Statistical stability for robust classes of maps with non-uniform expansion

José F. Alves and Marcelo Viana *

December 27, 2016

Abstract

We consider open sets of transformations in a manifold M, exhibiting non-uniformly expanding behaviour in some forward invariant domain $U \subset M$. Assuming that each transformation has a unique SRB measure in U, and some general uniformity conditions, we prove that the SRB measure varies continuously with the dynamics in the L^1-norm.

As an application we show that an open class of maps introduced in [V1] fits this situation, thus proving that the SRB measures constructed in [A] vary continuously with the map.

1 Introduction

In general terms, Dynamics has a twofold aim: to describe, for the majority of dynamical systems, the typical behaviour of trajectories, specially as time goes to infinity; to understand how this behaviour changes when the system is modified, and to what extent it is *stable* under small modifications. In this work we are primarily concerned with the latter problem.

A first fundamental concept of stability, *structural stability*, was formulated by Andronov and Pontryagin [AP]. It requires that the whole orbit structure remain unchanged under any small perturbation of the dynamical system: there exists a homeomorphism of the ambient manifold mapping trajectories of the initial system onto trajectories of the perturbed one, preserving the direction of time. In the early sixties,

*J.F.A. was partially supported by IMPA/CNPq and Fundação Calouste Gulbenkian. M.V. was partially supported by Pronex-Dynamical Systems, CNPq 001/2000, and Faperj.

Smale introduced the notion of *uniformly hyperbolic* (or *Axiom A*) system, having as one of his main goals to obtain a characterization of structural stability. Such a characterization was conjectured by Palis and Smale in [PS]: a diffeomorphism (or a flow) is structurally stable if and only if it is uniformly hyperbolic and satisfies the so-called strong transversality condition. Before that, structural stability had been proved for certain classes of systems, including Anosov and Morse-Smale systems. The "if" part of the conjecture was proved by Robbin, de Melo, Robinson in the mid-seventies. The converse remained a major open problem for yet another decade, until it was settled by Mañé for C^1 diffeomorphisms (perturbations are small with respect to the C^1 norm). The flow case was recently solved by Hayashi, also in the C^1 category. The C^k case, $k > 1$, is still open both for diffeomorphisms and for flows. See e.g. the book of Palis and Takens [PT] for precise definitions, references and a detailed historical account.

Despite these remarkable successes, structural stability proved to be too strong a requirement for many applications. Several important models, including e.g. Lorenz flows and Hénon maps, are not stable in the structural sense, yet key aspects of their dynamical behaviour clearly persist after small modifications of the system. Weaker notions of stability, with a similar topological flavour, were proposed throughout the sixties and the seventies, but they all turned out to be too restrictive.

More recently, increasing emphasis has been put on expressing stability in terms of persistence of *statistical properties* of the system. A natural formulation, the one that concerns us most in this work, corresponds to continuous variation of physical measures as a function of the dynamical system. Let us explain this in precise terms. We consider discrete-time systems, namely, smooth transformations $\varphi : M \to M$ on a manifold M. A Borel probability measure μ on M is a *Sinai-Ruelle-Bowen (SRB) measure* (or a *physical measure*), if there exists a positive Lebesgue measure set of points $z \in M$ for which

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

for any continuous function $f : M \to \mathbb{R}$. In other words, time averages of all continuous functions are given by the corresponding spatial averages computed with respect to μ, at least for a large set of initial states $z \in M$.

Let us suppose that φ admits a forward invariant region $U \subset M$, meaning that $\varphi(U) \subset U$, and there exists a (unique) SRB measure $\mu = \mu_\varphi$ supported in U such that (1) holds for Lebesgue almost every point $z \in U$. We say that φ is *statistically stable* (restricted to U) if similar facts are true for any C^k nearby map ψ, for some $k \geq 1$, and the map $\psi \mapsto \mu_\psi$, associating to each ψ its SRB measure μ_ψ, is continuous at

$\psi = \varphi$. For this definition, we consider in the space of Borel measures the usual weak* topology: two measures are close to each other if they assign close-by integrals to each continuous function. Thus, this notion of stability really means that time averages of continuous functions are only slightly affected when the system is perturbed.

Uniformly expanding smooth maps are well known to be statistically stable, and so are Axiom A diffeomorphisms, restricted to the basin of each attractor. On the other hand, not much is known in this regard outside the uniformly hyperbolic context. In the present work we propose an approach to proving statistical stability for certain robust (open) classes of non-uniformly expanding maps. Precise conditions will be given in the next subsection. For the time being, we just mention that our maps φ exhibit asymptotic expansion,

$$\lim_{n\to+\infty} \frac{1}{n} \log \|D\varphi^n(z)v\| > 0 \quad \text{for every } v \in T_z M,$$

at Lebesgue almost every point z in some forward invariant region U, but they are not uniformly expanding. Moreover, they admit a unique SRB measure which is an ergodic invariant measure absolutely continuous with respect to Lebesgue measure in U. These properties remain valid in a neighborhood of the initial map, and we prove that *the SRB measure μ_φ varies continuously with the mapping in this neighborhood.* In fact, our approach proves statistical stability in a strong sense: *the density $d\mu_\varphi/dm$ of μ_φ with respect to Lebesgue measure m varies continuously with φ as an L^1-function.*

To the best of our knowledge this is the first result of statistical stability for maps with non-uniform expansion. An application, and the example we had in mind when we started this work, are the maps with multidimensional non-uniform expansion introduced in [V1], and whose SRB measures were constructed in [A]. Using a very different approach, Dolgopyat [D] proved statistical stability and other ergodic properties for some open classes of diffeomorphisms having partially hyperbolic attractors whose central direction is *mostly contracting* (negative Lyapunov exponents). In that situation, cf. also Bonatti-Viana [BV], SRB measures are absolutely continuous with respect to Lebesgue measure along the strong-unstable (uniformly expanding) foliation of the attractor. Our systems in the present work are closer in spirit to partially hyperbolic attractors with *mostly expanding* central direction, in the sense of Alves-Bonatti-Viana [ABV]. Statistical stability for the latter systems has not yet been proved.

Acknowledgements. Besides our own institutions, we are grateful to the Paul Erdős Center in Budapest, where the final ideas of this work were put together.

1.1 Statement of results

Let $\varphi \colon M \to M$ be a map from some d-dimensional manifold into itself, S be some region in M, and $\phi : S \to S$ be a return map for φ in S. That is, there exists a countable partition $\mathcal{R} = \{R_i\}_i$ of a full Lebesgue measure subset of S, and there exists a function $h \colon \mathcal{R} \to \mathbb{Z}^+$ such that

$$\phi \mid R = \varphi^{h(R)} \mid R \quad \text{for each} \quad R \in \mathcal{R}.$$

For simplicity, we will assume that S is diffeomorphic to some bounded region $\tilde{S}$ of $\mathbb{R}^d$ (but similar arguments hold in general, using local charts). Then we can pretend that $S \subset \mathbb{R}^d$, through identifying it with $\tilde{S}$, and we do so.

We say that ϕ is a C^2 *piecewise expanding map* if the following conditions hold:

- The boundary of each R_i is piecewise C^2 (a countable union of C^2 hypersurfaces) and has finite $(d-1)$-dimensional volume.
- Each $\phi_i \equiv \phi \mid R_i$ is a C^2 bijection from the interior of R_i onto its image, admitting a C^2 extension to the closure of R_i.
- There is $0 < \sigma < 1$ such that $\|D\phi_i^{-1}\| < \sigma$ for every $i \geq 1$.

We say that ϕ has *bounded distortion* if:

- There is some $K > 0$ such that for every $i \geq 1$
$$\frac{\left\|D\left(J \circ \phi_i^{-1}\right)\right\|}{\left|J \circ \phi_i^{-1}\right|} < K,$$
where J is the Jacobian of ϕ.

Moreover, we say that ϕ has *long branches* if the images of the elements of the partition $\mathcal{R}$ satisfy the following geometric condition:

- There are constants $1 \geq \beta > \sigma/(1-\sigma)$ and $\rho > 0$ such that the boundary of each $\phi(R_i)$ has a tubular neighborhood of size ρ inside $\phi(R_i)$, and the C^2 components of the boundary of each $\phi(R_i)$ meet at angles greater than $\arcsin(\beta) > 0$.

It was shown in [A, Section 5] that every C^2 piecewise expanding map with bounded distortion and long branches has some invariant probability measure μ absolutely continuous with respect to Lebesgue measure on S (henceforth denoted m and assumed to be normalized). Then

$$\mu^* = \sum_{j=0}^{\infty} \varphi_*^j \left(\mu \mid \{h > j\}\right) \tag{2}$$

is an absolutely continuous invariant measure for φ. Moreover, the density $d\mu/dm$ of μ is in $L^p(S)$ for $p = d/d-1$. As a consequence, the measure μ^* is finite, as long as we have:

- The function h is in $L^q(S)$ for $q = d$ (this is taken so that $1/p + 1/q = 1$).

It was also observed in [A, Sections 5 and 6] that the absolutely continuous invariant measure μ^* may be taken ergodic (which implies that it is an SRB measure for φ) and, moreover, φ has finitely many such ergodic measures.

Now we state our first main result. Let $k \geq 2$ be fixed, and $\mathcal{U}$ be an open set of C^k transformations on M admitting a forward invariant compact region U. We endow $\mathcal{U}$ with the C^k topology. Assume that we may associate to each $\varphi \in \mathcal{U}$ a map $\phi_\varphi \colon S \to S$, a partition $\mathcal{R}_\varphi$ of a full Lebesgue measure subset of $S \subset U$, and a function $h_\varphi \colon \mathcal{R}_\varphi \to \mathbb{Z}^+$, such that ϕ_φ is a C^2 piecewise expanding map with bounded distortion and long branches, and $h_\varphi \in L^q(m)$. We consider elements φ_0 of $\mathcal{U}$ satisfying the following uniformity conditions:

(U1) Given any integer $N \geq 1$ and any $\epsilon > 0$, there is $\delta = \delta(\epsilon, N) > 0$ such that for $j = 1, \ldots, N$

$$\|\varphi - \varphi_0\|_{C^k} < \delta \Rightarrow m\big(\{h_\varphi = j\}\Delta\{h_{\varphi_0} = j\}\big) < \epsilon,$$

where Δ represents symmetric difference of two sets.

(U2) Given $\epsilon > 0$, there are $N \geq 1$ and $\delta = \delta(\epsilon, N) > 0$ for which

$$\|\varphi - \varphi_0\|_{C^k} < \delta \Rightarrow \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q < \epsilon,$$

where $\mathcal{X}_{\{h_\varphi > j\}}$ denotes the characteristic function of the set $\{h_\varphi > j\}$.

(U3) Constants σ, K, β, ρ as above may be chosen uniformly in a C^k neighborhood of φ_0.

We also assume that the maps in a neighborhood of φ_0 satisfy the following non-degeneracy condition: given any $\epsilon > 0$ there exists $\delta > 0$ such that

$$m(E) \leq \delta \quad \Rightarrow \quad m(\varphi^{-1}(E)) \leq \epsilon \tag{3}$$

for any measurable subset E of S and any φ in $\mathcal{U}$. This can often be enforced by requiring some jet of order $l \leq k$ of φ_0 to be everywhere non-degenerate.

Theorem A. *Let $\mathcal{U}$ be as above, and suppose that every $\varphi \in \mathcal{U}$ admits a unique SRB measure μ_φ in U. Then*

1. *μ_φ is absolutely continuous with respect to the Lebesgue measure m;*
2. *if $\varphi_0 \in \mathcal{U}$ satisfies (U1), (U2), (U3) then φ_0 is statistically stable in a strong sense: the map*
$$\mathcal{U} \ni \varphi \mapsto \frac{d\mu_\varphi}{dm}$$
is continuous, with respect to the L^1-norm, at $\varphi = \varphi_0$.

We observe that under assumption (U1), condition (U2) can be reformulated in equivalent terms as:

(U2') Given $\epsilon > 0$, there is $\delta > 0$ for which
$$\|\varphi - \varphi_0\|_{C^k} < \delta \Rightarrow \|h_\varphi - h_{\varphi_0}\|_q < \epsilon.$$

A proof of this equivalence will be given in Remark 3.5 at the end of Section 3.

Our next results state that the assumptions of Theorem A do correspond to robust classes of smooth maps in some manifolds.

Theorem B. *There exists a non-empty open set $\mathcal{N}$ in the space of C^3 transformations from $S^1 \times \mathbb{R}$ into itself such that: every $\varphi \in \mathcal{N}$ admits a C^2 piecewise expanding return map ϕ_φ with bounded distortion and long branches, the return time h_φ is in $L^2(m)$, and conditions (U1)–(U3) are satisfied.*

The open set $\mathcal{N}$ we exhibit for the proof of this result is the one constructed in [V1]. As pointed out in that paper, the choice of the cylinder $S^1 \times \mathbb{R}$ as ambient space is rather arbitrary, the construction extends easily to more general manifolds. In what follows we briefly describe the set $\mathcal{N}$, referring the reader to [V1] and Section 4 for more details.

Let d be some large integer: $d \geq 16$ suffices, but is far from being optimal. Let $a_0 \in (1, 2)$ be such that the critical point $x = 0$ is pre-periodic under iteration by the quadratic map $q(x) = a_0 - x^2$ (again, this is far too strong a requirement on the parameter a_0). Let $b : S^1 \to \mathbb{R}$ be a Morse function, for instance, $b(t) = \sin(2\pi t)$. Note that $S^1 = \mathbb{R}/\mathbb{Z}$. For each $\alpha > 0$, consider the map $\varphi_\alpha : S^1 \times \mathbb{R} \to S^1 \times \mathbb{R}$ given by $\varphi_\alpha(\theta, x) = (\hat{g}(\theta), \hat{f}(\theta, x))$, where $\hat{g}$ is the uniformly expanding map of the circle defined by $\hat{g}(\theta) = d\theta \pmod{\mathbb{Z}}$, and $\hat{f}(\theta, x) = a(\theta) - x^2$ with $a(\theta) = a_0 + \alpha b(\theta)$. We shall take $\mathcal{N}$ to be a small C^3 neighborhood of φ_α, for some (fixed) sufficiently small α.

It is easy to check that for α small enough there is an interval $I \subset (-2, 2)$ for which $\varphi_\alpha(S^1 \times I)$ is contained in the interior of $S^1 \times I$. Thus, any map φ uniformly close to φ_α has $U = S^1 \times I$ as a forward invariant region, and so φ has an attractor inside this invariant region, which is precisely the set

$$\Lambda = \bigcap_{n \geq 0} \varphi^n(U).$$

Observe also that (3) holds in this context, as long as the neighborhood $\mathcal{N}$ is chosen sufficiently small. Indeed, denoting $J\varphi = \det D\varphi$, we have

$$J\varphi_\alpha = D\hat{g}\,\frac{\partial \hat{f}}{\partial x} \quad \text{and} \quad \frac{\partial J\varphi_\alpha}{\partial x} = D\hat{g}\,\frac{\partial^2 \hat{f}}{\partial x^2}\,.$$

Our assumptions give that the last expression is bounded away from zero. So, choosing $\mathcal{N}$ small enough, there exists $c_1 > 0$ such that $|\partial J\varphi_\alpha/\partial x| \geq c_1$ for any $\varphi \in \mathcal{N}$. Consequently, $m(\varphi^{-1}(E)) \leq \operatorname{const} m(E)^{1/2}$ for any $\varphi \in \mathcal{N}$ and any measurable set E.

It was shown in [A] that the maps $\varphi \in \mathcal{N}$ admit (finitely many) SRB measures, which are ergodic absolutely continuous invariant measures. To be able to apply Theorem A to this open set $\mathcal{N}$, we also have to show that the SRB measure is unique for each $\varphi \in \mathcal{N}$. This will follow from a stronger fact that we state in the next theorem.

We say that φ is *topologically mixing* if for every open set $A \subset S^1 \times I$ there is some $n = n(A) \in \mathbb{Z}^+$ for which $\varphi^n(A) = \Lambda$. Moreover, φ is *ergodic with respect to Lebesgue measure* if for every Borel subset $B \subset S^1 \times I$ such that $\varphi^{-1}(B) = B$, either B or $(S^1 \times I) \setminus B$ have Lebesgue measure equal to zero. Clearly, if φ is ergodic with respect to Lebesgue measure then it has at most one SRB measure: any basin has full Lebesgue measure in $S^1 \times I$.

Theorem C. *Let $\mathcal{N}$ be as described above. Then the transformations $\varphi \in \mathcal{N}$ are topologically mixing and ergodic with respect to Lebesgue measure.*

This work is organized as follows. Theorem A is proved in Sections 2 and 3. The proof of Theorem B occupies Sections 4 and 5. Finally, Theorem C is proved in Sections 6 and 7.

2 Absolute continuity

The proof of Theorem A uses the notion of variation for functions in any dimension. For $f \in L^1(\mathbb{R}^d)$ with compact support, the *variation* of f is

$$\operatorname{var}(f) = \sup\left\{ \int_{\mathbb{R}^d} f \operatorname{div}(g)\,dm \colon g \in C_0^1(\mathbb{R}^d, \mathbb{R}^d),\, \|g\|_0 \leq 1 \right\},$$

where $C_0^1(\mathbb{R}^d, \mathbb{R}^d)$ is the set of C^1 maps from $\mathbb{R}^d$ to $\mathbb{R}^d$ with compact support, and $\| \, \|_0$ is the supremum norm in $C_0^1(\mathbb{R}^d, \mathbb{R}^d)$. If f is a C^1 map, then $\mathrm{var}(f)$ coincides with $\int \|Df\| dm$ (see e.g. [G, Example 1.2]). We consider the space of *bounded variation* functions

$$BV(\mathbb{R}^d) = \left\{ f \in L^1(\mathbb{R}^d) : \ \mathrm{var}(f) < +\infty \right\}.$$

The following general results about bounded variation functions are used in the sequel.

Proposition 2.1. *Given $f \in BV(\mathbb{R}^d)$, there is a sequence $(f_n)_n$ of C^∞ maps such that*

$$\lim_{n\to\infty} \int |f - f_n| dm = 0 \quad \textit{and} \quad \lim_{n\to\infty} \int \|Df_n\| dm = var(f).$$

Proof. See [G, Theorem 1.17]. □

Proposition 2.2. *If $(f_k)_k$ is a sequence of functions in $BV(\mathbb{R}^d)$ such that there is a constant $K_0 > 0$ for which*

$$\mathrm{var}(f_k) \le K_0 \quad \textit{and} \quad \int |f_k| dm \le K_0 \quad \textit{for every } k,$$

then $(f_k)_k$ has a subsequence converging in the L^1-norm to an f_0 with $\mathrm{var}(f_0) \le K_0$.

Proof. See [G, Theorem 1.19]. □

Proposition 2.3. *Let $f \in BV(\mathbb{R}^d)$ and take $p = d/(d-1)$. Then*

$$\|f\|_p \le K_1 \, \mathrm{var}(f),$$

where $K_1 > 0$ is a constant depending only on d.

Proof. See [G, Theorem 1.28]. □

Let $\varphi \in \mathcal{U}$ and ϕ be as in the statement of Theorem A. We introduce the transfer operator $\mathcal{L}_\phi : L^1(S) \longrightarrow L^1(S)$ associated to ϕ, defined by

$$\mathcal{L}_\phi f = \sum_{i=1}^{\infty} \frac{f \circ \phi_i^{-1}}{|J \circ \phi_i^{-1}|} \mathcal{X}_{\phi(R_i)}.$$

By change of variables

$$\int (\mathcal{L}_\phi f) g dm = \int f (g \circ \phi) dm, \tag{4}$$

whenever the integrals make sense. In particular, each fixed point of $\mathcal{L}_\phi$ is the density of an absolutely continuous ϕ-invariant finite measure. We also use the fact that $\mathcal{L}$ never expands L^1 norms:

$$\int |\mathcal{L}_\phi f| dm \leq \int \mathcal{L}_\phi |f| dm = \int |f| dm,$$

The next lemma, a Lasota-Yorke type inequality for maps in $BV(\mathbb{R}^d)$, plays a crucial role in the proof of the existence of fixed points for $\mathcal{L}_\phi$.

Lemma 2.4. *For any $\varphi \in \mathcal{U}$, there are constants $0 < \lambda < 1$ and $K_2 > 0$ such that*

$$\operatorname{var}(\mathcal{L}_\phi f) \leq \lambda \operatorname{var}(f) + K_2 \int |f| dm$$

for every $f \in BV(\mathbb{R}^d)$. Moreover, λ and K_2 may be chosen uniform in a neighborhood of any $\varphi_0 \in \mathcal{U}$ that satisfies (U3).

Proof. The first part is [A, Lemma 5.4]. The argument gives $\lambda = \sigma(1 + 1/\beta)$ and $K_2 = K + 1/(\beta\rho) + K\beta$. So both may be taken uniform in a whole neighborhood of any map φ_0 satisfying (U3). $\square$

Consider for each $k \geq 1$ the function

$$f_k = \frac{1}{k} \sum_{j=0}^{k-1} \mathcal{L}_\phi^j 1.$$

Using (4) and the fact that $f_k \geq 0$ we have $\int |f_k| dm = 1$ for $k \geq 1$. By Lemma 2.4, $\operatorname{var}(f_k) \leq K_3$ for $k \geq 1$, where $K_3 = \operatorname{var}(\mathcal{X}_S) + K_2 \sum_{k=0}^{\infty} \lambda^k + 1$. It follows from Proposition 2.2 that $(f_k)_k$ has a subsequence converging in the L^1-norm to some ρ with $\operatorname{var}(\rho) \leq K_3$. Hence, $\mu_\phi = \rho m$ is an absolutely continuous ϕ-invariant probability measure. ¿From this it is deduced in [A, Section 6] that

$$\mu_\varphi^* = \sum_{j=0}^{\infty} \varphi_*^j \left(\mu_\phi \mid \{h_\varphi > j\}\right) \tag{5}$$

is an absolutely continuous φ-invariant finite measure.

Lemma 2.5. *Given any $f \in L^1(\mathbb{R}^d)$, then the sequence $1/n \sum_{j=0}^{n-1} \mathcal{L}_\phi^j f$ has some accumulation point in $L^1(\mathbb{R}^d)$ that is a function with variation bounded by $4K_2\|f\|_1$.*

Proof. Let $f \in L^1(\mathbb{R}^d)$ and take a sequence $(f_n)_n$ in $BV(\mathbb{R}^d)$ converging to f in the L^1-norm. It is no restriction to assume that $\|f_n\|_1 \leq 2\|f\|_1$ for every $n \geq 1$ and we do it. For each $n \geq 1$ we have

$$\operatorname{var}(\mathcal{L}_\phi^j f_n) \leq \lambda^j \operatorname{var}(f_n) + K_2\|f_n\|_1 \leq 3K_2\|f\|_1$$

for large j. So, taking k large enough we have

$$\operatorname{var}\left(\frac{1}{k}\sum_{j=0}^{k-1}\mathcal{L}_\phi^j f_n\right) \leq 4K_2\|f\|_1.$$

Moreover

$$\Big\|\frac{1}{k}\sum_{j=0}^{k-1}\mathcal{L}_\phi^j f_n\Big\|_1 \leq \frac{1}{k}\sum_{j=0}^{k-1}\|\mathcal{L}_\phi^j f_n\|_1 \leq 2\|f\|_1$$

for every $j \geq 1$. It follows from Proposition 2.2 that there exists some $\hat{f}_n \in BV(\mathbb{R}^d)$ and a sequence $(k_i)_i$ for which

$$\lim_{i\to\infty}\Big\|\frac{1}{k_i}\sum_{j=0}^{k_i-1}\mathcal{L}_\phi^j f_n - \hat{f}_n\Big\|_1 = 0$$

and, moreover, $\operatorname{var}(\hat{f}_n) \leq 4K_2\|f\|_1$. Now we apply the same argument to the sequence $(\hat{f}_n)_n$ in order to obtain a subsequence $(n_l)_l$ such that $(\hat{f}_{n_l})_l$ converges in the L^1-norm to some $\hat{f}$ with $\operatorname{var}(\hat{f}) \leq 4K_2\|f\|_1$. Since

$$\Big\|\frac{1}{k}\sum_{j=0}^{k-1}\mathcal{L}_\phi^j f_{n_l} - \hat{f}\Big\|_1 \leq \Big\|\frac{1}{k}\sum_{j=0}^{k-1}\mathcal{L}_\phi^j f_{n_l} - \hat{f}_{n_l}\Big\|_1 + \|\hat{f}_{n_l} - \hat{f}\|_1,$$

there is some sequence $(k_l)_l$ for which

$$\lim_{l\to\infty}\Big\|\frac{1}{k_l}\sum_{j=0}^{k_l-1}\mathcal{L}_\phi^j f_{n_l} - \hat{f}\Big\|_1 = 0.$$

On the other hand,

$$\Big\|\frac{1}{k_l}\sum_{j=0}^{k_l-1}\left(\mathcal{L}_\phi^j f_{n_l} - \mathcal{L}_\phi^j f\right)\Big\|_1 \leq \frac{1}{k_l}\sum_{j=0}^{k_l-1}\|f_{n_l} - f\|_1 = \|f_{n_l} - f\|_1$$

and this last term goes to 0 as $l \to \infty$. Finally, this implies that

$$\lim_{l\to\infty} \Big\|\frac{1}{k_l}\sum_{j=0}^{k_l-1} \mathcal{L}_\phi^j f - \hat f\Big\|_1 = 0,$$

thus proving that $\hat f$ is an accumulation point for the sequence $1/n \sum_{j=0}^{n-1} \mathcal{L}_\phi^j f$. □

Observe that any accumulation point $\hat f$ of a sequence as in the lemma is a fixed point for the transfer operator.

Corollary 2.6. *Given any ϕ-invariant set $A \subset S$ with positive Lebesgue measure, there is an absolutely continuous ϕ-invariant probability measure $\mu_A = f_A m$ for which $\mu_A(A) = 1$. Moreover, f_A may be taken with* $\mathrm{var}(f_A) \le 4K_2$.

Proof. Let $A \subset S$ be a ϕ-invariant set with positive Lebesgue measure. Considering in the previous lemma $f = \mathcal{X}_A \in L^1(\mathbb{R}^d)$, we find $f_A \in BV(\mathbb{R}^d)$ and a sequence $(k_l)_l$ for which $\mathrm{var}(f_A) \le 4K_2\|\mathcal{X}_A\|_1 \le 4K_2$ and

$$\lim_{l\to\infty} \Big\|\frac{1}{k_l}\sum_{j=0}^{k_l-1} \mathcal{L}_\phi^j \mathcal{X}_A - f_A\Big\|_1 = 0.$$

In particular $\|f_A\|_1 = m(A) > 0$. Then $\mu_A = (f_A/m(A))m$ is a probability, and it is ϕ-invariant because f_A is a fixed point of $\mathcal{L}_\phi$. Since

$$m(A)\mu_A(S\setminus A) = \lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1}\int_{S\setminus A} \mathcal{L}_\phi^j \mathcal{X}_A\, dm = \lim_{l\to\infty}\frac{1}{k_l}\sum_{j=0}^{k_l-1}\int_S (\mathcal{X}_{(S\setminus A)}\circ\phi^j)\mathcal{X}_A\, dm = 0$$

we have that μ_A gives full weight to A, thus concluding the proof of the result. □

Corollary 2.7. *There is a constant $\widehat K(d) > 0$ such that if $A \subset S$ is a ϕ-invariant set with positive Lebesgue measure, then $m(A) \ge \widehat K(d)$.*

Proof. Let $A \subset S$ be a ϕ-invariant set with positive Lebesgue measure and $\mu_A = f_A m$ a measure as in Corollary 2.6. Since $f_A \in BV(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ (recall Proposition 2.3) and μ_A gives full weight to A, it follows from Minkowski's inequality that

$$1 = \int_A f_A dm \le \|f_A\|_p \cdot \|\mathcal{X}_A\|_q \le K_1 4K_2\, m(A)^{1/d}.$$

We take $\widehat K(d) = (K_1 4K_2)^{-d}$. □

Lemma 2.8. *There are ergodic absolutely continuous φ-invariant measures $\mu_1^*, \dots, \mu_r^*$ supported in U, and positive numbers $\alpha_1, \dots, \alpha_r$, such that $\alpha_1 + \cdots + \alpha_r = 1$ and $\mu_\varphi^* = \alpha_1\mu_1^* + \cdots + \alpha_r\mu_r^*$.*

Proof. If μ_φ^* is ergodic, it is enough to take $r = 1$, $\alpha_1 = 1$, and $\mu_1^* = \mu_\varphi^*$. Otherwise, there exists some φ-invariant set A such that $0 < \mu_\varphi^*(A) < 1$. Let us observe that $A \cap S$ is necessarily ϕ-invariant:

$$\phi^{-1}(A \cap S) = \{x \in S : \phi(x) \in A\} = \bigcup_{j \geq 1} \left(\varphi^{-j}(A) \cap \{h_\varphi = j\}\right) = A \cap S,$$

Because of the assumption $\mu_\varphi^*(A) > 0$ and the definition of μ_φ^* in (5), there exists $j \geq 0$ such that $\mu_\phi(\varphi^{-j}(A) \cap \{h = j\}) > 0$. Then, $\mu_\phi(A) = \mu_\phi(\varphi^{-j}(A))$ is also positive. Since μ_ϕ is supported in S, this is the same as saying that $\mu_\phi(A \cap S) > 0$. Then, by absolute continuity, $m(A \cap S) > 0$. So, by Corollary 2.7, we have

$$m(A \cap S) \geq \widehat{K}(d). \tag{6}$$

Now, either A is minimal, in the sense that there is no φ-invariant set $B \subset A$ with $\mu_\varphi^*(A) > \mu_\varphi^*(B) > 0$, or else we apply the same arguments as before, with B and $A \setminus B$ in the place of A. Of course, all this can be said about the complement $M \setminus A$ as well. The important point is that at all stages we have an uniform lower bound as in (6). Thus this subdivision must stop after a finite number of steps. That is, we find a decomposition of M into a finite number of φ-invariant sets $A_1, \dots, A_r$ with positive μ_φ-measure, such that $m(A_i \cap S) \geq \widehat{K}(d)$ for $1 \leq i \leq r$ and, most important, each A_i is minimal in the above sense. Define $\alpha_i = \mu_\varphi^*(A_i)$ and μ_i^* to be the restriction of μ_φ^* to A_i, divided by α_i. Clearly, each μ_i is absolutely continuous and φ-invariant (because A_i is φ-invariant). Moreover, μ_i^* is ergodic, because A_i was taken minimal. $\square$

The first part of Theorem A is contained in the following consequence of the previous lemma.

Proposition 2.9. *If $\varphi \in \mathcal{U}$ has a unique SRB measure μ_φ supported in U, then $\mu_\varphi^* = \mu_\varphi$.*

Proof. Each of the measures μ_i^* in Lemma 2.8 is an SRB measure supported in U. Therefore, the assumption implies that $r = 1$ and $\mu_\varphi^* = \mu_1^* = \mu_\varphi$. $\square$

3 Statistical stability

Now we prove that under the assumptions of Theorem A the density of the measure μ_φ^* varies continuously with the map φ, in the L^1-norm. Let φ_0 be some map in $\mathcal{U}$ satisfying (U1), (U2), (U3), and $(\varphi_n)_n$ be a sequence of maps in $\mathcal{U}$ converging to φ_0 in the C^k topology. Let ϕ_0 be the return map of φ_0, with return time h_0 (cf. the definition of $\mathcal{U}$). Let μ_0 be an absolutely continuous ϕ_0-invariant probability measure and μ_0^* be the φ_0-invariant measure obtained from it as in (2). We represent by ρ_0 the density of μ_0. Moreover, we denote by ϕ_n, h_n, μ_n, μ_n^*, ρ_n the corresponding objects for each φ_n. Our goal is to prove that μ_n^* converges to μ_0^* as n goes to infinity.

We begin by noting that, as a consequence of our construction,

$$\operatorname{var}(\rho_n) \le K_3 \quad \text{and} \quad \int \rho_n dm \le 1$$

for every $n \ge 1$ (recall Lemma 2.4). Thus, by Proposition 2.2, the sequence of densities $(\rho_n)_n$ is relatively compact with respect to the L^1 norm: any subsequence contains another subsequence which is L^1 convergent. This means that we only have to prove that $(\mu_{n_i}^*)_i$ converges to μ_0^* for every subsequence $(n_i)_i$ such that $(\rho_{n_i})_i$ converges in the L^1-norm to some function ρ_∞.

Let $(\rho_{n_i})_i$ and ρ_∞ be as above. The previous remark also gives $\operatorname{var}(\rho_\infty) \le K_3$. We consider $\mu_\infty = \rho_\infty m$ and define

$$\mu_\infty^* = \sum_{j=0}^{\infty} \varphi_*^j \left(\mu_\infty \mid \{h_0 > j\}\right).$$

We want to show that the densities of μ_n^* with respect to the Lebesgue measure converge in the L^1-norm to the density of μ_∞^* and, moreover, the measure μ_∞^* coincides with μ_0^*. We start with some auxiliary lemmas.

Lemma 3.1. *There is $K_4 = K_4(d) > 0$ such that, for any $f \in BV(\mathbb{R}^d)$ and any C^1 embedding $\psi\colon D \to \mathbb{R}^d$ of a compact domain $D \subset \mathbb{R}^d$,*

$$\int_D |f \circ \psi - f| dm \le K_4 \|\psi - id\|_0^d \operatorname{var}(f).$$

Proof. We start by proving the result when f is a continuous piecewise affine map. More precisely, we suppose that the support Δ of f can be decomposed into a finite number of domains $\Delta_1, \dots, \Delta_N$ such that the gradient ∇f of f is constant on each Δ_i. We define

$$D_1 = \{(x,z) \in \mathbb{R}^{d+1} \colon x \in D \text{ and } z \in [f(x), f(\psi(x))]\}$$

and let D_2 be the horizontal $\|\psi - id\|_0$-neighborhood of the graph of f. That is,

$$D_2 = \{(x,z) \in \mathbb{R}^{d+1} : z = f(y) \text{ for some } y \in \mathbb{R}^d \text{ with } \|x - y\| \leq \|\psi - id\|_0\}.$$

We claim that $D_1 \subset D_2$. Indeed, given $(x,z) \in D_1$, and since $z \in [f(x), f(\psi(x))]$, by the continuity of f there is y in the straight line segment $[x, \psi(x)]$ such that $z = f(y)$. Taking $t = y - x$ we have $\|t\| \leq \|\psi(x) - x\|$, which proves the claim. Now

$$\int_D |f \circ \psi - f| dm = \int_D \int_{[f(x), f(\psi(x))]} 1 \, dz dm(x) = \operatorname{vol}(D_1) \leq \operatorname{vol}(D_2).$$

For each $i = 1, \cdots, N$, let H_i be the horizontal $\|\psi - id\|_0$-neighborhood of the graph of $f \mid \Delta_i$. Clearly, $D_2 \subset H_1 \cup \cdots \cup H_N$, and so

$$\operatorname{vol}(D_2) \leq \sum_{i=1}^{N} \operatorname{vol}(H_i).$$

Letting $\nabla_i f$ denote the value of the gradient of $f \mid \Delta_i$, we have

$$\operatorname{vol}(H_i) \leq K_4 \|\psi - id\|_0^d \, \|\nabla_i f\| \operatorname{vol}(\Delta_i),$$

where $K_4 > 0$ is the volume of the unit ball in $\mathbb{R}^d$. Consequently,

$$\int_D |f \circ \psi - f| dm \leq K_4 \|\psi - id\|_0^d \sum_{i=1}^{N} \|\nabla_i f\| \operatorname{vol}(\Delta_i).$$

Taking into account that in this case

$$\sum_{i=1}^{N} \|\nabla_i f\| \operatorname{vol}(\Delta_i) = \int \|\nabla f\| dm = \operatorname{var}(f),$$

we obtain the result for any continuous piecewise affine map.

The next step is to deduce the result for any C^1 map f. For this we take a sequence $(f_n)_n$ of continuous piecewise affine maps such that

$$\|f - f_n\|_0 \to 0 \quad \text{and} \quad \|Df - Df_n\|_0 \to 0 \quad \text{as} \quad n \to \infty$$

(the derivatives Df_n are defined only in the interior of the smoothness domains). For instance, we may consider a sequence of triangulations $\mathcal{T}_n$ (of some cube covering the support of f), with diameters going to zero as $n \to \infty$, and let f_n be the unique piecewise affine function that coincides with f on the vertices of $\mathcal{T}_n$. Then

$$\int_D |f \circ \psi - f| dm = \lim_{n\to\infty} \int_D |f_n \circ \psi - f_n| dm$$

and

$$\mathrm{var}(f) = \int \|Df\| dm = \lim_{n\to\infty} \int \|Df_n\| dm = \lim_{n\to\infty} \mathrm{var}(f_n).$$

So, the previous case implies that the conclusion of the lemma holds also for f.

For the general case, we know by Proposition 2.1 that given $f \in BV(\mathbb{R}^d)$ there is a sequence $(f_n)_n$ of C^1 maps for which

$$\lim_{n\to\infty} \int |f - f_n| dm = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathrm{var}(f_n) = \mathrm{var}(f). \tag{7}$$

We have

$$\int_D |f \circ \psi - f| dm \le \int_D |f \circ \psi - f_n \circ \psi| dm + \int_D |f_n \circ \psi - f_n| dm + \int_D |f_n - f| dm.$$

Since

$$\int_D |f_n \circ \psi - f \circ \psi| dm = \int_{\psi(D)} |f_n - f| \cdot \Psi dm \le \|\Psi\|_0 \int |f_n - f| dm$$

where $\Psi = 1/|\det D\psi| \circ \psi^{-1}$, the result for general $f \in BV(\mathbb{R}^d)$ follows from (7) and the previous case. $\square$

At this point we also introduce the transfer operator $\mathcal{L}_\varphi$ associated to $\varphi \in \mathcal{U}$, defined for each $f \in L^1(\mathbb{R}^d)$ as

$$\mathcal{L}_\varphi f(y) = \sum_{x \in \varphi^{-1}(y)} \frac{f(x)}{|\det D\varphi(x)|}. \tag{8}$$

The function $\mathcal{L}_\varphi f(y)$ fails to be defined only when y is a critical value of φ. We have

$$\int (\mathcal{L}_\varphi f) g \, dm = \int f (g \circ \varphi) \, dm$$

for every $f, g \in L^1(\mathbb{R}^d)$ such that integrals make sense.

Lemma 3.2. *Given $\varphi_0 \in \mathcal{U}$ and $\epsilon > 0$, there is $\delta > 0$ such that for any $\varphi \in \mathcal{U}$ with $\|\varphi - \varphi_0\|_{C^1} < \delta$ we have*

$$\int |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm \le \epsilon \left(\operatorname{var}(f) + \|f\|_1\right),$$

for every $f \in BV(\mathbb{R}^d)$ with support contained in S.

Proof. Our assumptions, namely the existence of a piecewise expanding return map, imply that the critical set of φ_0 (the set of points where φ_0 fails to be a local diffeomorphism) intersects S in a zero Lebesgue measure set. Given any $\epsilon_1 > 0$, define $\mathcal{C}(\epsilon_1)$ as the ϵ_1-neighborhood of this intersection. Clearly, $m(\varphi(\mathcal{C}(\epsilon_1))) \le \operatorname{const} m(\mathcal{C}(\epsilon_1))$ for some constant that may be taken uniform in a C^1 neighborhood of φ_0. So, using (3) we may fix ϵ_1 small enough so that

$$m\big(\varphi_1^{-1}(\varphi_2(\mathcal{C}(\epsilon_1))\big) \le \frac{1}{2}\left(\frac{\epsilon}{8K_1}\right)^q, \tag{9}$$

for every φ_1, φ_2 in some neighborhood of φ_0, where K_1 is the constant in Proposition 2.3 and $q = d$. We decompose $S \setminus \mathcal{C}(\epsilon_1)$ into a finite collection $\mathcal{D}(\varphi_0)$ of domains of injectivity of φ_0. Observe that if φ is close enough to φ_0, in the C^1 sense, then $\mathcal{C}(\epsilon_1)$ also contains the critical set of φ. Hence, we may define a corresponding collection $\mathcal{D}(\varphi)$ of domains of injectivity for φ in $S \setminus \mathcal{C}(\epsilon_1)$, and there is a natural bijection associating to each $D_0 \in \mathcal{D}(\varphi_0)$ a unique $D \in \mathcal{D}(\varphi)$ such that the Lebesgue measure of $D \Delta D_0$ is small. Observe that $\mathcal{L}_\varphi$ is supported in

$$\varphi(S) = \varphi(\mathcal{C}(\epsilon_1)) \cup \bigcup_{D \in \mathcal{D}(\varphi)} \varphi(D),$$

and analogously for φ_0. So,

$$\begin{aligned}
\int |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm \;\; &\le \;\; \int_{\varphi_0(\mathcal{C}(\epsilon_1)) \cup \varphi(\mathcal{C}(\epsilon_1))} \big(|\mathcal{L}_\varphi f| + |\mathcal{L}_{\varphi_0} f|\big) dm && (10)\\
&+ \sum_{D_0 \in \mathcal{D}(\varphi_0)} \int_{\varphi_0(D_0) \cap \varphi(D)} |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm && (11)\\
&+ \sum_{D_0 \in \mathcal{D}(\varphi_0)} \int_{\varphi_0(D_0) \Delta \varphi(D)} \big(|\mathcal{L}_\varphi f| + |\mathcal{L}_\varphi f_0|\big) dm, && (12)
\end{aligned}$$

where D always denotes the element of $\mathcal{D}(\varphi)$ associated to each $D_0 \in \mathcal{D}(\varphi_0)$. Let us estimate the expressions on the right hand side of this inequality.

Let us start with (10). For notational simplicity, we write $E = \varphi_0(\mathcal{C}(\epsilon_1)) \cup \varphi(\mathcal{C}(\epsilon_1))$. Then

$$\int_E |\mathcal{L}_\varphi f| dm \le \int \mathcal{X}_E \big(\mathcal{L}_\varphi |f|\big) dm = \int (\mathcal{X}_E \circ \varphi)|f| dm.$$

It follows from Minkowski's inequality, Proposition 2.3, and (9) that

$$\int (\mathcal{X}_E \circ \varphi)|f| dm \le m\big(\varphi^{-1}(E)\big)^{1/q} \|f\|_p \le \frac{\epsilon}{8K_1} K_1 \operatorname{var}(f) = \frac{\epsilon}{8} \operatorname{var}(f).$$

The case $\varphi = \varphi_0$ gives a similar bound for the second term in (10). So,

$$\int_{\varphi_0(\mathcal{C}(\epsilon_1)) \cup \varphi(\mathcal{C}(\epsilon_1))} \big(|\mathcal{L}_\varphi f| + |\mathcal{L}_{\varphi_0} f|\big) dm \le \frac{\epsilon}{4} \operatorname{var}(f). \tag{13}$$

Making the change of variables $y = \varphi_0(x)$ in (11), we may rewrite it as

$$\int_{\widehat{D}_0} \left| \frac{f}{|\det D\varphi|} \circ (\varphi^{-1} \circ \varphi_0) - \frac{f}{|\det D\varphi_0|} \right| \cdot |\det D\varphi_0| dm,$$

where $\widehat{D}_0 = \varphi_0^{-1}\big(\varphi_0(D_0) \cap \varphi(D)\big) = D_0 \cap \big(\varphi_0^{-1} \circ \varphi\big)(D)$. For notational simplicity, we introduce $\psi = \varphi^{-1} \circ \varphi_0$. The previous expression is bounded by

$$\int_{\widehat{D}_0} \left(|f \circ \psi - f| \cdot \frac{|\det D\varphi_0|}{|\det D\varphi| \circ \psi} + |f| \cdot \left| \frac{|\det D\varphi_0|}{|\det D\varphi| \circ \psi} - 1 \right| \right) dm.$$

Choosing $\delta > 0$ sufficiently small, the assumption $\|\varphi - \varphi_0\|_{C^1} < \delta$ implies

$$\left| \frac{|\det D\varphi_0|}{|\det D\varphi| \circ \psi} - 1 \right| \le \epsilon \quad \text{and so} \quad \frac{|\det D\varphi_0|}{|\det D\varphi| \circ \psi} \le 2$$

on $S \setminus \mathcal{C}(\epsilon_1)$ (which contains $\widehat{D}_0$). Hence, using Lemma 3.1,

$$\begin{aligned} \int_{\varphi_0(D_0) \cap \varphi(D)} |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm &\le 2 \int_{\widehat{D}_0} |f \circ \psi - f| dm + \epsilon \int |f| dm \\ &\le 2K_4 \|\psi - id\|_0^d \operatorname{var}(f) + \epsilon \int |f| dm. \end{aligned}$$

Reducing $\delta > 0$, we can make $\|\psi - id\|_0^d$ arbitrarily small, so that

$$\int_{\varphi_0(D_0) \cap \varphi(D)} |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm \le \frac{\epsilon}{4} \operatorname{var}(f) + \epsilon \|f\|_1. \tag{14}$$

We estimate the terms in (12) in much the same way as we did for (10). For each D_0 let E be $\varphi_0(D_0)\Delta\varphi(D)$. The properties of the transfer operator, followed by Minkowski's inequality, yield

$$\begin{aligned}\int_E |\mathcal{L}_\varphi f| dm &\leq \int \mathcal{X}_E\big(\mathcal{L}_\varphi |f|\big) dm \\ &= \int \mathcal{X}_{\varphi^{-1}(E)} |f| dm \leq m\big(\varphi^{-1}(E)\big)^{1/q} \|f\|_p \,.\end{aligned}$$

Fix $\epsilon_2 > 0$ such that $\#\mathcal{D}(\varphi_0)4\epsilon_2 < \epsilon$. Taking δ sufficiently small, we may ensure that the Lebesgue measure of all the sets

$$\varphi^{-1}(E) = \varphi^{-1}(\varphi_0(D_0)\Delta\varphi(D))$$

is small enough so that, using also Proposition 2.3, the right hand side is less than $\epsilon_2 \operatorname{var}(f)$. In this way we get

$$\int_{\varphi_0(D_0)\Delta\varphi(D)} \big(|\mathcal{L}_{\varphi_0} f| + |\mathcal{L}_\varphi f|\big) dm \leq 2\epsilon_2 \operatorname{var}(f) \tag{15}$$

(the second term on the left is estimated in the same way as the first one). Putting (13), (14), (15) together, we obtain

$$\int |\mathcal{L}_\varphi f - \mathcal{L}_{\varphi_0} f| dm \leq \big(\frac{\epsilon}{2} + \#\mathcal{D}(\varphi_0)2\epsilon_2\big) \operatorname{var}(f) + \epsilon\|f\|_1$$

and this is smaller than $\epsilon(\operatorname{var}(f) + \|f\|_1)$. $\square$

Proposition 3.3. $\dfrac{d\mu^*_{n_i}}{dm}$ *converges to* $\dfrac{d\mu^*_\infty}{dm}$ *in the* L^1*-norm.*

Proof. We are going to prove that given $\epsilon > 0$ there is $\delta > 0$ for which

$$\left\| \frac{d\mu^*_{n_i}}{dm} - \frac{d\mu^*_\infty}{dm} \right\|_1 < \epsilon \quad \text{whenever} \quad \|\varphi_{n_i} - \varphi_0\|_{C^1} < \delta.$$

We have

$$\mu^*_\infty = \sum_{j=0}^{\infty} (\varphi_0^j)_* \left(\mu_\infty \mid \{h_0 > j\}\right) \quad \text{and} \quad \mu^*_{n_i} = \sum_{j=0}^{\infty} (\varphi_{n_i}^j)_* \left(\mu_{n_i} \mid \{h_{n_i} > j\}\right). \tag{16}$$

By (U2) there is an integer $N \geq 1$ and $\delta = \delta(\epsilon, N) > 0$ for which

$$\|\varphi - \varphi_0\|_{C^1} < \delta \Rightarrow \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q < \frac{\epsilon}{4K_1K_3}. \tag{17}$$

In what follows we take $i \geq 1$ to be sufficiently large so that $\|\varphi_{n_i} - \varphi_0\| < \delta$. We split each one of the sums in (16) as

$$\mu_\infty^* = \sum_{j=0}^{N} \nu_{\infty,j} + \eta_{\infty,N} \quad \text{and} \quad \mu_{n_i}^* = \sum_{j=0}^{N} \nu_{n_i,j} + \eta_{n_i,N}, \tag{18}$$

where

$$\nu_{\infty,j} = (\varphi_0)_*^j \left(\mu_\infty \mid \{h_0 > j\}\right), \quad \eta_{\infty,N} = \sum_{j=N+1}^{\infty} (\varphi_0)_*^j \left(\mu_\infty \mid \{h_0 > j\}\right),$$

and $\nu_{n_i,j}$ and $\eta_{n_i,N}$ are defined similarly, with φ_{n_i} , μ_{n_i} , h_{n_i} in the place of φ, μ_∞ , h_0 , respectively. We have

$$\eta_{\infty,N}(M) = \sum_{j=N}^{\infty} \mu_\infty(\{h_0 > j\}) = \sum_{j=N}^{\infty} \int \rho_\infty \mathcal{X}_{\{h_0 > j\}} dm \leq \|\rho_\infty\|_p \cdot \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_0 > j\}} \Big\|_q,$$

and, analogously,

$$\eta_{n_i,N}(M) \leq \|\rho_{n_i}\|_p \cdot \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_{n_i} > j\}} \Big\|_q$$

which together with Proposition 2.3 and (17) yield

$$\left\| \frac{d\eta_{n_i,N}}{dm} - \frac{d\eta_{\infty,N}}{dm} \right\|_1 \leq \eta_{n_i,N}(M) + \eta_{\infty,N}(M) < \epsilon/2. \tag{19}$$

On the other hand, for $j = 1, \ldots, N$

$$\left\| \frac{d\nu_{n_i,j}}{dm} - \frac{d\nu_{\infty,j}}{dm} \right\|_1 = \left\| \mathcal{L}_{\varphi_{n_i}^j} (\rho_{n_i} \mathcal{X}_{\{h_{n_i} > j\}}) - \mathcal{L}_{\varphi_0^j} (\rho_\infty \mathcal{X}_{\{h_0 > j\}}) \right\|_1. \tag{20}$$

Denote

$$A = \|\mathcal{L}_{\varphi_{n_i}^j} (\rho_{n_i} \mathcal{X}_{\{h_{n_i} > j\}}) - \mathcal{L}_{\varphi_{n_i}^j} (\rho_\infty \mathcal{X}_{\{h_0 > j\}})\|_1$$

and

$$B = \|\mathcal{L}_{\varphi_{n_i}^j}(\rho_\infty \mathcal{X}_{\{h_0>j\}}) - \mathcal{L}_{\varphi_0^j}(\rho_\infty \mathcal{X}_{\{h_0>j\}})\|_1.$$

Here we also use the transfer operators for the iterated maps $\varphi_{n_i}^j$ and φ_0^j defined in the same way as for φ in (8). Then

$$\begin{aligned} A &\leq \|\rho_{n_i}\mathcal{X}_{\{h_{n_i}>j\}} - \rho_\infty \mathcal{X}_{\{h_0>j\}}\|_1 \\ &\leq \|\rho_{n_i}\mathcal{X}_{\{h_{n_i}>j\}} - \rho_\infty \mathcal{X}_{\{h_{n_i}>j\}}\|_1 + \|\rho_\infty \mathcal{X}_{\{h_{n_i}>j\}} - \rho_\infty \mathcal{X}_{\{h_0>j\}}\|_1 \\ &\leq \|\rho_{n_i} - \rho_\infty\|_1 + \|\rho_\infty(\mathcal{X}_{\{h_{n_i}>j\}} - \mathcal{X}_{\{h_0>j\}})\|_1 \end{aligned}$$

and the last term is bounded by $\|\rho_\infty\|_p \|\mathcal{X}_{\{h_{n_i}>j\}} - \mathcal{X}_{\{h_0>j\}}\|_q$. Taking into account (U1), we get $A \leq \epsilon/(4N)$ if i is sufficiently large. Using Proposition 3.2 we also get $B \leq \epsilon/(4N)$, for large i. It follows that (20) is less than $A + B \leq \epsilon/(2N)$ for each $1 \leq j \leq N$. Thus the sum over all these j's is less than $\epsilon/2$. Together with (19), this completes the proof of the proposition. □

Proposition 3.4. *μ_∞^* is a φ_0-invariant measure.*

Proof. It follows from Proposition 3.3 that $(\mu_{n_i}^*)_i$ converges to μ_∞^* in the weak* topology. Hence, given any $f\colon M \to \mathbb{R}$ continuous we have

$$\int f d\mu_{n_i}^* \to \int f d\mu_\infty^* \quad \text{when} \quad i \to \infty.$$

On the other hand, since $\mu_{n_i}^*$ is φ_{n_i}-invariant we have

$$\int f d\mu_{n_i}^* = \int (f \circ \varphi_{n_i}) d\mu_{n_i}^* \quad \text{for every } i.$$

So, it suffices to prove that

$$\int (f \circ \varphi_{n_i}) d\mu_{n_i}^* \to \int (f \circ \varphi_0) d\mu_\infty^* \quad \text{when} \quad i \to \infty. \tag{21}$$

We have

$$\begin{aligned} &\left|\int (f \circ \varphi_{n_i}) d\mu_{n_i}^* - \int (f \circ \varphi_0) d\mu_\infty^*\right| \leq \\ &\qquad \left|\int (f \circ \varphi_{n_i}) d\mu_{n_i}^* - \int (f \circ \varphi_0) d\mu_{n_i}^*\right| + \left|\int (f \circ \varphi_0) d\mu_{n_i}^* - \int (f \circ \varphi_0) d\mu_\infty^*\right|. \end{aligned}$$

Since $f \circ \varphi_{n_i} - f \circ \varphi_0$ is uniformly close to zero in the compact set U, when i is large, the first term in the sum above is close to zero for i sufficiently large. On the other hand, since $(\mu_{n_i}^*)_i$ converges to μ_∞^* in the weak* topology we also have that the second term in the sum above is close to zero if i is large. □

It follows from this last result and the uniqueness of the absolutely continuous φ-invariant measure that $\mu_\infty^* = \mu_0^*$. So, Proposition 3.3 really states that the measures $\mu_{n_i}^*$ have densities converging in the L^1-norm to the density of μ_0^*. This completes the proof of Theorem A.

Remark 3.5. We also check that conditions (U2') and (U2) are equivalent if we assume (U1). First we prove that (U2') implies (U2). Let $\epsilon > 0$ be some small number and take $N \geq 1$ in such a way that $\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_{\varphi_0} > j\}} \|_q < \epsilon/3$. We have

$$
\begin{aligned}
\Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q &= \Big\| h_\varphi - h_{\varphi_0} + h_{\varphi_0} - \sum_{j=0}^{N-1} \mathcal{X}_{\{h_{\varphi_0} > j\}} + \sum_{j=0}^{N-1} \mathcal{X}_{\{h_{\varphi_0} > j\}} - \sum_{j=0}^{N-1} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q \\
&\leq \|h_\varphi - h_{\varphi_0}\|_q + \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_{\varphi_0} > j\}} \Big\|_q + \sum_{j=0}^{N-1} \Big\| \mathcal{X}_{\{h_{\varphi_0} > j\}} - \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q,
\end{aligned}
$$

and so, if we take $\delta = \delta(N, \epsilon) > 0$ sufficiently small then, under assumptions (U2') and (U1), the first and third terms in the sum above can be made smaller than $\epsilon/3$. This gives the conclusion of condition (U2).

For the converse, let $\epsilon > 0$ be some small number, and $N \geq 1$ and $\delta = \delta(N, \epsilon) > 0$ be taken in such a way that the conclusion of (U2) holds for $\epsilon/3$. We have

$$
\begin{aligned}
\|h_\varphi - h_{\varphi_0}\|_q &= \Big\| h_\varphi - \sum_{j=0}^{N-1} \mathcal{X}_{\{h_\varphi > j\}} + \sum_{j=0}^{N-1} \big(\mathcal{X}_{\{h_\varphi > j\}} - \mathcal{X}_{\{h_{\varphi_0} > j\}} \big) + \sum_{j=0}^{N-1} \mathcal{X}_{\{h_{\varphi_0} > j\}} - h_{\varphi_0} \Big\|_q \\
&\leq \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q + \sum_{j=0}^{N-1} \| \mathcal{X}_{\{h_\varphi > j\}} - \mathcal{X}_{\{h_{\varphi_0} > j\}} \|_q + \Big\| \sum_{j=N}^{\infty} \mathcal{X}_{\{h_{\varphi_0} > j\}} \Big\|_q.
\end{aligned}
$$

By the choices of N and δ, the first and third terms in the last sum above are smaller than $\epsilon/3$. Moreover, by condition (U1), reducing δ we can make the second term also smaller than $\epsilon/3$, thus obtaining the conclusion of (U2').

4 Return maps and hyperbolic returns

Now we start the proof of Theorem B. As mentioned before, the open class φ of systems we consider at this points is the one described in [V1]. A suitable construction of a return map ϕ for each $\varphi \in \mathcal{N}$ was given in [A]. On the other hand, for Theorem C we shall need some features that do not follow directly from that construction, but may be obtained from a slight modification of it. In the present section, besides reviewing the arguments in [A], we explain how it can be modified to yield those additional features (Lemma 4.9 below). This is based on a notion of hyperbolic return, that we introduce below.

For the sake of clarity, we start by assuming that the map φ has the special form

$$\varphi(\theta, x) = (g(\theta), f(\theta, x)), \quad \text{with} \quad \partial_x f(\theta, x) = 0 \quad \text{if and only if} \quad x = 0, \tag{22}$$

and describe how ϕ is obtained for each C^2 map φ satisfying

$$\|\varphi - \varphi_\alpha\|_{C^2} \leq \alpha \quad \text{on} \quad S^1 \times I. \tag{23}$$

In Section 5 we explain how the conclusions extend to the general case, using the existence of a central invariant foliation, and we verify conditions (U1), (U2), (U3).

Our estimates on the derivative rely on a statistical analysis of the returns of orbits to the neighborhood $S^1 \times (-\sqrt{\alpha}, \sqrt{\alpha}\,)$ of the critical set $\{x = 0\}$. For this, we introduce the partition $\mathcal{Q}$ of I (up to a zero Lebesgue measure set) into the following intervals:

$$I_r = (\sqrt{\alpha}e^{-r}, \sqrt{\alpha}e^{-(r-1)}\,) \;\text{ for }\; r \geq 1, \quad \text{and} \quad I_r = -I_{-r} \;\text{ for }\; r \leq -1,$$

$$I_{0^+} = \big(I \setminus [-\sqrt{\alpha}, \sqrt{\alpha}\,]\big) \cap \mathbb{R}^+ \quad \text{and} \quad I_{0^-} = \big(I \setminus [-\sqrt{\alpha}, \sqrt{\alpha}\,]\big) \cap \mathbb{R}^-.$$

This induces corresponding partitions on each fiber $\{\theta\} \times I$, whose elements we also denote as I_r and $I_{0^\pm}$, since this abuse of language never leads to ambiguity. In what follows, $\alpha > 0$ is a sufficiently small number independent of any other constant involved in the arguments. Furthermore, we indicate which of the constants depend on α. Given $(\theta, x) \in S^1 \times I$ and $j \geq 0$ we define $(\theta_j, x_j) = \varphi^j(\theta, x)$.

Lemma 4.1. *There are $C_1 > 1$ and $0 < \eta < 1/4$ such that for every small α there is an integer $N = N(\alpha)$ satisfying*

1. *If $|x| < 2\sqrt{\alpha}$, then $\prod_{j=0}^{N-1} |\partial_x f(\theta_j, x_j)| \geq |x|\alpha^{-1+\eta}$.*
2. *If $|x| < 2\sqrt{\alpha}$, then $|x_j| > \sqrt{\alpha}$ for every $j = 1, \ldots, N$.*
3. *$C_1^{-1} \log(1/\alpha) \leq N \leq C_1 \log(1/\alpha)$.*

Proof. See [V1, Lemma 2.4] and [A, Lemma 2.1]. □

Lemma 4.2. *There are $\tau > 1$, $C_2 > 0$ and $\delta > 0$ such that for $(\theta, x) \in S^1 \times I$ and $k \geq 1$ the following holds:*

1. *If* $|x_0|, \dots, |x_{k-1}| \geq \sqrt{\alpha}$*, then* $\prod_{j=0}^{k-1} |\partial_x f(\theta_j, x_j)| \geq C_2 \sqrt{\alpha} \tau^k$*.*
2. *If* $|x_0|, \dots, |x_{k-1}| \geq \sqrt{\alpha}$ *and* $|x_k| < \delta$*, then* $\prod_{j=0}^{k-1} |\partial_x f(\theta_j, x_j)| \geq C_2 \tau^k$*.*

Proof. See [V1, Lemma 2.5]. □

The constants C_1, η, τ, C_2, and δ in these two lemmas depend only on the quadratic map q, and so they may be taken uniform in the whole $\mathcal{N}$.

Now, for each integer $j \geq 0$ we define

$$r_j(\theta, x) = \begin{cases} |r| & \text{if } \varphi^j(\theta, x) \in S^1 \times I_r \text{ with } |r| \geq 1; \\ 0 & \text{if } \varphi^j(\theta, x) \notin S^1 \times [-\sqrt{\alpha}, \sqrt{\alpha}]. \end{cases} \tag{24}$$

We say that $\nu \geq 0$ is a *return* for (θ, x) if $\varphi^\nu(\theta, x) \in S^1 \times (-\delta, \delta)$. In particular, $r_j(\theta, x) = 0$ unless j is a return for (θ, x). Let $n \geq 1$ and $0 \leq \nu_1 \leq \cdots \leq \nu_s \leq n$ be the returns of (θ, x) from time 0 to n. It follows from Lemma 4.1 that for each $1 \leq i \leq s$

$$\prod_{j=\nu_i}^{\nu_i+N-1} |\partial_x f(\theta_j, x_j)| \geq e^{-r_{\nu_i}(\theta,x)} \alpha^{-1/2+\eta},$$

and from the second part of Lemma 4.2

$$\prod_{j=0}^{\nu_1-1} |\partial_x f(\theta_j, x_j)| \geq C_2 \tau^{\nu_1} \quad \text{and} \quad \prod_{j=\nu_i+N}^{\nu_{i+1}-1} |\partial_x f(\theta_j, x_j)| \geq C_2 \tau^{\nu_{i+1}-\nu_i-N}.$$

Assume n is a return, that is, $\nu_s = n$. Then, putting the previous estimates together, we get

$$\prod_{j=0}^{n-1} |\partial_x f(\theta_j, x_j)| \geq C_2^{-1} \exp\left(4cn - \sum_{j \in G_n(\theta,x)} r_j(\theta, x)\right) \tag{25}$$

where

$$G_n(\theta, x) = \left\{1 \leq \nu_i \leq n-1 : r_{\nu_i}(\theta, x) \geq \left(\frac{1}{2} - 2\eta\right) \log \frac{1}{\alpha}\right\}. \tag{26}$$

See e.g. [A, Section 2].

A key fact in this construction is that the exponent on the right hand side of (25) is positive, except for a set of initial points (θ, x) whose measure decreases very rapidly with n. More precisely, let

$$E_n = \big\{(\theta, x) \in S^1 \times I : \sum_{j \in G_n(\theta,x)} r_j(\theta, x) > 2cn\big\}. \tag{27}$$

Then, cf. (16) and (17) in [V1, Section 2.4], there are constants $C, \gamma > 0$ such that

$$m\big(E_n\big) \le C e^{-\gamma\sqrt{n}} \tag{28}$$

for every sufficiently large n, only depending on α. Note that (25) gives

$$\|D\varphi^n(\theta, x)(0, 1)\| \ge C_2^{-1} e^{2cn} \ge e^{cn}$$

for any $(\theta, x) \in (S^1 \times I) \setminus E_n$ and n sufficiently large.

Following [A], we fix $0 < \epsilon < c/2$ and say that $n \ge 1$ is a *hyperbolic time* for $(\theta, x) \in S^1 \times I$ if

$$\sum_{\substack{i \in G_n(\theta,x) \\ k \le i < n}} r_i(\theta, x) < (c + \epsilon)(n - k) \quad \text{for every} \quad 0 \le k < n.$$

Let $p \ge 1$ be some sufficiently large integer: the precise condition will be recalled in a while, it involves only the expansion rates of the maps $\hat{g}$ and $\hat{f}$. Let H be the set of points that has at least one hyperbolic time greater or equal than p. Decompose $H = \cup_{n \ge p} H_n$, where each H_n is the set of points whose first hyperbolic time greater or equal to p is n. By [A, Proposition 2.5], there is a positive integer $n_0 = n_0(p, \epsilon) \ge p$ such that

$$(S^1 \times I) \setminus (H_p \cup \cdots \cup H_n) \subset E_n \quad \text{for every} \quad n \ge n_0\,. \tag{29}$$

Starting from this, [A] constructed a special partition $\mathcal{R}$ of (a full Lebesgue measure subset of) $S^1 \times I$ into rectangles, whose elements are the domains of smoothness of the corresponding return map ϕ. Let $\mathcal{Q}$ be the partition of I described above, and $\mathcal{P}_n$, $n \ge 1$, be the sequence of Markov partitions of S^1 defined as follows. Let $S^1 = \mathbb{R}/\mathbb{Z}$ have the orientation induced by the usual order in $\mathbb{R}$ and θ_0 be the fixed point of g close to $\theta = 0$. Define $\mathcal{P}_n$ by

- $\mathcal{P}_1 = \{[\theta_{j-1}, \theta_j) : 1 \le j \le d\}$, where $\theta_0, \theta_1, \ldots, \theta_d = \theta_0$ are the pre-images of θ_0 under g (ordered according to the orientation of S^1).

- $\mathcal{P}_{n+1} = \{\text{connected components of } g^{-1}(\omega) : \omega \in \mathcal{P}_n\}$ for each $n \geq 1$.

[A, Section 3] explains how the elements of $\mathcal{P}_p \times \mathcal{Q}$, any large $p \geq 1$, can be successively subdivided (according to the itineraries of points relative to the horizontal strips $S^1 \times I_*$, $I_* \in \mathcal{Q}$) to obtain a partition $\mathcal{R}$ of a full Lebesgue measure subset of $S^1 \times I$ with the following properties. $\mathcal{R}$ may be written as a union $\mathcal{R} = \cup_{n \geq p} \mathcal{R}_n$ satisfying

$$H_n \subset \bigcup_{R \in \mathcal{R}_n} R \quad \text{and} \quad R \cap H_n \neq \emptyset \quad \text{for every} \quad R \in \mathcal{R}_n. \tag{30}$$

The elements of $\mathcal{R}_n$, any $p \geq n$, are rectangles of the form $\omega \times J$, with ω belonging to $\mathcal{P}_n$ and J a subinterval of I_* for some $I_* \in \mathcal{Q}$.

Now one defines $h : \mathcal{R} \to \mathbb{Z}^+$, by setting $h(R) = n \geq p$ for each $R \in \mathcal{R}_n$. Moreover, $\phi(x) = \varphi^{h(R)}(x)$ for every $x \in R$ and $R \in \mathcal{R}_n$. It was shown in [A] that, as long as p is chosen sufficiently large, ϕ is a C^2 piecewise expanding map with bounded distortion and long branches. In particular, cf. Proposition 3.8 in that paper, the images $\phi(R)$ of the rectangles $R \in \mathcal{R}$ have sizes bounded away from zero by some constant depending only on α.

For the proof of Theorem C we need a stronger form of this conclusion that, basically, corresponds to having this lower bound independent of α. The precise property we need is stated in Lemma 4.9 below. In order to get it, we give an alternative construction of a partition and a return map. This goes along the same lines as before, except that hyperbolic times are replaced by the following notion. We say that $n \geq 1$ is a *hyperbolic return* for $(\theta, x) \in S^1 \times I$ if n is both a hyperbolic time and a return for (θ, x).

Corresponding to (28) and (29), we are going to prove that the Lebesgue measure of the set of points that have no hyperbolic returns smaller than some large integer n decays at least as fast as $Ce^{-c\sqrt{n}}$. We use the following result, that is interesting by itself.

Lemma 4.3. *There exist constants $\hat{C} = \hat{C}(\alpha) > 0$ and $\hat{\tau} = \hat{\tau}(\alpha) > 1$ such that the Lebesgue measure of the set of points (θ, x) that have no return up to time l is bounded by $\hat{C}\hat{\tau}^{-l}$, for all $l \geq 1$.*

Proof. Let $\theta \in S^1$ and J be any subinterval of I with length $|J| \geq \delta/2$. For each $j \geq 0$, we denote J_j the interval defined by $\varphi^j(\{\theta\} \times J) = \{\hat{g}^j(\theta)\} \times J_j$. Let $R \geq 0$ be the smallest iterate for which J_R intersects $(-\sqrt{\alpha}, \sqrt{\alpha})$. By Lemma 4.2, the length of the J_j grows exponentially fast up to time R. So, R is bounded from above by a constant R_0 that does not depend on the interval J, only on α and the lower bound $\delta/2$ for the initial length. Then we can decompose J_R into (at most) three subintervals J^-, J^0, J^+ as follows. If δ belongs to J_R, we take $J^+ = J_R \cap \{x > \delta/2\}$. Otherwise J^+ is the empty set. Analogously, $J^- = J_R \cap \{x < -\delta/2\}$ if $-\delta$ belongs J_R, and $J^- = \emptyset$ otherwise. Finally J^0 is the complement of $J^+ \cup J^-$ in J_R. Therefore,

1. J^0 is contained in $(-\delta, \delta)$;
2. $|J^0| > \delta/4$ or else J^0 is the whole J_R;
3. $|J^\pm| > \delta/2$ or else they are empty.

Property 1 means that R is a return for all the points falling in J^0. Property 2 implies that such points constitute a definite fraction of the initial interval J. Here we are using the fact that the map $f^R(\theta, \cdot)$ has bounded distortion on J. To see that this is so, observe that the iterates $f^j(\theta, \cdot)$, $1 \leq j \leq R$, are uniformly expanding on J, by Lemma 4.2. Moreover, these iterates remain outside a fixed neighborhood $(-\sqrt{\alpha}, \sqrt{\alpha})$ of the critical point 0. Since $\log \partial_x f$ is Lipschitz continuous outside this neighborhood, with uniform Lipschitz constants, the bounded distortion property follows along well-known lines. So far, we have shown that a definite fraction $\tau_0 > 0$ of the initial interval J attains a return in not more than R_0 iterations. Furthermore, by property 3 above, we may repeat the argument recurrently for each of the remaining intervals $J^\pm$, with θ replaced by $\hat{g}^R(\theta)$. Using the bounded distortion property for successive iterates, we conclude that after jR_0 iterates, any $j \geq 1$, all but a fraction $(1-\tau_0)^j$ of the initial interval J has already gone through at least one return. This proves the lemma. □

Lemma 4.4. *Given any sufficiently large n, then almost every $(\theta, x) \in (S^1 \times I) \setminus E_n$ has some hyperbolic return. Moreover, Lebesgue almost every $(\theta, x) \in (S^1 \times I)$ has infinitely many hyperbolic returns.*

Proof. It was shown that [A, Proposition 2.5] that any $(\theta, x) \in (S^1 \times I) \setminus E_n$ has some hyperbolic time, as long as n is large enough. On the other hand, if n is a hyperbolic time for (θ, x) and $l > n$ is the next return for (θ, x) after n (recall the previous lemma) then l is a hyperbolic return for (θ, x), since $r_j(\theta, x) = 0$ for $j = n+1, \cdots, l-1$. This proves the first statement. The second one follows from the same argument, using the Remark 2.6 in [A] that Lebesgue almost every point in $S^1 \times I$ has infinitely many hyperbolic times. □

Similarly to what we have done before, let $p \geq 1$ be some fixed large integer, and define H^* the set of points that has at least one hyperbolic return greater or equal than p. We decompose $H^* = \cup_{n \geq p} H_n^*$, where each H_n^* is the set of points whose first hyperbolic return greater or equal than p is n.

Proposition 4.5. *There is an integer* $n_1 = n_1(p, \epsilon) \geq p$ *and constants* $C_0, \gamma_0 > 0$ *such that for each* $n \geq n_1$

$$m\big((S^1 \times I) \setminus (H_p^* \cup \cdots \cup H_n^*)\big) \leq C_0 e^{-\gamma_0 \sqrt{n}}.$$

Proof. Take $n \geq \max\{2p, n_0\}$ and let $l = [n/2]$. Let $(\theta, x) \in H_l$ be such that some $1 \leq k \leq l$ is a return for $\varphi^l(\theta, x)$. Let us take k minimum. Then, $k + l$ is a return for (θ, x) and, as observed already in the proof of Lemma 4.4, it is a hyperbolic return for (θ, x). Our choices of l and k imply that $p \leq k + l \leq n$. So, we have shown that the set of points $(\theta, x) \in H_l$ such that $\varphi^l(\theta, x)$ has some return prior to time l is contained in $H_p^* \cup \cdots \cup H_n^*$. The same argument remains valid for any j between p and l. Hence, defining

$$B_l = \bigcup_{j=p}^{l} \{(\theta, x) \in H_j \colon \varphi^j(\theta, x) \text{ has no returns from time 1 to } l\}$$

we have $(H_p \cup \cdots \cup H_l) \setminus B_l \subset (H_p^* \cup \cdots \cup H_n^*)$ and so

$$m\big((S^1 \times I) \setminus (H_p^* \cup \cdots \cup H_n^*)\big) \leq m\big((S^1 \times I) \setminus (H_p \cup \cdots \cup H_l)\big) + m(B_l),$$

Taking into account (29) and (28) above, it suffices to show that $m(B_l)$ also decays rapidly when n increases. We define, for each $j \geq p$ and $R_j \in \mathcal{R}_j$,

$$R_j(l) = \{(\theta, x) \in R_j \colon \varphi^j(\theta, x) \text{ has no returns from time 1 to } l\}.$$

Using (30) we obtain

$$m(B_l) \leq \sum_{j=p}^{l} \sum_{R_j \in \mathcal{R}_j} m\big(R_j(l)\big). \tag{31}$$

Fixing some $R_j \in \mathcal{R}_j$ and $(\theta_0, x_0) \in R_j$ we deduce from Proposition 4.7

$$m\left(\varphi^j\big(R_j(l)\big)\right) = \int_{R_j(l)} |J(\theta, x)|\, dm(\theta, x) \geq \frac{1}{\Delta} |J(\theta_0, x_0)|\, m\big(R_j(l)\big).$$

Similarly, $m\left(\varphi^j(R_j)\right) \leq \Delta\, |J(\theta_0, x_0)|\, m(R_j)$. Hence

$$\frac{m\big(R_j(l)\big)}{m(R_j)} \leq \Delta^2 \frac{m\left(\varphi^j\big(R_j(l)\big)\right)}{m\left(\varphi^j(R_j)\right)}. \tag{32}$$

It follows from the fact that $\varphi^j \mid R_j(l)$ is a diffeomorphism that the iterates of points in $\varphi^j\big(R_j(l)\big)$ do not hit the critical region $S^1 \times [-\sqrt{\alpha}, \sqrt{\alpha}]$ from time 1 to l. As a consequence of Lemma 4.3

$$m\left(\varphi^j\big(R_j(l)\big)\right) \le \hat{C}\hat{\tau}^{-l}. \tag{33}$$

On the other hand, by [A, Proposition 3.8], there is some absolute constant $\delta > 0$ such that $m\left(\varphi^j\big(R_j\big)\right) \ge \delta$. Combining this with (32) and (33) we obtain

$$m\big(R_j(l)\big) \le \frac{\Delta^2 C}{\delta}\tau^{-l} m(R_j)$$

which together with (31) gives

$$m(B_l) \le \sum_{j=p}^{l} \sum_{R_j \in \mathcal{R}_j} \frac{\Delta^2 C}{\delta}\tau^{-l} m(R_j) \le (l-p)\frac{\Delta^2 C}{\delta}\tau^{-l} \le n\frac{\Delta^2 C}{\delta}\tau^{-n/2}.$$

□

Now we proceed just as in [A, Section 3], with the H_n^* in the role of the sets H_n. That is, we construct a new partition $\mathcal{R}^*$ of a full Lebesgue measure subspace of $S^1 \times I$ into rectangles, using hyperbolic returns instead of hyperbolic times. This new partition may also be written as a union $\mathcal{R}^* = \cup_{n \ge p} \mathcal{R}_n^*$ with the sets $\mathcal{R}_n^*$ defined inductively and satisfying

$$H_n^* \subset \bigcup_{R \in \mathcal{R}_n^*} R \quad \text{and} \quad R \cap H_n^* \ne \emptyset \quad \text{for every} \quad R \in \mathcal{R}_n^*.$$

Furthermore, for each $n \ge p$, rectangles in $\mathcal{R}_n^*$ also have the form $\omega \times J$, with ω belonging to $\mathcal{P}_n$ and J a subinterval of I_* for some $I_* \in \mathcal{Q}$. We define a map $h^* : \mathcal{R}^* \to \mathbb{Z}^+$, by putting $h^*(R) = n \ge p$ for each $R \in \mathcal{R}_n^*$. This also gives a new return map ϕ^* defined by $\phi^* \mid R = \varphi^{h^*(R)} \mid R$ for every $R \in \mathcal{R}^*$.

It follows from (22) that for each $n \ge 1$ there is a map $F_n \colon S^1 \times I \to I$ such that $\varphi^n(\theta, x) = (g^n(\theta), F_n(\theta, x))$ for every $(\theta, x) \in S^1 \times I$. Let (θ, x) belong to $R^* \in \mathcal{R}$ and $h = h^*(R)$. Then

$$D\varphi^h(\theta, x) = \begin{pmatrix} \partial_\theta g^h(\theta) & 0 \\ \partial_\theta F_h(\theta, x) & \partial_x F_h(\theta, x) \end{pmatrix}.$$

By [A, Lemma 4.1], there is some constant $C_3 > 0$ such that for every $(\theta, x) \in S^1 \times I$ we have $|\partial_\theta F_h(\theta, x)| \le C_3 |\partial_\theta g^h(\theta)|$. Then

$$\|D\varphi^{-h}(\varphi^h(\theta, x))\| \le \max\left\{|\partial_\theta g^h(\theta)|^{-1} + C_3|\partial_x F_h(\theta, x)|^{-1}, |\partial_x F_h(\theta, x)|^{-1}\right\}.$$

Moreover, $|\partial_\theta g^h(\theta)|^{-1} \leq (d-\alpha)^{-h}$ and, from Lemma 4.9,

$$|\partial_x F_h(\theta, x)|^{-1} \leq C_2 \exp(-(2c-\epsilon)h).$$

Hence,

$$\|D\varphi^{-h}(\varphi^h(\theta, x))\| \leq (d-\alpha)^{-h} + (1+C_3)C_2 \exp(-(2c-\epsilon)h), \tag{34}$$

At this point we can specify the choice of the integer p: we take $p \geq 1$ large enough so that the induced map ϕ^* associated to φ is an expanding map in the sense of the definition given in Subsection 1.1 (recall that $h \geq p$).

Remark 4.6. The constants C_0 and γ_0 that we found in Proposition 4.5 depend on C, γ, and the quadratic map q. Moreover, the integer n_1 only depends on the previous constants and the integer $p \geq 1$. Note also that our choice of p only depends on the expansion rates of the maps $\hat{g}$ and $\hat{f}$. In particular, p may be taken independent of the map $\varphi \in \mathcal{N}$.

Proposition 4.7. *There is some constant $\Delta > 1$ such that for every $n \geq p$, $R \in \mathcal{R}_n^*$ and $(\theta, x), (\sigma, y) \in R$ we have*

$$\frac{1}{\Delta} \leq \left| \frac{J(\theta, x)}{J(\sigma, y)} \right| \leq \Delta,$$

where J is the Jacobian of $\varphi^n \mid R$.

Proof. As in [A, Proposition 4.2],

$$\left\| D\left(\log \left| J \circ \phi^{-1} \right|\right) \right\| = \frac{\left\| D\left(J \circ \phi^{-1}\right) \right\|}{\left|\left(J \circ \phi^{-1}\right)\right|}$$

is bounded by some constant C_1 that depends only on c, α, and bounds on the partial derivatives of f and g of first and second order. Fix some $R \in \mathcal{R}_n^*$ with $n \geq p$ and let $\phi = \varphi^n \mid R$. We have

$$\left| \frac{J(\theta, x)}{J(\sigma, y)} \right| = \exp\left(\log \left| \left(J \circ \phi^{-1}\right)(\phi(\theta, x)) \right| - \log \left| \left(J \circ \phi^{-1}\right)(\phi(\sigma, y)) \right| \right)$$

and

$$\left| \log \left| \left(J \circ \phi^{-1}\right)(\phi(\theta, x)) \right| - \log \left| \left(J \circ \phi^{-1}\right)(\phi(\sigma, y)) \right| \right| \leq \left\| D\left(\log \left| J \circ \phi^{-1} \right|\right)(\tau, z) \right\| \cdot C_2$$

for some $(\tau, z) \in \phi(S)$, where $C_2 > 0$ depends only on the diameter of $S^1 \times I$. So, it suffices to take $\Delta = C_1 C_2$. ∎

Remark 4.8. For future use, let us point out that the distortion bound Δ given by the proof does not depend on the choice of p.

Finally, we get the following result that is needed for the proof of Proposition 6.2. The whole point is that we have the right to use (25), which comes from the second statement of Lemma 4.2, because all the iterates φ^n in the definition of ϕ end at returns.

Lemma 4.9. *Let $(\theta, x) \in R$ for some $R \in \mathcal{R}^*$. Then for every $j = 0, \cdots, h^*(R) - 1$ we have*

$$\prod_{i=j}^{h^*(R)-1} |\partial_x f(\theta_i, x_i)| \geq C_2^{-1} \exp\big((2c - \epsilon)(h(R) - j)\big).$$

Proof. Analogous to [A, Lemma 3.7] with (25) in the place of (9) in [A]. $\square$

5 Uniformity conditions

An important feature of this construction, cf. [V1, Section 2.5], is that it remains valid for any map ψ close enough to φ, with uniform bounds on the measure of the exceptional sets $E_n(\psi)$:

$$m(E_n(\psi)) \leq C e^{-\gamma\sqrt{n}} \quad \text{for every } n \geq 1$$

where C and γ may be taken uniform (that is, constant) in a whole C^3 neighborhood of φ. Let us explain this last point, since it is not explicitly addressed in the previous papers. One consequence is that Proposition 4.5 holds in the whole open set $\mathcal{N}$, with uniform constants C_0 and γ_0 (recall Remark 4.6).

As explained in [V1, Section 2.5], it follows from the methods of [HPS] that any map ψ sufficiently close to φ admits a unique invariant central foliation $\mathcal{F}^c$ of $S^1 \times I$ by smooth curves uniformly close to vertical segments. This is because the vertical foliation is invariant and normally expanding for the map φ. In addition, the space of leaves of $\mathcal{F}^c$ is homeomorphic to a circle, and the map induced by ψ in it is topologically conjugate to $\hat{g}$. The previous analysis can then be carried out in terms of the expansion of ψ along this central foliation $\mathcal{F}^c$. More precisely, $|\partial_x f(\theta, x)|$ is replaced by

$$|\partial_c f(\theta, x)| \equiv |D\psi(\theta, x) v_c(\theta, x)|,$$

where $v_c(\theta, x)$ represents a norm 1 vector tangent to the foliation at (θ, x). The previous observations imply that v_c is uniformly close to $(0, 1)$ if ψ is close to φ. Moreover, cf. [V1, Section 2.5], it is no restriction to suppose $|\partial_c f(\theta, 0)| \equiv 0$ (incidentally, this is the only place where we need our maps to be C^3), so that $\partial_c f(\theta, x) \approx |x|$, as in the unperturbed case; recall (22). Defining $r_j(\theta, x)$ and $E_n = E_n(\psi)$ in the same way as before, cf. (24), we obtain an analog of (25):

$$\|D\psi^n(\theta, x)v_c(\theta, x)\| = \prod_{j=0}^{n-1} |\partial_c f(\theta_j, x_j)| \geq C_2^{-1} \exp\left(4cn - \sum_{j \in G_n(\theta,x)} r_j(\theta, x)\right),$$

for every (θ, x). We define $E_n(\psi)$ in the same way as $E_n = E_n(\varphi)$, recall (27), and then

$$\|D\psi^n(\theta, x)v_c(\theta, x)\| \geq e^{cn} \quad \text{for all } (\theta, x) \in (S^1 \times I) \setminus E_n.$$

The arguments in [V1, Section 2.4] apply with $|\partial_c f|$ in the place of $|\partial_x f|$, proving that the Lebesgue measure of $E_n(\psi)$ satisfies the bound in (28). The constants C and γ produced by these arguments depend only on α, which is fixed, and on estimates obtained in the previous sections of that paper. So, to see that these constants are indeed uniform in a neighborhood of φ, it suffices to check that the same is true for those preparatory estimates. This is clear in the case of the results of Section 2.1 (Lemmas 2.1 and 2.2, and Corollary 2.3), because they only involve one iterate of the map. Let us point out that the definition of admissible curve for ψ is just the same as for the unperturbed map φ. A continuity argument can be applied also to Section 2.2, but it is more subtle. The key observation is that, although the statements of Lemmas 2.4 and 2.5 involve an unbounded number of iterates, their proofs are based on analyzing bounded stretches of orbits. Finally, the results in Section 2.4 (Lemmas 2.6 and 2.7), involve not more than $M \approx \log(1/\alpha)$ iterates. So, once more by continuity, their estimates remain valid in a neighborhood of φ. We have concluded the observation that the bound (28) on the Lebesgue measure of the exceptional set E_n holds uniformly in a neighborhood of the map.

Now we are able to show that conditions (U1)-(U3) are satisfied by every element of $\mathcal{N}$, as long as we take the open set $\mathcal{N}$ sufficiently small.

(U1) The construction of the partition that leads to the map h_φ is based on the itineraries of points through the horizontal strips $S^1 \times I_*$ with $I_* \in \mathcal{Q}$, according to the expanding behaviour of the iterates of φ at hyperbolic returns. Since these hyperbolic returns depend only on a finite number of iterates of the map φ, by continuity, we can perform the construction of the partition in such a way that for some fixed integer N the Lebesgue measure of $\{h_\varphi = j\}$ varies continuously with the map φ for $j \leq N$.

(U2) For every $\varphi \in \mathcal{N}$ and any fixed large integer $N \geq 1$,

$$\Big\| \sum_{j\geq N} \mathcal{X}_{\{h_\varphi > j\}} \Big\|_q \leq \sum_{j\geq N} m\big(\{h_\varphi > j\}\big)^{1/q} \leq \sum_{j\geq N} C_0^{1/q} e^{-(\gamma_0/q)\sqrt{n}},$$

by Proposition 4.5. The right hand side can be made arbitrarily small by taking N sufficiently large.

(U3) The proof of [A, Proposition 4.2] gives a distortion constant K which is uniform in the whole $\mathcal{N}$. The constant σ is given by (34), which may be taken uniformly smaller than 1 for every $\varphi \in \mathcal{N}$. By [A, Corollary 3.3] the constant β is uniformly bounded away from zero, as long as α and the open set $\mathcal{N}$ are taken small enough. Finally, the proof of [A, Proposition 3.8] shows that ρ may be taken bounded from below by a constant only depending on α.

Remark 5.1. The following comments are to clarify the presentation of Lemma 2.6 in [V1], they are not used in the present work. We refer the reader to [V1] for the setting and notations. The conclusion of the lemma is contained in Corollary 2.3 of [V1], when r is large enough so that $|J(r-2)| \ll \sqrt{\alpha}$. In particular, it is enough to consider the case when r smaller than $(1/2 + 2\eta)\log(1/\alpha)$ (and larger than $(1/2 - 2\eta)\log(1/\alpha)$, cf. statement of the lemma). The function $k(r)$ defined in page 73 of [V1] can not exceed $M \approx \log(1/\alpha)$. So, the arguments at the end of that page actually prove that either $k(r) \geq \text{const}\, r$ or $k(r) \approx M$. However, under the above restriction on r, the latter possibility also implies $k(r) \geq \text{const}\, r$. In this way, the conclusion of the lemma follows in all the cases.

6 Topological mixing

Now we start the proof of Theorem C. In this section we prove that the maps in $\mathcal{N}$ are topologically mixing. This will then be used to show that these maps are ergodic with respect to Lebesgue measure.

It is convenient to introduce a new coordinate system (σ, y) in $S^1 \times I$, related to the original (θ, x) in the following way. As we have seen, every $\varphi \in \mathcal{N}$ admits an invariant central foliation whose leaves are smooth submanifolds close to line segments $\{\theta = \text{const}\}$. It was also mentioned in the previous section that the critical set of φ may be supposed to coincide with $S^1 \times \{0\}$. By definition, (σ, y) represents the point where the central leaf through the point $(\theta = \sigma, x = 0)$ intersects the circle $S^1 \times \{y\}$. Thus, central leaves correspond to vertical line segments $\{\sigma = \text{const}\}$ in these new coordinates. Moreover, the map φ has the form $\varphi(\sigma, y) = (\tilde{g}(\sigma), \tilde{f}_\sigma(y))$. Since the central foliation is usually not transversely smooth, the map $\tilde{g}(\cdot)$ is only continuous, and $\tilde{f}_\sigma$ depends only continuously on the variable σ. On the other hand, the leaves themselves are at least C^2 (initially they are C^3, but the change of coordinates that brings the critical set to $\{x = 0\}$ is only C^2). This ensures that every $\tilde{f}_\sigma(\cdot)$ is a C^2 map, moreover, it is C^2 close to the one-dimensional quadratic map $q(\cdot)$. Our arguments in the sequel always refer to the coordinates (σ, y).

Recall that the attractor Λ of a map $\varphi \in \mathcal{N}$ is defined as the intersection of all forward images of $S^1 \times I$:

$$\Lambda = \bigcap_{n \geq 0} \varphi^n(S^1 \times I).$$

Lemma 6.1. *Λ coincides with $\varphi^2(S^1 \times I)$, if the interval I is properly chosen.*

Proof. For the map q we have that the interval $J = [q^2(0), q(0)]$ is forward invariant, and $q^3(0)$ is in the interior of J. Then we may take $I \subset (-2, 2)$ slightly larger than J, so that $q(I)$ is contained in the interior of I and $q^2(I) = J$. We fix I once and for all, and reason by perturbation. Using that every $\tilde{f}_\sigma$ is C^2 close to q, and that its critical point is located at $y = 0$, we conclude that the first image $\tilde{f}_\sigma(I)$ is contained in I, and the second one $\tilde{f}_\sigma^2(I)$ coincides with the vertical segment

$$J(\sigma) = \{\tilde{g}^2(\sigma)\} \times [\tilde{f}_\sigma^2(0), \tilde{f}_{\tilde{g}(\sigma)}(0)].$$

By induction, it follows that $\tilde{f}_\sigma^n(I)$ coincides with $J(\tilde{g}^{n-2}(\sigma))$ for every $n \geq 2$. Thus, for any $\tau \in S^1$ and $n \geq 2$,

$$\varphi^n(S^1 \times I) \cap (\{\tau\} \times I) = \bigcup_{\tau'} J(\tau') = \varphi^2(S^1 \times I) \cap (\{\tau\} \times I)$$

where the union is taken over all the τ' in S^1 such that $\tilde{g}^2(\tau') = \tau$. This proves that Λ coincides with $\varphi^2(S^1 \times I)$. $\square$

The main result in this section is the following proposition. Here $\mathcal{R}^*$ is the partition, and $h^*(\cdot)$ is the function defined in the second part of Section 4.

Proposition 6.2. *There is an integer $M = M(\alpha)$ such that for any $S \in \mathcal{R}^*$,*

$$\left|\varphi^{h^*(S)+M}(S)\right| = \Lambda.$$

Proof. The proof is divided into four steps. First we prove that the height of $\varphi^{h^*(S)}(S)$ (the length of its intersection with vertical lines $\{\sigma = \text{const}\}$) is larger than const $\alpha^{1-2\eta}$. Then we show that a vertical segment of length const $\alpha^{1-2\eta}$ becomes a vertical segment with length const $\sqrt{\alpha}$, after a finite number of iterates. In the third step we show that, again after a finite number of iterates, the length of such segments becomes larger than some constant independent of α. In the final step we use the fact that $\tilde{f}$ is close to q, and the map q is topologically mixing, to obtain the result.

Step 1. *There is a constant $\Delta_1 > 0$ such that for every $S \in \mathcal{R}^*$ and $\sigma \in \pi_1(S)$,*

$$\left|\varphi^{h^*(S)}(S_\sigma)\right| \geq \Delta_1 \alpha^{1-2\eta}, \quad \text{where } S_\sigma = S \cap (\{\sigma\} \times I).$$

Here $\pi_1(\sigma, y) = \sigma$. The proof is analogous to [A, Proposition 3.8], using Lemma 4.9 above in the place of [A, Lemma 3.7]. It is at this point that we use the alternative construction described in the second part of Section 4.

Step 2. *There is a constant $\Delta_2 > 0$ and an integer $M_1 = M_1(\alpha)$ such that, given any $\sigma \in S^1$ and any interval $J \subset I$ with $|J| \geq \Delta_1 \alpha^{1-2\eta}$, there is $n \leq M_1$ such that*

$$|\varphi^n(\{\sigma\} \times J)| \geq \Delta_2 \sqrt{\alpha}.$$

We take $\Delta_2 = 1$. Let $R_0 \geq 0$ be the first integer for which $\tilde{f}_\sigma^{R_0}(J)$ intersects $(-\sqrt{\alpha}, \sqrt{\alpha})$. According to Lemma 4.2, the length of the iterates of J grows exponentially fast up to time R_0. In particular, since $|J|$ is bounded from below by a power of α, we must have $R_0 \leq \text{const} \log(1/\alpha)$. If $\tilde{f}_\sigma^{R_0}(J)$ is not contained in $(-2\sqrt{\alpha}, 2\sqrt{\alpha})$ then its length is larger than $\sqrt{\alpha}$, and we may take $n = R_0$. Otherwise, if $\tilde{f}_\sigma^{R_0}(J) \subset (-2\sqrt{\alpha}, 2\sqrt{\alpha})$, Lemma 4.2 gives

$$|\tilde{f}_\sigma^{R_0}(J)| \geq C_2 \tau^{R_0} |J| \geq C_2 |J| \geq C_2 \Delta_1 \alpha^{1-2\eta}.$$

In particular, there exists a subinterval J_1 of $\tilde{f}_\sigma^{R_0}(J)$ such that

$$J_1 \cap \big(-\frac{C_2 \Delta_1}{4} \alpha^{1-2\eta}, \frac{C_2 \Delta_1}{4} \alpha^{1-2\eta}\big) = \emptyset \quad \text{and} \quad |J_1| \geq \frac{C_2}{4} |J|\,.$$

Let $\sigma_1 = \tilde{g}(\sigma)$. Then, by Lemma 4.1, there exists $N \leq \text{const} \log(1/\alpha)$ such that

$$|\tilde{f}_{\sigma_1}^N(J_1)| \geq \frac{C_2 \Delta_1}{4} \alpha^{1-2\eta}\, \alpha^{-1+\eta} |J_1| \geq \frac{C_2^2 \Delta_1}{16}\, \alpha^{-\eta} |J|\,. \tag{35}$$

We assume that α is small enough so that the expression on the right hand side is larger than $2|J|$. Then we may repeat this procedure all over again, with $\sigma_2 = \tilde{g}^{R_0+N}(\sigma)$

in the place of σ, and $J_2 = \tilde{f}_\sigma^{R_0+N}$ in the place of J. In this way, we construct sequences $J_0 = J, J_2, \ldots, J_{2l}$, of vertical segments, $\sigma_0 = \sigma, \sigma_2, \ldots, \sigma_{2l}$ of points in S^1, and $R_0, R_2, \ldots, R_{2l-2}$, of integers, such that

$$|J_{2i+2}| > 2|J_{2i}| \quad \text{and} \quad J_{2i+2} \subset \tilde{f}_{\sigma_{2i}}^{R_{2i}+N}(J_{2i})$$

for every $0 \le i < l$. Because the lengths grow by a factor of 2 at each step, we must eventually reach a situation where $J_{2l+1} = \tilde{f}_{\sigma_{2l}}^{R_{2l}}(J_{2l})$ is not contained in $(-2\sqrt{\alpha}, 2\sqrt{\alpha})$. Here $R_{2l} \ge 0$ is the first iterate so that J_{2l+1} intersects $(-\sqrt{\alpha}, \sqrt{\alpha})$. Then the length of J_{2l+1} is larger than $\sqrt{\alpha}$, and so we may take $n = R_0 + N + R_2 + \cdots + N + R_{2l}$. Observe that $l \le \text{const} \log(1/|J|) \le \text{const} \log(1/\alpha)$, because the lengths grow exponentially fast. Therefore, using that N and the R_i are also bounded by $\text{const} \log(1/\alpha)$, we get that $n \le \text{const} \log^2(1/\alpha)$. So, we may choose $M_1 = \text{const} \log^2(1/\alpha)$.

Step 3. *There is a constant $\Delta_3 > 0$ and an integer $M_2 = M_2(\alpha)$ such that, given any $\sigma \in S^1$ and any interval $J \subset I$ with $|J| \ge \Delta_2\sqrt{\alpha}$, there exists $n \le M_2$ such that*

$$|\varphi^n(\{\sigma\} \times J)| \ge \Delta_3.$$

Arguing as in Step 2 we obtain an analog of (35)

$$|\tilde{f}_\sigma^{R_0+N}(J)| \ge \frac{\Delta_2 C_2}{4} \sqrt{\alpha}\, \alpha^{-1+\eta} \frac{C_2|J|}{4} \ge \frac{C_2^2 \Delta_2^2}{16} \alpha^\eta.$$

Let $R_1 \ge 1$ be the first integer for which $\tilde{f}_\sigma^{R_0+N+R_1}(J)$ intersects $(-\sqrt{\alpha}, \sqrt{\alpha})$. We fix small constants $0 < \delta_1 < \delta_0 < \delta$, *independent of* α, according to conditions that will appear in a little while. If $\tilde{f}_\sigma^{R_0+N+R_1}(J)$ is not contained in $(-\delta_1, \delta_1)$ then its length is larger than $\delta_1 - \sqrt{\alpha} > \delta_1/2$, and our claim is proved. Otherwise, it is contained in $(-\delta, \delta)$, and so we may use Lemma 4.2 to conclude that

$$|\tilde{f}_\sigma^{R_0+N+R_1}(J)| \ge C_2 \tau^{R_1} \frac{C_2^2 \Delta_2^2}{16} \alpha^\eta \ge 4C_3 \alpha^\eta,$$

with $C_3 = C_2^3 \Delta_2^2/64$. Then there is some connected component $\bar{J}$ of the difference $\tilde{f}_\sigma^{R_0+N+R_1}(J) \setminus (-\sqrt{\alpha}, \alpha)$ such that the length of $\bar{J}$ is larger than $2C_3\alpha^\eta - \sqrt{\alpha} > C_3\alpha^\eta$ (α is small and $\eta < 1/2$). In what follows, $\bar{\sigma} = \tilde{g}^{R_0+N+R_1}(\sigma)$. Let us outline how we obtain the claim in the only remaining case, that is, when $\bar{J}$ is contained in $(-\delta_1, \delta_1)$. The detailed argument will follow.

Recall that q was chosen so that the critical point 0 is pre-periodic. Let $l \ge 1$ be the smallest integer for which $z = q^l(0)$ is a periodic point of q. By [S], this periodic point must be repelling. We take δ_1 much smaller than δ_0, so that $q^l(-\delta_1, \delta_1)$ is contained in the $(\delta_0/2)$-neighborhood of z. Then, assuming α is small, $\tilde{f}^l_{\bar{\sigma}}(\bar{J})$ is contained in the δ_0-neighborhood of z. The key observation is that

$$(36) \qquad |\tilde{f}^l_{\bar{\sigma}}(\bar{J})| \geq \text{const}\,\alpha^\eta \quad \text{whereas} \quad \text{dist}(z, \tilde{f}^l_{\bar{\sigma}}(\bar{J})) \leq \text{const}\,\alpha.$$

The last inequality follows from the fact that $\bar{J}$ has $\pm\sqrt{\alpha}$ as a boundary point, which implies that the distance from $\tilde{f}_{\bar{\sigma}}(J)$ to the critical value $\tilde{f}_{\bar{\sigma}}(0)$ is bounded by const α. Since α^η is much larger than α, (36) means that $\tilde{f}^l_{\bar{\sigma}}(\bar{J})$ contains several fundamental domains associated to the periodic point z of q. This property being preserved under iteration, the first iterate of $\tilde{f}^l_{\bar{\sigma}}(\bar{J})$ that is not completely inside the δ_0-neighborhood of z must contain at least one fundamental domain with some boundary point of that neighborhood in it. Now, the length of such a fundamental domain is of order δ_0. So, at this point J and $\bar{J}$ have become larger than some constant independent of α, as claimed.

Now we give the details. Let $k \geq 1$ be the period of z, and $\rho^k = |(q^k)'(z)|$. As already mentioned, by [S] we must have $\rho > 1$. Fix $\rho_1, \rho_2 > 0$ with $\rho_1 < \rho < \rho_2$ and $\rho_1 > \rho_2^{1-\eta/2}$, and take $\delta_0 > 0$ small enough that

$$\rho_1^k < |D\tilde{f}^k_\tau(y)| < \rho_2^k, \quad \text{whenever} \quad |y - z| < \delta_0$$

for any $\tau \in S^1$, and assuming α is sufficiently small. Observe that $q^j(0)$ is never zero, for any $j > 0$. Thus, fixing δ_1 sufficiently small right from the start, we ensure that $\tilde{f}^j_\sigma(y)$ remains outside a fixed neighborhood of zero, for all $0 \leq j \leq l$ and $y \in \bar{J}$. Then,

$$(37) \qquad |D\tilde{f}^l_{\bar{\sigma}}(y)| \geq \text{const}\,|y|,$$

for all $y \in \bar{J}$. Consequently, for some $y \in \bar{J}$,

$$|\tilde{f}^l_{\bar{\sigma}}(\bar{J})| = |D\tilde{f}^l_{\bar{\sigma}}(y)|\,|\bar{J}| \geq \text{const}\,|y|\,C_3\alpha^\eta \geq \text{const}\,\alpha^{1/2+\eta}.$$

For any $y \in \bar{J}$ and $i \geq 0$, let $d_i = |y_{l+ki} - z|$, where $(\sigma_j, y_j) = \varphi(\bar{\sigma}, y)$. As already mentioned, we suppose $\delta_1 > 0$ and α sufficiently small so that

$$|y| < \delta_1 \quad \Rightarrow \quad d_0 \leq Cy^2 + C\alpha < \delta_0.$$

If (σ, y) and $i \geq 1$ are such that $|y| < \delta_1$ and $d_0, \ldots, d_{i-1} < \delta_0$, then $d_i \leq \rho_2^k d_{i-1} + C\alpha$ and so, inductively,

$$d_i \leq (1 + \rho_2^k + \cdots + \rho_2^{k(i-1)})C\alpha + \rho_2^{ki} d_0 \leq \rho_2^{ki}(C\alpha + Cy^2).$$

In particular, for $y = \pm\sqrt{\alpha}$ we have $d_i \le \rho_2^{ki} C\alpha$. Let $N_0 \ge 1$ be the smallest integer for which $\rho_2^{kN_0} C\alpha \ge \delta_0/2$, where C is as in the previous inequality. This choice of N_0 implies

$$d_i < \delta_0/2 \quad \text{for} \quad i = 0, \ldots, N_0 - 1. \tag{38}$$

Now we consider the following alternative cases:

1. Suppose $\tilde{f}_{\bar{\sigma}}^{l+ki}(\bar{J}) \subset (z - \delta_0, z + \delta_0)$ for every $i \in \{0, \cdots, N_0 - 1\}$.

   This implies that (recall that $\eta < 1/4$)

$$\begin{aligned} |\tilde{f}_{\bar{\sigma}}^{l+kN_0}(\bar{J})| &\ge \rho_1^{kN_0} |\tilde{f}_{\bar{\sigma}}^{l}(\bar{J})| \ge \text{const}\, \rho_2^{(1-\eta/2)kN_0} \alpha^{1/2+\eta} \\ &\ge \text{const}\, \alpha^{-1+\eta/2} \alpha^{1/2+\eta} \ge \text{const}\, \alpha^{-1/8} \gg 1. \end{aligned}$$

   So this case can not really happen.

2. There is $i \in \{0, \cdots, N_0 - 1\}$ such that $\tilde{f}_{\bar{\sigma}}^{l+ki}(\bar{J}) \not\subset (z - \delta_0, z + \delta_0)$.

   Since $d_i \le \delta_0/2$, it follows that

$$|\tilde{f}_{\bar{\sigma}}^{l+ki}(\bar{J})| \ge \delta_0 - \delta_0/2 > \delta_1/2.$$

Therefore, have shown that we may take $\Delta_3 = \delta_1/2$, $n = R_0 + N + R_1 + l + ki$, and $M_2 = R_0 + N + R_1 + l + kN_0$. Moreover, $M_2 \le \text{const} \log(1/\alpha)$.

Step 4. *There is an integer M_3 such that if $J \subset I$ is an interval with $|J| \ge \Delta_3$ then, for every $\sigma \in S^1$,*

$$|\varphi^{M_3}(\{\sigma\} \times J)| = (\{\tilde{g}^{M_3}(\sigma)\} \times I) \cap \Lambda.$$

Since the quadratic map q is such that the critical point is pre-periodic, the pre-orbit of the repelling fixed point P is dense in I (this is because q has no wandering intervals, see [MS]). Then there is some integer $n_1 \ge 1$ such that $q^{-n_1}(P)$ intersects every interval of length $\Delta_3/3$. It follows that for every interval $J \subset I$ with $|J| \ge \Delta_3$ the image $q^{n_1}(J)$ contains a neighborhood of P with a definite size, depending only on Δ_3. After a finite number of iterates $n_2 \ge 1$ this neighborhood becomes the whole interval $q^2(I) = [q^2(0), q(0)]$. Let $M_3 = n_1 + n_2 + 1$. Then, by continuity,

$$\varphi^{M_3}(\{\sigma\} \times J) = \{\tilde{g}^{M_3}(\sigma)\} \times J(\tilde{g}^{M_3-2}(\sigma)) = (\{\tilde{g}^{M_3}(\sigma)\} \times I) \cap \Lambda.$$

where $J(\tau) = [\tilde{f}_\tau^2(0), \tilde{f}_{\tilde{g}(\tau)}(0)]$, as in the proof of Lemma 6.1.

So, it suffices to take $M(\alpha) = M_1(\alpha) + M_2(\alpha) + M_3$ to complete the proof of Proposition 6.2. $\square$

Now we may prove that the maps $\varphi \in \mathcal{N}$ are topologically mixing. The idea is to apply the previous proposition to a sequence of partitions with diameters going to zero. Recall that in the definition of $\mathcal{R}^*$ we started by fixing a large integer p, then all the return times $h^*(\cdot)$ were taken larger than p. By (34), the diameter

$$\operatorname{diam}(\mathcal{R}^*) = \sup\{\operatorname{diam}(R) : R \in \mathcal{R}^*\},$$

can be made arbitrarily small by increasing p. Thus, now we allow p to vary, and for each value of p we denote by $\mathcal{R}^*(p)$ the partition and by $h_p^*(\cdot)$ the return time corresponding to each p.

Then, let A be an arbitrary open subset of $S^1 \times I$. Since the diameter of $\mathcal{R}^*(p)$ converges to zero when p goes to infinity, we may find $p \geq 1$ and $S \in \mathcal{R}^*(p)$ such that $S \subset A$. Fix p and S, then take M as in Proposition 6.2. We get that there is $n \leq h_p^*(S) + M$ such that $\varphi^n(A) = \Lambda$. This proves the topological mixing property.

7 Ergodicity

Now we prove that the maps $\varphi \in \mathcal{N}$ are ergodic with respect to Lebesgue measure. We start by proving a few auxiliary results.

Lemma 7.1. *Let B be a Borel subset of $S^1 \times I$ such that $\varphi^{-1}(B) = B$.*

1. *If $m(B \cap \Lambda) = 0$ then $m(B) = 0$.*
2. *If $\varphi^n(R) = \Lambda$ for some $n \geq 1$ and $R \subset S^1 \times I$, then $B \cap \Lambda \subset \varphi^n(B \cap R)$.*

Proof. We have $B = \varphi^{-2}(B) = \varphi^{-2}\left(B \cap \varphi^2(S^1 \times I)\right) = \varphi^{-2}\left(B \cap \Lambda\right)$, because the attractor $\Lambda = \varphi^2(S^1 \times I)$. Since the pre-image of a zero Lebesgue measure set also has zero Lebesgue measure, the first claim is an immediate consequence.

Now suppose $\varphi^n(R) = \Lambda$ and let $x \in B \cap \Lambda$. Then there is some $z \in R$ for which $\varphi^n(z) = x$. Moreover, z is in $\varphi^{-n}(B) = B$. Hence $x \in \varphi^n(B \cap R)$. $\square$

Lemma 7.2. *Let μ be a finite measure on a metric space X, and $\mathcal{P} = \{P_1, \ldots, P_r\}$ be a partition of X into Borel subsets. Assume that $(\mathcal{S}_n)_{n\geq 1}$ are partitions of X such that diam $(\mathcal{S}_n) \to 0$ when $n \to \infty$. Then, for each $n \geq 1$ there is a partition $\{Q_1^n, \ldots, Q_r^n\}$ of X such that for $i = 1, \ldots, r$*

1. *Q_i^n is a union of atoms of $\mathcal{S}_n$.*
2. $\lim_{n\to\infty} \mu(Q_i^n \triangle P_i) = 0$.

Proof. Take an arbitrary $\epsilon > 0$. Since μ is a regular measure, there are compact sets $K_1, \dots, K_r \subset X$ with

$$K_i \subset P_i \quad \text{and} \quad \mu(P_i \setminus K_i) < \epsilon$$

for $i = 1, \dots, m$. Let

$$\delta = \inf_{i \neq j} d(K_i, K_j) > 0$$

and take $n_0 \geq 1$ such that $\mathrm{diam}(\mathcal{S}_n) < \delta/2$ for $n \geq n_0$. For $n < n_0$ the Q_i^n are totally arbitrary. ¿From now on we suppose $n \geq n_0$.

Note that each $S \in \mathcal{S}_n$ intersects at most one K_i. We let Q_i^n be the union of all the elements of $\mathcal{S}_n$ that intersect K_i for each i. Those $S \in \mathcal{S}_n$ that do not intersect any of the K_i are distributed among the Q_i^n, in an arbitrary way. We have

$$\begin{aligned} \mu(Q_i^n \triangle P_i) &= \mu(Q_i^n \setminus P_i) + \mu(P_i \setminus Q_i^n) \\ &\leq \mu(X \setminus \cup_{i=1}^r K_i) + \mu(P_i \setminus K_i) \ \leq \ (r+1)\epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary and r is fixed, we have proved the result. □

Corollary 7.3. *Let μ and $\mathcal{S}_n$, $n \geq 1$, be as in Lemma 7.2, and B be a Borel subset of X with $\mu(B) > 0$. Then, for every $\epsilon > 0$ there is an integer $n_\epsilon \geq 1$ such that for each $n \geq n_\epsilon$ there is some $S \in \mathcal{S}_n$ with*

$$\mu(B^c \cap S) < \epsilon\mu(S).$$

Proof. Assume by contradiction that there are $\epsilon_0 > 0$ and a sequence of integers $(n_k)_{k\geq 1}$ going to $+\infty$ such that

$$\mu(B^c \cap S_{n_k}) \geq \epsilon_0 \mu(S_{n_k}). \tag{39}$$

for all $k \geq 1$ and all $S_{n_k} \in \mathcal{S}_{n_k}$. We know, from Lemma 7.2, that for every $k \geq 1$ there is a partition $\{Q_1^{n_k}, Q_2^{n_k}\}$ of X such that $Q_1^{n_k}$, $Q_2^{n_k}$ are unions of atoms of $\mathcal{S}_{n_k}$, and

$$\lim_{k\to+\infty} \mu(Q_1^{n_k} \triangle B) = 0 \quad \text{and} \quad \lim_{k\to+\infty} \mu(Q_2^{n_k} \triangle B^c) = 0.$$

Take $\epsilon = \epsilon_0\mu(B)/(1+\epsilon_0)$ and $k_0 \geq 1$ sufficiently large so that

$$\mu(Q_1^{n_{k_0}} \triangle B) < \epsilon \quad \text{and} \quad \mu(Q_2^{n_{k_0}} \triangle B^c) = \epsilon. \tag{40}$$

Writing $\mathcal{S}_{n_{k_0}} = \{S_i\}_{i\in\mathbb{N}}$, then there is some $\mathbb{I} \subset \mathbb{N}$ for which

$$Q_1^{n_{k_0}} = \bigcup_{i\in\mathbb{I}} S_i \quad \text{and} \quad Q_2^{n_{k_0}} = \bigcup_{i\in\mathbb{N}\setminus\mathbb{I}} S_i\,.$$

¿From (39) we have, in particular, $\mu(B^c \cap S_i) \geq \epsilon_0 \mu(S_i)$ for every $i \in \mathbb{I}$. So, summing over all $i \in \mathbb{I}$, we find

$$\mu(B^c \cap Q_1^{n_{k_0}}) \geq \epsilon_0 \mu(Q_1^{n_{k_0}}). \tag{41}$$

Finally, from (40) and (41) we get

$$\epsilon > \mu(B^c \cap Q_1^{n_{k_0}}) \geq \epsilon_0 \mu(Q_1^{n_{k_0}}) \geq \epsilon_0 \left(\mu(B) - \mu(B \setminus Q_1^{n_{k_0}})\right) > \epsilon_0(\mu(B) - \epsilon),$$

which contradict our choice of ϵ. $\square$

Now we are in a position to prove the ergodicity of the maps $\varphi \in \mathcal{N}$ with respect to the Lebesgue measure. Let B be a Borel subset of $S^1 \times I$ with $\varphi^{-1}(B) = B$ and positive Lebesgue measure. We need to prove that the Lebesgue measure of $B^c = (S^1 \times I) \setminus B$ is equal to zero. By the first part of Lemma 7.1, it suffices to prove that $m(B^c \cap \Lambda) = 0$. We use the partitions $(\mathcal{R}^*(p))_p$ introduced in the previous section. Recall that the diameter of $\mathcal{R}^*(p)$ goes to zero as $p \to \infty$. Take any $\epsilon > 0$ small. By Corollary 7.3, there are $p \geq 1$ and $S \in \mathcal{R}(p)$ for which

$$m(B^c \cap S) < \epsilon m(S).$$

Let p and S be fixed and $h = h_p^*(S)$. According to Proposition 6.2, we have $\varphi^{h+M}(S) = \Lambda$. Thus, using the second part of Lemma 7.1,

$$m(B^c \cap \Lambda) \leq m\left(\varphi^{h+M}(B^c \cap S)\right).$$

By Proposition 4.7, we have

$$m\left(\varphi^h(B^c \cap S)\right) = \int_{B^c \cap S} |J(\theta, x)|\, dm(\theta, x) \leq \Delta \left|J_{n_\epsilon}(\theta_0, x_0)\right| m\left(B^c \cap S\right),$$

for any $(\theta_0, x_0) \in S$. Recall, from Remark 4.8, that the distortion bound Δ does not depend on p. Similarly,

$$m\left(\varphi^h(S)\right) \geq \frac{1}{\Delta} \left|J(\theta_0, x_0)\right| m\left(S\right).$$

Hence

$$m\left(\varphi^h(B^c \cap S)\right) \leq \frac{m\left(\varphi^h(B^c \cap S)\right)}{m\left(\varphi^h(S)\right)} \leq \Delta^2 \frac{m\left(B^c \cap S\right)}{m\left(S\right)} \leq \Delta^2 \epsilon.$$

On the other hand,

$$m\left(\varphi^{h+M}(B^c \cap S)\right) \leq 4^M (d + \alpha)^M m\left(\varphi^h(B^c \cap S)\right).$$

Altogether, this shows that $m(B^c \cap \Lambda) \leq 4^M (d + \alpha)^M \Delta^2 \epsilon$. Since M is fixed and ϵ is arbitrarily small, $m(B^c \cap \Lambda)$ must be zero. So the proof of Theorem C is complete.

References

[A] J. F. Alves, *SRB measures for non-hyperbolic systems with multidimensional expansion*, Ann. Sci. de l'ENS 33 (2000), 1–32.

[ABV] J. F. Alves, C. Bonatti, M. Viana, *SRB measures for partially hyperbolic systems with mostly expanding central direction*, Invent. Math. 140 (2000), 351-398.

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

[BV] C. Bonatti, M. Viana, *SRB measures for partially hyperbolic systems with mostly contracting central direction*, Israel J. Math. 115 (2000), 157-193.

[D] D. Dolgopyat, *On dynamics of mostly contracting diffeomorphisms*, preprint 1998.

[G] E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkäuser Verlag, Basel-Boston, Mass., 1984.

[HPS] M. Hirsch, C. Pugh, M. Shub, *Invariant manifolds*, Lect. Notes in Math. 583, Springer Verlag, 1977.

[LY] A. Lasota and J.A. Yorke, *On the existence of invariant measures for piecewise monotonic maps*, Trans. Amer. Math. Soc. **186** (1973), 481-488.

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

[PS] J. Palis and S. Smale, *Structural stability theorems*, in *Global Analysis*, Proc. Sympos. Pure Math. XIV (Berkeley 1968), Amer. Math. Soc., 223–232, 1970.

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

[S] D. Singer, *Stable orbits and bifurcations of maps of the interval*, SIAM J. Appl. Math. **35** (1978), 260-267.

[V1] M. Viana, *Multidimensional nonhyperbolic attractors*, Publ. Math. IHES **85** (1997), 63-96.

[V2] M. Viana, *Stochastic dynamics of deterministic systems*, Lect. Notes XXI Braz. Math Colloq., IMPA, 1997.

José Ferreira Alves
Centro de Matemática da Universidade do Porto
Praça Gomes Teixeira, 4099-002 Porto, Portugal
jfalves@fc.up.pt

Marcelo Viana
IMPA, Estrada Dona Castorina 110
22460-320 Rio de Janeiro, Brazil
viana@impa.br