Lemma 4.11. *Let ν be any R-invariant probability measure absolutely continuous along unstable leaves. Then ν is ergodic for the return map R, and its basin*

$$B(\nu) = \{z \in \tilde{X} : \lim \frac{1}{n} \sum_{j=0}^{n-1} \delta_{R^k(z)} = \nu\}$$

intersects every long unstable leaf γ on a full m_γ-measure subset.

Proof. The main step is the following sublemma. A set $F \subset X$ is called $\tilde{\mathcal{H}}$-saturated if it consists of entire stable sets $\tilde{\Gamma} \in \tilde{\mathcal{H}}$.

Sublemma. *Let $F \subset \tilde{X}$ be an R-invariant $\tilde{\mathcal{H}}$-saturated measurable set such that $m_\alpha(\alpha \cap F) > 0$ for some long unstable leaf α. Then $m_\beta(\beta \setminus F) = 0$ for every long unstable leaf β.*

Proof. Let w be a density point of $\alpha \cap F$, for the arc-length measure m_α. By the Markov and bounded distortion properties in Lemma 4.5 and Proposition 4.7, we may find segments $\xi_i \ni w$ inside α such that some iterate $R^{l_i}(\xi_i) = \alpha_i$ is a long unstable leaf, and

$$\frac{m_{\alpha_i}(\alpha_i \setminus F)}{m_{\alpha_i}(\alpha_i)} = \frac{m_{\alpha_i}(R^{l_i}(\xi_i \setminus F))}{m_{\alpha_i}(R^{l_i}(\xi_i))} \leq C \frac{m_\alpha(\xi_i \setminus F)}{m_\alpha(\xi_i)} \to 0 \tag{18}$$

as $i \to \infty$. Suppose there existed some leaf β with $m_\beta(\beta \setminus F) > 0$. Then we could apply the same arguments to the complement F^c of F, which is also R-invariant and $\tilde{\mathcal{H}}$-saturated, to find a sequence of long unstable leaves β_j such that $m_{\beta_j}(\beta_j \setminus F^c)$ converges to zero. In particular, $m_{\beta_j}(\beta_j \setminus F^c) < \varepsilon_0/2$ for large j, where $\varepsilon_0 > 0$ is as in Proposition 3.6. It would follow that $m_{\beta_j}(\beta_j \cap F^c \cap \tilde{H}) > \varepsilon_0/2$. Then, using the Lipschitz property in (13),

$$m_\alpha(\alpha \setminus F) \geq m_\alpha(\alpha \cap F^c \cap \tilde{H}) \geq \frac{1}{2} m_{\beta_j}(\beta_j \cap F^c \cap \tilde{H}) > \frac{\varepsilon_0}{4}$$

for every α, which would contradict (18). Therefore, we must have $m_\beta(\beta \setminus F) = 0$ for every long unstable leaf β, as claimed in the sublemma. □

In order to prove Lemma 4.11, let $\varphi : \tilde{X} \to \mathbb{R}$ be a continuous function, and θ be any real number. Let E be the set of points z such that the forward time-average of φ on the R-orbit of z converges and is less than θ. Then E is R-invariant, and it is $\tilde{\mathcal{H}}$-saturated because the $\tilde{\Gamma} \in \tilde{\mathcal{H}}$ are stable sets for R. Suppose $\nu(E) > 0$. As ν is absolutely continuous along unstable leaves, we must have $m_\alpha(\alpha \cap E) > 0$ for some long unstable leaf α. It follows from the sublemma that $m_\beta(\beta \setminus E) = 0$ for every long unstable leaf β. Using absolute continuity once more, we get that E has full ν-measure: $\nu(\tilde{X} \setminus E) = 0$. This shows that time-averages of continuous functions are constant ν-almost everywhere. Therefore, ν is ergodic.

The proof of the last statement in the lemma is similar. By ergodicity, the basin of ν has full ν-measure, and so it intersects some long unstable leaf on a positive arc-length measure set. By the sublemma, the intersection really has full measure, for every long unstable leaf. □

Remark 4.12. The same argument proves that ν is ergodic for every R^k, with $k \geq 1$.

The next proposition summarizes the main facts in this section:

Proposition 4.13. *The map R has exactly one R-invariant probability measure ν absolutely continuous along unstable leaves. Moreover, ν is* *ergodic, its density is bounded by K, and its basin intersects every long unstable leaf on a full arc-length measure subset.*

4.8. **SRB measure for the attractor.** Define μ to be the saturation of $\mu_0 = p_*\nu$ under f, that is,

$$\mu = \sum_{j=0}^{\infty} f_*^j p_* \left(\nu \mid \{e > j\}\right). \tag{19}$$

Lemma 4.14. *The measure μ is finite, f-invariant and ergodic.*

Proof. Corollary 4.4 says that $m_\gamma(\gamma \cap \{e > j\}) \le Ce^{-cj}$ for every long unstable leaf γ. From Corollary 4.10 we deduce $\nu(\{e > j\}) \le KCe^{-cj}$ for all $j \ge 1$. Thus the series (19) converges, and defines a finite measure.

Since $f^j = R$ on $\{e = j\}$, and the measure ν is R-invariant,

$$\sum_{j=1}^{\infty} f_*^j p_* \big(\nu \mid \{e = j\}\big) = \sum_{j=1}^{\infty} p_* R_* \big(\nu \mid \{e = j\}\big) = p_* R_* \nu = p_* \nu.$$

The f-invariance of μ is an easy consequence: writing

$$\begin{aligned} f_*\mu &= \sum_{j=0}^{\infty} f_*^{j+1} p_* \big(\nu \mid \{e > j\}\big) \\ &= \sum_{i=1}^{\infty} f_*^{i} p_* \big(\nu \mid \{e > i\}\big) + \sum_{i=1}^{\infty} f_*^{i} p_* \big(\nu \mid \{e = i\}\big), \end{aligned}$$

we conclude that

$$f_*\mu = \sum_{i=1}^{\infty} f_*^{i} p_* \big(\nu \mid \{e > i\}\big) + p_*\nu = \sum_{i=0}^{\infty} f_*^{i} p_* \big(\nu \mid \{e > i\}\big) = \mu.$$

Finally, let $E \subset \Lambda$ be an f-invariant measurable set. Then the pre-image $p^{-1}(E) = p^{-1}(E \cap X)$ is R-invariant. By Lemma 4.11, either $\nu\big(p^{-1}(E)\big) = 0$ or $\nu\big(\tilde{X} \setminus p^{-1}(E)\big) = 0$. In the first case,

$$\mu(E) = \sum_{j=0}^{\infty} f_*^j p_* \big(\nu \mid \{e > j\}\big)(E) = \sum_{j=0}^{\infty} \nu\big(p^{-1}(E) \cap \{e > j\}\big) = 0.$$

In the second case we get $\mu(\Lambda \setminus E) = 0$, by the same argument applied to the complement of E. This proves that μ is ergodic. □

By Lemma 4.3, every connected component of X may be decomposed into finitely many segments ξ_i each of which lifts to a finite number segments $\xi_{i,j} \subset \tilde{X}$, that are projected diffeomorphically onto ξ_i by $p : \tilde{X} \to X$. See Figure 5.

Let $\mathcal{Q}_0$ be the partition of X into the segments ξ_i. Let $X_0 = X$ and

$$X_s = f^s(X) \setminus \cup_{j=0}^{s-1} f^j(X), \tag{20}$$

for each $s \geq 1$. Define $\mathcal{Q}_s$ to be the partition of X_s whose atoms are the sets $f^s(\xi_i) \setminus \cup_{j=0}^{s-1} f^j(X)$ with $\xi_i \in \mathcal{Q}_0$. Finally, let $\mathcal{Q} = \cup_{s=0}^{\infty} \mathcal{Q}_s$. It follows from the construction that $\mathcal{Q}$ is a measurable partition, in the sense of [29]: it is a countable product of finite partitions.

We say that a measure η on $\cup_{s=0}^{\infty} f^s(X)$ is *absolutely continuous along unstable manifolds* if its conditional measures $\{\eta_Q : Q \in \mathcal{Q}\}$ for the partition $\mathcal{Q}$ are absolutely continuous with respect to arc-length, almost everywhere.

Remark 4.15. The measure $\mu_0 = p_* \nu$ is absolutely continuous along unstable manifolds. To see this, consider the partition $\tilde{\mathcal{Q}}_0$ of $\tilde{X}$ into the segments $\xi_{i,j}$ above. Each long unstable leaf γ contains a finite number of elements of $\tilde{\mathcal{Q}}_0$. Thus, the conditional measures $\{\nu_\gamma : \gamma \in \mathcal{U}\}$ of ν for the partition $\mathcal{U}$ are finite convex combinations of the conditional measures $\{\tilde{\nu}_{\tilde{Q}} : \tilde{Q} \in \tilde{\mathcal{Q}}_0\}$ of ν for $\tilde{\mathcal{Q}}$. Since the ν_γ are absolutely continuous with respect to arc-length, the same is true for almost every $\tilde{\nu}_{\tilde{Q}}$. Moreover, p projects each element of $\tilde{\mathcal{Q}}_0$ diffeomorphically to some $Q \in \mathcal{Q}_0$, in a finite-to-1 fashion. In particular, the conditional measures $\{\mu_{0,Q} : Q \in \mathcal{Q}_0\}$ of $\mu_0 = p_* \nu$ for $\mathcal{Q}_0$ are finite convex combinations of the images $p_*(\tilde{\nu}_{\tilde{Q}})$. It follows that the $\mu_{0,Q}$ are almost everywhere absolutely continuous with respect to arc-length, as claimed.

Lemma 4.16. *The measure μ is absolutely continuous along unstable manifolds.*

Proof. The proof has two steps. First we consider the part μ_X of μ sitting in X, corresponding to returns to X prior to the escaping time.

More precisely, let $\{r_i\}$ be the sequence of return time functions: $r_i(z) = r_i(z, \gamma)$ is the i:th element of the set of times $r \geq 1$ for which $f^r(z)$ is in X. By convention, $r_i(z) = \infty$ if z returns less than i times. We also set $r_0(z, \gamma) = 0$ at all points. Let

$$\mu_X = \sum_{i=0}^{\infty} \mu_i \quad \text{where} \quad \mu_i = \sum_{j=0}^{\infty} f_*^j p_* \big(\nu \mid \{j = r_i \ \& \ r_i < e\}\big) \tag{21}$$

for each $i \geq 0$. Each $\nu \mid \{j = r_i \ \& \ r_i < e\} \leq \nu$ is absolutely continuous along unstable leaves, by Lemma 4.8. Then, cf. Remark 4.15, its image under p_* is a measure in X absolutely continuous along unstable manifolds. Then the same is true for $f_*^j p_* \big(\nu \mid \{j = r_i \ \& \ r_i < e\}\big)$, because f is a diffeomorphism (and because $j = r_i$ is a return time).

Using Lemma 4.8, we conclude that every μ_i is absolutely continuous along unstable manifolds, and then so is μ_X.

In the second step we derive the same conclusion for μ itself. Observe that μ may be written as

$$\sum_{i,j=0}^{\infty} f_*^j p_*\big(\nu \mid \{j = r_i \ \&\ r_i < e\}\big) + f_*^j p_*\big(\nu \mid \{r_i < j < r_{i+1} \ \&\ r_i < e\}\big)$$
$$= \sum_{i=0}^{\infty} \mu_i + \sum_{i=0}^{\infty}\sum_{s=1}^{\infty}\sum_{k=0}^{\infty} f_*^{k+s} p_* \nu_i(s,k)$$

with $\nu_i(s,k) = \nu \mid \{r_i < k+s < r_{i+1} \ \&\ r_i = k \ \&\ r_i < e\}\}$. Writing

$$\mu_i(s) = \sum_{k=0}^{\infty} f_*^k \mu_i(s,k) \quad \text{and} \quad \mu_X(s) = \sum_{i=0}^{\infty} \mu_i(s), \tag{22}$$

for each $s \geq 1$, we obtain

$$\mu = \sum_{i=0}^{\infty} \mu_i + \sum_{i=0}^{\infty}\sum_{s=1}^{\infty} f_*^s \mu_i(s) = \mu_X + \sum_{s=1}^{\infty} f_*^s \mu_X(s).$$

It is clear from their definitions in (21) and (22) that $\mu_i(s) \leq \mu_i$ for every $i \geq 0$, and so $\mu_X(s) \leq \mu_X$, for every $s \geq 1$. So, by Lemma 4.8, every $\mu_X(s)$ is absolutely continuous along unstable manifolds. Then every $f_*^s \mu_X(s)$ is absolutely continuous along unstable manifolds, as f is a diffeomorphism. Finally, by construction, each $f_*^s \mu_X(s)$ sits on the set X_s defined by (20). Since these sets are two-by-two disjoint, it follows that the sum μ is also absolutely continuous along unstable manifolds. □

Corollary 4.17. *The normalization $\mu_* = \mu/\mu(\Lambda)$ of μ is the SRB measure of the Hénon-like diffeomorphism f.*

Proof. By Lemmas 4.14 and 4.16, the probability measure μ_* is f-invariant, ergodic, and absolutely continuous along unstable manifolds. Moreover, it has one positive Lyapunov exponent: μ_* gives positive weight to X, and every long unstable leaf contained in X is exponentially contracted by negative iterates. The other Lyapunov exponent is negative, because the diffeomorphism f is area-contracting. It follows from general non-uniform hyperbolicity theory [25, 26] that the union of the stable manifolds through points in the basin of μ has positive area. Since this union is still contained in the basin, this shows that μ_* is SRB measure for f. □

5. Itineraries for random perturbations

We are going to associate to each pair $(z, \mathbf{g})$ of initial points z and $\mathbf{g} \in \Omega_\varepsilon^{\mathbb{N}}$ an *itinerary* $\{i_j(z, \mathbf{g}) : j \geq 0\}$, as well as a sequence of *free returns* $\{n_j(z, \mathbf{g}) : j \geq 0\}$. As before, we denote the random orbit $z_j = g_j \dots g_1(z)$, $j \geq 0$ by $\mathbf{z}$. Since all maps $g \in \Omega_\varepsilon$ are invertible, the correspondence between $\mathbf{z}$ and $(z, \mathbf{g})$ is one-to-one and we can use either notation whenever convenient.

To some extent, the symbols $i_j(z, \mathbf{g}) = (\tilde{\zeta}_j, r_j, l_j, \epsilon_j, \mathbf{y}_j)$ have the same meaning as in the deterministic case. To begin with, $\tilde{\zeta}_j$ is a critical point of the unperturbed map f, near the iterate z_{n_j}.

If z_{n_j} is not too close to $\tilde{\zeta}_j$, then (r_j, l_j, ϵ_j) also have the same meaning as before, namely, they describe the position of z_{n_j+1} with respect to the long stable leaves $\Gamma_{r,l}$ of the critical value $f(\tilde{\zeta}_j)$. In this case, the random iterates z_{n_j+i} remain close to the unperturbed orbit $f^i(\tilde{\zeta}_j)$ all the way through the deterministic bound period of z_{n_j}, so that the main estimates of the deterministic case remain valid for the z_{n_j+i}, up to the next free return n_{j+1}. In this case $\mathbf{y}_j$ has no role; for completeness we set it to be 0.

The main difference occurs when z_{n_j} is close to $\tilde{\zeta}_j$: distance $< \varepsilon^{1-\theta_0}$ for some small $\theta_0 > 0$. We call this an *ϵ-situation*. In this case the deterministic bound period is too long, and accumulated random effects become important before it is over. We can still define a bound period for the random orbit z_{n_j+i}, as we shall see, but it depends mostly on the noise level, not on the position of z_{n_j+1} relative to the critical value.

According to assumption (H1), the point z_{n_j+1} is almost surely in a domain $\Lambda_{\varepsilon, z_{n_j}}$ that may be laminated into nearly horizontal curves, such that most laminae are not too small. We take $\mathbf{y}_j$ to be the lamina that contains z_{n_j+1}. On the other hand, r_j, l_j, and ϵ_j have no role; for completeness we let $r_j = l_j = \epsilon_j = 0$.

Another main ingredient is to find a suitable binding point for the random orbit $\mathbf{z}$ at the next return n_{j+1}. For this purpose we introduce a *capture construction* for random perturbations: we find a segment L of the unstable manifold $W^u(P)$ whose deterministic trajectory shadows the random orbit on a time interval $[n_{j+1} - \tau, n_{j+1}]$. Then we take the binding point for $\mathbf{z}$ to coincide with the binding point of $f^\tau(L)$ for the unperturbed system f.

The precise definition of itineraries for random orbits follows. The noise level ε is fixed throughout this section.

5.1. Itineraries: Step zero.

For our purposes it is enough to consider itineraries for $(z, \mathbf{g})$, where $z = z_0$ belongs to some segment γ_0 of the unstable manifold of P. For definiteness we pick $\gamma_0 = f^{-1}(\Omega)$. Recall that Ω is the unstable side of the domain Δ_0 that has P as an endpoint.

As a first stage, we describe how to define the symbol $i_0(z, \mathbf{g})$ for $z \in \gamma_0$. Let ζ_0 be the critical point contained in $f^{-1}(\Omega)$ (compare Proposition 2.2). We take the binding point to be $\tilde{\zeta}_0 = \zeta_0$ and we define $n_0(z, \mathbf{g}) = 0$. Let $s(\varepsilon) \in \mathbb{N}$ be defined by

$$e^{-1}\varepsilon \le e^{-2s(\varepsilon)} < e\varepsilon.$$

(1) If z_1 is to the right of $\Gamma_{s(\varepsilon)}$ then $i_0(z, \mathbf{g}) = (\zeta_0, 0, 0, 0, \mathbf{y}_0)$ where $\mathbf{y}_0$ is the lamina of $\Lambda_{\varepsilon, z_0}$ that contains z_1. See Figure 6. We refer to this case as an *ε-situation.*

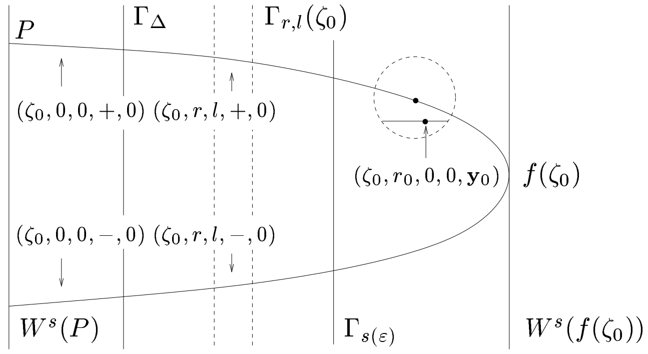

FIGURE 6.

(2) If z_1 is to the left of $\Gamma_{s(\varepsilon)}$ (we call this a deterministic situation):
   (a) $i_0(z, \mathbf{g}) = (\zeta_0, r, l, \pm, 0)$ if z_1 is in the region of Δ_0 in between $\Gamma_{r,l}(\zeta_0)$ and $\Gamma_{r,l-1}(\zeta_0)$, with $(r, l) > (\Delta, 0)$ the sign $+/-$ corresponding to whether z_0 is to the right or to the left of the critical point;
   (b) $i_0(z, \mathbf{g}) = (\zeta_0, 0, 0, \pm, 0)$ if z_1 is in between Γ_Δ and $W^s_{loc}(P)$, the sign $+/-$ corresponding to whether z_0 is to the right or to the left of the critical point.

5.2. **Step 1: Bound period.** We will now describe the first inductive step. We start from a curve γ_1 of either of the following kinds:

- γ_1 is some lamina of $\Lambda_{\varepsilon, z_0}$; this corresponds to a first symbol $i_0(z, \mathbf{g})$ of type (1).
- γ_1 is the image under fixed g_1 of a sub-segment of γ_0 corresponding to prescribed first symbol $i_0(z, \mathbf{g})$ of type (2).

In cases (1) and (2a) we define the bound period associated to the return n_0 as $[1, p_1]$, where p_1 is the largest integer so that

$$|f^j(\tilde{\zeta}_0) - g_2^{j-1}(z_1)| \le e^{-\beta j} \quad \text{for all } \mathbf{g} \in \Omega_\varepsilon^{\mathbb{N}},\ z_1 \in \gamma_1, \text{ and } 1 \le j \le p_1.$$

Here $g_2^{j-1} = g_j \circ \cdots \circ g_2$. In case (2b) we simply take $p_1 = 0$.

We have two basic lemmas concerning the distortion properties of an expanded vector, that may be thought of as extensions of Proposition 2.8 to the present random setting. The first lemma corresponds to ε-situations and the second one to deterministic situations. The symbol $\approx$ means that the quotient between the two expressions is bounded above and below by constants C and c, respectively.

Lemma 5.1. *Suppose γ_1 corresponds to case (1) above. Then for every $z_1 \in \gamma_1$ and $\mathbf{g} \in \Omega_\varepsilon^{\mathbb{N}}$,*

(a) $p_1 \approx \log \frac{1}{\varepsilon}$;

(b) $\left\|Dg_2^j(z_1)\,(1,0)\right\| \approx \left\|Df^j(f(\tilde{\zeta}_0))\,(1,0)\right\| \geq e^{cj}$ *for* $1 \leq j \leq p_1$;

c) $\|Dg_2^{p_1}(z_1)(1,0)\| \geq \varepsilon^{-9/10}$.

Lemma 5.2. *Suppose γ_1 corresponds to case (2a) above. Then for every $z_1 \in \gamma_1$ and $\mathbf{g} \in \Omega_\varepsilon^{\mathbb{N}}$,*

a) $p_1 \approx r$;

b) $\left\|Dg_2^j(z_1)\,(1,0)\right\| \approx \|Df^j(f(\tilde{\zeta}_0))\,(1,0)\| \geq e^{cj}$ *for* $1 \leq j \leq p_1$;

c) $\|Dg_2^{p_1}(z_1)\,(1,0)\| \geq ce^{2r(9/10)}$;

d) $\left\|Dg^{p_1+1}(g_1^{-1}(z_1))\,(1,0)\right\| \geq \sigma_1^{(p_1+1)/3}$.

Proofs of Lemmas 5.1 and 5.2. We only outline the arguments, since all the ingredients are well-known by now. Indeed, the two statements are higher-dimensional versions of, e.g., Lemmas 5.3 and 4.4 in [9]. Moreover, distortion bounds of this kind have been obtained before in higher dimensions, for instance in [22, Lemma 10.5], and the same estimates apply here.

First of all, the definition of bound period implies that

$$\operatorname{dist}(g_2^{j-1}(z_1), f^j(\tilde{\zeta}_0)) \leq e^{-\beta j} \ll e^{-\alpha j}$$

is exponentially smaller than the distance from $f^j(\tilde{\zeta}_0)$ to the critical set. One deduces that

$$\|Dg_2^j(z_1)(1,0)\| \approx \|Df^j(f(\tilde{\zeta}_0))(1,0)\| \approx \|Df^j(z_1)(1,0)\|$$

for all $1 \leq j \leq p_1$, as claimed in part (b) of either lemma. An important point here is that the derivatives of all maps g_i are Lipschitz continuous, with uniform Lipschitz constant. Recall Remark 1.4.

Now observe that $\sup_{\mathbf{g}} \operatorname{dist}(g_2^j(z_1), f^{j+1}(\tilde{\zeta}_0))$ is given, essentially, by

$$\begin{aligned} \varepsilon + \|Df(f^j(\tilde{\zeta}_0))\|\varepsilon + \cdots + \|Df^j(f(\tilde{\zeta}_0))\|\varepsilon& \\ + \|Df^j(f(\tilde{\zeta}_0))(1,0)\| \operatorname{horiz} \operatorname{dist}(z_1, f(\tilde{\zeta}_0)).& \end{aligned} \tag{23}$$

In the setting of Lemma 5.1, $\operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0)) \le C\varepsilon$ and so the length of the bound period is determined, essentially, by the effect of the random noise:

$$\|Df^{p_1}(f(\tilde{\zeta}_0))\|\varepsilon \approx e^{-\beta p_1}.$$

Using that the norm is between $\sigma_1^{cp_1}$ and 4^{p_1}, this gives

$$p_1 \approx \log\frac{1}{\varepsilon} \quad \text{and also} \quad \|Df^{p_1}(f(\tilde{\zeta}_0))\| \ge \varepsilon^{1-\theta}$$

where θ is close to zero if β is. The first relation is claim (a) in the lemma. Claim (c) follows from the second inequality and claim (b).

On the contrary, Lemma 5.2 corresponds to the case when the bound period is short enough so that the effect of random noise is negligible. In more precise terms, $\operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0)) \ge c\varepsilon$ and so the leading term in (23) is the last one. Hence,

$$\|Df^{p_1}(f(\tilde{\zeta}_0))\| \cdot \operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0)) \approx e^{-\beta p_1}. \tag{24}$$

Using the upper and lower bounds on the norm in the same way as before, we deduce claim (a):

$$p_1 \approx -\log\big(\operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0))\big) \approx r.$$

Moreover, using part (b) and (24), we get claim (c):

$$\|Dg_2^{p_1}(z_1)\| \approx \|Df^{p_1}(f(\tilde{\zeta}_0))\| \ge \big[\operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0))\big]^{-9/10} \approx e^{2r(9/10)}.$$

and claim (d):

$$\begin{aligned}\|Dg^{p_1+1}(z_0)(1,0)\|^2 &\approx \|Dg_2^{p_1}(z_1)(1,0)\|^2 \operatorname{dist}(z_0, \tilde{\zeta}_0)^2 \\ &\approx \|Df^{p_1}(f(\tilde{\zeta}_0))(1,0)\|^2 \operatorname{horiz\,dist}(z_1, f(\tilde{\zeta}_0)) \\ &\ge c\sigma_1^{p_1} e^{-\beta(p_1+1)} \ge \sigma_1^{2(p_1+1)/3},\end{aligned}$$

as long as β is sufficiently small. □

5.3. **Step 1: The capture argument.** In all cases, we define the next free return $n_1(z, \mathbf{g})$ as the first iterate $n_1 > p_1$ for which $\gamma_{n_1} = g_2^{n_1-1}(\gamma_1)$ intersects the domain $\{(x,y) : |x| < \delta\}$. We need to define a binding point for the random leaf γ_{n_1}. The key idea, contained in the following lemma, is that we may approximate γ_{n_1} by a free segment L of the unstable manifold $W^u(P)$ of the unperturbed map. Then we use the binding point of L, for the deterministic map f, as the binding point of the random leaf.

Lemma 5.3. *There exists a free segment $L = L(n_1, \gamma_{n_1})$ of the unstable manifold $W^u(P)$ that is $\varepsilon^{1-\theta_0}$-close to γ_{n_1} in the C^1 topology.*

Proof. It is assumed that ε is small with respect to all other constants involved in the arguments.

Case 1: The previous return is an ε-situation.

Define $q_0 \geq 1$ by $b^{q_0} \approx \varepsilon$. Let $\tau_1 = n_1 - q_0$. Assuming b is small,

$$q_0 \approx \frac{\log \varepsilon}{\log b} \ll \log \frac{1}{\varepsilon} \approx p_1 \qquad \text{and} \qquad 4^{10 q_0} \approx \varepsilon^{\frac{10 \log 4}{\log b}} \ll \varepsilon^{-\theta_0}.$$

We distinguish two sub-cases:

Case 1a: There are no returns during $[\tau_1, n_1)$.

Let $\gamma_{\tau_1} = g_2^{\tau_1 - 1}(\gamma_1)$. For each $0 \leq j \leq q_0$ and $z_{\tau_1} \in \gamma_{\tau_1}$,

$$\operatorname{dist}\left(f^j(z_{\tau_1}), g^j_{\tau_1+1}(z_{\tau_1})\right) \leq 4^{j-1}\varepsilon + \cdots + 4\varepsilon + \varepsilon \leq 4^j \varepsilon \leq \varepsilon^{1-\theta_0} \ll \delta.$$

So, by Proposition 2.5, every $z_{\tau_1} \in \gamma_{\tau_1}$ is expanding up to time q_0, for the unperturbed map f. Let $\Gamma = \Gamma(z_{\tau_1})$ be contracting leaves of order q_0, for the unperturbed map f, through the points $z_{\tau_1} \in \gamma_{\tau_1}$. We may suppose that γ_{τ_1} is far from the tips of the generation zero segment G_0 of $W^u(P)$. Indeed, replacing γ_{τ_1} by its its second iterate, if necessary, we guarantee that the distance to the tips is $> c\,(2 - a_2)$. Recall that we consider parameters in an interval $[a_1, a_2]$ with $a_2 < 2$, and note that replacing q_0 by $q_0 - 2$ is harmless for what follows. Then $\Gamma(z_{\tau_1})$ intersects G_0 at a point $\eta = \eta(z_{\tau_1})$. Let $L_0 = \{\eta(z_{\tau_1}) : \ z_{\tau_1} \in \gamma_{\tau_1}\} \subset G_0$ be the nearly horizontal segment captured in this way. See Figure 7.

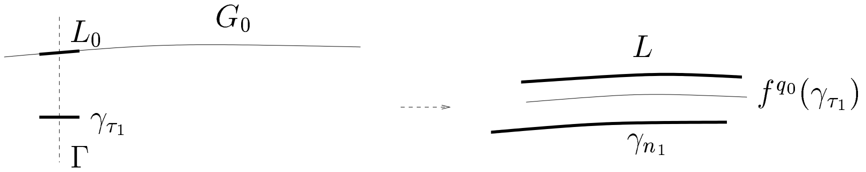

FIGURE 7.

We claim that $L = f^{q_0}(L_0)$ is a free segment of W^u. Indeed, from $\operatorname{dist}(f(\tilde{\zeta}_0), L_0) \geq c\,(2 - a_2)$ and Proposition 2.8 we get that the points of L_0 remain in a bound state during at most $c \log \frac{1}{2 - a_2}$ iterates, which is much less than q_0 if ε is small. Moreover, since L_0 has no returns in the first q_0 iterates, once it becomes free it remains free up to time q_0, as claimed. Finally, using Proposition 2.3,

$$\begin{aligned} \operatorname{dist}(L, \gamma_{n_1}) &\leq \operatorname{dist}\left(L, f^{q_0}(\gamma_{\tau_1})\right) + \operatorname{dist}\left(f^{q_0}(\gamma_{\tau_1}), \gamma_{n_1}\right) \\ &\leq (Cb)^{q_0} + 4^{q_0}\varepsilon \leq \varepsilon^{1-\theta_0} \end{aligned}$$

as stated in the lemma.

Case 1b: Suppose there are returns in $[\tau_1, n_1)$.

By the definition of n_1, necessarily $\tau_1 \le p_1$. Note that n_1 is a free iterate for ζ_0. So, by [7, Lemma 6.6], there is some favourable position $\tau_2 \in [n_1 - 3q_0, \tau_1)$. This means that for every $\tau_2 + j \in [\tau_2, n_1)$ the distance from $f^{\tau_2+j}(\zeta_0)$ to the critical set is at least λ_0^j, where $\lambda_0 = e^{-36}$ say. As a consequence, $f^{\tau_2}(\zeta_0)$ is expanding up to time $n_1 - \tau_2$. Since $3q_0 \ge n_1 - \tau_2 \ge n_1 - \tau_1 = q_0$, the previous calculation holds with τ_1 replaced by τ_2. We proceed in just the same way as before.

Case 2: The previous return is a deterministic situation.

Let $q_0 \ge 1$ and $\tau_1 = n_1 - q_0$ be defined as before. We distinguish three sub-cases:

Case 2a: Suppose there are no returns in $[\tau_1, n_1)$.

We proceed in just the same way as in Case 1a.

Case 2b: Suppose $\tau_1 \le p_1$ and $p_1 \ge 3q_0$.

This is analogous to Case 1b: take $\tau_2 \in [\mu_1 - 3q_0, \tau_1)$ to be a favourable position and replace γ_{τ_1} by $\gamma_{\tau_2} = g_2^{\tau_2 - 1}(\gamma_1)$. Then proceed as before.

Case 2c: Suppose $\tau_1 \le p_1 \le 3q_0$.

The hypothesis implies $n_1 \le 4q_0$. By Lemma 5.2, every point $z_1 \in \gamma_1$ is expanding up to time p_1. Since there are no returns in (p_1, n_1), z_1 is also expanding up to time $n_1 - 1$. Take $L_0 = f(\gamma_0)$ and $L = f^{n_1-1}(L_0)$, where γ_0 is the unstable segment introduces in Section 3.3. Using $n_1 \le 4q_0$ we get that

$$\operatorname{dist}\left(g^{n_1-1}(\gamma_1), f^{n_1-1}(L_0)\right) \le (Cb)^{q_0} + 4^{4q_0}\varepsilon \le \varepsilon^{1-\theta_0},$$

in just the same way as before.

Notice that in all the cases we have obtained a C^0 bound for the distance between L and γ_{n_1}. In order to get a C^1 it is enough to combine, through Hadamard's lemma, this C^0 bound with the fact that the two curves have uniformly bounded C^2 norm. The latter is contained in Lemma 5.4 below. $\square$

5.4. **Step 1: Binding point.** By Proposition 2.7, there exists a critical point $\tilde\zeta_1$ such that L is in tangential position with respect to $\tilde\zeta_1$. By definition this is the binding point of γ_{n_1}.

Recall that $s(\varepsilon)$ is defined by $e^{-2s(\varepsilon)} \approx \varepsilon$. Let $\Gamma_{s(\varepsilon)}$ be the corresponding long stable leaf for the critical point $\tilde\zeta_1$. The next lemma says that points for which n_1 is not an ε-situation are in tangential position relative to the binding point.

Lemma 5.4. *If $z_{n_1+1} = g_{n_1+1}(z_{n_1})$ is to the left of the long stable leaf $\Gamma_{s(\varepsilon)}$ then $(z_{n_1}, t(z_{n_1}))$ is in tangential position relative to $\tilde\zeta_1$.*

Proof. The divide the argument into three cases, depending on the situation in the definition of the itinerary. The first step is to show that γ_{n_1} is a nearly horizontal curve. This is clear in case (2b), because the curve γ_1 is already fairly horizontal, and there are no returns in the time interval $[1, n_1)$. In case (2a), the bound period corresponds to the one of the unperturbed dynamics, and so the argument is just the same as in Lemma 3.3 above, which is Lemma 3.6 from [8]. Finally, a similar argument applies also in case (1), observing that the curve γ_1, a lamina of $\Lambda_{\varepsilon,z_0}$, is already nearly horizontal.

To conclude the proof observe that if z_{n_1+1} is as in the statement then

$$\operatorname{dist}(z_{n_1}, \tilde{\zeta}_1) \geq c\varepsilon^{1/2} \gg \varepsilon^{1-\theta_0} \geq \operatorname{dist}(L, \gamma_{n_1}).$$

Since L is in tangential position to $\tilde{\zeta}_1$, the claim follows. □

This lemma is crucial for what follows: it ensures that, in the absence of ε-situations, the same estimates as in the unperturbed case remain true for these random iterations, only with slightly worse constants.

5.5. **Step 1: Conclusion.** Now we are in a position to define $i_1(z, \mathbf{g})$. Fix $i_0 = i_0(z, \mathbf{g})$ and let $n_1 = n_1(z, \mathbf{g})$ be as above. Let $\tilde{\zeta}_1$ be the binding point of γ_{n_1}, as defined above, and $\{\Gamma_{r,l}\}$ be the sequence of long stable leaves associated to $\tilde{\zeta}_1$. Recall that Γ_Δ is independent of the critical point, and that

Essentially, we define

1. If z_{n_1+1} is to the right of $\Gamma_{s(\varepsilon)}$ then $i_1(z, \mathbf{g}) = (\tilde{\zeta}_1, r_1, 0, 0, \mathbf{y}_1)$ where $\mathbf{y}_1$ is the lamina of $\Lambda_{\varepsilon, z_{n_1}}$ that contains z_{n_1+1}.
2. If z_{n_1+1} is to the left of $\Gamma_{s(\varepsilon)}$:
   - (a) $i_1(z, \mathbf{g}) = (\tilde{\zeta}_1, r, l, \pm, 0)$ if z_{n_1+1} is in the region of Δ_0 in between $\Gamma_{r,l}(\tilde{\zeta}_1)$ and $\Gamma_{r,l-1}(\tilde{\zeta}_1)$, with $(r, l) > (\Delta, 0)$ the sign $+/-$ corresponding to whether z_{n_1} is to the right or to the left of the critical point;
   - (b) $i_1(z, \mathbf{g}) = (\tilde{\zeta}_1, 0, 0, \pm, 0)$ if z_{n_1+1} is between Γ_Δ and $W^s_{loc}(P)$, the sign $+/-$ corresponding to whether z_{n-1} is to the right or to the left of the critical point.

However, we adjoin segments that do not fully cross from $\Gamma_{r,l}$ to $\Gamma_{r,l+1}$ to their adjacent curve segment(s). If γ_{n_1+1} crosses at most one of the long stable curves $\Gamma_{r,l}$ we say that n_1 is an *inessential situation*, otherwise we call it an *essential situation*.

The curves L as in Lemma 5.3 are called *shadowing leaves*.

5.6. **General step.** The general step of the definition of itineraries for the random process is entirely analogous to Step 1 that we have just described. Suppose that $n_s(z, \mathbf{g})$ and $i_s(z, \mathbf{g})$ have been defined for $0 \le s \le k$. Consider the random curves γ_{n_k+1} located near the critical value of either of the following types:

- if $i_k(z, \mathbf{g})$ corresponds to an ε-situation this curve is a lamina of $\Lambda_{\varepsilon, z_{n_k}}$;
- otherwise $\gamma_{n_k+1} = g_{n_k+1}(\gamma_{n_k})$.

Assume the capture construction has been carried out, and a binding point $\tilde{\zeta}_k$ has been defined as explained above. The bound period p_{k+1} to the binding point $\tilde{\zeta}_k$ is defined in the same way as before. Using Lemma 5.4, at time n_k, we get the analogs of Lemmas 5.1 and 5.2 for p_{k+1}. Then we let n_{k+1} be the first return after time $n_k + p_{k+1}$. Finally, we prove the statement corresponding to Lemma 5.4 at time n_{k+1}, by the same arguments as for $k = 0$.

5.7. **The large deviations argument.** By analogy to the unperturbed case, we say that a free return n_k is a *random escape situation* for $(z, \mathbf{g})$ if the corresponding random leaf γ_{n_k+1} stretches across $\mathcal{R}_\Delta$ extending at least $\delta/10$ to either side. We need to show that long escaping times are exponentially unlikely also in the random setting. For that we must reproduce the basic large deviations estimate (cf. Lemma 4.1)

$$\operatorname{Prob}(r_1, \dots, r_l) \le C^l \, e^{-c(r_1 + \cdots + r_l)}$$

for every choice of r_j and uniform constants C and $c > 0$. More precisely, we need the proposition that is stated next.

For $k \ge 0$, let $\eta_k = \gamma_{n_k+1}$ be a random curve close to the critical value as constructed while defining itineraries: either η_k is a lamina of $\Lambda_{\varepsilon, z_{n_k}}$ or $\eta_k = g_{n_k+1}(\gamma_{n_k})$ where γ_{n_k} is a $C^2(b)$ curve corresponding to fixed values of $i_s(z, \mathbf{g}) = (\tilde{\zeta}_k, r_k, l_k, \epsilon_k, 0)$ for all $0 \le s \le k$. Let $\hat{m}_0$ be normalized arc-length measure on η_k.

Proposition 5.5. *Let $\mathcal{P}(\rho_1, \dots, \rho_k; \eta_k)$ be the total $\hat{m}_0 \times \nu_\varepsilon^{\mathbb{N}}$-probability of the set of pairs $(z, \mathbf{g})$ with $g^{n_k+1}(z) \in \eta_k$ and such that $n_{k+j}(z, \mathbf{g})$ is a deterministic situation and $i_{k+j}(z, \mathbf{g}) = (\cdot, \rho_j, \cdot, \cdot, \cdot)$ for every $1 \le j \le l$. Then*

$$\mathcal{P}(\rho_1, \dots, \rho_l; \eta_k) \le \frac{C^l}{\operatorname{length}(\eta_k)} \, e^{-c(\rho_1 + \cdots + \rho_l)}$$

for all $\rho_1, \dots, \rho_l \in \mathbb{N}$. If η_k corresponds to an ε-situation, we may replace $\operatorname{length}(\eta_k)$ *by* $\operatorname{length}(\eta_k)\varepsilon^{-9/10}$ *in the denominator.*

Proof. A similar estimate was obtained before in [7, Section 2.2], for the deterministic case $g_i = f$. The proof there carries on to the present context, up to straightforward adaptations, because the time interval we deal with here contains no ε-situations, and so all returns are governed by Lemma 5.4. The last statement in the proposition follows from the same arguments, only starting at time $n_k + p_{k+1} + 1$: recall that during the bound period the random curve is expanded $\geq \varepsilon^{-9/10}$ if η_k corresponds to an ε-situation, by Lemma 5.1(c).

In particular, one has the following bounded distortion result, which is also of independent interest:

Lemma 5.6. *Let η_k be a random leaf, as introduced before. Suppose $\xi_1, \xi_2 \in \eta_k$ share the same itinerary up to time $n_k + 1 + n$. Then*

$$\|Dg^n_{n_k+1}(\xi)t(\xi_1)\| \leq C\|Dg^n_{n_k+1}(\xi)t(\xi_2)\|,$$

where $t(\,\cdot\,)$ denotes a norm 1 tangent vector to the random leaf.

The proof is again analogous to the deterministic case, see for instance Lemma 10.5 in [22]. □

Let $e_k(z, g)$ be the *escaping time* of a pair $(z, \mathbf{g})$ with $g^{n_k+1}(z) \in \eta_k$:

$$e_k(z, \mathbf{g}) = n_l(z, \mathbf{g}) - n_k(z, \mathbf{g})$$

where $l > k$ is minimum such that n_l is an escape situation for $(z, \mathbf{g})$. As a consequence of Proposition 5.5, one obtains the desired exponential estimate on the probability of large escaping times:

Corollary 5.7. *Let $\mathcal{P}(m; \eta_k)$ denote the $\hat{m}_0 \times \nu_\varepsilon^{\mathbb{N}}$-probability of the set of pairs $(z, \mathbf{g})$ with $g^{n_k+1}(z) \in \eta_k$ and such that $e_k(z, \mathbf{g}) > m$ and there are no ε-situations in $[n_k + 1, n_k + m]$. Then,*

$$\mathcal{P}(m; \eta_k) \leq \frac{C}{\operatorname{length}(\eta_k)} e^{-cm}$$

for all m. If η_k corresponds to an ε-situation we also have

$$\mathcal{P}(m; \eta_k) \leq \frac{C}{\operatorname{length}(\eta_k)\varepsilon^{-9/10}} e^{-c(m-p_{k+1})}.$$

Proof. This is completely analogous to the corresponding deterministic statement and so we may use the same proof. See [7, Section 2.2]. For the last statement, it suffices to take $n_k + p_{k+1} + 1$ as the starting time. □

When n_k is an escape situation the length of η_k is uniformly bounded from below, and the estimate given by Corollary 5.7 is analogous to the one one gets in the usual argument for the unperturbed map. The case when n_k is an ε-situation is more delicate, because η_k, a lamina of $\Lambda_{\varepsilon, z_{n_k}}$ may have arbitrarily small length. Assumption (H1) allows us to overcome this difficulty: small laminae have small total probability.

Corollary 5.8. *Suppose η_k corresponds to an ε-situation, and symbols $i_0, i_1, \ldots, i_k$. Then*

$$\mathcal{P}(m \mid i_0, \ldots, i_{k-1}, \tilde{\zeta}_k) \le C\varepsilon^{-1/10} e^{-c(m-p_{k+1})},$$

where the left hand side is the probability of $e_k(z, \mathbf{g}) > m$ and no ε-situations in the first m-iterates, conditioned to $i_0, i_1, \ldots, i_{k-1}$, and $\tilde{\zeta}_k$ (but not $\mathbf{y}_k$).

Proof. By Corollary 5.7 and hypothesis (H2),

$$\mathcal{P}(m \mid i_0, \ldots, i_{k-1}, \tilde{\zeta}_k) = \int \mathcal{P}(\mathbf{y}; m)\, d\mathbf{y} \le \int \frac{C\varepsilon^{9/10}}{\operatorname{length}(\mathbf{y})} e^{-c(m-p_{k+1})}\, d\mathbf{y},$$

where the integrals are over all laminae of $\Lambda_{\varepsilon, z_{n_k}}$. Hypothesis (H1) says that the set of laminae with length less than εe^{-s} has conditional probability less than $K e^{-(1+\kappa)s}$ for every $s \ge 0$. It follows that

$$\mathcal{P}(m \mid i_0, \ldots, i_{k-1}, \tilde{\zeta}_k) \le \sum_{s=0}^{\infty} C e^{s} \varepsilon^{-1/10} e^{-(1+\kappa)s} e^{-c(m-p_{k+1})}$$

and this is $\le C\varepsilon^{-1/10} e^{-c(m-p_{k+1})}$ because $\kappa > 0$. $\square$

6. Proof of the main theorem

6.1. **Uniqueness of the stationary measure.** The basin $B(\Lambda)$ is a neighbourhood of the Hénon-like attractor Λ; see [8, Section 5]. Thus, assuming ε is sufficiently small, all stationary measures obtained as accumulation points of

$$\frac{1}{n} \sum_{j=0}^{n-1} \mathcal{T}_\varepsilon^{j} \eta$$

are supported inside $B(\Lambda)$, for any measure η supported in Λ.

Lemma 6.1. *Under assumption (H3), the Markov chain has a unique stationary measure μ_ε supported inside the basin $B(\Lambda)$, and this measure is ergodic.*

Proof. In view of ergodic decomposition, cf. Section 1.2, we only have to prove that there exists a unique ergodic stationary measure.

Let μ_ε be any such measure, and $G(\mu_\varepsilon)$ be the set of points $z \in B(\Lambda)$ such that almost every random orbit $\mathbf{z}$ starting at z satisfies (5) for every continuous function φ. Hypothesis (H3) implies that the ball of radius $\rho(\varepsilon)$ around $f(w)$ is contained in $G(\mu_\varepsilon)$, for any $w \in G(\mu_\varepsilon)$. We have shown in [8, Section 5] that the stable manifold of the fixed point P is dense in the basin $B(\Lambda)$. It follows that $W^s(P)$ intersects the interior of $G(\mu_\varepsilon)$ in some point z. By the previous argument, $B_{\rho(\varepsilon)}(f^n(z))$ is contained in $G(\mu_\varepsilon)$ for every $n \geq 1$. Of course, $f^n(z)$ converges to P as $n \to \infty$. It follows that P is in the interior of $G(\mu_\varepsilon)$.

Now let ν_ε be any other ergodic stationary measure. By the previous paragraph, the intersection of $G(\mu_\varepsilon)$ and $G(\nu_\varepsilon)$ contains a neighbourhood of P. By (5), we have $\int \varphi\, d\mu_\varepsilon = \tilde{\varphi}(\mathbf{z}) = \int \varphi\, d\nu_\varepsilon$ for almost every random orbit $\mathbf{z}$ starting in this intersection, and every continuous function φ. This proves that $\nu_\varepsilon = \mu_\varepsilon$. □

6.2. **An upper bound for stationary measures.** This section contains the main estimate, Proposition 6.3, from which we shall deduce the statement of stochastic stability. The following terminology will be useful:

Definition 6.2. Given two Borel measures α and β on a manifold M, and a positive functional $r : C_0^1(M) \to \mathbb{R}$ on the space of C^1 functions with compact support, we write

$$\alpha \leq \beta + r(\,\cdot\,)$$

to mean that there is a measure $\tilde{\beta} \leq \beta$ such that

$$\left| \int \varphi d\alpha - \int \varphi d\tilde{\beta} \right| \leq r(\varphi) \quad \text{for all } \varphi \in C_0^1(M).$$

Let m_0 be the arc-length measure on the curve segment $\gamma_0 = f^{-1}(\Omega)$, normalized so as to be a probability measure. For every $\varepsilon > 0$ and $n \geq 1$, let

$$\mu_{\varepsilon,n} = \frac{1}{n} \sum_{j=1}^{n} \mathcal{T}_\varepsilon^j m_0 = \frac{1}{n} \sum_{j=1}^{n} \pi_{1*} \mathcal{F}_{\varepsilon*}^j (m_0 \times \nu_\varepsilon^{\mathbb{N}}). \tag{25}$$

For simplicity, we write $\mathcal{P}_\varepsilon = m_0 \times \nu_\varepsilon^{\mathbb{N}}$, and $\mathcal{T}_\varepsilon^j$ to mean $\pi_{1*}\mathcal{F}_{\varepsilon*}^j$ (a slight abuse of language).

Proposition 6.3. *There exist constants $C > 0$, $c > 0$, and*

(a) *measures $\{\lambda_{\varepsilon,n} : \varepsilon > 0, n \in \mathbb{N}\}$ on $\tilde{X}$, absolutely continuous on unstable leaves, with density and total mass bounded by C;*

(b) *measures $\{M_{\varepsilon,n,N} : \varepsilon > 0,\ N \in \mathbb{N},\ n \in \mathbb{N}\}$ on the attractor Λ, with total mass $\|M_{\varepsilon,n,N}\| \leq C \exp(-cN)$ for every ε, N, n;*

(c) *positive functionals $R_{\varepsilon,N} : C_0^1(\mathbb{R}^2) \to \mathbb{R}$ converging to zero pointwise when $\varepsilon \to 0$, for every fixed $N \in \mathbb{N}$;*

such that, for every $\varepsilon > 0$ and $n, N \in \mathbb{N}$,

$$\mu_{\varepsilon,n} \leq \sum_{s=0}^{\infty} f_*^s p_* (\lambda_{\varepsilon,n} \mid \{e(\,\cdot\,) > s\}) + M_{\varepsilon,n,N} + R_{\varepsilon,N}(\,\cdot\,)$$

Here $e(\,\cdot\,)$ denotes deterministic escape time, as defined in Section 4.1. In the proof we use the partitions $\mathcal{I}_n$ of $\gamma_0 \times \Omega_\varepsilon^{\mathbb{N}}$ defined by

- Points $(z, \mathbf{g})$ in each element of $\mathcal{I}_n$ have the same random itinerary up to time n.
- The sequence g_i is also prescribed up to time n, except for the map g_τ at the last ε-situation $\tau \le n$; the latter is arbitrary, if it is not for the fact that the corresponding symbol i_k is fixed.

Another useful sequence of partitions $\mathcal{J}_n$ is defined as follows:

- each element J of $\mathcal{J}_n$ is the union of all $I \in \mathcal{I}_n$ sharing the same last ε-situation τ, the same sequence of maps g_i for $i \neq \tau$, $i \le n$, and the same itinerary up to time τ.

Observe that this union is finite: each $I \in \mathcal{I}_n$ contained in $J \in \mathcal{J}_n$ is described by an itinerary in the time interval from $\tau + 1$ to n and, in the absence of ε-situations, there are only finitely many possible itineraries. We write $\gamma(i, I) = \pi_1 \mathcal{F}_\varepsilon^i(I)$ and $\gamma(i, J) = \pi_1 \mathcal{F}_\varepsilon^i(J)$ for each $I \in \mathcal{I}_n$, $J \in \mathcal{J}_n$, and $i \le n$. See Figure 8.

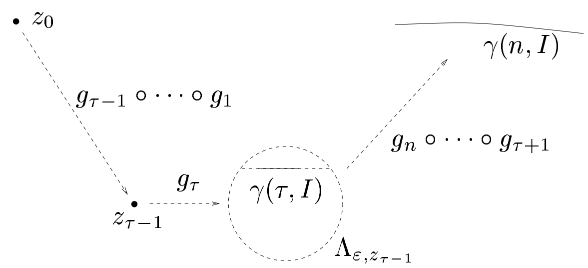

FIGURE 8.

These partitions are designed so that each $\gamma(\tau, J)$ coincides with a lamina $\mathbf{y}$ of $\Lambda_{\varepsilon, z_{\tau-1}}$. Moreover, the iterate $\mathcal{T}_\varepsilon^\tau(\mathcal{P}_\varepsilon \mid J)$ coincides with the conditional probability $\psi_{\mathbf{y}} m_{\mathbf{y}}$ of $p_\varepsilon(\cdot \mid z_{\tau-1})$ along $\mathbf{y}$. If J is the element of $\mathcal{J}_n$ that contains I,

$$\gamma(n, I) = (g_n \circ \cdots \circ g_{\tau+1}) \gamma(\tau, I),$$

with $\gamma(\tau, I)$ a sub-segment of $\mathbf{y} = \gamma(\tau, J)$. We represent by $(\psi_{\mathbf{y}} m_{\mathbf{y}})_I$ the *restriction* of $\psi_{\mathbf{y}} m_{\mathbf{y}}$ to $\gamma(\tau, I)$. We call *relative weight* of I with respect to J the quantity

$$\mathrm{P}_\varepsilon(I \mid J) = \frac{\operatorname{length}(\gamma(\tau, I))}{\operatorname{length}(\gamma(\tau, J))}.$$

Finally, let $B_{\varepsilon,n}$ be the quotient measure of $\mathcal{P}_\varepsilon$ relative to $\mathcal{J}_n$.

Proof of Proposition 6.3. We split (25) along the partition $\mathcal{J}_n$:

$$\begin{aligned}\mu_{\varepsilon,n} &= \frac{1}{n}\sum_{j=1}^{n}\int_{\mathcal{J}_j}\mathcal{T}_\varepsilon^j(\mathcal{P}_\varepsilon \mid J)\,dB_{\varepsilon,j}(J)\\ &= \frac{1}{n}\sum_{j=1}^{n}\sum_{\tau=0}^{j}\int_{\mathcal{J}_{j,\tau}}\mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})\,dB_{\varepsilon,j}(J),\end{aligned}$$

where $\mathcal{J}_{j,\tau}$ is the subset of $\mathcal{J}_j$ for which τ is the last ε-situation. So,

$$\mu_{\varepsilon,n} = \frac{1}{n}\sum_{j=1}^{n}\sum_{\tau=0}^{j}\int_{\mathcal{J}_{j,\tau}}\sum_{I\subset J}\mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I\,dB_{\varepsilon,j}(J), \tag{26}$$

where the last sum is over all $I \in \mathcal{I}_j$ contained in J. We denote by $\mu_{\varepsilon,n}^0$ the expression obtained restricting the sum to those $I \subset J$ having no escape situations $\nu \in [\tau, j]$, and by $\mu_{\varepsilon,n}^e$ the expression obtained restricting the sum to the terms for which such a ν does exist. Thus

$$\mu_{\varepsilon,n} = \mu_{\varepsilon,n}^0 + \mu_{\varepsilon,n}^e. \tag{27}$$

We are going to derive appropriate bounds for each of the two terms.

Our bound on the first term μ_ε^0 is given by the following

Lemma 6.4. *The total mass of the measure $\mu_{\varepsilon,n}^0$ is less than $C\varepsilon^{1/10}$, for all small $\varepsilon > 0$.*

We shall deduce this result from two auxiliary sublemmas.

Sublemma. *$\mathcal{P}_\varepsilon(\mathcal{E}(\tau)) \le C\varepsilon^{1/5}$ for any $\tau \ge 0$, where $\mathcal{E}(\tau)$ denotes the set of pairs $(z, \mathbf{g})$ for which τ is an ε-situation.*

Proof. For $\tau = 0$ just note that the curve γ_0 is long and, for any g_1, only a fraction $\le C\varepsilon^{1/2} \ll \varepsilon^{1/5}$ of it is mapped to the right of $\Gamma_{s(\varepsilon)}$. Now suppose $\tau = n_{k+1}$ for some $k \ge 0$. We distinguish three cases, according to the nature of the previous free return n_k.

If n_k is a deterministic situation, then $\text{length}(\gamma_{n_k}) \ge ce^{-r_k} \ge c\varepsilon^{1/2}$. Using the expansion during the bound period granted by Lemma 5.2(c), as well as the expansion during the subsequent free period, we conclude that

$$\text{length}(\gamma_{n_{k+1}}) \ge ce^{-2r_k}e^{2r_k(9/10)} \ge ce^{-2r_k/10} \ge c\varepsilon^{1/10}.$$

Since only a sub-segment of length $\le C\varepsilon^{1/2}$ can be mapped to the right of $\Gamma_{s(\varepsilon)}$, we get that ε-situation has conditional probability $\le C\varepsilon^{1/2-1/10} < \varepsilon^{1/5}$.

Now let n_k be an ε-situation with length$(\mathbf{y}_k) \geq \varepsilon^{6/5}$. This is similar to the previous case. Indeed, using the expansion in Lemma 5.1(c) we conclude that

$$\text{length}(\gamma_{n_{k+1}}) \geq c\varepsilon^{6/5}\varepsilon^{-9/10} \geq c\varepsilon^{3/10},$$

so that the fraction to the right of $\Gamma_{s(\varepsilon)}$ is less than $C\varepsilon^{1/2-3/10} \leq C\varepsilon^{1/5}$.

Finally, suppose that n_k is an ε-situation with length$(\mathbf{y}_k) < \varepsilon^{6/5}$. By hypotheses (H1), (H2) this possibility has conditional probability $\leq C\varepsilon^{(1+\kappa)/5}$, conditioned to any given itinerary prior to time n_k. Thus this case contributes a total probability $\leq C\varepsilon^{(1+\kappa)/5} < \varepsilon^{1/5}$. □

Let $p(\varepsilon) = C\log(1/\varepsilon)$ be the upper bound, given by Lemma 5.1(a), for the duration of the bound period following an ε-situation.

Sublemma. *We have $\mathcal{P}_\varepsilon(\mathcal{E}(\tau, m)) \leq C\varepsilon^{-1/10}e^{-c(m-p(\varepsilon))}\mathcal{P}_\varepsilon(\mathcal{E}(\tau))$ for all $m \geq 1$, where $\mathcal{E}(\tau, m)$ denotes the set of pairs $(z, \mathbf{g})$ for which τ is an ε-situation and there are neither escape situations nor ε-situations in the time interval $[\tau+1, \tau+m]$.*

Proof. Fix the itinerary and the sequence of maps g_i for all times $< \tau$, and fix also the binding point at time τ. By Corollary 5.8, the conditional probability of having neither ε-situations nor escape situations in the first m iterates is less than $C\varepsilon^{-1/10}e^{-c(m-p(\varepsilon))}$. Integrating over all choices of the itinerary and binding point, we get the statement. □

Proof of Lemma 6.4. Let $1 \leq \tau \leq j$ be fixed. By definition, every $I \subset J$, $J \in \mathcal{J}_{j,\tau}$ is contained in $\mathcal{E}(\tau, j-\tau)$. Thus, the mass of the measure

$$\int_{\mathcal{J}_{j,\tau}} \sum_{I \subset J} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \, dB_{\varepsilon,j}(J)$$

is bounded by the probability of $\mathcal{E}(j, j-\tau)$. Thus, using the last sublemma above, the total mass of $\mu^0_{\varepsilon,n}$ is bounded by

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n}\sum_{\tau=0}^{j} \mathcal{P}_\varepsilon\left(\mathcal{E}(\tau, j-\tau)\right) \\
&\qquad \leq \frac{1}{n}\sum_{j=1}^{n}\big[p(\varepsilon) + \sum_{\tau=0}^{j-p(\varepsilon)} C\varepsilon^{-1/10}e^{-c(j-\tau-p(\varepsilon))}\big]\mathcal{P}_\varepsilon\left(\mathcal{E}(\tau)\right).
\end{aligned}$$

Note that for the first $p(\varepsilon)$ terms we used $\mathcal{P}_\varepsilon\left(\mathcal{E}(\tau, j-\tau)\right) \leq 1$. From the first of the sublemmas above, we find that the right hand side is less than

$$\frac{1}{n}\sum_{j=1}^{n}\big[C\log\frac{1}{\varepsilon} + \sum_{i=0}^{\infty} C\varepsilon^{-1/10}e^{-ci}\big]\varepsilon^{1/5} \leq C\varepsilon^{1/10},$$

as claimed in the lemma. $\square$

Now we proceed to bound $\mu^e_{\varepsilon,n}$. For this purpose, we split the sum in (26) according to the value ν of the last escape situation:

$$\mu^e_{\varepsilon,n} = \frac{1}{n}\sum_{j=1}^{n}\sum_{\nu=1}^{j}\sum_{\tau=0}^{\nu}\int_{\mathcal{J}_{j,\tau}}\sum_{I\subset J(\nu)} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \, dB_{\varepsilon,j}(J). \tag{28}$$

First we deal with the terms for which $j=\nu$. In this case $\gamma(\nu,I)$ is an escaping random leaf, cf. Section 5. The capture construction in Lemma 5.3 provides an escaping leaf $L(\nu,I)$ of the unperturbed map f, close to $\gamma(\nu,I)$ in the C^1 sense. Let $m_{L(\nu,I)}$ be the arc-length measure along $L(\nu,I)$.

Lemma 6.5. *Let $j=\nu$. There are $C>0$, independent of ε, τ, ν, I, J, and positive functionals $r_\varepsilon(\,\cdot\,)$ independent of τ, ν, I, J, such that*

$$\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_y m_y)_I \le C\,\mathrm{P}_\varepsilon(I \mid J)\big(m_{L(\nu,I)} + r_\varepsilon(\,\cdot\,)\big).$$

and $\lim\limits_{\varepsilon\to 0} r_\varepsilon(\,\cdot\,)=0$.

Proof. This is a consequence of hypothesis (H2), the distortion control provided by Lemma 5.6, and the capture procedure. Indeed,

$$\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I = (g_\nu\circ\cdots g_{\tau+1})_*(\psi_{\mathbf{y}} m_{\mathbf{y}})_I\,.$$

By (H2), the density $\psi_{\mathbf{y}}$ is bounded by $K/\operatorname{length}(\mathbf{y})$. The distortion lemma implies that the derivative of $g_\nu\circ\cdots g_{\tau+1}$ along $\gamma(\tau,I)$ is comparable, up to a bounded factor, to

$$\frac{\operatorname{length}(\gamma(\nu,I))}{\operatorname{length}(\gamma(\tau,I))}$$

at every point of $\gamma(\tau,I)$. It follows that the measure $\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I$ is absolutely continuous with respect to arc length $m_{\gamma(\nu,I)}$ along $\gamma(\nu,I)$, with density bounded by

$$C\frac{K}{\operatorname{length}(\mathbf{y})}\frac{\operatorname{length}(\gamma(\tau,I))}{\operatorname{length}(\gamma(\nu,I))} \le C\frac{\operatorname{length}(\gamma(\tau,I))}{\operatorname{length}(\mathbf{y})} = C\,\mathrm{P}_\varepsilon(I\mid J). \tag{29}$$

The inequality uses the fact that $\operatorname{length}(\gamma(\nu,I))$ is uniformly bounded from below, because ν is an escape situation.

The capture construction gives that $\gamma(\nu,I)$ is C^1-close to $L(\nu,I)$, with a bound on the distance that goes uniformly to zero when ε goes to zero. This implies

$$m_{\gamma(\nu,I)} \le m_{L(\nu,I)} + r_\varepsilon(\,\cdot\,) \tag{30}$$

for some positive functional $r_\varepsilon(\,\cdot\,)$, depending only on the C^1 distance between the two curves. In particular, $\lim_{\varepsilon\to 0} r_\varepsilon(\,\cdot\,) = 0$. The lemma follows from (29)-(30). □

Now we consider $j = \nu + s$ for $s \geq 1$. Let $\tilde{I}$ represent any of the subsets of J obtained by further restricting the itinerary up to time ν, and for which ν is an escape situation. By definition, the $\tilde{I}$ are pairwise disjoint, and every $I \in \mathcal{I}_j$ such that ν is the last escape situation before j is contained in some $\tilde{I}$. The weight of $\tilde{I}$ relative to J is

$$\mathrm{P}_\varepsilon(\tilde{I} \mid J) = \frac{\mathrm{length}(\gamma(\tau, \tilde{I}))}{\mathrm{length}(\gamma(\tau, J))}.$$

Let $L(\nu, \tilde{I})$ be the deterministic leaf assigned to the random escaping leaf $\gamma(\nu, \tilde{I})$ by Lemma 5.3.

Lemma 6.6. *Let $j = \nu + s$. There is $C > 0$, independent of ε, τ, ν, s, $\tilde{I}$, J, and there are positive functionals $r_{\varepsilon,s}(\,\cdot\,)$, independent of τ, ν, $\tilde{I}$, J, such that the sum over $I \subset \tilde{I}$*

$$\sum_{I \subset \tilde{I}} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \leq C\,\mathrm{P}_\varepsilon(\tilde{I} \mid J)\big(f_*^s(m_{L(\nu,\tilde{I})} \mid \{e(\,\cdot\,) > s\}) + r_{\varepsilon,s}(\,\cdot\,)\big)$$

and $\lim_{\varepsilon\to 0} r_{\varepsilon,s}(\,\cdot\,) = 0$ for each fixed s.

Proof. The first step of the proof is an estimate at time ν, similar to the proof of Lemma 6.5. Exactly the same arguments as in the proof of (29), with $\tilde{I}$ in the place of I, give that the measure

$$\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_{\tilde{I}} = (g_\nu \circ \cdots \circ g_{\tau+1})_*(\psi_{\mathbf{y}} m_{\mathbf{y}})_{\tilde{I}}$$

is absolutely continuous with respect to arc-length $m_{\gamma(\nu,\tilde{I})}$ along $\gamma(\nu, \tilde{I})$, with density bounded by $C\,\mathrm{P}_\varepsilon(\tilde{I} \mid J)$. Thus, the sum

$$\sum_{I \subset \tilde{I}} \mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I$$

is bounded by the restriction of the measure $C\,\mathrm{P}_\varepsilon(\tilde{I} \mid J)\, m_{\gamma(\nu,\tilde{I})}$ to the union of the $\gamma(\nu, I)$ over all $I \subset \tilde{I}$. Since we are dealing only with partition elements I for which ν is the last escape situation before time $j = \nu + s$, this union is contained in the subset $\{e_\varepsilon(\,\cdot\,) > s\}$ of points in $\gamma(\nu, \tilde{I})$ whose random escaping time is larger than s. Summarizing, we have shown that

$$\sum_{I \subset \tilde{I}} \mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \leq C\,\mathrm{P}_\varepsilon(\tilde{I} \mid J)\big(m_{\gamma(\nu,\tilde{I})} \mid \{e_\varepsilon(\,\cdot\,) > s\}\big) \tag{31}$$

By Lemma 5.3, the curve $\gamma(\nu,\tilde I)$ is C^1-close to the shadowing leaf $L(\nu,\tilde I)$, with an upper bound on the distance that depends only on ε and goes to zero when ε goes to zero. Moreover, for fixed s, the subset $\{e_\varepsilon(\cdot)>s\}$ of $\gamma(\nu,\tilde I)$ is uniformly close to the subset $\{e(\cdot)>s\}$ of L if ε is small. This gives

$$(m_{\gamma(\nu,\tilde I)} \mid \{e_\varepsilon(\cdot)>s\}) \le (m_{L(\nu,\tilde I)} \mid \{e(\cdot)>s\}) + \rho'_{\varepsilon,s}(\cdot)$$

for a positive functional $\rho'_\varepsilon(\cdot)$ that goes to zero when ε goes to zero. Combining these two inequalities,

$$\sum_{I\subset\tilde I}\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I \le C\,\mathrm{P}_\varepsilon(\tilde I\mid J)\big(m_{L(\nu,\tilde I)}\mid\{e(\cdot)>s\}\big)+\rho'_{\varepsilon,s}(\cdot). \tag{32}$$

The last step in the proof is a continuity argument, to go from time ν to time $j=\nu+s$. Observe that the expression in the statement

$$\sum_{I\subset\tilde I}\mathcal{T}_\varepsilon^{j-E}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I=(g_j\circ\cdots\circ g_{\nu+1})_*\sum_{I\subset\tilde I}\mathcal{T}_\varepsilon^{\nu-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I\,. \tag{33}$$

Since all the maps g_i are ε-close to f in the C^1 topology,

$$\begin{aligned}(g_j\circ\cdots\circ g_{\nu+1})_*&(m_{L(\nu,\tilde I)}\mid\{e(\cdot)>s\})\\ &\le f_*^s(m_{L(\nu,\tilde I)}\mid\{e(\cdot)>s\})+\rho''_{\varepsilon,s}(\cdot)\end{aligned} \tag{34}$$

for some choice of a positive functional $\rho''_{\varepsilon,s}(\cdot)$ depending only on ε and s, and which goes to zero when ε goes to zero.

From (32)-(34) we immediately get the conclusion of the lemma. □

We are going to use Lemma 6.6 to estimate the terms in (28) for which $s=j-\nu$ is not too large. For the other terms we shall use the next lemma instead.

Lemma 6.7. *Let $j=\nu+s$. There are $C>0$ and $c>0$, independent of ε, τ, ν, s, J, such that the total mass of the measure*

$$\sum_{I\subset J(\nu)}\mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I$$

is bounded by Ce^{-cs}.

Proof. At this point this is a consequence of the fact that escape times have exponential tail, cf. Corollary 5.7. Indeed, (31) implies that the total mass of

$$\sum_{I\subset\tilde I}\mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}}m_{\mathbf{y}})_I$$

is bounded by

$$C\,\mathrm{P}_\varepsilon(\tilde I\mid J)m_{\gamma(\nu,\tilde I)}(\{e_\varepsilon(\cdot)>s\})$$

for each of the subsets $\tilde{I}$. Since $\gamma(\nu, \tilde{I})$ is an escaping random leaf, its length is uniformly bounded from zero. So, according to Corollary 5.7,

$$m_{\gamma(\nu,\tilde{I})}(\{e_\varepsilon(\,\cdot\,) > s\}) \le Ce^{-cs}.$$

Therefore, adding the previous inequality over all subsets $\tilde{I}$, we get that the total mass of

$$\sum_{I \subset J(\nu)} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I$$

is bounded by $\sum_{\tilde{I} \subset J} Ce^{-cs}\, \mathrm{P}_\varepsilon(\tilde{I} \mid J) \le Ce^{-cs}$, as claimed. □

Now we fix an integer $N \ge 1$, and we split (28) as

$$\mu^e_{\varepsilon,n} = \mu^1_{\varepsilon,n,N} + \mu^2_{\varepsilon,n,N},$$

where the first part corresponds to the sum of the terms with $j - \nu \ge N$, and the second one includes all the other terms in (28).

Lemma 6.7 allows us to show that the total mass of $\mu^1_{\varepsilon,n,N}$ goes uniformly to zero when N increases:

Corollary 6.8. *There are $C > 0$ and $c > 0$, independent of ε, n, N, such that the total mass of the measure*

$$\mu^1_{\varepsilon,n,N} = \frac{1}{n} \sum_{j=1}^{n} \sum_{\nu=1}^{j-N} \sum_{\tau=0}^{\nu} \int_{\mathcal{J}_{j,\tau}} \sum_{I \subset J(\nu)} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \, dB_{\varepsilon,j}(J)$$

is bounded by Ce^{-cN}.

Proof. Using Lemma 6.7, the total mass of $\mu^1_{\varepsilon,n,N}$ is bounded by

$$\frac{1}{n} \sum_{j=1}^{n} \sum_{\nu=1}^{j-N} \sum_{\tau=0}^{\nu} \int_{\mathcal{J}_{j,\tau}} Ce^{-c(j-\nu)} \, dB_{\varepsilon,j}(J)$$

Since the sets $\mathcal{J}_{j,\tau}$ are pairwise disjoint for each fixed j, and $dB_{\varepsilon,j}$ is a probability measure on $\mathcal{J}_j$, this is bounded by

$$\frac{1}{n} \sum_{j=1}^{n} \sum_{s=N}^{\infty} Ce^{-cs} \le Ce^{-cN}.$$

This proves the corollary. □

Next, we use Lemmas 6.5 and 6.6 to bound the measure

$$\mu^2_{\varepsilon,n,N} = \frac{1}{n} \sum_{j=1}^{n} \sum_{\nu=j-N+1}^{j} \sum_{\tau=0}^{\nu} \int_{\mathcal{J}_{j,\tau}} \sum_{I \subset J(\nu)} \mathcal{T}_\varepsilon^{j-\tau}(\psi_{\mathbf{y}} m_{\mathbf{y}})_I \, dB_{\varepsilon,j}(J).$$

Using that the $\mathcal{J}_{\varepsilon,\tau}$ corresponding to different values of τ are pairwise disjoint, we get that for each j and ν the sum in τ is bounded by

$$C\int_{\mathcal{J}_j}\sum_{\tilde{I}\subset J(\nu)} \mathrm{P}_\varepsilon(\tilde{I}\mid J)\big(f_*^s(m_{L(\nu,\tilde{I})}\mid\{e(\,\cdot\,)>s\})+r_{\varepsilon,s}(\,\cdot\,)\big)\,dB_{\varepsilon,j}(J)$$

where $s=j-\nu$. We write $r_{\varepsilon,0}(\,\cdot\,)=r_\varepsilon(\,\cdot\,)$. Note that for $s=0$ our notations I and $\tilde{I}$, as defined before, coincide.

The skew-product $\mathrm{P}_\varepsilon(\tilde{I}\mid J)\times dB_{\varepsilon,j}$ defines a measure in the space of $\tilde{I}$, with total mass bounded by 1. Let $dU_{\varepsilon,j}$ be its push-forward under the map

$$\tilde{I}\mapsto L(\nu,\tilde{I}),$$

defined from the set of all $\tilde{I}$ to the space of escaping leaves, $\nu=\nu(\tilde{I})$ being the escape situation associated to the definition of each $\tilde{I}$. Then $dU_{\varepsilon,j}$ is a measure in the space $\mathcal{U}$ of escaping leaves of f, with total mass bounded by 1, and the previous expression is bounded by

$$C\int_{\mathcal{U}} f_*^s(m_L\mid\{e(\,\cdot\,)>s\})\,dU_{\varepsilon,j}(L)+Cr_{\varepsilon,s}(\,\cdot\,).$$

So far we have shown that

$$\mu^2_{\varepsilon,n,N}\le\frac{1}{n}\sum_{j=1}^{N}\sum_{s=0}^{N-1}\Big[C\int_{\mathcal{U}} f_*^s(m_L\mid\{e(\,\cdot\,)>s\})\,dU_{\varepsilon,j}(L)+Cr_{\varepsilon,s}(\,\cdot\,)\Big].$$

It is time to introduce the measure $\lambda_{\varepsilon,n}$ on the set $\tilde{X}$ defined by

$$\lambda_{\varepsilon,n}=\frac{1}{n}\sum_{j=1}^{n}C\int_{\mathcal{U}} m_L\,dU_{\varepsilon,j}(L).$$

The previous inequality becomes

$$\mu^2_{\varepsilon,n,N}\le\sum_{s=0}^{N-1} f_*^s(\lambda_{\varepsilon,n}\mid\{e(\,\cdot\,)>s\})+C\sum_{s=0}^{N-1} r_{\varepsilon,s}(\,\cdot\,). \tag{35}$$

It is clear from the definition that $\lambda_{\varepsilon,n}$ is absolutely continuous along unstable leaves, with density bounded by the constant C. Moreover, the total mass is bounded by the constant C, because every leaf L has length ≤ 1, and the total mass of $dU_{\varepsilon,j}$ is bounded by 1. Now define,

$$M_{\varepsilon,n,N}=\mu^1_{\varepsilon,n,N}\quad\text{and}\quad R_{\varepsilon,N}(\,\cdot\,)=\mu^0_{\varepsilon,n}+C\sum_{s=0}^{N-1} r_{\varepsilon,s}(\,\cdot\,).$$

Corollary 6.8 says that the total mass of $M_{\varepsilon,n,N}$ decreases exponentially with N. Lemmas 6.4–6.6 imply that $\lim_{\varepsilon\to 0}R_{\varepsilon,N}(\,\cdot\,)=0$ for every fixed

N. Moreover,

$$\mu_{\varepsilon,n} \leq \sum_{s=0}^{\infty} f_*^s(\lambda_{\varepsilon,n} \mid \{e(\cdot) > s\}) + M_{\varepsilon,n,N} + R_{\varepsilon,N}(\cdot),$$

as claimed in Proposition 6.3. □

6.3. **Stochastic stability.** Now we are ready to prove Theorem A. We start from the conclusion of Proposition 6.3:

$$\mu_{\varepsilon,n} \leq \sum_{s=0}^{\infty} f_*^s p_*(\lambda_{\varepsilon,n} \mid \{e(\cdot) > s\}) + M_{\varepsilon,n,N} + R_{\varepsilon,N}(\cdot).$$

Making $n \to \infty$ along a suitable subsequence,

- $\mu_{\varepsilon,n}$ accumulates on the unique stationary measure μ_ε;
- $\lambda_{\varepsilon,n}$ accumulates on some measure λ_ε on $\tilde{X}$, absolutely continuous along unstable leaves with density and total mass bounded by C;
- $M_{N,\varepsilon,n}$ accumulates on a measure $M_{N,\varepsilon}$ with total mass bounded by $C\exp(-cN)$.

Keeping N fixed and making $\varepsilon \to 0$ along a suitable subsequence of any given sequence,

- μ_ε accumulates on some measure μ_0, which must be f-invariant (see [19, Theorem 1.1]);
- λ_ε accumulates on some measure λ on $\tilde{X}$, absolutely continuous along unstable leaves with density and total mass bounded by C;
- $M_{N,\varepsilon}$ accumulates on some measure M_N whose total mass is less than $C\exp(-cN)$;

We also have $R_{N,\varepsilon}(\cdot) \to 0$, pointwise. In this way we get

$$\mu_0 \leq \sum_{s=0}^{\infty} f_*^s p_*(\lambda \mid \{e(\cdot) > s\}) + M_N \quad \text{for all } N \geq 1.$$

Finally, making N go to infinity, we obtain that

$$\mu_0 \leq \lambda_0 \quad \text{where} \quad \lambda_0 = \sum_{s=0}^{\infty} f_*^s p_*(\lambda \mid \{e(\cdot) > s\}).$$

Just as in Remark 4.15, the measure $p_*\lambda$ is absolutely continuous along unstable manifolds in X. Then, as in Lemma 4.16, the saturation λ_0 is also absolutely continuous along unstable manifolds. It follows that the f-invariant measure μ_0 is absolutely continuous along unstable manifolds. Since there exists a unique such probability measure, cf. Theorem 2.9, we conclude that μ_0 coincides with the SRB measure μ_* in Corollary 4.17.

This completes the proof of Theorem A.

References

[1] J. F. Alves and V. Araújo. Stochastic stability for robust classes of non-uniformly expanding maps. *Astérisque.*

[2] A. Andronov and L. Pontryagin. Systèmes grossiers. *Dokl. Akad. Nauk. USSR*, 14:247–251, 1937.

[3] V. Araújo. Attractors and time averages for random maps. *Annales de l'Inst. Henri Poincaré - Analyse Non-linéaire*, 17:307–369, 2000.

[4] L. Arnold. *Random dynamical systems.* Springer-Verlag, 1998.

[5] A. Avila and C. G. Moreira. Statistical properties of unimodal maps: smooth families with negative Schwarzian derivative. *Astérisque.*

[6] V. Baladi and M. Viana. Strong stochastic stability and rate of mixing for unimodal maps. *Ann. Sci. École Norm. Sup.*, 29:483–517, 1996.

[7] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Annals of Math.*, 133:73–169, 1991.

[8] M. Benedicks and M. Viana. Solution of the basin problem for Hénon-like attractors. *Invent. Math.*, 143:375–434, 2001.

[9] M. Benedicks and L.-S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergod. Th. & Dynam. Sys.*, 12:13–37, 1992.

[10] M. Benedicks and L.-S. Young. SBR-measures for certain Hénon maps. *Invent. Math.*, 112:541–576, 1993.

[11] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Astérisque*, 261:13–56, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[12] P. Collet. Ergodic properties of some unimodal mappings of the interval. Technical report, Institute Mittag-Leffler, 1984.

[13] W. de Melo and S. van Strien. *One-dimensional dynamics.* Springer Verlag, 1993.

[14] L. J. Díaz, J. Rocha, and M. Viana. Strange attractors in saddle-node cycles: prevalence and globality. *Invent. Math.*, 125:37–74, 1996.

[15] 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.

[16] A. Katok and Yu. Kifer. Random perturbations of transformations of an interval. *J. Analyse Math.*, 47:193–237, 1986.

[17] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94:313–333, 1982.

[18] Yu. Kifer. *Ergodic theory of random perturbations.* Birkhäuser, 1986.

[19] Yu. Kifer. *Random perturbations of dynamical systems.* Birkhäuser, 1988.

[20] R. Mañé. A proof of the C^1 stability conjecture. *Publ. Math. I.H.E.S.*, 66:161–210, 1988.

[21] R. Metzger. Stochastic stability for contracting Lorenz maps and flows. *Comm. Math. Phys.*, 212:277–296, 2000.

[22] L. Mora and M. Viana. Abundance of strange attractors. *Acta Math.*, 171:1–71, 1993.

[23] 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.

[24] J. Palis and F. Takens. *Hyperbolicity and sensitive-chaotic dynamics at homoclinic bifurcations*. Cambridge University Press, 1993.

[25] Ya. Pesin. Families of invariant manifolds corresponding to non-zero characteristic exponents. *Math. USSR. Izv.*, 10:1261–1302, 1976.

[26] C. Pugh and M. Shub. Ergodic attractors. *Trans. Amer. Math. Soc.*, 312:1–54, 1989.

[27] J. Robbin. A structural stability theorem. *Annals of Math.*, 94:447–493, 1971.

[28] C. Robinson. Structural stability of vector fields. *Annals of Math.*, 99:154–175, 1974. Errata in Annals of Math. 101:368, 1975.

[29] V.A. Rokhlin. On the fundamental ideas of measure theory. *A. M. S. Transl.*, 10:1–52, 1962. Transl. from Mat. Sbornik 25 (1949), 107–150.

[30] W. Rudin. *Real and complex analysis*. McGraw-Hill, 3 edition, 1987.

[31] Ya. Sinai. Gibbs measures in ergodic theory. *Russian Math. Surveys*, 27:21–69, 1972.

[32] Ph. Thieullen, C. Tresser, and L.-S. Young. Positive Lyapunov exponent for generic one-parameter families of unimodal maps. *J. Anal. Math.*, 64:121–172, 1994.

[33] Q. Wang and L.-S. Young. Strange attractors with one direction of instability. *Comm. Math. Phys.*, 218:1–97, 2001.

[34] L.-S. Young. Stochastic stability of hyperbolic attractors. *Ergod. Th. & Dynam. Sys.*, 6:311–319, 1986.

Royal Institute of Technology (KTH), Department of Mathematics, S-100 44 Stockholm, Sweden

E-mail address: michaelb@math.kth.se

Instituto de Matemática Pura e Aplicada (IMPA), Est. D. Castorina 110, Jardim Botânico, 22460-320 Rio de Janeiro, Brazil

E-mail address: viana@impa.br