J. Inst. Math. Jussieu (2021) **20**(1), 305–329 305
doi:10.1017/S1474748019000185

STABLE SETS OF CERTAIN NON-UNIFORMLY HYPERBOLIC HORSESHOES HAVE THE EXPECTED DIMENSION

CARLOS MATHEUS[1], JACOB PALIS[2] AND JEAN-CHRISTOPHE YOCCOZ[3]

[1] *Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539), F-93439, Villetaneuse, France* (matheus@impa.br)

[2] *IMPA, Estrada D. Castorina, 110, CEP 22460-320, Rio de Janeiro, RJ, Brazil* (jpalis@impa.br)

[3] *Collège de France, 3, Rue d'Ulm, Paris, CEDEX 05, France* (jean-c.yoccoz@college-de-france.fr)

(Received 9 August 2018; revised 6 March 2019; accepted 8 March 2019; first published online 4 April 2019)

Abstract We show that the stable and unstable sets of non-uniformly hyperbolic horseshoes arising in some heteroclinic bifurcations of surface diffeomorphisms have the value conjectured in a previous work by the second and third authors of the present paper. Our results apply to first heteroclinic bifurcations associated with horseshoes with Hausdorff dimension <22/21 of conservative surface diffeomorphisms.

Keywords: heteroclinic bifurcations; non-uniformly hyperbolic horseshoes; Hausdorff dimension

2010 *Mathematics subject classification*: Primary 37C29; 37E30
Secondary 28D20

Contents

1. Introduction

In 2009, the second and third authors of the present paper proved in [5] that the semi-local dynamics of first heteroclinic bifurcations associated with 'slightly thick' horseshoes of surface diffeomorphisms usually can be described by the so-called *non-uniformly hyperbolic horseshoes*.

In this article, we pursue the studies of Palis–Yoccoz [5] and Matheus–Palis [3] on the Hausdorff dimensions of the stable and unstable sets of non-uniformly hyperbolic horseshoes.

In order to state our main result (Theorem 1.4), first, we need to recall the setting of the work by Palis–Yoccoz [5].

1.1. Heteroclinic bifurcations in Palis–Yoccoz regime

Fix a smooth diffeomorphism $g_0 : M \to M$ of a compact surface M. Assume that p_s and p_u are periodic points of g_0 in *distinct* orbits such that $W^s(p_s)$ and $W^u(p_u)$ meet *tangentially* and *quadratically* at some point q. Suppose that K is a horseshoe of g_0 such that $p_s, p_u \in K$ and $q \in M \setminus K$, and, for some neighborhoods[1] U of K and V of the orbit $\mathcal{O}(q)$, the maximal invariant set of $U \cup V$ is $K \cup \mathcal{O}(q)$. In summary, g_0 has a first heteroclinic tangency at q associated with periodic points p_s, p_u of a horseshoe K.

Let $(g_t)_{|t|<t_0}$ be a one-parameter family of smooth diffeomorphisms of M *generically* unfolding the first heteroclinic tangency of g_0 described in the previous paragraph. Assume that the continuations of $W^s(p_s)$ and $W^u(p_u)$ have no intersection near q for $-t_0 < t < 0$ but have two transverse intersections near q for $0 < t < t_0$.

Denote by $K_{g_t} := \bigcap_{n\in\mathbb{Z}} g_t^{-n}(U)$ the hyperbolic continuation of K. In our context, it is not hard to describe the maximal invariant set

$$\Lambda_{g_t} := \bigcap_{n\in\mathbb{Z}} g_t^{-n}(U \cup V) \tag{1.1}$$

in terms of K_{g_t} *when* $-t_0 < t \leqslant 0$; indeed, $\Lambda_{g_t} = K_{g_t}$ when $-t_0 < t < 0$ and $\Lambda_{g_0} = K \cup \mathcal{O}(q)$.

On the other hand, the study of Λ_{g_t} for $0 < t < t_0$ represents an important challenge when the Hausdorff dimension of the initial horseshoe $K = K_{g_0}$ is larger than one.

[1]It is shown in Appendix A that it is often the case that the particular choices of U and V are not very relevant.

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

In their paper [5], Palis and Yoccoz studied *strongly regular parameters* $0 < t < t_0$ whenever K_{g_0} is *slightly thick*, i.e.,

$$(d_s^0 + d_u^0)^2 + (\max\{d_s^0, d_u^0\})^2 < (d_s^0 + d_u^0) + \max\{d_s^0, d_u^0\}, \tag{1.2}$$

where d_s^0 and d_u^0 (respectively) are the transverse Hausdorff dimensions of the invariant sets $W^s(K_{g_0})$ and $W^u(K_{g_0})$ (respectively). In this setting, Palis and Yoccoz proved that any strongly regular parameter t has the property that Λ_{g_t} is a *non-uniformly hyperbolic horseshoe*, and, moreover, the strongly regular parameters are abundant near $t = 0$:

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathrm{Leb}_1(\{0 < t < \varepsilon : t \text{ is a strongly regular parameter}\}) = 1.$$

(Here Leb_1 is the one-dimensional Lebesgue measure.)

Remark 1.1. This result of Palis and Yoccoz is a semi-local dynamical result; indeed, Appendix A (by C. G. Moreira and the first two authors of this paper) shows that it is often the case that $U \cup V$ can be chosen to be of almost full Lebesgue measure.

We refer the reader to the original paper [5] for the precise definitions of *strongly regular parameters* and *non-uniformly hyperbolic horseshoes*. For the purpose of this article, we will discuss *some* features of non-uniformly hyperbolic horseshoes in Section 2.

For the time being, we recall only that non-uniformly hyperbolic horseshoes are *saddle-like* sets.

Theorem 1.2 (Cf. [5, Theorem 6] and [3, Theorem 1.2]). *Under the previous assumptions, if t is a strongly regular parameter, then*

$$\mathrm{HD}(W^s(\Lambda_{g_t})) < 2 \quad \text{and} \quad \mathrm{HD}(W^u(\Lambda_{g_t})) < 2,$$

where HD *stands for the Hausdorff dimension. In particular, Λ_{g_t} does not contain attractors or repellers.*

As it turns out, this result leaves open the *exact* calculation of the quantities $\mathrm{HD}(W^s(\Lambda_{g_t}))$; in fact, Palis and Yoccoz conjectured in [5, p. 14] that the stable sets of non-uniformly hyperbolic horseshoes have Hausdorff dimensions very close or perhaps equal to the *expected* dimension $1 + d_s$, where d_s is a certain number close to d_s^0 measuring the transverse dimension of the stable set of the 'main non-uniformly hyperbolic part' of Λ_{g_t}.

Remark 1.3. The proof of Theorem 1.2 *never* allows to show that $W^s(\Lambda_{g_t})$ has the expected dimension; see [3, Remark 8].

In this article, we give the following (partial) answer to this conjecture.

1.2. Statement of the main theorem

We show that the conjecture stated above is true *at least* when the transverse dimensions d_s^0 and d_u^0 of the stable and unstable sets of the initial horseshoe K_{g_0} satisfy a *stronger* constraint than (1.2).

Theorem 1.4. *In the same setting of Theorem 1.2, denote*

$$\beta^*(d_s^0, d_u^0) := \frac{(1-\min\{d_s^0, d_u^0\})(d_s^0+d_u^0)}{\max\{d_s^0, d_u^0\}(\max\{d_s^0, d_u^0\}+d_s^0+d_u^0-1)}.$$

In addition to (1.2) (i.e., $\beta^*(d_s^0, d_u^0) > 1$), let us also assume that the transverse dimensions d_s^0 and d_u^0 of the stable and unstable sets of the initial horseshoe K_{g_0} satisfy

$$\beta^*(d_s^0, d_u^0) \leqslant 1+\min\left\{\frac{-\log|\lambda(p_s)|}{\log|\mu(p_s)|}, \frac{\log|\mu(p_u)|}{-\log|\lambda(p_u)|}\right\}, \tag{1.3}$$

where $\lambda(p_\alpha)$ and $\mu(p_\alpha)$ are the stable and unstable eigenvalues of the periodic point p_α, respectively, for $\alpha = s, u$, and

$$\beta^*(d_s^0, d_u^0) > \tfrac{5}{3}. \tag{1.4}$$

Then, for any strongly regular parameter t, one has

$$\mathrm{HD}(W^s(\Lambda_{g_t})) = 1 + d_s,$$

where $0 < d_s = d_s(g_t) < 1$ is a certain quantity close to d_s^0 given by the transverse dimensions of the lamination $\widetilde{\mathcal{R}}_+^\infty$ of stable curves associated with the well-behaved parts of Λ_{g_t} (see [5, pages 12, 13 and 14]).

Remark 1.5. Analogously to [5], there is a symmetry between past and future in our arguments. Thus, the analog of Theorem 1.4 for the unstable set $W^u(\Lambda_{g_t})$ holds after exchanging the roles of d_s^0 and d_u^0.

Remark 1.6. Of course, we believe that conditions (1.3) and (1.4) are not necessary for the validity of the conclusion of Theorem 1.4, but our proof of this result in Section 3 does not allow us to bypass these technical conditions. We hope to come back to this issue in the future.

Remark 1.7. Condition (1.3) is automatic in the *conservative* case (when g_0 preserves a smooth area form). Indeed, the multipliers $\lambda(p_\alpha)$ and $\mu(p_\alpha)$ verify $\lambda(p_\alpha)\mu(p_\alpha) = 1$ in this situation, so that (1.3) becomes the requirement $\beta^*(d_s^0, d_u^0) \leqslant 2$ which is always true when $d_s^0 + d_u^0 > 1$.

Similarly, condition (1.3) is automatic if K_{g_0} is a product of two affine Cantor sets K^s and K^u of the real line obtained from affine maps with constant dilatations $1/\lambda > 1$ and $\mu > 1$ sending two finite collections of $\ell \in \mathbb{N}$ disjoint closed subintervals of $[0, 1]$ surjectively on their convex hull $[0, 1]$. In fact, it is well known that the transverse Hausdorff dimensions of such a horseshoe K_{g_0} are $d_s^0 = \log\ell/\log(1/\lambda)$ and $d_u^0 = \log\ell/\log\mu$, so that requirement (1.3) becomes

$$\beta^*(d_s^0, d_u^0) \leqslant 1+\min\left\{\frac{\log(1/\lambda)}{\log\mu}, \frac{\log\mu}{\log(1/\lambda)}\right\} = 1+\min\left\{\frac{d_s^0}{d_u^0}, \frac{d_u^0}{d_s^0}\right\}$$
$$= \frac{d_s^0+d_u^0}{\max\{d_s^0, d_u^0\}}$$

which is always valid when $d_s^0 + d_u^0 > 1$.

In summary, it is 'often' the case that condition (1.3) is *less* restrictive than condition (1.4) in 'many' applications of Theorem 1.4.

Remark 1.8. A natural question closely related to the statement of Theorem 1.4 is: given a strongly regular parameter t, what is the Hausdorff dimension of the non-uniformly hyperbolic horseshoe Λ_{g_t} *itself*? Of course, it is reasonable to conjecture that a non-uniformly hyperbolic horseshoe Λ_{g_t} has the 'expected' dimension $\mathrm{HD}(\Lambda_{g_t}) = d_s + d_u$. In this direction, let us observe that Theorem 1.4 implies only that $\mathrm{HD}(\Lambda_{g_t}) \leqslant 1 + \min\{d_s, d_u\}$ (since $\Lambda_{g_t} = W^s(\Lambda_{g_t}) \cap W^u(\Lambda_{g_t})$), but this is still far from the 'expected' value (as $d_s + d_u < 1 + \min\{d_s, d_u\}$). We plan to address elsewhere the question of computing $\mathrm{HD}(\Lambda_{g_t})$ for strongly regular parameters t.

For the sake of comparison[2] of conditions (1.2) and (1.4), we plot below (using *Mathematica*) the portions of the regions

$$\mathcal{D} = \{(d_s^0, d_u^0) \in [0, 1] \times [0, 1] : (d_s^0, d_u^0) \text{ satisfies } (1.4)\}$$

and

$$\mathcal{PY} = \{(d_s^0, d_u^0) \in [0, 1] \times [0, 1] : (d_s^0, d_u^0) \text{ satisfies } (1.2)\}$$

below[3] the diagonal $\Delta = \{(d_s^0, d_u^0) \in \mathbb{R}^2 : 1/2 < d_s^0 = d_u^0 < 1\}$ (cf. Figure 1).

We have that $\mathcal{D}$ occupies slightly more than 3% of $\mathcal{PY}$:

$$\frac{\mathrm{area}(D)}{\mathrm{area}(\mathcal{PY})} = 0.030136\ldots.$$

These regions intersect the diagonal segment

$$\Delta = \{(d_s^0, d_u^0) \in \mathbb{R}^2 : 1/2 < d_s^0 = d_u^0 < 1\}$$

along

$$\mathcal{PY} \cap \Delta = \{(d_s^0, d_u^0) \in \mathbb{R}^2 : 1/2 < d_s^0 = d_u^0 < 3/5\}$$

and

$$\mathcal{D} \cap \Delta = \{(d_s^0, d_u^0) \in \mathbb{R}^2 : 1/2 < d_s^0 = d_u^0 < 11/21\}.$$

Remark 1.9. The symmetry in Remark 1.5 actually implies that $\mathrm{HD}(W^s(g_t)) = 1 + d_s$ and $\mathrm{HD}(W^u(g_t)) = 1 + d_u$ if t is a strongly regular parameter, (d_s^0, d_u^0) satisfies (1.3) and (d_s^0, d_u^0) belongs to the region $\mathcal{D}$.

1.3. Outline of the proof of the main result

Recall from the Palis–Yoccoz paper [5] that the stable set $W^s(\Lambda)$ of a non-uniformly hyperbolic horseshoe Λ can be written as the disjoint union of an *exceptional part* $\mathcal{E}^+$ and a lamination with $C^{1+\mathrm{Lip}}$-leaves, Lipschitz holonomy, and transverse Hausdorff dimension $0 < d_s < 1$ close to the stable dimension d_s^0 of the initial horseshoe K_{g_0}.

[2] In view of Remark 1.7, we can 'ignore' (1.3) (in some examples) when trying to compare the restrictions imposed in Theorems 1.2 and 1.4.

[3] The other portion is obtained by reflection along the diagonal.

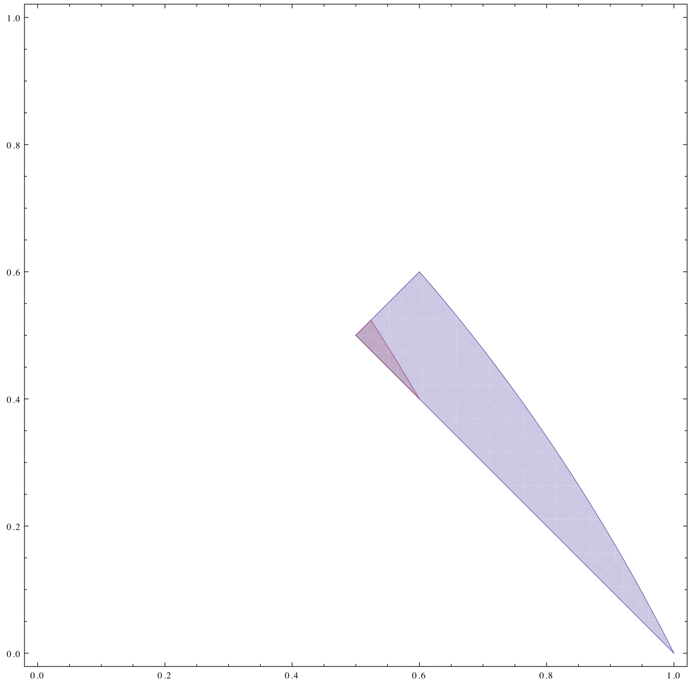

Figure 1. $\mathcal{D}$ (in red) inside $\mathcal{PY}$ (in blue).

Hence, the proof of Theorem 1.4 is reduced to show that the Hausdorff dimension of the exceptional part $\mathcal{E}^+$ of $W^s(\Lambda)$ is $\mathrm{HD}(\mathcal{E}^+) < 1 + d_s$.

By definition, the points of $\mathcal{E}^+$ visit a sequence of 'strips' $(P_k)_{k\in\mathbb{N}}$ whose 'widths' decay *doubly exponentially* fast (cf. [5, Lemma 24]). In particular, by fixing $k \in \mathbb{N}$ large and by decomposing the strip P_k into squares, we obtain a covering of very small diameter of the *image* of $\mathcal{E}^+$ under some positive iterate of the dynamics.

It was shown in [3] that the covering of the images of $\mathcal{E}^+$ in the previous paragraph can be used to prove that $\mathrm{HD}(\mathcal{E}^+) < 2$. More concretely, the negative iterates of the dynamics take the covering of P_k back to $\mathcal{E}^+$ while alternating between affine-like (hyperbolic) iterates and a fixed folding map. In principle, the folding effect accumulates very quickly, but if we *ignore* the action of folding map by replacing all 'parabolic shapes' by 'fat strips', then we obtain a cover of $\mathcal{E}^+$ with small diameter and *controlled* cardinality thanks to the double exponential decay of P_j's. As it turns out, this suffices to establish $\mathrm{HD}(\mathcal{E}^+) < 2$, but this strategy does *not* yield $\mathrm{HD}(\mathcal{E}^+) < 1 + d_s$ (cf. Remark 1.3).

For this reason, in the proof of Theorem 1.4, we do *not* completely ignore the 'parabolic shapes' mentioned above. In fact, we estimate the contribution of the parabolic shapes inside the P_j's to the Hausdorff dimension of $\mathcal{E}^+$ in terms of the derivative and Jacobian of the dynamics thanks to an *analytical* lemma (cf. Lemma 3.4) saying that the Hausdorff measure of scale 1 of the image $f(\mathbb{D}^2)$ of the unit disk $\mathbb{D}^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\}$ under a C^1-map $f : \mathbb{D}^2 \to \mathbb{R}^2$ is bounded *by interpolation* of the C^0-norms of the derivative and Jacobian of f. Also, we prove that this estimate is sufficient to derive $\mathrm{HD}(\mathcal{E}^+) < 1 + d_s$ when the double exponential rate of decays of widths of P_j's is adequate (namely, (1.4) holds). Furthermore, we prove the analytical lemma by decomposing dyadically $f(\mathbb{D}^2)$ and by interpreting the d-Hausdorff measure of $f(\mathbb{D}^2)$ as a L^d-norm. In this way, for $1 < d < 2$, we can estimate this L^d-norm by interpolation between certain L^1 and L^2 norms that are naturally controlled by the derivatives and Jacobians of f.

In summary, the novelty in the proof of Theorem 1.4 (in comparison with Theorem 1.2) is the application of the analytical lemma described above to control the Hausdorff measure of $\mathcal{E}^+$.

Remark 1.10. Similar to [3], our main result holds for the *same* strongly regular parameters of Palis–Yoccoz [5].

Remark 1.11. The arguments outlined above provide sequences of good coverings of the stable and unstable sets $W^s(\Lambda)$ and $W^u(\Lambda)$ permitting to calculate their Hausdorff dimensions. However, in relation with Remark 1.8, let us observe that it is not obvious how to *combine* these sequences to produce good coverings of the non-uniformly hyperbolic horseshoe Λ *itself* allowing to compute its Hausdorff dimension. In fact, the naive idea of taking intersections of elements of coverings of $W^s(\Lambda)$ and $W^u(\Lambda)$ in order to produce a cover of $\Lambda = W^s(\Lambda) \cap W^u(\Lambda)$ does not work *directly* because of the possible 'lack of transversality' (especially near $\mathcal{E}^+ \cap \mathcal{E}^-$) that allows for a potentially bad geometry of such coverings of Λ.

1.4. Organization of the paper

We divide the remainder of this article into two parts. In Section 2, we recall some facts from the Palis–Yoccoz article [5]. In Section 3, we prove an analytical lemma (cf. Lemma 3.4) about Hausdorff measures of planar sets and we apply it to get Theorem 1.4.

2. Preliminaries

In this section, we review some basic properties of the non-uniformly horseshoes introduced in [5] (see also [3, Section 2]).

2.1. Strongly regular parameters

Let $0 < \varepsilon_0 \ll \tau \ll 1$ be two very small constants, and define a sequence of scales $\varepsilon_{k+1} = \varepsilon_k^{1+\tau}$, $k \in \mathbb{N}$. The inductive scheme in [5] defining the *strongly regular parameters* goes as follows. The initial *candidate* interval is $I_0 = [\varepsilon_0, 2\varepsilon_0]$. The kth step of induction consists in dividing the selected candidate intervals of the previous step into $\lfloor \varepsilon_k^{-\tau} \rfloor$ disjoint candidates of lengths ε_{k+1}. These new candidates are submitted to a *strong regularity test* and we select for the $(k+1)$th step of induction *only* the candidates passing this test.

By definition, $t \in I_0 = [\varepsilon_0, 2\varepsilon_0]$ is a *strongly regular parameter* whenever $\{t\} = \bigcap_{k \in \mathbb{N}} I_k$, where $I_0 \supset \cdots \supset I_k \supset \ldots$ are selected candidate intervals.

The strong regularity tests are relevant for two reasons (at least). First, they are rich enough to ensure several nice properties of 'non-uniform hyperbolicity' of Λ_{g_t} for strongly regular parameters $t \in I_0$. Second, they are sufficiently flexible to allow the presence of many strongly regular parameters; by [5, Corollary 15], the set of strongly regular parameters $t \in I_0 = [\varepsilon_0, 2\varepsilon_0]$ has Lebesgue measure $\geqslant \varepsilon_0(1 - 3\varepsilon_0^{\tau^2})$.

The notion of strong regularity tests is intimately related to an adequate class $\mathcal{R}(I)$ of *affine-like iterates* attached to each candidate interval I.

In the next three subsections, we briefly recall the construction of $\mathcal{R}(I)$.

2.2. Semi-local dynamics of heteroclinic bifurcations

We fix *geometrical Markov partitions* of the horseshoes K_{g_t} depending smoothly on g_t. In other terms, we choose a finite system of smooth charts $I^s_a \times I^u_a \to R_a \subset M$ indexed by a finite alphabet $a \in \mathcal{A}$ with the properties that these charts depend smoothly on g_t, the intervals I^s_a and I^u_a are compact, the rectangles R_a are disjoint, the horseshoe K_{g_t} is the maximal invariant set of the interior of $R := \bigcup_{a\in\mathcal{A}} R_a$ and the family $(K_{g_t} \cap R_a)_{a\in\mathcal{A}}$ is a Markov partition of K_{g_t}. Moreover, we assume that no rectangle meets the orbits of p_s and p_u at the same time.

Remark 2.1. The intervals $I^s_a = [x^-_a, x^+_a]$, $I^u_a = [y^-_a, y^+_a]$, $a \in \mathcal{A}$, above can be replaced by slightly *larger* intervals $J^s_a = [x^-_a - C^{-1}, x^+_a + C^{-1}]$, $J^u_a = [y^-_a - C^{-1}, y^+_a + C^{-1}]$ (where $C = C(g_0) \geqslant 1$ is a large constant) without changing any of the properties in the previous paragraph. This fact will be used later during the discussion of affine-like iterates.

The Markov partition $(K_{g_t} \cap R_a)_{a\in\mathcal{A}}$ allows to topologically conjugate the dynamics of g_t on K_{g_t} and the subshift of finite type of $\mathcal{A}^{\mathbb{Z}}$ with transitions

$$\mathcal{B} := \{(a, a') \in \mathcal{A}^2 : g_0(R_a) \cap R_{a'} \cap K_{g_0} \neq \emptyset\}.$$

Furthermore, for each g_t with $t > 0$, we have a compact lenticular region $L_u \subset R_{a_u}$ (near the initial heteroclinic tangency point $q \in M \setminus K$ of g_0) bounded by a piece of the unstable manifold of p_u and a piece of the stable manifold of p_s. Moreover, L_u moves outside R for $N_0 - 1$ iterates of g_t before entering R (for some integer $N_0 = N_0(g_0) \geqslant 2$) because no rectangle meets both orbits of p_s and p_u. The image $L_s = g_t^{N_0}(L_u)$ of L_u under $G := g_t^{N_0}|_{L_u}$ defines another lenticular region $L_s \subset R_{a_s}$ and the regions $g^i(L_u)$, $0 \leqslant i \leqslant N_0$ are called *parabolic tongues*.

Let $\widehat{R} := R \cup \bigcup_{0<i<N_0} g^i(L_u)$. By definition, the set Λ_g introduced in (1.1) is the maximal invariant set of $\widehat{R}$, i.e., $\Lambda_{g_t} = \bigcap_{n\in\mathbb{Z}} g_t^{-n}(\widehat{R})$.

The dynamics of g_t on $\widehat{R}$ is driven by the transition maps

$$g_{aa'} = g_t|_{R_a \cap g^{-1}(R_{a'})} : R_a \cap g_t^{-1}(R_{a'}) \to g_t(R_a) \cap R_{a'}, \quad (a, a') \in \mathcal{B},$$

related to the Markov partition R and the folding map $G = g_t^{N_0}|_{L_u} : L_u \to L_s$ between the parabolic tongues.

Qualitatively speaking, the transitions $g_{aa'}$ correspond to 'affine' hyperbolic maps; for our choices of charts, $g_{aa'}$ contracts 'almost vertical' directions and expands 'almost horizontal' directions. Of course, this hyperbolic structure can be destroyed by the folding map G, and this phenomenon is the source of non-hyperbolicity of Λ_{g_t}.

For this reason, the notion of non-uniformly hyperbolic horseshoes is defined in [5] in terms of a certain 'affine-like' iterate of g_t. Before entering into this discussion, let us quickly overview the notion of *affine-like maps*.

2.3. Generalities on affine-like maps

Let I_0^s, I_0^u, I_1^s and I_1^u be compact intervals with coordinates x_0, y_0, x_1 and y_1. A diffeomorphism F from a *vertical strip*

$$P := \{(x_0, y_0) : \varphi^-(y_0) \leqslant x_0 \leqslant \varphi^+(y_0)\} \subset I_0^s \times I_0^u$$

onto a *horizontal strip*

$$Q := \{(x_1, y_1) : \psi^-(x_1) \leqslant y_1 \leqslant \psi^+(x_1)\} \subset I_1^s \times I_1^u$$

is *affine-like* whenever the projection from the graph of F to $I_0^u \times I_1^s$ is a diffeomorphism onto $I_0^u \times I_1^s$.

By definition, an affine-like map F has an *implicit representation* (A, B), i.e., there are smooth maps A and B on $I_0^u \times I_1^s$ such that $F(x_0, y_0) = (x_1, y_1)$ if and only if $x_0 = A(y_0, x_1)$ and $y_1 = B(y_0, x_1)$.

For our purposes, we shall consider *exclusively* affine-like maps satisfying a *cone condition* and a *distortion estimate*. More concretely, let $\lambda > 1$, $u_0 > 0$, $v_0 > 0$ with $1 < u_0 v_0 \leqslant \lambda^2$ and $D_0 > 0$ be the constants fixed in [5, page 32]; their choices depend solely on g_0.

An affine-like map $F(x_0, y_0) = (x_1, y_1)$ with implicit representation (A, B) satisfies a (λ, u, v) *cone condition* if

$$\lambda|A_x| + u_0|A_y| \leqslant 1 \quad \text{and} \quad \lambda|B_y| + v_0|B_x| \leqslant 1,$$

where A_x, A_y, B_x, B_y are the first-order partial derivatives of A and B. Also, an affine-like map $F(x_0, y_0) = (x_1, y_1)$ with implicit representation (A, B) satisfies a $2D_0$ *distortion condition* if

$$\partial_x \log|A_x|, \partial_y \log|A_x|, A_{yy}, \partial_y \log|B_y|, \partial_x \log|B_y|, B_{xx}$$

are uniformly bounded by $2D_0$.

Remark 2.2. The widths of the domain P and the image Q of an affine-like map $F : P \to Q$ with implicit representation (A, B) are

$$|P| := \max|A_x| \quad \text{and} \quad |Q| := \max|B_y|.$$

The widths satisfy $|P| \leqslant C \min|A_x|$ and $|Q| \leqslant C \min|B_y|$, where $C = C(g_0) \geqslant 1$.

The transitions $g_{aa'}$ associated with the Markov partition R of the horseshoe K_{g_t} are affine-like maps satisfying the cone and distortion conditions with parameters $(\lambda, u_0, v_0, 2D_0)$; see [5, Section 3.4].

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

Moreover, we can build new affine-like maps using the so-called *simple* and *parabolic* compositions of two affine-like maps.

Given compact intervals I_j^s, I_j^u, $j = 0, 1, 2$, and two affine-like maps $F : P \to Q$ and $F' : P' \to Q'$ with domains $P \subset I_0^s \times I_0^u$ and $P' \subset I_1^s \times I_1^u$ and images $Q \subset I_1^s \times I_1^u$ and $Q' \subset I_2^s \times I_2^u$ satisfying the (λ, u_0, v_0) cone condition, the map $F'' = F' \circ F$ from $P'' = P \cap F^{-1}(P')$ to $Q'' = Q \cap P'$ is an affine-like map satisfying the (λ^2, u_0, v_0) cone condition (see [5, Section 3.3]). The map F'' is the *simple composition* of F and F'.

Given compact intervals I_j^s, I_j^u, $j = 0, 1$, and two affine-like maps $F_0 : P_0 \to Q_0$, $F_1 : P_1 \to Q_1$ from vertical strips $P_0 \subset I_0^s \times I_0^u$, $P_1 \subset I_{a_s}^s \times I_{a_s}^u$ to horizontal strips $Q_0 \subset I_{a_u}^s \times I_{a_u}^u$, $Q_1 \subset I_1^s \times I_1^u$, we can introduce a quantity $\delta(Q_0, P_1)$ roughly measuring the distance between Q_0 and the tip of the parabolic strip $G^{-1}(P_1)$ (where G is the folding map); see [5, Section 3.5]. If

$$\delta(Q_0, P_1) > (1/b)(|P_1| + |Q_0|)$$

and the implicit representations of F_0 and F_1 satisfy the bound

$$\max\{|(A_1)_y|, |(A_1)_{yy}|, |(B_0)_x|, |(B_0)_{xx}|\} < b$$

for an adequate constant $b = b(g_0) > 0$, the composition $F_1 \circ G \circ F_0$ defines two affine-like maps $F^{\pm} : P^{\pm} \to Q^{\pm}$ with domains $P^{\pm} \subset P_0$ and $Q^{\pm} \subset Q_1$ called the *parabolic compositions* of F_0 and F_1.

2.4. The class $\mathcal{R}(I)$ of certain affine-like iterates

Given a parameter interval $I \subset [\varepsilon_0, 2\varepsilon_0]$, a triple $(P, Q, n) = (P_t, Q_t, n)_{t \in I}$ is called an *I-persistent affine-like iterate* if $P_t \subset R_a$ and $Q_t \subset R_{a'}$ are vertical and horizontal strips, respectively, varying smoothly with $t \in I$, $n \geqslant 0$ is an integer such that $g_t^n|_{P_t} : P_t \to Q_t$ is an affine-like map for all $t \in I$ and $g_t^m(P_t) \subset \widehat{R}$ for each $0 \leqslant m \leqslant n$.

Given a candidate parameter interval I, it is assigned in [5, Section 5.3] a class $\mathcal{R}(I)$ of certain I-persistent affine-like iterates verifying seven requirements, (R1)–(R7):

(R1) the transitions $g_{aa'} : R_a \cap g_t^{-1}(R_{a'}) \to g_t(R_a) \cap R_{a'}$, $(a, a') \in \mathcal{B}$, belong to $\mathcal{R}(I)$,

(R2) each $(P, Q, n) \in \mathcal{R}(I)$ is an I-persistent affine-like iterate satisfying the (λ, u_0, v_0) cone condition and the $2D_0$ distortion condition,

(R3) the class $\mathcal{R}(I)$ is stable under simple compositions,

(R4) denote by P_s, respectively Q_u, the smallest cylinder of the Markov partition of K_{g_t} containing L_s, respectively L_u; if $(P, Q, n) \in \mathcal{R}(I)$ and $P \subset P_s$, then $|A_y|, |A_{yy}| \leqslant C\varepsilon_0$ for all $t \in I$; similarly, if $(P, Q, n) \in \mathcal{R}(I)$ and $Q \subset Q_u$, then $|B_x|, |B_{xx}| \leqslant C\varepsilon_0$ for all $t \in I$,

(R5) the class $\mathcal{R}(I)$ is stable under certain *allowed* parabolic compositions (cf. [5, page 33]),

(R6) each $(P, Q, n) \in \mathcal{R}(I)$ with $n > 1$ is obtained from simple or allowed parabolic compositions of shorter elements,

(R7) if the parabolic composition of $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I)$ is allowed, then

$$\delta(Q_0, P_1) \geqslant (1/C)(|P_1|^{1-\eta} + |Q_0|^{1-\eta})$$

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

where $\delta(Q_0, P_1)$ is the distance between Q_0 and the tip of $G^{-1}(P_1)$, $C = C(g_0) \geqslant 1$, and the parameter η relates to ε_0 and τ via the condition $0 < \varepsilon_0 \ll \eta \ll \tau < 1$.

Furthermore, [5, Theorem 1] ensures that the class $\mathcal{R}(I)$ satisfying (R1)–(R7) is *unique*.

For technical reasons, we will need to work with *extensions* of the elements $\mathcal{R}(I)$. More concretely, we consider the intervals J_a^s, J_a^u from Remark 2.1 and we denote by $S := \bigcup_{a\in\mathcal{A}} S_a$ the geometric Markov partition associated with smooth charts $J_a^s \times J_a^u \to S_a$. We say that $(\widetilde{P}, \widetilde{Q}, n)$ *extends* $(P, Q, n) \in \mathcal{R}(I)$ if $(\widetilde{P}, \widetilde{Q}, n)$ is an affine-like map with respect to S satisfying the (λ, u_0, v_0) cone condition and the $3D_0$ distortion condition such that the restriction of $(\widetilde{P}, \widetilde{Q}, n)$ to R is (P, Q, n). Note that if $(\widetilde{P}, \widetilde{Q}, n)$ extends (P, Q, n), then $\widetilde{P}$ is a strip of width $\leqslant C|P|$ containing a $C^{-1}|P|$-neighborhood of P and $\widetilde{Q}$ is a strip of width $\leqslant C|Q|$ containing a $C^{-1}|Q|$-neighborhood of Q (where $C = C(g_0) \geqslant 1$), thanks to the cone and distortion conditions.

Proposition 2.3. *Each element $(P, Q, n) \in \mathcal{R}(I)$ admits an extension.*

Proof. Consider the subclass $\mathcal{S}(I)$ of $\mathcal{R}(I)$ consisting of elements admitting an extension. We want to show that $\mathcal{S}(I) = \mathcal{R}(I)$, and, in view of [5, Theorem 1], it suffices to check that $\mathcal{S}(I)$ verifies the requirements (R1)–(R7).

The fact that the transitions $g_{aa'}$ can be extended was already observed in Remark 2.1. In particular, $\mathcal{S}(I)$ satisfies (R1).

The requirements (R2), (R4) and (R7) for $\mathcal{S}(I)$ are automatic (because they concern geometric properties of $(P, Q, n) \in \mathcal{S}(I) \subset \mathcal{R}(I)$ themselves).

The condition (R3) for $\mathcal{S}(I)$ holds because the simple composition of (P_0, Q_0, n_0), $(P_1, Q_1, n_1) \in \mathcal{S}(I)$ is extended by the simple composition of the extensions of (P_0, Q_0, n_0) and (P_1, Q_1, n_1).

If $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{S}(I)$ satisfy the transversality requirement $Q_0 \pitchfork_I P_1$ (from [5, page 34]) allowing parabolic composition, then their extensions $(\widetilde{P}_0, \widetilde{Q}_0, n_0)$, $(\widetilde{P}_1, \widetilde{Q}_1, n_1)$ verify the same transversality requirement after replacing the constant 2 in (T1), (T2), (T3) in [5, page 34] by 7/4. From this fact and the discussion of parabolic compositions in [5, Sections 3.5 and 3.6], one sees that the parabolic composition of $(\widetilde{P}_0, \widetilde{Q}_0, n_0)$ and $(\widetilde{P}_1, \widetilde{Q}_1, n_1)$ is an extension of the parabolic composition of (P_0, Q_0, n_0) and $(P_1, Q_1, n_1) \in \mathcal{S}(I)$. Therefore, $\mathcal{S}(I)$ satisfies (R5).

At this point, it remains to check (R6) for $\mathcal{S}(I)$. For this sake, we recall (from [5, Section 5.5]) that $\mathcal{R}(I_0)$ consists of all affine-like iterates associated with the horseshoe K_{g_t}. In particular, $\mathcal{S}(I_0) = \mathcal{R}(I_0)$ thanks to our discussion so far. On the other hand, if I is a candidate interval distinct from I_0 and $\mathcal{S}(\widetilde{I}) = \mathcal{R}(\widetilde{I})$ for the smallest candidate interval $\widetilde{I}$ containing I, then we can apply the *structure theorem* (cf. [5, Theorem 2]) to write any element $(P, Q, n) \in \mathcal{R}(I)$ not coming from $\mathcal{R}(\widetilde{I})$ as the allowed parabolic compositions of shorter elements $(P_0, Q_0, n_0), \ldots, (P_k, Q_k, n_k) \in \mathcal{R}(\widetilde{I})$, $k > 0$. Since $\mathcal{S}(\widetilde{I}) = \mathcal{R}(\widetilde{I})$, we conclude that $\mathcal{S}(I)$ verifies (R6). □

2.5. Strong regularity tests

A candidate parameter interval I is tested for several quantitative conditions on the family of so-called *bicritical elements* of $\mathcal{R}(I)$. If a candidate interval I passes this strong regularity test, then all bicritical elements $(P, Q, n) \in \mathcal{R}(I)$ are thin in the sense that

$$|P| < |I|^{\beta}, \quad |Q| < |I|^{\beta},$$

where $\beta > 1$ depends only on g_0; more precisely, one imposes the mild condition that

$$1 < \beta < 1 + \min\{\omega_s, \omega_u\}, \tag{2.1}$$

where $\omega_s = -\frac{\log|\lambda(p_s)|}{\log|\mu(p_s)|}$ and $\omega_u = -\frac{\log|\mu(p_u)|}{\log|\lambda(p_u)|}$ with $\mu(p_s)$, $\mu(p_u)$ denoting the unstable eigenvalues of the periodic points p_s, p_u and $\lambda(p_s)$, $\lambda(p_u)$ denoting the stable eigenvalues of the periodic points p_s, p_u, and the important condition that

$$1 < \beta < \frac{(1 - \min\{d_s^0, d_u^0\})(d_s^0 + d_u^0)}{\max\{d_s^0, d_u^0\}(\max\{d_s^0, d_u^0\} + d_s^0 + d_u^0 - 1)} := \beta^*(d_s^0, d_u^0) \tag{2.2}$$

(cf. [5, Remark 8]).

2.6. Non-uniformly hyperbolic horseshoes and their stable sets

Let us fix *once and for all* a strongly regular parameter $t \in I_0 = [\varepsilon_0, 2\varepsilon_0]$, i.e., $\{t\} = \bigcap_{m\in\mathbb{N}} I_m$ for some decreasing sequence I_m of candidate intervals passing the strong regularity tests. In the sequel, $g_t = g$ denotes the corresponding dynamical system.

We define $\mathcal{R} := \bigcup_{m\in\mathbb{N}} \mathcal{R}(I_m)$, and, given a decreasing sequence of vertical strips P_k associated with some affine-like iterates $(P_k, Q_k, n_k) \in \mathcal{R}$, we say that $\omega = \bigcap_{k\in\mathbb{N}} P_k$ is a *stable curve*.

The set of stable curves is denoted by $\mathcal{R}_+^{\infty}$. The union of stable curves

$$\widetilde{\mathcal{R}}_+^{\infty} := \bigcup_{\omega\in\mathcal{R}_+^{\infty}} \omega$$

is a lamination by C^{1+Lip} curves with Lipschitz holonomy (cf. [5, Section 10.5]).

The set $\mathcal{R}_+^{\infty}$ is naturally partitioned in terms of *prime elements* of $\mathcal{R}$. More precisely, $(P, Q, n) \in \mathcal{R}$ is called a prime element if it is not the simple composition of two shorter elements. This notion allows to write $\mathcal{R}_+^{\infty} := \mathcal{D}_+ \cup \mathcal{N}_+$, where $\mathcal{N}_+$ is the set of stable curves contained in infinitely many prime elements and $\mathcal{D}_+^{\infty}$ is the complement of $\mathcal{N}_+$.

If $\omega \in \mathcal{D}_+$ is a stable curve such that $(P, Q, n) \in \mathcal{R}$ is the thinnest prime element containing ω, then $g^n(\omega)$ is contained in a stable curve $\omega' := T^+(\omega) \in \mathcal{R}_+^{\infty}$. In this way, we obtain a partially defined dynamics T^+ on $\mathcal{R}_+^{\infty} = \mathcal{D}_+ \cup \mathcal{N}_+$. The map $T^+ : \mathcal{D}_+ \to \mathcal{R}_+^{\infty}$ is Bernoulli and uniformly expanding with countably many branches; see [5, Section 10.5].

These hyperbolic features of T_+ permit to introduce a one-parameter family of transfer operators L_d whose dominant eigenvalues $\lambda_d > 0$ detect the transverse Hausdorff dimension of the lamination $\widetilde{\mathcal{R}}_+^{\infty}$, i.e., $\widetilde{\mathcal{R}}_+^{\infty}$ has Hausdorff dimension $1 + d_s$ where d_s is the unique value of d with $\lambda_d = 1$ (cf. [5, Theorem 4]).

The set $\{z \in W^s(\Lambda) : g^n(z) \in \widetilde{\mathcal{R}}_+^{\infty}$ for some $n \geqslant 0\}$ is the so-called *well-behaved part* of the stable set $W^s(\Lambda_g)$.

Following [5, Section 11.6], we write

$$W^s(\Lambda) = \bigcup_{n\geqslant 0} g^{-n}(W^s(\Lambda, \widehat{R}) \cap R))$$

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

and we split the local stable set $W^s(\Lambda, \widehat{R}) \cap R$ into its well-behaved part and its *exceptional part*:

$$W^s(\Lambda, \widehat{R}) \cap R := \bigcup_{n \geqslant 0} (W^s(\Lambda, \widehat{R}) \cap R \cap g^{-n}(\widetilde{\mathcal{R}}_+^\infty)) \cup \mathcal{E}^+$$

where

$$\mathcal{E}^+ := \{z \in W^s(\Lambda, \widehat{R}) \cap R : g^n(z) \notin \widetilde{\mathcal{R}}_+^\infty \text{ for all } n \geqslant 0\}. \tag{2.3}$$

Since g is a diffeomorphism and the C^{1+Lip}-lamination $\widetilde{\mathcal{R}}_+^\infty$ has transverse Hausdorff dimension $0 < d_s < 1$, we deduce that the Hausdorff dimension of the stable set $W^s(\Lambda)$ is the following.

Proposition 2.4. $\mathrm{HD}(W^s(\Lambda)) = \max\{1 + d_s, \mathrm{HD}(\mathcal{E}^+)\}$.

For the study of $\mathrm{HD}(\mathcal{E}^+)$, it is important to recall that the exceptional set $\mathcal{E}^+$ has a natural decomposition in terms of the successive passages through the so-called *parabolic cores* of vertical strips (cf. [5, Section 11.7]).

More precisely, the parabolic core $c(P)$ of $(P, Q, n) \in \mathcal{R}$ is the set of points of $W^s(\Lambda, \widehat{R})$ belonging to P but not to any *child*[4] of P. If we denote by $\mathcal{C}_-$ the set of elements $(P_0, Q_0, n_0) \in \mathcal{R}$ with $c(P_0) \neq \emptyset$, then

$$\mathcal{E}^+ = \bigcup_{(P_0, Q_0, n_0) \in \mathcal{C}_-} \mathcal{E}^+(P_0),$$

where $\mathcal{E}^+(P_0) := \mathcal{E}^+ \cap c(P_0)$.

Since $(P_0, Q_0, n_0) \in \mathcal{C}_-$ implies that $g^{n_0}(\mathcal{E}^+(P_0)) \subset Q_0 \cap L_u \cap \mathcal{E}^+$ and $G(g^{n_0}(\mathcal{E}^+(P_0)) = g^{n_0+N_0}(\mathcal{E}^+(P_0)) \subset L_s \cap \mathcal{E}^+$, we can write

$$\mathcal{E}^+(P_0) := \bigcup_{(P_1, Q_1, n_1) \in \mathcal{C}_-} \mathcal{E}^+(P_0, P_1),$$

where $\mathcal{E}^+(P_0, P_1) := \{z \in \mathcal{E}^+(P_0) : g^{n_0+N_0}(z) \in c(P_1)\}$.

In general, we can inductively define

$$\mathcal{E}^+(P_0, \ldots, P_k) = \bigcup_{(P_{k+1}, Q_{k+1}, n_{k+1}) \in \mathcal{C}_-} \mathcal{E}^+(P_0, \ldots, P_k, P_{k+1})$$

so that

$$\mathcal{E}^+ = \bigcup_{(P_0, P_1, \ldots, P_k) \text{ admissible}} \mathcal{E}^+(P_0, \ldots, P_k),$$

where $(P_0, \ldots, P_k)$ is *admissible* whenever $\mathcal{E}^+(P_0, \ldots, P_k) \neq \emptyset$.

The admissibility condition on $(P_0, \ldots, P_{k+1})$ is a severe geometrical constraint on the elements $(P_i, Q_i, n_i) \in \mathcal{R}$: for example, $(P_0, Q_0, n_0) \in \mathcal{C}_-$,

$$\max\{|P_1|, |Q_1|\} \leqslant \varepsilon_0^\beta \tag{2.4}$$

[4] P' is a child of P if P' is the vertical strip associated with some $(P', Q', n') \in \mathcal{R}$ obtained by simple compositions of (P, Q, n) with the transition maps $g_{aa'}$ of the Markov partition of the horseshoe K_g or parabolic composition of (P, Q, n) with some element of $\mathcal{R}$ (cf. [5, Section 6.2]).

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

and, for $\widetilde{\beta} := \beta(1-\eta)(1+\tau)^{-1}$,

$$\max\{|P_{j+1}|, |Q_{j+1}|\} \leqslant C|Q_j|^{\widetilde{\beta}} \tag{2.5}$$

for all $j \geqslant 1$ (cf. [5, Lemma 24]).

Hence, by taking $1 < \widehat{\beta} < \widetilde{\beta}$, the admissibility condition implies that

$$\max\{|P_j|, |Q_j|\} \leqslant \varepsilon_0^{\widehat{\beta}^j} \tag{2.6}$$

(for ε_0 sufficiently small). Therefore, the widths of the strips P_j and Q_j confining the dynamics of $\mathcal{E}^+$ decay doubly exponentially fast.

2.7. Hausdorff measures

Given a bounded subset X of the plane, $0 \leqslant d \leqslant 2$, and $\delta > 0$, the d-Hausdorff measure $m_\delta^d(X)$ at scale $\delta > 0$ of X is the infimum over open coverings $(U_i)_{i\in I}$ of X with diameter $\operatorname{diam}(U_i) < \delta$ of the following quantity:

$$\sum_{i\in I} \operatorname{diam}(U_i)^d.$$

In other terms, $m_\delta^d(X)$ is the d-Hausdorff measure at scale $\delta > 0$ of X. Observe that

$$m_\delta^d\left(\bigcup_{\alpha\in\mathbb{N}} X_\alpha\right) \leqslant \sum_{\alpha\in\mathbb{N}} m_\delta^d(X_\alpha).$$

In this context, the Hausdorff dimension of X is

$$\operatorname{HD}(X) := \inf\{d \in [0,2] : m^d(X) = 0\}.$$

3. The expected Hausdorff dimension of $W^s(\Lambda)$

By Proposition 2.4, the proof of Theorem 1.4 is reduced to the following.

Theorem 3.1. *In the setting of Theorem 1.4,* $\operatorname{HD}(\mathcal{E}^+) < 1 + d_s$.

For the proof of this theorem, we need some facts about the Hausdorff measures of images of maps with bounded geometry.

3.1. Planar maps with bounded geometry

We start with a lemma about the Hausdorff measure at scale 1 of the image of the unit disk $\mathbb{D}^2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\}$ under a map with bounded geometry:

Lemma 3.2. *Let $K \geqslant 1$, $L \geqslant 1$ and $f : \mathbb{D}^2 \to \mathbb{R}^2$ be a C^1 diffeomorphism onto its image such that $\|Df\| \leqslant K$ and $|\operatorname{Jac}(f)| := |\det Df| \leqslant L$. Then, there is a universal constant C (e.g., $C = 170\pi$) such that for all $1 \leqslant d \leqslant 2$, we have*

$$\inf_{\substack{(U_i) \text{ covers } f(\mathbb{D}^2),\\ \operatorname{diam}(U_i)\leqslant\sqrt{2}\,\forall i}} \sum_i \operatorname{diam}(U_i)^d \leqslant C \cdot \max\{K, L\}^{2-d} L^{d-1}.$$

Proof. Fix $0 < \varepsilon_0 < 1$ small enough so that f has an extension to the disk $\mathbb{D}^2_{1+\varepsilon_0} := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant (1+\varepsilon_0)^2\}$. By a slight abuse of notation, we still denote such an extension by f.

Given $0 < \varepsilon < \varepsilon_0$, let $U_{1+\varepsilon} = f(\mathbb{D}^2_{1+\varepsilon})$ and $\partial U_{1+\varepsilon}$ be its boundary. For later use, we set $K_\varepsilon := \sup_{\mathbb{D}^2_{1+\varepsilon}} \|Df\|$ and $L_\varepsilon := \sup_{\mathbb{D}^2_{1+\varepsilon}} |\mathrm{Jac}(f)|$. For $k \geqslant 0$ integer, let $\mathcal{Q}_k$ be the collection of squares in the plane of side $1/2^k$ and vertices on $\mathbb{Z}^2/2^k$. Let $\mathcal{C}_0^{(\varepsilon)}$ be the set of squares Q in $\mathcal{Q}_0$ such that

$$\mathrm{area}(Q \cap U_{1+\varepsilon}) \geqslant \tfrac{1}{5} \cdot \mathrm{area}(Q).$$

For $k > 0$, let $\mathcal{C}_k^{(\varepsilon)}$ be the set of squares Q in $\mathcal{Q}_k$ such that Q is not contained in some $Q' \in \mathcal{C}_l^{(\varepsilon)}$, $l < k$ and

$$\mathrm{area}(Q \cap U_{1+\varepsilon}) \geqslant \tfrac{1}{5} \cdot \mathrm{area}(Q).$$

Remark 3.3. In this construction, we are implicitly assuming that $U_{1+\varepsilon} = f(\mathbb{D}^2_{1+\varepsilon})$ is not entirely contained in a dyadic square $Q \in \bigcup_{k=0}^{\infty} \mathcal{Q}_k$. Of course, there is no loss of generality in this assumption; if $U_{1+\varepsilon} \subset Q$ for some $Q \in \mathcal{Q}_k$, then the lemma follows from the trivial bound $\mathrm{diam}(Q)^d \leqslant \sqrt{2}^d$.

Note that $f(\mathbb{D}^2)$ is contained in the interior of $U_{1+\varepsilon}$. In particular, each point of $f(\mathbb{D}^2)$ belongs to some dyadic square contained in $U_{1+\varepsilon}$. Hence, $(U_i^{(\varepsilon)})_{i\in\mathbb{N}} := \bigcup_{k=0}^{\infty} \mathcal{C}_k^{(\varepsilon)}$ is a covering of $f(\mathbb{D}^2)$ with $\mathrm{diam}(U_i^{(\varepsilon)}) \leqslant \sqrt{2}$ and

$$\sum_i \mathrm{diam}(U_i^{(\varepsilon)})^d = \sum_{k=0}^{\infty} N_k^{(\varepsilon)} \left(\frac{1}{2^k}\right)^d,$$

where $N_k^{(\varepsilon)} := (\sqrt{2})^d \# \mathcal{C}_k^{(\varepsilon)}$. By considering this expression as an L^d-norm and by applying interpolation between the L^1 and L^2 norms, we see that

$$\sum_{k=0}^{\infty} N_k^{(\varepsilon)} \left(\frac{1}{2^k}\right)^d \leqslant \left(\sum_{k=0}^{\infty} \frac{N_k^{(\varepsilon)}}{2^k}\right)^{2-d} \left(\sum_{k=0}^{\infty} \frac{N_k^{(\varepsilon)}}{(2^k)^2}\right)^{d-1}. \tag{3.1}$$

We estimate these L^1 and L^2 norms as follows. First, we have

$$\begin{aligned}\sum_k \frac{N_k^{(\varepsilon)}}{(2^k)^2} &= (\sqrt{2})^d \sum_k \sum_{Q \in \mathcal{C}_k^{(\varepsilon)}} \mathrm{area}(Q) \\ &\leqslant 5(\sqrt{2})^d \sum_k \sum_{Q \in \mathcal{C}_k^{(\varepsilon)}} \mathrm{area}(Q \cap U_{1+\varepsilon}) \\ &\leqslant 10\, \mathrm{area}(U_{1+\varepsilon}) \\ &\leqslant 10\pi \cdot (1+\varepsilon)^2 L_\varepsilon, \end{aligned} \tag{3.2}$$

for any $1 \leqslant d \leqslant 2$. From the previous estimate, we obtain that $N_0^{(\varepsilon)} \leqslant 10\pi(1+\varepsilon)^2 L_\varepsilon$.

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

On the other hand, we *claim* that there exists an universal constant $c' > 0$ (e.g., $c' = 1/20$) such that for any $k > 0$ and $Q \in \mathcal{C}_k$, we have

$$\text{length}(\partial U_{1+\varepsilon} \cap Q) \geqslant c'/2^k.$$

This claim implies

$$\begin{aligned}\sum_{k>0} \frac{N_k^{(\varepsilon)}}{2^k} &= (\sqrt{2})^d \sum_{k>0} \sum_{Q\in\mathcal{C}_k^{(\varepsilon)}} \frac{1}{2^k} \\ &\leqslant (\sqrt{2})^d c'^{-1} \sum_{k>0} \sum_{Q\in\mathcal{C}_k^{(\varepsilon)}} \text{length}(\partial U_{1+\varepsilon} \cap Q) \\ &\leqslant 2(\sqrt{2})^d c'^{-1} \text{length}(\partial U_{1+\varepsilon}) \\ &\leqslant 8\pi c'^{-1}(1+\varepsilon) K_\varepsilon,\end{aligned}$$

for any $1 \leqslant d \leqslant 2$. Hence,

$$\begin{aligned}\textstyle\sum_{k\geqslant 0} \frac{N_k^{(\varepsilon)}}{2^k} = N_0^{(\varepsilon)} + \sum_{k>0} \frac{N_k^{(\varepsilon)}}{2^k} &\leqslant 10\pi(1+\varepsilon)^2 L_\varepsilon + 8\pi c'^{-1}(1+\varepsilon)K_\varepsilon \\ &\leqslant (10\pi + 8\pi c'^{-1})(1+\varepsilon)^2 \max\{K_\varepsilon, L_\varepsilon\}.\end{aligned} \tag{3.3}$$

Thus, in view of (3.3), (3.2) and (3.1), since $0 < \varepsilon < \varepsilon_0$ is arbitrary and $K_0 := \lim_{\varepsilon\to 0} K_\varepsilon$, $L_0 := \lim_{\varepsilon\to 0} L_\varepsilon$ satisfy $K_0 \leqslant K$, $L_0 \leqslant L$, the Lemma follows (with $C = 170\pi$ when $c' = 1/20$) once we prove the claim.

To show the claim, we observe that if $\text{length}(\partial U_{1+\varepsilon} \cap Q) < c'/2^k$, then $\partial U_{1+\varepsilon} \cap Q$ is contained in a $c'/2^k$-neighborhood of ∂Q (thanks to Remark 3.3). So, the complement of this neighborhood (whose area is $(1-2c')^2 \cdot \text{area}(Q)$) is either contained in $U_{1+\varepsilon}$ or disjoint from $U_{1+\varepsilon}$. This contradicts the definition of $\mathcal{C}_k^{(\varepsilon)}$ if $c' > 0$ is small enough (e.g., $c' = 1/20$). □

After scaling, we obtain the following version of the previous lemma.

Lemma 3.4. *Let $K \geqslant L \geqslant 1$ and $f : \mathbb{D}_r^2 \to \mathbb{R}^2$ be a C^1 diffeomorphism from $\mathbb{D}_r^2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant r^2\}$ on its image such that $\|Df\| \leqslant K$ and $|\text{Jac}(f)| \leqslant L$. Then, there is an universal constant C (e.g., $C = 170\pi$) such that for all $1 \leqslant d \leqslant 2$, we have*

$$\inf_{\substack{(U_i) \text{ covers } f(\mathbb{D}_r^2), \\ \text{diam}(U_i) \leqslant \frac{Lr\sqrt{2}}{K}\,\forall i}} \sum_i \text{diam}(U_i)^d \leqslant C \cdot r^d \cdot K^{2-d} \cdot L^{d-1}.$$

3.2. Application of Lemma 3.4 to the proof of Theorem 3.1

Consider again the decomposition

$$\mathcal{E}^+ = \bigcup_{(P_0,\ldots,P_k) \text{ admissible}} \mathcal{E}^+(P_0, \ldots, P_k)$$

and let us estimate $m_{s_k}^d(\mathcal{E}^+(P_0, \ldots, P_k))$. For this sake, recall that the admissibility condition on (P_i, Q_i, n_i), $i = 0, \ldots, k$ implies that

$$g^{n_0+N_0+\cdots+n_{k-1}+N_0}(\mathcal{E}^+(P_0, \ldots, P_k))$$

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

is contained in a rectangular region of width $C|Q_k|^{\frac{(1-\eta)}{2}}|P_k|$ and height $C|Q_{k-1}|^{\frac{(1-\eta)}{2}}$ (cf. [5, proof of Proposition 62] and the beginning of [3, proof of Lemma 3.2]).

In order to alleviate the notations, we denote $g^{n_i}|_{P_i} := F_i$, $G := g^{N_0}|_{L_u}$, $F^{(k)} := G \circ F_{k-1} \circ \cdots \circ G \circ F_0$, and we write $\delta_j := |Q_j|$ for $j = 0, \ldots, k-1$. In this language, we have that

$$F^{(k)}(\mathcal{E}^+(P_0, \ldots, P_k)) = g^{n_0+N_0+\cdots+n_{k-1}+N_0}(\mathcal{E}^+(P_0, \ldots, P_k))$$

is contained in a rectangular region of width $C|Q_k|^{\frac{(1-\eta)}{2}}|P_k|$ and height $C\delta_{k-1}^{\frac{(1-\eta)}{2}}$. Let us divide this rectangular region into $N_k := C^2 \frac{\delta_{k-1}^{(1-\eta)/2}}{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}$ disks of diameters $C|Q_k|^{\frac{(1-\eta)}{2}}|P_k|$, and let us denote by $\mathcal{O}_k$ the subcollection of such disks intersecting $F^{(k)}(\mathcal{E}^+(P_0, \ldots, P_k))$.

Recall that Proposition 2.3 says that the affine-like iterate F_i can be extended to an affine-like iterate $\widetilde{F}_i$ with domain $\widetilde{P}_i$ in such a way that $|\widetilde{P}_i| \leqslant C|P_i|$ and a $C^{-1}|P_i|$-neighborhood of P_i is included in the domain $\widetilde{P}_i$ of F_i. Given a square $S \in \mathcal{O}_k$, we have that its pre-image under G contains a point of $F^{(k)}(\mathcal{E}^+(P_0, \ldots, P_k))$ and its diameter is $C|Q_k|^{\frac{(1-\eta)}{2}}|P_k|$. Since $\max\{|P_k|, |Q_k|\} \leqslant C|Q_{k-1}|^{\widetilde{\beta}}$ with $\widetilde{\beta} > 1$ (cf. (2.5)), the pre-image of S under G is contained in a $C^{-1}|Q_{k-1}|$-neighborhood of Q_{k-1}, and, hence, it is contained in $\widetilde{Q}_{k-1}$. Therefore, the pre-image of S under $G \circ \widetilde{F}_{k-1}$ contains a point of $F^{(k-1)}(\mathcal{E}^+(P_0, \ldots, P_k))$ and its diameter is $\leqslant C\frac{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}{|Q_{k-1}|}$. Hence, the pre-image of S under $G \circ \widetilde{F}_{k-1} \circ G$ is contained in a $C^{-1}|Q_{k-2}|$-neighborhood of Q_{k-2} and, *a fortiori*, in $\widetilde{Q}_{k-2}$ whenever

$$C\frac{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}{|Q_{k-1}|} \leqslant C^{-1}|Q_{k-2}|.$$

Since $\max\{|P_j|, |Q_j|\} \leqslant |Q_{j-1}|^{\widetilde{\beta}}$, the inequality above holds when

$$\widetilde{\beta}\left(\frac{(3-\eta)}{2}\widetilde{\beta} - 1\right) > 1, \ \text{i.e.,} \ \frac{(3-\eta)}{2} > \frac{1}{\widetilde{\beta}} + \frac{1}{\widetilde{\beta}^2}.$$

In this case, the pre-image of S under $G \circ \widetilde{F}_{k-1} \circ G \circ \widetilde{F}_{k-2}$ contains a point of $F^{(k-2)}(\mathcal{E}^+(P_0, \ldots, P_k))$ and its diameter is $\leqslant C\frac{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}{|Q_{k-1}||Q_{k-2}|}$. By induction, the pre-image of S under $G \circ \widetilde{F}_{k-1} \circ G \circ \cdots \circ \widetilde{F}_{j+1} \circ G$ is contained in a $C^{-1}|Q_j|$ of Q_j, and, *a fortiori*, in $\widetilde{Q}_j$, whenever

$$C^{k-j}\frac{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}{|Q_{k-1}|\ldots|Q_{j+1}|} \leqslant |Q_j|.$$

Since $\max\{|P_\ell|, |Q_\ell|\} \leqslant |Q_{\ell-1}|^{\widetilde{\beta}}$, the inequality above holds when

$$\frac{(3-\eta)}{2} > \frac{1}{\widetilde{\beta}} + \cdots + \frac{1}{\widetilde{\beta}^{k-j}}.$$

In this case, the pre-image of S under $G \circ \widetilde{F}_{k-1} \circ \cdots \circ G \circ \widetilde{F}_j$ contains a point of $F^{(j)}(\mathcal{E}^+(P_0, \ldots, P_k))$ and its diameter is $\leqslant C\frac{|Q_k|^{\frac{(1-\eta)}{2}}|P_k|}{|Q_{k-1}|\ldots|Q_j|}$.

In particular, we have that $\mathcal{E}^+(P_0, \ldots, P_k)$ is covered by the pre-images under $\widetilde{F}^{(k)} := G \circ \widetilde{F}_{k-1} \circ \cdots \circ G \circ \widetilde{F}_0$ of the disks in $\mathcal{O}_k$ *whenever* we can take $\widetilde{\beta}$ with $\sum_{\ell=1}^{\infty} \widetilde{\beta}^{-\ell} < 3/2$, i.e., $\widetilde{\beta} > 5/3$. Observe that, from the definitions, such a choice is possible if the quantity β in (2.1) and (2.2) satisfies $\beta > 5/3$. Since assumption (1.3) in Theorem 1.4 says that constraint (2.1) is superfluous, β can be taken arbitrarily close to

$$\beta^* := \beta^*(d_s^0, d_u^0) = \frac{(1 - \min\{d_s^0, d_u^0\})(d_s^0 + d_u^0)}{\max\{d_s^0, d_u^0\}(\max\{d_s^0, d_u^0\} + d_s^0 + d_u^0 - 1)}$$

and, hence, the property $\beta > 5/3$ is ensured by hypothesis (1.4).

Our plan to estimate $m_{s_k}^d(\mathcal{E}^+(P_0, \ldots, P_k))$ is to apply Lemma 3.4 to the image of each of these disks under the map $(\widetilde{F}^{(k)})^{-1}$. Therefore, let us estimate the Lipschitz constant and the Jacobian of this map on these squares.

Lemma 3.5. *On the disks of the collection $\mathcal{O}_k$, one has*

$$|\text{Jac}((\widetilde{F}^{(k)})^{-1})| \leqslant C^k \prod_{i=0}^{k-1} \frac{|P_i|}{|Q_i|} \leqslant C^k |P_0| \delta_{k-1}^{-1} := L_k.$$

Proof. The Jacobian determinant of an affine-like map from a vertical strip P to a horizontal strip Q with implicit representation (A, B) is

$$C^{-1}|P|/|Q| \leqslant A_x^{-1} B_y \leqslant C|P|/|Q|$$

(see Remark 2.2).

By definition, $(\widetilde{F}^{(k)})^{-1} = (G \circ \widetilde{F}_{k-1} \circ \cdots \circ G \circ \widetilde{F}_0)^{-1}$, where $G = g^{N_0}$ is the folding map (a fixed map with uniformly bounded Jacobian) and $\widetilde{F}_i$ are the affine-like maps $g^{n_i}|_{\widetilde{P}_i} : \widetilde{P}_i \to \widetilde{Q}_i$ with $|\widetilde{P}_i| \leqslant C|P_i|$ and $|\widetilde{Q}_i| \leqslant C|Q_i|$. Therefore,

$$|\text{Jac}((\widetilde{F}^{(k)})^{-1})| = |\text{Jac}((G \circ \widetilde{F}_{k-1} \circ \cdots \circ G \circ \widetilde{F}_0)^{-1})| \leqslant C^k \prod_{i=0}^{k-1} \frac{|P_i|}{|Q_i|}.$$

Since $|P_i| \leqslant C|Q_{i-1}|^{\widetilde{\beta}}$ with $\widetilde{\beta} > 1$ (cf. (2.5)), it follows that

$$|\text{Jac}((\widetilde{F}^{(k)})^{-1})| \leqslant C^k \prod_{i=0}^{k-1} \frac{|P_i|}{|Q_i|} \leqslant C^k \frac{|P_0|}{|Q_{k-1}|} := C^k |P_0| \delta_{k-1}^{-1}.$$

This proves the lemma. □

Lemma 3.6. *On the disks of the collection $\mathcal{O}_k$, one has*

$$\|D(\widetilde{F}^{(k)})^{-1}\| \leqslant C^k \delta_{k-1}^{-1} (\delta_{k-2} \ldots \delta_0)^{-(1+\eta)/2} := K_k.$$

Proof. Let u_k be a unit vector at a point x_k of a disk in $\mathcal{O}_k$. We define inductively

$$y_{j-1} = G^{-1}(x_j), \quad x_j = \widetilde{F}_j^{-1}(y_j)$$

and

$$v_{j-1} = DG^{-1}(x_j) u_j, \quad u_j = D\widetilde{F}_j^{-1}(y_j) v_j.$$

Observe that $\|v_{k-1}\| \sim 1$.

Given an affine-like map $F : P \to Q$, the vector field on Q obtained by pushing forward by F the horizontal direction on P is called the *horizontal direction in the affine-like sense*.

We will prove by induction on j that the following two facts:

$$\|u_{k-j}\| \leqslant C^j \delta_{k-1}^{-1}(\delta_{k-2}\ldots\delta_{k-j})^{-(1+\eta)/2},$$

and, moreover, if the angle of v_{k-j} with the horizontal direction in the affine-like sense is at most $\delta_{k-j}^{1/2}$, one has

$$\|u_{k-j}\| \leqslant C^j(\delta_{k-1}\ldots\delta_0)^{-(1+\eta)/2}.$$

For this sake, we consider three cases:

- $\|u_{k-j}\| \leqslant 1$: this means that the angle of v_{k-j} with the horizontal direction in the affine-like sense is $\leqslant C|Q_{k-1}|$; in this case, the estimate follows by induction on j.
- $\|u_{k-j}\| > 1$: in this case, we have

$$\|v_{k-j-1}\| \sim \|u_{k-j}\| \leqslant C|Q_{k-j}|^{-1}\|u_{k-j+1}\|$$

  and the angle of v_{k-j-2} with the horizontal direction in the affine-like sense is at most $|Q_{k-j-1}|^{(1-\eta)/2}$ (compared with the calculations in [5, page 192]).
- $\|u_{k-j}\| > 1$ and the angle of v_{k-j-1} with the 'horizontal' direction is $\leqslant C|Q_{k-j-1}|^{(1-\eta)/2}$. In this case, we have

$$\|v_{k-j-2}\| \sim \|u_{k-j-1}\| \leqslant C|Q_{k-j-1}|^{-(1+\eta)/2}\|u_{k-j}\|.$$

Since $\|u_k\| = 1$ and $\|v_{k-1}\| \sim 1$, this completes the argument. □

By plugging Lemmas 3.5 and 3.6 into Lemma 3.4 for each of the squares $Q \in \mathcal{O}_k$, we obtain

$$m_{s_k}^d((\widetilde{F}^{(k)})^{-1}(Q)) \leqslant Cr_k^d \cdot K_k^{2-d} \cdot L_k^{d-1},$$

where $1 \leqslant d \leqslant 2$, $K_k = C^k\delta_{k-1}^{-1}(\delta_{k-2}\ldots\delta_0)^{-(1+\eta)/2}$, $L_k = C^k|P_0|\delta_{k-1}^{-1}$, $r_k = C|Q_k|^{\frac{(1-\eta)}{2}}|P_k|$, and $s_k = L_k r_k/K_k$. This gives

$$m_{s_k}^d(\mathcal{E}^+(P_0,\ldots,P_k)) \leqslant C^k N_k \cdot r_k^d \cdot K_k^{2-d} \cdot L_k^{d-1},$$

where $N_k = C^2\delta_{k-1}^{(1-\eta)/2}/r_k$. This estimate can be rewritten as

$$m_{s_k}^d(\mathcal{E}^+(P_0,\ldots,P_k)) \leqslant \frac{C^k|P_0|^{d-1}|P_k|^{d-1}|Q_k|^{(d-1)(1-\eta)/2}}{|Q_{k-1}|^{(1+\eta)/2}(|Q_{k-2}|\ldots|Q_0|)^{(2-d)(1+\eta)/2}}. \tag{3.4}$$

At this point, it is useful to recall that $\max\{|P_j|, |Q_j|\} \leqslant C|Q_{j+1}|^{\widetilde{\beta}}$ for $j \geqslant 0$ (cf. (2.5)), where $\widetilde{\beta} = \beta(1-\eta)(1+\tau)^{-1}$ is close to the parameter β satisfying constraints (2.1) and (2.2). Furthermore, assumption (1.3) in Theorem 1.4 says that the constraint (2.1) is superfluous, so that we can take β arbitrarily close to

$$\beta^* := \beta^*(d_s^0, d_u^0) = \frac{(1-\min\{d_s^0, d_u^0\})(d_s^0 + d_u^0)}{\max\{d_s^0, d_u^0\}(\max\{d_s^0, d_u^0\} + d_s^0 + d_u^0 - 1)}.$$

From these facts, we can use (3.4) to prove the following lemma:

Lemma 3.7. *For an appropriate choice of* $d = 1 + d_s^0 - o(1)$*, one has*

(a) $m_{s_k}^d(\mathcal{E}^+(P_0,\ldots,P_k)) \leqslant C^k |P_0|^{d-1}|Q_k|^{\frac{(d-1)(1-\eta)}{2}}$ *whenever* $\beta^* \cdot d_s^0 > \frac{1}{2} + \frac{1-d_s^0}{2(\beta^*-1)}$*;*

(b) $m_{s_k}^d(\mathcal{E}^+(P_0,\ldots,P_k)) \leqslant C^k |P_0|^{d-1}|Q_k|^{\frac{(d-1)(3-\eta)}{2} - \frac{1}{2\widetilde{\beta}} - \frac{(2-d)}{2\widetilde{\beta}(\widetilde{\beta}-1)}}$ *whenever* $\beta^* \cdot d_s^0 \leqslant \frac{1}{2} + \frac{1-d_s^0}{2(\beta^*-1)}$.

Proof. By (3.4), our task is to control

$$\frac{|P_k|^{d-1}|Q_k|^{(d-1)(1-\eta)/2}}{|Q_{k-1}|^{(1+\eta)/2}(|Q_{k-2}|\ldots|Q_0|)^{(2-d)(1+\eta)/2}}$$

for $d - 1 = d_s^0 - o(1)$.

On the other hand, since $|Q_{k-1}| \leqslant C|Q_{k-2}|^{\widetilde{\beta}} \leqslant \cdots \leqslant C^{k-1-j}|Q_j|^{\widetilde{\beta}^{k-1-j}}$, $|P_k| \leqslant C|Q_{k-1}|^{\widetilde{\beta}}$ and $\sum_{j=0}^{k-2} \frac{1}{\widetilde{\beta}^{k-1-j}} \leqslant \frac{1}{\widetilde{\beta}-1}$, we see that

- if $\widetilde{\beta}$ is close to β^* and $\beta^* \cdot d_s^0 > \frac{1}{2} + \frac{1-d_s^0}{2(\beta^*-1)}$, then

$$\begin{aligned}\frac{|P_k|^{d_s^0}|Q_k|^{d_s^0(1-\eta)/2}}{|Q_{k-1}|^{\frac{1+\eta}{2}}(|Q_{k-2}|\ldots|Q_0|)^{\frac{(1-d_s^0)(1+\eta)}{2}}} &\leqslant C^k|Q_{k-1}|^{\widetilde{\beta}d_s^0-(1+\eta)(\frac{1}{2}+\frac{1-d_s^0}{2(\widetilde{\beta}-1)})}|Q_k|^{\frac{d_s^0(1-\eta)}{2}}\\ &\leqslant C^k|Q_k|^{(d-1)(1-\eta)/2};\end{aligned}$$

- if $\widetilde{\beta}$ is close to β^* and $\beta^* \cdot d_s^0 \leqslant \frac{1}{2} + \frac{1-d_s^0}{2(\beta^*-1)}$, then

$$\begin{aligned}\frac{|P_k|^{d_s^0}|Q_k|^{d_s^0(1-\eta)/2}}{|Q_{k-1}|^{(1+\eta)/2}(|Q_{k-2}|\ldots|Q_0|)^{(1-d_s^0)(1+\eta)/2}} &\leqslant C^k \frac{|Q_k|^{d_s^0(1-\eta)/2}}{|Q_{k-1}|^{(1+\eta)(\frac{1}{2}+\frac{1-d_s^0}{2(\widetilde{\beta}-1)})-\widetilde{\beta}d_s^0}}\\ &\leqslant C^k|Q_k|^{\frac{d_s^0(3-\eta)}{2}-(1+\eta)(\frac{1}{2\widetilde{\beta}}-\frac{1-d_s^0}{2\widetilde{\beta}(\widetilde{\beta}-1)})}.\end{aligned}$$

This completes the proof of the lemma (for $d - 1 = d_s^0 - o(1)$). □

This lemma enables us to complete the proof of Theorem 3.1.

Proof of Theorem 3.1. Take $d = 1 + d_s^0 - o(1)$. The decomposition

$$\mathcal{E}^+ = \bigcup_{(P_0,\ldots,P_k)\text{ admissible}} \mathcal{E}^+(P_0,\ldots,P_k),$$

the fact that the number of admissible sequences $(P_0,\ldots,Q_0)$ with fixed extremities P_0 and Q_k is $\leqslant C|Q_k|^{-C\eta}$ (cf. [5, page 193]), and Lemma 3.7 imply that

$$m_{s_k}^d(\mathcal{E}^+) \leqslant \sum_{\substack{P_0 \text{ with } Q_0 \text{ critical,}\\ Q_k \text{ critical}}} |P_0|^{d-1}|Q_k|^{e(d)-C\eta} \tag{3.5}$$

for all $k \in \mathbb{N}$, where

$$e(d) = \begin{cases} \dfrac{d-1}{2}, & \text{if } \beta^* \cdot d_s^0 > \dfrac{1}{2} + \dfrac{1-d_s^0}{2(\beta^*-1)}\\ \dfrac{3(d-1)}{2} - \dfrac{1}{2\widetilde{\beta}} - \dfrac{(2-d)}{2\widetilde{\beta}(\widetilde{\beta}-1)}, & \text{if } \beta^* \cdot d_s^0 \leqslant \dfrac{1}{2} + \dfrac{1-d_s^0}{2(\beta^*-1)}.\end{cases}$$

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

By Hölder's inequality, it follows from (3.5) that

$$m^d_{s_k}(\mathcal{E}^+) \leqslant \Bigg(\sum_{P \text{ with } Q \text{ critical}} |P|^{(d-1)p}\Bigg)^{1/p} \Bigg(\sum_{Q \text{ critical}} |Q|^{(e(d)-C\eta)q}\Bigg)^{1/q}$$

for any $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ (and $k \in \mathbb{N}$).

As it is explained in [5, pages 186, 187 and 188], the two series above are uniformly convergent and, hence,

$$m^d_{s_k}(\mathcal{E}^+) \leqslant C < \infty \quad \forall k \in \mathbb{N}, \tag{3.6}$$

for the following choices of parameters:

$$(d-1)p = \rho_s \sim d^0_s \tag{3.7}$$

and

$$(e(d) - C\eta)q = -\frac{\sigma}{1+\tau} + \tau d^*_u + \tau \sim d^0_s + d^0_u - 1, \tag{3.8}$$

where d^*_u, σ and ρ_s (respectively) are the quantities defined in pages 135 and 138 (respectively) of [5].

Since $s_k \to 0$ as $k \to \infty$, we proved that

$$\mathrm{HD}(\mathcal{E}^+) \leqslant d$$

for d satisfying (3.7) and (3.8). In particular, our task is reduced to prove that we can take $d < 1 + d^0_s$ verifying these constraints.

Note that the value of d verifying (3.7) and (3.8) is *already* imposed by the extra relation $1/p + 1/q = 1$.

More precisely,

(i) if $\beta^* \cdot d^0_s > \frac{1}{2} + \frac{1-d^0_s}{2(\beta^*-1)}$, then $e(d) = \frac{d-1}{2}$; therefore, the relations (3.7), (3.8) and $1/p + 1/q = 1$ imply that $(d-1)$ is close to

$$(d-1) \sim \left(\frac{1}{d^0_s} + \frac{1}{2(d^0_s + d^0_u - 1)}\right)^{-1}$$

(ii) if $\beta^* \cdot d^0_s \leqslant \frac{1}{2} + \frac{1-d^0_s}{2(\beta^*-1)}$, then $e(d) = \frac{3(d-1)}{2} - \frac{1}{2\widetilde{\beta}} - \frac{(2-d)}{2\widetilde{\beta}(\widetilde{\beta}-1)}$; therefore, the value of $(d-1)$ satisfying the constraints above is close to

$$(d-1) \sim \frac{(d^0_s + d^0_u - 1) + \frac{1}{2\beta^*} + \frac{1}{2\beta^*(\beta^*-1)}}{\frac{3}{2} + \frac{(d^0_s+d^0_u-1)}{d^0_s} + \frac{1}{2\beta^*(\beta^*-1)}}$$

(here, we are using that $\widetilde{\beta}$ is close to β^*).

In the first case (item (i)), we always have that $d - 1 < d^0_s$ because

$$\left(\frac{1}{d^0_s} + \frac{1}{2(d^0_s + d^0_u - 1)}\right)^{-1} < d^0_s.$$

In the second case (item (ii)), the fact that $d-1 < d_s^0$ is a direct consequence of our main assumption (1.4) in Theorem 1.4; indeed, a simple calculation reveals that the inequality

$$\frac{(d_s^0+d_u^0-1)+\frac{1}{2\beta^*}+\frac{1}{2\beta^*(\beta^*-1)}}{\frac{3}{2}+\frac{(d_s^0+d_u^0-1)}{d_s^0}+\frac{1}{2\beta^*(\beta^*-1)}} < d_s^0$$

is equivalent to $1+\frac{1-d_s^0}{(\beta^*-1)} < 3\beta^* d_s^0$. Since this inequality holds when $\beta^*(d_s^0, d_u^0) > 5/3$ (i.e., our assumption (1.4) in Theorem 1.4), the proof of Theorem 3.1 (and Theorem 1.4) is complete. □

Acknowledgements. We are grateful to the following institutions for their hospitality during the preparation of this article: Collège de France, Instituto de Matemática Pura e Aplicada (IMPA) and Kungliga Tekniska högskolan (KTH). The authors were partially supported by the Balzan Research Project of J. Palis, the French ANR grant 'DynPDE' (ANR-10-BLAN 0102) and the Brazilian CAPES grant (88887.136371/2017-00).

Appendix A. Large open sets generating non-uniformly hyperbolic horseshoes by C. Matheus, C. G. Moreira and J. Palis

In this appendix, we show that a non-uniformly hyperbolic horseshoe of an area-preserving real-analytic diffeomorphism is the maximal invariant subset of open sets of almost full Lebesgue measure.

More concretely, in our note [2], we proved that the non-uniformly hyperbolic horseshoes of Palis–Yoccoz [5] occur for many members of the standard family $\varphi_\lambda(x, y) = (-y + 2x + \lambda\sin(2\pi x), x)$ on the two-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. In fact, it was shown that, for all $k \in \mathbb{R}$ sufficiently large, there exists a subset $L \subset (k - \frac{4}{k^1/3}, k + \frac{4}{k^{1/3}})$ of positive Lebesgue measure such that, for all $r \in L$, the maximal invariant subset $\Lambda_r = \bigcap_{n\in\mathbb{Z}} \varphi_r^{-n}(U_r)$ is a non-uniformly hyperbolic horseshoe for a certain choice of open set $U_r \subset \mathbb{T}^2$ with total area $\frac{25}{256} + O(\frac{1}{k^{2/3}})$.

One of our goals here is to show that the open sets U_r above (whose areas are about 9.7% of the total area of the two-torus) can be replaced by open sets of almost full area.

Actually, it is not hard to see that this fact is a consequence[5] of the following general statement.

Theorem A.1. *Let $\varphi : M^2 \to M^2$ be an aperiodic diffeomorphism of a compact surface M^2. Suppose that φ possesses a non-uniformly hyperbolic horseshoe Λ. Then, for each $\varepsilon > 0$, there exists an open set W such that $M^2 \setminus W$ has area $<\varepsilon$ and $\Lambda = \bigcap_{n\in\mathbb{Z}} \varphi^{-n}(W)$.*

The proof of this result takes two steps. In Section A.1, we construct an open set of almost full area whose maximal invariant subset is empty; more concretely, we build a high 'Kakutani–Rokhlin tower' via an elementary probabilistic argument (à la Erdös),

[5]The maps φ_λ are aperiodic for $\lambda > 0$ because its powers are not the identity (as the origin is a hyperbolic fixed point), so that the set of periodic points must have zero Lebesgue measure (and, actually, Hausdorff dimension $\leqslant 1$) by real analyticity.

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

so that the desired open set is obtained by deleting the base from the tower. After that, in Section A.2, we 'add' this open set of almost full area to the definition of our non-uniformly hyperbolic set; since the maximal invariant subset of this open set is empty, we end up by obtaining exactly the same non-uniformly hyperbolic horseshoe as the maximal invariant subset of an open set of almost full area.

Remark A.2. After the first version of this appendix was ready, J. Bochi communicated to us that Theorem A.1 can also be derived (by slightly different methods) from [1, Theorem 2 and Remark 2].

A.1. Large open sets with empty maximal invariant subsets

Lemma A.3. *Under the same assumptions of Theorem A.1, for each* $\varepsilon > 0$*, there exists* $N \in \mathbb{N}$ *and an open set* V *such that* $M^2 \setminus V$ *has area* $<\varepsilon$ *and*

$$\bigcap_{|n|\leqslant N} \varphi^{-n}(V) = \emptyset.$$

Proof. For the sake of simplicity, we restrict ourselves to the case $M^2 = \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ equipped with the Lebesgue measure, Leb.

Let $\varepsilon > 0$ be given and consider $N \in \mathbb{N}$ large. Since φ is aperiodic, the compact set $K_N := \{x \in M^2 : \varphi^m(x) = x$ for some $m \leqslant N\}$ has zero Lebesgue measure. Thus, we can fix $\delta > 0$ such that $\mathrm{Leb}(V_\delta(K)) < \varepsilon/2$. Furthermore, given such a $\delta > 0$, we can choose $\delta/2 > \mu > 0$ such that if $y \in M^2 \setminus V_\delta(K)$, then $\varphi^{-j}(y) \in M^2 \setminus V_{2\mu}(K)$ for each $0 \leqslant j < N$. Finally, given $\mu > 0$, we can select $\mu > \eta > 0$ such that if $z \in M^2 \setminus V_\mu(K)$, then the sets $\varphi^j(\overline{B(z, \eta)})$, $0 \leqslant j < N$, are pairwise disjoints.

Given $Y \subset M^2 \setminus V_\delta(K)$, we claim that

$$\int_{M^2\setminus V_\mu(K)} \mathrm{Leb}\left(Y \cap \bigcup_{j=0}^{N-1} \varphi^j(B(x,\eta))\right) dx = N\pi\eta^2 \mathrm{Leb}(Y). \tag{A.1}$$

Indeed, note that $\mathrm{Leb}\ (Y \cap \bigcup_{j=0}^{N-1} \varphi^j(B(x,\eta))) = \sum_{j=0}^{N-1} \mathrm{Leb}(Y \cap \varphi^j(B(x,\eta)))$ for all $x \in M^2 \setminus V_\mu(K)$, so that

$$\int_{M^2\setminus V_\mu(K)} \mathrm{Leb}\left(Y \cap \bigcup_{j=0}^{N-1} \varphi^j(B(x,\eta))\right) dx = \sum_{j=0}^{N-1} \int_{M^2\setminus V_\mu(K)} \int_Y \chi_{\varphi^j(B(x,\eta))}(y)\, dy\, dx.$$

By Fubini's theorem, we have

$$\int_{M^2\setminus V_\mu(K)} \mathrm{Leb}\left(Y \cap \bigcup_{j=0}^{N-1} \varphi^j(B(x,\eta))\right) dx = \sum_{j=0}^{N-1} \int_Y \int_{M^2\setminus V_\mu(K)} \chi_{B(\varphi^{-j}(y),\eta)}(x)\, dx\, dy.$$

Since $B(\varphi^{-j}(y),\eta)\subset M^2\setminus V_\mu(K)$ for $0\leqslant j<N$ (because $y\in Y\subset M^2\setminus V_\delta(K)$ and $\eta<\mu$), we get that

$$\int_{M^2\setminus V_\mu(K)} \mathrm{Leb}\left(Y\cap\bigcup_{j=0}^{N-1}\varphi^j(B(x,\eta))\right)dx = \sum_{j=0}^{N-1}\int_Y \mathrm{Leb}(B(\varphi^{-j}(y),\eta))\,dy$$
$$= \sum_{j=0}^{N-1}\int_Y \pi\eta^2\,dy = N\pi\eta^2\mathrm{Leb}(Y).$$

In other terms, we showed (A.1).

Next, we affirm that, for each $m\in\mathbb{N}$, there are $x_1,\ldots,x_m\in M^2$ such that

$$\mathrm{Leb}\left((M^2\setminus V_\delta(K))\Bigg\backslash\bigcup_{i=1}^{m}\bigcup_{j=0}^{N-1}\varphi^j(B(x_i,\eta))\right)\leqslant(1-\pi N\eta^2)^m. \tag{A.2}$$

In fact, let us prove this fact by induction: for $m=0$, the affirmation is obvious; assuming that it holds for m, we employ (A.1) with

$$Y=Y_m:=(M^2\setminus V_\delta(K))\Bigg\backslash\bigcup_{i=1}^{m}\bigcup_{j=0}^{N-1}\varphi^j(B(x_i,\eta))$$

in order to obtain $x_{m+1}\in M^2$ such that

$$\mathrm{Leb}\left(Y_m\cap\bigcup_{j=0}^{N-1}\varphi^j(B(x_{m+1},\eta))\right)\geqslant\pi N\eta^2\mathrm{Leb}(Y_m)$$

and, *a fortiori*,

$$\mathrm{Leb}\left((M^2\setminus V_\delta(K))\Bigg\backslash\bigcup_{i=1}^{m+1}\bigcup_{j=0}^{N-1}\varphi^j(B(x_i,\eta))\right)=\mathrm{Leb}\left(Y_m\Bigg\backslash\bigcup_{j=0}^{N-1}\varphi^j(B(x_{m+1},\eta))\right)$$
$$\leqslant(1-\pi N\eta^2)\mathrm{Leb}(Y_m)$$
$$\leqslant(1-\pi N\eta^2)^{m+1},$$

so that the induction argument is complete.

Finally, let us construct the open set V satisfying the properties in the statement of the lemma. In this direction, we apply (A.2) with $m:=\lfloor\frac{1}{\pi\sqrt{N}\eta^2}\rfloor$ and we set

$$V:=\bigcup_{i=1}^{m}\bigcup_{j=0}^{N-1}\varphi^j(B(x_i,\eta))\Bigg\backslash\bigcup_{i=1}^{m}\overline{B(x_i,\eta)}.$$

Since $\mathrm{Leb}(V_\delta(K))<\varepsilon/2$, $\mathrm{Leb}(\bigcup_{i=1}^m\overline{B(x_i,\eta)})\leqslant m\pi\eta^2\leqslant 1/\sqrt{N}$ and, by (A.2), $\mathrm{Leb}(Y_m)\leqslant(1-\pi N\eta^2)^m\sim e^{-\frac{\pi N\eta^2}{\pi\sqrt{N}\eta^2}}=e^{-\sqrt{N}}$, we have that

$$\mathrm{Leb}(M^2\setminus V)\leqslant\frac{\varepsilon}{2}+\frac{1}{\sqrt{N}}+e^{-\sqrt{N}}<\varepsilon.$$

Also, $\bigcap_{j=0}^{N-1}\varphi^j(V)=\emptyset$ (by definition). This proves the lemma. □

https://doi.org/10.1017/S1474748019000185 Published online by Cambridge University Press

A.2. Proof of Theorem A.1

Let $U \subset M^2$ be an open set whose maximal invariant subset $\Lambda = \bigcap_{n\in\mathbb{Z}} \varphi^{-n}(U)$ is a non-uniformly hyperbolic horseshoe associated with an aperiodic diffeomorphism φ.

Given $\varepsilon > 0$, consider the integer $N \in \mathbb{N}$ and the open subset $V \subset M^2$ provided by Lemma A.3.

Since Λ is compact, we can select a neighborhood $\widetilde{U}$ of Λ such that $\bigcup_{n=-N}^{N} \varphi^{-n}(\widetilde{U}) \subset U$.

Let $W := \widetilde{U} \cup V$. Note that $M^2 \setminus W$ has area $<\varepsilon$ (because $V \subset W$). Thus, the proof of the theorem will be complete once we show that

$$X := \bigcap_{n\in\mathbb{Z}} \varphi^{-n}(W) = \Lambda.$$

Observe that $\Lambda \subset X$, so that our task is reduced to prove that $X \subset \Lambda$. For this sake, we consider $x \in X$. Since $\bigcap_{|n|\leqslant N} \varphi^{-n}(V) = \emptyset$, there exists $|n| \leqslant N$ such that $\varphi^n(x) \in \widetilde{U}$, and, *a fortiori*, $x \in U$. In other words, we showed that $X \subset U$. By invariance, we get the desired conclusion, namely $X \subset \bigcap_{n\in\mathbb{Z}} \varphi^{-n}(\widetilde{U}) = \Lambda$.

References

1. A. Avila and J. Bochi, A generic C^1 map has no absolutely continuous invariant probability measure, *Nonlinearity* **19** (2006), 2717–2725.
2. C. Matheus, C. G. Moreira and J. Palis, Non-uniformly hyperbolic horseshoes in the standard family, *C. R. Math. Acad. Sci. Paris* **356** (2018), 146–149.
3. C. Matheus and J. Palis, An estimate on the Hausdorff dimension of stable sets of non-uniformly hyperbolic horseshoes, *Discrete Contin. Dyn. Syst.* **38** (2018), 431–448.
4. C. Matheus, J. Palis and J.-C. Yoccoz, The Hausdorff dimension of stable sets of non-uniformly hyperbolic horseshoes, work in progress.
5. J. Palis and J.-C. Yoccoz, Non-uniformly hyperbolic horseshoes arising from bifurcations of Poincaré heteroclinic cycles, *Publ. Math. Inst. Hautes Études Sci.* **110** (2009), 1–217.