

CONTINUITY PROPERTIES OF LYAPUNOV EXPONENTS FOR SURFACE DIFFEOMORPHISMS

JÉRÔME BUZZI, SYLVAIN CROVISIER, AND OMRI SARIG

ABSTRACT. We study the entropy and Lyapunov exponents of invariant measures μ for smooth surface diffeomorphisms f , as functions of (f, μ) . The main result is an inequality relating the discontinuities of these functions. One consequence is that for a C^∞ surface diffeomorphisms, on any set of ergodic measures with entropy bounded away from zero, continuity of the entropy implies continuity of the exponents. Another consequence is the upper semi-continuity of the Hausdorff dimension on the set of ergodic invariant measures with entropy bounded away from zero. We also obtain a new criterion for the existence of SRB measures with positive entropy.

INTRODUCTION

Entropy and Lyapunov exponents play a major role in the study of differentiable dynamical systems, and their dependence on the measure and the map is of great interest. This dependence is sometimes continuous, but not always (for entropy, see [33, 35, 12, 9, 17, 10], and for Lyapunov exponents, see [44, 20, 6, 7, 1, 47]). While there are many works relating the values of the entropy to the values of the Lyapunov exponents [43, 39, 31, 27], the relation between the (dis)continuity of these objects as functions of the measure and the diffeomorphism has not yet been studied. The purpose of this work is to fill this gap, in the smooth two-dimensional case. For instance, we show:

Theorem. *Let f be a C^∞ diffeomorphism of a compact surface without boundary. Let ν_1, ν_2, \dots be ergodic measures for f , which converge in the weak-* topology to an ergodic measure μ with positive entropy.*

If the entropy of ν_k converges to the entropy of μ , then the Lyapunov exponents of ν_k converge to the Lyapunov exponents of μ .

This has strong dynamical consequences, some of which we will discuss here, and some of which we will discuss in a companion paper. For example, we have the following application to the problem of the existence of SRB measures. Let $\delta^u(\mu)$ denote the unstable dimension of μ in the sense of Ledrappier and Young (see Section 1.5).

Corollary. *Let f be a C^∞ diffeomorphism of a compact surface without boundary. If there exist ergodic invariant probability measures ν_k , with entropy uniformly bounded away from 0, and such that $\delta^u(\nu_k) \rightarrow 1$, then f admits an ergodic SRB measure with positive entropy.*

Date: March 4, 2021.

2010 Mathematics Subject Classification. 37C40, 37D30, 37A35, 37D35.

Keywords. Dynamical systems; smooth ergodic theory; entropy; Lyapunov exponents.

J.B. was partially supported by the ISDEEC project ANR-16-CE40-0013.

S.C. was partially supported by the ERC project 692925 *NUHGD*.

Part of this work was done when O.S. was visiting Université Paris-Sud and IHÉS, and he would like to thank these institutions for their hospitality and excellent working conditions. O.S. also acknowledges partial support of ISF grant 1149/18.

For other consequences, including the upper semi-continuity of the unstable dimension and of the Hausdorff dimension of ergodic measures with positive entropy, see Section 1.5.

These results follow from inequalities between the multiplicative size of the defects in continuity of the entropy and the top Lyapunov exponent. These inequalities, which are the main results of this work, are described in detail in the next section.

1. MAIN RESULTS

Throughout this paper, M is a two-dimensional compact C^∞ Riemannian manifold without boundary. Let $\text{Diff}^r(M)$ denotes the class of C^r diffeomorphisms on M (see §3.1).

Suppose μ is an f -invariant probability measure for some $f \in \text{Diff}^1(M)$. The *Kolmogorov-Sinai entropy* of μ will be denoted by $h(f, \mu)$. Almost every $x \in M$ has two well-defined Lyapunov exponents $\lambda^-(f, x) \leq \lambda^+(f, x)$. The *Lyapunov exponents* of μ are defined by

$$\lambda^-(f, \mu) := \lambda_\mu^- := \int \lambda^-(f, x) d\mu, \quad \lambda^+(f, \mu) := \lambda_\mu^+ := \int \lambda^+(f, x) d\mu.$$

We are interested in the regularity of $(f, \mu) \mapsto h(f, \mu)$ and $(f, \mu) \mapsto \lambda^\pm(f, \mu)$.

In the C^∞ case, these functions are semi-continuous. Specifically, suppose $f_k, f \in \text{Diff}^\infty(M)$, ν_k are ergodic f_k -invariant measures, $f_k \rightarrow f$ in C^∞ and $\nu_k \rightarrow \mu$ weak-*. Then

$$\begin{aligned} \limsup_{k \rightarrow \infty} h(f_k, \nu_k) &\leq h(f, \mu), \\ \limsup_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) &\leq \lambda^+(f, \mu), \quad \liminf_{k \rightarrow \infty} \lambda^-(f_k, \nu_k) \geq \lambda^-(f, \mu). \end{aligned} \tag{1}$$

(See Section 1.6 for the history of these results.)

By (1) and Ruelle's inequality, if $\limsup_k h(f_k, \nu_k) > 0$, then $\lambda^+(f, \mu) \geq h(f, \mu) \geq \limsup_k h(f_k, \nu_k) > 0$, and

$$\frac{\limsup_{k \rightarrow \infty} \lambda^+(f_k, \nu_k)}{\lambda^+(f, \mu)}, \quad \frac{\limsup_{k \rightarrow \infty} h(f_k, \nu_k)}{h(f, \mu)} \in (0, 1].$$

We call these quantities the *discontinuity ratios*, and think of them as measures for the difference between the two sides of the inequalities in (1).

We will provide inequalities relating the discontinuity ratio of the entropy to the discontinuity ratio of λ^+ . (It is enough to consider λ^+ , because $\lambda^-(f, \mu) = -\lambda^+(f^{-1}, \mu)$.)

1.1. The ergodic C^∞ case. Our results are simplest and strongest when the maps are C^∞ and the limiting measure is ergodic:

Theorem A. *For every $k \geq 1$, let $f_k \in \text{Diff}^\infty(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose*

- $\lim_k \lambda^+(f_k, \nu_k)$ and $\lim_k h(f_k, \nu_k)$ exist and are positive,
- f_k converge in the C^∞ topology to a diffeomorphism $f \in \text{Diff}^\infty(M)$,
- ν_k converge weak-* to a probability measure μ (necessarily f -invariant).

If μ is f -ergodic, then $\lim_{k \rightarrow \infty} \frac{h(f_k, \nu_k)}{h(f, \mu)} \leq \lim_{k \rightarrow \infty} \frac{\lambda^+(f_k, \nu_k)}{\lambda^+(f, \mu)}$.

The following result is an immediate consequence of this and (1), and was the original aim of our work:

Corollary 1.1. *For every $k \geq 1$, let $f_k \in \text{Diff}^\infty(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose $f_k \rightarrow f$ in the C^∞ topology, and $\nu_k \rightarrow \mu$ weak-* where μ is an f -ergodic invariant measure with positive entropy. If $h(f_k, \nu_k) \rightarrow h(f, \mu)$, then:*

$$\lambda^+(f_k, \nu_k) \rightarrow \lambda^+(f, \mu) \text{ and } \lambda^-(f_k, \nu_k) \rightarrow \lambda^-(f, \mu).$$

The result mentioned in the introduction is the special case $f_1 = f_2 = \dots = f$.

1.2. The ergodic C^r case. The following result extends Theorem A to the C^r case up to an extra term similar to what happens in Yomdin's theory. We define the *asymptotic dilation* of a C^1 map $f : M \rightarrow M$ to be

$$\lambda(f) := \lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df^n\|_{\text{sup}}, \text{ where } \|Df\|_{\text{sup}} := \sup_{x \in M} \sup_{\substack{v \in T_x M \\ v \neq 0}} \frac{\|Df_x v\|_{f(x)}}{\|v\|_x}. \quad (2)$$

Since M is compact, $\lambda(f)$ is independent of the choice of a Riemannian metric $\|\cdot\|_x$.

Theorem B. *Fix $r > 2$. For every $k \geq 1$, let $f_k \in \text{Diff}^r(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose*

- $\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k)$ and $\lim_{k \rightarrow \infty} h(f_k, \nu_k)$ exist and are positive,
- $f_k \rightarrow f$ in the C^r -topology,
- ν_k converge weak-* to a probability measure μ .

If μ is f -ergodic and has positive entropy, then

$$\frac{\lim_{k \rightarrow \infty} h(f_k, \nu_k)}{h(f, \mu)} - \frac{1}{h(f, \mu)} \frac{\lambda(f) + \lambda(f^{-1})}{r-1} \leq \frac{\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k)}{\lambda^+(f, \mu)}.$$

By [10], the condition $h(f, \mu) > 0$ holds once $\lim_{k \rightarrow \infty} h(f_k, \nu_k) > \frac{\min(\lambda(f), \lambda(f^{-1}))}{r}$.

As we will explain in Section 1.4, the smoothness index r above does not need to be an integer.

1.3. The non-ergodic case. The assumption that the limiting measure is ergodic is often difficult to check, and we now explain what can be said in its absence (a more general but also more technical result, Theorem D, will be given in section 7).

Theorem C. *Fix $r > 2$. For every $k \geq 1$, let $f_k \in \text{Diff}^r(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose*

- $\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k)$ and $\lim_{k \rightarrow \infty} h(f_k, \nu_k)$ exist and are positive,
- $f_k \rightarrow f$ in the C^r -topology,
- $\nu_k \rightarrow \mu$ weak-* for some f -invariant probability measure μ (perhaps non-ergodic).

If $\lim_{k \rightarrow \infty} h(f_k, \nu_k) > \frac{\lambda(f) + \lambda(f^{-1})}{r-1}$ then there exist $\beta \in (0, 1]$, and two f -invariant probability measures μ_0, μ_1 with $h(f, \mu_1) > 0$ such that $\mu = (1 - \beta)\mu_0 + \beta\mu_1$ and

$$\frac{\lim_{k \rightarrow \infty} h(f_k, \nu_k)}{h(f, \mu_1)} - \frac{1}{h(f, \mu_1)} \frac{\lambda(f) + \lambda(f^{-1})}{r-1} \leq \beta = \frac{\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k)}{\lambda^+(f, \mu_1)}. \quad (3)$$

Moreover $\lambda^+(f, x) > 0 \geq \lambda^-(f, x)$ for μ_1 -a.e. $x \in M$.

Note that for C^∞ diffeomorphisms, the term $(\lambda(f) + \lambda(f^{-1}))/r$ can be replaced by zero, because the theorem can be applied with r arbitrarily large.

The decomposition $\mu = (1 - \beta)\mu_0 + \beta\mu_1$ depends on the sequences $(f_k)_{k \geq 1}$, $(\nu_k)_{k \geq 1}$, and not just on their limits. We give a heuristic description of this decomposition in Section 2.2.

1.4. Additional comments. We now supplement Theorems A, B and C by some examples, comments, strengthenings, and generalizations. The proofs can be found in Section 8.

1.4.1. Examples of discontinuities. Theorem A is sharp in the following sense:

Example 1.2. *For every $0 < \alpha \leq \beta \leq 1$, there exist a C^∞ surface diffeomorphism f and a sequence of ergodic and invariant measures ν_k converging weak-* to an ergodic invariant probability measure μ such that $\lim_k h(f, \nu_k) > 0$, and*

$$\lim_{k \rightarrow \infty} h(f, \nu_k)/h(f, \mu) = \alpha, \quad \lim_{k \rightarrow \infty} \lambda^+(f, \nu_k)/\lambda^+(f, \mu) = \beta.$$

1.4.2. Variant inequality. Theorem A does not use the symmetry between a diffeomorphism and its inverse. When $0 < -\lambda^-(f, \mu) < \lambda^+(f, \mu)$, this symmetry yields a sharper bound:

Corollary 1.3. *Under the assumptions of Theorem A,*

$$\lambda^+(f, \mu) - \lim_k \lambda^+(f_k, \nu_k) \leq |\lambda^-(f, \mu)| \left(1 - \frac{\lim_k h(f_k, \nu_k)}{h(f, \mu)} \right). \quad (4)$$

1.4.3. Sequences of non-ergodic measures. Our results can be extended to the case when the invariant measures ν_k are not ergodic, but this requires stronger assumptions on the Lyapunov exponents of ν_k , which in the non-ergodic case are functions and not constants. See Corollary 8.5.

1.4.4. Lifted version. f induces a dynamical system \widehat{f} on the projective tangent bundle \widehat{M} , see Section 3.2. It turns out that $\mu = (1 - \beta)\mu_0 + \beta\mu_1$ is a projection of a decomposition of a limit point $\widehat{\mu} := \lim \widehat{\nu}_{k_i}$, where $\widehat{\nu}_k$ are lifts of ν_k to \widehat{f} -invariant measures on \widehat{M} . The decomposition of $\widehat{\mu}$ contains more information than the decomposition of μ , and leads to a stronger statement, Theorem D, in Section 7. This strengthening is essential to the proof of Corollary 8.5 on the case when ν_k are not ergodic.

1.4.5. Convergence of C^r -diffeomorphisms. In finite differentiability, Theorem D allows a weaker convergence assumption (denoted by $f_k \xrightarrow{r-\text{bd}} f$, see Section 3.1), and r does not have to be an integer.

1.4.6. Entropy upper semi-continuity in the C^r case. It is well-known that for C^r -diffeomorphisms the entropy may fail to be upper semi-continuous, but the defect in upper semi-continuity can be bounded (see the discussion in Section 1.6 below). This bound manifests itself in Theorems B and C in the expression $\frac{\lambda(f) + \lambda(f^{-1})}{r-1}$. But our proof gives a slightly stronger bound $\frac{\lambda(\widehat{f})}{r-1}$, in terms of the dynamics of \widehat{f} on the projective tangent bundle, see §3.2, §3.5, and Theorem D.

In some special cases, even this stronger bound can be improved. For instance, if $\lambda^+(f_k, \nu_k) \rightarrow \lambda^+(f, \mu)$, then

$$\lim_{k \rightarrow \infty} h(f_k, \nu_k) \leq h(f, \mu) + \frac{\min\{\lambda(f), \lambda(f^{-1})\}}{r}, \quad (5)$$

which is stronger than the assertion of Theorem B. When $f_1 = f_2 = \dots = f$, (5) is a refinement of a classical inequality of Yomdin and Newhouse; it follows from bounds on the tail entropy, which were first written explicitly in [10], and which are consequences of the Downarowicz variational principle [16]. When the sequence (f_k) is non-constant, it follows from a bound on robust tail entropy, which can be shown using techniques in [11]. We thank David Burguet for explaining this to us.

1.5. Applications to dimension theory and to SRB measures. The Hausdorff dimension of a Borel measure μ on M is defined to be the infimum of the Hausdorff dimensions of all Borel sets of full μ -measure [49, p.115]. We denote this by $\text{HD}(\mu)$.

Corollary 1.4. *For every $k \geq 1$, let $f_k \in \text{Diff}^\infty(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose $f_k \rightarrow f$ in the C^∞ topology, $\nu_k \rightarrow \mu$ weak-*, and $\liminf_{k \rightarrow \infty} h(f_k, \nu_k) > 0$. If μ is ergodic, then $\limsup_{k \rightarrow \infty} \text{HD}(\nu_k) \leq \text{HD}(\mu)$.*

Proof. One can always take a subsequence such that $\lim_k h(f_k, \nu_k) > 0$, $\lim_k \lambda^+(f_k, \nu_k) > 0$ and $\lim_k \lambda^-(f_k, \nu_k) < 0$ exist, and such that $\text{HD}(\nu_k)$ converges to the limsup of the initial sequence. In [49], Young gives the following formula for the Hausdorff dimension:

$$\text{HD}(\mu) = h(f, \mu) \left(1/\lambda^+(f, \mu) + 1/|\lambda^-(f, \mu)| \right). \quad (6)$$

Writing $\lim_k \frac{h(f_k, \nu_k)}{\lambda^+(f, \nu_k)} = \left(\frac{h(f, \mu)}{\lambda^+(f, \mu)} \right) \left(\frac{\lim_k h(f_k, \nu_k)}{h(f, \mu)} \right) \left(\frac{\lim_k \lambda^+(f_k, \nu_k)}{\lambda^+(f, \mu)} \right)^{-1}$, we conclude from Theorem A that $\lim_k \frac{h(f_k, \nu_k)}{\lambda^+(f, \nu_k)} \leq \frac{h(f, \mu)}{\lambda^+(f, \mu)}$. Working with f_k^{-1} and f^{-1} we obtain in a similar way that $\lim_k \frac{h(f_k, \nu_k)}{|\lambda^-(f, \nu_k)|} \leq \frac{h(f, \mu)}{|\lambda^-(f, \mu)|}$. \square

Suppose $f \in \text{Diff}^2(M)$. An f invariant probability measure is called a *Sinai-Ruelle-Bowen (SRB) measure*, if $\lambda^+(f, x) > 0$ μ -a.e., and if the conditional measures of μ obtained by disintegrating it with respect to a measurable partition subordinated to the lamination by unstable manifolds are a.e. absolutely continuous with respect to the induced Riemannian measures. We recall two classical characterizations of SRB measures from [31] and [27].

Suppose $\lambda^+(f, x) > 0$ μ -almost everywhere. The *geometric pressure* of μ is

$$P^u(\mu) = P^u(f, \mu) := h(f, \mu) - \lambda^+(f, \mu).$$

By Ruelle's inequality, $P^u(\mu) \leq 0$ for all invariant probability measures, and by Ledrappier-Young [31], $P^u(\mu) = 0$ iff μ is an SRB measure.

Next suppose μ is ergodic and $\lambda^+(f, \mu) > 0$. The *unstable dimension* of μ is

$$\delta^u(f, \mu) := \frac{h(f, \mu)}{\lambda^+(f, \mu)}.$$

By Ledrappier-Young Theory [27], $\delta^u(f, \mu)$ is the a.s. value of the Hausdorff dimension of the conditional measures of μ on a measurable partition subordinated to the local unstable manifolds, and μ is an SRB measure iff $\delta^u(\mu) = 1$.

Theorem A immediately implies the following:

Corollary 1.5. *For every $k \geq 1$, let $f_k \in \text{Diff}^\infty(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose $f_k \rightarrow f$ in the C^∞ topology, $\nu_k \rightarrow \mu$ weak-*, and $\lim h(f_k, \nu_k)$ exists and is positive. If μ is ergodic, then:*

- (1) $\limsup_{k \rightarrow \infty} \delta^u(\nu_k) \leq \delta^u(\mu)$
- (2) $\limsup_{k \rightarrow \infty} P^u(f_k, \nu_k) \leq P^u(f, \mu) \frac{\lim_{k \rightarrow \infty} h(f_k, \nu_k)}{h(f, \mu)}$
- (3) *if $P^u(f_k, \nu_k) \rightarrow 0$ or $\delta^u(\nu_k) \rightarrow 1$, then μ is an SRB measure.*

Proof. The proof of Corollary 1.4 also shows (1); (1) \Rightarrow (2) is a simple algebraic manipulation; and (1)+(2) \Rightarrow (3) by the Ledrappier-Young characterizations of SRB measures as measures with zero pressure and/or unstable dimension equal to one. \square

We can remove the assumption that μ is ergodic, using Theorem C:

Corollary 1.6. *For every $k \geq 1$, let $f_k \in \text{Diff}^\infty(M)$ and let ν_k be an f_k -ergodic invariant measure. Suppose $f_k \rightarrow f$ in the C^∞ topology, $\nu_k \rightarrow \mu$ weak-*, and $\lim h(f_k, \nu_k)$ exists and is positive. Then there are ergodic components μ', μ'' of μ satisfying $\lambda^+(\mu'), \lambda^+(\mu'') > 0$ and*

$$\delta^u(\mu'') \geq \lim_{k \rightarrow \infty} \delta^u(\nu_k), \quad P^u(f, \mu') \geq \lim_{k \rightarrow \infty} P^u(f_k, \nu_k).$$

Proof. Consider the decomposition $\mu = (1-\beta)\mu_0 + \beta\mu_1$ and take suitable ergodic components of μ_1 . \square

The following statement implies the corollary in the introduction.

Corollary 1.7. *Let f be a C^∞ diffeomorphism of a compact smooth surface without boundary, and fix some $h > 0$. The following are equivalent:*

- (i) *f admits an SRB measure with entropy at least h ;*
- (ii) *$\sup\{P^u(f, \mu) : \mu \text{ ergodic measure for } f \text{ s.t. } h(f, \mu) \geq h\} = 0$;*
- (iii) *$\sup\{\delta^u(\mu) : \mu \text{ ergodic measure for } f \text{ s.t. } h(f, \mu) \geq h\} = 1$.*

Proof. (i) \Rightarrow (ii) is due to Ledrappier & Strelcyn [29], and (ii) \Leftrightarrow (iii) is trivial.

To see (ii) \Rightarrow (i), we take a sequence of f -ergodic measures ν_k with $h(f, \nu_k) \rightarrow h' \geq h$ and $P^u(f, \nu_k) \rightarrow 0$. We select a subsequence $(\nu_{k_i})_{i \geq 1}$ s.t. $\lambda^+ := \lim \lambda^+(\nu_{k_i})$ and $\mu := \lim \nu_{k_i}$ exist. By Ruelle's inequality, $\lambda^+ \geq h' > 0$, and by Theorem C, $\mu = \beta\mu_1 + (1-\beta)\mu_0$ where

$$\lim_{i \rightarrow \infty} \lambda^+(f, \nu_{k_i}) = \beta\lambda^+(f, \mu_1), \text{ and } \lim_{i \rightarrow \infty} h(f, \nu_{k_i}) \leq \beta h(f, \mu_1).$$

It follows that $0 = \lim_{i \rightarrow \infty} P^u(f, \nu_{k_i}) \leq \beta P^u(f, \mu_1) \leq 0$, where the last inequality is Ruelle's inequality for μ_1 . Since $\beta\lambda^+(f, \mu_1) = \lim \lambda^+(f, \nu_{k_i}) = \lambda^+ \neq 0$, it must be the case that $\beta \neq 0$, and $P^u(f, \mu_1) = 0$. If $\mu_1 = \int \mu'_\xi d\xi$ is the ergodic decomposition of μ_1 , then $\int P^u(f, \mu'_\xi) d\xi = P^u(f, \mu_1) = 0$. By Ruelle's inequality, the integrand is non-positive, so $P^u(f, \mu'_\xi) = 0$ for μ_1 -a.e. ergodic component. At the same time, $h(f, \mu_1) \geq \beta^{-1} \lim h(f_k, \nu_k) > h'$, so some of these ergodic components must have entropy $\geq h'$. Thus μ_1 has ergodic components with entropy bigger than h , and zero pressure. By Ledrappier-Young Theory, these are SRB measures with entropy bigger than h , and (i) is proved. \square

Notice that it is essential in this proof to be able to deal with non-ergodic limits, since we have no control of $\lim \nu_k$.

1.6. Related works. In this paper we relate the continuity properties of the entropy to those of the Lyapunov exponents. The continuity of these objects has been studied separately before in several works, which we now recall.

Entropy. In general, the entropy map $(f, \mu) \mapsto h(f, \mu)$ is not lower semi-continuous, even in the uniformly hyperbolic case. For example, it is easy to construct sequences of atomic measures (with zero entropy) on a basic set, which converge to limits with positive entropy.

However, for C^∞ diffeomorphisms on compact manifolds, the entropy map is upper semi-continuous: This is due to Newhouse [35]. For C^r diffeomorphisms with finite r , even upper semi-continuity may fail (for examples in dimension four see [33], and for examples in dimension two see [14]). However, the (additive) defect in semi-continuity:

$$D(f, \mu) := \limsup_{(g, \nu) \rightarrow (f, \mu)} h(g, \nu) - h(f, \mu)$$

can be bounded from above by $\min(\lambda(f), \lambda(f^{-1}))/r$, using Yomdin theory [35, 12, 10, 11]. A subject of more recent interest is the loss of semi-continuity due to non-compactness. This has been studied for countable Markov shifts [22, 23], geodesic flows on non-compact

homogeneous spaces [18, 25], and geodesic flows on non-compact manifolds with negative sectional curvatures [21, 42].

Lyapunov exponents. The top Lyapunov exponent map $(f, \mu) \mapsto \lambda^+(f, \mu)$ varies continuously for uniformly hyperbolic systems on surfaces. It even depends analytically on the diffeomorphism f [44]. Moreover if μ_{\max} is the unique measure of maximal entropy of a mixing Anosov surface diffeomorphism, then [24] (see also [41, 45]) implies that $|\lambda^+(\mu_{\max}) - \lambda^+(\nu)| \leq c\sqrt{|h(f, \mu_{\max}) - h(f, \nu)|}$, where c only depends on f .

In the non-uniformly hyperbolic case, the situation is different. For example, [6] proves that among conservative systems the Lyapunov exponents of the volume measure are discontinuous when the diffeomorphism varies in the C^1 -topology, unless they vanish.

We are not aware of other general results on the continuity of the Lyapunov exponents for general non-uniformly hyperbolic surface diffeomorphisms.

By contrast, much is known on the continuity of Lyapunov exponents of random products of independent identically distributed $SL(2, \mathbb{R})$ matrices, as functions of the underlying Bernoulli process, see [20, 7]. More general Hölder continuous matrix cocycles with holonomies are considered in [2], and a higher-dimensional extension has been announced in [47].

Dimension. L.-S. Young gave the famous formula (6) for the dimension of hyperbolic invariant measures in [49] in terms of the entropy and the Lyapunov exponents of the measure. For further dimension theoretic properties of hyperbolic invariant measures, see [4] and [3]. The continuity of the dimension of invariant sets and measures for hyperbolic systems have been considered in numerous works, for instance [37] proves that basic sets on surface have a Hausdorff dimension which varies continuously with the diffeomorphism, [5] proves that the supremum of the Hausdorff dimensions of ergodic measures on such a basic set is attained by a measure of maximal dimension and [3] discusses some non-uniformly hyperbolic cases.

2. A HEURISTIC OVERVIEW OF THE PROOF

All our results follow from Theorem C, and the remainder of the paper is dedicated to the proof of this theorem. Here we give a heuristic overview of the proof, in the special case when $f_1 = f_2 = \dots = f$ is a C^∞ diffeomorphism.

2.1. The origin of the discontinuities in λ^+ . As Furstenberg discovered, the Lyapunov exponents are easier to study in terms of the *projective dynamics* $\hat{f}(x, E) = (f(x), Df_x(E))$ on the projective tangent bundle

$$\widehat{M} := \{(x, E) : x \in M, E \subset T_x M \text{ is a one-dimensional linear space}\}.$$

Indeed by Ledrappier's work, a Lyapunov exponent of an f -ergodic measure μ is simply the integral of the continuous function

$$\varphi(x, E) := \log \|Df_x|_E\|,$$

with respect to the lift of μ to the bundle of the associated Oseledets spaces.

Suppose ν_k are ergodic measures with positive entropy such that $\nu_k \rightarrow \mu$ weak*, and suppose for the moment that μ is ergodic and with positive entropy. Since $\dim(M) = 2$, ν_k have two simple Lyapunov exponents, and there are exactly two ergodic lifts $\widehat{\nu}_k^+$ and $\widehat{\nu}_k^-$, one carried by the bundle \mathcal{E}^u of unstable Oseledets spaces, and the other carried by the bundle \mathcal{E}^s of the stable Oseledets spaces. (The third bundle \mathcal{E}^0 associated to the zero exponent has

measure zero for all lifts of ν_k .) Hence,

$$\lambda^\pm(f, \nu_k) = \int \varphi d\hat{\nu}_k^\pm.$$

Suppose $\hat{\nu}_k^+$ converge weak-star on \widehat{M} to an \hat{f} -invariant probability measure $\hat{\mu}$ (this is true for a subsequence). Since $\varphi : \widehat{M} \rightarrow \mathbb{R}$ is continuous,

$$\lim_{k \rightarrow \infty} \lambda^+(f, \nu_k) = \lim_{k \rightarrow \infty} \hat{\nu}_k(\varphi) = \hat{\mu}(\varphi).$$

The limiting measure $\hat{\mu}$ is a lift of μ , but this does not have to be the lift of μ to \mathcal{E}^u , $\hat{\mu}^+$. If $\hat{\mu}(\varphi) \neq \hat{\mu}^+(\varphi)$, then $\lim_{k \rightarrow \infty} \lambda^+(f, \nu_k) \neq \lambda^+(f, \mu)$.

It is certainly possible that $\hat{\mu} \neq \hat{\mu}^+$: The Oseledets bundle \mathcal{E}^u carrying the lifts $\hat{\nu}_k^+$ is not necessarily bounded away from \mathcal{E}^s , and some mass $0 \leq \rho \leq 1$ on \mathcal{E}^u can escape to \mathcal{E}^s .

Escape of mass to \mathcal{E}^s is reflected in long stretches of time when ν_k -typical orbits do not experience the exponential growth of \mathcal{E}^u -directions predicted by λ^+ . Instead, they see, temporarily, exponential decay at rate λ^- , cancelling some of the previous growth. If $\hat{\mu}^+, \hat{\mu}^-$ denote the two ergodic lifts, we must have $\hat{\mu} = (1 - \rho)\hat{\mu}^+ + \rho\hat{\mu}^-$ and thus,

$$\lim_{k \rightarrow \infty} \lambda^+(f, \nu_k) = (1 - \rho)\hat{\mu}^+(\varphi) + \rho\hat{\mu}^-(\varphi) = \lambda_\mu^+ - \rho(\lambda_\mu^+ - \lambda_\mu^-).$$

In the language of Theorem C (and since $\mu = \mu_1$ by ergodicity), the discontinuity ratio β is:

$$\beta := \frac{\lim_k \lambda^+(f, \nu_k)}{\lambda^+(f, \mu)} = 1 - \rho(1 - \lambda_\mu^- / \lambda_\mu^+).$$

(A different description of β will be given below.) So if μ is ergodic, then β is a function of ρ , whence of the amount of mass which escapes to \mathcal{E}^s .

In the case where μ is not ergodic, the different ergodic components of μ have to be considered, and some of them may have zero Lyapunov exponents. The way in which ν_k -typical orbits approximate those ergodic components determine the possible cancellations. So if μ is not ergodic, then β may depend on the entire sequence (ν_k) , not just on its limit μ .

2.2. Neutral blocks, the decomposition of μ , and the parameter β . Recall the measurable \hat{f} -invariant decomposition $\mathcal{X} = \mathcal{E}^s \cup \mathcal{E}^u \cup \mathcal{E}^0$ defined by the Oseledets theorem according to the sign of the limit $\frac{1}{n} \log \|Df_x^n|_E\|$. It has full measure with respect to any \hat{f} -invariant measure. To get quantitative estimates, we select compact subsets $K^* \subset \mathcal{E}^*$, for $*$ $\in \{s, u, 0\}$, from which the contraction, expansion, or “central” behavior of the sequence $\|Df_x^n|_E\|$ are uniformly controlled, and such that the $\hat{\mu}$ -measure of $K := K^s \cup K^u \cup K^0$ is close to 1. Since each \mathcal{E}^* is invariant, we can choose these compact sets to be nearly invariant: Points in a very small neighborhood stay close for a long time.

Hence, if \hat{x}_0 is a $\hat{\nu}_k^+$ -typical point for some very large k , its orbit under \hat{f} spends nearly all its time close to K and every visit in a small neighborhood of $K^s \cup K^0$ is the beginning of a long period of uniform contraction (or weak expansion/contraction). One expects no entropy creation not only during this period, but also during the “recovery period” which follows, i.e., until the expansion predicted by the Lyapunov exponent of ν_k cancels this period of contraction (or weak expansion/contraction).

We select such long time intervals along the orbit of \hat{x}_0 in the following greedy way. Fixing $\alpha > 0$ small and L large, an (α, L) -*neutral block* is a maximal interval of integers $(n_0, \dots, n_0 + \ell)$ such that $\ell \geq L$ and

$$\|Df_{f^{n_0}(x_0)}^n|_{E^u}\| \leq \exp(\alpha(n - n_0)) \text{ for all } 0 < n \leq \ell.$$

We will check that indeed, there is very little if any entropy creation during neutral blocks.

Our estimates will be in terms of the distribution of these long neutral blocks. Let $\widehat{x}_k = (x_k, E^u(x_k))$ be $\widehat{\nu}_k^+$ -generic points and let $\mathfrak{N}_{\alpha,L}(\widehat{x}_k)$ denote the union of all (α, L) -neutral blocks of the orbit of \widehat{x}_k . In Section 6 we show that it is possible to choose a subsequence $k_i \rightarrow \infty$ so that following limits make sense weak-* on \widehat{M} for $(\widehat{\nu}_{k_1}^+ \times \widehat{\nu}_{k_2}^+ \times \cdots)$ -a.e. $(\widehat{x}_{k_1}, \widehat{x}_{k_2}, \dots)$:

$$\begin{aligned}\widehat{m}_0 &:= \lim_{\alpha \rightarrow 0} \lim_{L \rightarrow \infty} \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j \in [0, N) \cap \mathfrak{N}_{\alpha, L}(\widehat{x}_k)} \delta_{\widehat{f}^j(\widehat{x}_{k_i})} \right) \\ \widehat{m}_1 &:= \lim_{\alpha \rightarrow 0} \lim_{L \rightarrow \infty} \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j \in [0, N) \setminus \mathfrak{N}_{\alpha, L}(\widehat{x}_k)} \delta_{\widehat{f}^j(\widehat{x}_{k_i})} \right).\end{aligned}$$

Notice that the sum of the two limits in the brackets is a.s. $\widehat{\nu}_{k_i}^+$, because this is the limit of the empirical measure of \widehat{x}_{k_i} , and \widehat{x}_{k_i} are all a.s. $\widehat{\nu}_{k_i}^+$ -generic. So

$$\widehat{m}_0 + \widehat{m}_1 = \lim_{k \rightarrow \infty} \widehat{\nu}_k^+ = \widehat{\mu}.$$

The measures $\widehat{m}_0, \widehat{m}_1$ are \widehat{f} -invariant. We will see that $\int \varphi d\widehat{m}_0 = 0$, and that \widehat{m}_1 is carried by \mathcal{E}^u . The decomposition $\mu = \beta\mu_1 + (1 - \beta)\mu_0$ in Theorem C is defined by

$$\beta := 1 - \widehat{\mu}_0(\widehat{M}), \quad (1 - \beta)\mu_0 := \widehat{\pi}_*(\widehat{\mu}_0), \quad \text{and} \quad \beta\mu_1 := \widehat{\pi}_*(\widehat{\mu} - \widehat{\mu}_0),$$

where $\widehat{\pi} : \widehat{M} \rightarrow M$ is the natural projection. Note that β is indeed the discontinuity ratio $\lim_k \lambda^+(f, \nu_k) / \lambda^+(f, \mu_1)$ and the quantity $1 - \beta$ coincides with the fraction of the time spent in maximal neutral blocks. The measures μ_0, μ_1 and β depend not just on $\widehat{\mu}$, but also on the way the measures $\widehat{\nu}_k^+$ accumulate on $\widehat{\mu}$.

2.3. Upper bound on the entropy. To complete the proof of the theorem it remains to show that $\lim_{k \rightarrow \infty} h(f, \nu_k) \leq \beta h(f, \mu_1)$. This is the heart of the proof, and where most of the difficulties lie. We use Ledrappier-Young Theory and Yomdin Theory.

- *Ledrappier-Young theory* bounds $h(f, \nu_k)$ by the exponential rate of growth of the minimal number of (n, ϵ) -balls needed to cover a definite fraction of a local unstable manifold $W_{loc}^u(x_k)$, where x_k is a fixed ν_k -typical point and the scale ϵ tends to zero. The “fraction” is measured using the conditional measure $\nu_{x_k}^u$ of ν_k on $W_{loc}^u(x_k)$. In particular, it suffices to follow points $x \in W_{loc}^u(x_k)$ with $T_x W_{loc}^u(x_k) = \mathcal{E}^u(x)$.
- *Yomdin theory* provides tools for controlling the number of (n, ϵ) -balls needed for such covers, for C^r maps. Instead of working with (n, ϵ) -balls, one works with parametrized pieces of unstable manifolds which lie inside (n, ϵ) -balls and which have uniformly bounded C^r size, and Yomdin Theory allows to bound the number of such pieces. Here the regularity assumptions on f come into play. The expression $\frac{\lambda(f) + \lambda(f^{-1})}{r-1}$ in (3) is due to Yomdin theory (see section 2.4).

Let us sketch our argument for the upper bound on the entropy using the neutral blocks. Since the unstable lift of $\nu_{x_k}^u$ -almost every point is $\widehat{\nu}_k^+$ -typical, neutral blocks represent roughly a fraction $1 - \beta$ of their time. During a neutral block, typical points on a small piece of $f^n(W_{loc}^u(x_k))$ do not separate much, therefore this piece remains small (or can be kept small by a subdivision into a small exponential number of pieces). For the rest of the time, these subcurves follow the ergodic components of μ_1 , hence they experience entropic

separation at an exponential rate given by $h(f, \mu_1)$. Since the time outside neutral blocks is a proportion β of the total time, this leads to the bound

$$h(f, \nu_k) \leq \beta h(f, \mu_1).$$

This argument explains the link between the entropy bound and the semicontinuity defect of the Lyapunov exponents.

This sketch glosses over several difficulties. We will only comment on the main issue: How to use non-expansion of the linearization Df at $(x_k, E^u(x_k))$ during a neutral block, to infer non-expansion of the map f itself on a small piece of $W_{loc}^u(x_k)$ during this neutral block. The difficulty is in controlling Df on $(x'_k, E^u(x'_k))$ for x'_k close to x_k .

2.4. Control of the expansion during neutral blocks. This is one of the most delicate points in the proof. To deduce the non-expansion of the small piece of the unstable manifold containing this point, we need to know that not only the diameter of this curve is small but that its tangent is almost constant too. This forces us to work with pieces of $W_{loc}^u(x_k)$ whose lifts to \widehat{M} are also small: The size in the fiber of \widehat{M} measures the variability of the tangent directions.

How small is small enough? To use information on Df to control what happens on $W_{loc}^u(x_k)$, we need the fluctuations of the tangent direction along any piece to be smaller than some $\widehat{\varepsilon} > 0$, determined (mostly) by the modulus of continuity of Df . Using the uniform continuity of the measurable unstable bundle on a set of large measure, we find an $\varepsilon > 0$ such that if the diameter of the projection to M is less than ε , then the fluctuation of the tangent is smaller than $\widehat{\varepsilon}$.

The price we pay for this solution is that we need to work with different scales in M and along the fibers of the bundle $\widehat{M} \rightarrow M$. This leads us to introduce *fibred* $(n, \varepsilon, \widehat{\varepsilon})$ -balls, and to work with Yomdin theory for $\widehat{f} : \widehat{M} \rightarrow \widehat{M}$. When dealing with diffeomorphisms of finite regularity, there is an additional price to pay: If f is C^r , then \widehat{f} is only C^{r-1} , and this accounts for the extra term from Yomdin theory

$$\frac{\widehat{\lambda}(\widehat{f})}{r-1} \leq \frac{\lambda(f) + \lambda(f^{-1})}{r-1} \quad \text{in (3).}$$

2.5. Organisation of the paper. The different ingredients of the proof appear as follows in the text.

Section 3: background on tangent dynamics and Lyapunov exponents.

Section 4: results from Ledrappier-Young and Yomdin theories on the entropy in differentiable dynamics.

Section 5: reparametrization lemmas estimating the entropy from neutral blocks and other time intervals.

Section 6: neutral decomposition of typical orbits.

Section 7: proof of the technical version of our main theorem.

Section 8: proof of the remaining statements.

A remark on style. Our constructions, estimating entropy for a sequence of measures converging to a nonergodic one, require many parameters. We have chosen to make the dependences as explicit as possible to help the reader check that there is no circular argument.

2.6. Standing notations for the duration of the paper. We collect here some notations that we will use frequently below.

- $|X|$ or $\text{Card}(X)$: the cardinality of a set X .

- M is a compact Riemannian C^∞ manifold without boundary, with tangent bundle TM , tangent spaces $T_x M$, and Riemannian norm $\|\cdot\|_x$. Derivatives of maps $f : M \rightarrow M$ are denoted by $Df : TM \rightarrow TM$ or $Df_x : T_x M \rightarrow T_{f(x)} M$.
- $h(f, \mu)$, λ_μ^\pm , λ_x^\pm , $TM = E^+ \oplus E^-$: the entropy, average and pointwise Lyapunov exponents, Oseledec splitting associated to a measure μ (also denoted $\lambda_\mu^{s/u}$, $E^{s/u}$ when the measure is hyperbolic of saddle type), see section 3.3.
- \widehat{M} , \widehat{f} , $\widehat{x} = (x, E)$: the projective tangent bundle, the lift of a diffeomorphism f and of a point x , see section 3.2.
- $B_f(x, n, \varepsilon)$, $r_f(n, \varepsilon, X)$: an (n, ε) -Bowen ball for f and the (n, ε) -covering number of a set $X \subset M$, see section 4.2.
- $\|Df\|_{\text{sup}}$: the sup-norm of the tangent map, see section 1.2.
- $\lambda(f) := \lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df^n\|_{\text{sup}}$: the asymptotic dilation.
- $\|f\|_{C^r}$: the C^r size of f , see section 3.1.
- $Q_{r,N}(f)$: the supremum of the C^r sizes of f, f^2, \dots, f^N and of the C^{r-1} semi-norm of $\widehat{f}, \widehat{f}^2, \dots, \widehat{f}^N$ see section 4.6.
- If v is a vector, then $\mathbb{R}.v := \{tv : t \in \mathbb{R}\}$.

3. TANGENT DYNAMICS AND THE SEMI-CONTINUITY OF LYAPUNOV EXPONENTS

Let M be a smooth compact Riemannian surface without boundary.

3.1. Review of the C^r size of maps. Let U be an open subset of \mathbb{R}^n .

Given $k \in \mathbb{N}$, we say that a map $F : U \rightarrow \mathbb{R}^d$ is C^k if for all $\omega \in (\mathbb{N} \cup \{0\})^n$ such that $|\omega| := \omega_1 + \dots + \omega_n = k$, the partial derivative

$$\partial^\omega F := \frac{\partial^{\omega_1 + \dots + \omega_n} F}{\partial^{\omega_1} x_1 \dots \partial^{\omega_n} x_n}$$

exists and is continuous on U . For any compact subset $K \subset U$, we then define the C^k size

$$\|F\|_{C^k, K} := \max_{1 \leq |\omega| \leq k} \max_{x \in K} \|\partial^\omega F(x)\|.$$

Given $\alpha \in (0, 1)$, we say that a map F is C^α if the following quantity is finite for any compact set $K \subset U$,

$$\|F\|_{C^\alpha, K} := \sup_{\substack{x, y \in K \\ x \neq y}} \frac{\|F(x) - F(y)\|}{\|x - y\|^\alpha}.$$

Given $r > 1$ which is not an integer, we decompose it as $r = k + \alpha$, with $k = \lfloor r \rfloor$ and $\alpha \in (0, 1)$. We say that F is C^r if it is C^k and each partial derivative $\partial^\omega F$, $|\omega| = k$ is C^α . For any compact set $K \subset U$, we define the C^r size

$$\|F\|_{C^r, K} := \|F\|_{C^k, K} + \max_{|\omega|=k} \|\partial^\omega F\|_{C^\alpha, K}.$$

Let Ω be a compact subset of \mathbb{R}^n which is equal to the closure of its interior (we mostly need $[0, 1]^n$). A map $F : \Omega \rightarrow \mathbb{R}^d$ is C^r if F has a C^r extension to an open neighborhood of Ω . In this case, the C^r size of F on Ω is

$$\|F\|_{C^r} := \sup_{K \subset \text{int}(\Omega)} \|F\|_{C^r, K}.$$

This (finite) quantity is independent of the extension of F to the neighborhood of Ω . Notice that the C^r size of a constant function is zero.

A C^r *structure* on a smooth manifold N is defined by a maximal atlas \mathfrak{A} with C^r changes of coordinates. A smooth manifold equipped with a C^r structure \mathfrak{A} is called a C^r manifold. A finite subset of \mathfrak{A} which covers N is called a C^r atlas of N .

Let N_1, N_2 be two compact C^r manifolds (later this will be M, \widehat{M} or the circle S^1), and let \mathcal{A}_i be finite C^r atlases of N_i . Let Ω be a compact subset of N_1 equal to the closure of its interior. We say that $f: \Omega \rightarrow N_2$ is a C^r *map* if each map $\chi_2^{-1} \circ f \circ \chi_1$, where χ_i ranges over \mathcal{A}_i , is C^r . The C^r *size* of f is:

$$\|f\|_{C^r} := \max_{\chi_1 \in \mathcal{A}_1, \chi_2 \in \mathcal{A}_2} \|\chi_2^{-1} \circ f \circ \chi_1\|_{C^r} < \infty.$$

Again, the constant map has size zero.

The quantity $\|f\|_{C^r}$ depends on the choice of atlases \mathcal{A}_i , but if N_i are compact, then finite atlases induce equivalent C^r sizes. In case $N_1 = S^1$, we will always use the Euclidean atlas.

Suppose $f_k, f \in \text{Diff}^r(M)$ and $1 \leq r < \infty$. We will say that f_k converges to f *uniformly in a C^r -bounded way*, if $f_k \rightarrow f$ uniformly, and $\sup_{k \geq 1} \|f_k\|_{C^r} < \infty$. We write in this case

$$f_k \xrightarrow{r\text{-bd}} f.$$

If M is compact, $f_k, f \in \text{Diff}^\infty(M)$, and $f_k \rightarrow f$ in C^∞ , then $f_k \xrightarrow{r\text{-bd}} f$ for all r finite.

The Arzela-Ascoli theorem implies the following.

Lemma 3.1. *Let N_1, N_2 be compact C^r manifolds, and $f, f_1, f_2, \dots: N_1 \rightarrow N_2$ be a collection of C^r maps such that (f_k) converges to f uniformly, and $\sup_k \|f_k\|_{C^r} < \infty$. Then (f_k) converges to f in the C^ℓ -topology for any $\ell < r$, $\ell \in \mathbb{N}$.*

Thus, if for some real $r > 1$ s.t. $r \notin \mathbb{N}$, $f, f_1, f_2, \dots \in \text{Diff}^r(M)$ where M is a compact manifold, then $f_k \xrightarrow{r\text{-bd}} f$ implies that $f_k \rightarrow f$ in the $C^{\lfloor r \rfloor}$ -topology.

3.2. The projective tangent bundle. Let

$$P_x M := \{E : E \text{ is a one-dimensional linear subspace of } T_x M\}.$$

$P_x M$ is the quotient of $T_x M \setminus \{0\}$ by the equivalence relation $v \sim w \iff \exists \lambda \neq 0, v = \lambda w$. It can also be viewed as the image of $\{v \in T_x M : \|v\|_x = 1\}$ by the two-to-one map $v \mapsto \text{Span}\{v\}$. These identifications allow us to endow $P_x M$ with a topology and with a smooth structure, and to identify the tangent spaces $T_E(P_x M)$ with $\{w \in T_x M : w \perp E\}$. We can also pull back the induced Riemannian inner product on $\{w \in T_x M : w \perp E\}$ to an inner product on $T_E(P_x M)$. This endows $P_x M$ with a Riemannian structure. The resulting Riemannian distance on $P_x M$ is simply $\text{dist}(E_1, E_2) = |\angle(E_1, E_2)|$. With this structure, $P_x M$ is isometric to the circle with perimeter π .

The *projective tangent bundle* (or just “projective bundle”) of M is the bundle $(\widehat{M}, \widehat{\pi}, M)$ where $\widehat{\pi}: \widehat{M} \rightarrow M$ is the natural projection $\widehat{\pi}(x, E) = x$, and

$$\widehat{M} := \{(x, E) : x \in M, E \text{ is a one-dimensional linear subspace of } T_x M\} = \bigsqcup_{x \in M} P_x M.$$

\widehat{M} is a smooth compact three-dimensional manifold. We endow it with the Riemannian metric $\sqrt{ds^2 + d\theta^2}$, where ds is the length element on M and $d\theta$ is the length element on $P_x M$. Points in \widehat{M} will be denoted by $\widehat{x} = (x, E)$.

Let $f: M \rightarrow M$ be a C^1 diffeomorphism. The *canonical lift* of f is the homeomorphism $\widehat{f}: \widehat{M} \rightarrow \widehat{M}$ given by

$$\widehat{f}(x, E) = (f(x), Df_x(E)). \quad (7)$$

If f is of class C^r , then \widehat{f} is of class C^{r-1} . Notice that $\widehat{\pi} \circ \widehat{f} = f \circ \widehat{\pi}$, and

$$\widehat{f}^n(x, E) = (f^n(x), (Df^n)_x(E)).$$

Every \widehat{f} -invariant probability measure $\widehat{\nu}$ on \widehat{M} projects to an f -invariant probability measure ν on M given by

$$\nu := \widehat{\pi}_*(\widehat{\nu}) := \widehat{\nu} \circ \pi^{-1}.$$

We call ν the *projection* of $\widehat{\nu}$, and $\widehat{\nu}$ a *lift* of ν . In what follows, when we “lift”, we always mean an \widehat{f} -invariant lift. The following lemma is a well-known consequence of the compactness of \widehat{M} .

Lemma 3.2. (1) *Every f -invariant probability measure ν has at least one lift $\widehat{\nu}$.*
 (2) *If ν is f -ergodic and $\widehat{\nu}$ lifts ν , then a.e. ergodic component of $\widehat{\nu}$ is a lift of ν .*
Hence every ergodic f -invariant probability measure has at least one ergodic lift.

3.3. Review of Lyapunov exponents. We review some facts on Lyapunov exponents in dimension two (see [46, Theorems 3.12 and 3.14]). Suppose $f \in \text{Diff}^1(M)$ and μ is an f -invariant Borel probability measure. Oseledets’ theorem asserts that for μ -a.e. x , $\lim_{|n| \rightarrow \infty} \frac{1}{n} \log \|Df_x^n v\|$ exists for all $v \in T_x M \setminus \{0\}$. The possible values of the limit are called the *Lyapunov exponents* of x . There are at most two such values. We denote them by $\lambda^+(f, x)$, $\lambda^-(f, x)$, or λ_x^+ , λ_x^- , with the convention

$$\lambda^+(f, x) \geq \lambda^-(f, x).$$

If $\lambda^+(f, x) \neq \lambda^-(f, x)$ then $T_x M = E^+(x) \oplus E^-(x)$ where

$$E^\pm(x) := \left\{ v \in T_x M \setminus \{0\} : \lim_{|n| \rightarrow \infty} \frac{1}{n} \log \|Df_x^n v\| = \lambda^\pm(f, x) \right\} \cup \{0\}.$$

The decomposition $T_x M = E^+(x) \oplus E^-(x)$ is called the *Oseledets splitting*.

If μ is *ergodic*, the functions $\lambda^+(f, x) \geq \lambda^-(f, x)$ are equal μ -almost everywhere to constants called the *Lyapunov exponents* of μ and denoted by $\lambda^+(f, \mu)$, $\lambda^-(f, \mu)$. If μ is not ergodic, the Lyapunov exponents of μ are defined by

$$\lambda^+(f, \mu) := \int \lambda^+(f, x) d\mu(x), \quad \lambda^-(f, \mu) := \int \lambda^-(f, x) d\mu(x).$$

In both cases, $\lambda^+(f, \mu) \geq \lambda^-(f, \mu)$.

By the subadditive ergodic theorem, the largest Lyapunov exponent also satisfies

$$\lambda^+(f, \mu) = \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|Df_y^n\| d\mu(y) = \inf_n \frac{1}{n} \int \log \|Df_y^n\| d\mu(y). \quad (8)$$

Throughout this paper, an ergodic invariant probability measure μ is called *hyperbolic* if one of its Lyapunov exponents is positive, and the other is negative (sometimes this is called hyperbolic of saddle-type). If μ is hyperbolic, we sometimes write $\lambda^u = \lambda^+$, $\lambda^s = \lambda^-$, $E^u = E^+$ and $E^s = E^-$.

3.4. Semi-continuity of Lyapunov exponents. We will use the dynamics of the projective tangent bundle to study the semi-continuity properties of $(f, \mu) \mapsto \lambda^+(f, \mu)$ (and by symmetry of $(f, \mu) \mapsto \lambda^-(f, \mu)$). The principal tool is the function

$$\varphi : \widehat{M} \rightarrow \mathbb{R}, \quad \varphi(x, E) = \log \|Df_x|_E\|.$$

Notice that if f is a C^1 diffeomorphism, then φ is bounded and uniformly continuous. We will make frequent use of the following identity:

$$\log \|Df^n|_E\| = \sum_{k=0}^{n-1} (\varphi \circ \hat{f}^k)(x, E) \quad ((x, E) \in \widehat{M}). \quad (9)$$

This is because E is a one-dimensional subspace of $T_x M$, and therefore by the chain rule $\|Df_x^n|_E\| = \prod_{k=0}^{n-1} \|Df_{f^k(x)}|_{Df_x^k(E)}\| = \prod_{k=0}^{n-1} \exp[(\varphi \circ \hat{f}^k)(x, E)]$. Equation (9) presents the subadditive cocycle $\log \|Df^n|_E\|$ for f as an additive cocycle for \hat{f} . See [28, Prop. 5.1 on p. 328] for a proof of a more general fact (and [19, Lemma 8.7] for the first use of a related idea).

Lemma 3.3. *Suppose $f \in \text{Diff}^1(M)$, and μ is an ergodic f -invariant probability measure. Then:*

- (1) λ is a Lyapunov exponent of μ iff μ has an \hat{f} -ergodic lift $\hat{\mu}$ s.t. $\int \varphi d\hat{\mu} = \lambda$;
- (2) If μ has two different Lyapunov exponents, then it has exactly two ergodic \hat{f} -invariant lifts:

$$\hat{\mu}^+ := \int_{\widehat{M}} \delta_{(x, E^+(x))} d\mu(x), \text{ and } \hat{\mu}^- := \int_{\widehat{M}} \delta_{(x, E^-(x))} d\mu(x).$$

Moreover $\int \varphi d\hat{\mu}^\pm = \lambda^\pm(f, \mu)$.

For any \hat{f} -invariant probability measure, it will be convenient to denote

$$\hat{\lambda}(\hat{f}, \hat{\mu}) := \int_{\widehat{M}} \varphi d\hat{\mu}.$$

When μ is hyperbolic, the lifts $\hat{\mu}^+, \hat{\mu}^-$ are called the unstable and stable lifts of μ .

Proof. If μ has equal Lyapunov exponents, (1) follows from eq. (9) and the ergodic theorem. Otherwise, by Oseledets theorem, there are two a.e. defined sections $x \mapsto E_x^\pm$ s.t. $Df_x E_x^\pm = E_{f(x)}^\pm$. Every \hat{f} -invariant probability measure carried by the graph of an invariant section $x \mapsto E_x$ is ergodic, and coincides with $\int_M \delta_{(x, E_x)} d\mu(x)$. So (1) and (2) follow from Lemma 3.2 and (9) (see [28]). \square

Corollary 3.4. *Suppose $f \in \text{Diff}^1(M)$ and μ is an f -invariant probability measure (not necessarily ergodic) s.t. $\lambda^+(f, x) > \lambda^-(f, x)$ for μ -a.e. $x \in M$. Then any \hat{f} -invariant lift $\hat{\mu}$ of μ is carried by $\text{graph}(E^+) \cup \text{graph}(E^-)$, and there are unique \hat{f} -invariant lifts $\hat{\mu}^+, \hat{\mu}^-$ s.t. $\hat{\mu}^+(\text{graph}(E^+)) = 1, \hat{\mu}^-(\text{graph}(E^-)) = 1$.*

Proof. By a general Borel construction, the graphs of E^+, E^- are measurable. There are unique lifts $\hat{\mu}^+, \hat{\mu}^-$ of μ to $\text{graph}(E^+), \text{graph}(E^-)$, and it is easy to check using the identity $Df_x(E_x^\pm) = E_{f(x)}^\pm$ that $\hat{\mu}^\pm$ are \hat{f} -invariant.

Now let $\hat{\mu}$ be an arbitrary lift of μ and consider its ergodic decomposition $\hat{\mu} = \int \hat{\mu}_\xi dm$. For almost every ξ , the ergodic measure $\mu_\xi = \hat{\mu}_\xi \circ \hat{\pi}^{-1}$ has two different exponents, hence, by Lemma 3.3, its lift $\hat{\mu}_\xi$ is some combination $a(\xi)\hat{\mu}_\xi^+ + (1-a(\xi))\hat{\mu}_\xi^-$ of its lifts $\hat{\mu}_\xi^\pm$ where $0 \leq a(\xi) \leq 1$. Observe that $a(\xi) = \hat{\mu}_\xi(\text{graph}(E^+))$, hence the function $\xi \mapsto a(\xi)$ is measurable. It follows that

$$\hat{\mu} = \int a(\xi)\hat{\mu}_\xi^+ + (1-a(\xi))\hat{\mu}_\xi^- dm.$$

In particular, any lift $\hat{\mu}$ is carried by the union of the graphs of E^+ and E^- . \square

For each $n \geq 1$, $(f, \mu) \mapsto \int \log \|Df_x^n\| d\mu(x)$ is continuous as a function on $\text{Diff}^1(M) \times \{\text{probability measure on } M\}$. Eq. (8) now give us the following “folklore” fact:

Theorem 3.5 (Upper semicontinuity). *Let $f_k \in \text{Diff}^1(M)$ be diffeomorphisms with ergodic invariant probability measures ν_k . If $f_k \xrightarrow{c^1} f$ and $\nu_k \xrightarrow{w^*} \mu$, then the largest Lyapunov exponents $\lambda^+(f_k, \nu_k)$ satisfy $\limsup_k \lambda^+(f_k, \nu_k) \leq \lambda^+(f, \mu)$.*

The next result computes the defect in continuity $\lambda^+(f, \mu) - \limsup \lambda^+(f_k, \nu_k)$ in terms of the dynamics on the projective bundle. It relates the defect in continuity to the escape of some of the mass of the lifts to $\text{graph}(E^+)$ to the vicinity of $\text{graph}(E^-)$.

Let $\widehat{\mu}$ be an \widehat{f} -invariant probability measure with ergodic decomposition $\widehat{\mu} = \int_{\Omega} \widehat{\mu}_{\xi} dm$. The projection μ of $\widehat{\mu}$ to M has ergodic decomposition $\mu = \int_{\Omega} \mu_{\xi} dm$, where $\mu_{\xi} = \widehat{\mu}_{\xi} \circ \pi^{-1}$. Note that in general, the map $\widehat{\mu}_{\xi} \mapsto \mu_{\xi}$ is not injective.

We split the set of ergodic components $\widehat{\mu}_{\xi}$ ($\xi \in \Omega$) by considering whether they are carried by the invariant line bundles E^+ or E^- or by a subset of \widehat{M} where these bundles are not defined:

$$\begin{aligned} \Omega^- &:= \{\xi \in \Omega : \lambda^-(f, \mu_{\xi}) < \lambda^+(f, \mu_{\xi}), \widehat{\mu}_{\xi} = \widehat{\mu}_{\xi}^-\} \\ \Omega^+ &:= \{\xi \in \Omega : \lambda^-(f, \mu_{\xi}) < \lambda^+(f, \mu_{\xi}), \widehat{\mu}_{\xi} = \widehat{\mu}_{\xi}^+\} \\ \Omega^0 &:= \{\xi \in \Omega : \lambda^-(f, \mu_{\xi}) = \lambda^+(f, \mu_{\xi})\}. \end{aligned}$$

Theorem 3.6 (Defect in continuity). *Let M be a compact smooth boundaryless surface. For each $k \geq 1$, fix $f_k \in \text{Diff}^1(M)$ and ergodic f_k -invariant measures ν_k with $\lambda^-(f_k, \nu_k) < \lambda^+(f_k, \nu_k)$. Let $\widehat{\nu}_k^+$ be the ergodic lift to \widehat{M} carried by the bundle E^+ . Suppose $f_k \xrightarrow{c^1} f$, $\nu_k \xrightarrow{w^*} \mu$, and $\widehat{\nu}_k^+ \xrightarrow{w^*} \widehat{\mu}$. Considering the ergodic decompositions $\widehat{\mu} = \int_{\Omega} \widehat{\mu}_{\xi} dm$, $\mu = \int_{\Omega} \mu_{\xi} dm$ and defining $\Omega = \Omega^+ \cup \Omega^- \cup \Omega^0$ as above, we have*

$$\lim_{k \rightarrow \infty} \lambda^+(f, \nu_k) = \lambda^+(f, \mu) - \int_{\Omega^-} [\lambda^+(f, \mu_{\xi}) - \lambda^-(f, \mu_{\xi})] dm.$$

In the special case when μ is ergodic, we have the following (see Lemma 3.3):

- (1) If $\lambda^+(f, \mu) = \lambda^-(f, \mu)$ or $\widehat{\mu} = \widehat{\mu}^+$, then $\lim_k \lambda^+(f_k, \nu_k) = \lambda^+(f, \mu)$.
- (2) If $\lambda^+(f, \mu) < \lambda^-(f, \mu)$ and $\widehat{\mu} \neq \widehat{\mu}^+$, then $\lim_k \lambda^+(f_k, \nu_k) < \lambda^+(f, \mu)$.
More precisely, there is a unique $0 < a \leq 1$ such that $\widehat{\mu} = a\widehat{\mu}^- + (1-a)\widehat{\mu}^+$ and

$$\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \lambda^+(f, \mu) - a(\lambda^+(f, \mu) - \lambda^-(f, \mu)).$$

If ν_k and μ are hyperbolic (of saddle type), then $a \neq 1$.

Proof. Using Lemma 3.3, we see that

$$\int \varphi d\widehat{\mu}_{\xi} = \begin{cases} \int \varphi d\widehat{\mu}_{\xi}^+ = \lambda^+(f, \mu_{\xi}) & \text{if } \xi \in \Omega^+ \\ \int \varphi d\widehat{\mu}_{\xi}^- = \lambda^-(f, \mu_{\xi}) & \text{if } \xi \in \Omega^- \\ \lambda^+(f, \mu_{\xi}) & \text{if } \xi \in \Omega^0, \end{cases} \quad (10)$$

$$\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \lim_{k \rightarrow \infty} \int \varphi d\widehat{\nu}_k^+ = \int \varphi d\widehat{\mu} = \iint \varphi d\widehat{\mu}_{\xi} dm. \quad (11)$$

Substituting (10) in (11), we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) &= \int_{\Omega^-} \lambda^-(f, \mu_\xi) dm + \int_{\Omega^+ \cup \Omega^0} \lambda^+(f, \mu_\xi) dm \\ &= \int_{\Omega^-} \lambda^-(f, \mu_\xi) dm + \int_{\Omega} \lambda^+(f, \mu_\xi) dm - \int_{\Omega^-} \lambda^+(f, \mu_\xi) dm \\ &= \lambda^+(f, \mu) - \int_{\Omega^-} [\lambda^+(f, \mu_\xi) - \lambda^-(f, \mu_\xi)] dm, \end{aligned}$$

which proves the theorem for general, possibly non-ergodic, limits μ .

When μ is ergodic, we apply equation (11) and Lemma 3.3:

- (1) If $\lambda^+(f, \mu) = \lambda^-(f, \mu)$ or $\hat{\mu} = \hat{\mu}^+$, then $\int \varphi d\hat{\mu} = \lambda^+(f, \mu)$ and item (1) follows.
- (2) Otherwise, μ has two different exponents and $m(\Omega^+) < 1$. Necessarily, $m(\Omega^0) = 0$, $a := m(\Omega^-) \in (0, 1]$ and the lift $\hat{\mu}$ can be written as $(1-a)\hat{\mu}^+ + a\hat{\mu}^-$. So

$$\begin{aligned} \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) &= \int \varphi d\hat{\mu} = (1-a)\lambda^+(f, \mu) + a\lambda^-(f, \mu) \\ &= \lambda^+(f, \mu) - a(\lambda^+(f, \mu) - \lambda^-(f, \mu)) \\ &< \lambda^+(f, \mu). \end{aligned}$$

Finally, note that, if $a = 1$ and μ is hyperbolic, then $\lim_k \lambda^+(f_k, \nu_k) = \lambda^-(f, \mu) < 0$, and ν_k are not hyperbolic of saddle type for k large enough. \square

Notice that the defect in continuity originates at Ω^- , the set of ergodic components of $\lim \hat{\nu}_k^+$ which are carried by $\text{graph}(E^-)$. This confirms the heuristic that discontinuity in Lyapunov exponents is due to the asymptotic escape of mass from $\text{graph}(E^+)$, which carries $\hat{\nu}_k^+$, to $\text{graph}(E^-)$, which carries $\hat{\mu}_\xi$ for $\xi \in \Omega^-$.

3.5. A bound for the asymptotic dilation of \hat{f} .

Lemma 3.7. *For any C^2 diffeomorphism f of a surface M , $\lambda(\hat{f}) \leq \lambda(f) + \lambda(f^{-1})$.*

Proof. Working locally in charts, we identify the iterates f^n locally with diffeomorphisms $F_i: U_i \rightarrow \mathbb{R}^2$ defined on open subsets of \mathbb{R}^2 . We choose the charts so that the change of coordinates distorts the metric by a factor of less than 2. The lift $\hat{f}^n \in \text{Diff}^1(\hat{M})$ is identified with:

$$\hat{F}_i(x, v) = \left(F_i(x), \frac{(DF_i)_x v}{\|(DF_i)_x v\|} \right).$$

In what follows, we omit the first factor.

The differential of \hat{F}_i can be computed in a straightforward way (writing $a * b$ for the scalar product of two vectors in \mathbb{R}^2):

$$\begin{aligned} (D\hat{F}_i)_{(x,v)}(y, w) &= \frac{(DF_i)_x w}{\|(DF_i)_x v\|} + \frac{(D^2 F_i)_x(y) \cdot v}{\|(DF_i)_x v\|} \\ &\quad - \left(\frac{(DF_i)_x v}{\|(DF_i)_x v\|} * \frac{(DF_i)_x w + (D^2 F_i)_x(y) \cdot v}{\|(DF_i)_x v\|} \right) \frac{(DF_i)_x(v)}{\|(DF_i)_x v\|}. \end{aligned} \tag{12}$$

Thus,

$$\|(D\hat{F}_i)_{(x,v)}\| \leq 2(\|(DF_i)_x\| \|(DF_i^{-1})_{f(x)}\| + \|(D^2 F_i)_x\| \|(DF_i^{-1})_{f(x)}\|).$$

We apply this to some iterate of f on M , remembering the distortion in the metric:

$$\|D\hat{f}^n\|_{\sup} \leq 32\|Df^n\|_{\sup}\|Df^{-n}\|_{\sup} + 16\|Df^{-n}\|_{\sup}\|D^2 f^n\|_{\sup}$$

Let $0 < \varepsilon < 1/4$. Fix n an integer so large that $32 \leq e^{\varepsilon n/4}$ and

$$\|Df^n\|_{\sup} \leq e^{n(\lambda(f)+\varepsilon/8)} \text{ and } \|Df^{-n}\|_{\sup} \leq e^{n(\lambda(f^{-1})+\varepsilon/8)}.$$

Therefore,

$$\lambda(\widehat{f}) \leq \lambda(f) + \lambda(f^{-1}) + \varepsilon/2 + \log \left(1 + e^{-\lambda(f)} \|D^2 f^n\|_{\sup}^{1/n} \right)$$

By dilating the metric on M , we can ensure that $\|D^2 f^n\|_{\sup}^{1/n} \leq (\varepsilon/4)e^{\lambda(f)}$ without changing the asymptotic dilations. Thus $\lambda(\widehat{f}) \leq \lambda(f) + \lambda(f^{-1}) + \varepsilon$. As ε is arbitrarily small, the claim follows. \square

These computations allow the following control of the C^{r-1} size of the lift of a C^r diffeomorphism.

Lemma 3.8. *For every real $2 < r < \infty$, there is a constant $A = A(r)$ with the following property. For any $g \in \text{Diff}^r(M)$ with lift $\widehat{g} \in \text{Diff}^{r-1}(\widehat{M})$,*

$$\|\widehat{g}\|_{C^{r-1}} \leq A (\|g\|_{C^r} \cdot \|Dg^{-1}\|_{\sup})^A.$$

Proof. We first consider the k th derivative of f in charts for the maximal integer $k \leq r$. A straightforward induction on the integer $k \geq 2$ based on eq. (12) shows that the $(k-1)$ th differential of \widehat{F} at some point $(x, v) \in \widehat{M}$ can be written as a linear combination of terms:

$$\frac{1}{\|D_x F.v\|^p} (D^{\alpha_1} F * D^{\beta_1} F) \dots (D^{\alpha_j} F * D^{\beta_j} F) \cdot D^\gamma F$$

where $p, \alpha_1, \beta_1, \dots, \alpha_j, \beta_j, \gamma$ are integers, and $\alpha_1, \beta_1, \dots, \gamma \leq k$. The coefficients of this linear combination depend only on k .

If $k = r$, the claim is immediate. If $\alpha := r - k > 0$, recall that the C^r size is the sum of the C^k size and α -Hölder size of the k -th derivative. A further computation using the above expression gives the required bound for the Hölder constant of order α of $D^k \widehat{g}$. \square

4. ENTROPY FORMULAS AND REPARAMETRIZATIONS

We saw in last section that the defect in continuity of $(f, \mu) \mapsto \lambda^+(f, \mu)$ can be described in terms of the canonical lift $\widehat{f} : \widehat{M} \rightarrow \widehat{M}$. In this section we develop tools for studying the entropy map $(f, \mu) \mapsto h(f, \mu)$ in terms of $\widehat{f} : \widehat{M} \rightarrow \widehat{M}$.

Specifically, we will show that the entropy of hyperbolic measures on M can be studied in terms of the exponential rate of growth in C^r -complexity of $\widehat{f}^n \circ \widehat{\sigma}$, where $\widehat{\sigma} : [0, 1] \rightarrow \widehat{M}$ is the curve $\widehat{\sigma}(t) = (\sigma(t), \mathbb{R}.\sigma'(t))$ and $\sigma : [0, 1] \rightarrow M$ is a smooth parameterization of a local unstable manifold.

4.1. Review of entropy and the ergodic decomposition. This section collects several classical facts on the entropy theory of non-ergodic measures. For proofs and details, see [15, chap. 13].

Consider a compact metric space X together with a continuous map T preserving an invariant Borel probability measure μ and the σ -algebra \mathcal{X} of Borel subsets of X . The ergodic decomposition of μ with respect to T is:

$$\mu = \int_X \mu_x d\mu(x)$$

where $\mu_x := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{0 \leq k < n} \delta_{T^k x}$ (the weak-* limit exists almost everywhere by the ergodic theorem).

The map $x \mapsto \mu_x$ is μ -measurable with respect to the σ -algebra \mathcal{I} of invariant measurable subsets; for μ -a.e. $x \in X$, μ_x is a T -invariant and ergodic Borel probability measure and for every Borel μ -absolutely integrable u , $x \mapsto \mu_x(u)$ belongs to $L^1(\mu)$ and $\mu(u) = \int_X \mu_x(u) d\mu(x)$. The metric entropy of μ and μ_x are related by

$$h(T, \mu) = \int_X h(T, \mu_x) d\mu(x).$$

Suppose ξ is a countable measurable partition of X with finite mean entropy $H_\mu(\xi) := -\sum_{A \in \xi} \mu(A) \log \mu(A)$. Let $\xi_n := \{\bigcap_{i=0}^{n-1} T^{-i} A_i : A_i \in \xi\}$. It is a classical fact that $h(T, \mu, \xi) := \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\xi_n)$ exists.

Similarly, one defines $h(T, \mu_x, \xi)$. The function $x \mapsto h(T, \mu_x, \xi)$ is defined μ -a.e., is measurable with respect to the μ -completion of \mathcal{I} . Let $\xi_n(x)$ denote the atom of ξ_n which contains x . The Shannon-McMillan-Breiman theorem, in its version for non-ergodic measures [15, (13.4)], states that,

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mu(\xi_n(x)) = h(T, \mu_x, \xi) \quad \text{both } \mu\text{-a.e. and in } L^1(\mu). \quad (13)$$

In addition, we have the following identity: [15, (13.3)]: $h(T, \mu, \xi) = \int_X h(T, \mu_x, \xi) d\mu(x)$. In particular, $h(T, \mu, \xi) \leq \bar{h}(T, \mu, \xi)$, where

$$\bar{h}(T, \mu) := \text{ess-sup}_{x \in X} h(T, \mu_x). \quad (14)$$

We call $\bar{h}(T, \mu)$ the *essential supremum entropy* of μ :

4.2. Bowen and Katok entropy formulas. Let $T : X \rightarrow X$ be a continuous map on a compact metrix space X . An (n, ϵ) -Bowen ball is a set of the form

$$B_T(x, n, \epsilon) := \{y \in X : \forall 0 \leq k < n, d(T^k y, T^k x) < \epsilon\}.$$

The (n, ϵ) -covering number of a subset $Z \subset X$, is

$$r_T(n, \epsilon, Z) := \min\{|C| : \bigcup_{x \in C} B_T(x, n, \epsilon) \supset Z\}.$$

Bowen [8] defined the topological entropy of a (possibly non-invariant) set $Z \subset X$ for T to be

$$h_{\text{top}}(T, Z) = \lim_{\epsilon \rightarrow 0} h_{\text{top}}(T, Z, \epsilon) \quad \text{with} \quad h_{\text{top}}(T, Z, \epsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_T(n, \epsilon, Z), \quad (15)$$

and showed that the topological entropy of T is $h_{\text{top}}(T) = h_{\text{top}}(T, X)$.

Katok gave a similar formula for the metric entropy of an invariant measure. Let μ be an invariant probability measure. For every $\gamma \in (0, 1)$, let

$$r_T(n, \epsilon, \mu, \gamma) := \inf\{r_T(n, \epsilon, Z) : Z \subset X \text{ measurable s.t. } \mu(Z) > \gamma\}.$$

He showed that if μ is ergodic, then $h(T, \mu) = \lim_{\lambda \rightarrow 1} \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_T(n, \epsilon, \mu, \gamma)$. Katok's proof in [26] also works in the non-ergodic case, if we replace the usual Shannon-McMillan-Breiman Theorem by (13). The result is that for a general (possibly non-ergodic) invariant probability measure μ ,

$$\bar{h}(T, \mu) = \lim_{\lambda \rightarrow 1} \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_T(n, \epsilon, \mu, \gamma). \quad (16)$$

Here $\bar{h}(T, \mu)$ is the essential supremum entropy from (14).

4.3. The lift to the projective tangent bundle preserves entropy. This is a standard consequence of the following theorem [30].

Theorem 4.1 (Ledrappier-Walters). *Let (X, T) , (Y, S) be continuous self-maps of compact metric spaces and let $\pi : (Y, S) \rightarrow (X, T)$ be a topological factor map s.t.*

$$\forall x \in X \quad h_{\text{top}}(S, \pi^{-1}(x)) = 0. \quad (17)$$

Then, for any S -invariant Borel probability measure ν on Y , $h(S, \nu) = h(T, \pi_\nu)$.*

Thus by the variational principle, in the setup of the theorem, $h_{\text{top}}(T) = h_{\text{top}}(S)$.

Corollary 4.2. *Suppose $f \in \text{Diff}^1(M)$ and \widehat{f} is the canonical extension of f to the projective bundle \widehat{M} . Then $h_{\text{top}}(\widehat{f}) = h_{\text{top}}(f)$, and for every \widehat{f} -invariant probability measure $\widehat{\nu}$ with projection ν , $h(\widehat{f}, \widehat{\nu}) = h(f, \nu)$ and $\bar{h}(\widehat{f}, \widehat{\nu}) = \bar{h}(f, \nu)$.*

Proof. We check condition (17) for $T = f$, $S = \widehat{f}$, and apply the previous theorem. \widehat{M} is a topological bundle over M , and its fibers $P_x M$ are homeomorphic to circles. The map $\widehat{f} : P_x M \rightarrow P_{f(x)} M$ is a homeomorphism. For every $\varepsilon > 0$, one can find partitions ξ_x of $P_x M$ into a bounded number of arcs with diameter at most ε . It is easy to see that $\xi_x \vee \widehat{f}^{-1} \xi_{\widehat{f}x} \vee \dots \vee \widehat{f}^{-n+1} \xi_{\widehat{f}^{n-1}x}$ has cardinality at most $\sum_{k=0}^{n-1} \text{Card}(\xi_{\widehat{f}^k x}) = O(n)$. It follows that $h_{\text{top}}(\widehat{f}, P_x M) = 0$ for all x . \square

4.4. Bowen and Katok entropy formulas on the bundle \widehat{M} . We need a variant of the Bowen and Katok entropy formulas which uses a different type of Bowen balls, which are better adapted to the bundle structure of \widehat{M} .

Recall the natural projection $\widehat{\pi} : \widehat{M} \rightarrow M$, $\widehat{\pi}(x, E) = x$. The *fibred ball* with center $\widehat{x} \in \widehat{M}$ and scales $\varepsilon, \widehat{\varepsilon} > 0$ is the set

$$B(\widehat{x}, \varepsilon, \widehat{\varepsilon}) := \{\widehat{y} \in \widehat{M} : d(\widehat{x}, \widehat{y}) < \widehat{\varepsilon} \text{ and } d(\widehat{\pi}(\widehat{x}), \widehat{\pi}(\widehat{y})) < \varepsilon\}.$$

A set $S \subset \widehat{M}$ has *fibred size* $\leq (\varepsilon, \widehat{\varepsilon})$ if $S \subset B(\widehat{x}, \varepsilon, \widehat{\varepsilon})$ for some \widehat{x} .

Suppose $f \in \text{Diff}^1(M)$ and \widehat{f} is the canonical extension of f to \widehat{M} . The *fibred $(n, \varepsilon, \widehat{\varepsilon})$ -Bowen ball* with center $\widehat{x} \in \widehat{M}$, size $(\varepsilon, \widehat{\varepsilon})$ and length n is the set

$$B_{\widehat{f}}(\widehat{x}, n, \varepsilon, \widehat{\varepsilon}) := \{\widehat{y} \in \widehat{M} : \forall 0 \leq k < n, d(\widehat{f}^k(\widehat{x}), \widehat{f}^k(\widehat{y})) < \widehat{\varepsilon} \text{ and } d(f^k(\widehat{\pi}(\widehat{x})), f^k(\widehat{\pi}(\widehat{y}))) < \varepsilon\}. \quad (18)$$

The $(n, \varepsilon, \widehat{\varepsilon})$ -*covering number* of a subset $Z \subset \widehat{M}$ is the minimal number of fibred $(n, \varepsilon, \widehat{\varepsilon})$ -Bowen balls whose union contains Z . It is denoted by

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, Z).$$

Clearly if $\varepsilon_1 \leq \varepsilon_2$ and $\widehat{\varepsilon}_1 \leq \widehat{\varepsilon}_2$, then $r_{\widehat{f}}(n, \varepsilon_1, \widehat{\varepsilon}_1, Z) \geq r_{\widehat{f}}(n, \varepsilon_2, \widehat{\varepsilon}_2, Z)$.

Similarly, given an ergodic measure $\widehat{\mu}$ of \widehat{f} and a number $0 < \gamma < 1$, the $(n, \varepsilon, \widehat{\varepsilon}, \gamma)$ -*covering number* of $\widehat{\mu}$ is the minimal number of fibred $(n, \varepsilon, \widehat{\varepsilon})$ -Bowen balls whose union has $\widehat{\mu}$ -measure at least γ . It is denoted by

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\mu}, \gamma).$$

If $\varepsilon_1 \leq \varepsilon_2$, $\widehat{\varepsilon}_1 \leq \widehat{\varepsilon}_2$, and $\gamma_1 \geq \gamma_2$, then $r_{\widehat{f}}(n, \varepsilon_1, \widehat{\varepsilon}_1, \widehat{\mu}, \gamma_1) \geq r_{\widehat{f}}(n, \varepsilon_2, \widehat{\varepsilon}_2, \widehat{\mu}, \gamma_2)$.

Proposition 4.3. *Fix $f \in \text{Diff}^1(M)$ with canonical lift \widehat{f} . We have:*

(1) Bowen's formula: For every $\widehat{\varepsilon} > 0$,

$$h_{\text{top}}(f) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{M}) = \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{M})$$

(2) Katok's formula: Suppose μ is an f -ergodic invariant measure, and let $\widehat{\mu}$ be an \widehat{f} -ergodic lift of μ . Then for every $\widehat{\varepsilon} > 0$ and $0 < \gamma < 1$,

$$h(f, \mu) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\mu}, \gamma),$$

If μ is f -invariant, but possibly not ergodic, then for every lift $\widehat{\mu}$,

$$\overline{h}(f, \mu) = \lim_{\gamma \rightarrow 1} \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\mu}, \gamma),$$

Proof. We follow the proof of [8, Thm 17]. We need the following claim.

Claim 4.4. For every $\widehat{\varepsilon}, \alpha > 0$, there are $C, \varepsilon_* > 0$ such that

$$\forall \varepsilon \in (0, \varepsilon_*), \forall n \geq 1, \forall x \in M, \quad r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\pi}^{-1}B_f(x, n, \varepsilon)) \leq Ce^{\alpha n}.$$

Proof of the Claim. Let $\widehat{\varepsilon}, \alpha > 0$. Note first that

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\pi}^{-1}B_f(x, n, \varepsilon)) \leq r_{\widehat{f}}(n, \widehat{\varepsilon}, \widehat{\pi}^{-1}B_f(x, n, \varepsilon)),$$

hence it is enough to bound the latter. Since $h_{\text{top}}(\widehat{f}, \widehat{\pi}^{-1}(x)) = 0$, for each $x \in M$ there is a smallest integer $n_x \geq 1$ such that $r_{\widehat{f}}(n_x, \widehat{\varepsilon}/2, \widehat{\pi}^{-1}(x)) \leq e^{\alpha n_x}$.

Recall that a set-valued function F from a topological space X to the set of subsets of a topological set Y is called *upper semi-continuous*, if for every $E \subset Y$ closed, $\{x \in X : F(x) \cap E \neq \emptyset\}$ is closed. The continuity of $\widehat{\pi}$ and the compactness of \widehat{M} implies that $x \mapsto \widehat{\pi}^{-1}(x)$ is upper semi-continuous.

It follows that if $\widehat{\pi}^{-1}(x)$ is contained in some open set U (say the union of a minimal cover by fibered Bowen balls), then $\widehat{\pi}^{-1}(y)$ is contained in U for all y sufficiently close to x . Hence there is an $r_x > 0$ such that

$$r_{\widehat{f}}(n_x, \widehat{\varepsilon}/2, \widehat{\pi}^{-1}B(x, r_x)) \leq r_{\widehat{f}}(n_x, \widehat{\varepsilon}/2, \widehat{\pi}^{-1}(x)) \leq e^{\alpha n_x}. \quad (19)$$

Using a compactness argument, we see that $n_* := \sup\{n_x : x \in M\}$ is finite and that one can arrange for $\varepsilon_* := \inf\{r_x : x \in M\}$ to be positive.

Let $\varepsilon \in (0, \varepsilon_*)$, $n \geq 1$ and $x \in M$. Define $t_0(x) := 0$ and $t_{i+1}(x) := t_i(x) + n_{f^{t_i}(x)}$. Choose $i \geq 0$ maximal such that $t_i(x) \leq n$. Note that $0 \leq n - t_i(x) < n_*$. Thus, setting $x_j := f^{t_j}(x)$ and $n_j := n_{x_j}$,

$$B_f(x, n, \varepsilon) = \bigcap_{j=0}^{i-1} f^{-t_j(x)} B_f(x_j, n_j, \varepsilon) \cap f^{-t_i(x)} B(x_i, n - t_i(x), \varepsilon).$$

Let $\{z_1, \dots, z_C\}$ be a $(n_* - 1, \widehat{\varepsilon}/2)$ -cover of \widehat{M} with cardinality C . For each $0 \leq j < i$, let $\{y_{j1}, \dots, y_{jm}\}$ with $m \leq e^{\alpha n_j}$ be the $(n_j, \widehat{\varepsilon}/2)$ -cover of $\widehat{\pi}^{-1}B(x_j, n_j, \varepsilon)$ implied by eq. (19). Then $\widehat{\pi}^{-1}B_f(x, n, \varepsilon)$ is contained in the following union:

$$\bigcup_{\substack{k, k_0, \dots, k_{i-1} \\ k_j \leq \exp \alpha n_j \\ k \leq C}} \bigcap_{j=0}^{i-1} \widehat{f}^{-t_j(x)} B_{\widehat{f}}(y_{jk_j}, n_j, \widehat{\varepsilon}/2) \cap \widehat{f}^{-t_i(x)} B_{\widehat{f}}(z_k, n - t_i(x), \widehat{\varepsilon}/2).$$

Any two points contained in the same term of this union are not $(n, \widehat{\varepsilon})$ -separated. So the cardinality of any $(n, \widehat{\varepsilon})$ -separated subset of $\widehat{\pi}^{-1}B_f(x, n, \varepsilon)$ is bounded by

$$C \times \prod_{j=0}^{i-1} e^{\alpha n_j} \leq C e^{\alpha n}.$$

Since a maximal $(n, \widehat{\varepsilon})$ -separated subset is the set of centers of a cover by $(n, \widehat{\varepsilon})$ -Bowen balls, $r_{\widehat{f}}(n, \widehat{\varepsilon}, \widehat{\pi}^{-1}B_f(x, n, \varepsilon)) \leq C e^{\alpha n}$, which proves the claim.

The claim implies that for any $\widehat{\varepsilon}, \alpha > 0$, there are $C, \varepsilon_* > 0$ such that, for any $0 < \varepsilon \leq \varepsilon_*$ and any $Z \subset M$,

$$r_f(n, \varepsilon, Z) \leq r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\pi}^{-1}Z) \leq C e^{\alpha n} r_f(n, \varepsilon, Z).$$

The proposition now follows from the classical identities (15) and (16). \square

4.5. Curves and C^r reparametrizations. The entropy of a diffeomorphism can be related to the exponential rate of growth in C^r complexity of the iterates of a local unstable manifold. To do this we need to control the curvature, and for this purpose it is useful to lift the curve to projective bundle \widehat{M} and study its iterations there. Here we develop the tools needed for doing this.

Definition 4.5. A C^r curve $\sigma : [0, 1] \rightarrow M$ is regular if its derivative $\sigma'(t)$ never vanishes. In this case, it has a canonical lift $\widehat{\sigma} : [0, 1] \rightarrow \widehat{M}$ defined by

$$\widehat{\sigma}(t) := (\sigma(t), \mathbb{R} \cdot \sigma'(t)).$$

Here and throughout, $\mathbb{R} \cdot \sigma'(t) \equiv \text{span}\{\sigma'(t)\}$ is a linear subspace of $T_x M$. By regularity, $\mathbb{R} \cdot \sigma'(t)$ is one-dimensional.

Definition 4.6. Fix $r \geq 2$ and let $\varepsilon, \widehat{\varepsilon}$ be two positive numbers. A regular C^r curve $\sigma : [0, 1] \rightarrow M$ has C^r size (or just size) less than $(\varepsilon, \widehat{\varepsilon})$ if

$$\|\sigma\|_{C^r} < \varepsilon \text{ and } \|\widehat{\sigma}\|_{C^{r-1}} < \widehat{\varepsilon}.$$

The curve has diameter less than $(\varepsilon, \widehat{\varepsilon})$ if

$$\text{diam}_M(\sigma([0, 1])) < \varepsilon \text{ and } \text{diam}_{\widehat{M}}(\widehat{\sigma}([0, 1])) < \widehat{\varepsilon}.$$

To say that a curve has finite C^r size implies that it is regular and C^r .

Remark 4.7. If a curve σ is parametrized by length (i.e., $\|\sigma'(t)\| = 1$ for all t), then it has size at most $(\|\sigma\|_{C^r}, C\|\sigma'\|_{C^{r-1}})$ for some constant $C > 0$ which depends on the choice of C^{r-1} atlas of \widehat{M} used to define C^{r-1} size (see section 3.1).

Suppose $\sigma : [0, 1] \rightarrow M$ is a curve. If we cut $[0, 1]$ into small intervals $[a_i, b_i]$ and reparametrize $\sigma|_{[a_i, b_i]}$ by $\sigma \circ \psi_i$ where $\psi_i : [0, 1] \rightarrow [a_i, b_i]$ is an affine bijection, then $\|\sigma \circ \psi_i\|_{C^r} \leq \kappa \|\sigma\|_{C^r}$, $\|\widehat{\sigma} \circ \psi_i\|_{C^r} \leq \kappa \|\widehat{\sigma}\|_{C^r}$, where $\kappa = |a_i - b_i| < 1$. Cutting sufficiently finely, we can obtain covering by pieces with affine reparametrizations with C^r size as small as we wish.

Yomdin measured the C^r complexity of a curve (more generally a set) by counting how many reparametrized pieces with C^r size less than 1 are needed to cover it. We adapt this to the projective dynamics:

Definition 4.8. Let $\sigma : [0, 1] \rightarrow M$ be a C^r curve. A reparametrization of σ is a non-constant affine map $\psi : [0, 1] \rightarrow [0, 1]$.

A family of reparametrizations of σ over a subset $T \subset [0, 1]$ is a collection \mathcal{R} of reparametrizations such that $T \subset \bigcup_{\psi \in \mathcal{R}} \psi([0, 1])$.

Let $f \in \text{Diff}^r(M)$ and let σ be a regular C^r curve. We will be interested in families of reparametrizations which remain bounded in C^r size after application of f^n for certain n . Specifically, fix numbers $\varepsilon, \widehat{\varepsilon} > 0$, an integer $N \geq 1$, and $T \subset [0, 1]$.

Definition 4.9. A reparametrization ψ of σ is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n , if there exists an increasing sequence $(n_0, n_1, \dots, n_\ell)$ such that

- $n_0 = 0$, $n_\ell = n$, and $n_j - n_{j-1} \leq N$ for each $1 \leq j \leq \ell$,
- for each $0 \leq j \leq \ell$ the curve $f^{n_j} \circ \sigma \circ \psi$ has C^r size less than $(\varepsilon, \widehat{\varepsilon})$.

We call the integers n_j the admissible times.

A family \mathcal{R} of reparametrizations of σ over T is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n , if each $\psi \in \mathcal{R}$ is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n .

Lemma 4.10 (Concatenation of reparametrizations). *Let $f \in \text{Diff}^r(M)$, $r \geq 2$, and consider a regular C^r curve σ on M , and $T \subset [0, 1]$. Suppose that*

- \mathcal{R} is a family of reparametrizations of σ over T ;
- \mathcal{R} is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n ;
- for each $\psi \in \mathcal{R}$, there is a family \mathcal{R}_ψ of reparametrizations of $f^n \circ \sigma \circ \psi$ over $\psi^{-1}(T)$, which is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n' .

Then $\mathcal{R}' := \bigcup_{\psi \in \mathcal{R}} \{\psi \circ \phi : \phi \in \mathcal{R}_\psi\}$ is a family of reparameterizations of σ over T , and \mathcal{R}' is $(C^r, f, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time $n + n'$.

Proof. Since $\psi, \phi : [0, 1] \rightarrow [0, 1]$ are non-constant and affine, $\psi \circ \phi : [0, 1] \rightarrow [0, 1]$ is non-constant and affine. Next,

$$\bigcup_{\psi \in \mathcal{R}} \bigcup_{\phi \in \mathcal{R}_\psi} (\psi \circ \phi)[0, 1] = \bigcup_{\psi \in \mathcal{R}} \psi \left(\bigcup_{\phi \in \mathcal{R}_\psi} \phi[0, 1] \right) \supseteq \bigcup_{\psi \in \mathcal{R}} \psi(\psi^{-1}(T)) = \bigcup_{\psi \in \mathcal{R}} T \cap \psi[0, 1] = T,$$

So \mathcal{R}' is a family of reparametrizations of $(\sigma, \widehat{\sigma})$ over T .

To see that \mathcal{R}' is admissible, fix $\psi \in \mathcal{R}, \phi \in \mathcal{R}_\psi$, and choose admissible times $0 = n_0 < n_1 < \dots < n_\ell = n$ and $0 = n'_0 < n'_1 < \dots < n'_m = n'$ for ψ and ϕ . Let

$$n_{\ell+k} := n + n'_k \quad (k = 0, \dots, m).$$

We claim that $0 = n_0 < n_1 < \dots < n_{\ell+m} = n + n'$ are admissible times for $\psi \circ \phi$. That the gaps are no larger than N is clear. If $j = \ell + 1, \dots, \ell + m$, then $f^{n_j} \circ \psi \circ \phi$ is a regular curve with size $< (\varepsilon, \widehat{\varepsilon})$, because $\phi \in \mathcal{R}_\psi$ and \mathcal{R}_ψ is admissible. If $j \leq \ell$, then using the fact that $\phi' = c$ with c a constant s.t. $|c| \leq 1$, we find that $\|f^{n_j} \circ \sigma \circ \psi \circ \phi\|_{C^r} \leq \|f^{n_j} \circ \sigma \circ \psi\|_{C^r} \cdot |c| \leq \|f^{n_j} \circ \sigma \circ \psi\|_{C^r}$. This is less than ε , because of the admissibility of \mathcal{R} . Similarly $\|\widehat{f}^{n_j} \circ \widehat{\sigma} \circ \psi \circ \phi\|_{C^{r-1}} \leq \|\widehat{f}^{n_j} \circ \widehat{\sigma} \circ \psi\|_{C^{r-1}} < \widehat{\varepsilon}$. \square

Lemma 4.11 (Length of reparametrizations). *Let $f \in \text{Diff}^r(M)$, $r \geq 2$. For any $\eta > 0$ and $N \geq 1$, there exist $\varepsilon_* > 0$ and a C^2 -neighborhood \mathcal{U}_* of f in $\text{Diff}^r(M)$ with the following property.*

Consider $g \in \mathcal{U}_$, $\varepsilon, \widehat{\varepsilon} \in (0, \varepsilon_*)$, a regular C^r curve σ , and a reparametrization φ of σ which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n . Then for any $(x, E) \in \widehat{\sigma}[0, 1]$,*

$$\text{Length}(g^n \circ \sigma \circ \varphi) < e^{\eta n/10} \|Dg_x^n|_E\| \text{Length}(\sigma \circ \varphi).$$

Proof. \widehat{M} is compact, so one can find $\varepsilon_* > 0$ and a small C^2 -neighborhood \mathcal{U}_* of f in $\text{Diff}^r(M)$ such that for all $g \in \mathcal{U}_*$ and $\widehat{x}, \widehat{y} \in \widehat{M}$ satisfying $d(\widehat{x}, \widehat{y}) < \varepsilon_*$,

$$\log \|Dg|_{\widehat{g}^k(\widehat{x})}\| \leq \log \|Dg|_{\widehat{g}^k(\widehat{y})}\| + \frac{\eta}{10} \quad (0 \leq k \leq N).$$

Let us consider $g \in \mathcal{U}_*$, a curve σ and a reparametrization φ as in the statement, with admissibility times n_0, \dots, n_ℓ . One gets for any $0 \leq i < \ell$,

$$\forall (x, E), (y, F) \in \widehat{g}^{n_i} \circ \widehat{\sigma}[0, 1], \quad \|Dg_y^{n_{i+1}-n_i}|_F\| < e^{(n_{i+1}-n_i)\eta/10} \|Dg_x^{n_{i+1}-n_i}|_E\|,$$

which immediately implies the conclusion of the lemma. \square

Admissible families of reparametrizations yield covers by fibered Bowen balls with size of the same order of magnitude.

Lemma 4.12 (Bowen covers from admissible reparametrizations). *Let $f \in \text{Diff}^r(M)$, $r \geq 2$, $T \subset [0, 1]$, and σ be a regular C^r curve. Fix $\varepsilon_*, \widehat{\varepsilon}_* > 0$ and $N \geq 1$. Let \mathcal{R} be a family of reparametrizations of σ over T which is $(C^r, f, N, \varepsilon_*, \widehat{\varepsilon}_*)$ -admissible up to time n . Then, for every $\varepsilon, \widehat{\varepsilon} > 0$,*

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\sigma}(T)) \leq \frac{2\widehat{\varepsilon}_* \|D\widehat{f}\|_{\sup}^N}{\min(\varepsilon, \widehat{\varepsilon})} |\mathcal{R}|.$$

Proof. Let $L := \|D\widehat{f}\|_{\sup}$. $L \geq 1$ since \widehat{f} is surjective.

Let $\rho := \frac{\min(\varepsilon, \widehat{\varepsilon})}{\widehat{\varepsilon}_* L^N}$, and construct a ρ -dense set $C_\rho \subset [0, 1]$ s.t. $|C_\rho| \leq \frac{2}{\rho}$. Set $C_\rho(\mathcal{R}) := \{(\widehat{\sigma} \circ \psi)(t') : t' \in C_\rho, \psi \in \mathcal{R}\}$.

Fix $\psi \in \mathcal{R}$ with admissible times $0 = n_0 < n_1 < \dots < n_\ell = n$. Then $n_{j+1} - n_j \leq N$ and the regular curve $f^{n_j} \circ \sigma \circ \psi$ has C^r size at most $(\varepsilon_*, \widehat{\varepsilon}_*)$, whence

$$\|D\widehat{f}^k \circ \sigma \circ \psi\|_{\sup} \leq \widehat{\varepsilon}_* L^{k-n_j} \quad \text{for } k \in [n_j, n_{j+1}).$$

In particular, $\|D\widehat{f}^k \circ \sigma \circ \psi\|_{\sup} \leq \widehat{\varepsilon}_* L^N$ for all $0 \leq k \leq n$.

For every $\widehat{x} \in \widehat{\sigma}(T \cap \psi[0, 1])$ with $\widehat{x} = \widehat{\sigma} \circ \psi(t)$, there is $\widehat{y} = \widehat{\sigma} \circ \psi(t')$ with $t' \in C_\rho$ such that $|t - t'| < \rho$, whence

$$d(\widehat{f}^k(\widehat{x}), \widehat{f}^k(\widehat{y})) < \rho \widehat{\varepsilon}_* L^N \quad \text{and} \quad d(f^k(\widehat{\pi}(\widehat{x})), f^k(\widehat{\pi}(\widehat{y}))) < \rho \widehat{\varepsilon}_* L^N \quad \text{for all } k = 0, \dots, n.$$

(the second inequality follows from the first). Notice that $\widehat{y} \in C_\rho(\mathcal{R})$.

All this shows that $\widehat{\sigma}(T) \subset \bigcup_{\widehat{y} \in C_\rho(\mathcal{R})} B_{\widehat{f}}(\widehat{y}, n, \rho \widehat{\varepsilon}_* L^N, \rho \widehat{\varepsilon}_* L^N) \subset \bigcup_{\widehat{y} \in C_\rho(\mathcal{R})} B_{\widehat{f}}(\widehat{y}, n, \varepsilon, \widehat{\varepsilon})$. So

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\sigma}(T)) \leq |C_\rho(\mathcal{R})| \leq |\mathcal{R}| \cdot |C_\rho| \leq \frac{2}{\rho} |\mathcal{R}| \leq \frac{2\widehat{\varepsilon}_*}{\min(\varepsilon, \widehat{\varepsilon})} \|D\widehat{f}\|_{\sup}^N |\mathcal{R}|. \quad \square$$

4.6. Yomdin estimates. In this section we discuss a converse to Lemma 4.12: Covers by Bowen balls generate admissible reparametrizations with cardinality of the same order of magnitude. This result is much more delicate than Lemma 4.12, and requires Yomdin's Theorem [48]. Here is the tool we need from Yomdin's work, in a form adapted to our setup. Let

$$Q_r(g) := \max(\|g\|_{C^r}, \|\widehat{g}\|_{C^{r-1}}), \quad Q_{r,N}(g) := \max_{n=1,2,\dots,N} Q_r(g^n). \quad (20)$$

Theorem 4.13 (Yomdin). *Given real numbers $2 \leq r < \infty$ and $Q > 0$, there are $\Upsilon = \Upsilon(r) > 0$ and $\varepsilon_Y = \varepsilon_Y(r, Q) > 0$ with the following properties. For every*

- C^r diffeomorphism $g : M \rightarrow M$ such that $Q_r(g) \leq Q$
- regular C^r curve σ with C^r size at most $(\varepsilon, \widehat{\varepsilon})$ with $0 < \varepsilon, \widehat{\varepsilon} \leq \varepsilon_Y$,
- $\widehat{x} \in \widehat{\sigma}[0, 1]$, and $T := \{t \in [0, 1] : \widehat{g}(\widehat{\sigma}(t)) \in B(\widehat{g}(\widehat{x}), \varepsilon, \widehat{\varepsilon})\}$.

there exists a family \mathcal{R} of reparametrizations of σ over T such that

- (1) *for every $\psi \in \mathcal{R}$, $g \circ \sigma \circ \psi$ is a regular C^r curve with C^r size at most $(\varepsilon, \widehat{\varepsilon})$,*
- (2) $|\mathcal{R}| \leq \Upsilon \|D\widehat{g}\|_{\sup}^{1/(r-1)}$.

Proof. The reader who will compare this theorem to Yomdin's original statement in [48] will find that the two results are nearly the same, except for the following differences: (1) Yomdin considered the more general case of $\sigma : [0, 1]^\ell \rightarrow M$ whereas we restrict to $\ell = 1$; (2) Yomdin did not specify that all reparametrizations are affine as we do; (3) Our result allows r to be real, not just an integer; and (4) We use $(n, \varepsilon, \widehat{\varepsilon})$ balls in \widehat{M} , whereas Yomdin used (n, ε) balls in M .

In the special case of curves $\ell = 1$, Yomdin's proof works verbatim with affine reparametrizations, see, e.g., [48, p. 297–298]. The extension to non-integer smoothness is also simple and well-known, see, e.g., [13, p.133].

To deal with (4), we apply the Yomdin's original theorem twice, first for (g, σ) on M and then for the lift $(\widehat{g}, \widehat{\sigma})$ on \widehat{M} . This yields numbers $\Upsilon_i = \Upsilon_i(r)$ ($i = 1, 2$) and $\varepsilon_Y = \varepsilon_Y(r, Q) > 0$ as follows. Suppose $0 < \varepsilon, \widehat{\varepsilon} \leq \varepsilon_Y$, $g \in \text{Diff}^r(M)$, $Q_r(g) \leq Q$, and $\sigma : [0, 1] \rightarrow M$ has C^r size at most $(\varepsilon, \widehat{\varepsilon})$. Then

- (1) There is a family \mathcal{R}_1 of reparametrizations of $g \circ \sigma$ over T s.t. $|\mathcal{R}_1| \leq \Upsilon_1 \|Dg\|_{\sup}^{1/r}$, and so that each $\psi \in \mathcal{R}_1$ is affine, contracting, and $\|g \circ \sigma \circ \psi\|_{C^r} \leq \varepsilon$.
- (2) There is a family \mathcal{R}_2 of reparametrizations of $\widehat{g} \circ \widehat{\sigma}$ over T s.t. $|\mathcal{R}_2| \leq \Upsilon_2 \|D\widehat{g}\|_{\sup}^{\frac{1}{r-1}}$ and so that each $\psi \in \mathcal{R}_2$ is affine, contracting, and $\|\widehat{g} \circ \widehat{\sigma} \circ \psi\|_{C^r} \leq \widehat{\varepsilon}$.

Each family of reparametrizations \mathcal{R}_i generates a cover of T by the intervals $\psi([0, 1])$, $\psi \in \mathcal{R}_i$. Without loss of generality, the interiors of these intervals are pairwise disjoint (otherwise discard some of them and shrink the rest by composing the reparametrizations by affine contractions). We define

$$\mathcal{R} := \{\phi_{\psi_1, \psi_2} \mid (\psi_1, \psi_2) \in \mathcal{R}_1 \times \mathcal{R}_2 \text{ s.t. } \psi_1((0, 1)) \cap \psi_2((0, 1)) \neq \emptyset\}$$

where $\phi_{\psi_1, \psi_2} : [0, 1] \rightarrow \psi_1([0, 1]) \cap \psi_2([0, 1])$ is an affine diffeomorphism. The image of this new family of reparametrizations contains the intersection of the images of \mathcal{R}_1 and \mathcal{R}_2 , hence \mathcal{R} is a family of reparametrizations of σ over T .

As the reparametrizations are affine and contracting, we see that

$$\|g \circ \sigma \circ \phi_{\psi_1, \psi_2}\|_{C^r} \leq \|g \circ \sigma \circ \psi_1\|_{C^r} < \varepsilon$$

Likewise for $\widehat{g} \circ \widehat{\sigma} \circ \phi_{\psi_1, \psi_2}$. Since $\widehat{g} \circ \widehat{\sigma} \circ \phi_{\psi_1, \psi_2}$ coincides with the lift of $g \circ \sigma \circ \phi_{\psi_1, \psi_2}$, item (1) of the theorem holds.

Next, using the order structure on the interval, it is not difficult to show that $|\mathcal{R}| \leq |\mathcal{R}_1| + |\mathcal{R}_2| - 1 \leq \Upsilon_1 \|Dg\|_{\sup}^{1/r} + \Upsilon_2 \|D\widehat{g}\|_{\sup}^{\frac{1}{r-1}}$. Obviously $\|Dg\|_{\sup}^{1/r} \leq \|D\widehat{g}\|_{\sup}^{1/(r-1)}$, so item (2) holds with $\Upsilon := \Upsilon_1 + \Upsilon_2$. \square

Corollary 4.14 (Existence of admissible reparametrizations). *For $2 \leq r < \infty$, $Q > 0$, let $\Upsilon(r)$, $\varepsilon_Y(r, Q)$ be the constants from Theorem 4.13. Suppose*

- $g \in \text{Diff}^r(M)$ and $Q_{r,N}(g) < Q$,
- $\sigma : [0, 1] \rightarrow M$ is a regular curve with C^r size $\leq (\varepsilon, \widehat{\varepsilon})$, where $0 < \varepsilon, \widehat{\varepsilon} < \varepsilon_Y$,
- $N, n \geq 1$, and $T \subset [0, 1]$.

Then there exists a family \mathcal{R} of reparametrizations of σ over T , which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to the time n , and with cardinality

$$|\mathcal{R}| \leq C(r, \widehat{g}) r_{\widehat{g}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{g} \circ \widehat{\sigma}(T)),$$

where $C(r, \widehat{g}) = \Upsilon(r)^{\lceil \frac{n}{N} \rceil} \|D\widehat{g}\|_{\sup}^{\frac{\lfloor \frac{n}{N} \rfloor}{r-1}} \|D\widehat{g}^{n-N} \lfloor \frac{n}{N} \rfloor\|_{\sup}^{\frac{1}{r-1}}$.

Proof. Fix $n \geq 1$ and divide with remainder $n = qN + p$, $q \geq 0$, $p = 0, \dots, N - 1$.

Let $\ell := r_{\widehat{g}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{g} \circ \widehat{\sigma}(T))$, then there exists a cover of $\widehat{g} \circ \widehat{\sigma}(T)$ by ℓ fibered $(n, \varepsilon, \widehat{\varepsilon})$ -Bowen balls $B_i := B(\widehat{g}(\widehat{x}_i), n, \varepsilon, \widehat{\varepsilon})$. For each B_i , we will construct an admissible family of reparametrizations of σ over $T_i := (\widehat{g} \circ \widehat{\sigma})^{-1}(B_i)$, and then take the union over i .

Step 0. If $p = 0$ move to step 1. Otherwise proceed as follows.

Fix $1 \leq i \leq \ell$. Yomdin's theorem for g^p , σ and \widehat{x}_i gives a family of reparametrizations \mathcal{R}_0 of σ over

$$T_i^0 := \{t \in [0, 1] : \widehat{g}^p(\widehat{\sigma}(t)) \in B(\widehat{g}^p(\widehat{x}_i), \varepsilon, \widehat{\varepsilon})\},$$

which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time p (the admissible times are $0, p$), and such that $|\mathcal{R}_0| \leq \Upsilon \|D\widehat{g}^p\|_{\sup}^{1/(r-1)}$. Notice that $T_i^0 \supset T_i$. If $q = 0$, we have a reparametrization up to time n , and we stop.

Step 1. Fix $\psi \in \mathcal{R}_0$ and apply Yomdin's theorem to g^N , $g^p \circ \sigma \circ \psi$, and $\widehat{g}^p(\widehat{x}_i)$. The result is a family \mathcal{R}_ψ of reparametrizations of $g^p \circ \sigma \circ \psi$ over

$$T_i^1(\psi) := \{t \in [0, 1] : \widehat{g}^{N+p}[(\widehat{\sigma} \circ \psi)(t)] \in B(\widehat{g}^{N+p}(\widehat{x}_i), \varepsilon, \widehat{\varepsilon})\}$$

which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time N (the admissible times are $0, N$), and with cardinality $|\mathcal{R}_\psi| \leq \Upsilon \|D\widehat{g}^N\|_{\sup}^{1/(r-1)}$.

Notice that $T_i^1(\psi) \supset \psi^{-1}(T_i)$, therefore the concatenation

$$\mathcal{R}_1 := \{\psi \circ \phi : \psi \in \mathcal{R}_0, \phi \in \mathcal{R}_\psi\}$$

is an admissible family of reparametrizations of σ over T_i up to time $N + p$, with admissible times $0, p, N + p$ and cardinality

$$|\mathcal{R}_1| \leq \Upsilon^2 \|D\widehat{g}^N\|_{\sup}^{1/(r-1)} \|D\widehat{g}^p\|_{\sup}^{1/(r-1)}.$$

If $q = 1$, we have a reparametrization up to time n , and we stop.

Otherwise we continue as before to a “step 2” which applies Yomdin's theorem to g^N , $g^{N+p} \circ \sigma \circ \psi$ and $\widehat{g}^{N+p}(\widehat{x}_i)$.

Eventually, at step q , we arrive to a family of reparametrizations \mathcal{R}_q over T_i which is admissible up to time $qN + p = n$, and which has cardinality

$$|\mathcal{R}_q| \leq \Upsilon^{q+1} \|D\widehat{g}^N\|_{\sup}^{q/(r-1)} \|D\widehat{g}^p\|_{\sup}^{1/(r-1)} \equiv \Upsilon^{\lceil n/N \rceil} \|D\widehat{g}^N\|_{\sup}^{\lfloor \frac{n/N}{r-1} \rfloor} \|D\widehat{g}^{n-N\lfloor n/N \rfloor}\|_{\sup}^{1/(r-1)}.$$

Taking the union over $i = 1, \dots, \ell$, we obtain the family of reparametrizations over $\bigcup T_i \supset T$, as required. \square

4.7. Entropy and growth of C^r complexity of unstable manifolds. In this section f is a C^r diffeomorphism, $r > 1$, of a surface M and ν is an ergodic hyperbolic probability measure of saddle type, ie $\lambda^s := \lambda^-(f, \nu)$ is strictly negative, and $\lambda^u := \lambda^+(f, \nu)$ is strictly positive. Pesin's Unstable Manifold Theorem says that ν -a.e. x belongs to an unstable manifold $W^u(x)$, which is an injectively immersed C^r curve and is characterized as:

$$W^u(x) = \{y : \limsup_{n \rightarrow +\infty} \frac{1}{n} \log d(f^{-n}(x), f^{-n}(y)) < 0\}.$$

A measurable partition ξ is *subordinated* to the unstable lamination W^u of ν if for ν -almost every $x \in M$, the atom $\xi(x)$ is a neighborhood of x inside the curve $W^u(x)$ and ξ is increasing: Every atom of $f(\xi)$ is a union of atoms of ξ . By [29], such measurable partitions exist.

Since ξ is measurable, Rokhlin's disintegration theorem applies, and for ν -a.e. x there exists a probability measure ν_x^u on $\xi(x)$ so that

$$\nu = \int \nu_x^u d\nu(x).$$

The family $\{\nu_x^u\}$ is not unique, but given ξ , any two families like that are equal outside a set of x of measure zero. Therefore it is not a serious abuse of terminology to call ν_x^u the conditional measure on $\xi(x)$.

In this section we use the entropy theory of Ledrappier and Young [27] and especially the following corollary established by Zang (see [50, Remark 1.8]) to show that the entropy of ν can be bounded by the exponential rate of growth of the C^r complexity of the curve $f^n(W_{loc}^u(x))$, as quantified in the previous section using admissible C^r reparametrizations up to time n .

Theorem 4.15 (Y. Zang). *Let us consider $f \in \text{Diff}^r(M)$ with $r > 1$, an ergodic hyperbolic probability measure ν , and a system of conditional measures $\{\nu_x^u\}$ on local unstable manifolds. Then for ν -a.e. $x \in M$, the measure ν_x^u satisfies:*

$$h(f, \nu) = \inf_{\gamma > 0} \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log r_f(n, \varepsilon, \nu_x^u, \gamma).$$

The difference between this result and Ledrappier-Young theory is that Zang assumes C^r smoothness for some $r > 1$ and hyperbolicity, whereas Ledrappier and Young assume C^2 smoothness, but no hyperbolicity.

Corollary 4.16. *Let $f \in \text{Diff}^2(M)$, and let ν be an ergodic hyperbolic measure with a system of conditional measures $\{\nu_x^u\}$ on local unstable manifolds. For any $F \subset M$ with positive ν -measure, for ν -a.e. $x_0 \in F$, and for any choice of*

- $\sigma: [0, 1] \rightarrow W^u(x_0)$, a regular C^r curve,
- $T \subset [0, 1]$, a set such that $\nu_{x_0}^u(\sigma(T) \cap F) > 0$,

the following holds. If \mathcal{R}_n ($n \geq 1$) are families of reparametrizations of σ over T which are $(C^r, f, N, \varepsilon_, \widehat{\varepsilon}_*)$ -admissible up to time n for some (any) $\varepsilon_*, \widehat{\varepsilon}_* > 0, N \geq 1$ independent of n , then $h(f, \nu) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{R}_n|$.*

Proof. Fix $\varepsilon_*, \widehat{\varepsilon}_*, N > 0$ and let \mathcal{R}_n be families of admissible reparametrizations as in the statement. Let $\widehat{\sigma}: [0, 1] \rightarrow \widehat{M}$ be the canonical lift of the regular curve σ . Fix $\varepsilon, \widehat{\varepsilon} > 0$ arbitrarily small. By Lemma 4.12,

$$r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\sigma}(T)) \leq C|\mathcal{R}_n|,$$

where $C := C(f, \varepsilon, \widehat{\varepsilon}, \widehat{\varepsilon}_*, N) := \frac{2\widehat{\varepsilon}_* \|D\widehat{f}\|_{\sup}^N}{\min(\varepsilon, \widehat{\varepsilon})}$ is independent of n .

By the definition of $r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\sigma}(T))$, there exist $\widehat{x}_1, \dots, \widehat{x}_\ell \in \widehat{M}$ with $\ell \leq C|\mathcal{R}_n|$ such that $\bigcup_{i=1}^\ell B_{\widehat{f}}(\widehat{x}_i, n, \varepsilon, \widehat{\varepsilon}) \supset \widehat{\sigma}(T)$. Necessarily

$$\bigcup_{i=1}^\ell B_f(\pi(\widehat{x}_i), n, \varepsilon) \supset \sigma(T).$$

In particular, $r_f(n, \varepsilon, \sigma(T)) \leq r_{\widehat{f}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{\sigma}(T)) \leq C|\mathcal{R}_n|$, whence by Zang's Theorem, $h(f, \nu) \leq \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log r_f(n, \varepsilon, \sigma(T) \cap F) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{R}_n|$. \square

5. MAIN REPARAMETRIZATION LEMMAS

This section collects our main technical results on the existence of admissible families of reparametrizations of pieces of unstable manifolds.

The point is to produce families with cardinality as small as possible. The first result provides admissible families of reparametrizations of local unstable manifolds, with cardinality controlled in terms of the entropy. The second result, which is much more subtle, produces much smaller families of reparametrizations for the subset of the local unstable stable where there is little expansion up to some iterate, see Definition 6.1.

5.1. Statements. Throughout this section M is a compact smooth surface without boundary, $f : M \rightarrow M$ is a diffeomorphism, $\widehat{f} : \widehat{M} \rightarrow \widehat{M}$ is the canonical lift (7), and $\widehat{\mu}$ is a (possibly non-ergodic!) \widehat{f} -invariant probability measure, which projects to an f -invariant measure μ . $Q_{r,N}(f)$ is given by (20) in the previous section, \bar{h} is the essential entropy (14), and $\lambda(\widehat{f})$ is the asymptotic dilation of \widehat{f} , see (2) and §3.5.

The statements of the following two propositions should be formally understood as stating the existence of functions N_1, \bar{n}_1, γ_0 , and N_0 with values in $(0, \infty)$ such that the following stated properties hold.

Proposition 5.1. *Let us consider $f \in \text{Diff}^r(M)$ with $2 \leq r < \infty$, an \widehat{f} -invariant probability $\widehat{\mu}$, some real numbers $Q, \eta, \gamma, \varepsilon, \widehat{\varepsilon} > 0$, and integers N, n . Assume that:*

- $N \geq N_1(r, f, \eta)$,
- $0 < \varepsilon, \widehat{\varepsilon} \leq \varepsilon_Y(r, Q)$ (the constant in Yomdin's theorem 4.13), and
- $n \geq \bar{n}_1(f, \widehat{\mu}, \eta, \gamma, N, \varepsilon, \widehat{\varepsilon})$.

Then there are:

- a C^2 neighborhood $\mathcal{U}_1(f, \eta, \gamma, \varepsilon, \widehat{\varepsilon}, N, n)$ of f in $\text{Diff}^r(M)$,
- an open set $\widehat{U}_1(f, \widehat{\mu}, \eta, \gamma, \varepsilon, \widehat{\varepsilon}, n) \subset \widehat{M}$ with $\widehat{\mu}(\widehat{U}_1) > 1 - \gamma^2$ and $\widehat{\mu}(\partial \widehat{U}_1) = 0$,

such that the following property holds:

- (*) *For any $g \in \mathcal{U}_1$ with $Q_{r,N}(g) < Q$ and for any regular curve σ with C^r size at most $(\varepsilon, \widehat{\varepsilon})$, there is a family \mathcal{R} of reparametrizations over $\widehat{\sigma}^{-1}(\widehat{U}_1)$ s.t.*
- \mathcal{R} is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n ,
 - $|\mathcal{R}| \leq \exp \left[n \left(\bar{h}(f, \mu) + \frac{\lambda(\widehat{f})}{r-1} + \eta \right) \right]$,

The following and key estimate applies to the part of the local unstable manifold which does not (initially) see much expansion. More precisely, suppose g is a diffeomorphism with canonical lift \widehat{g} , and let $\alpha > 0$. An *orbit segment* with length n is a string $\vartheta := (\widehat{x}, \widehat{g}(\widehat{x}), \dots, \widehat{g}^{n-1}(\widehat{x}))$.

Definition 5.2. *An orbit segment $\vartheta = (\widehat{x}, \widehat{g}(\widehat{x}), \dots, \widehat{g}^{n-1}(\widehat{x}))$ is α -neutral, if, denoting $\widehat{x} = (x, E)$, we have $\|Dg_x^m|_E\| \leq e^{\alpha m}$ for every $1 \leq m \leq n$.*

Proposition 5.3. *Let us consider $f \in \text{Diff}^r(M)$ with $2 \leq r < \infty$, an \widehat{f} -invariant probability $\widehat{\mu}$, some real numbers $Q, \eta, \gamma, \delta > 0$, and an integer N . Assume that:*

- $0 < \gamma \leq \gamma_0(r, f, \eta)$,
- $N \geq N_0(r, f, \eta, \gamma)$.

Then there are:

- $0 < \varepsilon, \widehat{\varepsilon} \leq \delta$.
- a C^2 neighborhood $\mathcal{U}_0(f, \widehat{\mu}, Q, \eta, \gamma, \delta, N)$ of f in $\text{Diff}^r(M)$,
- an open set $\widehat{U}_0(f, \widehat{\mu}, Q, \eta, \gamma, \delta, N) \subset \widehat{M}$ with $\widehat{\mu}(\widehat{U}_0) > 1 - \gamma^2$ and $\widehat{\mu}(\partial \widehat{U}_0) = 0$,

- an integer $\bar{n}_0 := \bar{n}_0(f, \hat{\mu}, Q, \eta, \gamma, N, \delta) \geq 1$,

such that the following property holds:

- (**) For any $g \in \mathcal{U}_0$ with $Q_{r,N}(g) < Q$, any regular curve σ with C^r size at most $(\varepsilon, \hat{\varepsilon})$, and any $n \geq \bar{n}_0$ there is a family \mathcal{R}_n of reparametrizations over

$$T := \hat{\sigma}^{-1} \left\{ \hat{x}: (\hat{x}, \dots, \hat{g}^{n-1}(\hat{x})) \text{ is } \frac{\eta}{10}\text{-neutral and } \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\hat{g}^j(\hat{x})}(\hat{U}_0) > 1 - \gamma \right\}, \quad (21)$$

such that

- \mathcal{R}_n is $(C^r, g, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time n ,
- $|\mathcal{R}_n| \leq \exp \left[n \left(\frac{\lambda(\hat{f})}{r-1} + \eta \right) \right]$.

Unlike Proposition 5.1, here the upper bound has no entropic term. Indeed, in the C^∞ case, the exponential rate of growth tends to zero with η . This low complexity is due to the neutrality of the piece of σ we are parametrizing.

The proofs of these two propositions may be skipped at the first reading.

5.2. Proof of Proposition 5.1. Let $r \in [2, \infty)$, $f \in \text{Diff}^r(M)$, $Q, \eta, \gamma > 0$, and consider a \hat{f} -invariant probability measure $\hat{\mu}$. Note that the lift $\hat{f} \in \text{Diff}^{r-1}(\hat{M})$ is uniquely defined by f and depends continuously on $f \in \text{Diff}^r(M)$. Recall the number $\Upsilon := \Upsilon(r)$ given by Yomdin's Theorem 4.13.

- Fix an integer $N_1 = N_1(r, f, \eta) \geq 1$ such that for all $N \geq N_1$,

$$\Upsilon < \exp\left(\frac{\eta N}{10}\right) \quad \text{and} \quad \frac{1}{N} \log \|D\hat{f}^N\|_{\text{sup}} < \lambda(\hat{f}) + \frac{\eta}{10}. \quad (22)$$

This is possible since $\lambda(\hat{f}) = \lim_{N \rightarrow +\infty} \frac{1}{N} \log \|D\hat{f}^N\|_{\text{sup}}$.

- Fix $N \geq N_1$.
- Set $\varepsilon_Y := \varepsilon_Y(r, Q)$ as in Yomdin's Theorem 4.13.
- Let $\varepsilon, \hat{\varepsilon}$ be arbitrary in $(0, \varepsilon_Y)$.
- Pick an integer $\bar{n}_1 = \bar{n}_1(f, \hat{\mu}, \eta, \gamma, N, \varepsilon, \hat{\varepsilon}) > N$ using Proposition 4.3(2) such that for any $n \geq \bar{n}_1$,

$$\frac{1}{n} \log r_{\hat{f}}(n, \frac{\varepsilon}{2}, \frac{\hat{\varepsilon}}{2}, \hat{\mu}, 1 - \gamma^2) < \bar{h}(\hat{f}, \hat{\mu}) + \frac{\eta}{4}, \quad (23)$$

$$\log \|D\hat{f}\|_{\text{sup}}^N < \frac{n\eta}{10}. \quad (24)$$

- Let n be some integer larger than \bar{n}_1 .
- Let $\mathcal{U}_1 = \mathcal{U}_1(f, \eta, N, n, \varepsilon, \hat{\varepsilon}) \subset \text{Diff}^r(M)$ be a small enough C^2 neighborhood of f in $\text{Diff}^r(M)$ such that the lift \hat{g} of any $g \in \mathcal{U}_1$ satisfies:

$$\|D\hat{g}^N\|_{\text{sup}} \leq e^{\eta N/10} \|D\hat{f}^N\|_{\text{sup}}, \quad \|D\hat{g}\|_{\text{sup}} \leq e^{\eta/10} \|D\hat{f}\|_{\text{sup}}, \quad (25)$$

and so that every fibered $(n, \frac{\varepsilon}{2}, \frac{\hat{\varepsilon}}{2})$ -Bowen ball for \hat{f} is contained in a fibered $(n, \varepsilon, \hat{\varepsilon})$ -Bowen ball for \hat{g} .

By the regularity of Borel measures, there is a compact set $\hat{K} = \hat{K}(f, \hat{\mu}, \eta, \gamma, n, \varepsilon, \hat{\varepsilon})$ with $\hat{\mu}(\hat{K}) > 1 - \gamma^2$. By eq. (23) there is a neighborhood \hat{U}_1 of \hat{K} , such that $\hat{g}(\hat{U}_1)$ is contained in the union of a collection \mathcal{C} of fibered $(n, \frac{\varepsilon}{2}, \frac{\hat{\varepsilon}}{2})$ -Bowen balls for \hat{f} with cardinality at most $\exp(n(\bar{h}(f, \hat{\mu}) + \frac{\eta}{4}))$. Passing to a smaller open set containing \hat{K} , we can ensure that $\hat{\mu}(\partial \hat{U}_1) = 0$.

Suppose $g \in \mathcal{U}_1$, $Q_{r,N}(g) < Q$, and let σ be a regular C^r curve with C^r size at most $(\varepsilon, \widehat{\varepsilon})$. By construction, we can cover $\widehat{g}(\widehat{\sigma}[0, 1] \cap \widehat{U}_1)$ using only the $(n, \frac{\varepsilon}{2}, \frac{\widehat{\varepsilon}}{2})$ -Bowen balls for \widehat{f} from the collection \mathcal{C} . Every ball in \mathcal{C} is contained in some $(n, \varepsilon, \widehat{\varepsilon})$ -Bowen ball for \widehat{g} . Thus,

$$r_{\widehat{g}}(n, \varepsilon, \widehat{\varepsilon}, \widehat{g}(\widehat{\sigma}[0, 1] \cap \widehat{U}_1)) \leq \exp(n(\bar{h}(f, \widehat{\mu}) + \frac{\eta}{4})).$$

Since $\varepsilon, \widehat{\varepsilon} < \varepsilon_Y$, we can apply Yomdin theory in the form of Corollary 4.14. The result is a family \mathcal{R} of reparametrizations of σ over $\widehat{\sigma}^{-1}\widehat{U}_1$ which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n , and such that

$$|\mathcal{R}| \leq \Upsilon^{\lceil n/N \rceil} \|D\widehat{g}^N\|_{\sup}^{\frac{\lceil n/N \rceil}{r-1}} \|D\widehat{g}\|_{\sup}^{\frac{N}{r-1}} \exp(n(\bar{h}(f, \widehat{\mu}) + \frac{\eta}{4})).$$

Using (22), (24) and (25), one gets $|\mathcal{R}| \leq \exp(n(\bar{h}(f, \widehat{\mu}) + \frac{\lambda(\widehat{f})}{r-1} + \eta))$. \square

5.3. Proof of Proposition 5.3. The proof splits into two parts: In steps 0-6 we select the parameters $\gamma_0, N_0, \varepsilon, \widehat{\varepsilon}, \mathcal{U}_0, \widehat{U}_0, \bar{n}_0$; In steps 7-11 we build admissible families of reparametrizations as in the statement, and estimate their cardinality.

Fix $r \in [2, \infty)$, $f \in \text{Diff}^r(M)$, $Q, \eta, \delta > 0$, and consider an \widehat{f} -invariant probability measure $\widehat{\mu}$. Let $\Upsilon := \Upsilon(r)$ given by Yomdin's Theorem 4.13. Let

$$H(t) := t \ln \frac{1}{t} + (1-t) \ln \frac{1}{1-t}. \quad (26)$$

Step 0 (Preliminary choices). – Choose $\gamma_0 = \gamma_0(r, f, \eta) > 0$ such that:

$$\forall 0 < \gamma \leq \gamma_0, \quad 3\Upsilon \leq \exp(\frac{\eta}{10\gamma}), \quad H(4\gamma) < \frac{\eta}{10}, \quad \|D\widehat{f}\|_{\sup} < \exp(\frac{\eta}{10\gamma}). \quad (27)$$

– Let $0 < \gamma < \gamma_0$ be arbitrary.

– Fix an integer $N_0 = N_0(r, f, \gamma, \eta) > \max(\frac{1}{\gamma}, \frac{10}{\eta})$ such that (cf. (2), §3.5):

$$\forall N \geq N_0, \quad \frac{1}{N} \log \|D\widehat{f}^N\|_{\sup} < \lambda(\widehat{f}) + \frac{\eta}{10}. \quad (28)$$

– Let $N \geq N_0$ be arbitrary.

– Fix $\varepsilon_Y = \varepsilon_Y(r, Q)$ as in Yomdin's Theorem 4.13.

– Let $\mathcal{V} = \mathcal{V}(f, \eta, N)$ be a C^2 neighborhood of f in $\text{Diff}^r(M)$ such that for $g \in \mathcal{V}$,

$$\|D\widehat{g}^k\|_{\sup} \leq e^{\eta/10} \|D\widehat{f}^k\|_{\sup} \quad (k = 1, \dots, N). \quad (29)$$

Step 1 (Decomposition of $\widehat{\mu}$). $\widehat{\mu}$ projects to an f -invariant probability measure. Let $\lambda_x^\pm := \lambda^\pm(f, x)$ denote the Lyapunov exponents. These are well-defined μ -a.e., but since we are assuming nothing on $\widehat{\mu}$ they could be equal on a set of positive measure. Consider the invariant measurable subset

$$M_\# := \{x \in M \mid \lambda_x^- \text{ and } \lambda_x^+ \text{ are defined and distinct}\}.$$

We decompose $\widehat{\mu}$ as a barycenter of two invariant probability measures:

$$\widehat{\mu} = a\widehat{\mu}_\# + (1-a)\widehat{\mu}_0,$$

where $a := \mu(M_\#)$, $a\widehat{\mu}_\# := \widehat{\mu}(\cdot \cap \widehat{\pi}^{-1}(M_\#))$ and $(1-a)\widehat{\mu}_0 := \widehat{\mu}(\cdot \cap \widehat{\pi}^{-1}(M \setminus M_\#))$. In the following we assume $a \in (0, 1)$. Indeed, in the special case where $a = 0$ (resp. $a = 1$), we simply write $\widehat{\mu} = \widehat{\mu}_0$ (resp. $\widehat{\mu} = \widehat{\mu}_\#$) and the proof of Proposition 5.3 becomes simpler and can be easily obtained by adapting the general case.

Let μ_0 and $\mu_\#$ be the projections of $\widehat{\mu}_0, \widehat{\mu}_\#$ to M . These are f -invariant measures; μ_0 -a.e. x has two equal Lyapunov exponents; and $\mu_\#$ -a.e. x has two different Lyapunov exponents.

Step 2 (Compact subsets $\widehat{K}_+, \widehat{K}_-$ approximating $\widehat{\mu}_\#$). As in section 3.3, the Oseledets splitting induces two \widehat{f} -invariant measurable sections $x \mapsto (x, E^\pm(x))$, $M \rightarrow \widehat{M}$, defined $\mu_\#$ -a.e.

– By Lusin's theorem there exists a compact set $K_\# = K_\#(f, \widehat{\mu}, \gamma, N)$ inside $M_\#$ such that $\mu_\#(K_\#) > 1 - \gamma^2$, and so that the functions

$$x \mapsto (x, E^+(f^k(x))), \quad x \mapsto (x, E^-(f^k(x))) \quad (0 \leq k \leq N)$$

are continuous on $K_\#$. By Lemma 3.3, $\widehat{\mu}_\#$ is carried by $\text{graph}(E^+) \cup \text{graph}(E^-)$. Therefore, the sets

$$\widehat{K}^+ := \text{graph}(E^+|_{K_\#}) \text{ and } \widehat{K}^- := \text{graph}(E^-|_{K_\#})$$

are compact and disjoint, and $\widehat{\mu}_\#(\widehat{K}^+ \cup \widehat{K}^-) = \mu_\#(K_\#) > 1 - \gamma^2$. We set

$$\widehat{K}_\# := \widehat{\pi}^{-1}(K_\#) = \widehat{K}_- \sqcup \widehat{K}_+.$$

Every $z \in K_\#$ has two lifts: $\widehat{z}^+ \in \widehat{K}^+$, and $\widehat{z}^- \in \widehat{K}^-$. It is easy to see that $\widehat{f}^k(\widehat{z}^\pm) = (f^k(z), E^\pm(f^k(z)))$.

Step 3 (Control of N iterates starting near $\widehat{K}_\#$).

– There exist $\varepsilon_* = \varepsilon_*(r, f, \eta, N) > 0$ and a C^2 neighborhood $\mathcal{U}_* = \mathcal{U}_*(f, \eta, N)$ of f in $\text{Diff}^r(M)$ such that the conclusion of Lemma 4.11 holds.

– Let $\widehat{\varepsilon} = \widehat{\varepsilon}(r, f, Q, \eta, \delta, N, K_\#) := \frac{1}{10} \min(\delta, \varepsilon_Y, \varepsilon_*, \text{dist}(\widehat{K}_+, \widehat{K}_-))$. Note that the latter distance is positive since \widehat{K}_+ and \widehat{K}_- are disjoint compact sets.

– By construction of $K_\#$, there is $\varepsilon = \varepsilon(f, \widehat{\varepsilon}, K_\#, N) \in (0, \widehat{\varepsilon})$ such that if $x, y \in K_\#$ satisfy $d(x, y) \leq \varepsilon$, then

$$\forall 0 \leq k \leq N, \quad d(\widehat{f}^k(\widehat{x}^+), \widehat{f}^k(\widehat{y}^+)) < \frac{\widehat{\varepsilon}}{2} \text{ and } d(\widehat{f}^k(\widehat{x}^-), \widehat{f}^k(\widehat{y}^-)) < \frac{\widehat{\varepsilon}}{2}.$$

By the choice of $\widehat{\varepsilon}$, if $\widehat{x}, \widehat{y} \in \widehat{K}_\#$ satisfy $d(\widehat{x}, \widehat{y}) \leq \widehat{\varepsilon}$, then either $\widehat{x} = \widehat{x}^+, \widehat{y} = \widehat{y}^+$ or $\widehat{x} = \widehat{x}^-, \widehat{y} = \widehat{y}^-$. Consequently, for any $\widehat{x}, \widehat{y} \in \widehat{M}$,

$$\left. \begin{array}{l} \widehat{x}, \widehat{y} \in \widehat{K}_\# \\ d(\widehat{x}, \widehat{y}) \leq \widehat{\varepsilon}, \quad d(x, y) \leq \varepsilon \end{array} \right\} \implies \forall 0 \leq k \leq N, \quad d(\widehat{f}^k(\widehat{x}), \widehat{f}^k(\widehat{y})) < \frac{\widehat{\varepsilon}}{2}.$$

Below, $B(S, \beta)$ denotes the β -neighborhood of a subset S .

– Since \widehat{M} is compact, there exist $\delta_{**} = \delta_{**}(f, N, \varepsilon, \widehat{\varepsilon}) > 0$ and a C^2 neighborhood $\mathcal{U}_{**} = \mathcal{U}_{**}(f, N, \varepsilon, \widehat{\varepsilon})$ of f in $\text{Diff}^r(M)$ such that

$$\left. \begin{array}{l} g \in \mathcal{U}_{**} \\ \widehat{x}, \widehat{y} \in B(\widehat{K}_\#, \delta_{**}) \\ d(x, y) \leq \varepsilon, \quad d(\widehat{x}, \widehat{y}) \leq \widehat{\varepsilon} \end{array} \right\} \implies \forall 0 \leq k \leq N, \quad d(\widehat{g}^k(\widehat{x}), \widehat{g}^k(\widehat{y})) < \widehat{\varepsilon}. \quad (30)$$

– Choose an open set $\widehat{W}_\# = \widehat{W}_\#(\widehat{K}_\#, \delta_{**})$ such that $\widehat{K}_\# \subset \widehat{W}_\# \subset B(\widehat{K}_\#, \delta_{**})$ and $\widehat{\mu}(\partial \widehat{W}_\#) = 0$. By (30), if σ is a regular curve with C^r size at most $(\varepsilon, \widehat{\varepsilon})$, then for all $g \in \mathcal{U}_{**}$,

$$\forall 0 \leq k \leq N \quad \text{diam}(\widehat{g}^k(\sigma \cap \widehat{W}_\#)) < \widehat{\varepsilon}. \quad (31)$$

Step 4 (Control of n_* iterates starting near $\hat{\mu}_0$). Recall that μ_0 has equal Lyapunov exponents almost everywhere. By Ruelle's inequality μ_0 must have zero entropy. It follows that $\bar{h}(f, \mu_0) = 0$.

– By Proposition 4.3, $\lim_{\bar{\gamma} \rightarrow 1} \lim_{\bar{\varepsilon} \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_{\hat{f}}(n, \bar{\varepsilon}, \frac{\bar{\varepsilon}}{2}, \hat{\mu}_0, \bar{\gamma}) = 0$. If $0 < \bar{\varepsilon} < \frac{\varepsilon}{2}$ and $\gamma < \bar{\gamma} < 1$, then $r_{\hat{f}}(n, \frac{\varepsilon}{2}, \frac{\bar{\varepsilon}}{2}, \hat{\mu}, \gamma) \leq r_{\hat{f}}(n, \bar{\varepsilon}, \frac{\bar{\varepsilon}}{2}, \hat{\mu}_0, \bar{\gamma})$. It follows that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log r_{\hat{f}}(n, \frac{\varepsilon}{2}, \frac{\bar{\varepsilon}}{2}, \hat{\mu}_0, \bar{\gamma}) = 0.$$

So we can find a large integer $n_* = n_*(f, \hat{\mu}, \eta, \gamma, \varepsilon, \bar{\varepsilon}, N)$ that is a multiple of N and a compact set $\hat{K}_0 := \hat{K}_0(\hat{f}, \hat{\mu}, \eta, \gamma, \varepsilon, \bar{\varepsilon}, n_*) \subset M$ with $\hat{\mu}_0(\hat{K}_0) > 1 - \gamma^2$, s.t.

$$\frac{1}{n_*} \log r_{\hat{f}}(n_*, \frac{\varepsilon}{2}, \frac{\bar{\varepsilon}}{2}, \hat{f}(\hat{K}_0)) < \frac{\eta}{10}.$$

– By continuity, we can choose a neighborhood $\widehat{W}_0 = \widehat{W}_0(f, \eta, \varepsilon, \bar{\varepsilon}, n_*, \hat{K}_0)$ of \hat{K}_0 with $\hat{\mu}(\partial \widehat{W}_0) = 0$, and a C^2 neighborhood $\mathcal{U}_{***} = \mathcal{U}_{***}(f, \eta, \varepsilon, \bar{\varepsilon}, n_*, \hat{K}_0)$ of f in $\text{Diff}^r(M)$ so that:

$$\forall g \in \mathcal{U}_{***}, \quad \frac{1}{n_*} \log r_{\hat{g}}(n_*, \varepsilon, \bar{\varepsilon}, \hat{g}(\widehat{W}_0)) < \frac{\eta}{10}. \quad (32)$$

Step 5 (\widehat{U}_0 and a decomposition of long typical orbit segments). Let

$$\widehat{U}_0 := \widehat{W}_0 \cup \widehat{W}_{\#}.$$

Then $\widehat{U}_0 = \widehat{U}_0(r, f, \hat{\mu}, Q, \eta, \gamma, \delta, N)$, $\hat{\mu}(\widehat{U}_0) > 1 - \gamma^2$ and $\hat{\mu}(\partial \widehat{U}_0) = 0$.

Claim 5.4. *Suppose $n > n_*/\gamma$. Any orbit segment $(\hat{x}, \hat{g}(\hat{x}), \dots, \hat{g}^{n-1}(\hat{x}))$ which spends a proportion of time larger than $1 - \gamma$ in \widehat{U}_0 can be decomposed into:*

- (a) orbit segments of length n_* with initial point in \widehat{W}_0 ,
- (b) orbit segments of length N with initial point in $\widehat{W}_{\#}$,
- (c) orbit segments of length 1, of total number less than $2\gamma n$.

(\widehat{W}_0 and $\widehat{W}_{\#}$ are not necessarily disjoint, so the decomposition may not be unique.)

Proof. A decomposition as in the statement is completely characterized by the increasing sequence of times $(n_0, n_1, \dots, n_{\ell})$, where $\hat{g}^{n_i}(\hat{x})$ is the initial point of the i -th segment. (So $n_0 = 0$, $n_{\ell} = n$.)

We set $n_0 = 0$ and define the sequence inductively. Assuming that $n_i < n$ has already been defined, we set

- (a) $n_{i+1} := n_i + n_*$ if $n_i + n_* \leq n$ and $\hat{g}^{n_i}(\hat{x}) \in \widehat{W}_0 \setminus \widehat{W}_{\#}$,
- (b) $n_{i+1} := n_i + N$ if $n_i + N \leq n$ and $\hat{g}^{n_i}(\hat{x}) \in \widehat{W}_{\#}$,
- (c) $n_{i+1} := n_i + 1$ otherwise.

We stop when $n_{i+1} = n$.

Since $n_* \geq N$, the times n_i such that $\hat{g}^{n_i}(\hat{x}) \in \widehat{U}_0$ but which are not associated to case (a) or (b) must satisfy $n_i > n - n_*$. Since $n > n_*/\gamma$ and since $(\hat{x}, \dots, \hat{g}^{n-1}(\hat{x}))$ spends a proportion of time larger than $1 - \gamma$ in \widehat{U}_0 , the set of times n_i corresponding to case (c) has size smaller than $2\gamma n$. \square

The sequence of times $\theta := (n_0, n_1, \dots, n_{\ell})$ obtained in the previous claim is called *type of a decomposition*. Recall that $H(t) = t \ln \frac{1}{t} + (1-t) \ln \frac{1}{1-t}$.

Claim 5.5. *There exists $n_H := n_H(\gamma)$ such that for all $n > n_H$, the number of possible types θ is less than $\exp(H(4\gamma)n)$.*

Proof. By our choices of N, N_0 and n_* , we have $n_*, N > 1/\gamma$. Hence there can be at most γn times n_i such that $n_{i+1} - n_i \in \{n_*, N\}$. Since there are also at most $2\gamma n$ times n_i such that $n_{i+1} - n_i = 1$, we must have $\ell \leq \lfloor 3\gamma n \rfloor$.

The number of types is thus bounded by $\sum_{\ell=1}^{\lfloor 3\gamma n \rfloor} \binom{n}{\ell-1}$. Since $3\gamma < \frac{1}{2}$, this is less than $3\gamma n \binom{n}{\lfloor 3\gamma n \rfloor}$, which by de Moivre's approximation is less than $\exp[nH(3\gamma) + o(n)]$. Since $4\gamma < \frac{1}{2}$, we have $H(3\gamma) < H(4\gamma)$, and the claim follows. \square

Step 6 (Definition of \mathcal{U}_0, \bar{n}_0). We fix the last parameters of our construction. Recall the C^2 neighborhoods of f in $\text{Diff}^r(M)$ introduced in Lemma 4.11 and eqs. (29), (30), (32).

- Let $\mathcal{U}_0 := \mathcal{U}_0(r, f, \hat{\mu}, Q, \eta, \gamma, \varepsilon, \hat{\varepsilon}, n_*, N) := \mathcal{V} \cap \mathcal{U}_* \cap \mathcal{U}_{**} \cap \mathcal{U}_{***}$.
- Define $\bar{n}_0 = \bar{n}_0(f, \hat{\mu}, \eta, \gamma, \delta, N) := \max\{n_*/\gamma, n_H\}$.

Step 7 (An inductive scheme). Now we fix some $g \in \mathcal{U}_0$ with $Q_{r,N}(g) < Q$, a regular C^r curve σ with C^r size at most $(\varepsilon, \hat{\varepsilon})$ and $n > \bar{n}_0$. We need to bound the minimal cardinality of a family \mathcal{R}_n of reparametrizations of σ which are $(C^r, g, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time n , over the set

$$T = \hat{\sigma}^{-1} \left\{ \hat{x}: (\hat{x}, \hat{g}(\hat{x}), \dots, \hat{g}^{n-1}(\hat{x})) \text{ is } \frac{\eta}{10}\text{-neutral and } \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\hat{g}^j(\hat{x})}(\hat{U}_0) \geq 1 - \gamma \right\}.$$

For each type $\theta = (n_0, \dots, n_\ell)$, we introduce the corresponding subset

$$T_\theta := T \cap \hat{\sigma}^{-1} \left\{ \hat{x}: (\hat{x}, \hat{g}(\hat{x}), \dots, \hat{g}^{n-1}(\hat{x})) \text{ has type } \theta \right\}.$$

Then T is the union of T_θ over all possible type θ .

Fixed some type $\theta = (n_0, \dots, n_\ell)$. We will build by induction a family $\mathcal{R}_{n_i}^\theta$ of reparametrizations ψ of σ over T_θ satisfying the following properties:

- (i) admissibility: $\mathcal{R}_{n_i}^\theta$ is $(C^r, g, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time n_i ;
- (ii) small cardinality: if $i \geq 1$,

$$\begin{aligned} |\mathcal{R}_{n_i}^\theta| &\leq \exp\left(\left(\frac{\lambda(\hat{f})}{r-1} + \frac{7\eta}{10}\right)n_i\right) |\mathcal{R}_{n_{i-1}}^\theta| && \text{when } n_i - n_{i-1} \geq N, \\ |\mathcal{R}_{n_i}^\theta| &\leq \exp(\eta/\gamma) |\mathcal{R}_{n_{i-1}}^\theta| && \text{otherwise;} \end{aligned}$$

- (iii) small length: for each $\psi \in \mathcal{R}_{n_i}^\theta$ and any $(x, E) \in \hat{\sigma} \circ \psi([0, 1])$,

$$\text{Length}(g^{n_i} \circ \sigma \circ \psi) < \varepsilon e^{-\frac{\eta}{10}n_i} \|Dg_x^{n_i}|_E\|.$$

At the end of the construction, one obtains a family $\mathcal{R}_n^\theta := \mathcal{R}_{n_\ell}^\theta$ over T_θ which is admissible up to time n . Then one can take the union over all θ and finish the construction.

We begin the construction by defining $\mathcal{R}_0^\theta := \{\text{Id}\}$: This meets our requirements because $n_0 = 0$ and σ has C^r size at most $(\varepsilon, \hat{\varepsilon})$.

Now we assume by induction that $\mathcal{R}_{n_i}^\theta$ has been constructed, and we build $\mathcal{R}_{n_{i+1}}^\theta$. The construction uses the concatenation procedure described in Lemma 4.10. For each $\psi \in \mathcal{R}_{n_i}^\theta$ we have to build a family \mathcal{R}_ψ of reparametrizations of the curve $g^{n_i} \circ \sigma \circ \psi$ over $\psi^{-1}(T_\theta)$ with the following properties:

- (i') \mathcal{R}_ψ is $(C^r, g, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time $n_{i+1} - n_i$,
- (ii') $\log |\mathcal{R}_\psi|$ is bounded by $(\frac{\lambda(\hat{f})}{r-1} + \frac{7\eta}{10})(n_{i+1} - n_i)$ if $n_{i+1} - n_i \geq N$ and by $\frac{\eta}{\gamma}$ otherwise,

(iii') for each $\varphi \in \mathcal{R}_\psi$ and $(x, E) \in \widehat{\sigma} \circ \psi \circ \varphi([0, 1])$,

$$\text{Length}(g^{n_{i+1}} \circ \sigma \circ \psi \circ \varphi) < \varepsilon e^{-\frac{\eta}{10} n_{i+1}} \|Dg_x^{n_{i+1}}|_E\|.$$

The family $\mathcal{R}_{n_{i+1}}^\theta := \{\psi \circ \varphi, \psi \in \mathcal{R}_{n_i}^\theta, \varphi \in \mathcal{R}_\psi\}$ then satisfies (i–iii) above.

– Given a type $\theta = (n_0, \dots, n_\ell)$, an integer $i \in \{0, \dots, \ell - 1\}$ and a reparametrization $\psi \in \mathcal{R}_{n_i}^\theta$, the construction of the families \mathcal{R}_ψ depends on which of the following cases from Claim 5.4 holds for n_i :

Case (a): $\widehat{g}^{n_i}(\widehat{x}) \in \widehat{W}_0$ and $n_{i+1} - n_i = n_*$.

Case (b): $\widehat{g}^{n_i}(\widehat{x}) \in \widehat{W}_\#$ and $n_{i+1} - n_i = N$.

Case (c): $n_{i+1} - n_i = 1$.

The three cases are discussed in steps 8–10 below.

In order to simplify the notations, we set

$$\sigma' := g^{n_i} \circ \sigma \circ \psi \text{ and } T' := \psi^{-1}(T_\theta).$$

Note that σ' has C^r size at most $(\varepsilon, \widehat{\varepsilon})$; moreover the induction assumption (iii) gives for each $(x, E) \in \widehat{g}^{-n_i} \circ \sigma'([0, 1])$,

$$\text{Length}(\sigma') < \varepsilon e^{-\frac{\eta}{10} n_i} \|Dg_x^{n_i}|_E\|. \quad (33)$$

Step 8 (Case (a)): In this case $n_{i+1} - n_i = n_*$ and $\widehat{\sigma}'(T') \subset \widehat{W}_0$. By eq. (32),

$$r_{\widehat{g}}(n_*, \varepsilon, \widehat{\varepsilon}, \widehat{g} \circ \sigma'(T')) \leq e^{\eta n_*/10}. \quad (34)$$

The integer n_* is a multiple of N . Corollary 4.14 of Yomdin's theorem yields a family \mathcal{R}_ψ^0 of reparametrizations of σ' over T' which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n_* and with cardinality:

$$|\mathcal{R}_\psi^0| \leq \Upsilon^{n_*/N} \|D\widehat{g}^N\|_{\sup}^{n_*/(r-1)N} r_{\widehat{g}}(n_*, \varepsilon, \widehat{\varepsilon}, \widehat{g} \circ \sigma'(T')).$$

Combining with (27), (28), (29), (34) and $N > \frac{1}{\gamma}$, we get $\log |\mathcal{R}_\psi^0| \leq (\frac{\lambda(\widehat{f})}{r-1} + \frac{4\eta}{10})n_*$.

The conclusion of Lemma 4.11, together with (33) gives that, for each $\varphi \in \mathcal{R}_\psi^0$ and $(x, E) \in \widehat{g}^{-n_i} \circ \sigma'([0, 1])$,

$$\text{Length}(g^{n_*} \circ \sigma' \circ \varphi) < \varepsilon e^{-\frac{\eta}{10} n_{i+1}} \|D_x g^{n_{i+1}}|_E\| \cdot e^{\frac{\eta}{5} n_*}. \quad (35)$$

In order to compensate for the factor $e^{\frac{\eta}{5} n_*}$, we subdivide $[0, 1]$ into intervals I_1, \dots, I_m with length less or equal to $e^{-\frac{\eta}{5} n_*}$. Since $n_* \geq N \geq N_0 > \frac{10}{\eta}$, $e^{\frac{\eta}{10} n_*} > 2$, whence

$$m \leq \lceil e^{\frac{\eta}{5} n_*} \rceil < 2e^{\frac{\eta}{5} n_*} < e^{\frac{3\eta}{10} n_*}.$$

Let $\chi_j: [0, 1] \rightarrow I_j$ be affine bijections, and let

$$\mathcal{R}_\psi := \{\varphi \circ \chi_j, \varphi \in \mathcal{R}_\psi^0, j = 1, \dots, m\}.$$

The cardinality of $\log |\mathcal{R}_\psi|$ is thus bounded as required by $(\frac{\lambda(\widehat{f})}{r-1} + \frac{7\eta}{10})n_*$, so (ii') holds. Property (iii') follows from (35) and the choice of m and χ_j , and (i') follows from Lemma 4.10.

Step 9 (Case (b)): In this case $n_{i+1} - n_i = N$ and $\widehat{\sigma}'(T') \subset \widehat{W}_\#$. We combine Lemma 4.11 with (33) and get that, for each $(x, E) \in \widehat{g}^{-n_i} \circ \widehat{\sigma}'([0, 1])$,

$$\text{Length}(g^N \circ \sigma') < \varepsilon e^{-\frac{\eta}{10}n_{i+1}} \|D_x g^{n_{i+1}}|_E\| \cdot e^{\frac{\eta}{5}N}.$$

One can thus subdivide $[0, 1]$ into intervals I_1, \dots, I_m with $m \leq \lceil e^{\frac{\eta}{5}N} \rceil \leq e^{\frac{3\eta}{10}N}$, such that:

$$\text{Length}(g^N \circ \sigma'(I_j)) < \varepsilon e^{-\frac{\eta}{10}n_{i+1}} \|D_x g^{n_{i+1}}|_E\| \quad \text{for } (x, E) \in \widehat{g}^{-n_i} \circ \widehat{\sigma}'([0, 1]). \quad (36)$$

We can focus on the intervals I_j such that $I_j \cap T' \neq \emptyset$. Fixing such an I_j , there exists $\widehat{x} = (x, E) \in (\widehat{g}^{-n_i} \circ \widehat{\sigma}')(I_j)$, such that $(\widehat{x}, \widehat{g}(\widehat{x}), \dots, \widehat{g}^{n_{i+1}-1}(\widehat{x}))$ is $\frac{\eta}{10}$ -neutral. In particular, $\|D_x g^{n_{i+1}}|_E\| \leq e^{\frac{\eta}{10}n_{i+1}}$, whence

$$\text{Length}(g^N \circ \sigma'(I_j)) < \varepsilon e^{-\frac{\eta}{10}n_{i+1}} e^{\frac{\eta}{10}n_{i+1}} = \varepsilon.$$

On the other hand since $\widehat{\sigma}'(T') \subset \widehat{W}_\#$, eq. (31) implies:

$$\text{diam}(\widehat{g}^N \circ \widehat{\sigma}'(T' \cap I_j)) < \widehat{\varepsilon}.$$

We have shown that the image $\widehat{g}^N \circ \widehat{\sigma}'(T' \cap I_j)$ is contained in a $(\varepsilon, \widehat{\varepsilon})$ -ball. We can apply Yomdin's Theorem 4.13 and obtain a family \mathcal{R}_j of reparametrizations φ of σ' over $T' \cap I_j$ with cardinality at most $\Upsilon \|D\widehat{g}^N\|_{\sup}^{1/(r-1)}$ such that each curve $g^N \circ \sigma' \circ \varphi$ has C^r size at most $(\varepsilon, \widehat{\varepsilon})$. Consequently, \mathcal{R}_j is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time N .

The union $\mathcal{R}_\psi := \bigcup_j \mathcal{R}_j$ is thus a family of reparametrizations φ of σ' over T' which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time N . By (36) they satisfy the bound (iii') of the induction scheme (Step 7) on the length of $\sigma' \circ \varphi$. Combining the bounds on m and $|\mathcal{R}_j|$, one bounds the cardinality of \mathcal{R}_ψ by

$$|\mathcal{R}_\psi| < \Upsilon e^{\frac{3\eta}{10}N} \|D\widehat{g}^N\|_{\sup}^{1/(r-1)}.$$

which by (27), (29) and $N > \frac{1}{\gamma}$, is bounded by $\exp(\frac{\lambda(\widehat{f})}{r-1}N + \frac{6\eta}{10}N)$ as required.

Step 10 (Case (c)): In this case $n_{i+1} - n_i = 1$. By our assumptions on σ' , we have $\text{Length}(\widehat{\sigma}') < \widehat{\varepsilon}$ and $\text{Length}(\sigma') < \varepsilon e^{-\frac{\eta}{10}n_i} \|Dg_x^{n_i}|_E\|$ for each $(x, E) \in \widehat{g}^{-n_i} \circ \widehat{\sigma}'([0, 1])$. One can thus decompose $[0, 1]$ into intervals I_1, \dots, I_m with $m \leq \lceil \|D\widehat{g}\|_{\sup} \rceil + \lceil e^{\frac{\eta}{10}} \|Dg\|_{\sup} \rceil - 1$ such that:

- (a) $\text{Length}(g \circ \sigma'(I_j)) < \varepsilon e^{-\frac{\eta}{10}n_{i+1}} \|Dg_x^{n_{i+1}}|_E\|$ for each $(x, E) \in \widehat{g}^{-n_i} \circ \widehat{\sigma}'([0, 1])$,
- (b) $\text{Length}(\widehat{g} \circ \widehat{\sigma}'(I_j)) < \widehat{\varepsilon}$.

In particular, for each $1 \leq j \leq m$, the image $\widehat{g} \circ \widehat{\sigma}'(I_j)$ is contained in a $(\varepsilon, \widehat{\varepsilon})$ -ball. We can thus apply Yomdin's Theorem 4.13 and obtain a family \mathcal{R}_j of reparametrizations of σ' over I_j which is $(C^r, g, 1, \varepsilon, \widehat{\varepsilon})$ -admissible up to time 1 and with cardinality at most $\Upsilon \|D\widehat{g}\|_{\sup}^{1/(r-1)}$.

The union $\mathcal{R}_\psi := \bigcup_j \mathcal{R}_j$ is thus a family of reparametrizations φ of σ' over $[0, 1]$ which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time 1. By property (a) above they satisfy the bound (iii') of the induction scheme (Step 7) on the length of $\sigma' \circ \varphi$. Note that $m \leq 3e^{\frac{\eta}{10}} \|D\widehat{g}\|_{\sup}^{1/(r-1)}$, hence \mathcal{R}_ψ has cardinality smaller than

$$|\mathcal{R}_\psi| < 3\Upsilon e^{\frac{\eta}{10}} \|D\widehat{g}\|_{\sup}^{1+1/(r-1)},$$

which by (27) and (29), is bounded by $\exp(\eta/\gamma)$.

Step 11 (Completion of the proof). Steps 7–10 provide the construction of the family \mathcal{R}_n^θ for each type θ . The inductive bounds (ii) for $|\mathcal{R}_{n_i}^\theta|/|\mathcal{R}_{n_{i-1}}^\theta|$ imply

$$|\mathcal{R}_n^\theta| \leq \exp \left(\frac{\lambda(\widehat{f})}{r-1} n + \frac{7\eta}{10} n + \frac{\eta}{\gamma} A_c(\theta) \right),$$

where $A_c(\theta)$ is the number of times n_i belongs to case (c) for the type θ .

Let \mathcal{R}_n denote the union of \mathcal{R}_n^θ over all possible types θ . This is a family of reparametrizations of σ over $T = \bigcup_\theta T_\theta$, which is $(C^r, g, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time n .

Since $A_c(\theta) \leq 2\gamma n$ for all θ (by Claim 5.4), since the number of types θ is bounded by $\exp(H(4\gamma)n)$ (by Claim 5.5), and since $H(4\gamma) < \frac{\eta}{10}$ (by our choice of γ , see (27)), this gives

$$\begin{aligned} |\mathcal{R}_n| &\leq \exp \left(\frac{\lambda(\widehat{f})}{r-1} n + \frac{7\eta}{10} n + \frac{\eta}{10\gamma} \cdot 2\gamma n + H(4\gamma)n \right) \\ &\leq \exp \left(\frac{\lambda(\widehat{f})}{r-1} n + \frac{7\eta}{10} n + \frac{\eta}{5} n + \frac{\eta}{10} n \right) = \exp \left(\frac{\lambda(\widehat{f})}{r-1} n + \eta n \right). \end{aligned}$$

This concludes the proof of Proposition 5.3. \square

6. THE NEUTRAL DECOMPOSITION

Let f be a homeomorphism on a compact metric space X . We denote the point mass measure at $x \in X$ by δ_x . Given $\mathfrak{N} \subset \mathbb{N}$, let

$$\mu_{x,n}^{\mathfrak{N}} := \frac{1}{n} \sum_{j \in [0,n) \cap \mathfrak{N}} \delta_{f^j(x)}.$$

The weak-* limit points of $(\mu_{x,n}^{\mathfrak{N}})_{n \geq 1}$ are called the \mathfrak{N} -empirical measures of x .

Definition 6.1. Suppose $\varphi: X \rightarrow \mathbb{R}$ is continuous, $\alpha > 0$ and $L \geq 1$. An (α, L) -neutral block of (x, f, φ) is an interval of integers $(n_0, n_0 + 1, \dots, n_1 - 1)$ s.t.

- $n_1 - n_0 \geq L$, and
- $\varphi(f^{n_0}(x)) + \varphi(f^{n_0+1}(x)) + \dots + \varphi(f^{n_1-1}(x)) \leq \alpha \cdot (n - n_0)$ for all $n_0 < n \leq n_1$.

We denote by $\mathfrak{N}_{\alpha,L}(x, f, \varphi)$ the union of the (α, L) -neutral blocks of (x, f, φ) .

Any interval of integers which is a union of two neutral blocks is still a neutral block. Therefore, if $\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(x))$, $\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=-n}^{-1} \varphi(f^k(x)) > \alpha$, then $\mathfrak{N}_{\alpha,L}(x, f, \varphi)$ is a disjoint union of (finite) maximal neutral blocks.

Proposition 6.2. Let f, f_1, f_2, \dots be homeomorphisms of a compact metric space X , and let $\varphi, \varphi_1, \varphi_2, \dots$ be continuous functions on X such that $f_k \rightarrow f$ and $\varphi_k \rightarrow \varphi$ uniformly. For each k , let ν_k be an ergodic probability for f_k such that $\int \varphi_k d\nu_k \geq 0$. Then there exist a subsequence (ν_{k_i}) and (positive) measures m_0, m_1 such that:

- (i) Both m_0 and m_1 are f -invariant.
- (ii) The subsequence (ν_{k_i}) converges weak-* to $m_0 + m_1$.
- (iii) For any neighborhoods $V = (V_0, V_1)$ of m_0, m_1 , if $\alpha < \alpha_*(V)$, $L \geq L_*(V)$, $i \geq i_*(V, \alpha, L)$, then for ν_{k_i} -a.e. x , the $\mathfrak{N}_{\alpha,L}(x, f_{k_i}, \varphi_{k_i})$ -empirical measures belong to V_0 , and the $(\mathbb{N} \setminus \mathfrak{N}_{\alpha,L}(x, f_{k_i}, \varphi_{k_i}))$ -empirical measures belong to V_1 .
- (iv) $\int \varphi dm_0 = 0$.
- (v) For m_1 -almost every point x , the limit of $\frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x))$ is positive.

Remark 6.3. The measures m_0, m_1 are not normalized, but there are f -invariant probabilities μ_0, μ_1 such that $m_0 = (1 - \beta)\mu_0$ and $m_1 = \beta\mu_1$, where $\beta = m_1(X)$.

Example 1. The following constructions show that the decomposition $m_0 + m_1$ depends on the sequence (ν_k) and not just on its limit.

Let $X = \{-2, 1, 2\}^{\mathbb{Z}}$, $f_k = f$ the left shift, and $\varphi_k(x) := \varphi(x) = x_0$. For each $k \geq 1$, consider the periodic sequence $p^{(k)}$ with period

$$\underbrace{-2, \dots, -2}_k, \underbrace{+2, \dots, +2}_k, \underbrace{+1, \dots, +1}_k.$$

Let ν_k be the unique shift invariant probability measure on the orbit of $p^{(k)}$. Let $\delta_s :=$ the probability measure concentrated on $(\dots s, s, s \dots)$. It is easy to see that ν_k converges to $\mu := \frac{1}{3}(\delta_{-2} + \delta_{+1} + \delta_{+2})$

If $0 < \alpha < 1$, $L \geq 1$, and $k \geq L/2$, then the maximal (α, L) -neutral blocks of $(p^{(k)}, f_k, \varphi_k)$ are $[n_0, n_1] \cap \mathbb{Z}$ where

$$(p_{n_0}^{(k)}, \dots, p_{n_1-1}^{(k)}) = (\underbrace{-2, \dots, -2}_k, \underbrace{+2, \dots, +2}_k, \underbrace{+1, \dots, +1}_\ell),$$

with $\ell = \lfloor \frac{2\alpha}{1-\alpha} k \rfloor$. So $m_0 = \frac{1}{3}(\delta_{-2} + \delta_{+2})$ and $m_1 = \frac{1}{3}\delta_{+1}$.

Now consider the measures ν'_k obtained from the periodic sequence $q^{(k)}$ with period $\underbrace{-2, \dots, -2}_k, \underbrace{+1, \dots, +1}_k, \underbrace{+2, \dots, +2}_k$. These measures also converge to μ . But now, if $0 < \alpha < 1$, $L \geq 1$ and $k > L/2$, then the maximal neutral blocks of $(q^{(k)}, f_k, \varphi_k)$ are $[n_0, n_1] \cap \mathbb{Z}$ where

$$(q_{n_0}^{(k)}, \dots, q_{n_1-1}^{(k)}) = (\underbrace{-2, \dots, -2}_k, \underbrace{+1, \dots, +1}_k, \underbrace{+2, \dots, +2}_\ell),$$

with $\ell = \lfloor \frac{2\alpha+1}{2-\alpha} k \rfloor$. So $m_0 = \frac{1}{3}\delta_{-2} + \frac{1}{3}\delta_{+1} + \frac{1}{6}\delta_{+2}$ and $m_1 = \frac{1}{6}\delta_{+2}$.

Example 2: Suppose (f_k) converges to f in $\text{Diff}^r(M)$, and ν_k are f_k -invariant measures which converge to an f -invariant measure μ . Assume the limiting measure $\mu := \lim \nu_k$ is ergodic and hyperbolic of saddle type. Let $\lambda_\mu^- < 0, \lambda_\mu^+$ be the Lyapunov exponents of μ .

Consider the unstable lifts $\widehat{\nu}_k^u$ to the fibered bundle. Passing to a subsequence, we may assume that $\widehat{\nu}_k^u$ converge weak-* to a limit $\widehat{\mu}$ (a lift of μ). Let $\varphi_k : \widehat{M} \rightarrow \mathbb{R}$ be $\varphi_k(x, E) := \log \|(Df_k)_x|_E\|$. We apply Proposition 6.2 to $\widehat{f}_k, \varphi_k, \widehat{\nu}_k^u$, obtaining a decomposition $\widehat{\mu} = m_0 + m_1 = (1 - \beta)\widehat{\mu}_0 + \beta\widehat{\mu}_1$. On the other hand, since μ is ergodic and hyperbolic,

$$\widehat{\mu} = a\widehat{\mu}^- + (1 - a)\widehat{\mu}^+$$

where $\widehat{\mu}^\pm$ are the unique lifts of μ to $\text{graph}(E^\pm)$. So $\widehat{\mu}_0, \widehat{\mu}_1 \ll \widehat{\mu}^+ + \widehat{\mu}^-$, whence by the ergodicity of $\widehat{\mu}^\pm$, $\widehat{\mu}_0, \widehat{\mu}_1$ are convex combinations of $\widehat{\mu}^+, \widehat{\mu}^-$. By (v), $\widehat{\mu}_1$ has no $\widehat{\mu}^-$ component, so $\widehat{\mu}_1 = \widehat{\mu}^+$. By (iv), if $\widehat{\mu}_0 = b\widehat{\mu}^- + (1 - b)\widehat{\mu}^+$ then necessarily $0 = \int \varphi d\widehat{\mu}_0 = b\lambda_\mu^- + (1 - b)\lambda_\mu^+$, whence $b = \lambda_\mu^+ / (\lambda_\mu^+ + |\lambda_\mu^-|)$. It follows that

$$\widehat{\mu}_1 = \widehat{\mu}^+, \quad \widehat{\mu}_0 = \frac{\lambda_\mu^+ \widehat{\mu}^- + |\lambda_\mu^-| \widehat{\mu}^+}{\lambda_\mu^+ + |\lambda_\mu^-|}.$$

To finish the calculation of $m_0 = (1 - \beta)\widehat{\mu}_0, m_1 = \beta\widehat{\mu}_1$ it remains to determine β . To do this, we substitute the formulas for $\widehat{\mu}_0, \widehat{\mu}_1$ in the identity $a\widehat{\mu}^- + (1 - a)\widehat{\mu}^+ = \widehat{\mu} = (1 - \beta)\widehat{\mu}_0 + \beta\widehat{\mu}_1$. Since $\widehat{\mu}^\pm$ are ergodic, the coefficient of $\widehat{\mu}^-$ on both sides must be equal, which leads to $\beta = 1 - a \frac{\lambda_\mu^+ + |\lambda_\mu^-|}{\lambda_\mu^+}$. Looking at case (2) of Theorem 3.6, we recognize that

$$\beta = \frac{1}{\lambda^+(f, \mu)} \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k).$$

Proof of Proposition 6.2. Without loss of generality, (ν_k) converges weak-* to an f -invariant probability measure μ (otherwise pass to a suitable subsequence). We abbreviate $\mathfrak{N}_{\alpha,L}^k(x) := \mathfrak{N}_{\alpha,L}(x, f_k, \varphi_k) \subset \mathbb{Z}$ and define

$$N_{\alpha,L}^k := \{x \in X : 0 \in \mathfrak{N}_{\alpha,L}^k(x)\}. \quad (37)$$

This is a measurable set. We call it the (α, L) -neutral set of (φ_k, f_k) . Let $\chi_{\alpha,L}^k$ denote the indicator function of $N_{\alpha,L}^k$. Since ν_k is ergodic, for ν_k -a.e. point x ,

$$\lim_{n \rightarrow \infty} \mu_{x,n}^{\mathfrak{N}_{\alpha,L}^k(x)} = \chi_{\alpha,L}^k \nu_k \text{ in the weak-* topology.}$$

Claim 6.4. *There exists an increasing sequence of integers (k_i) such that*

$$\forall (\alpha, L) \in (0, 1] \times \mathbb{N} \quad \lim_{i \rightarrow \infty} \chi_{\alpha,L}^{k_i} \nu_{k_i} \text{ exists}$$

in the weak- topology. We write these limits as $m_{\alpha,L}$.*

Proof of the claim. Fix some countable dense set $E \subset (0, 1]$. By compactness and a diagonal argument, there is an increasing sequence $k_i \rightarrow \infty$ such that the following limits exist in the weak-* topology:

$$\forall (\alpha, L) \in E \times \mathbb{N} \quad \lim_{p \rightarrow \infty} \chi_{\alpha,L}^{k_i} \nu_{k_i} = m_{\alpha,L}.$$

Let us check that this can be extended to all $(\alpha, L) \in (0, 1] \times \mathbb{N}$, maybe after passing to a subsequence. Indeed, select a countable family (u_j) of nonnegative continuous functions which generate a countable dense algebra over \mathbb{Q} in $C^0(M)$ (the space of continuous real-valued functions on M with the supremum norm). Fix L and u_j . The function $\alpha \in E \mapsto m_{\alpha,L}(u_j)$ is non-decreasing on E , and therefore extends uniquely to a left-continuous function $\alpha \in [0, 1] \rightarrow \tilde{m}_{\alpha,L}(u_j)$. The discontinuity points form a countable set $D_{L,j}$. Again by monotonicity with respect to α ,

$$m_{\alpha,L}(u_j) = \tilde{m}_{\alpha,L}(u_j) = \lim_i \chi_{\alpha,L}^{k_i} \nu_{k_i}(u_j)$$

at every $\alpha \in [0, 1] \setminus D_{L,j}$. By a further extraction of a subsequence, we ensure that $\chi_{\alpha,L} \nu_{k_i}$ converge for all (α, L) in the countable set $\bigcup_{L \in \mathbb{N}, j \geq 1} D_{L,j}$. \square

We return to the proof of Proposition 5.2. To simplify notation, from now on (ν_k) will denote the subsequence (ν_{k_i}) .

The following weak-* limit exists by monotonicity:

$$m_0 := \lim_{\substack{\alpha \rightarrow 0 \\ L \rightarrow \infty}} m_{\alpha,L} = \inf_{\alpha > 0, L \geq 1} m_{\alpha,L}.$$

We set $m_1 := \mu - m_0$. Since $0 \leq \chi_{\alpha,L}^k \nu_k \leq \nu_k$, it follows that $0 \leq m_0 \leq \mu$ so that both m_0 and m_1 are positive measures.

Neutral blocks have length at least L , therefore for every continuous function u , we have $(\chi_{\alpha,L}^k \nu_k)(u - u \circ f_k) \leq (2/L)\|u\|_{\sup}$. Since $u \circ f - u \circ f_k \rightarrow 0$ uniformly, $m_{\alpha,L}(u - u \circ f) \leq (2/L)\|u\|_{\sup}$. It follows that m_0 is f -invariant. So is $m_1 = \mu - m_0$. This proves items (i) and (ii).

Item (iii) is a simple consequence of the construction.

We turn to (iv). For any function $\psi: X \rightarrow \mathbb{R}$, we define

$$S_n^k \psi := \sum_{j=0}^{n-1} \psi \circ f_k^j \text{ and } S_n^\infty \psi := \sum_{j=0}^{n-1} \psi \circ f^j.$$

For every x , we decompose $\mathfrak{N}_{\alpha,L}^k(x) \cap [0, \infty)$ into maximal disjoint intervals:

$$\mathfrak{N}_{\alpha,L}^k(x) \cap [0, \infty) = \bigsqcup_{i \geq 1} [a_i, a_i + b_i).$$

Since ν_k is ergodic, for ν_k -a.e. x ,

$$\begin{aligned} (\chi_{\alpha,L}^k \nu_k)(\varphi_k) &= \int \chi_{\alpha,L}^k \varphi_k d\nu_k = \lim_{n \rightarrow \infty} \frac{1}{n} S_n^k(\chi_{\alpha,L}^k \varphi_k)(x) \\ &= \lim_{j \rightarrow \infty} \frac{1}{a_j + b_j} \sum_{i \leq j} (S_{b_i}^k \varphi_k)(f_k^{a_i}(x)). \end{aligned}$$

Each interval $[a_i, a_i + b_i)$ is a maximal (α, L) -neutral block except possibly the initial one, if it contains 0. The first block contributes $C_0(x)/n \rightarrow 0$ to the limit. The other blocks are all maximal neutral blocks, and satisfy the bounds

$$\alpha(b_i + 1) - \varphi_k(f_k^{a_i+b_i}(x)) < (S_{b_i}^k \varphi_k)(f_k^{a_i}(x)) \leq \alpha b_i.$$

The first inequality comes from the maximality of the block, the second is the definition of neutrality. Summing over $i = 1, \dots, j$, we obtain the bounds

$$C_0(x) + \alpha \left(\sum_{1 \leq i \leq j} b_i + j \right) - j \sup_{x,k} \varphi_k(x) < (S_{b_j+a_j}^k \chi_{\alpha,L} \varphi_k)(x) \leq C_0(x) + \alpha \sum_{1 \leq i \leq j} b_i.$$

Since each such complete neutral block has length at least L , there are $j \leq n/L$ maximal (α, L) -blocks in $[0, a_j + b_j)$. Dividing by $a_j + b_j \geq \sum_{i \leq j} b_i$ and discarding some nonnegative terms from the lower bound, we obtain in the limit $j \rightarrow \infty$,

$$-\sup_{x,k} \varphi_k(x)/L < (\chi_{\alpha,L} \nu_k)(\varphi_k) \leq \alpha.$$

Passing to the limit $\alpha \rightarrow 0, L \rightarrow \infty$ and recalling that $\varphi_k \rightarrow \varphi$ uniformly, we obtain item (iv): $m_0(\varphi) = 0$.

We prove item (v) by contradiction, assuming that

$$\gamma := \frac{1}{2} m_1(\{x : \lim_{n \rightarrow \infty} (1/n) (S_n^\infty \varphi)(x) \leq 0\}) > 0.$$

There are $\alpha_0 > 0, L_0 < \infty$ such that, for $0 < \alpha \leq \alpha_0, L \geq L_0$,

$$|m_{\alpha,L}(X) - m_0(X)| < \frac{\gamma}{100}. \quad (38)$$

Given $K \geq 0$, let

$$\begin{aligned} V_0^k(K) &:= \{x \in M \mid \exists 0 \leq a \leq K \text{ s.t. } [-a, 0] \cap \mathbb{Z} \text{ is } (\alpha_0, L_0)\text{-neutral for } (x, f_k, \varphi_k)\}, \\ W_0^\infty(K) &:= \{x \in M \mid \exists 0 \leq a \leq K \text{ s.t. } [-a, 0] \cap \mathbb{Z} \text{ is } (\alpha_0/2, L_0)\text{-neutral for } (x, f, \varphi)\}. \end{aligned}$$

These are closed sets, and $W_0^\infty(K) \subset V_0^k(K)$ for all large k . We can ensure that $\alpha_0/2$ is not a member of the countable set of α 's such that

$$\mu(\{x \in X : \exists n \geq 1 \ (1/n) S_n^\infty \varphi(x) = \alpha\}) > 0.$$

In particular, $\mu(\partial W_0^\infty(K)) = 0$ for each integer $K \geq 1$. Therefore, for any $K \geq 1$,

$$\lim_{\alpha, L} \lim_k \nu_k(W_0^\infty(K) \setminus N_{\alpha,L}^k) = \lim_{\alpha, L} \lim_k [(1 - \chi_{\alpha,L}^k) \nu_k](W_0^\infty(K)) = m_1(W_0^\infty(K)). \quad (39)$$

By the ergodic theorem, for μ -a.e. $x \in M$ such that $\lim_{n \rightarrow \infty} (1/n) S_n^\infty \varphi(x) \leq 0$, it is also the case that $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=-n+1}^0 \varphi(f^j(x)) \leq 0$. For such x , the Pliss lemma [40], [32, Ch.

IV.11] yields arbitrarily large integers $a \geq 0$ such that $[-a, 0]$ is $\alpha_0/2$ -neutral. In particular, fixing $K_0 \geq L_0$ large enough,

$$m_1(W_0^\infty(K_0)) > \gamma.$$

Hence there exist $0 < \alpha_1 \leq \alpha_0$ and $L_1 > \max(200K_0/\gamma, L_0)$ such that for all k large enough,

$$\nu_k(V_0^k(K_0) \setminus N_{\alpha_1, L_1}^k) \geq \nu_k(W_0^\infty(K_0) \setminus N_{\alpha_1, L_1}^k) > \gamma. \quad (40)$$

Since ν_k is ergodic, for ν_k -a.e. x , the set $\mathfrak{N}_1 := \mathfrak{N}_{\alpha_1, L_1}^k(x)$ of visits under iterations of f_k to the (α_1, L_1) -neutral set N_{α_1, L_1}^k has density

$$d(\mathfrak{N}_1) := \lim_{n \rightarrow \infty} \frac{|\mathfrak{N}_1 \cap [0, n-1]|}{n} = \nu_k(N_{\alpha_1, L_1}^k) \xrightarrow{k \rightarrow \infty} m_{\alpha_1, L_1}(X).$$

So by eq. (38), for k large enough and ν_k -a.e. x ,

$$d(\mathfrak{N}_1) > m_0(X) - \gamma/100. \quad (41)$$

Similarly, the set $\mathfrak{N}_0 := \mathfrak{N}_{\alpha_0, L_0}^k(x)$ has density

$$d(\mathfrak{N}_0) = \nu_k(N_{\alpha_0, L_0}^k) < m_0(X) + \gamma/100. \quad (42)$$

for k large enough and ν_k -a.e. x . Finally, let \mathfrak{V} denotes the set of j such that $f^j(x) \in V_0^k(K_0) \setminus N_{\alpha_1, L_1}^k$. Then by eq. (40), for k large enough and ν_k -a.e. x ,

$$d(\mathfrak{V}) = \nu_k(V_0^k(K_0) \setminus N_{\alpha_1, L_1}^k) > \gamma. \quad (43)$$

We will show that (41)–(43) lead to a contradiction.

By definition of $V_0^k(K_0)$, each $j \in \mathfrak{V}$ is the last element of an (α_0, L_0) -neutral block $I(j)$ with length $\leq K_0$ (we do not claim that this block is maximal). Let

$$\mathfrak{J} := \bigcup \{I(j) : j \in \mathfrak{V}, I(j) \cap \mathfrak{N}_1 = \emptyset\}, \quad \mathfrak{J}' := \bigcup \{I(j) : j \in \mathfrak{V}, I(j) \cap \mathfrak{N}_1 \neq \emptyset\}.$$

We claim that the upper asymptotic density $\bar{d}(\mathfrak{J}') := \limsup \frac{1}{n} |\mathfrak{J}' \cap [0, n]|$ is less than $\gamma/100$. To see this note that if $j \in \mathfrak{V}$ and $I(j) \cap \mathfrak{N}_1 \neq \emptyset$, then $j \notin \mathfrak{N}_1$ (by definition of \mathfrak{V}) and since $L_0 < L_1$, $I(j)$ contains the last element of a maximal sub-interval of \mathfrak{N}_1 . The interval with length $2K_0$ centered at this last element must contain $I(j)$. Since the number of maximal sub-intervals of $\mathfrak{N}_1 \cap [0, n]$ is bounded by n/L_1 , the upper asymptotic density of \mathfrak{J}' is no more than $2K_0/L_1 < \gamma/100$.

It follows that the upper asymptotic density of \mathfrak{J} is at least

$$\bar{d}(\mathfrak{J}) \geq d(\mathfrak{V}) - \frac{\gamma}{100} > \frac{99}{100}\gamma,$$

and since \mathfrak{N}_1 and \mathfrak{J} are disjoint $\bar{d}(\mathfrak{N}_1 \cup \mathfrak{J}) = d(\mathfrak{N}_1) + \bar{d}(\mathfrak{J}) > m_0(X) + \frac{98}{100}\gamma$.

But \mathfrak{N}_1 and \mathfrak{J} are a union of (α_0, L_0) -neutral blocks, so $\mathfrak{N}_1 \cup \mathfrak{J} \subset \mathfrak{N}_0$, whence by eq. (42), $\bar{d}(\mathfrak{N}_1 \cup \mathfrak{J}) < m_0(X) + \gamma/100$. This contradiction proves item (v). \square

7. PROOF OF THE MAIN THEOREM

We recall the notation $\widehat{\lambda}(\widehat{f}, \widehat{\mu}) := \int_{\widehat{M}} \log \|Df_x|_E\| d\widehat{\mu}(x, E)$. In this section, we prove the following stronger version of Theorem C.

Theorem D. *Fix a real number $r > 2$. For every $k \geq 1$, let $f_k \in \text{Diff}^r(M)$ and let ν_k be an f_k -ergodic measure. Let $\widehat{\nu}_k$ be an \widehat{f}_k -ergodic lift satisfying $\widehat{\lambda}(\widehat{f}_k, \widehat{\nu}_k) = \lambda^+(f_k, \nu_k)$ such that:*

- *the limits $\lim_k \lambda^+(f_k, \nu_k)$ and $\lim_k h(f_k, \nu_k)$ exist and $\lim_k \lambda^+(f_k, \nu_k) \geq 0$,*
- *$f_k \xrightarrow{r\text{-}bd} f$ for some $f \in \text{Diff}^r(M)$ (i.e. $f_k \rightarrow f$ uniformly and $\sup_k \|f_k\|_{C^r} < \infty$),*
- *$\widehat{\nu}_k \xrightarrow{w^*} \widehat{\mu}$ for some \widehat{f} -invariant probability measure $\widehat{\mu}$ on \widehat{M} , perhaps non-ergodic.*

Then there exist $\beta \in [0, 1]$, two f -invariant measures μ_0, μ_1 with \hat{f} -invariant lifts $\hat{\mu}_0, \hat{\mu}_1$ s.t.

$$\hat{\mu} = (1 - \beta)\hat{\mu}_0 + \beta\hat{\mu}_1, \quad (44)$$

$$\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \beta \lambda^+(f, \mu_1), \quad (45)$$

$$\lim_{k \rightarrow \infty} h(f_k, \nu_k) - \frac{1}{r-1} \lambda(\hat{f}) \leq \beta h(f, \mu_1). \quad (46)$$

Moreover:

- If $\beta > 0$, then $\hat{\lambda}(\hat{f}, \hat{\mu}_1) = \lambda^+(f, \mu_1)$ and $\lambda^+(f, x) > 0$ for μ_1 -a.e. x .
- If $\beta < 1$, then $\hat{\lambda}(\hat{f}, \hat{\mu}_0) = 0$.

Note that when ν_k is hyperbolic, the measure $\hat{\nu}_k$ above is simply the unstable lift $\hat{\nu}_k^+$.

7.1. Reductions. We assume the setting of Theorem D. There is no loss of generality in assuming that r is finite, since the C^∞ case follows from the C^r case by letting $r \rightarrow \infty$. By Lemma 3.1 and since $f_k \xrightarrow{r\text{-bd}} f$ with $r > 2$,

$$f_k \rightarrow f \text{ in the } C^2 \text{ topology and } \hat{f}_k \rightarrow \hat{f} \text{ in the } C^1 \text{ topology.} \quad (47)$$

Let $h := \lim_{k \rightarrow \infty} h(f_k, \nu_k)$. By Ruelle's inequality and (47), h is bounded by $\sup_k \|Df_k\|_{\sup} < \infty$. It is clearly non-negative. The theorem has a simple proof when $h = 0$:

Proof of Theorem D when $h = 0$. In this case eq. (46) is trivial. Since $\log \|Dg(x)|_E\|$ depends continuously on (x, E) and g , one gets $0 \leq \lim_k \lambda^+(f_k, \nu_k) = \lim_k \lambda(f_k, \hat{\nu}_k) = \lambda(f, \hat{\mu})$.

If all ergodic components $\hat{\mu}'$ of $\hat{\mu}$ satisfy $\hat{\lambda}(f, \hat{\mu}') > 0$, it is enough to take $\hat{\mu}_0 = \hat{\mu}_1 = \hat{\mu}$ and fix $\beta = 1$. If some ergodic components of $\hat{\mu}$ satisfy $\hat{\lambda}(f, \hat{\mu}) \leq 0$ and since $\hat{\lambda}(f, \hat{\mu}) > 0$, one can decompose $\hat{\mu} = (1 - \beta)\hat{\mu}_1 + \beta\hat{\mu}_0$ where $\hat{\mu}_0, \hat{\mu}_1$ are two \hat{f} -invariant measures such that $\hat{\lambda}(f, \hat{\mu}_0) = 0$ and all ergodic components $\hat{\mu}'$ of $\hat{\mu}_1$ satisfy $\hat{\lambda}(f, \hat{\mu}') > 0$. \square

Henceforth, we assume that

$$h := \lim_{k \rightarrow \infty} h(f_k, \nu_k) > 0, \text{ and } h(f_k, \nu_k) > 0 \text{ for all } k.$$

In particular, each measure ν_k is hyperbolic, i.e. has one positive and one negative Lyapunov exponent. Note that it is enough to prove the theorem for any convenient further subsequence.

7.2. The decomposition of the limiting measure. Theorem D is stated in terms of the properties of a special decomposition $\mu = (1 - \beta)\mu_0 + \beta\mu_1$ of the $\mu = \lim \nu_k$. In this section we construct β, μ_0 and μ_1 .

The idea is to apply Proposition 6.2 to a suitable sequence of measures. By Ruelle's inequality and the reduction to the case $h(f_k, \nu_k) > 0$, ν_k must be f_k -hyperbolic of saddle type. Let $\hat{\nu}_k^+$ denote the unstable lift of ν_k to \widehat{M} , and let \hat{f}_k, \hat{f} be the lifts of f_k, f to \widehat{M} . Define $\varphi_k, \varphi : \widehat{M} \rightarrow \mathbb{R}$ by

$$\varphi_k(x, E) := \log \|Df_k|_E\|, \varphi(x, E) := \log \|Df|_E\|.$$

We apply Proposition 6.2 to $\widehat{M}, \hat{f}_k, \varphi_k, \hat{\nu}_k^+$. (The proposition is applicable, because by eq. (47), $\hat{f}_k \rightarrow \hat{f}$ in $\text{Diff}^1(\widehat{M})$ and $\varphi_k \rightarrow \varphi$ uniformly on \widehat{M} , and because by Lemma 3.3 and Ruelle's inequality, $\int \varphi_k d\hat{\nu}_k = \lambda^+(f_k, \nu_k) \geq h(f_k, \nu_k) > 0$.)

Proposition 6.2 gives us a subsequence $\{k_i\}$ and two finite positive measures \hat{m}_0, \hat{m}_1 with the following properties.

- (i) $\hat{m}_i \circ \hat{f}^{-1} = \hat{m}_i$.

- (ii) $\widehat{\nu}_{k_i}^+ \xrightarrow{i \rightarrow \infty} \widehat{m}_0 + \widehat{m}_1 =: \widehat{\mu}$. The limit $\widehat{\mu}$ is \widehat{f} -invariant and lifts μ .
- (iii) Suppose $\widehat{V}_0, \widehat{V}_1$ are weak-* open sets of measures such that $\widehat{V}_i \ni \widehat{m}_i$, then there are $\alpha_*(\widehat{V}_0, \widehat{V}_1) \in (0, 1)$ and $L_*(\widehat{V}_0, \widehat{V}_1) \geq 1$ as follows. If $0 < \alpha < \alpha_*$ and $L > L_*$, then for all $k_i > k_*(\widehat{V}_0, \widehat{V}_1, \alpha, L)$, for $\widehat{\nu}_{k_i}$ -a.e. $\widehat{x} \in \widehat{M}$,
 - the $\mathfrak{N}_{\alpha, L}(\widehat{x}, \widehat{f}_{k_i}, \varphi_{k_i})$ -empirical measures of \widehat{x} belong to \widehat{V}_0 ,
 - the $\mathbb{N} \setminus \mathfrak{N}_{\alpha, L}(\widehat{x}, \widehat{f}_{k_i}, \varphi_{k_i})$ -empirical measures of \widehat{x} belong to \widehat{V}_1 .
- (iv) $\int \varphi d\widehat{m}_0 = 0$.
- (v) For \widehat{m}_1 -a.e. point \widehat{x} , $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(\widehat{f}^j(\widehat{x})) > 0$.
- (vi) $h(f_{k_i}, \nu_{k_i}) \xrightarrow{i \rightarrow \infty} h > 0$ and $h(f_{k_i}, \nu_{k_i}) > 0$ for all i .
- (vii) $\lambda^+(f_{k_i}, \nu_{k_i}) \xrightarrow{i \rightarrow \infty} \int \varphi d\widehat{\mu}$.

Parts (i)–(v) are in Proposition 6.2; Part (vi) is the reduction in §7.1; and Part (vii) is because $\lambda^+(f_{k_i}, \nu_{k_i}) = \int \varphi_{k_i} d\widehat{\nu}_{k_i}^+$ (by Lemma 3.3), $\varphi_{k_i} \rightarrow \varphi$ uniformly, and $\widehat{\nu}_{k_i}^+ \rightarrow \widehat{\mu}$ weak-*. To keep the notation as simple as possible we will henceforth assume without loss of generality that $\{\nu_k\} = \{\nu_{k_i}\}$.

Let

$$\beta := 1 - \widehat{m}_0(\widehat{M}) = \widehat{m}_1(\widehat{M}).$$

$$\widehat{\mu}_i := \frac{1}{\widehat{m}_i(\widehat{M})} \widehat{m}_i, \text{ or any invariant probability measure if } \widehat{m}_i(\widehat{M}) = 0.$$

$\mu_i :=$ the projections of $\widehat{\mu}_i$ to the corresponding f -invariant measures on M .

Notice that $\widehat{\mu} = (1 - \beta)\widehat{\mu}_0 + \beta\widehat{\mu}_1$, $\mu = (1 - \beta)\mu_0 + \beta\mu_1$, and $0 \leq \beta \leq 1$.

Claim 7.1. *If $\beta < 1$, then $\widehat{\mu}_0(\varphi) = \widehat{m}_0(\varphi)/(1 - \beta) = 0$.*

Proof. This is property (iv). □

Claim 7.2. *$\beta \neq 0$. Consequently, $\widehat{m}_1(\widehat{M}) \neq 0$, and μ_1 is a probability measure.*

Proof. Assume by contradiction that $\beta = 0$. Then $\widehat{\mu} = \widehat{m}_0$, and

$$\begin{aligned} 0 &= \int \varphi d\widehat{m}_0 = \int \varphi d\widehat{\mu}, \quad \text{by (iv)} \\ &= \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k), \quad \text{by (vii) and the assumption that } \{k_i\} = \{k\} \\ &\geq \lim_{k \rightarrow \infty} h(f_k, \nu_k) > 0, \quad \text{by Ruelle's inequality and (vi).} \end{aligned}$$

This contradiction shows that $\beta \neq 0$. □

Claim 7.3. *μ_1 -a.e. x has one positive and one non-positive Lyapunov exponent.*

Proof. By (v), the definition of φ , and Fubini's Theorem, for μ_1 -a.e. $x \in M$, there exists a one-dimensional subspace $E \subset T_x M$ such that

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df^n|_E\| = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ \widehat{f}^j)(x, E) > 0.$$

It follows that μ_1 -a.e. x has at least one positive Lyapunov exponent.

Assume by contradiction that the claim is false, then there is an f -invariant set Ω of positive μ_1 -measure such that every $x \in \Omega$ has two (possibly equal) positive Lyapunov exponents. Recall the following well-known fact:

FACT: If $\dim(M) = 2$, then μ_1 -almost every x with two positive Lyapunov exponents has an open neighborhood U_x such that $\lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df_y^{-n}\| < 0$ for all $y \in U_x$.

Proof. Pesin's local stable manifold theorem [38, Thm 2.2.1] implies that μ_1 -almost every point x admits a neighborhood U_x and constants $C > 0$ and $\kappa \in (0, 1)$ such that for any $y \in U_x$ and $n \geq 0$,

$$d(f^{-n}(x), f^{-n}(y)) \leq C\kappa^n.$$

Since the orbit of x is recurrent, this implies that the forward orbit of x converges towards a periodic orbit O and in fact must coincide with that periodic orbit, again by recurrence. As a consequence, x is a hyperbolic sink so $\|Df_x^{-N}\| < \frac{1}{2}$ for some $N \geq 1$ such that $f^N(x) = x$. \square

This fact enables us to build an open set U such that

$$\mu_1(U) > 0, \mu_1(\partial U) = \mu_0(\partial U) = 0, \lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df_y^{-n}\| < 0 \text{ for all } y \in U.$$

By (ii), $\nu_k \xrightarrow[k \rightarrow \infty]{w^*} (1 - \beta)\mu_0 + \beta\mu_1$, and $\beta \neq 0$ by claim 7.2. So $\nu_k(U) > 0$ for all k large enough. But this is a contradiction, because the ν_k have one negative Lyapunov exponent, so that $\lim_{n \rightarrow +\infty} \frac{1}{n} \log \|Df_y^{-n}\| > 0$ ν_k -almost everywhere. \square

Claim 7.4. $\hat{\mu}_1$ is the unstable lift $\hat{\mu}_1^+$ of μ_1 .

Proof. By Claim 7.3 and the Oseledets theorem, for μ_1 -a.e. x , $T_x M = E^u(x) \oplus E^c(x)$, where $Df_x E^*(x) = E^*(f(x))$, ($*$ = u, c). By Corollary 3.4, μ_1 has a unique lift $\hat{\mu}_1^+$ to $\text{graph}(E^u)$, and all other lifts charge some part of $\text{graph}(E^c)$. Since $\lim(1/n) \sum_{j=0}^{n-1} \varphi(\hat{f}^j(\hat{x})) \leq 0$ on $\text{graph}(E^c)$, property (v) forces $\hat{\mu}_1 = \hat{\mu}_1^+$ a.e. \square

7.3. Proof of Theorem D part (1). We compare the exponents of ν_k and μ .

$$\begin{aligned} \beta\lambda^+(f, \mu_1) &= \beta \int \varphi d\hat{\mu}_1^+ \text{ by Lemma 3.3} \\ &= \beta \int \varphi d\hat{\mu}_1 = \int \varphi d\hat{m}_1 \text{ by claim 7.4 and definition of } \hat{m}_1 \\ &= \int \varphi d\hat{m}_1 + \int \varphi d\hat{m}_0 = \int \varphi d\hat{\mu} \text{ by (iv) and definition of } \hat{\mu} \\ &= \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) \text{ by (vii) and the convention } \{\nu_{k_i}\} = \{\nu_k\}. \end{aligned}$$

Thus $\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \beta\lambda^+(f, \mu_1)$, as required.

7.4. Proof of Theorem D part (2). We now come to the heart of the proof of Theorem D.

Step 1 (The decomposition $\mu_1 = \sum a_c \mu_{1,c}$). We decompose μ_1 into invariant measures $\mu_{1,c}$ all of whose ergodic components have nearly the same entropy.

Let $\Upsilon := \Upsilon(r)$ be as in Yomdin's Theorem 4.13, and let $\gamma_0(r, f, \eta) > 0$ be as in Proposition 5.3. We fix $\eta, \gamma > 0$ arbitrarily small, and $\ell \in \mathbb{N}$ arbitrarily large as follows. First we choose $\eta > 0$; then we take an integer ℓ such that

$$\ell\eta > \bar{h}(f, \mu_1); \tag{48}$$

and then we then fix $0 < \gamma < \min(\gamma_0(r, f, r), 1/20)$ such that for each $k \geq 1$,

$$10\gamma \left(\log(2\Upsilon) + \frac{r}{r-1} \log \|D\hat{f}_k\|_{\sup} \right) < \eta, \tag{49}$$

$$(\ell + 10)\gamma < 1, \quad \text{and} \quad \bar{h}(f, \mu_1)\ell\gamma < \eta. \quad (50)$$

By (48), we can decompose $\mu_1 = \sum_{c=1}^{\ell'} a_c \mu_{1,c}$, $a_c \in (0, 1]$, $\sum_{c=1}^{\ell'} a_c = 1$ where $\ell' \leq \ell$ and $\mu_{1,c}$ are f -invariant probability measures such that:

- For $c \neq c'$, the measures $\mu_{1,c}$ and $\mu_{1,c'}$ are mutually singular;
- For each c there is a number h_c such that all the ergodic components of $\mu_{1,c}$ have entropy in $[h_c, h_c + \eta)$;
- $h(f, \mu_1) - \eta < \sum_c a_c h_c \leq h(f, \mu_1)$.

Since $\mu_{1,c} \ll \mu_1$, Claim 7.3 implies that $\mu_{1,c}$ -a.e. x has one positive and one non-negative Lyapunov exponent. By Corollary 3.4, $\mu_{1,c}$ has an unstable lift $\hat{\mu}_{1,c}^+$ carried by $\text{graph}(E^u)$. By Claim 7.4, this gives

$$\hat{\mu}_1 = \hat{\mu}_1^+ = \sum_{c=1}^{\ell'} a_c \hat{\mu}_{1,c}^+.$$

Step 2 (The neutral segment parameters \hat{U}_0, \bar{n}_0).

- Let $N_1(r, f, \eta)$ and $N_0(r, f, \eta, \gamma)$ be as in Propositions 5.1 and 5.3.
- Fix N larger than $N_1(r, f, \eta)$ and $N_0(r, f, \eta, \gamma)$.
- Let $Q := \sup_{k \geq 1} Q_{r,N}(f_k) + 1$, with $Q_{r,N}(\cdot)$ defined as in (20). This supremum is finite because, for any $n = 1, \dots, N$, any $k \geq 1$,

$$\|\hat{f}_k^n\|_{C^{r-1}} \leq A(\|f_k^n\|_{C^r} \cdot \|Df_k^{-n}\|_{\text{sup}})^A \leq B(\|f_k\|_{C^r} \cdot \|Df_k^{-1}\|_{\text{sup}}^n)^B$$

for some $A = A(r)$ by Lemma 3.8, and $B = B(r, n)$ by the formulas for the differential of a composition. Since $f_k \xrightarrow{r\text{-bd}} f$, the factors $\|f_k\|_{C^r}$, $k \geq 1$, are bounded. Since $f_k \rightarrow f$ in $\text{Diff}^1(M)$, $\|Df_k^{-1}\|_{\text{sup}}$ converges to $\|Df^{-1}\|_{\text{sup}}$ and is therefore bounded too.

- Let $\varepsilon_Y := \varepsilon_Y(r, Q)$ be as in Yomdin's Theorem 4.13. We also set $\delta := \varepsilon_Y$.

With these choices of η, γ, δ, N , we apply Proposition 5.3 to $\hat{\mu}_0$ and obtain:

- some numbers $0 < \varepsilon, \hat{\varepsilon} < \delta$,
- an integer \bar{n}_0 ,
- a C^2 neighborhood \mathcal{U}_0 of f in $\text{Diff}^r(M)$,
- and an open set \hat{U}_0 satisfying $\hat{\mu}_0(\hat{U}_0) > 1 - \gamma^2$ and $\hat{\mu}_0(\partial\hat{U}_0) = 0$,

such that property (**) holds.

By further reducing \hat{U}_0 , we can also ensure that

$$\hat{\mu}(\partial\hat{U}_0) = 0. \quad (51)$$

Step 3 (Expanding segment parameters $\hat{U}_{1,c}, n_{1,c}$). Having fixed $N, \varepsilon, \hat{\varepsilon}$ as above, Proposition 5.1 now associates to each $\hat{\mu}_{1,c}$, with $1 \leq c \leq \ell'$, an integer $\bar{n}_{1,c} := \bar{n}_1(f, \hat{\mu}_{1,c}, \eta, \gamma, N, \varepsilon, \hat{\varepsilon})$. We then introduce the integers

$$n_1 := \max(\{\bar{n}_{1,c} : 1 \leq c \leq \ell'\} \cup \{1/\gamma\}),$$

$$n_{1,c} := n_1 + c.$$

We construct open sets $\hat{U}_{1,c}$ and an integer k_0 with the following properties:

- (a) $\hat{\mu}_{1,c}(\hat{U}_{1,c}) > 1 - \gamma^2$ and $\hat{\mu}(\partial\hat{U}_{1,c}) = 0$.
- (b) For all f_k with $k > k_0$, for any regular curve σ with C^r size at most $(\varepsilon, \hat{\varepsilon})$, there exists a family of reparametrizations \mathcal{R} of σ over $\hat{\sigma}^{-1}(\hat{U}_{1,c})$ such that
 - (b1) \mathcal{R} is $(C^r, f_k, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time $n_{1,c} := n_1 + c$,

- (b2) $\frac{1}{n_1+c} \log |\mathcal{R}| \leq h_c + \frac{\lambda(\widehat{f}_k)}{r-1} + \eta$.
(c) For all \widehat{f}_k with $k > k_0$, for any different $1 \leq c, c' \leq \ell$, and for any $0 \leq j \leq n_{1,c}$,

$$\widehat{f}_k^j(\text{closure}(\widehat{U}_{1,c})) \cap \text{closure}(\widehat{U}_{1,c'}) = \emptyset.$$

- (d) If $c \neq c'$, then $\widehat{\mu}_{1,c}(\widehat{U}_{1,c'}) < \gamma^2$.

Construction. For each c , we apply Proposition 5.1 to f , $\mu_{1,c}$ and to the parameters $\eta, \gamma, \varepsilon, \widehat{\varepsilon}, N$ and $n = n_{1,c}$. This gives an open set $\widehat{U}_{1,c}$ s.t.

- (a') $\widehat{\mu}_{1,c}(\widehat{U}_{1,c}) > 1 - \gamma^2$ and $\widehat{\mu}_{1,c}(\partial\widehat{U}_{1,c}) = 0$.
(b') For all g sufficiently close to f in C^2 -topology such that $Q_{r,N}(g) < Q$ and for any regular C^r curve σ with C^r size at most $(\varepsilon, \widehat{\varepsilon})$, there exists a family of reparametrizations \mathcal{R} of σ over $\widehat{\sigma}^{-1}(\widehat{U}_{1,c})$ satisfying (b1) and (b2).

Choose k'_0 so that (b') holds for all $g = f_k$ with $k > k'_0$, for all $1 \leq c \leq \ell$.

By assumption, the measures $\widehat{\mu}_{1,c}$ (for $1 \leq c \leq \ell'$) are mutually singular and there exist pairwise disjoint f -invariant measurable sets X_c such that $\widehat{\mu}_{1,c}(X_{c'})$ equals one when $c = c'$, and zero otherwise. Using (a'), one constructs compact sets $\widehat{K}_{1,c} \subset X_c \cap \widehat{U}_{1,c}$ such that $\widehat{\mu}_{1,c}(\widehat{K}_{1,c}) > 1 - \gamma^2$. Necessarily $\widehat{f}^j(\widehat{K}_{1,c}) \cap \widehat{K}_{1,c'} \subset X_c \cap X_{c'} = \emptyset$ for all different $1 \leq c, c' \leq \ell'$ and every $0 \leq j \leq n_{1,c}$. So

$$\min \left\{ \text{dist}(\widehat{f}^j(\widehat{K}_{1,c}), \widehat{K}_{1,c'}) : 1 \leq c, c' \leq \ell', c \neq c', j = 1, \dots, n_{1,c} \right\} > 0.$$

This inequality remains true if one replaces \widehat{f} by \widehat{f}_k with k large enough and the compact sets $\widehat{K}_{1,c}$ by small enough neighborhoods $\widehat{U}'_{1,c}$. We may choose those neighborhoods so that $\widehat{\mu}(\partial\widehat{U}'_{1,c}) = 0$. Replacing each $\widehat{U}_{1,c}$ by its intersection with $\widehat{U}'_{1,c}$, we obtain sets satisfying both the conclusion (*) of Proposition 5.1 and:

$$\widehat{f}_k^j(\widehat{U}_{1,c}) \cap \widehat{U}_{1,c'} = \emptyset \quad (1 \leq c \neq c' \leq \ell', \quad \forall 0 \leq j \leq n_{1,c})$$

for k large enough. We replace the sets $\widehat{U}_{1,c}$ by these new $\widehat{U}_{1,c}$. Moreover (a') and (b') are preserved.

Notice that if $c \neq c'$, then $\widehat{U}_{1,c'} \cap \widehat{U}_{1,c} = \emptyset$, so $\widehat{\mu}_{1,c}(\widehat{U}_{1,c'}) < \widehat{\mu}_{1,c}(\widehat{M} \setminus \widehat{K}_{1,c}) < \gamma^2$. Then, all the properties (a)-(d) hold.

Step 4 (weak-* neighborhoods of \widehat{m}_0 and \widehat{m}_1). We construct weak-* open neighborhoods $\widehat{V}_0, \widehat{V}_1$ of the measures $\widehat{m}_0, \widehat{m}_1$ s.t.

$$\begin{aligned} \widehat{m} \in \widehat{V}_0 &\Rightarrow \begin{cases} |\widehat{m}(\widehat{U}_0) - \widehat{m}_0(\widehat{U}_0)| < \gamma^2, \\ |\widehat{m}(\widehat{M}) - \widehat{m}_0(\widehat{M})| < \gamma^2, \end{cases} \\ \widehat{m} \in \widehat{V}_1 &\Rightarrow \begin{cases} |\widehat{m}(\widehat{U}_{1,c}) - \widehat{m}_1(\widehat{U}_{1,c})| < \gamma^2, \quad 1 \leq c \leq \ell', \\ |\widehat{m}(\widehat{M}) - \widehat{m}_1(\widehat{M})| < \gamma^2. \end{cases} \end{aligned} \quad (52)$$

Such neighborhoods exist since $\widehat{m}_i(\partial\widehat{U}_{1,c}) = 0$ and $\widehat{m}_i(\partial\widehat{U}_0) = 0$ from the property (a) in step 3 and (51).

Step 5 (Neutral block parameters α, L, k_*). Recall the integer k_0 obtained in step 3. Using property (iii) in §7.2, one finds $\alpha \in (0, \eta/10)$, $L \geq 1$ satisfying

$$L > 2\gamma^{-1} \max\{\overline{n}_0, n_{1,1}, \dots, n_{1,\ell'}\}, \quad (53)$$

and $k_* = k_*(\alpha, L) \geq k_0$ such that for $k > k_*$ and $\widehat{\nu}_k^+$ -a.e. $\widehat{x} \in \widehat{M}$,

- the $\mathfrak{N}_{\alpha,L}(\widehat{x}, \widehat{f}_k, \varphi_k)$ -empirical measures are in a compact subset of \widehat{V}_0 ;
- the $\mathbb{N} \setminus \mathfrak{N}_{\alpha,L}(\widehat{x}, \widehat{f}_k, \varphi_k)$ -empirical measures are in a compact subset of \widehat{V}_1 .

These compact sets will give the extra margin necessary to deal with boundary terms (see the proof of Lemma 7.5 below).

Step 6 (Decomposition of orbits into orbit segments). Recall that the orbit segment of \widehat{f}_k with length t and initial point \widehat{x} is the string $(\widehat{x}, \widehat{f}_k(\widehat{x}), \dots, \widehat{f}_k^{t-1}(\widehat{x}))$. It is associated with the measure

$$\widehat{\mu}_{\widehat{f}_k, \widehat{x}}^t := \frac{1}{t} \sum_{j=0}^{t-1} \delta_{\widehat{f}_k^j(\widehat{x})}.$$

An orbit segment will be called *neutral* if $(0, 1, \dots, t-1)$ is an (α, L) -neutral block of $(\widehat{x}, \widehat{f}_k, \varphi_k)$ as defined in Section 6, i.e. if $t \geq L$ and if $\widehat{x} = (x, E)$ satisfies:

$$\|Df_k^m|_E\| \leq e^{\alpha m} \quad \text{for all } 0 < m \leq t. \quad (54)$$

Using the open sets $\widehat{U}_0, \widehat{U}_{1,c}$ and the integers $n_{1,c}$ defined at steps 2 and 3, we introduce $\ell' + 2$ classes of orbit segments $(\widehat{x}, \widehat{f}_k(\widehat{x}), \dots, \widehat{f}_k^{t-1}(\widehat{x}))$:

- (a) **Segments with color** $1 \leq c \leq \ell'$: orbit segments such that $\widehat{x} \in \widehat{U}_{1,c}$ and $t = n_{1,c}$.
- (b) **Blank segments**: neutral orbit segments such that $\widehat{\mu}_{\widehat{f}_k, \widehat{x}}^t(\widehat{U}_0) \geq 1 - \gamma$.
- (c) **Fillers**: orbit segments with length $t = 1$.

The class of an orbit segment as above can be recognized from its length t : If $t = 1$, it is a filler, if $t \in [n_1 + 1, n_1 + \ell']$, it is colored with color $t - n_1$, and if t is larger than L , then it is blank, see (53). So these $\ell' + 2$ classes are disjoint.

Lemma 7.5. *For all $k > k_*$ and for $\widehat{\nu}_k^+$ -a.e. \widehat{x} , there exists $n_k(\widehat{x}) \in \mathbb{N}$ such that all the orbit segments $(\widehat{x}, \widehat{f}_k(\widehat{x}), \dots, \widehat{f}_k^{n-1}(\widehat{x}))$ with $n \geq n_k(\widehat{x})$ can be decomposed into:*

- (a) *colored segments of total length at most $\beta a_c n + \gamma n$, for each color c ,*
- (b) *blank segments of total length at least $(1 - \beta)n - 4\gamma n$,*
- (c) *fillers of total length at most $6\gamma n$.*

Proof. By the reduction in section 7.1, the ergodic measures ν_k have positive entropy, and therefore the $\widehat{\nu}_k$ measure of \widehat{f}_k -periodic points is zero. Thus it is sufficient to consider non-periodic \widehat{x} only. Orbit segments of non-periodic points can be identified with the non-ordered sets of points they contain without any loss of information, because there is only one way to order them to get an orbit segment. We will therefore feel free to abuse terminology and treat orbit segments as sets, subject to the usual set-theoretic operations.

Given an orbit segment $\vartheta := (\widehat{x}, \dots, \widehat{f}_k^{n-1}(\widehat{x}))$, we are going to build a decomposition

$$\underbrace{(\widehat{f}_k^{t_0}(\widehat{x}), \dots, \widehat{f}_k^{t_1-1}(\widehat{x}))}_{\vartheta_1}; \dots; \underbrace{(\widehat{f}_k^{t_{m-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_m-1}(\widehat{x}))}_{\vartheta_m} \quad (55)$$

where $t_0 = 0 < t_1 < \dots < t_m = n$ and each segment $\vartheta_i := (\widehat{f}_k^{t_{i-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_i-1}(\widehat{x}))$ is either a colored segment, a blank segment, or a filler.

We call the sequence (t_0, \dots, t_m) the *type of the decomposition* since it determines not only how the orbit segment is divided but to which class each segment belong.

By analogy with Section 6, a neutral sub-segment of ϑ is called *maximal*, if it does not lie in a strictly longer neutral sub-segment of ϑ . Let $\mathcal{S}_{\text{neut}}(\vartheta)$ denote the collection of all maximal neutral sub-segments of ϑ . It is not difficult to see that every neutral sub-segment

of ϑ is contained in some element of $\mathcal{S}_{neut}(\vartheta)$, and that the segments in $\mathcal{S}_{neut}(\vartheta)$ are pairwise disjoint.

DECOMPOSITION. We define t_i inductively beginning with $t_0 := 0$. Assuming that $0 \leq t_{i-1} < n$ has been defined, we consider the following three possibilities:

- *Case (a).* There exists $1 \leq c \leq \ell'$ such that $\widehat{f}^{t_{i-1}}(\widehat{x}) \in \widehat{U}_{1,c}$, the orbit segment $(\widehat{f}_k^{t_{i-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_{i-1}+n_{1,c}-1}(\widehat{x}))$ does not intersect any segment in $\mathcal{S}_{neut}(\vartheta)$, and $t_{i-1} + n_{1,c} \leq n$. The color c is uniquely defined because the $\widehat{U}_{1,c}$ are disjoint. We set $t_i := t_{i-1} + n_{1,c}$. The resulting orbit segment $\vartheta_i := (\widehat{f}_k^{t_{i-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_i-1}(\widehat{x}))$ with length $n_{1,c}$ is a segment with color c .
- *Case (b).* There exists T such that $\vartheta_i := (\widehat{f}_k^{t_{i-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_{i-1}+T-1}(\widehat{x})) \in \mathcal{S}_{neut}(\vartheta)$, and $|\vartheta_i \cap \widehat{U}_0| > (1 - \gamma)|\vartheta_i|$. The integer T is unique by maximality and, by the definition of $\mathcal{S}_{neut}(\vartheta)$, it satisfies $T \geq L$ and $t_{i-1} + T \leq n$. We set $t_i := t_{i-1} + T$. Then $\vartheta_i = (\widehat{f}_k^{t_{i-1}}(\widehat{x}), \dots, \widehat{f}_k^{t_i-1}(\widehat{x}))$, and ϑ_i is a blank segment with length T .
- *Case (c).* There are no such T or c . In this case we set $t_i := t_{i-1} + 1$, and $\vartheta_i := (\widehat{f}_k^{t_{i-1}}(\widehat{x}))$. This is a filler.

These cases are mutually exclusive and at least one of them must happen (case (b) implies that $\vartheta_i \in \mathcal{S}_{neut}(\vartheta)$, excluding case (a), and case (c) happens iff case (a) and case (b) both fail), and in all cases, $t_i \leq n$ since $t_{i-1} < n$. Thus we have defined $t_{i+1} \in (t_i, n]$ unambiguously.

The inductive process stops with $t_m = n$. The result is a decomposition of ϑ as in eq. (55) into blank segments, colored segments, and fillers.

SIZE ESTIMATES. We now fix a $\widehat{\nu}_k$ -typical \widehat{x} , a large n , set $\vartheta := (\widehat{x}, \dots, \widehat{f}^{n-1}(\widehat{x}))$, and estimate the total size of the fillers, blank segments, and the segments of given color in ϑ . “Typical” means that our estimates apply to a set of full $\widehat{\nu}_k$ -measure, and the “largeness” of n is allowed to depend on \widehat{x} .

Let Neut be the union of all neutral sub-segments of ϑ and Neut^c its complement:

$$\text{Neut} := \bigcup \mathcal{S}_{neut}(\vartheta), \quad \text{Neut}^c := \vartheta \setminus \text{Neut}.$$

Clearly $\text{Neut} \subseteq \mathfrak{N}_{\alpha,L}(\widehat{x}, \widehat{f}_k, \varphi_k) \cap [0, n)$, but the sets could be different, because the neutral segments in $\mathfrak{N}_{\alpha,L}(\widehat{x}, \widehat{f}_k, \varphi_k)$ which contains 0 or n may have a non- (α, L) -neutral intersection with $[0, n)$. However, it is not difficult to see that $\frac{1}{n} |\text{Neut} \Delta (\mathfrak{N}_{\alpha,L}(\widehat{x}, \widehat{f}_k, \varphi_k) \cap [0, n))| \xrightarrow{n \rightarrow \infty} 0$. Therefore, for $\widehat{\nu}_k$ -a.e. \widehat{x} there exists $n_k(\widehat{x})$ such that for all $n > n_k(\widehat{x})$ and $k > k_*$ (cf. step 5)

$$\widehat{m}'_0 := \frac{1}{n} \sum_{\widehat{y} \in \text{Neut}} \delta_{\widehat{y}} \in \widehat{V}_0, \quad \widehat{m}'_1 := \frac{1}{n} \sum_{\widehat{y} \in \text{Neut}^c} \delta_{\widehat{y}} \in \widehat{V}_1.$$

Recall that $\widehat{m}_1 = \beta \sum_c a_c \widehat{\mu}_{1,c}$ with $\sum_c a_c = 1$ and $0 < \beta \leq 1$. Since $\widehat{m}'_1 \in \widehat{V}_1$, eq. (52) and items (a) and (c) of Step 3 imply:

$$\begin{aligned} \widehat{m}'_1(\widehat{M} \setminus \bigcup_{c'} \widehat{U}_{1,c'}) &= \widehat{m}'_1(\widehat{M}) - \sum_{c'} \widehat{m}'_1(\widehat{U}_{1,c'}) < \widehat{m}_1(\widehat{M}) + \gamma^2 - \sum_{c'} (\widehat{m}_1(\widehat{U}_{1,c'}) - \gamma^2) \\ &\leq \beta \left(1 - \sum_{c'} a_{c'} (1 - \gamma^2) \right) + (1 + \ell') \gamma^2 \\ &\leq (\ell + 2) \gamma^2, \text{ because the number of colors } \ell' \text{ is at most } \ell. \end{aligned} \tag{56}$$

Let C_c denote the union of all colored segments with color c ; let B denote the union of all blank segments; and let F denote the union of all fillers.

(a) *Colored segments.* By construction, if $c \neq c'$, then $\bigcup_{j=0}^{n_{1,c}} \widehat{f}_k^j(\widehat{U}_{1,c}) \cap \widehat{U}_{1,c'} = \emptyset$. So if ϑ_i has color c , then $\vartheta_i \subset \text{Neut}^c \setminus \bigcup_{c' \neq c} \widehat{U}_{1,c'}$, and

$$|C_c| \leq |\text{Neut}^c \setminus \bigcup_{c' \neq c} \widehat{U}_{1,c'}| \leq |\text{Neut}^c \setminus \bigcup_{c'} \widehat{U}_{1,c'}| + |\text{Neut}^c \cap \widehat{U}_{1,c}| \\ \leq n \cdot \widehat{m}'_1(\widehat{M} \setminus \bigcup_{c'} \widehat{U}_{1,c'}) + n \cdot \widehat{m}'_1(\widehat{U}_{1,c}) < n(\ell + 2)\gamma^2 + n[\widehat{m}_1(\widehat{U}_{1,c}) + \gamma^2],$$

by eqs. (52) and (56). By step 3 (d), $\widehat{m}_1(\widehat{U}_{1,c}) < \beta(a_c + \gamma^2)$. Substituting this in the above and using (50) give

$$|C_c| \leq n\beta a_c + n(\ell + 4)\gamma^2 \leq n\beta a_c + n\gamma.$$

(b) *Blank segments.* By definition, every blank segment is neutral, so $B \subset \text{Neut}$, and $|B| = |\text{Neut}| - |\text{Neut} \setminus B|$. By (52) and the definition of \widehat{m}'_0 ,

$$|\text{Neut}| = n \cdot \widehat{m}'_0(\widehat{M}) \geq n(\widehat{m}_0(\widehat{M}) - \gamma^2) = n(1 - \beta - \gamma^2).$$

The set $\text{Neut} \setminus B$ is the union of the maximal neutral orbit segments which visit \widehat{U}_0 with frequency less than $1 - \gamma$. Thus $\gamma \cdot |\text{Neut} \setminus B| < \widehat{m}'_0(\widehat{M} \setminus \widehat{U}_0)n$. By (52) and the bound $\widehat{\mu}_0(\widehat{U}_0) > 1 - \gamma^2$ in step 2,

$$\widehat{m}'_0(\widehat{M} \setminus \widehat{U}_0) < \widehat{m}_0(\widehat{M}) - \widehat{m}_0(\widehat{U}_0) + 2\gamma^2 < (1 - \beta) - (1 - \beta)(1 - \gamma^2) + 2\gamma^2 = \gamma^2(3 - \beta),$$

so $|\text{Neut} \setminus B| \leq \gamma^{-1} \widehat{m}'_0(\widehat{M} \setminus \widehat{U}_0)n < \gamma(3 - \beta)n$. It follows that $|B| > (1 - \beta)n - 4\gamma n$.

(c) *Fillers.* By construction, a filler is a segment of length one (\widehat{y}) such that one of the following holds:

- (i) \widehat{y} does not belong to a colored segment or to a segment in $\mathcal{S}_{neut}(\vartheta)$;
- (ii) \widehat{y} belongs to a segment in $\mathcal{S}_{neut}(\vartheta)$, but this segment is not a blank segment;
- (iii) \widehat{y} belongs to a segment of length $n_{1,c}$ which begins at $\widehat{U}_{1,c}$, but it fails to be a colored segment because it extends beyond the right endpoint of ϑ ;
- (iv) \widehat{y} belongs to a segment of length $n_{1,c}$ which begins at $\widehat{U}_{1,c}$, but it fails to be a colored segment because it intersects an element of $\mathcal{S}_{neut}(\vartheta)$.

The fillers of type (i) belong to $\text{Neut}^c \setminus \bigcup_c \widehat{U}_{1,c}$, so their cardinality is bounded by eq. (56): $|\text{Neut}^c \setminus \bigcup_c \widehat{U}_{1,c}| = n \cdot \widehat{m}'_1(\widehat{M} \setminus \bigcup_c \widehat{U}_{1,c}) < (2 + \ell)\gamma^2 n < \gamma n$

The fillers of type (ii) belong to $\text{Neut} \setminus B$. As we saw above this means that their cardinality is less than $\gamma(3 - \beta)n < 3\gamma n$.

The number of fillers of type (iii) is clearly bounded by the maximum length of a colored segment $\max_c n_{1,c} = n_1 + \ell' \leq n_1 + \ell$. This can be assumed to be less than γn when n is large enough.

It remains to control the fillers of type (iv). Fix $\vartheta_0 \in \mathcal{S}_{neut}(\vartheta)$, and suppose \widehat{y} belongs to a colored segment which intersects ϑ_0 . All colored segments have lengths at most $n_1 + \ell'$, therefore \widehat{y} must belong to one of two segments of length $n_1 + \ell'$ adjacent to the endpoints of ϑ_0 . This gives the following bound for the number of fillers of type (iv): $2(n_1 + \ell') \cdot |\mathcal{S}_{neut}(\vartheta)|$. Recalling that $\mathcal{S}_{neut}(\vartheta)$ consists of disjoint sub-segments of ϑ , each with length at least L , we find that

$$|\mathcal{S}_{neut}(\vartheta)| \leq \frac{n}{L}.$$

Thus by (53), the number of fillers of type (iv) is at most $\frac{2(n_1 + \ell')}{L} n < \gamma n$.

It follows that the total length of the fillers is $|\mathbf{F}| < 6\gamma n$. \square

Step 7 (A bound on the number of decomposition types). In the previous step we decomposed orbit segments of typical points with length n large enough into colored segments, blank segments and fillers.

Let $\theta = (t_0, t_1, \dots, t_m)$ be the type of the decomposition, see (55) and the discussion which follows it. Here we bound the number of possible types. As always, let $H(t) := t \log \frac{1}{t} + (1-t) \log \frac{1}{1-t}$ for $0 < t < 1$.

Claim 7.6. *There exists $n_H := n_H(\gamma)$ such that the number of types of decompositions of all \hat{f}_k -orbit segments as in Lemma 7.5 with length $n > n_H$ and k arbitrary is at most $\exp[nH(10\gamma)]$.*

Proof. A decomposition of an orbit segment with length n has

- at most γn blank segments (because these have lengths $\geq L > 1/\gamma$),
- at most γn colored segments (because these have lengths $\geq n_1 + 1 > 1/\gamma$),
- and at most $6\gamma n$ fillers (by Lemma 7.5).

This gives a total of at most $\lfloor 8\gamma n \rfloor$ segments.

So every type $\theta = (t_0, \dots, t_m)$ has length $m < \lfloor 8\gamma n \rfloor + 1$. Since $t_0 = 0, t_m = n$, there can be at most $\sum_{m=1}^{\lfloor 8\gamma n \rfloor} \binom{n}{m-1}$ different types. Since $8\gamma < 1/2$, the sum is bounded by $8\gamma n \binom{n}{\lfloor 8\gamma n \rfloor}$. By De Moivre's approximation, this is less than $\exp[nH(8\gamma) + o(n)]$ as $n \rightarrow \infty$.

The claim follows, because $10\gamma < 1/2$ so $H(8\gamma) < H(10\gamma)$. \square

Step 8 (Conditional measures and choice of N_k, F_k). The measures ν_k are assumed to be f_k -ergodic, and by the reductions in section 7.1 they have positive entropy. So by Ruelle's inequality, each ν_k is a hyperbolic measure.

As explained in Section 4.7, one can introduce a measurable partition subordinated to the unstable lamination of ν_k and associate to it a system of conditionals probability measures $\nu_{k,x}^u$.

We fix $N_k \geq 1$ and a Borel set $F_k \subset M$ with $\nu_k(F_k) > \frac{1}{2}$ such that for every point $x \in F_k$ and for the diffeomorphism f_k :

- x has a well-defined unstable manifold, an immersed C^r curve $W^u(x) \subset M$;
- $\nu_{k,x}^u$ is well-defined and x belongs to the support of the restriction of $\nu_{k,x}^u$ to F_k ;
- $\hat{x} := (x, E^u(x))$ satisfies Lemma 7.5 with $n_k(\hat{x}) \leq N_k$. In particular for each $n \geq N_k$, the orbit segment $(\hat{x}, \hat{f}_k(\hat{x}), \dots, \hat{f}_k^{n-1}(\hat{x}))$ has a decomposition as in Lemma 7.5. Let $\theta = \theta(x, n)$ be the type of decomposition.

Step 9 (Construction of reparametrizations). Choose a point $x \in F_k$ which satisfies Corollary 4.16.

Let $\sigma: [0, 1] \rightarrow W^u(x)$ be a regular C^r -curve which parametrizes a neighborhood of x in $W^u(x)$ in the intrinsic topology, and which has C^r size at most $(\varepsilon, \hat{\varepsilon})$. By the choice of F_k , $T := \sigma^{-1}(F_k)$ has positive measure for $\nu_{k,x}^u$.

Fix $n \geq N_k$, and let $\varepsilon, \hat{\varepsilon}$ and N be as in step 2. Our aim is to construct a particular family of reparametrizations \mathcal{R}_n of σ over T , which is $(C^r, f_k, N, \varepsilon, \hat{\varepsilon})$ -admissible up to time n . In later steps, we will estimate the cardinality of \mathcal{R}_n and use Corollary 4.16 to obtain the upper bound for $h(f_k, \nu_k)$ which completes the proof of the theorem.

We begin by fixing a type $\theta := (t_0, t_1, \dots, t_m)$ with $t_m = n$, and constructing a family of reparametrizations \mathcal{R}_n^θ of σ admissible up to time n over the set

$$T_\theta := \sigma^{-1}\{y \in F_k \text{ with type } \theta\}.$$

Then we will take the union over all possible types and obtain the family \mathcal{R}_n of reparametrizations over T .

\mathcal{R}_n^θ is obtained inductively by defining families $\mathcal{R}_{t_i}^\theta$ of parametrizations of σ over T_θ , which are $(C^r, f_k, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time t_i . The base of the induction is defined by taking $\mathcal{R}_0^\theta := \{\text{Id}\}$. This parametrization is admissible, because σ has C^r size at most $(\varepsilon, \widehat{\varepsilon})$. After m steps, we will obtain the family $\mathcal{R}_n^\theta = \mathcal{R}_{\theta, t_m}$, which is an admissible family of reparametrizations over T_θ , up to time n .

INDUCTION STEP: We build $\mathcal{R}_{t_i}^\theta$, assuming $\mathcal{R}_{t_{i-1}}^\theta$ was already constructed. We proceed by concatenation (see Lemma 4.10). We will set

$$\mathcal{R}_{t_{i-1}}^\theta := \{\psi \circ \varphi : \psi \in \mathcal{R}_{t_{i-1}}^\theta, \varphi \in \mathcal{R}(\psi, t_{i-1})\}$$

for well-chosen families $\mathcal{R}(\psi, t_{i-1})$ which parametrize $f_k^{t_{i-1}} \circ \sigma \circ \psi$ over $\psi^{-1}(T_\theta)$ in a $(C^r, f_k, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time $t_i - t_{i-1}$, and which we now construct.

Fix $\psi \in \mathcal{R}_{t_{i-1}}^\theta$ and let $\sigma' := f_k^{t_{i-1}} \circ \sigma \circ \psi$ and $T' := \psi^{-1}(T_\theta)$. By the induction hypothesis, $\mathcal{R}_{t_{i-1}}^\theta$ is admissible, therefore σ' has C^r size at most $(\varepsilon, \widehat{\varepsilon})$.

By the definition of T_θ , the orbit segments $(\widehat{y}, \widehat{f}_k(\widehat{y}), \dots, \widehat{f}_k^{t_i - t_{i-1}}(\widehat{y}))$ have the same type for every $\widehat{y} \in \widehat{\sigma}'(T')$: If $t_i - t_{i-1} = n_{1,c}$ they are all colored segments with color c ; if $t_i - t_{i-1} \geq L$ they are all blank; and if $t_i - t_{i-1} = 1$ they are all fillers. See step 6. Our construction of $\mathcal{R}(\psi, t_{i-1})$ depends on the case:

Case (a): $t_i - t_{i-1} = n_{1,c}$. In this case, $(\widehat{f}_k^{t_{i-1}}(\widehat{y}), \dots, \widehat{f}_k^{t_i - 1}(\widehat{y}))$ are colored segments with the same color c for all $\widehat{y} \in \widehat{\sigma}(T_\theta)$. Thus $\widehat{f}_k^{t_{i-1}}(\widehat{\sigma}(T_\theta)) \subset \widehat{U}_{1,c}$. Applying Proposition 5.1 to $\widehat{\mu}_{1,c}, n_{1,c}$ and σ' , we obtain a family $\mathcal{R}(\psi, t_{i-1})$ of reparametrizations φ over the set T' which is $(C^r, f_k, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time $t_i - t_{i-1}$, and which satisfies the cardinality bound

$$\frac{1}{t_i - t_{i-1}} \log |\mathcal{R}(\psi, t_{i-1})| \leq \bar{h}(f, \mu_{1,c}) + \frac{\lambda(\widehat{f})}{r-1} + \eta \leq h_c + \frac{\lambda(\widehat{f})}{r-1} + 2\eta$$

(Recall that the entropy of every ergodic component of $\mu_{1,c}$ is in $[h_c, h_c + \eta)$.)

Case (b): $t_i - t_{i-1} \geq L$. In this case $(\widehat{f}_k^{t_{i-1}}(\widehat{y}), \dots, \widehat{f}_k^{t_i - 1}(\widehat{y}))$ are blank for all $\widehat{y} \in \widehat{\sigma}(T_\theta)$. By the choice of α, L in Step 5, these segments are $\frac{\eta}{10}$ -neutral, and their lengths are larger or equal to \bar{n}_0 . Consequently, the set T' is contained in the set controlled by Proposition 5.3, when applied to the diffeomorphism $g = f_k$ and to the curve σ' (see eq. (21)). Hence, this proposition gives us a family $\mathcal{R}(\psi, t_{i-1})$ of reparametrizations φ of σ' over the set T' which is $(C^r, f_k, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time $t_i - t_{i-1}$, and which satisfies the cardinality bound

$$\frac{1}{t_i - t_{i-1}} \log |\mathcal{R}(\psi, t_{i-1})| \leq \frac{\lambda(\widehat{f})}{r-1} + \eta.$$

Case (c): $t_i - t_{i-1} = 1$. As σ' has C^r size at most $(\varepsilon, \widehat{\varepsilon})$,

$$r_{\widehat{f}_k}(\varepsilon, \widehat{\varepsilon}, 1, \widehat{f}_k(\widehat{\sigma}(T'))) \leq \|D\widehat{f}_k\|_{\sup} + 1 \leq 2\|D\widehat{f}_k\|_{\sup}.$$

Since $\varepsilon, \widehat{\varepsilon}$ have been chosen smaller than ε_Y , Corollary 4.14 of Yomdin's Theorem applies and provides a family $\mathcal{R}(\psi, t_i)$ of reparametrizations φ over the set T_θ which are $(C^r, f_k, N, \varepsilon, \widehat{\varepsilon})$ -admissible up to time $t_i - t_{i-1} = 1$, and which satisfy the cardinality bound $|\mathcal{R}(\psi, t_{i-1})| \leq \Upsilon \|D\widehat{f}_k\|_{\sup}^{1/(r-1)} \times 2\|D\widehat{f}_k\|_{\sup}$, hence

$$\frac{1}{t_i - t_{i-1}} \log |\mathcal{R}(\psi, t_{i-1})| \leq \log(2\Upsilon) + \frac{r}{r-1} \log \|D\widehat{f}_k\|_{\sup}.$$

This completes the inductive step.

Step 10 (Cardinality of \mathcal{R}_n). The families of reparametrizations obtained in step 9 satisfy $|\mathcal{R}(\psi, t_i)| \leq \exp(\kappa_i(\theta)(t_i - t_{i-1}))$, where

$$\kappa_i(\theta) := \begin{cases} h_c + (r-1)^{-1}\lambda(\widehat{f}) + 2\eta & \text{if } t_i - t_{i-1} = n_{1,c}, \quad (\text{case a}), \\ (r-1)^{-1}\lambda(\widehat{f}) + \eta & \text{if } t_i - t_{i-1} \geq L, \quad (\text{case b}), \\ \log(2\Upsilon) + \frac{r}{r-1} \log \|D\widehat{f}_k\|_{\sup} & \text{if } t_i - t_{i-1} = 1, \quad (\text{case c}). \end{cases}$$

It follows that $|\mathcal{R}_n^\theta| \leq \exp\left(\sum_{i=1}^m \kappa_i(\theta)(t_i - t_{i-1})\right)$.

The total length of the blank segments is (trivially) less than n , and the total lengths of the segments with color c and fillers is respectively, less than $\beta a_c n + \gamma n$ and $6\gamma n$, by Lemma 7.5. Denoting the total length of colored segments with color c (resp. blank segments, fillers) by N_c (resp. N_b , N_f), we find that

$$\begin{aligned} \sum_{i=1}^m \kappa_i(\theta)(t_i - t_{i-1}) &\leq \sum_c h_c N_c + \frac{\lambda(\widehat{f})}{r-1} \left(\sum_c N_c + N_b \right) + 2\eta \sum_c N_c + \eta N_b \\ &\quad + \left(\log(2\Upsilon) + \frac{r}{r-1} \log \|D\widehat{f}_k\|_{\sup} \right) N_f. \end{aligned}$$

Using the trivial bounds $\sum_c N_c + N_b = n - N_f \leq n$, $N_b \leq n$, and the bounds $N_c \leq \beta a_c n + \gamma n$, $N_f \leq 6\gamma n$ from Lemma 7.5, we find that

$$\begin{aligned} \sum_{i=1}^m \kappa_i(\theta)(t_i - t_{i-1}) &\leq n \sum_c h_c [\beta a_c + \gamma] + n \frac{\lambda(\widehat{f})}{r-1} + 2\eta n \\ &\quad + 6\gamma n \left(\log(2\Upsilon) + \frac{r}{r-1} \log \|D\widehat{f}_k\|_{\sup} \right). \end{aligned}$$

Recall that we chose a_c, h_c, ℓ and γ so that

- $\sum_c a_c h_c \leq h(f, \mu_1)$ by the choice of the decomposition of μ_1 in Step 1;
- $\gamma \sum_c h_c \leq \ell \gamma \max\{h_c\} < \eta$ by eq. (50);
- $6\gamma(\log(2\Upsilon) + \frac{r}{r-1} \log \|D\widehat{f}_k\|_{\sup}) < \eta$ by eq. (49).

Hence, $|\mathcal{R}_n^\theta| \leq \exp\left(\beta h(f, \mu_1)n + \frac{\lambda(\widehat{f})}{r-1}n + 4\eta n\right)$.

Recalling that $\mathcal{R}_n := \bigcup_\theta \mathcal{R}_n(\theta)$ and the number of types is bounded for all n large enough by $\exp[nH(10\gamma)]$, we conclude that

$$|\mathcal{R}_n| \leq \exp\left(\beta h(f, \mu_1)n + \frac{\lambda(\widehat{f})}{r-1}n + 4\eta n + H(10\gamma)n\right).$$

Step 11 (Completion of the proof). By Corollary 4.16, for all k large enough

$$h(f_k, \nu_k) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{R}_n| \leq \beta h(f, \mu_1) + \frac{\lambda(\widehat{f})}{r-1} + (4\eta + H(10\gamma)).$$

Passing to the limits in the order $k \rightarrow \infty$, $\gamma \rightarrow 0$, $\eta \rightarrow 0$ gives the second part of Theorem D, and completes its proof. \square

8. SUPPLEMENTS

We prove here the additional properties mentioned in Section 1.4.

8.1. Discontinuities: construction of the Example 1.2. Let us recall that two transitive hyperbolic sets K_1, K_2 are *homoclinically related* if a stable manifold of K_1 has a transverse intersection point with an unstable manifold of K_2 and a stable manifold of K_2 has a transverse intersection point with an unstable manifold of K_1 . In this case there exists a transitive hyperbolic set that contains K_1 and K_2 .

If O is a periodic orbit, we will denote by μ_O the invariant probability measure supported on O . We say that a sequence of periodic orbits (O_k) converges weak-* to a measure μ , if the sequence of measures (μ_{O_k}) converges weak-* towards μ .

We say that a C^∞ diffeomorphism f_0 belongs to the *Newhouse domain* if there exist an attracting region U where $|\det Df_0| < 1$, a transitive hyperbolic, locally maximal set $K \subset U$ (not reduced to a periodic orbit) and a C^∞ neighborhood \mathcal{U} of f_0 such that for any diffeomorphism $f \in \mathcal{U}$ the hyperbolic continuation of K (still denoted by K) admits a stable manifold and an unstable manifold with a non-transverse intersection. The Newhouse domain is open by definition, and non-empty by [34].

We prove the following more precise version of Example 1.2:

Proposition 8.1. *The Newhouse domain in $\text{Diff}^\infty(M)$ contains a dense G_δ subset of diffeomorphisms f with the following property. For any pair of numbers $0 < \alpha \leq \beta \leq 1$, there is a sequence of ergodic measures (ν_k) converging weak-* to a measure μ with $h(f, \mu) > 0$ and such that:*

$$\lim h(f, \nu_k) = \alpha h(f, \mu) \text{ and } \lim \lambda^+(f, \nu_k) = \beta \lambda^+(f, \mu).$$

Remark 8.2. One can choose for μ any invariant probability measure with positive entropy, ergodic or not, and carried by the hyperbolic set K associated with the Newhouse domain of f .

Lemma 8.3. *There is a dense G_δ subset of the Newhouse domain in $\text{Diff}^\infty(M)$, made of diffeomorphisms f with the following property. For any periodic orbit P contained in the hyperbolic set K associated to f , there exists a sequence of hyperbolic periodic orbits O_k homoclinically related to P which converge weak-* towards P and satisfy $\lambda^+(O_k) \rightarrow 0$.*

Proof. Let $f_0 \in \mathcal{U}$. By an application of Baire's argument, it is enough to find $f \in C^\infty$ close to f_0 with a periodic orbit O homoclinically related to P which is weak-* close to P and has a top Lyapunov exponent close to 0. We sketch the proof which uses classical arguments on the behavior near homoclinic tangencies, and we refer to [36] for further details. In order to simplify the presentation, we assume that P is fixed and the eigenvalues $0 < \lambda < 1 < \mu$ of $Df(P)$ are positive. By dissipation, $\lambda \cdot \mu < 1$.

Since the stable (resp. unstable) manifold of P is dense in the stable (resp. unstable) lamination of K , and since f_0 belongs to the Newhouse domain, one can perturb f_0 in such a way that P exhibits a quadratic homoclinic tangency $z \in W_{loc}^s(P)$. One can also assume that the eigenvalues λ, μ are non-resonant, so that by Sternberg's theorem, there exists a smooth chart on a neighborhood $U \simeq [-1, 1]^2$ of P , where the dynamics is linear: On $[-1, 1] \times [-\mu^{-1}, \mu^{-1}]$, f coincides with the map $\mathcal{L}: (x, y) \mapsto (\lambda \cdot x, \mu \cdot y)$. The local manifolds $W_{loc}^s(P)$ and $W_{loc}^u(P)$ coincide with $\{y = 0\}$ and $\{x = 0\}$. Moreover z has a preimage $z' \in W_{loc}^u(P)$ by an iterate f^N and one denotes by \mathcal{T} the map induced by f^N from a neighborhood of z' to z . The unstable manifold at z is locally a graph $\{(x, \varphi(x))\}$ and by a suitable rescaling of the axis of U , one can require that $D^2\varphi \simeq 1$ near z .

Let us fix $\delta > 0$ small. When n is large one considers a rectangle

$$R = z + [-\delta, \delta] \times [a - \delta\mu^{-n}, a + \delta\mu^{-n}],$$

where a is chosen such that $z' = (0, a \cdot \mu^n)$. Note that $C^{-1} \leq a \cdot \mu^n \leq C$ where C depends on the Sternberg linearization domain U , but not on n .

The rectangle R is mapped by f^{n+N} to a thin curved rectangle $\mathcal{T} \circ \mathcal{L}^n(R)$ whose width is of the order of $\delta\lambda^n$, hence much smaller than the width of R . One perturbs f near z' in such a way that the transition map \mathcal{T} is composed with a vertical translation. The tip of the image can thus be adjusted to be at distance $L \cdot a$ from the rectangle R where L is a large constant independent from n . Therefore $f^{n+N}(R)$ crosses $\{y = 0\}$ and also R with a slope s close to $L \cdot a$ (since $D^2\varphi \simeq 1$). See Figure 1.

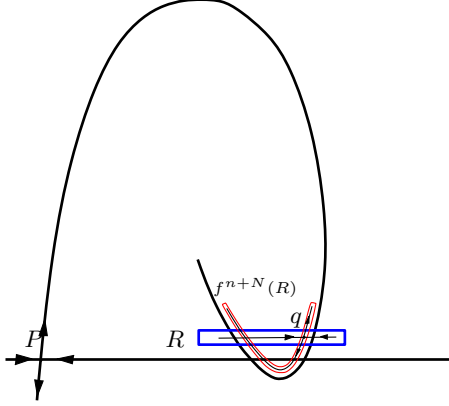


FIGURE 1. Return map near an homoclinic tangency.

Moreover $R \cap f^{n+N}(R)$ contains a periodic point q whose unstable direction is dilated at the period by a factor of the order of $s \exp(-\lambda^+ n) \simeq L \cdot a \cdot \mu^n$, which is close to a large constant (comparable to L). As the period $n + N$ of q can be chosen arbitrarily large, the unstable Lyapunov exponent of q is close to 0.

Note that the unstable manifold of q crosses $f^n(R)$ along its largest dimension (see Figure 1), hence crosses $W_{loc}^s(P)$. The local stable manifold of q is a graph which crosses R horizontally. The image $f^n(W_{loc}^s(P))$ is close to $f^n(R)$, crosses R , and then the local stable manifold of Q . Hence P and the orbit of q are homoclinically related. As the n first iterates of q belong to the linearization domain U , the orbit of q spends an arbitrarily large proportion of time in any neighborhood of P , as the period $n + N$ goes to infinity, proving that the invariant probability measure supported on the orbit of q gets arbitrarily close to P in the weak-* topology. \square

We will also need the following fact:

Lemma 8.4. *Let Λ be a locally maximal, hyperbolic compact set carrying an invariant probability measure m , not necessarily ergodic. Then there exist ergodic invariant probability measures m_k carried by Λ which converges to m in the weak-* topology and in entropy: $\lim_k h(f, m_k) = h(f, m)$.*

Sketch of proof. This is routine, even if we could not locate an exact reference. Observe that it is enough to show this for a transitive subshift of finite type Σ . Given an invariant probability measure on Σ , approximate it by a Markov measure with finite memory N . Taking N sufficiently large, we can make this approximation arbitrarily close, both weak-*

and in entropy. By a small modification of the transition probabilities we can make the measure fully supported on Σ , and therefore ergodic. \square

Proof of Proposition 8.1. For convenience, we fix some distance d on the space of Borel probability measures of M , compatible with the weak-* topology. Let f be a diffeomorphism with a locally maximal transitive hyperbolic set K as given by Lemma 8.3. Since K is not reduced to a single periodic orbit, it carries invariant probability measures with positive entropy. We choose any one of them. Lemma 8.3 yields a sequence $(O_k)_{k \geq 1}$ of hyperbolic periodic orbits homoclinically related to K such that $d(\mu_{O_k}, \mu) < 1/k$ and $|\lambda^+(O_k)| < 1/k$.

Fix $k \geq 1$. Let P be a periodic orbit in K so close to μ that $d(\mu_P, \mu) < 1/k$ and $|\lambda^+(f, \mu_P) - \lambda^+(f, \mu)| < 1/k$ (by continuity of the unstable bundle over K).

Let Λ be a transitive, hyperbolic, locally maximal invariant set containing $K \cup O_k$. Define $m = \alpha\mu + (\beta - \alpha)\mu_P + (1 - \beta)\mu_{O_k}$ on Λ . Now Lemma 8.4 yields ν_k such that:

- $|h(f, \nu_k) - \alpha h(f, \mu)| = |h(f, \nu_k) - h(f, m)| < 1/k$;
- $d(\nu_k, \mu) < d(\nu_k, m) + 1/k < 2/k$;
- $|\lambda^+(f, \nu_k) - \beta \lambda^+(f, \mu)| < |\lambda^+(f, \nu_k) - \lambda^+(f, m)| + 1/k < 2/k$.

The sequence $(\nu_k)_{k \geq 1}$ is as claimed. \square

8.2. Variant inequality on Lyapunov exponents: proof of Corollary 1.3. By the Oseledec's theorem, $\lambda^+(f, \mu) + \lambda^-(f, \mu) = \int \log |\det Df| d\mu$, which is continuous with respect to (f, μ) in the $C^1 \times \text{weak-}^*$ topology. Since the sum is continuous, the discontinuities in the summands must cancel out, whence

$$\lambda^+(f, \mu) - \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = - \left(\lambda^-(f, \mu) - \lim_{k \rightarrow \infty} \lambda^-(f_k, \nu_k) \right). \quad (57)$$

Now, Theorem A is equivalent to the statement

$$\lambda^+(f, \mu) - \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) \leq \lambda^+(f, \mu) \left(1 - \frac{\lim_k h(f_k, \nu_k)}{h(f, \mu)} \right). \quad (58)$$

Applying this to f_k^{-1} and noting (57) and $\lambda^-(f, \mu) = -\lambda^+(f^{-1}, \mu)$, we obtain (4). Since $0 < -\lambda^-(f, \mu) < \lambda^+(f, \mu)$, this is stronger than the conclusion of Theorem A. \square

8.3. Sequences of non-ergodic measures. We state and prove a version of Theorem D removing its assumption that the converging measures ν_k are ergodic.

Corollary 8.5. *Fix a real number $r > 2$. For every $k \geq 1$, let $f_k \in \text{Diff}^r(M)$, ν_k be an f_k -invariant measure, not necessarily ergodic, and $\hat{\nu}_k$ be a lift satisfying $\hat{\lambda}(\hat{f}_k, \hat{\nu}_k) = \lambda^+(f_k, \nu_k)$. Let us assume that:*

- (1) $\lim_k \lambda^+(f_k, \nu_k)$ and $\lim_k h(f_k, \nu_k)$ exist and $\liminf_k \int \min(\lambda^+(f_k, x), 0) d\nu_k(x) = 0$,
- (2) $f_k \xrightarrow{r-bd} f$ for some $f \in \text{Diff}^r(M)$,
- (3) $\hat{\nu}_k \xrightarrow{w^*} \hat{\mu}$ for some \hat{f} -invariant probability measure $\hat{\mu}$ (perhaps non-ergodic) on \widehat{M} .

Then there exist $\beta \in [0, 1]$, two f -invariant measures μ_0, μ_1 with \hat{f} -invariant lifts $\hat{\mu}_0, \hat{\mu}_1$ s.t.

- (a) $\hat{\mu} = (1 - \beta)\hat{\mu}_0 + \beta\hat{\mu}_1$;
- (b) $\lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \beta \lambda^+(f, \mu_1)$;
- (c) $\lim_{k \rightarrow \infty} h(f_k, \nu_k) - \frac{1}{r-1} \lambda(\hat{f}) \leq \beta \left(h(f, \mu_1) + \frac{\lambda(f)}{r} \right)$;
- (d) if $\beta < 1$, $\hat{\lambda}(\hat{f}, \hat{\mu}_0) = 0$;
- (e) if $\beta > 0$, then $\hat{\lambda}(\hat{f}, \hat{\mu}_1) = \lambda^+(f, \mu_1)$ and $\lambda^+(f, x) > 0$ for μ_1 -a.e. x .

Note that if $f_k \in \text{Diff}^\infty(M)$ and $f_k \rightarrow f$ in C^∞ , then the corollary applies for all $r > 2$, and Property (c) becomes $\lim_{k \rightarrow \infty} h(f_k, \nu_k) \leq \beta h(f, \mu_1)$.

Note also that the conclusions (a)-(e) are the same as in Theorem D, except for the extra term $\lambda(f)/r$ on the right hand side. See the following remark on this term.

Remark 8.6. Our proof relies on discretizing the ergodic decompositions of the measures ν_k and applying Theorem D to the atoms thus defined and taking a limit. This limiting process is responsible for the term $\lambda(f)/r$ in the entropy estimate in Property (c).

Using the decomposition in the projective bundle is the key to avoid any such loss in the Lyapunov estimate eq. (b) and is therefore essential for our proof of this generalization.

Proof. Let $\mathbb{P}(M)$ denote the set of Borel probability measures on M , and let d be the L^1 -Wasserstein distance over $\mathbb{P}(M)$. This distance is compatible with the weak-* topology and satisfies $d(\sum_{i=1}^N \alpha_i \mu_i, \sum_{i=1}^N \alpha_i \nu_i) \leq \sum_{i=1}^N \alpha_i d(\mu_i, \nu_i)$ for all convex combinations.

We fix some $\varepsilon > 0$ and discretize the ergodic decompositions

$$\nu_k = \int_X \nu_{k,\xi} dP_k(\xi).$$

By compactness of $\mathbb{P}(M)$, there are measurable partitions $X = X_{k,1}^\varepsilon \sqcup \dots \sqcup X_{k,N^\varepsilon}^\varepsilon$ with number of elements N^ε independent of k , and with the following property for every $1 \leq i \leq N^\varepsilon$: There is an \widehat{f}_k -ergodic measure $\widehat{\nu}_{k,i}^\varepsilon$ with projection $\nu_{k,i}^\varepsilon$ satisfying $\widehat{\lambda}(\widehat{f}_k, \widehat{\nu}_{k,i}^\varepsilon) = \lambda^+(f_k, \nu_{k,i}^\varepsilon)$ and, for P_k -a.e. $\xi \in X_{k,i}^\varepsilon$,

$$d(\nu_{k,\xi}, \nu_{k,i}^\varepsilon) < \varepsilon, \quad |h(f_k, \nu_{k,\xi}) - h(f_k, \nu_{k,i}^\varepsilon)| < \varepsilon, \quad \text{and} \quad (59)$$

$$\lambda^+(f_k, \nu_{k,\xi}) - \frac{1}{k} < \lambda^+(f_k, \nu_{k,i}^\varepsilon) < \lambda^+(f_k, \nu_{k,\xi}) + \varepsilon. \quad (60)$$

Passing to a subsequence, we may assume without loss of generality the existence of the following limits:

$$\widehat{\mu}_i^\varepsilon := \lim_k \widehat{\nu}_{k,i}^\varepsilon, \quad \lim_k h(f_k, \nu_{k,i}^\varepsilon), \quad \lim_k \lambda^+(f_k, \nu_{k,i}^\varepsilon), \quad \alpha_i^\varepsilon := \lim_k P_k(X_{k,i}^\varepsilon).$$

For each $i = 1, \dots, N$, we set $\mu_i^\varepsilon := \widehat{\pi}_*(\widehat{\mu}_i^\varepsilon)$. By inequalities (60) and by our assumption $\liminf_k \int \min(\lambda^+(f_k, x), 0) d\nu_k(x) = 0$, it follows that either $\alpha_i^\varepsilon = 0$ or $\lim_k \lambda^+(f_k, \nu_{k,i}^\varepsilon) \geq 0$. One can thus apply Theorem D to the sequence $(f_k, \widehat{\nu}_{k,i}^\varepsilon)_{k \geq 1}$ converging to $(f, \widehat{\mu}_i^\varepsilon)$ and obtain a decomposition

$$\widehat{\mu}_i^\varepsilon = (1 - \beta_i^\varepsilon) \widehat{\mu}_{0,i}^\varepsilon + \beta_i^\varepsilon \widehat{\mu}_{1,i}^\varepsilon$$

for some $0 \leq \beta_i^\varepsilon \leq 1$ and $\widehat{\mu}_{0,i}^\varepsilon, \widehat{\mu}_{1,i}^\varepsilon$ \widehat{f} -invariant measures such that

$$\lim_k h(f_k, \nu_{k,i}^\varepsilon) \leq \beta_i^\varepsilon h(f, \mu_{1,i}^\varepsilon) + \frac{\lambda(\widehat{f})}{r-1} \quad \text{and} \quad \lim_k \lambda^+(f_k, \nu_{k,i}^\varepsilon) = \beta_i^\varepsilon \lambda^+(f, \mu_{1,i}^\varepsilon) = \beta_i^\varepsilon \widehat{\mu}_{1,i}^\varepsilon(\varphi).$$

We collect all the pieces, setting:

$$\widehat{\nu}_k^\varepsilon := \sum_{i=1}^{N^\varepsilon} \alpha_{k,i}^\varepsilon \widehat{\nu}_{k,i}^\varepsilon, \quad \widehat{\mu}^\varepsilon := \sum_{i=1}^{N^\varepsilon} \alpha_i^\varepsilon \widehat{\mu}_i^\varepsilon, \quad \beta^\varepsilon := \sum_{i=1}^{N^\varepsilon} \alpha_i^\varepsilon \beta_i^\varepsilon, \quad \widehat{\mu}_s^\varepsilon := \sum_{i=1}^{N^\varepsilon} \alpha_i^\varepsilon \widehat{\mu}_{s,i}^\varepsilon \quad \text{for } s = 0, 1.$$

We denote by $\nu_k^\varepsilon, \mu^\varepsilon, \mu_s^\varepsilon$ the projections by $\widehat{\pi}$. We have that:

$$\widehat{\mu}^\varepsilon = (1 - \beta^\varepsilon) \widehat{\mu}_0^\varepsilon + \beta^\varepsilon \widehat{\mu}_1^\varepsilon.$$

By eq. (59),

$$d(\nu_k^\varepsilon, \nu_k) \leq \varepsilon, \quad |h(f_k, \nu_k^\varepsilon) - h(f_k, \nu_k)| < \varepsilon, \quad |\lambda^+(f_k, \nu_k^\varepsilon) - \lambda^+(f_k, \nu_k)| < \varepsilon.$$

Using that the entropy and average exponents are affine functions, we get

$$\lim_k h(f_k, \nu_k) \leq \liminf_k h(f_k, \nu_k^\varepsilon) + \varepsilon \leq \beta^\varepsilon h(f, \mu_1^\varepsilon) + \frac{\lambda(\widehat{f})}{r-1} + \varepsilon \quad (61)$$

and

$$\left| \lim_k \lambda^+(f_k, \nu_k) - \lim_k \lambda^+(f_k, \nu_k^\varepsilon) \right| \leq \varepsilon \text{ so } \left| \lim_k \lambda^+(f_k, \nu_k) - \beta^\varepsilon \lambda^+(f, \mu_1^\varepsilon) \right| \leq \varepsilon. \quad (62)$$

To conclude, we pick a sequence of numbers $(\varepsilon_j)_{j \geq 1}$ decreasing to zero along which the three sequences $\widehat{\mu}^\varepsilon$, $\widehat{\mu}_0^\varepsilon$ and $\widehat{\mu}_1^\varepsilon$ converge to measures $\widehat{\mu}$, $\widehat{\mu}_0$ and $\widehat{\mu}_1$. Since $d(\mu^\varepsilon, \mu) \leq \varepsilon$, we must have $\widehat{\pi}_* \widehat{\mu} = \mu$. We define $\mu_s := \pi_*(\widehat{\mu}_s)$ for $s = 0, 1$. Property (a) follows by continuity.

By Yomdin theory (see the discussion after eq. (5)),

$$\lim_j h(f, \mu_1^{\varepsilon_j}) \leq h(f, \mu_1) + \lambda(f)/r,$$

yielding Property (c). The decomposition converges in the projective bundle, hence, recalling that $\varphi(x, E) := \log \|D_x f|_E\|$ is a continuous function,

$$\lim_j \lambda^+(f, \mu_1^{\varepsilon_j}) = \lim_j \widehat{\mu}_1^{\varepsilon_j}(\varphi) = \lambda^+(f, \mu_1).$$

Property (b) follows. For each i , one has $\widehat{\mu}_{0,i}^\varepsilon(\varphi) = 0$ once $\beta_i^\varepsilon \neq 0$; this implies Property (d).

We now turn to Property (e). We have $\beta \widehat{\lambda}(\widehat{f}, \widehat{\mu}_1) = \widehat{\lambda}(\widehat{f}, \widehat{\mu}) = \lim_{k \rightarrow \infty} \widehat{\lambda}(f_k, \widehat{\nu}_k)$ and by Property (b): $\lim_{k \rightarrow \infty} \widehat{\lambda}(f_k, \widehat{\nu}_k) = \lim_{k \rightarrow \infty} \lambda^+(f_k, \nu_k) = \beta \lambda^+(f, \mu_1)$. This gives the first part, assuming $\beta > 0$.

Note that $\lambda^+(f, x) \geq 0$ for μ -a.e. x since otherwise the ergodic decomposition of μ would contain a source as an atom and therefore ν_k would contain the same atom with uniform weight for large k , in contradiction to our assumption $\liminf_k \int \min(\lambda^+(f_k, x), 0) d\nu_k(x) = 0$. Let

$$\widehat{Z} := \{\widehat{x} \in \widehat{M} : \lambda^+(f, \widehat{\pi}(\widehat{x})) = 0\}.$$

We assume $0 < \widehat{\mu}_1(\widehat{Z}) < 1$ since it is otherwise easy to conclude. We set:

$$\beta' := \beta(1 - \widehat{\mu}_1(\widehat{Z})), \quad (1 - \beta')\widehat{\mu}'_0 := (1 - \beta)\widehat{\mu}_0 + \beta \frac{\widehat{\mu}_1(\cdot \cap \widehat{Z})}{\widehat{\mu}_1(\widehat{Z})}, \quad \beta'\widehat{\mu}'_1 := \beta \frac{\widehat{\mu}_1(\cdot \setminus \widehat{Z})}{1 - \widehat{\mu}_1(\widehat{Z})}$$

obtaining a new decomposition $\widehat{\mu} = (1 - \beta')\widehat{\mu}'_0 + \beta'\widehat{\mu}'_1$. Note that, setting $\mu'_1 := \widehat{\pi}_* \widehat{\mu}'_1$,

$$\beta' h(f, \mu'_1) = \beta h(f, \mu_1) \text{ and } \beta' \lambda^+(f, \mu'_1) = \beta \lambda^+(f, \mu_1).$$

since the probability measure $\frac{\mu_1(\cdot \cap Z)}{\mu_1(Z)}$ has both zero entropy and zero top Lyapunov exponent. This concludes the proof of the Corollary 8.5. \square

REFERENCES

- [1] A. Avila and M. Viana. Extremal Lyapunov exponents: an invariance principle and applications. *Invent. Math.* **181** (2010), 115–189.
- [2] L. Backes, A. Brown, and C. Butler. Continuity of Lyapunov exponents for cocycles with invariant holonomies. *J. Mod. Dyn.* **12** (2018), 223–260.
- [3] L. Barreira and K. Gelfert. Dimension estimates in smooth dynamics: a survey of recent results. *Ergodic Theory Dynam. Systems* **31** (2011), 641–671.
- [4] L. Barreira, Y. Pesin, and J. Schmeling. Dimension and product structure of hyperbolic measures. *Ann. of Math.* **149** (1999), 755–783.
- [5] L. Barreira and C. Wolf. Measures of maximal dimension for hyperbolic diffeomorphisms. *Comm. Math. Phys.* **239** (2003), 93–113.
- [6] J. Bochi. Genericity of zero Lyapunov exponent. *Ergodic Theory Dynam. Systems* **22** (2002), 1667–1696.

- [7] C. Bocker-Neto and M. Viana. Continuity of Lyapunov exponents for random two-dimensional matrices. *Ergodic Theory Dynam. Systems* **37** (2017), 1413–1442.
- [8] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* **153** (1971), 401–414.
- [9] M. Boyle and T. Downarowicz. The entropy theory of symbolic extensions. *Invent. Math.* **156** (2004), 119–161.
- [10] D. Burguet. Symbolic extensions in intermediate smoothness on surfaces. *Ann. Sci. Éc. Norm. Supér.* **45** (2012), 337–362.
- [11] D. Burguet. Usc/fibred entropy structure and applications. *Dyn. Syst.* **32** (2017), 391–409.
- [12] J. Buzzi. *Représentation markovienne des applications régulières de l'intervalle*. PhD thesis, Université Paris-Sud, Orsay, 1995.
- [13] J. Buzzi. Intrinsic ergodicity of smooth interval maps. *Israel J. Math.* **100** (1997), 125–161.
- [14] J. Buzzi. C^r surface diffeomorphisms with no maximal entropy measure. *Ergodic Theory Dynam. Systems* **34** (2014), 1770–1793.
- [15] M. Denker, C. Grillenberger, and K. Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin-New York, 1976.
- [16] T. Downarowicz. Entropy structure. *J. Anal. Math.* **96** (2005), 57–116.
- [17] T. Downarowicz and S. Newhouse. Symbolic extensions and smooth dynamical systems. *Invent. Math.* **160** (2005), 453–499.
- [18] M. Einsiedler, S. Kadyrov, and A. Pohl. Escape of mass and entropy for diagonal flows in real rank one situations. *Israel J. Math.* **210** (2015), 245–295.
- [19] H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.* **108** (1963), 377–428.
- [20] H. Furstenberg and Y. Kifer. Random matrix products and measures on projective spaces. *Israel J. Math.* **46** (1983), 12–32.
- [21] G. Iommi, F. Riquelme, and A. Velozo. Entropy in the cusp and phase transitions for geodesic flows. *Israel J. Math.* **225** (2018), 609–659.
- [22] G. Iommi, M. Todd, and A. Velozo. Escape of entropy for countable markov shifts. ArXiv:1908.10741.
- [23] G. Iommi, M. Todd, and A. Velozo. Upper semi-continuity of entropy in non-compact settings. *Math. Res. Lett.* **27** (2020), 1055–1077.
- [24] S. Kadyrov. Effective equidistribution of periodic orbits for subshifts of finite type. *Colloq. Math.* **149** (2017), 93–101.
- [25] S. Kadyrov and A. Pohl. Amount of failure of upper-semicontinuity of entropy in non-compact rank-one situations, and Hausdorff dimension. *Ergodic Theory Dynam. Systems* **37** (2017), 539–563.
- [26] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.* **51** (1980), 137–173.
- [27] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math.* **122** (1985), 540–574.
- [28] F. Ledrappier. Quelques propriétés des exposants caractéristiques. *Lecture Notes in Math.* **1097** (1984), 185–202. École d'été de probabilités de Saint-Flour XII, 1982.
- [29] F. Ledrappier and J.-M. Strelcyn. A proof of the estimation from below in Pesin's entropy formula. *Ergodic Theory Dynam. Systems* **2** (1982), 203–219.
- [30] F. Ledrappier and P. Walters. A relativised variational principle for continuous transformations. *J. London Math. Soc.* **16** (1977), 568–576.
- [31] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. *Ann. of Math.* **122** (1985), 509–539.
- [32] R. Mañé. *Ergodic theory and differentiable dynamics*, volume 8 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, 1987.
- [33] M. Misiurewicz. Diffeomorphism without any measure with maximal entropy. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **21** (1973), 903–910.
- [34] S. Newhouse. The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms. *Publ. Math. Inst. Hautes Études Sci.* **50** (1979), 101–151.
- [35] S. Newhouse. Entropy in smooth dynamical systems. *Proceedings of the International Congress of Mathematicians, Kyoto 1990* (1991), 1285–1294.
- [36] J. Palis and F. Takens. *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations*, volume 35 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1993.
- [37] J. Palis and M. Viana. On the continuity of Hausdorff dimension and limit capacity for horseshoes. *Lecture Notes in Math* **1331** (1988), 150–160. Dynamical systems, Valparaiso 1986.

- [38] Y. Pesin. Families of invariant manifolds that correspond to nonzero characteristic exponents. *Izv. Akad. Nauk SSSR Ser. Mat.* **40** (1976), 1332–1379.
- [39] Y. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Math. Surveys* **32** (1977), 55–114.
- [40] V. A. Pliss. On a conjecture of Smale. *Differencial' nye Uravnenija* **8** (1972), 268–282.
- [41] F. Polo. *Equidistribution in chaotic dynamical systems*. PhD thesis, The Ohio State University, 2011.
- [42] F. Riquelme and A. Velozo. Escape of mass and entropy for geodesic flows. *Ergodic Theory Dynam. Systems* **39** (2019), 446–473.
- [43] D. Ruelle. An inequality for the entropy of differentiable maps. *Bol. Soc. Brasil. Mat.* **9** (1978), 83–87.
- [44] D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Adv. in Math.* **32** (1979), 68–80.
- [45] R. Rühr. Effectivity of uniqueness of the maximal entropy measure on p -adic homogeneous spaces. *Ergodic Theory Dynam. Systems* **36** (2016), 1972–1988.
- [46] M. Viana. *Lectures on Lyapunov exponents*, volume 145 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2014.
- [47] M. Viana. (dis)continuity of Lyapunov exponents. *Ergodic Theory Dynam. Systems* **40** (2020), 577–611.
- [48] Y. Yomdin. Volume growth and entropy. *Israel J. Math.* **57** (1987), 285–300.
- [49] L.-S. Young. Dimension, entropy and Lyapunov exponents. *Ergodic Theory Dynam. Systems* **2** (1982), 109–124.
- [50] Y. Zang. Entropies and volume growth of unstable manifolds. ArXiv:1912.13189.

J. BUZZI, Laboratoire de Mathématiques d’Orsay, CNRS - UMR 8628. Université Paris-Saclay, 91405 Orsay, France. *E-mail address*: `jerome.buzzi@math.cnrs.fr`

S. CROVISIER, Laboratoire de Mathématiques d’Orsay, CNRS - UMR 8628. Université Paris-Saclay, 91405 Orsay, France. *E-mail address*: `sylvain.crovisier@math.u-psud.fr`

O. SARIG, Weizmann Institute of Science, 234 Herzl Street, 7610001 Rehovot, Israel. *E-mail address*: `omsarig@gmail.com`