Ergod. Th. & Dynam. Sys., (2022), **42**, 1284–1326
doi:10.1017/etds.2021.43

Thermodynamics of smooth models of pseudo-Anosov homeomorphisms

DOMINIC VECONI

Department of Mathematics, Penn State University, State College, Pennsylvania 16802, USA
(e-mail: dkv5049@psu.edu)

(*Received* 29 *December* 2019 *and accepted in revised form* 7 *November* 2020)

Dedicated to the memory of Anatole Katok

Abstract. We develop a thermodynamic formalism for a smooth realization of pseudo-Anosov surface homeomorphisms. In this realization, the singularities of the pseudo-Anosov map are assumed to be fixed, and the trajectories are slowed down so the differential is the identity at these points. Using Young towers, we prove existence and uniqueness of equilibrium states for geometric t-potentials. This family of equilibrium states includes a unique SRB measure and a measure of maximal entropy, the latter of which has exponential decay of correlations and the central limit theorem.

Key words: nonuniform hyperbolicity, pseudo-Anosov diffeomorphisms, thermodynamic formalism, smooth ergodic theory
2020 Mathematics Subject Classification: 37C05, 37C40, 37D25, 37D35 (Primary); 37A50, 37C86, 37E30 (Secondary)

1. *Introduction*

In [**18**], Thurston classified linear automorphisms of the torus into three classes, according to the eigenvalues of the automorphism $A \in \mathrm{SL}(2, \mathbb{Z})$:

- diagonalizable automorphisms with eigenvalues of modulus 1 (rotations);
- non-diagonalizable automorphisms (Dehn twists);
- automorphisms with eigenvalues of modulus $\neq 1$ (Anosov diffeomorphisms).

In this same work, Thurston went on to classify homeomorphisms of any surface up to isotopy class. The principle was quite similar, and is now known as the Nielson–Thurston classification of elements of mapping class groups. This is summarized in the following theorem.

THEOREM 1.1. *Let M be a compact orientable surface, and let $f : M \to M$ be a homeomorphism. Then f is isotopic to a homeomorphism F satisfying exactly one of the* *following three conditions.*

- *F is a rotation: there is an integer n for which $F^n \equiv \mathrm{Id}$.*
- *F is reducible: there is a closed curve in M which F leaves invariant.*
- *F is pseudo-Anosov.*

Of these three isotopy classes, from a dynamical systems perspective, the pseudo-Anosov maps are the most interesting. The most familiar example of a pseudo-Anosov map is the Arnold 'cat map' of the two-dimensional torus $\mathbb{T}^2$, which is in fact an Anosov diffeomorphism. No other surface admits an Anosov diffeomorphism, but pseudo-Anosov homeomorphisms of surfaces besides $\mathbb{T}^2$ form an analogy of Anosov maps to other surfaces. Like their Anosov cousins, pseudo-Anosov maps admit a pair of transverse foliations of the state space, and the map uniformly contracts points along the leaves of one foliation and uniformly dilates points along the leaves of the other. In the traditional definition of a pseudo-Anosov homeomorphism (see §2), the contraction and dilation factors are constant and inverses of each other, similarly to a hyperbolic toral automorphism such as the cat map. (Accordingly, these maps are often referred to as 'linear pseudo-Anosov maps'; see, for example, [7].) The primary difference between Anosov and pseudo-Anosov maps is the presence of finitely many singularities in the foliations. These are points where three or more leaves of one of the foliations meet at a single point. These leaves are known as 'prongs' of the singularity. The constant rate of contraction and expansion along the transverse foliations means the map is globally smooth except at the singularities. Pseudo-Anosov homeomorphisms have found their way into almost every field of geometry, such as Teichmüller theory and algebraic geometry. However, the ergodic properties of globally smooth realizations of pseudo-Anosov maps remain a relatively undeveloped area of study.

In [8], Gerber and Katok produced a C^∞ realization of pseudo-Anosov homeomorphisms by slowing down the trajectories near the isolated singularities. The result is a surface diffeomorphism that is uniformly hyperbolic away from a finite set of fixed-point singularities, but whose differential slows down to the identity at these fixed points, thus admitting Lyapunov exponents of zero. These smooth pseudo-Anosov models also admit continuous foliations whose leaves are smooth except at the fixed singular points. Pseudo-Anosov diffeomorphisms constructed in this way are analogues of the one-dimensional Manneville–Pomeau map of the unit interval to compact surfaces of arbitrary genus (see [15]), in that they admit finitely many fixed-point singularities where the differential slows down to the identity, but the map exhibits uniform hyperbolicity away from these singularities.

To discuss the ergodic properties of these pseudo-Anosov diffeomorphisms, we use techniques and results from thermodynamic formalism. Thermodynamic formalism has been used to study ergodic and geometric properties of several classes of non-uniformly hyperbolic and non-uniformly expanding maps. One objective of thermodynamic formalism is to determine the existence and uniqueness of probability measures known as *Sinai–Ruelle–Bowen (SRB) measures*. These are invariant measures that admit positive Lyapunov exponents almost everywhere, and have absolutely continuous conditional measures on unstable submanifolds (see §4). They are also known as 'physical measures',

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

in the sense that the set of points $x \in M$ for which we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \varphi(f^n(x)) = \int \varphi\, d\mu \quad \text{for any } \varphi \in C^0(M),$$

has positive measure. More generally, one also may consider *equilibrium measures* for a given potential $\varphi \in C^0(M)$. Equilibrium measures are mathematical generalizations of Gibbs distributions in statistical physics, which minimize the Helmholtz free energy of a physical system. Within thermodynamic formalism, Helmholtz free energy is replaced with the topological pressure $P_f(\varphi) = \sup\{h_\mu(f) + \int \varphi\, d\mu : \mu \in \mathcal{M}_f\}$, where $h_\mu(f)$ is the metric entropy of f with respect to μ, and $\mathcal{M}_f$ is the space of f-invariant Borel probability measures on the manifold M. Equilibrium measures, in other words, are invariant probability measures that maximize the sum of the metric entropy of f and the space average of φ with respect to μ. The two most important equilibrium measures are SRB measures (for which the potential is the negative log of the unstable Jacobian, or $\varphi_1(x) = -\log\det|Df_x|_{E^u(x)}|$), and measures of maximal entropy (for which the potential is $\varphi_0 \equiv 0$).

One of the earliest applications of thermodynamic formalism was in studying the ergodic theory of uniformly hyperbolic and Axiom A diffeomorphisms (see, for example, [2]). Since then, the theory of thermodynamic formalism has proven useful in other contexts. For example, the one-dimensional Manneville–Pomeau maps $f : [0, 1] \to [0, 1]$, defined by $f(x) = x(1 + ax^\alpha) \bmod 1$ for $a > 0$, $\alpha > 0$, have been extensively studied as classic examples of one-dimensional non-uniformly expanding maps (see, for example, [14, 20] for some recent work on the infinite ergodic theory of Manneville–Pomeau maps). Additionally, in [5], Climenhaga, Pesin, and Zelerowicz proved existence of equilibrium measures for a broad class of potential functions in the partially hyperbolic setting. These equilibrium measures include, in particular, a unique measure of maximal entropy and a unique SRB measure. Finally, in [3], Buzzi, Crovisier, and Sarig showed that any surface diffeomorphism admits at most finitely many ergodic measures of maximal entropy, and that there is a unique such measure in the topologically transitive case. Our results are a special instance of this setting, and develop further statistical and ergodic properties of the measure of maximal entropy and other equilibrium states.

In this paper we effect a thermodynamic formalism for these pseudo-Anosov diffeomorphisms. Specifically, given a pseudo-Anosov diffeomorphism g of a compact surface M, we consider the family of geometric t-potentials $\varphi_t(x) = -t \log|Dg|_{E^u(x)}|$ parametrized by $t \in \mathbb{R}$, where $E^u(x)$ is the stable subspace at the point $x \in M$. Our main result, Theorem 4.1, claims that there is a number $t_0 < 0$ such that for every $t \in (t_0, 1)$, there is a unique equilibrium measure μ_t for φ_t that is Bernoulli, has exponential decay of correlations, and satisfies the central limit theorem with respect to a class of functions containing all Hölder continuous functions on M. We also show that the pressure function $t \mapsto P_g(\varphi_t)$ is real analytic in the open interval $(t_0, 1)$. Since the pseudo-Anosov diffeomorphism g is topologically conjugate to a pseudo-Anosov homeomorphism f, their topological entropies agree, and since f has a unique measure of maximal entropy, so does g. We denote this measure μ_0, for the potential $\varphi_0 \equiv 0$. As a corollary to Theorem 4.1, we obtain a thorough description of the statistical properties of μ_0. Furthermore, we prove that the map g has a unique SRB measure, and we describe its ergodic properties. We emphasize that a phase transition occurs at $t = 1$: in addition to the SRB measure, there is a family of ergodic equilibrium measures for φ_1 composed of convex combinations of Dirac measures at the singularities.

The techniques we employ to establish our results are similar to those used by Pesin, Senti, and Zhang in [13] to effect thermodynamic formalism of the Katok map. The latter is an area-preserving diffeomorphism of the torus with non-zero Lyapunov exponents. Similarly to the smooth pseudo-Anosov models, the Katok map is obtained by slowing down trajectories near the origin to produce an indifferent fixed point (that is, a fixed point of the map whose differential is equal to the identity). However, there are substantial differences between the Katok map of the torus and the Gerber–Katok smooth pseudo-Anosov models. These include the following.

- The Katok map acts on the torus, and thus can be lifted to $\mathbb{R}^2$, while pseudo-Anosov maps do not in general admit a lift to $\mathbb{R}^2$. The lift of the Katok map to $\mathbb{R}^2$ plays an essential role in simplifying the analysis in [13], and some adjustments to this argument are required to carry out similar analysis of globally smooth pseudo-Anosov diffeomorphisms.
- The foliations of pseudo-Anosov diffeomorphisms are singular. In particular, the singularities do not admit a locally stable or unstable subspace forming a curve, but rather forming the prongs that meet at the singularity. Furthermore, one cannot use coordinate charts whose interiors contain the singularities if the coordinates correspond to the stable and unstable foliations. Instead, the analysis must be performed in stable and unstable sectors whose vertices are the singularities (see §3).
- Whereas the slowdown function used to construct the Katok map depends only on the radius of the slowed-down neighborhood, the choice of slowdown function of the pseudo-Anosov homeomorphism depends on the number of prongs of the singularity. This affects the analysis of the behavior of the trajectories near the singularities.

The development of thermodynamics of the Katok map in [13] uses the technology of Young diffeomorphisms, which are generalizations of hyperbolic maps. The definition of Young diffeomorphisms relies on hyperbolicity of an induced map on a small subset of the state space with local hyperbolic product structure. This induced map can be carried over to a derived dynamical system on the corresponding Rokhlin tower. The thermodynamics of Young diffeomorphisms have been thoroughly investigated in [12, 17]. Young towers have been used to study thermodynamic and ergodic properties of a variety of non-uniformly hyperbolic dynamical systems (see [4]), including almost Anosov toral diffeomorphisms (see [19]).

This paper is structured as follows. In §2 we define pseudo-Anosov homeomorphisms and discuss some of their dynamical properties, including measure invariance and Markov partitions. In §3 we describe the smooth models of pseudo-Anosov homeomorphisms and state some important dynamical and topological properties of these maps. In §4 we state our main results. Section 5 is devoted to the study of dynamics near the singularities and includes some technical calculations needed to prove our main result. Some of these calculations are similar to those performed in [13, §5] but require some modifications and adjustments. Section 6 gives a brief survey of the thermodynamic properties of Young diffeomorphisms and inducing schemes we will be using. Section 7 proves that our smooth models of pseudo-Anosov homeomorphisms are Young diffeomorphisms, and finally §8 uses this fact to prove our main results.

2. *Preliminaries*

We begin with a discussion on measured foliations of a compact two-dimensional C^∞ Riemannian manifold M, where we assume M is without boundary. Our exposition is adapted from the presentation in [1, §6.4]. For the reader's convenience, we have restated their exposition here and have included additional details and remarks on the notation concerning pseudo-Anosov maps and their behavior.

Definition 2.1. A *measured foliation with singularities* is a triple $(\mathcal{F}, S, \nu)$, where:

- $S = \{x_1, \ldots, x_m\}$ is a finite set of points in M, called *singularities*;
- $\mathcal{F} = \widetilde{\mathcal{F}} \uplus \mathcal{S}$ is a partition of M, where $\mathcal{S}$ is a partition of S into points and $\widetilde{\mathcal{F}}$ is a smooth foliation of $M \setminus S$;
- ν is a *transverse measure*; in other words, ν is a measure defined on each curve on M transverse to the leaves of $\widetilde{\mathcal{F}}$.

The triple satisfies the following properties.

(1) There is a finite atlas of C^∞ charts $\phi_k : U_k \to \mathbb{C}$ for $k = 1, \ldots, \ell$, $\ell \geq m$.

(2) For each $k = 1, \ldots, m$, there is a number $p = p(k) \geq 3$ of elements of $\widetilde{\mathcal{F}}$ meeting at $x_k \in S$ (these elements are called *prongs* of x_k) such that:
  (a) $\phi_k(x_k) = 0$ and $\phi_k(U_k) = D_{a_k} := \{z \in \mathbb{C} : |z| \leq a_k\}$ for some $a_k > 0$;
  (b) if $C \in \widetilde{\mathcal{F}}$, then the components of $C \cap U_k$ are mapped by ϕ_k to sets of the form
$$\{z \in \mathbb{C} : \mathrm{Im}(z^{p/2}) = \text{constant}\} \cap \phi_k(U_k);$$
  (c) the measure $\nu|U_k$ is the pullback under ϕ_k of
$$|\mathrm{Im}(dz^{p/2})| = |\mathrm{Im}(z^{(p-2)/2}dz)|.$$

(3) For each $k > m$, we have:
  (a) $\phi_k(U_k) = (0, b_k) \times (0, c_k) \subset \mathbb{R}^2 \approx \mathbb{C}$ for some $b_k, c_k > 0$;
  (b) if $C \in \widetilde{\mathcal{F}}$, then components of $C \cap U_k$ are mapped by ϕ_k to lines of the form
$$\{z \in \mathbb{C} : \mathrm{Im}\, z = \text{constant}\} \cap \phi_k(U_k);$$
  (c) the measure $\nu|U_k$ is given by the pullback of $|\mathrm{Im}\, dz|$ under ϕ_k.

An archetypal singularity with $p = 3$ prongs is shown in Figure 1.

Remark 2.2. Henceforth, we refer to the C^∞ curves that are elements of $\mathcal{F}$ as 'leaves (of the foliation)'; in particular, despite the technical fact that the singleton sets of singularities $\{x_1\}, \ldots, \{x_k\}$ are elements of $\mathcal{F}$, we do not refer to these points when we refer to 'leaves of the foliation'.

Remark 2.3. The transverse measure ν is not a measure on M itself, in the measure-theoretic sense of the word. What ν is measuring is the 'distance traveled'

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

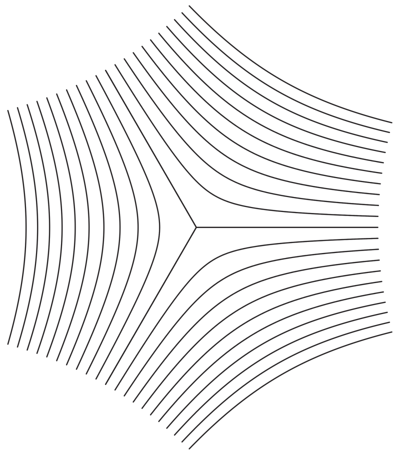

FIGURE 1. A three-pronged singularity of a measured foliation with singularities.

transverse to the leaves of the foliation, similarly to how the 1-form dx measures distance traveled transverse to the leaves $\{x = x_0\}$. To make this more explicit, properties (2) and (3) in the above definition ensure that ν is holonomy-invariant. In particular, if γ and γ' are isotopic curves in $M \setminus S$ transverse to the leaves of $\mathcal{F}$, and the initial points of γ and γ' lie in the same leaf $\mathcal{F}_0$ and the terminal points lie in the same leaf $\mathcal{F}_1$, then $\nu(\gamma) = \nu(\gamma')$.

Definition 2.4. A surface homeomorphism f of a manifold M is *pseudo-Anosov* if there are measured foliations with singularities $(\mathcal{F}^s, S, \nu^s)$ and $(\mathcal{F}^u, S, \nu^u)$ (with the same finite set of singularities $S = \{x_1, \ldots, x_m\}$) and an atlas of C^∞ charts $\phi_k : U_k \to \mathbb{C}$ for $k = 1, \ldots, \ell$, $\ell > m$, satisfying the following properties.

(1) f is differentiable, except on S.
(2) For each $x_k \in S$, $\mathcal{F}^s$ and $\mathcal{F}^u$ have the same number $p(k)$ of prongs at x_k.
(3) The leaves of $\mathcal{F}^s$ and $\mathcal{F}^u$ intersect transversally at non-singular points.
(4) Both measured foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ are f-invariant.
(5) There is a constant $\lambda > 1$ such that

$$f(\mathcal{F}^s, \nu^s) = (\mathcal{F}^s, \nu^s/\lambda) \quad \text{and} \quad f(\mathcal{F}^u, \nu^u) = (\mathcal{F}^u, \lambda\nu^u).$$

(6) For each $k = 1, \ldots, m$, we have $x_k \in U_k$, and $\phi_k : U_k \to \mathbb{C}$ satisfies:
   (a) $\phi_k(x_k) = 0$ and $\phi_k(U_k) = D_{a_k}$ for some $a_k > 0$;
   (b) if C is a curve leaf in $\mathcal{F}^s$, then the components of $C \cap U_k$ are mapped by ϕ_k to sets of the form

$$\{z \in \mathbb{C} : \mathrm{Re}(z^{p/2}) = \text{constant}\} \cap D_{a_k};$$

   (c) if C is a curve leaf in $\mathcal{F}^u$, then the components of $C \cap U_k$ are mapped by ϕ_k to sets of the form

$$\{z \in \mathbb{C} : \mathrm{Im}(z^{p/2}) = \text{constant}\} \cap D_{a_k};$$

   (d) the measures $\nu^s|U_k$ and $\nu^u|U_k$ are given by the pullbacks of

$$|\mathrm{Re}(dz^{p/2})| = |\mathrm{Re}(z^{(p-2)/2}dx)|$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

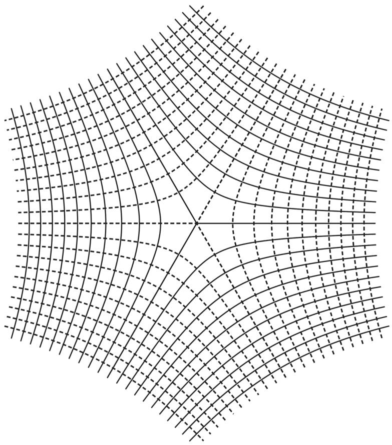

FIGURE 2. A singular neighborhood with a three-pronged singularity. The solid lines and broken lines respectively represent the stable and unstable foliations $\mathcal{F}^s$ and $\mathcal{F}^u$, for example.

and

$$|\mathrm{Im}(dz^{p/2})| = |\mathrm{Im}(z^{(p-2)/2}dx)|$$

under ϕ_k, respectively.

(7) For each $k > m$, we have:

(a) $\phi_k(U_k) = (0, b_k) \times (0, c_k) \subset \mathbb{R}^2 \approx \mathbb{C}$ for some $b_k, c_k > 0$;

(b) if C is a curve leaf in $\mathcal{F}^s$, then components of $C \cap U_k$ are mapped by ϕ_k to lines of the form

$$\{z \in \mathbb{C} : \mathrm{Re}\, z = \text{constant}\} \cap \phi_k(U_k);$$

(c) if C is a curve leaf in $\mathcal{F}^u$, then components of $C \cap U_k$ are mapped by ϕ_k to lines of the form

$$\{z \in \mathbb{C} : \mathrm{Im}\, z = \text{constant}\} \cap \phi_k(U_k);$$

(d) the measures $\nu^s|U_k$ and $\nu^u|U_k$ are given by the pullbacks of $|\mathrm{Re}\, dz|$ and $|\mathrm{Im}\, dz|$ under ϕ_k, respectively.

For $k = 1, \ldots, m$, we call the neighborhood $U_k \subset M$ described in part (6) of this definition a *singular neighborhood*, and for $k > m$, we call U_k a *regular neighborhood* (see Figure 2).

Remark 2.5. The notation $f(\mathcal{F}^u, \nu^u) = (\mathcal{F}^u, \lambda\nu^u)$ means two things. First, it means that if γ is a subset of a leaf of $\mathcal{F}^u$, then so is $f(\gamma)$, and in particular, so is $f^{-1}(\gamma)$. Second, it means if γ is an open interval in $\mathcal{F}^s$, or more generally any arc in M transverse to the foliation $\mathcal{F}^u$, then $\nu^u(f^{-1}(\gamma)) = \lambda\nu^u(\gamma)$. That is, $f_*\nu^u = \lambda\nu^u$, with $f_*\nu^u$ the pushforward transverse measure. Likewise for the notation $f(\mathcal{F}^s, \nu^s) = (\mathcal{F}^s, \nu^s/\lambda)$. So points on the same $\mathcal{F}^s$-leaf contract in the ν^u-measure by a factor of λ, and points on the same $\mathcal{F}^u$-leaf dilate in the ν^s-measure by a factor of λ.

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Remark 2.6. Since f is a homeomorphism, f permutes the singularities; that is, the singular set S is f-invariant. However, our arguments assume the singularities are fixed under the pseudo-Anosov homeomorphism. If the singularities are not fixed points, one could consider an appropriate iterate of f and study the dynamics of this iterate, arriving at the same results.

We state a few important properties of pseudo-Anosov homeomorphisms we will use over the course of our arguments.

PROPOSITION 2.7. *Let $f : M \to M$ be a pseudo-Anosov homeomorphism. For $x \in M \setminus S$, $T_x M = T_x \mathcal{F}^s(x) \oplus T_x \mathcal{F}^u(x)$, and in these coordinates, $Df_x(\xi^s, \xi^u) = (\xi^s/\lambda, \lambda\xi^u)$, where ξ^s and ξ^u are non-zero vectors in $T_x \mathcal{F}^s(x)$ and $T_x \mathcal{F}^u(x)$, $\mathcal{F}^s(x)$ and $\mathcal{F}^u(x)$ represent the curve containing x in the respective foliation, and λ is the dilation factor for f.*

Proof. This follows immediately from the definition of pseudo-Anosov diffeomorphisms after a calculation in coordinates (see Remark 2.5). □

PROPOSITION 2.8. (See [6, Exposé 10]) *A pseudo-Anosov surface homeomorphism $f : M \to M$ preserves a smooth invariant probability measure ν defined locally as the product of ν^s on $\mathcal{F}^u$-leaves with ν^u on $\mathcal{F}^s$-leaves. In any coordinate chart of M, this probability measure ν has a density with respect to the measure induced by the Lebesgue measure on $\mathbb{R}^2$, and this density vanishes at singularities.*

PROPOSITION 2.9. (See [6, Exposé 10]) *Every pseudo-Anosov homeomorphism of a surface M admits a finite Markov partition of arbitrarily small diameter. Conjugated to the symbolic system induced by this Markov partition, with the measure ν as in the preceding proposition, (M, f, ν) is Bernoulli.*

3. *Pseudo-Anosov diffeomorphisms*

Generally speaking, pseudo-Anosov homeomorphisms as defined in Definition 2.4 are differentiable everywhere except at the singularities x_k with $p(k) \geq 3$. This is a consequence of the fact that f contracts (respectively, expands) points in the stable (respectively, unstable) leaves of the foliation, so the differential of f cannot possibly be linear at the singularities.

In this section we construct a surface diffeomorphism $g : M \to M$ that is topologically conjugate to the pseudo-Anosov homeomorphism f, and whose differential at the singularity is the identity. (Since we assume the singularities are fixed, this is a reasonable statement.)

Before proceeding with the construction, we point out that some literature refers to the maps defined in Definition 2.4 as 'pseudo-Anosov diffeomorphisms', despite the fact that these maps are not differentiable at the singularities. To avoid any confusion, we reserve the word 'diffeomorphism' only for those maps that are differentiable on all of M, and use the phrase 'pseudo-Anosov homeomorphism' for the maps described in Definition 2.4.

Let $x_k \in S$, let $p = p(x_k)$, and let $\phi_k : U_k \to \mathbb{C}$ be the chart described in part (6) of Definition (2.4). The *stable* and *unstable prongs* at x_k are the leaves P^s_{kj} and P^u_{kj}, $j = 0, \ldots, p-1$ of $\mathcal{F}^s$ and $\mathcal{F}^u$, respectively, whose endpoints meet at x_k. Locally, they are

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

given by

$$P^s_{kj} = \phi_k^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_k,\ \tau = \frac{2j+1}{p}\pi\right\},$$
$$P^u_{kj} = \phi_k^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_k,\ \tau = \frac{2j}{p}\pi\right\}.$$

For simplicity, assume $f(P^s_{kj}) \subseteq P^s_{kj}$ for all $j = 1, \ldots, p$. Furthermore, we define the *stable* and *unstable sectors* at x_k to be the regions in U_k bounded by the stable (respectively, unstable) prongs:

$$S^s_{kj} = \phi_k^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_k,\ \frac{2j-1}{p}\pi \le \tau \le \frac{2j+1}{p}\pi\right\},$$
$$S^u_{kj} = \phi_k^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_k,\ \frac{2j}{p}\pi \le \tau \le \frac{2j+2}{p}\pi\right\}.$$

The strategy for creating our diffeomorphism g is adapted from [1, §6.4.2]. In each stable sector, we apply a 'slowdown' of the trajectories, followed by a change of coordinates ensuring the resulting diffeomorphism g preserves the measure induced by a convenient Riemannian metric.

Let $F : \mathbb{C} \to \mathbb{C}$ be the map $s_1 + is_2 \mapsto \lambda s_1 + is_2/\lambda$. Note F is the time-1 map of the vector field V given by

$$\dot{s}_1 = (\log\lambda)s_1, \quad \dot{s}_2 = -(\log\lambda)s_2.$$

Let $0 < r_1 < r_0 < \min\{a_1, \ldots, a_\ell\}$, and define $\widetilde{r}_0$ and $\widetilde{r}_1$ by $\widetilde{r}_j = (2/p)r_j^{p/2}$ for $j = 0, 1$ and for each $p = p(k)$. Define a 'slowdown' function Ψ_p for the p-pronged singularity on the interval $[0, \infty)$ so that:

(1) $\Psi_p(u) = (p/2)^{(2p-4)/p}u^{(p-2)/p}$ for $u \le \widetilde{r}_1^2$;

(2) Ψ_p is C^∞ except at 0;

(3) $\dot{\Psi}_p(u) \ge 0$ for $u > 0$;

(4) $\Psi_p(u) = 1$ for $u \ge \widetilde{r}_0^2$.

Consider the vector field V_{Ψ_p} on $D_{\widetilde{r}_0} \subset \mathbb{C}$ defined by

$$\dot{s}_1 = (\log\lambda)s_1\Psi_p(s_1^2 + s_2^2) \quad \text{and} \quad \dot{s}_2 = -(\log\lambda)s_2\Psi_p(s_1^2 + s_2^2). \tag{3.1}$$

Let G_p be the time-1 map of the vector field V_{Ψ_p}. Assume r_1 is chosen to be small enough so that $G_p = F$ on a neighborhood of the boundary of $D_{\widetilde{r}_0}$, and assume r_0 is chosen to be small enough so that the open neighborhood $\mathcal{U}_0 := \bigcup_{k=1}^m \phi_k^{-1}(D_{r_0})$ of S is disjoint from the open set $\bigcup_{k=m+1}^{\ell} \phi_k^{-1}(D_{a_k})$. We also define the open neighborhood $\widetilde{\mathcal{U}}_0 := \bigcup_{k=1}^m \phi_k^{-1}(D_{\widetilde{r}_0}) \subset \mathcal{U}_0$, as well as $\mathcal{U}_1$ and $\widetilde{\mathcal{U}}_1$ defined analogously with D_{r_1} and $D_{\widetilde{r}_1}$, respectively.

Let $\widetilde{a}_k = (2/p)a_k^{p/2}$, and define the coordinate change $\Phi_{kj} : \phi_k S^s_{kj} \to \{z : \mathrm{Re} z \ge 0\} \cap D_{\widetilde{a}_k}$ by

$$\Phi_{kj}(z) = (2/p)z^{p/2} = w = s_1 + is_2.$$

Define $g : M \to M$ by $g(x) = f(x)$ for $x \notin \mathcal{U}_0$, and meanwhile for $1 \le k \le m$, $1 \le j \le p(k)$, define g on each sector $S^s_{kj} \cap \phi_k^{-1}(D_{r_0})$ by

$$g(x) = \phi_k^{-1}\Phi_{kj}^{-1}G_p\Phi_{kj}\phi_k(x).$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

PROPOSITION 3.1. (See [1]) *The map g defined above is well defined on the unstable prongs and singularity. It is in fact a diffeomorphism topologically conjugate to f, and for any $\varepsilon > 0$, r_0 and r_1 can be chosen so that $\|f - g\|_{C^0} < \varepsilon$. In particular, g admits a Markov partition of arbitrarily small diameter.*

Next we define a Riemannian metric $\zeta = \langle\cdot,\cdot\rangle$ on $M \setminus S$ with respect to which the map g is invariant. In the stable sector $S^s_{kj} \cap \phi_k^{-1}(D_{\widetilde{a}_k})$, we consider the coordinates $w = s_1 + is_2$ given by $\Phi_{kj} \circ \phi_k$ defined above. Outside of this neighborhood, we use the coordinates $z = s_1 + is_2$. In both sets of coordinates, the stable and unstable transversal measures are $\nu^s = |ds_1|$ and $\nu^u = |ds_2|$. On stable sectors in $M \setminus S$, we define the Riemannian metric ζ to be the pullback of $(ds_1^2 + ds_2^2)/\Psi_p(s_1^2 + s_2^2)$ under $\Phi_{kj} \circ \phi_k$. In regular neighborhoods (U_k, ϕ_k), we define $\zeta = \phi_k^*(ds_1^2 + ds_2^2)$. Since $\widetilde{r}_0$ is chosen so that $\phi_k^{-1}(D_{\widetilde{r}_0})$ is disjoint from regular neighborhoods, and $\Psi_p(u) \equiv 1$ for $u \geq \widetilde{r}_0^2$, ζ is consistently defined on chart overlaps. One can further show that ζ agrees with the Euclidean metric in $\phi_k^{-1}(D_{\widetilde{r}_0})$. So ζ can be extended to a Riemannian metric on all of M.

PROPOSITION 3.2. (See [1]) *Letting $z = t_1 + it_2$ be the coordinates given by (ϕ_k, U_k), $1 \leq k \leq m$, the Riemannian metric ζ is actually the Euclidean metric $dt_1^2 + dt_2^2$. In particular, the diffeomorphism $g : M \to M$ is μ_1-area-preserving, where μ_1 is the volume determined by ζ.*

For stable sectors S^s_{kj}, we use the coordinates $w = \Phi^s_{kj}(z) = s_1 + is_2$, and in regular neighborhoods U_k, $k \geq m$, we use the coordinates $z = s_1 + is_2$. Then s_1 represents the coordinate in the unstable foliation, and s_2 is the coordinate in the stable foliation. Define the coordinates (ξ_1, ξ_2) in each tangent space T_xM, $x \in M \setminus S$, to be the coordinates with respect to

$$(\Phi_{kj} \circ \phi_k)_*^{-1}\left(\Psi_p(s_1^2 + s_2^2)\frac{\partial}{\partial s_i}\right), \quad i = 1, 2, \tag{3.2}$$

in each stable sector, and with respect to $(\phi_k)_*^{-1}(\partial/\partial s_i)$, $i = 1, 2$, in each regular neighborhood. For $x \in M \setminus S$, let C_x^+ be the cone in T_xM bounded by the lines $\xi_1 = \pm\xi_2$, respectively, and containing the tangent line to the $\mathcal{F}^u$ leaf through x. Respectively define C_x^- to be the cone containing the $\mathcal{F}^s$ leaf.

PROPOSITION 3.3. (See [1]) *For $x \in M \setminus S$, the cones C_x^+, C_x^- satisfy the following assertions.*

(1) *C_x^+ and C_x^- depend continuously on $x \in M \setminus S$.*
(2) *C_x^+ (respectively, C_x^-) is strictly invariant under Dg (respectively, Dg^{-1}) on $x \in M \setminus S$.*
(3) *For each $x \in M \setminus S$, the intersections*

$$E^u(x) := \bigcap_{n=0}^{\infty} Dg^n C^+_{g^{-n}(x)} \quad \text{and} \quad E^s(x) := \bigcap_{n=0}^{\infty} Dg^{-n} C^-_{g^n(x)}$$

are one-dimensional subspaces of T_xM. Moreover, if $x \in M \setminus S$ is on an unstable leaf, then $E^u(x)$ is tangent to the unstable leaf (and similarly for $E^s(x)$ on a stable leaf).

(4) *$E^u(x)$ and $E^s(x)$ depend continuously on $x \in M \setminus S$.*

We will need a stronger condition on cone invariance. For $x \in M \setminus S$ and for $0 < \alpha < 1$, define the families of cones $K^+(x)$ and $K^-(x)$ by

$$K^+(x) = \{v = (\xi_1, \xi_2) \in T_x M : |\xi_2| < \alpha|\xi_1|\},$$
$$K^-(x) = \{v = (\xi_1, \xi_2) \in T_x M : |\xi_1| < \alpha|\xi_2|\}.$$

In the original construction of pseudo-Anosov diffeomorphisms yielding Proposition 3.3, we have $\alpha = 1$. But for certain later arguments, we will require $\alpha < 1$.

LEMMA 3.4. *There exists a $0 < \alpha_0 < 1$ such that for all $\alpha_0 < \alpha < 1$, and for all $x \in M$,*

$$Dg_x K^+(x) \subseteq K^+(g(x)) \quad \text{and} \quad Dg^{-1}_{g(x)} K^-(g(x)) \subseteq K^-(x).$$

Proof. We prove invariance only for $K^+(x)$; the invariance of the stable cones is proven similarly by considering g^{-1}. Assume $x \in \widetilde{\mathcal{U}}_0$, as the result is clearly true outside of $\widetilde{\mathcal{U}}_0$. Consider the vector field (3.1) defined on $\mathbb{C}$. The variational equations for (3.1) give us

$$\frac{d\zeta_1}{dt} = \log\lambda((\Psi_p(u) + 2s_1^2\dot{\Psi}_p(u))\xi_1 + 2s_1s_2\dot{\Psi}_p(u)\xi_2)$$

and

$$\frac{d\zeta_2}{dt} = -\log\lambda(2s_1s_2\dot{\Psi}_p(u)\xi_1 + (\Psi_p(u) + 2s_2^2\dot{\Psi}_p(u))\xi_2),$$

where $u := s_1^2 + s_2^2$. The 'slope' $\eta := \xi_2/\xi_1$ of a tangent vector in $\mathbb{C}$ changes under the flow of (3.1) as:

$$\frac{d\eta}{dt} = -2\log\lambda((1+\eta^2)s_1s_2\dot{\Psi}_p(u) + (\Psi_p(u) + (s_1^2+s_2^2)\dot{\Psi}_p)\eta). \tag{3.3}$$

Suppose $\widetilde{r}_1^2 \le u \le \widetilde{r}_0^2$. Since $\Psi_p > 0$, and $\dot{\Psi}_p > 0$ is decreasing, we have

$$\frac{\Psi_p(u)}{\dot{\Psi}_p(u)} \ge \frac{\Psi_p(\widetilde{r}_1^2)}{\dot{\Psi}_p(\widetilde{r}_1^2)} = \frac{p}{p-2}\widetilde{r}_1^2 \ge \frac{p}{p-2}\left(\frac{\widetilde{r}_1}{\widetilde{r}_0}\right)^2 u.$$

Meanwhile, if $0 < u < \widetilde{r}_1^2$, we have

$$\frac{\Psi_p(u)}{\dot{\Psi}_p(u)} = \frac{p}{p-2}u \ge \frac{p}{p-2}\left(\frac{\widetilde{r}_1}{\widetilde{r}_0}\right)^2 u.$$

If $\eta > 0$, this gives us

$$\begin{aligned}\frac{d\eta}{dt} &\le -2\log\lambda\dot{\Psi}_p(u)\left((1+\eta^2)s_1s_2 + \left(1 + \frac{p}{p-2}\left(\frac{\widetilde{r}_1}{\widetilde{r}_2}\right)^2\right)(s_1^2+s_2^2)\eta\right) \\ &= -2\log\lambda\dot{\Psi}_p(u)\left(\left(\left(1+\frac{p}{p-2}\left(\frac{\widetilde{r}_1}{\widetilde{r}_0}\right)^2\right)\eta - \frac{1}{2}(1+\eta^2)\right)(s_1^2+s_2^2) \right.\\ &\qquad \left. + \frac{1}{2}(1+\eta^2)(s_1+s_2)^2\right) \\ &\le -2\log\lambda\dot{\Psi}_p(u)\psi(\eta)(s_1^2+s_2^2),\end{aligned}$$

where $\psi(\eta) := p/(p-2)(\widetilde{r}_1/\widetilde{r}_2)^2 - \frac{1}{2}(\eta - 1)^2$. Since $\psi(1) > 0$, there is a $\alpha_0 \in (0, 1)$ with $\psi(\eta) > 0$ for $\alpha_0 < \eta < 1$. Therefore $d\eta/dt < 0$ for $\alpha_0 < \eta < 1$. For $\eta < 0$, we have

$$\begin{aligned}\frac{d\eta}{dt} &= 2\log\lambda((\Psi_p(u) + (s_1^2 + s_2^2)\dot{\Psi}_p(u))|\eta| - s_1 s_2(1+\eta^2)\dot{\Psi}_p(u)) \\ &\geq 2\log\lambda\dot{\Psi}_p(u)\bigg(\bigg(1 + \frac{p}{p-2}\bigg(\frac{\widetilde{r}_1}{\widetilde{r}_0}\bigg)^2\bigg)(s_1^2 + s_2^2)|\eta| - s_1 s_2(1+\eta^2)\bigg).\end{aligned}$$

A similar argument will show $d\eta/dt > 0$ for $-1 < \eta < -\alpha_0$. Letting $\alpha = \eta$, for $z \in \mathbb{C}$, we have $D(G_p)_z K_0^+(z) \subseteq K_0^+(G_p(z))$ and $D(G_p)^{-1}_{G_p(z)} K_0^-(G_p(z)) \subseteq K_0^-(z)$, where

$$\begin{aligned}K_0^+(z) &= \{(\zeta_1, \zeta_2) \in T_z\mathbb{C} : |\zeta_2| < \alpha|\zeta_1|\}, \\ K_0^-(z) &= \{(\zeta_1, \zeta_2) \in T_z\mathbb{C} : |\zeta_1| < \alpha|\zeta_2|\}.\end{aligned}$$

Note that α_0 does not depend on the distance of $z \in \mathbb{C}$ from 0. Applying the coordinate map $\phi_k^{-1} \circ \Phi_{kj}^{-1} : \{z : \mathrm{Re}(z) \geq 0\} \cap D_{\widetilde{a}_k} \to M$, the cones $K^+(x)$ and $K^-(x)$ defined using the coordinates in (3.2) for $T_x M$ satisfy the same invariance property as K_0^+ and K_0^-. This proves the lemma. □

4. *Main results*

We begin by defining the relevant ergodic properties under consideration. Given a continuous potential function $\varphi : M \to \mathbb{R}$, a probability measure μ_φ on M is an *equilibrium measure* for φ if

$$P_g(\varphi) = h_{\mu_\varphi}(g) + \int_M \varphi \, d\mu_\varphi,$$

where $h_{\mu_\varphi}(g)$ is the metric entropy of g with respect to μ_φ, and $P_g(\varphi)$ is the topological pressure of φ; that is, $P_g(\varphi)$ is the supremum of $h_\mu(g) + \int_M \varphi \, d\mu$ over all g-invariant probability measures μ on M.

A special instance of equilibrium measures are known as SRB measures. Given a (uniformly, non-uniformly, or partially) hyperbolic function $f : M \to M$ on a Riemannian manifold M, an f-invariant Borel probability measure μ on M is called an *SRB measure* if f admits positive Lyapunov exponents μ-almost everywhere, and if the conditional measures of μ on the unstable submanifolds are absolutely continuous with respect to the Riemannian leaf volume.

Additionally, we say that g has *exponential decay of correlations* with respect to a measure $\mu \in \mathcal{M}(g, M)$ and a class of functions $\mathcal{H}$ on M if there exists $\kappa \in (0, 1)$ such that for any $h_1, h_2 \in \mathcal{H}$,

$$\bigg|\int h_1(g^n(x)) h_2(x) \, d\mu(x) - \int h_1(x) \, d\mu(x) \int h_2(x) \, d\mu(x)\bigg| \leq C\kappa^n$$

for some $C = C(h_1, h_2) > 0$. Furthermore, g is said to satisfy the *central limit theorem* for a class $\mathcal{H}$ of functions if for any $h \in \mathcal{H}$ that is not a coboundary (namely, $h \neq h' \circ g - h'$ for any $h' \in \mathcal{H}$), there exists $\sigma > 0$ such that

$$\lim_{n\to\infty} \mu\bigg\{\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \bigg(h(g^i(x)) - \int h \, d\mu\bigg) < t\bigg\} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\tau^2/2\sigma^2} \, d\tau.$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

The family of potential functions we consider are the *geometric t-potentials* defined by $\varphi_t(x) = -t \log |Dg_x|_{E^u(x)}|$. Although the unstable distribution E^u does not continuously extend to the singularities, the differential Dg_{x_0} is the identity at each singularity x_0, so φ_t continuously extends to the singularities; in particular, $\varphi_t(x_0) = 0$ for each singularity x_0. So the geometric t-potential is well defined in this setting.

Our result shows there is a $t_0 < 0$ for which every $t \in (t_0, 1)$ admits a unique equilibrium state $\mu_{\varphi_t} =: \mu_t$ for the potential $\varphi_t : M \to \mathbb{R}$. When $t = 0$, $\varphi_0 \equiv 0$, so the equilibrium measure μ_0 satisfies $P_g(0) = h_{\mu_0}(g)$, and so μ_0 is the unique measure of maximal entropy for g.

We now state our main result.

THEOREM 4.1. *Consider a pseudo-Anosov diffeomorphism $g : M \to M$ on a compact Riemannian manifold M. The following statements hold.*

(1) *Given any $t_0 < 0$, we may take $r_0 > 0$ in the construction of g so that for any $t \in (t_0, 1)$, there is a unique equilibrium measure μ_t associated to φ_t. This equilibrium measure has exponential decay of correlations and satisfies the central limit theorem with respect to a class of functions containing all Hölder continuous functions on M, and is Bernoulli. Additionally, the pressure function $t \mapsto P_g(\varphi_t)$ is real analytic in the open interval $(t_0, 1)$.*
(2) *For $t = 1$, there are two classes of equilibrium measures associated to φ_1: convex combinations of Dirac measures concentrated at the singularities, and a unique invariant SRB measure μ.*
(3) *For $t > 1$, the equilibrium measures associated to φ_t are precisely the convex combinations of Dirac measures concentrated at the singularities.*

Remark 4.2. Uniqueness of the measure μ_t for $t \in (t_0, 1)$ implies that this measure is ergodic, but in fact Theorem 4.1 gives us that this measure is Bernoulli.

Remark 4.3. Taking $t = 0$, this theorem shows that the dynamical system (M, g) admits a unique measure of maximal entropy that is Bernoulli, has exponential decay of correlations, and satisfies the central limit theorem.

Remark 4.4. Although we know $t \mapsto P_g(\varphi_t)$ is real analytic in $(t_0, 1)$, we do not know about the behavior of $P_g(\varphi_t)$ for $t \leq t_0$. In particular, it is not known whether (M, g, φ_t) admits a phase transition at $t = t_0$.†

5. *Dynamics near singularities*

In this section, we discuss the dynamical properties of pseudo-Anosov diffeomorphisms, considering both their global behavior and their behavior near singularities. The thermodynamic constructions we will develop in §§6 and 7 require bounds on how quickly nearby orbits diverge from each other. For this reason, the estimates and inequalities

† For the Katok map, it is shown in [21] that for sufficiently small values of the parameters $\alpha > 0$ and $r > 0$, the Katok map has a unique equilibrium measure μ_t corresponding to the geometric potential φ_t for all values of $t < 1$.

collected in this section will become important tools to examine how nearby orbits behave in neighborhoods of the singularities.

Several of the technical calculations made here are similar to the calculations performed for the Katok map in [13]. However, they are carried out here for the reader's convenience, as well as the fact that the slowdown function in the Katok map uses different constants depending on the radius of the slowed-down neighborhood (by contrast, our slowdown function depends not on the radius of the slowdown, but on the number of prongs of the singularity).

Our first two technical estimates concern how long an orbit remains in a neighborhood of a singularity. Recall our definitions $\widetilde{r}_j = (2/p)r_j^{p/2}$ for $j = 0, 1$. In particular, $\widetilde{r}_0$ and $\widetilde{r}_1$ depend on p, and thus depend on k for $k = 1, \ldots, m$.

LEMMA 5.1. *There exists a $T_p > 0$, depending on p, λ, r_0, and r_1, so that for any solution $s(t)$ of (3.1) with $s(0) \in D_{\widetilde{r}_0}$,*

$$\max\{t > 0 : s(t) \in D_{\widetilde{r}_0} \setminus D_{\widetilde{r}_1}\} < T_p.$$

Proof. The value $s_1 s_2$ is invariant under the flow. If $s_1 s_2 \geq \frac{1}{2}\widetilde{r}_1^2$, then when $s_1 = s_2$, the minimum value of $s_1^2 + s_2^2$ is at least $\widetilde{r}_1^2$, and the trajectory never enters $D_{\widetilde{r}_1}$. If $s_1 s_2 < \frac{1}{2}\widetilde{r}_1^2$, the trajectory either will enter $D_{\widetilde{r}_1}$ or has already entered $D_{\widetilde{r}_1}$ and is on its way out of $D_{\widetilde{r}_0}$.

Case 1: $s_1 s_2 \geq \frac{1}{2}\widetilde{r}_1^2$. Since $\widetilde{r}_0^2 \geq s_1^2 + s_2^2 \geq s_2^2$, we have $\frac{1}{4}\widetilde{r}_1^4 \leq s_1^2 s_2^2 \leq s_1^2 \widetilde{r}_0^2$, so $s_1^2 \geq \widetilde{r}_1^4 / 4\widetilde{r}_0^2$. So, since Ψ_p is an increasing function,

$$\frac{d}{dt}(s_1^2) = 2s_1^2 \Psi_p(s_1^2 + s_2^2) \log \lambda \geq \frac{\widetilde{r}_1^4}{2\widetilde{r}_0^2} \Psi_p(\widetilde{r}_1^2) \log \lambda.$$

It follows that the time T it takes for s_1^2 to reach $\widetilde{r}_0^2$ from $s_1^2(0) \geq \widetilde{r}_1^4 / 4\widetilde{r}_0^2$ satisfies

$$T \leq \frac{\widetilde{r}_0^2 - (\widetilde{r}_1^4 / 4\widetilde{r}_0^2)}{(\widetilde{r}_1^4 / 2\widetilde{r}_0^2)\Psi_p(\widetilde{r}_1^2) \log \lambda} = \frac{4r_0^{2p} - r_1^{2p}}{2r_1^{3p-2} \log \lambda}.$$

Case 2: $s_1 s_2 < \frac{1}{2}\widetilde{r}_1^2$. Assume that $s_1 < s_2$, ensuring that the trajectory will enter $D_{\widetilde{r}_1}$. If we can prove there is a uniform time bound T before which this happens, then by symmetry of the vector field, the same T is an upper bound for the time it takes this trajectory to exit $D_{\widetilde{r}_0}$ when $s_1 > s_2$.

We will in fact establish a bound on how long it takes s_2^2 to decrease from $s_2^2(0)$ to $\frac{1}{2}\widetilde{r}_1^2$ when $s_1 < s_2$. For then, because $s_1 s_2 < \frac{1}{2}\widetilde{r}_1^2$, by the time $s_2^2 = \frac{1}{2}\widetilde{r}_1^2$, the trajectory will already have entered $D_{\widetilde{r}_1}$. So, $s_2^2 \geq \frac{1}{2}\widetilde{r}_1^2$, and since in this case $s_1^2 + s_2^2 \geq \frac{1}{2}\widetilde{r}_1^2$, we have

$$\frac{d}{dt}(s_2^2) = -2s_2^2 \Psi_p(s_1^2 + s_2^2) \log \lambda \leq -\widetilde{r}_1^2 \Psi_p\left(\frac{1}{2}\widetilde{r}_1^2\right) \log \lambda.$$

It follows that the time T it takes for s_2^2 to reach $\frac{1}{2}\widetilde{r}_1^2$ from $s_2^2(0) \leq \widetilde{r}_0^2$ satisfies

$$T \leq \frac{\widetilde{r}_0^2 - (1/2)\widetilde{r}_1^2}{\widetilde{r}_1^2 \Psi_p((1/2)\widetilde{r}_1^2) \log \lambda} = 2^{(p-2)/2} \frac{2r_0^p - r_1^p}{2r_1^{2p-2} \log \lambda}. \qquad \square$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

LEMMA 5.2. *There exists a $T \in \mathbb{Z}$, depending on r_0 and λ, so that for any $x \in \widetilde{\mathcal{U}}_0 := \bigcup_{k=1}^{m} \phi_k^{-1}(D_{\widetilde{r}_0}) \subset M$, we have*

$$\max \left\{ N > 0 : g^n(x) \in \bigcup_{k=1}^{m} \phi_k^{-1}(D_{\widetilde{r}_0} \setminus D_{\widetilde{r}_1}) \text{ for all } n = 0, \ldots N \right\} \leq T.$$

Proof. This follows from Lemma 5.1 after taking $T = \max\{T_{p(k)} : k = 1, \ldots, m\}$. □

Next, we will establish bounds on how quickly nearby points will diverge while remaining near the singularities. The main lemma that demonstrates this bound is Lemma 5.5.

LEMMA 5.3. *For $i, j = 1, 2$, define the functions $d_{ij} : D_{\widetilde{r}_1} \to \mathbb{R}$ by*

$$d_{ij}(s_1, s_2) = \frac{\partial^2}{\partial s_i \partial s_j}(s_2 \Psi_p(s_1^2 + s_2^2)).$$

Then

$$\max_{i,j=1,2} |d_{ij}(s_1, s_2)| \leq \frac{6p - 12}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} (s_1^2 + s_2^2)^{(p-4)/2p}.$$

Proof. Recall that for $u \leq \widetilde{r}_1^2$, we have $\Psi_p(u) = (p/2)^{(2p-4)/p} u^{(p-2)/p}$. So,

$$\frac{\partial}{\partial s_1}(s_2 \Psi_p(s_1^2 + s_2^2)) = \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} s_1 s_2 (s_1^2 + s_2^2)^{-2/p},$$

$$\frac{\partial}{\partial s_2}(s_2 \Psi_p(s_1^2 + s_2^2)) = \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} s_2^2 (s_1^2 + s_2^2)^{-2/p} + \left(\frac{p}{2}\right)^{(2p-4)/p} (s_1^2 + s_2^2)^{(p-2)/p}.$$

Note $|s_1|^2 \leq \sqrt{s_1^2 + s_2^2}$, and since $p \geq 3$,

$$-2 \leq -\frac{4 s_1^2}{p(s_1^2 + s_2^2)} \leq 0.$$

Therefore, for all $(s_1, s_2) \in D_{\widetilde{r}_1}$,

$$\begin{aligned}
|d_{11}(s_1, s_2)| &= \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} \left| \frac{\partial}{\partial s_1} s_1 s_2 (s_1^2 + s_2^2)^{-2/p} \right| \\
&= \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} \left| s_2 (s_1^2 + s_2^2)^{-2/p} - \frac{4}{p} s_1^2 s_2 (s_1^2 + s_2^2)^{-(p+2)/p} \right| \\
&= \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} |s_2| (s_1^2 + s_2^2)^{-2/p} \left| 1 - \frac{4 s_1^2}{p(s_1^2 + s_2^2)} \right| \\
&\leq \frac{2p-4}{p} \left(\frac{p}{2}\right)^{(2p-4)/p} (s_1^2 + s_2^2)^{(p-4)/2p}.
\end{aligned}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

A similar argument applies for $d_{12} = d_{21}$ and for d_{22}, though in d_{22} we use the estimate $-2 \leq 4s_1^2/3p(s_1^2+s_2^2)$ instead:

$$\begin{aligned}|d_{12}(s_1,s_2)| &= \frac{2p-4}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}|s_1|(s_1^2+s_2^2)^{-2/p}\left|1-\frac{4s_2^2}{p(s_1^2+s_2^2)}\right| \\ &\leq \frac{2p-4}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}(s_1^2+s_2^2)^{(p-4)/2p}, \\ |d_{22}(s_1,s_2)| &= \frac{6p-12}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}|s_2|(s_1^2+s_2^2)^{-2/p}\left|1-\frac{4s_2^2}{3p(s_1^2+s_2^2)}\right| \\ &\leq \frac{6p-12}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}(s_1^2+s_2^2)^{(p-4)/2p}. \qquad \square\end{aligned}$$

Let $s(t) = (s_1(t),\ s_2(t))$ be a solution to (3.1), and assume $s(t)$ is defined in the unique interval $[0, T]$ for which $G_p^{-1}(s(0)), G_p(s(T)) \notin D_{\tilde{r}_1}$ and $s(t) \in \overline{D_{\tilde{r}_1}}$ for $0 \leq t \leq T$. In particular, this means $s(0), s(T) \in \partial D_{\tilde{r}_1}$. (Recall that G_p is the time-1 map of the vector field (3.1).) Further denote $T_1 = T/2$, so that if $s_1(t) > 0$ and $s_2(t) > 0$ for $t \in [0, T]$, we have $s_1(t) \leq s_2(t)$ for $t \in [0, T_1]$ and $s_1(t) \geq s_1(t)$ for $t \in [T_1, T]$.

LEMMA 5.4. *Given a solution $s(t)$ to (3.1), and T and T_1 defined above, we have the following inequalities:*

(a) $|s_1(t)| \leq |s_1(b)|(1 + C_0 s_1(b)^{(2p-4)/p}(b-t))^{-p/(2p-4)}$, $0 \leq t \leq b \leq T$;
(b) $|s_2(t)| \leq |s_2(a)|(1 + C_0 s_2(a)^{(2p-4)/p}(t-a))^{-p/(2p-4)}$, $0 \leq a \leq t \leq T$;
(c) $|s_2(t)| \geq |s_2(a)|(1 + 2^{(p-2)/p}C_0 s_2(a)^{(2p-4)/p}(t-a))^{-p/(2p-4)}$, $0 \leq a \leq t \leq T_1$;
(d) $|s_1(t)| \geq |s_1(b)|(1 + 2^{(p-2)/p}C_0 s_1(b)^{(2p-4)/p}(b-t))^{-p/(2p-4)}$, $T_1 \leq t \leq b \leq T$.

Here $C_0 = (2p-4)/p(p/2)^{(2p-4)/p}\log\lambda$.

Proof. By symmetry, we may assume $s_1(t) > 0$ and $s_2(t) > 0$ for $t \in [0, T]$. Then, using the facts that $s_1^2 + s_2^2 \geq s_i^2$ for $i = 1, 2$, and that $\Psi_p(u) = (p/2)^{(2p-4)/p}u^{(p-2)/p}$ for $0 \leq u \leq \tilde{r}_1^2$, (3.1) implies

$$\begin{aligned}\frac{d}{dt}s_1(t) &\geq \left(\frac{p}{2}\right)^{(2p-4)/p}s_1(t)^{(3p-4)/p}\log\lambda, \\ \frac{d}{dt}s_2(t) &\leq -\left(\frac{p}{2}\right)^{(2p-4)/p}s_2(t)^{(3p-4)/p}\log\lambda.\end{aligned}$$

In particular, this gives us

$$\begin{aligned}s_1(t)^{-(3p-4)/p}\frac{d}{dt}s_1(t) &\geq \left(\frac{p}{2}\right)^{(2p-4)/p}\log\lambda, \\ s_2(t)^{-(3p-4)/p}\frac{d}{dt}s_2(t) &\leq -\left(\frac{p}{2}\right)^{(2p-4)/p}\log\lambda.\end{aligned}$$

Integrating these expressions between a and b, where $0 \leq a \leq b \leq T$, we get:

$$\begin{aligned}s_2(b)^{-(2p-4)/p} - s_2(a)^{-(2p-4)/p} &\geq C_0(b-a), \\ s_1(b)^{-(2p-4)/p} - s_1(a)^{-(2p-4)/p} &\leq -C_0(b-a),\end{aligned}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

where $C_0 = (2p-4)/p(p/2)^{(2p-4)/p} \log \lambda$. From assuming that $s_i(t) > 0$, $i = 1, 2$, we get inequalities (a) and (b).

Using the fact that $s_1(t) \leq s_2(t)$ for $0 \leq t \leq T_1 = \frac{1}{2}T$ and $s_1(t) \geq s_2(t)$ for $T_1 \leq t \leq T$, we get

$$s_1(t)^2 + s_2(t)^2 \leq 2s_2(t)^2, \quad 0 \leq t \leq T_1,$$
$$s_1(t)^2 + s_2(t)^2 \leq 2s_1(t)^2, \quad T_1 \leq t \leq T.$$

Once again, applying (3.1) yields

$$\frac{d}{dt}s_1(t) \leq 2^{(p-2)/p}\left(\frac{p}{2}\right)^{(2p-4)/p} s_1(t)^{(3p-4)/p} \log \lambda, \quad T_1 \leq t \leq T,$$
$$\frac{d}{dt}s_2(t) \geq -2^{(p-2)/p}\left(\frac{p}{2}\right)^{(2p-4)/p} s_2(t)^{(3p-4)/p} \log \lambda, \quad 0 \leq T_1 \leq T.$$

Using the same integration strategy from a to b as before gives us

$$s_1(b)^{-(2p-4)/p} - s_1(t)^{-(2p-4)/p} \geq -2^{(p-2)/p}C_0(b-t), \quad T_1 \leq t \leq b \leq T,$$
$$s_2(t)^{-(2p-4)/p} - s_2(a)^{-(2p-4)/p} \leq 2^{(p-2)/p}C_0(t-a), \quad 0 \leq a \leq t \leq T_1.$$

This gives us inequalities (c) and (d). □

Now suppose $\widetilde{s}(t) = (\widetilde{s}_1(t), \widetilde{s}_2(t))$ is another solution of (3.1) defined for $t \in [0, T]$. We will need an upper and lower bound for $\Delta s(t) := \widetilde{s}(t) - s(t)$. Let $\Delta s_j(t) = \widetilde{s}_j(t) - s_j(t)$, $j = 1, 2$.

LEMMA 5.5. *Suppose $s_1(t) \neq 0 \neq s_2(t)$ for $t \in [0, T]$ and that $\Delta s_2(t) > 0$ for $t \in [0, T]$. Suppose further that $0 < \alpha < 1$ satisfies:*
(1) $|\Delta s_1(t)| \leq \alpha \Delta s_2(t)$ *for $t \in [0, T]$;*
(2) $|\Delta s_2(0)/s_2(0)| \leq (1-\alpha)/72$.
Then,

$$\Delta s_2(t) \leq \frac{\Delta s_2(0)}{s_2(0)} s_2(t)(1 + 2^{(p-2)/p}C_0 s_2(0)^{(2p-4)/p} t)^{-\beta}, \qquad 0 \leq t \leq T_1,$$
$$\Delta s_2(t) \leq \frac{\Delta s_2(T_1)}{s_1(T_1)} s_1(t)\left(\frac{1 + 2^{(p-2)/p}C_0 s_1(b)^{(2p-4)/p}(b-t)}{1 + 2^{(p-2)/p}C_0 s_1(b)^{(2p-4)/p}(b-T_1)}\right)^{\beta}, \qquad T_1 \leq t \leq b \leq T,$$

where $\beta = 2^{-(3p-2)/p}(1-\alpha)$, and C_0 is the constant from Lemma 5.4. Furthermore, for $0 \leq a \leq T_1 \leq b \leq T$,

$$\|\Delta s(b)\| \leq \sqrt{1+\alpha^2}\frac{s_1(b)}{s_2(a)}\|\Delta s(a)\|. \tag{5.1}$$

Proof. Assume $s_j(t) > 0$ for $j = 1, 2$; the other cases follow by symmetry. Further denote $u = s_1^2 + s_2^2$ and $\widetilde{u} = \widetilde{s}_1^2 + \widetilde{s}_2^2$. Applying equation (3.1) to the second Lagrange remainder of the function $(s_1, s_2) \mapsto s_2\Psi_p(s_1^2 + s_2^2)$ centered at the point (s_1, s_2), we get

$$\begin{aligned}
\frac{d}{dt}\Delta s_2 &= -\log\lambda(\widetilde{s}_2\Psi_p(\widetilde{u}) - s_2\Psi_p(u)) \\
&= -\log\lambda\bigg(\frac{\partial}{\partial s_1}(s_2\Psi_p(u))\Delta s_1 + \frac{\partial}{\partial s_2}(s_2\Psi_p(u))\Delta s_2 \\
&\quad + \frac{1}{2}\sum_{j,k=1,2} d_{jk}(\xi_1, \xi_2)\Delta s_j\Delta s_k\bigg) \\
&= -\log\lambda\bigg(2s_1s_2\dot{\Psi}_p(u)\Delta s_1 + (\Psi_p(u) + 2s_2^2\dot{\Psi}_p(u))\Delta s_2 \\
&\quad + \frac{1}{2}\sum_{j,k=1,2} d_{jk}(\xi_1, \xi_2)\Delta s_j\Delta s_k\bigg),
\end{aligned}$$

where the d_{jk} are as in Lemma (5.3) and $\xi = (\xi_1, \xi_2) \in D_{\widetilde{r}_1}$ is such that ξ_j lies between s_j and $\widetilde{s}_j$ for $j = 1, 2$.It follows that

$$\begin{aligned}
\frac{d}{dt}\bigg(\frac{\Delta s_2}{s_2}\bigg) &= \frac{1}{s_2}\frac{d}{dt}\Delta s_2 - \frac{1}{s_2^2}\dot{s}_2\Delta s_2 \\
&= -\log\lambda\bigg(2s_1\dot{\Psi}_p(u)\Delta s_1 + \frac{1}{s_2}\Psi_p(u)\Delta s_2 + 2s_2\dot{\Psi}_p(u)\Delta s_2\bigg) \\
&\quad - \frac{\log\lambda}{2}\sum_{j,k=1,2} d_{jk}(\xi_1, \xi_2)\frac{\Delta s_j\Delta s_k}{s_2} + \log\lambda\frac{1}{s_2}\Psi_p(u)\Delta s_2 \\
&= -\frac{(2p-4)\log\lambda}{p}\bigg(\frac{p}{2}\bigg)^{(2p-4)/p} u^{-2/p}(s_1\Delta s_1 + s_2\Delta s_2) \\
&\quad - \frac{\log\lambda}{2}\sum_{j,k=1,2} d_{jk}(\xi_1, \xi_2)\frac{\Delta s_j\Delta s_k}{s_2}.
\end{aligned}$$

Suppose $0 \le t \le T_1$, so that $0 < s_1(t) \le s_2(t)$. Since $|\Delta s_1(t)| \le \alpha\Delta s_2(t)$ by assumption, we get

$$s_1\Delta s_1 + s_2\Delta s_2 \ge (-s_1\alpha + s_2)\Delta s_2 \ge (1-\alpha)s_2\Delta s_2.$$

Lemma 5.3 implies

$$\sum_{j,k} d_{jk}(\xi_1, \xi_2)\Delta s_j\Delta s_k \ge -\frac{24p-48}{p}\bigg(\frac{p}{2}\bigg)^{(2p-4)/p}(\xi_1^2 + \xi_2^2)^{(p-4)/2p}(\Delta s_2)^2. \tag{5.2}$$

It follows from the above two inequalities that

$$\begin{aligned}
\frac{d}{dt}\bigg(\frac{\Delta s_2}{s_2}\bigg) &\le -(1-\alpha)\frac{(2p-4)\log\lambda}{p}\bigg(\frac{p}{2}\bigg)^{(2p-4)/p}(s_1^2 + s_2^2)^{-2/p}s_2\Delta s_2 \\
&\quad + \frac{(12p-24)\log\lambda}{p}\bigg(\frac{p}{2}\bigg)^{(2p-4)/p}(\xi_1^2 + \xi_2^2)^{(p-4)/2p}\frac{(\Delta s_2)^2}{s_2}.
\end{aligned}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Since $s_1(t) \le s_2(t)$ for $0 \le t \le T_1$, we have $s_2^2 \le s_1^2 + s_2^2 \le 2s_2^2$. Therefore,

$$\begin{aligned}\frac{d}{dt}\left(\frac{\Delta s_2}{s_2}\right) &\le -(1-\alpha)\frac{(2p-4)\log\lambda}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}(s_1^2+s_2^2)^{(p-2)/p}\frac{s_2^2}{s_1^2+s_2^2}\frac{\Delta s_2}{s_2}\\ &\quad+\frac{(12p-24)\log\lambda}{p}\left(\frac{p}{2}\right)^{(2p-4)/p}s_2^{(2p-4)/p}\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p}\left(\frac{\Delta s_2}{s_2}\right)^2\\ &\le -(1-\alpha)\frac{(p-2)\log\lambda}{p}\left(\frac{ps_2}{2}\right)^{(2p-4)/p}\frac{\Delta s_2}{s_2}\\ &\quad+\frac{(12p-24)\log\lambda}{p}\left(\frac{ps_2}{2}\right)^{(2p-4)/p}\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p}\left(\frac{\Delta s_2}{s_2}\right)^2.\end{aligned}$$

Denoting $\kappa = \kappa(t) = (\Delta s_2/s_2)(t)$, we summarize:

$$\begin{aligned}\frac{d\kappa}{dt} &\le -(1-\alpha)\frac{(p-2)\log\lambda}{p}\left(\frac{ps_2}{2}\right)^{(2p-4)/p}\kappa\\ &\quad+\frac{(12p-24)\log\lambda}{p}\left(\frac{ps_2}{2}\right)^{(2p-4)/p}\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p}\kappa^2\\ &= -\frac{(p-2)\log\lambda}{p}\left(\frac{ps_2}{2}\right)^{(2p-4)/p}\kappa\left(1-\alpha-12\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p}\kappa\right).\end{aligned}\tag{5.3}$$

Note that $0 < s_2 \le \xi_2 \le \widetilde{s}_2 = s_2 + \Delta s_2$, and $\xi_1 \le s_1 + |\Delta s_1| \le s_2 + \alpha\Delta s_2$. Therefore,

$$\begin{aligned}1 \le \frac{\xi_2^2}{s_2^2} \le \frac{\xi_1^2+\xi_2^2}{s_2^2} &\le \frac{(s_2+\alpha\Delta s_2)^2+(s_2+\Delta s_2)^2}{s_2^2}\\ &= (1+\alpha\kappa)^2+(1+\kappa)^2 < 2(1+\kappa)^2.\end{aligned}\tag{5.4}$$

It follows that

$$\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p} \le \begin{cases}1 & \text{if } p = 3, 4,\\ (2(1+\kappa)^2)^{(p-4)/2p} & \text{if } p \ge 5.\end{cases}$$

Using assumption (2), we observe that

$$1-\alpha-12\left(\frac{\xi_1^2+\xi_2^2}{s_2^2}\right)^{(p-4)/2p}\kappa(0) \ge \frac{1-\alpha}{2}.$$

Equation (5.3) now implies

$$\left.\frac{d\kappa}{dt}\right|_{t=0} \le -\frac{(1-\alpha)(p-2)\log\lambda}{2p}\left(\frac{ps_2(0)}{2}\right)^{(2p-4)/p}\kappa(0) < 0.$$

So $\kappa(t)$ satisfies

$$0 < \kappa(t) < \frac{1-\alpha}{72}\tag{5.5}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

for $0 \leq t < \delta$ for a small number $\delta > 0$. The same arguments as before now imply

$$\frac{d\kappa}{dt} \leq -\frac{(1-\alpha)(p-2)\log\lambda}{2p}\left(\frac{ps_2(t)}{2}\right)^{(2p-4)/p}\kappa(t) < 0 \tag{5.6}$$

for $0 \leq t < \delta$. Since κ and s_2 are continuous and positive on $[0, T_1]$, the estimates (5.5) and (5.6) apply for $0 \leq t \leq T_1$. Applying Grönwall's inequality to (5.6) gives us, for $0 \leq t \leq T_1$,

$$\kappa(t) \leq \kappa(0)\exp\left(-\frac{(1-\alpha)(p-2)\log\lambda}{2p}\left(\frac{p}{2}\right)^{(2p-4)/p}\int_0^t s_2(\tau)^{(2p-4)/p}\,d\tau\right). \tag{5.7}$$

Applying the third inequality in Lemma 5.4 to this integral gives us

$$\begin{aligned}\int_0^t s_2(\tau)^{(2p-4)/p}\,d\tau &\geq \int_0^t s_2(0)^{(2p-4)/p}(1+2^{(p-2)/p}C_0s_2(0)^{(2p-4)/p}\tau)^{-1}\,d\tau\\ &= \frac{1}{2^{(p-2)/p}C_0}\log(1+2^{(p-2)/p}C_0s_2(0)^{(2p-4)/p}t).\end{aligned}$$

Recalling that $C_0 = (2p-4)/p(p/2)^{(2p-4)/p}\log\lambda$, (5.7) now becomes

$$\begin{aligned}\kappa(t) &\leq \kappa(0)\exp\left(-\frac{(1-\alpha)}{2^{(3p-2)/p}}\log(1+2^{(p-2)/p}C_0s_2(0)^{(2p-4)/p}t)\right)\\ &= \kappa(0)(1+2^{(p-2)/p}C_0s_2(0)^{(2p-4)/p}t)^{-\beta},\end{aligned} \tag{5.8}$$

giving us the first inequality of the lemma.

To prove the second inequality, arguing as before for $T_1 \leq t \leq T$, we get

$$\begin{aligned}\frac{d}{dt}\Delta s_2 = -\log\lambda\bigg(&\frac{\partial}{\partial s_1}(s_2\Psi_p(u))\Delta s_1 + \frac{\partial}{\partial s_2}(s_2\Psi_p(u))\Delta s_2\\ &+\frac{1}{2}\sum_{j,k=1,2}d_{jk}(\xi_1,\xi_2)\Delta s_j\Delta s_k\bigg)\end{aligned}$$

for $\xi = (\xi_1, \xi_2)$ satisfying $\min\{s_j, \widetilde{s}_j\} \leq \xi_j \leq \max\{s_j, \widetilde{s}_j\}$. Thus, using assumption (1) and positivity of s_i, $\dot{\Psi}_p$, and Δs_2,

$$\begin{aligned}\frac{d}{dt}\left(\frac{\Delta s_2}{s_1}\right) &= \frac{1}{s_1}\frac{d}{dt}\Delta s_2 - \frac{1}{s_1^2}\dot{s}_1\Delta s_2\\ &= -\log\lambda\left(2s_1s_2\dot{\Psi}_p(u)\frac{\Delta s_1}{s_1} + (2s_2^2\dot{\Psi}_p(u)+\Psi_p(u))\frac{\Delta s_2}{s_1}\right)\\ &\quad -\frac{1}{2}\log\lambda\sum_{j,k=1,2}d_{jk}(\xi_1,\xi_2)\frac{\Delta s_j\Delta s_k}{s_1} - \log\lambda\Psi_p(u)\frac{\Delta s_2}{s_1}\\ &\leq -2\log\lambda(\Psi_p(u) - \alpha s_1s_2\dot{\Psi}_p(u) + s_2^2\dot{\Psi}_p(u))\frac{\Delta s_2}{s_1}\\ &\quad -\frac{1}{2}\log\sum_{j,k}d_{j,k}(\xi_1,\xi_2)\frac{\Delta s_j\Delta s_k}{s_1}\\ &\leq -2\log\lambda(\Psi_p(u) - \alpha s_1s_2\dot{\Psi}_p(u))\frac{\Delta s_2}{s_1} - \frac{\log\lambda}{2}\sum_{j,k}d_{j,k}(\xi_1,\xi_2)\frac{\Delta s_j\Delta s_k}{s_1}.\end{aligned}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Observe that

$$\frac{\Psi_p}{\dot{\Psi}_p} - \alpha s_1 s_2 = \frac{p}{p-2}(s_1^2 + s_2^2) - \alpha s_1 s_2 \geq \left(\frac{p}{p-2} - \frac{\alpha}{2}\right)(s_1^2 + s_2^2)$$
$$\geq \frac{p(2-\alpha)}{2(p-2)}(s_1^2 + s_2^2).$$

It follows, in particular, that

$$\Psi_p(u) - \alpha s_1 s_2 \dot{\Psi}_p(u) \geq \left(\frac{p}{2}\right)^{(2p-4)/p} \frac{2-\alpha}{2}(s_1^2 + s_2^2)^{(p-2)/p}.$$

Furthermore, applying the inequality in (5.2), we get

$$\frac{d}{dt}\left(\frac{\Delta s_2}{s_1}\right) \leq -\log\lambda\left(\frac{p}{2}\right)^{(2p-4)/p}(2-\alpha)s_1^{(2p-4)/p}\frac{\Delta s_2}{s_1}$$
$$+ \log\lambda\left(\frac{p}{2}\right)^{(2p-4)/p} s_1^{(2p-4)/p}\frac{12(p-2)}{p}\left(\frac{\xi_1^2 + \xi_2^2}{s_1^2}\right)^{(p-4)/2p}\left(\frac{\Delta s_2}{s_1}\right)^2.$$

In particular, if we denote $\chi(t) = (\Delta s_2/s_1)(t)$, we find that

$$\frac{d\chi}{dt} \leq -\log\lambda\left(\frac{p}{2}\right)^{(2p-4)/p} s_1^{(2p-4)/p}\chi\left(2 - \alpha - \frac{12(p-2)}{p}\left(\frac{\xi_1^2+\xi_2^2}{s_1^2}\right)^{(p-4)/2p}\chi\right). \tag{5.9}$$

Recall that $\min\{s_j, \widetilde{s}_j\} \leq \xi_j \leq \max\{s_j, \widetilde{s}_j\}$, and that $\Delta s_j = \widetilde{s}_j - s_j$ for $j = 1, 2$. Therefore,

$$s_j - |\Delta s_j| \leq \xi_j \leq s_j + |\Delta s_j|.$$

In particular, since $|\Delta s_1| \leq \alpha \Delta s_2$ by assumption (1), we get

$$\xi_1^2 + \xi_2^2 \geq \xi_1^2 \geq (s_1 - |\Delta s_1|)^2 \geq (s_1 - \alpha\Delta s_2)^2 = s_1^2\left(1 - \frac{\alpha\Delta s_2}{s_1}\right)^2 \geq s_1^2(1-\chi)^2.$$

Furthermore, since $s_2(t) \leq s_1(t)$ whenever $T_1 \leq t \leq T$, we get

$$\frac{\xi_1^2 + \xi_2^2}{s_1^2} \leq \left(1 + \frac{|\Delta s_1|}{s_1}\right)^2 + \left(\frac{s_2}{s_1} + \frac{\Delta s_2}{s_1}\right)^2 \leq (1+\alpha\chi)^2 + (1+\chi)^2 < 2(1+\chi)^2.$$

It follows that

$$\left(\frac{\xi_1^2+\xi_2^2}{s_1^2}\right)^{(p-4)/2p} \leq \begin{cases} (1-\chi)^{(p-4)/p}, & p = 3, 4, \\ 2^{(p-4)/2p}(1+\chi)^{(p-4)/p}, & p \geq 5. \end{cases}$$

Since $s_1(T_1) = s_2(T_1)$, by the first estimate in this lemma and assumption (2), we find that

$$0 \leq \chi(T_1) = \frac{\Delta s_2(T_1)}{s_1(T_1)} = \frac{\Delta s_2(T_1)}{s_2(T_1)} \leq \frac{\Delta s_2(0)}{s_2(0)} \leq \frac{1-\alpha}{72}.$$

Again, applying assumption (2) gives us

$$2 - \alpha - \frac{12(p-2)}{p}\left(\frac{\xi_1^2+\xi_2^2}{s_1^2}\right)^{(p-4)/2p}\chi(T_1) \geq \frac{1-\alpha}{2}.$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

So (5.9) now becomes

$$\left.\frac{d\chi}{dt}\right|_{t=T_1} < -\frac{(1-\alpha)\log\lambda}{2}\left(\frac{p}{2}\right)^{(2p-4)/p} s_1(T_1)^{(2p-4)/p}\chi(T_1) < 0. \tag{5.10}$$

Repeating the argument for the first estimate in this lemma, we find that the inequalities in (5.10) hold for all $t \in [T_1, T]$. For $T_1 \le t \le b \le T$, by Grönwall's inequality and inequality (d) in Lemma 5.4, we get

$$\begin{aligned}
\chi(t) &\le \chi(T_1)\exp\left(-\frac{(1-\alpha)\log\lambda}{2}\left(\frac{p}{2}\right)^{(2p-4)/p}\int_{T_1}^{t} s_1(\tau)^{(2p-4)/p}\,d\tau\right)\\
&\le \chi(T_1)\exp\left(-\frac{(1-\alpha)\log\lambda}{2}\left(\frac{p}{2}\right)^{(2p-4)/p} s_1(b)^{(2p-4)/p}\right.\\
&\quad\left.\times\int_{T_1}^{t}(1+2^{(p-2)/p}C_0 s_1(T_1)^{(2p-4)/p} s_1(T_1)^{(2p-4)/p}(b-\tau))^{-1}\,d\tau\right)\\
&= \chi(T_1)\exp\left(\frac{p(1-\alpha)}{2^{(3p-2)/p}(p-2)}\log\left(\frac{1+2^{(p-2)/p}C_0 s_1(T_1)^{(2p-4)/p}(b-t)}{1+2^{(p-2)/p}C_0 s_1(T_1)^{(2p-4)/p}(b-T_1)}\right)\right)\\
&= \chi(T_1)\left(\frac{1+2^{(p-2)/p}C_0 s_1(T_1)^{(2p-4)/p}(b-t)}{1+2^{(p-2)/p}C_0 s_1(T_1)^{(2p-4)/p}(b-T_1)}\right)^{\beta p/(p-2)}.
\end{aligned}$$

The second estimate now follows.

To prove the final inequality, (5.6) and (5.10) show that $\kappa(a) \ge \kappa(T_1)$ and $\chi(T_1) \ge \chi(b)$ for $0 \le a \le T_1 \le b \le T$. More explicitly,

$$\frac{\Delta s_2(T_1)}{s_2(T_1)} \le \frac{\Delta s_2(a)}{s_2(a)} \quad\text{and}\quad \frac{\Delta s_2(b)}{s_1(b)} \le \frac{\Delta s_2(T_1)}{s_2(T_1)}.$$

Recalling that $s_2(T_1) = s_1(T_1)$, combining the above inequalities gives us

$$\Delta s_2(b) \le \frac{s_1(b)\Delta s_2(T_1)}{s_2(T_1)} \le \frac{s_1(b)\Delta s_2(a)}{s_2(a)}.$$

By the assumption that $|\Delta s_1| \le \alpha\Delta s_2$, we get

$$\Delta s_2 \le \|\Delta s\| \le \sqrt{1+\alpha^2}\,\Delta s_2,$$

and combining this inequality with the preceding one gives us the final inequality in the statement of the lemma. □

Our final estimate concerns the size of the angles between tangent vectors in the unstable cones near the singularities. This will be used in examining the distance between the unstable subspaces of nearby points in neighborhoods of the singularities.

Recall that the neighborhood $\widetilde{\mathcal{U}}_1$ of S is given by $\widetilde{\mathcal{U}}_1 = \bigcup_{k=1}^{m}\phi_k^{-1}(D_{\widetilde{r}_1})$. For $x \in \widetilde{\mathcal{U}}_1$, define

$$\gamma(x) = \max_{\substack{v,w\in K^+(x)\\ \|v\|=\|w\|=1}}\left\{\frac{\angle(Dg_x v, Dg_x w)}{\angle(v,w)}\right\} \tag{5.11}$$

and denote $\gamma_j(x) = \gamma(g^j(x))$ for $j \ge 0$.

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

LEMMA 5.6. *For every $x \in \widetilde{\mathcal{U}}_1$ with $g^j(x)$ in the same component of $\widetilde{\mathcal{U}}_1$ for $j = 0, \ldots, k$, we have*

$$\prod_{j=0}^{k-1} \gamma_j(x) \leq (1 + C_0 s_2(0)^{(2p-4)/p} k)^{-p/(p-2)},$$

where C_0 is the constant from Lemma 5.4.

Proof. Denote $z = \Phi_{kj}(\phi_k(x)) = (s_1(0), s_2(0))$, so that

$$(\Phi_{kj} \circ \phi_k)(g^j(x)) = (s_1(j), s_2(j)).$$

Consider a tangent vector $v = (\zeta_1, \zeta_2)$ in $\mathbb{C}$ along a trajectory of the vector field (3.1). Reparametrizing $\eta = \zeta_2/\zeta_1$ with respect to s_1 instead of t along this curve, equation (3.3) implies

$$\frac{d\eta}{ds_1} = \frac{d\eta}{dt}\left(\frac{ds_1}{dt}\right)^{-1} = -2\bigg((1+\eta^2)s_2 \frac{\dot{\Psi}_p(u)}{\Psi_p(u)} + \bigg(\frac{1}{s_1}\dot{\Psi}_p(u) + \frac{s_1^2 + s_2^2}{s_1}\frac{\dot{\Psi}_p(u)}{\Psi_p(u)}\bigg)\eta\bigg).$$

For $i = 1, 2$, let $\eta_i(s_1) = \eta_i(s_1, s_1(j), \eta_i^0)$ be a solution to this differential equation with initial condition $\eta_i(s_1(j)) = \eta_i^0$. Then

$$\frac{d}{dt}(\eta_1 - \eta_2) = -2\frac{1}{s_1}\bigg(1 + \frac{\dot{\Psi}_p(u)}{\Psi_p(u)}(s_1^2 + s_2^2 + s_1 s_2(\eta_1 + \eta_2))\bigg)(\eta_1 - \eta_2).$$

If $(\xi_1, \xi_2) = D(\Phi_{kj} \circ \phi_k)_z^{-1}(\zeta_1, \zeta_2) \in K^+(x)$, then $|\eta_i| < \alpha < 1$ for $i = 1, 2$ (see Lemma 3.4), so $\eta_1 + \eta_2 > -2$. Positivity of Ψ_p and $\dot{\Psi}_p$ now yields

$$\frac{d}{dt}(\eta_1 - \eta_2) \leq -2\frac{1}{s_1}\bigg(1 + \frac{\dot{\Psi}_p(u)}{\Psi_p(u)}(s_1 - s_2)^2\bigg)(\eta_1 - \eta_2),$$

and so by Grönwall's inequality,

$$\begin{aligned}
&|\eta_1(s_1(j+1)) - \eta_2(s_1(j+1))| \\
&\qquad \leq |\eta_1^0 - \eta_2^0| \exp\bigg(-2\int_{s_1(j)}^{s_1(j+1)} \frac{1}{s_1}\bigg(1 + \frac{\dot{\Psi}_p(u)}{\Psi_p(u)}(s_1 - s_2)^2\bigg)ds_1\bigg) \\
&\qquad \leq |\eta_1^0 - \eta_2^0| \exp\bigg(-2\int_{s_1(j)}^{s_1(j+1)} \frac{ds_1}{s_1}\bigg) \\
&\qquad = |\eta_1^0 - \eta_2^0|\bigg(\frac{s_1(j)}{s_1(j+1)}\bigg)^2 \\
&\qquad = |\eta_1^0 - \eta_2^0|\bigg(\frac{s_2(j+1)}{s_2(j)}\bigg)^2,
\end{aligned}$$

where the final equality follows from the fact that the trajectories lie on hyperbolas, and so the product $s_1 s_2$ is constant. Observe that if $v = (v_1, v_2)$ and $w = (w_1, w_2)$ are two vectors with $\eta_v = v_2/v_1$ and $\eta_w = w_2/w_1$, then

$$\angle(v, w) = |\arctan \eta_v - \arctan \eta_w|,$$

and so by concavity of $\eta \mapsto \arctan \eta$ and conformality of the coordinate map $\Phi_{kj} \circ \phi_k$,

$$\gamma_j(x) \leq \max_{\eta_1,\eta_2} \left\{ \frac{|\eta_1(s_1(j+1), s_1(j), \eta_1^0) - \eta_2(s_1(j+1), s_1(j), \eta_2^0)|}{|\eta_1^0 - \eta_2^0|} \right\}$$
$$\leq \left(\frac{s_2(j+1)}{s_2(j)} \right)^2.$$

It follows that

$$\prod_{j=0}^{k-1} \gamma_j(x) \leq \left(\frac{s_2(k)}{s_2(0)} \right)^2.$$

The desired result now follows from inequality (b) in Lemma 5.4, since by hypothesis $g^j(x)$ is in the same component of $\widetilde{U}_1$, hence $G_p^j(z) \in D_{\widetilde{r}_1}$ for $0 \leq j \leq k$. □

6. *Thermodynamics of Young diffeomorphisms*

Given a $C^{1+\alpha}$ diffeomorphism f on a compact Riemannian manifold M, we call an embedded C^1 disc $\gamma \subset M$ an *unstable disc* (respectively, *stable disc*) if for all $x, y \in \gamma$, we have $d(f^{-n}(x), f^{-n}(y)) \to 0$ (respectively, $d(f^n(x), f^n(y)) \to 0$) as $n \to +\infty$. A collection of embedded C^1 discs $\Gamma = \{\gamma_i\}_{i \in \mathcal{I}}$ is a *continuous family of unstable discs* if there is a Borel subset $K^s \subset M$ and a homeomorphism $\Phi : K^s \times D^u \to \bigcup_i \gamma_i$, where $D^u \subset \mathbb{R}^d$ is the closed unit disc for some $d < \dim M$, satisfying the following assertions.

- The assignment $x \mapsto \Phi|_{\{x\} \times D^u}$ is a continuous map from K^s to the space of C^1 embeddings $D^u \hookrightarrow M$, and this assignment can be extended to the closure $\overline{K^s}$.
- For every $x \in K^s$, $\gamma = \Phi(\{x\} \times D^u)$ is an unstable disc in Γ.

Thus the index set $\mathcal{I}$ may be taken to be $K^s \times \{0\} \subset K^s \times D^u$. We define *continuous families of stable discs* analogously.

A subset $\Lambda \subset M$ has *hyperbolic product structure* if there is a continuous family $\Gamma^u = \{\gamma_i^u\}_{i \in \mathcal{I}}$ of unstable discs and a continuous family $\Gamma^s = \{\gamma_j^s\}_{j \in \mathcal{J}}$ of stable discs such that:

- $\dim \gamma_i^u + \dim \gamma_j^s = \dim M$ for all i, j;
- the unstable discs are transversal to the stable discs, with an angle uniformly bounded away from 0;
- each unstable disc intersects each stable disc in exactly one point;
- $\Lambda = (\bigcup_i \gamma_i^u) \cap (\bigcup_j \gamma_j^s)$.

A subset $\Lambda_0 \subset \Lambda$ with hyperbolic product structure is an *s-subset* if the continuous family of unstable discs defining Λ_0 is the same as the continuous family of unstable discs for Λ, and the continuous family of stable discs defining Λ_0 is a subfamily Γ_0^s of the continuous family of stable discs defining Γ_0. In other words, if $\Lambda_0 \subset \Lambda$ has hyperbolic product structure generated by the families of stable and unstable discs given by Γ_0^s and Γ_0^u, then Λ_0 is an s-subset if $\Gamma_0^s \subseteq \Gamma^s$ and $\Gamma_0^u = \Gamma^u$. A *u-subset* is defined analogously.

Definition 6.1. A $C^{1+\alpha}$ diffeomorphism $f : M \to M$, with M a compact Riemannian manifold, is a *Young's diffeomorphism* if the following conditions are satisfied.

(Y1) There exists $\Lambda \subset M$ (called the *base*) with hyperbolic product structure, a countable collection of continuous subfamilies $\Gamma_i^s \subset \Gamma^s$ of stable discs, and positive integers τ_i, $i \in \mathbb{N}$, such that the s-subsets

$$\Lambda_i^s := \bigcup_{\gamma \in \Gamma_i^s} (\gamma \cap \Lambda) \subset \Lambda$$

are pairwise disjoint and satisfy:

(a) (*invariance*) for $x \in \Lambda_i^s$,

$$f^{\tau_i}(\gamma^s(x)) \subset \gamma^s(f^{\tau_i}(x)) \quad \text{and} \quad f^{\tau_i}(\gamma^u(x)) \supset \gamma^u(f^{\tau_i}(x)),$$

where $\gamma^{u,s}(x)$ denotes the (un)stable disc containing x; and

(b) (*Markov property*) $\Lambda_i^u := f^{\tau_i}(\Lambda_i^s)$ is a u-subset of Λ such that for $x \in \Lambda_i^s$,

$$f^{-\tau_i}(\gamma^s(f^{\tau_i}(x)) \cap \Lambda_i^u) = \gamma^s(x) \cap \Lambda \quad \text{and}$$
$$f^{\tau_i}(\gamma^u(x) \cap \Lambda_i^s) = \gamma^u(f^{\tau_i}(x)) \cap \Lambda.$$

(Y2) For $\gamma^u \in \Gamma^u$, we have

$$\mu_{\gamma^u}(\gamma^u \cap \Lambda) > 0 \quad \text{and} \quad \mu_{\gamma^u}(\mathrm{cl}((\Lambda \setminus \textstyle\bigcup_i \Lambda_i^s) \cap \gamma^u)) = 0,$$

where μ_{γ^u} is the induced Riemannian leaf volume on γ^u and $\mathrm{cl}(A)$ denotes the closure of A in M for $A \subseteq M$.

(Y3) There is $a \in (0, 1)$ so that for any $i \in \mathbb{N}$, we have:

(a) for $x \in \Lambda_i^s$ and $y \in \gamma^s(x)$,

$$d(F(x), F(y)) \leq a d(x, y);$$

(b) for $x \in \Lambda_i^s$ and $y \in \gamma^u(x) \cap \Lambda_i^s$,

$$d(x, y) \leq a d(F(x), F(y)),$$

where $F : \bigcup_i \Lambda_i^s \to \Lambda$ is the *induced map* defined by

$$F|_{\Lambda_i^s} := f^{\tau_i}|_{\Lambda_i^s}.$$

(Y4) Denote $J^u F(x) = \det |DF|_{E^u(x)}|$. There exist $c > 0$ and $\kappa \in (0, 1)$ such that:

(a) for all $n \geq 0$, $x \in F^{-n}(\bigcup_i \Lambda_i^s)$ and $y \in \gamma^s(x)$, we have

$$\left| \log \frac{J^u F(F^n(x))}{J^u F(F^n(y))} \right| \leq c\kappa^n;$$

(b) for any $i_0, \ldots, i_n \in \mathbb{N}$ with $F^k(x), F^k(y) \in \Lambda_{i_k}^s$ for $0 \leq k \leq n$ and $y \in \gamma^u(x)$, we have

$$\left| \log \frac{J^u F(F^{n-k}(x))}{J^u F(F^{n-k}(y))} \right| \leq c\kappa^k.$$

(Y5) There is some $\gamma^u \in \Gamma^u$ such that

$$\sum_{i=1}^{\infty} \tau_i \mu_{\gamma^u}(\Lambda_i^s) < \infty.$$

We say that the tower satisfies the *arithmetic condition* if the greatest common divisor of the integers $\{\tau_i\}$ is 1.

We use the following result to discuss thermodynamics of Young's diffeomorphisms, which was originally presented as Proposition 4.1 and Remark 4 in [13].

PROPOSITION 6.2. *Let $f : M \to M$ be a $C^{1+\alpha}$ diffeomorphism of a compact smooth Riemannian manifold M satisfying conditions (Y1)–(Y5), and assume τ is the first return time to the base of the tower. Then the following assertions hold.*

(1) *There exists an equilibrium measure μ_1 for the potential φ_1, which is the unique SRB measure.*

(2) *Assume that for some constants $C > 0$ and $0 < h < h_{\mu_1}(f)$, with $h_{\mu_1}(f)$ the metric entropy, we have*

$$S_n := \#\{\Lambda_i^s : \tau_i = n\} \leq Ce^{hn}$$

Define

$$\log \lambda_1 = \sup_{i \geq 1} \sup_{x \in \Lambda_i^s} \frac{1}{\tau_i} \log |J^u F(x)| \leq \max_{x \in M} \log |J^u f(x)| \tag{6.1}$$

and

$$t_0 = \frac{h - h_{\mu_1}(f)}{\log \lambda_1 - h_{\mu_1}(f)}. \tag{6.2}$$

Then, for every $t \in (t_0, 1)$, there exists a measure $\mu_t \in \mathcal{M}(f, Y)$, where $Y = \{f^k(x) : x \in \bigcup \Lambda_i^s,\ 0 \leq k \leq \tau(x) - 1\}$, which is a unique equilibrium measure for the potential φ_t.

(3) *Assume that the tower satisfies the arithmetic condition, and that there is $K > 0$ such that for every $i \geq 0$, every $x, y \in \Lambda_i^s$, and any $j \in \{0, \ldots, \tau_i\}$,*

$$d(f^j(x), f^j(y)) \leq K \max\{d(x, y), d(F(x), F(y))\}. \tag{6.3}$$

Then, for every $t_0 < t < 1$, the measure μ_t has exponential decay of correlations and satisfies the central limit theorem with respect to a class of functions which contains all Hölder continuous functions on M.

7. *Young towers over pseudo-Anosov diffeomorphisms*

Our argument that smooth pseudo-Anosov diffeomorphisms are Young diffeomorphisms requires the construction of a hyperbolic tower on pseudo-Anosov homeomorphisms first. We begin this section by constructing this hyperbolic tower, taking an element of the Markov partition of the pseudo-Anosov homeomorphism as the base of the tower.

We assume that our pseudo-Anosov homeomorphism f admits only one singularity; the analysis follows similarly with more singularities, but the notation becomes unwieldy due to the different numbers of prongs at each singularity. Therefore we state without proof that the arguments of this section imply that pseudo-Anosov diffeomorphisms admitting multiple singularities are also Young diffeomorphisms. An example of a pseudo-Anosov homeomorphism of the genus-2 torus admitting only one singularity may be found in [11].

By Proposition 2.9, a pseudo-Anosov surface homeomorphism $f : M \to M$ admits a Markov partition of arbitrarily small diameter. Let $\widetilde{\mathcal{P}}$ be such a Markov partition, and let

$\widetilde{P} \in \widetilde{\mathcal{P}}$ be an element of the Markov partition contained in a chart U_1 not intersecting with the chart U_0 of the singularity x_0. For $x \in \widetilde{P}$, let $\widetilde{\gamma}^s(x)$ and $\widetilde{\gamma}^u(x)$ respectively be the connected component of the intersection of the stable and unstable leaves with $\widetilde{P}$ containing x.

Let $\widetilde{\tau}(x)$ be the first return time of x to $\mathrm{Int}\widetilde{P}$ for $x \in \widetilde{P}$. For x with $\widetilde{\tau}(x) < \infty$, define

$$\widetilde{\Lambda}^s(x) = \bigcup_{y \in \widetilde{U}^u(x) \setminus \widetilde{A}^u(x)} \widetilde{\gamma}^s(y),$$

where $\widetilde{U}^u(x) \subseteq \widetilde{\gamma}^u(x)$ is an interval containing x, open in the induced topology of $\widetilde{\gamma}(x)$, and $\widetilde{A}^u(x) \subset \widetilde{U}^u(x)$ is the set of points that either lie on the boundary of the Markov partition, or never return to $\widetilde{P}$. One can show that the leaf volume of $\widetilde{A}^u(x)$ is 0, so that for each $y \in \widetilde{\Lambda}^s(x)$, the leaf volume of $\widetilde{\gamma}(y) \cap \widetilde{\Lambda}^s(x)$ is positive. We further choose our interval $U^u(x)$ so that:

- for $y \in \widetilde{\Lambda}^s(x)$, we have $\widetilde{\tau}(y) = \widetilde{\tau}(x)$; and
- for $y \in \widetilde{P}$ with $\widetilde{\tau}(x) = \widetilde{\tau}(y)$, we have $y \in \widetilde{\Lambda}(z)$ for some $z \in \widetilde{P}$.

One can show that the image under $\widetilde{f}^{\widetilde{\tau}(x)}$ of $\widetilde{\Lambda}^s(x)$ is a u-subset containing $\widetilde{f}^{\widetilde{\tau}(x)}(x)$, and that, for $x, y \in \widetilde{P}$ with finite return time, $\widetilde{\Lambda}^s(x)$ and $\widetilde{\Lambda}^s(y)$ are either disjoint or coinciding. As discussed in [**13**], this gives us a countable collection of disjoint sets $\widetilde{\Lambda}_i^s$ and numbers $\widetilde{\tau}_i$ for which the pseudo-Anosov homeomorphism $f : M \to M$ is a Young map, with s-sets $\widetilde{\Lambda}_i^s$, inducing times $\widetilde{\tau}_i$, and tower base

$$\widetilde{\Lambda} := \bigcup_{i=1}^{\infty} \mathrm{cl}(\widetilde{\Lambda}_i^s).$$

In the following theorem, conditions (Y1′)–(Y5′) are virtually identical to conditions (Y1)–(Y5) in Definition 6.1. They are reprinted in the following theorem because pseudo-Anosov homeomorphisms are not true diffeomorphisms, and thus by definition cannot satisfy conditions (Y1)–(Y5). However, analogous conditions may be established for pseudo-Anosov homeomorphisms, and these conditions will be used to show that globally smooth realizations of pseudo-Anosov diffeomorphisms (which are true diffeomorphisms) are Young diffeomorphisms.

THEOREM 7.1. *The set $\widetilde{\Lambda}$ defined above for the pseudo-Anosov homeomorphism $f : M \to M$ satisfies the following conditions.*

(Y1′) *$\widetilde{\Lambda}$ has hyperbolic product structure, and the sets $\{\widetilde{\Lambda}_i^s\}_{i \in \mathbb{N}}$ are pairwise disjoint s-subsets and satisfy:*

(a) *(*invariance*) for $x \in \widetilde{\Lambda}_i^s$,*

$$f^{\tau_i}(\gamma^s(x)) \subset \gamma^s(f^{\tau_i}(x)) \quad \text{and} \quad f^{\tau_i}(\gamma^u(x)) \supset \gamma^u(f^{\tau_i}(x)),$$

where $\gamma^{u,s}(x)$ denotes the (un)stable disc containing x; and

(b) *(*Markov property*) $\widetilde{\Lambda}_i^u := f^{\tau_i}(\Lambda_i^s)$ is a u-subset of $\widetilde{\Lambda}$ such that for $x \in \widetilde{\Lambda}_i^s$,*

$$f^{-\tau_i}(\gamma^s(f^{\tau_i}(x)) \cap \widetilde{\Lambda}_i^u) = \gamma^s(x) \cap \widetilde{\Lambda} \quad \text{and}$$
$$f^{\tau_i}(\gamma^u(x) \cap \widetilde{\Lambda}_i^s) = \gamma^u(f^{\tau_i}(x)) \cap \widetilde{\Lambda}.$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

(Y2′) *For $\gamma^u \in \Gamma^u$, we have*

$$\nu^s(\gamma^u \cap \widetilde{\Lambda}) > 0 \quad \textit{and} \quad \nu^s(\mathrm{cl}((\widetilde{\Lambda} \setminus \textstyle\bigcup_i \widetilde{\Lambda}_i^s) \cap \gamma^u)) = 0,$$

where ν^s is the transversal invariant measure with respect to the stable foliation $\mathcal{F}^s$ for f.

(Y3′) *There is $a \in (0, 1)$ so that for any $i \in \mathbb{N}$, we have:*

(a) *for $x \in \widetilde{\Lambda}_i^s$ and $y \in \gamma^s(x)$,*

$$d^s(F(x), F(y)) \leq a d^s(x, y);$$

(b) *for $x \in \widetilde{\Lambda}_i^s$ and $y \in \gamma^u(x) \cap \widetilde{\Lambda}_i^s$,*

$$d^u(x, y) \leq a d^u(F(x), F(y)),$$

where $F : \bigcup_i \widetilde{\Lambda}_i^s \to \widetilde{\Lambda}$ is the induced map *defined by*

$$F|_{\widetilde{\Lambda}_i^s} := f^{\tau_i}|_{\widetilde{\Lambda}_i^s}$$

and d^s and d^u are the distances in the stable and unstable leaves of the foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ in $\widetilde{P}$, given respectively by ν^u and ν^s.

(Y4′) *Denote $J^u F(x) = \det|DF|_{E^u(x)}|$. There exist $c > 0$ and $\kappa \in (0, 1)$ such that:*

(a) *for all $n \geq 0$, $x \in F^{-n}(\bigcup_i \widetilde{\Lambda}_i^s)$, and $y \in \gamma^s(x)$, we have*

$$\left| \log \frac{J^u F(F^n(x))}{J^u F(F^n(y))} \right| \leq c\kappa^n;$$

(b) *for any $i_0, \ldots, i_n \in \mathbb{N}$ with $F^k(x), F^k(y) \in \widetilde{\Lambda}_{i_k}^s$ for $0 \leq k \leq n$ and $y \in \gamma^u(x)$, we have*

$$\left| \log \frac{J^u F(F^{n-k}(x))}{J^u F(F^{n-k}(y))} \right| \leq c\kappa^k.$$

(Y5′) *There is some $\gamma^u \in \widetilde{\Gamma}^u$ such that*

$$\sum_{i=1}^{\infty} \tau_i \nu^s(\widetilde{\Lambda}_i^s \cap \gamma^u) < \infty.$$

Proof. Properties (Y1′), (Y3′), and (Y4′) all follow from Proposition 2.7. Property (Y2′) follows because $x \in \mathrm{cl}((\Lambda \setminus \bigcup_i \Lambda_i^s) \cap \gamma^u)$ implies that either $x \in \partial P$ or $\tau(x) = \infty$, both of which happen on a set of Lebesgue measure 0 (and the smooth measure for pseudo-Anosov homeomorphisms has density with respect to Lebesgue measure). And since τ is a first return time, (Y5′) follows from Kac's theorem. □

The next lemma gives a bound on the number S_n of distinct s-subsets $\widetilde{\Lambda}_i^s$ with a given inducing time $\widetilde{\tau}_i = n$. Since the pseudo-Anosov homeomorphism f is topologically conjugate to the smooth realization g, this will eventually give us an analogous bound on the number of distinct s-subsets for the base of the tower for g. (See condition (2) of Proposition 6.2.)

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

LEMMA 7.2. *There exists $h < h_{\text{top}}(f)$ such that $S_n \le e^{hn}$, where S_n is the number of s-sets $\widetilde{\Lambda}_i^s$ with inducing time $\widetilde{\tau}_i = n$.*

Proof. The proof is analogous to [13, Lemma 6.1], since pseudo-Anosov homeomorphisms admit finite Markov partitions. □

Let $H : M \to M$ be the conjugacy map so that $g \circ H = H \circ f$, and let $\mathcal{P} = H(\widetilde{\mathcal{P}})$, $P = H(\widetilde{P})$. Then $\mathcal{P}$ is a Markov partition for the pseudo-Anosov diffeomorphism (M, g), and P is a partition element. By continuity of H, we may assume the elements of $\mathcal{P}$ have arbitrarily small diameter. Furthermore, let $\Lambda = H(\widetilde{\Lambda})$. Then Λ has direct hyperbolic product structure with full length stable and unstable curves $\gamma^s(x) = H(\widetilde{\gamma}^s(x))$ and $\gamma^u(x) = H(\widetilde{\gamma}^u(x))$. Then $\Lambda_i^s = H(\widetilde{\Lambda}_i^s)$ are s-sets and $\Lambda_i^u = H(\widetilde{\Lambda}_i^u) = g^{\tau_i}(\Lambda_i^s)$, where $\tau_i = \widetilde{\tau}_i$ for each i, and $\tau(x) = \tau_i$ whenever $x \in \Lambda_i^s$.

Recall that $\mathcal{U}_0 = \bigcup_{k=1}^m \phi_k^{-1}(D_{r_0})$. If there is only one singularity, then $\mathcal{U}_0 = \phi_0^{-1}(D_{r_0})$. Given $Q > 0$, we can take r_0 in the construction of g to be so small and refine the partition $\widetilde{\mathcal{P}}$ so that the partition element $\widetilde{P}$ (and hence P) may be chosen so that

$$g^n(x) \notin \mathcal{U}_0 \quad \text{for any } 0 \le n \le Q, \tag{7.1}$$

and any x so that either $x \in P$, or $x \notin \mathcal{U}_0$ while $g^{-1}(x) \in \mathcal{U}_0$.

We now prove that the set $\Lambda = H(\widetilde{\Lambda})$ constructed above is the base of a Young tower on M for the diffeomorphism g. Properties (Y1), (Y2), and (Y5) are straightforward to verify. Our strategy in proving these conditions, along with (Y3), is similar to that used in [13], but we restate it here for the reader's convenience. The main difference between the argument used for these pseudo-Anosov diffeomorphisms and the Katok map comes in proving (Y4), where we use a local trivialization of our surface M as opposed to the universal cover of $\mathbb{T}^2$ by $\mathbb{R}^2$.

THEOREM 7.3. *The collection of s-subsets $\Lambda_i^s = H(\widetilde{\Lambda}_i^s)$ satisfies conditions (Y1)–(Y5), making the smooth pseudo-Anosov diffeomorphism $g : M \to M$ a Young diffeomorphism.*

Proof. Condition (Y1) follows from the corresponding properties of the pseudo-Anosov homeomorphism f since H is a topological conjugacy. The fact that $\mu_{\gamma^u}(\gamma^u \cap \Lambda) > 0$ follows from the corresponding property for the $\widetilde{\gamma}^u$ leaves. Suppose $x \in \mathrm{cl}((\Lambda \setminus \bigcup_i \Lambda_i^s) \cap \gamma^u)$. Then either x lies on the boundary of the Markov partition element P, or $\tau(x) = \infty$, and since both the Markov partition boundary and the set of $x \in P$ with $\tau(x) = \infty$ are Lebesgue null, we get condition (Y2). Condition (Y5) follows from Kac's formula, since the inducing times are first return times to the base of the tower.

To prove condition (Y3), define the *itinerary* $\mathcal{I}(x) = \{0 = n_0 < n_1 < \cdots < n_{2L+1} = \tau(x)\} \subset \mathbb{Z}$ of a point $x \in \Lambda$, with $L = L(x)$, so that $g^k(x) \in \mathcal{U}_0$ if and only if $n_{2j-1} \le k < n_{2j}$ for $j \ge 1$. Assume Λ is small enough so that $\mathcal{I}(x) = \mathcal{I}(y)$ whenever $y \in \gamma(x) \subset \Lambda$.

Let $x \in \Lambda_i^s$, $y \in \gamma^s(x) \subset \Lambda_i^s$. Denote $x_n = g^n(x)$ and $y_n = g^n(y)$. Note that $\gamma^s(x) \subset \mathcal{F}^s(x)$. By invariance of the stable and unstable measured foliations $\mathcal{F}^s$ and $\mathcal{F}^u$, y_n lies on the stable curve $\mathcal{F}^s(x_n)$ through x_n for every $n \ge 1$. For $n_{2j} \le n < n_{2j+1}$, $T_{x_n}\mathcal{F}^s(x_n) = E_{x_n}^s$ lies inside C_x^-; in fact one can show that $\mathcal{F}^s(x_n)$ is an admissible manifold. Thus the segment of $\mathcal{F}^s(x_n)$ joining x_n and y_n expands uniformly under the homeomorphism f^{-1}.

Due to our choice of the number Q, there is a number $\beta \in (0, 1)$ such that

$$d(x_{n_{2j+1}}, y_{n_{2j+1}}) \leq \beta^{n_{2j+1}-n_{2j}} d(x_{n_{2j}}, y_{n_{2j}}) \leq \beta^{Q} d(x_{n_{2j}}, y_{n_{2j}}). \tag{7.2}$$

Now we consider $n_{2j-1} \leq n < n_{2j}$. Let $[m_j^1, m_j^2] \subseteq [n_{2j-1}, n_{2j} - 1]$ be the largest interval (possibly empty) with x_n in the closure of $\widetilde{\mathcal{U}}_1 = \phi_0^{-1}(D_{\widetilde{r}_1}(0))$ for every $n \in [m_j^1, m_j^2]$. By virtue of Lemma 5.2, there is a uniform $T > 0$ with $m_j^1 - n_{2j-1} \leq T$ and $n_{2j} - m_j^2 \leq T$. Thus there is a constant $C > 0$ so that

$$d(x_{m_j^1}, y_{m_j^1}) \leq C d(x_{n_{2j-1}}, y_{n_{2j-1}}) \quad \text{and} \quad d(x_{n_{2j}}, y_{n_{2j}}) \leq C d(x_{m_j^2}, y_{m_j^2}). \tag{7.3}$$

Now, let $s(t)$ and $\widetilde{s}(t)$ be solutions to equation (3.1) with $s(0) = x_{m_j^1}$ and $\widetilde{s}(0) = y_{m_j^1}$. Assumption (1) of Lemma 5.5 is satisfied since y_n lies in the stable cone of x_n for every n, and assumption (2) can be assured if our choice of r_0 in the slowdown construction of the pseudo-Anosov diffeomorphism is chosen to be sufficiently small. So by the final inequality of this lemma, letting $a = m_j^1$ and $b = m_j^2$, we get

$$\|\Delta s(m_j^2)\| \leq \sqrt{1 + \alpha^2} \frac{s_1(m_j^2)}{s_2(m_j^1)} \|\Delta s(m_j^1)\|.$$

Let $\Delta_{kj} s(t) = \Phi_{kj}^{-1}(\widetilde{s}(t)) - \Phi_{kj}^{-1}(s(t))$. Because Φ_{kj} is uniformly bounded above and below, there is a constant $K > 0$ such that, for every t for which $\widetilde{s}(t)$ and $s(t)$ are defined,

$$K^{-1}\|\Delta_{kj} s(t)\| \leq \|\Delta s(t)\| \leq K\|\Delta_{kj} s(t)\|, \tag{7.4}$$

and since the Riemannian metric in $\mathcal{U}_0$ is given in coordinates by $dt_1^2 + dt_2^2 = (\Phi_{kj}^{-1})^*(ds_1^2 + ds_2^2)$, we get $\|\Delta_{kj} s(n)\| = d(x_n, y_n)$ for $n \in [m_j^1, m_j^2]$. Therefore, combining this observation with (7.4), (7.2), (7.3), and (5.1), we get

$$\begin{aligned} d(x_{n_{2j}}, y_{n_{2j}}) &\leq C K^2 \sqrt{1 + \alpha^2} \frac{s_1(m_j^2)}{s_2(m_j^1)} d(x_{m_j^1}, y_{m_j^1}) \\ &\leq C^2 K^2 \sqrt{1 + \alpha^2} \frac{s_1(m_j^2)}{s_2(m_j^1)} d(x_{n_{2j-1}}, y_{n_{2j-1}}) \\ &\leq C^2 K^2 \beta^Q \sqrt{1 + \alpha^2} \frac{s_1(m_j^2)}{s_2(m_j^1)} d(x_{n_{2j-2}}, y_{n_{2j-2}}). \end{aligned}$$

Since $s_1(m_j^2)$ and $s_2(m_j^1)$ are each of order r_0, their quotient is uniformly bounded, so assuming Q is sufficiently large, there is a $0 < \theta_1 < 1$ for which

$$d(x_{n_{2j}}, y_{n_{2j}}) \leq \theta_1 d(x_{n_{2j-2}}, y_{n_{2j-2}}) \tag{7.5}$$

and a similar bound holds for odd indices of the itinerary. It follows that

$$d(g^{\tau(x)}(x), g^{\tau(x)}(y)) \leq \theta_1^L d(x, y),$$

where L is determined by the itinerary $\mathcal{I}(x)$. Condition (Y3)(a) follows, and (Y3)(b) follows by the same argument applied to g^{-1}.

To prove condition (Y4), we prove condition (Y4)(a) and note that (Y4)(b) can be proved similarly by considering g^{-1} instead of g. We use the following general statement, originally presented as [13, Lemma 6.3]. $\square$

LEMMA 7.4. *Let* $\{A_n\}$, $\{B_n\}$, $0 \leq n \leq N$, *be two collections of linear transformations of* $\mathbb{R}^d$. *Given a subspace* $E \subset \mathbb{R}^d$, *let* $K = K(E, \theta)$ *denote the cone of angle* θ *around* E. *Assume the subspace* E *is such that:*

(a) $A_n(K) \subset K$ *for all* n;

(b) *there are* $\gamma_n > 0$ *such that for each* n, *and for any unit vectors* $v, w \in K$,
$$\angle(A_n v, A_n w) \leq \gamma_n \angle(v, w);$$

(c) *there are* $d > 0$ *and* $\delta_n > 0$ *such that for each* $n \geq 0$, *and every* $v \in K$,
$$\|A_n v - B_n v\| \leq d\delta_n \|A_n v\|;$$

(d) *there is* $c > 0$ *independent of* n *such that for every* $v \in K$,
$$\|A_n v\| \geq c\|v\|.$$

Then there is a $C > 0$, *independent of the choice of linear transformations* $\{A_n\}$ *and* $\{B_n\}$, *such that for every* $v, w \in K$,
$$\left| \log \frac{\|\prod_{n=0}^{N} A_n v\|}{\|\prod_{n=0}^{N} B_n w\|} \right| \leq C\left(d \sum_{n=0}^{N} \delta_n + \angle(v, w) \sum_{n=0}^{N} \prod_{k=0}^{n} \gamma_k \right). \tag{7.6}$$

Let $x \in P$ with $N := \tau(x) - 1 < \infty$, and let $y \in \gamma^s(x) \subset P$. For each $n \geq 0$, once again let $x_n = g^n(x)$ and $y_n = g^n(y)$, and in each tangent space $T_{x_n}M$, let $K_n^+ = K^+(x_n) \subset T_{x_n}M$ denote the cone of angle $\arctan \alpha$ around $E^u(x_n)$ described in Lemma 3.4. By this lemma, the sequence of cones $\{K_n^+\}$ is invariant under Dg. For each n, denote $\widetilde{A}_n = Dg_{x_n} : T_{x_n}M \to T_{x_{n+1}}M$ and $\widehat{B}_n = Dg_{y_n} : T_{y_n}M \to T_{y_{n+1}}M$. Further, since y_n lies on the stable leaf of x_n for all n, let $P_n : T_{y_n}M \to T_{x_n}M$ denote parallel translation along the segment of the stable leaf connecting y_n to x_n, and denote $\widetilde{B}_n = P_{n+1} \circ \widehat{B}_n \circ P_n^{-1} : T_{x_n}M \to T_{x_{n+1}}M$. Using the orthonormal coordinates (ξ_1, ξ_2) for $T_{x_n}M$ defined previously, so that ξ_1 denotes the unstable direction and ξ_2 denotes the stable direction (see the discussion preceding Proposition 3.3), we may isometrically identify each tangent space $T_{x_n}M$ with $\mathbb{R}^2$ with the Euclidean metric. Call this isometry $\Xi_n : T_{x_n}M \to \mathbb{R}^2$, and denote $A_n = \Xi_{n+1} \circ \widetilde{A}_n \circ \Xi_n^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ and $B_n = \Xi_{n+1} \circ \widetilde{B}_n \circ \Xi_n^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$. Also let $K = \Xi_n(K_n^+) \subset \mathbb{R}^2$. Since Ξ_n is an isometry and K_n^+ is a cone of angle $\arctan \alpha$ for each n, K is independent of n and is thus well defined. Finally, define the numbers $d = d(x, y)$, as well as
$$\gamma_n = \max_{\substack{v,w \in K \\ \|v\|=\|w\|=1}} \left\{ \frac{\angle(A_n v, A_n w)}{\angle(v, w)} \right\} \quad \text{and} \quad \delta_n = \frac{1}{d} \max_{v \in K \setminus \{0\}} \left\{ \frac{\|A_n v - B_n v\|}{\|A_n v\|} \right\}$$
for each $n \geq 0$.

The final step in proving our pseudo-Anosov diffeomorphism g is a Young's diffeomorphism relies on the following technical lemma. Its proof is somewhat similar to the proof of [13, Lemma 6.4], but requires some modifications related to the subtle differences in the slowdown function used in the Katok map as opposed to our pseudo-Anosov diffeomorphism g, as well as to the fact that the universal cover of a surface that is not a torus is not $\mathbb{R}^2$.

LEMMA 7.5. *The linear operators A_n and B_n, as well as the cone K, all satisfy the conditions of Lemma 7.4 using γ_n, δ_n, d, and $N = \tau(x) - 1$ defined above. Furthermore, there are constants $\widetilde{C} > 0$ and $0 < \theta_2 < 1$, independent of $x \in P$, such that*

$$\sum_{n=0}^{\tau(x)-1} \delta_n < \widetilde{C}, \quad \sum_{n=0}^{\tau(x)-1} \prod_{k=0}^{n} \gamma_k < \widetilde{C}, \quad \text{and} \quad \prod_{n=0}^{\tau(x)-1} \gamma_n < \theta_2.$$

Proof of Lemma 7.5. Condition (a) of Lemma 7.4 follows from the definition of A_n, the invariance of the cone family K_n^+ under $\widetilde{A}_n$, and the fact that $\Xi_n : T_{x_n}M \to \mathbb{R}^2$ is an isometry for every n. Conditions (b) and (c) of Lemma 7.4 follow from the definitions of γ_n and δ_n. Finally, condition (d) of Lemma 7.4 follows from the fact that g is a diffeomorphism and Ξ_n is an isometry, so $\|A_n\| = \|\Xi_{n+1} \circ Dg_{x_n} \circ \Xi_n^{-1}\|$ is uniformly bounded away from 0.

We begin by proving summability of δ_n. Assume $\operatorname{diam} P < \rho$, where ρ is the injectivity radius of M. Since $y_n \in \gamma^s(x_n)$ and $d(x_n, y_n) < \rho$, the tangent vector $v_n = (\exp_{x_n})|_{B(\rho,n)}^{-1}(y_n)$ lies in the stable cone $K_n^- \subset T_{x_n}M$, where $B(\rho, n) = \{v \in T_{x_n}M : \|v\| < \rho\}$. By symmetry of the vector field (3.1), we only need to consider the behavior of the trajectories $\{x_n\}$ and $\{y_n\}$ in the 'upper subsector' $S_j^s \cap S_j^u$, corresponding to the first quadrant in coordinates given by $\Phi_j \circ \phi_0$. (Here we denote by S_j^s, S_j^u, and Φ_j to be the subsets and functions described earlier as S_{kj}^s, S_{kj}^u, and Φ_{kj}, where we did not assume we only had one singularity.) Further assume $\widetilde{s}_2 := \operatorname{Im}(\Phi_j(\phi_0(y))) > s_2 := \operatorname{Im}(\Phi_j(\phi_0(x)))$, so that $\Delta s_2 := \widetilde{s}_2 - s_2 > 0$. Otherwise, exchange the sequences $\{x_n\}$ and $\{y_n\}$.

Recall the itinerary $\mathcal{I}(x) = \{0 = n_0 < n_1 < \cdots < n_{2L+1} = \tau(x)\} \subset \mathbb{Z}$ of the point $x \in \Lambda$, defined via $x_n \in \mathcal{U}_0$ if and only if $n_{2j-1} \le n < n_{2j}$. Consider $n_{2j} \le n < n_{2j+1}$, so $x_n \notin \mathcal{U}_0$. In coordinates, $g(s_1, s_2) = (\lambda s_1, \lambda^{-1} s_2)$, so $A_n = B_n$ are constant matrices, so $\delta_n = 0$.

Suppose now that $n_{2j+1} \le n < n_{2j+2}$. Denote by $D(s_1, s_2)$ the coefficient matrix of the variational equations of (3.1), given explicitly by

$$D(s_1, s_2) = \log \lambda \begin{bmatrix} \Psi_p(u) + 2s_1^2 \dot{\Psi}_p(u) & 2s_1 s_2 \dot{\Psi}_p(u) \\ -2s_1 s_2 \dot{\Psi}_p(u) & -\Psi_p(u) - 2s_2^2 \dot{\Psi}_p(u) \end{bmatrix}. \tag{7.7}$$

Let $s(t), \widetilde{s}(t) : [n, n+1] \to \mathbb{R}^2$ be solutions to (3.1) with initial condition $s(n) = x_n$ and $\widetilde{s}(n) = y_n$, and let $A_n(t)$ and $B_n(t)$ be the 2×2 Jacobian matrices

$$A_n(t) = d(\theta_t)((\Phi_{kj} \circ \phi_k)(x_n)) \quad \text{and} \quad B_n(t) = d(\theta_t)((\Phi_{kj} \circ \phi_k)(y_n)),$$

where $\theta_t : \mathbb{R}^2 \to \mathbb{R}^2$ is the time-t map of the flow of (3.1) on $\mathbb{R}^2$, for $n \le t \le n+1$. Then $A_n(1) = A_n$ and $B_n(1) = B_n$ from before, and $A_n(t)$ and $B_n(t)$ are the unique solutions to the systems of differential equations

$$\frac{dA_n(t)}{dt} = D(s(n+t))A_n(t) \quad \text{and} \quad \frac{dB_n(t)}{dt} = D(\widetilde{s}(n+t))B_n(t)$$

with initial conditions $A_n(0) = B_n(0) = \mathrm{Id}$. It follows that $A_n(t) - B_n(t)$ satisfies the differential equation

$$\frac{dA_n(t)}{dt} - \frac{dB_n(t)}{dt} = (D(s(n+t)) - D(\widetilde{s}(n+t)))A_n(t) + D(\widetilde{s}(n+t))(A_n(t) - B_n(t)).$$

Using the integrating factor $\exp \int_0^t D(\widetilde{s}(n+\tau))\,d\tau = B_n(t)$, this implies

$$A_n(t) - B_n(t) = B_n(t) \int_0^t B_n(t)^{-1}(D(s(n+t)) - D(\widetilde{s}(n+t)))A_n(t)\,d\tau. \tag{7.8}$$

Note that $\|D(s) - D(\widetilde{s})\| \le \|\partial D(\xi)\|\|\Delta s\|$, where $\partial D(s)$ denotes the total derivative of the matrix $D(s_1, s_2)$ and $\xi = (\xi_1, \xi_2)$, with $\min\{s_i, \widetilde{s}_i\} \le \xi_i \le \max\{s_i, \widetilde{s}_i\}$. This, in conjunction with (7.8) and Lemma 5.3, gives us

$$\begin{aligned}\|A_n - B_n\| &\le \|B_n(1)\| \sup_{0\le\tau\le1} \|B_n(\tau)^{-1}\|\|A_n(\tau)\|\|D(s(n+\tau)) - D(\widetilde{s}(n+\tau))\| \\ &\le \|B_n(1)\| \sup_{0\le\tau\le1} \|B_n(\tau)^{-1}\|\|A_n(\tau)\|\|\partial D(\xi(n+\tau))\|\|\Delta s(n+\tau)\| \\ &\le C_p \sup_{0\le\tau\le1} (\xi_1^2 + \xi_2^2)^{(p-4)/2p}(n+\tau)\|\Delta s(n+\tau)\|, \end{aligned} \tag{7.9}$$

where C_p is a constant that depends on p, but not on n (as the matrices $B_n(t)$ and $A_n(t)$ are uniformly bounded above and below in n and in t).

By condition (4) of Lemma 7.4 and the definition of δ_n,

$$\delta_n \le \frac{1}{cd(x,y)}\|A_n - B_n\| = \frac{1}{c}\frac{d(x_{n_{2j+1}}, y_{n_{2j+1}})}{d(x,y)}\frac{\|A_n - B_n\|}{d(x_{n_{2j+1}}, y_{n_{2j+1}})}.$$

We now claim that

$$\mathcal{D}_j := \sum_{n=n_{2j+1}}^{n_{2j+2}-1} \frac{\|A_n - B_n\|}{d(x_{n_{2j+1}}, y_{n_{2j+1}})} \le C, \tag{7.10}$$

where C is a constant independent of j. If this is true, then because $\delta_n = 0$ for $n_{2j} \le n < n_{2j+1}$, by (7.5),

$$\begin{aligned}\sum_{n=0}^{\tau(x)-1} \delta_n = \sum_{j=1}^{L}\sum_{n=n_{2j+1}}^{n_{2j+2}-1} \delta_n &= \sum_{j=1}^{L} \frac{1}{c}\frac{d(x_{n_{2j+1}}, y_{n_{2j+1}})}{d(x,y)} \sum_{n=n_{2j+1}}^{n_{2j+2}-1} \frac{\|A_n - B_n\|}{d(x_{n_{2j+1}}, y_{n_{2j+1}})} \\ &= \frac{C}{c}\sum_{j=1}^{L} \theta_1^j \le \widetilde{C},\end{aligned}$$

and because θ_1 is independent of $x, y \in P$, and c and C are both of order $\sup_n \|A_n\|$, $\widetilde{C}$ is also independent of our choice of x and y.

Recall that $[m_j^1, m_j^2] \subseteq [n2j+1, n_{2j+2} - 1]$ is the largest (possibly empty) interval of integers with $x_m \in D_{\widetilde{r}_1}$ for each $n \in [m_j^1, m_j^2]$, and $[m_j^1, T_j]$ is the largest time interval for which $s_1(t) \le s_2(t)$ for all $m_j^1 \le t \le T_j$. If $[m_j^1, m_j^2]$ is empty, then $s(t) \in (\Phi_{kj} \circ \phi_k)(D_{\widetilde{r}_0} \setminus D_{\widetilde{r}_1})$ for all $t \in [n_{2j+1}, n_{2j+2} - 1]$. In this instance, by Lemma 5.2, $n_{2j+2} - n_{2j+1} \le T$ is uniformly bounded, and hence (7.10) is a sum of uniformly boundedly many terms that are uniformly bounded, by (7.9).

Now suppose $[m_j^1, m_j^2]$ is non-empty. The sum in (7.10) splits into four different sums:

$$\mathcal{D}_j = \bigg(\sum_{n=n_{2j+1}}^{m_j^1-1} + \sum_{n=m_j^1}^{T_j-1} + \sum_{n=T_j}^{m_j^2} + \sum_{n=m_j^2+1}^{n_{2j+2}-1} \bigg) \frac{\|A_n - B_n\|}{d(x_{n_{2j+1}}, y_{n_{2j+1}})}. \tag{7.11}$$

We show that each of these sums is itself uniformly bounded. This is true for the first and fourth sum, because in these instances, $s(t)$ is in the annular region $(\Phi_{kj} \circ \phi_k)(D_{\widetilde{r}_0} \setminus D_{\widetilde{r}_1})$, and so the number of summands is uniformly bounded by Lemma 5.1.

To show this for the middle two sums, note that since $\widetilde{s}(t) \in \mathbb{R}^2$ is in the stable cone of $s(t)$ for all t in the domain, we have

$$|\Delta s_1| \leq \alpha \Delta s_2 \leq \Delta s_2. \tag{7.12}$$

First, suppose $m_j^1 \leq n \leq T_j - 1$, so that $s_1(t) \leq s_2(t)$. We would like to apply Lemma 5.5 in the interval $[m_j^1, n]$, so we require $\Delta s_2(m_j^1)/s_2(m_j^1) \leq (1-\alpha)/72$. This is attainable by choosing r_0 to be sufficiently small and Q in (7.1) to be sufficiently large. Applying Lemma 5.5 for $n \leq T_j - 1$, and $0 \leq \tau \leq 1$, we get

$$\begin{aligned} |\Delta s(n+\tau)| &\leq 2\Delta s_2(n+\tau) \\ &\leq 2\frac{\Delta s_2(m_j^1)}{s_2(m_j^1)} s_2(n+\tau)(1 + 2^{(p-2)/p} C_0 s_2(m_j^1)^{(2p-4)/p}(n+\tau-m_j^1))^{-\beta} \\ &\leq 2\frac{\Delta s_2(m_j^1)}{s_2(m_j^1)} s_2(n+\tau)(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n+\tau-m_j^1))^{-\beta} \end{aligned} \tag{7.13}$$

since $\beta = 2^{-(3p-2)/p}(1-\alpha) > 0$. Recalling that $\xi(t) = (\xi_1(t), \xi_2(t))$ is such that $\min\{s_i, \widetilde{s}_i\} \leq \xi_i \leq \max\{s_i, \widetilde{s}_i\}$ for $i = 1, 2$, (5.4) gives us

$$s_2^2(t) \leq (\xi_1^2 + \xi_2^2)(t) \leq 2(1+\kappa)^2 s_2^2(t) \leq C s_2^2(t)$$

as $\kappa = (\Delta s_2/s_2) \leq (1-\alpha)/72$. Estimates (7.9) and (7.13) give us

$$\begin{aligned} &\|A_n - B_n\| \\ &\quad \leq C\frac{\|\Delta s(m_j^1)\|}{s_2(m_j^1)} \sup_{0 \leq \tau \leq 1} s_2(n+\tau)^{(2p-4)/p}(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n+\tau-m_j^1))^{-\beta}, \end{aligned}$$

where we are using the fact that $|\Delta s_2| \leq \|\Delta s\|$. Applying Lemma 5.4(b) on the interval $[m_j^1, n+1]$ gives us

$$\begin{aligned} &\|A_n - B_n\| \\ &\quad \leq C\frac{\|\Delta s(m_j^1)\|}{s_2(m_j^1)} \sup_{0 \leq \tau \leq 1} s_2(m_j^1)^{(2p-4)/p}(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n+\tau-m_j^1))^{-1-\beta} \\ &\quad = C\|\Delta s(m_j^1)\| s_2(m_j^1)^{(p-4)/p}(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n-m_j^1))^{-1-\beta}. \end{aligned}$$

We make three observations. First, recalling that $n = m_j^1$ is the first time that $s(n)$ is within $\widetilde{r}_1$ of the origin, we observe that $s_2(m_j^1)$ is bounded above and below by a constant

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

multiple of $\widetilde{r}_1$, independent of $x \in \Lambda$ or $j = 1, \ldots, L$. Second, $\|\Delta s(m_j^1)\| = d(x_{m_j^1}, y_{m_j^1})$, by definition of our Riemannian metric in $\mathcal{U}_0$. Third, since Lemma 5.1 implies $m_j^1 - n_{2j+1}$ is bounded by a value independent of x or j, the value $d(x_{m_j^1}, y_{m_j^1})/(d(x_{2j+1}, y_{2j+1}))$ is uniformly bounded independently of $x, y \in \Lambda$ or $j \geq 1$. These three observations imply

$$\frac{\|A_n - B_n\|}{d(x_{2n+1}, y_{2n+1})} \leq C(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n - m_j^1))^{-1-\beta}.$$

Therefore,

$$\sum_{n=m_j^1}^{T_j-1} \frac{\|A_n - B_n\|}{d(x_{2n+1}, y_{2n+1})} \leq \sum_{n=m_j^1}^{\infty} C(1 + C_0 s_2(m_j^1)^{(2p-4)/p}(n - m_j^1))^{-1-\beta},$$

which is uniformly bounded in j. Therefore the second term in (7.11) is uniformly bounded in j.

Finally, we turn our attention to the case where $T_j \leq n \leq m_j^2$, where we have $s_1 \geq s_2$. By symmetry, we have that $T_j \geq (m_j^2 + m_j^1 - 2)/2$. By (7.12) and the second inequality in Lemma 5.5, we have

$$\begin{aligned}\|\Delta s(n+\tau)\| &\leq 2\Delta s_2(n+\tau)\\ &\leq 2\frac{\Delta s_2(T_j)}{s_1(T_j)} s_1(n+\tau)\left(\frac{1 + 2^{(p-2)/p} C_0 s_1(m_2^j)^{(2p-4)/p}(m_j^2 - n - \tau)}{1 + 2^{(p-2)/p} C_0 s_1(m_2^j)^{(2p-4)/p}(m_j^2 - T_j)}\right)^{\beta}.\end{aligned}$$

Since $\min\{s_i, \widetilde{s}_i\} \leq \xi_i \leq \max\{s_i, \widetilde{s}_i\}$ for $i = 1, 2$, we have $s_i - |\Delta s_i| \leq \xi_i \leq s_i + |\Delta s_i|$. In particular,

$$\xi_1^2 + \xi_2^2 \geq \xi_1^2 \geq (s_1 - |\Delta s_1|)^2 = s_1^2\left(1 - \frac{|\Delta s_1|}{s_1}\right)^2 \geq s_1^2\left(1 - \frac{\Delta s_2}{s_1}\right)^2 \geq C^{-1} s_1^2$$

and

$$\xi_1^2 + \xi_2^2 \leq (s_1 + |\Delta s_1|)^2 + (s_2 + |\Delta s_2|)^2 \leq 2(s_1 + \Delta s_2)^2 = 2s_1\left(1 + \frac{\Delta s_2}{s_1}\right)^2 \leq C s_1^2,$$

which both follow because $\Delta s_2/s_1$ is monotonically decreasing by (5.10). Together, these two estimates imply

$$(\xi_1(n+\tau)^2 + \xi_2(n+\tau)^2)^{(p-4)/2p} \leq C s_1(n+\tau)^{(p-4)/p}.$$

Applying (7.9) and inequality (a) in Lemma 5.4 to these inequalities gives us

$$\begin{aligned}\|A_n - B_n\| &\leq C \sup_{0\leq\tau\leq 1} [s_1(n+\tau)^{(p-4)/p}\|\Delta s(n+\tau)\|] \leq 2C\frac{\Delta s_2(T_j)}{s_1(T_j)}\\ &\quad\cdot \sup_{0\leq\tau\leq 1}\left[s_1(n+\tau)^{(2p-4)/p}\left(\frac{1 + 2^{(p-2)/p} C_0 s_1(m_j^2)^{(2p-4)/p}(m_j^2 - n - \tau)}{1 + 2^{(p-2)/p} C_0 s_1(m_j^2)^{(2p-4)/p}(m_j^2 - T_j)}\right)^{\beta}\right]\end{aligned}$$

$$\leq 2C\frac{\Delta s_2(T_j)}{s_1(T_j)}s_1(m_j^2)^{(2p-4)/p}$$
$$\cdot \sup_{0\leq\tau\leq 1}\left[\frac{(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-n-\tau))^{\beta-1}}{(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j))^{\beta}}\right].$$

By (5.6), since $s_1(m_j^2)$ and $s_2(m_j^1)$ are uniformly bounded,

$$\frac{|\Delta s_2(T_j)|}{s_1(T_j)}s_1(m_j^2)^{(2p-4)/p}=\frac{|\Delta s_2(T_j)|}{s_2(T_j)}s_1(m_j^2)^{(2p-4)/p}$$
$$\leq\frac{|\Delta s_2(m_j^1)|}{s_2(m_j^1)}s_1(m_j^2)^{(2p-4)/p}\leq C|\Delta s_2(m_j^1)|.$$

Furthermore, since $|\Delta s_2(m_j^1)|/(d(x_{n_{2j+1}},y_{n_{2j+1}}))$ is uniformly bounded, we finally obtain

$$\frac{\|A_n-B_n\|}{d(x_{n_{2j+1}},y_{n_{2j+1}})}\leq C\frac{(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-n))^{\beta-1}}{(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j))^{\beta}}.$$

Therefore,

$$\sum_{n=T_j}^{m_j^2}\frac{\|A_n-B_n\|}{d(x_{n_{2j+1}},y_{n_{2j+1}})}\leq C(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j))^{-\beta}$$
$$\times\sum_{n=T_j}^{m_j^2}(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-n))^{\beta-1}$$
$$\leq C(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j))^{-\beta}$$
$$\times\left(1+\int_0^{m_j^2-T_j}(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}\tau)^{\beta-1}\,d\tau\right)$$
$$\leq C(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j))^{-\beta}$$
$$\times\left(1+\frac{(1+2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}(m_j^2-T_j)^{(p-2)/p})^{\beta}}{2^{(p-2)/p}C_0s_1(m_j^2)^{(2p-4)/p}\beta}\right)$$
$$\leq C(1+(2^{(p-2)/p}\widetilde{r}_1^{(2p-4)/p}C_0\beta)^{-1}),$$

where the second inequality follows from the fact that the integrand is a decreasing function of τ, and the final inequality follows from the fact that $\widetilde{r}_1\leq s_1(m_j^2)$ by definition of m_j^2. Therefore the third sum of (7.11) is uniformly bounded. This completes the proof that δ_n is a summable sequence.

We now prove the estimates involving γ_k. For $n\in[n_{2j},n_{2j+1}-1]$, we have $x_n,y_n\notin\mathcal{U}_0$, where Dg_{x_n} and Dg_{y_n} are constant hyperbolic linear transformations. For these values for n, the maps contract angles uniformly, so there is a $\gamma>0$ for which $\gamma_n<\gamma<1$ for all n. For $n \in [m_j^1, m_j^2]$, we have $x_n \in \mathcal{U}_1$, so applying Lemma 5.6,

$$\prod_{n=m_j^1}^{m_j^2-1} \gamma_n \leq (1 + C_0 s_2(m_j^1)^{(2p-4)/p}(m_j^2 - m_j^1))^{-p/(p-2)}$$
$$\leq (1 + C(m_j^2 - m_j^1))^{-p/(p-2)},$$

since $s_2(m_j^1)$ is uniformly bounded. Because the interval of integers $[m_j^1, m_j^2]$ differs from $[n_{2j+1}, n_{2j+2} - 1]$ by a finite set, and the cardinality of this finite set is uniformly bounded in j by Lemma 5.1, there is a uniform constant $C' > 0$ for which

$$\prod_{j=n_{2j+1}}^{n_{2j+2}-1} \gamma_n \leq C'(1 + C(m_j^2 - m_j^1))^{-p/(p-2)} \leq C'.$$

In particular,

$$\prod_{n=n_{2j}}^{n_{2j+2}-1} \gamma_n \leq C' \gamma^{n_{2j+1}-n_{2j}} < \theta_3, \tag{7.14}$$

for some constant $\theta_3 > 0$. The third estimate of the lemma follows.

To prove the second and final estimates of the lemma, we observe that a similar estimate to (7.14) may be made with the upper limit replaced with $n_{2j+1} - 1$. In particular, for $n_{2j+1} \leq n \leq n_{2j+2} - 1$,

$$\prod_{k=n_{2j+1}}^{n} \gamma_j \leq C'(1 + C(n - n_{2j+1}))^{-p/(p-2)}$$

and

$$\prod_{n=n_{2j}}^{n_{2j+1}-1} \gamma_n < \theta_3'$$

for some $\theta_3' > 0$ that is uniformly bounded. Therefore,

$$\sum_{n=0}^{\tau(x)} \prod_{k=0}^{n} \gamma_k = \sum_{j=0}^{L(x)} \sum_{n=n_{2j}}^{n_{2j+2}-1} \prod_{k=0}^{n} \gamma_k = \sum_{j=0}^{L(x)} \bigg(\prod_{k=0}^{n_{2j}-1} \gamma_k \sum_{n=n_{2j}}^{n_{2j+2}-1} \prod_{k=n_{2j}}^{n} \gamma_k \bigg)$$
$$\leq \sum_{j=0}^{L(x)} \bigg(\theta_3^j \bigg(\sum_{n=n_{2j}}^{n_{2j+1}-1} \prod_{k=n_{2j}}^{n} \gamma_k + \prod_{k=n_{2j}}^{n_{2j+1}-1} \gamma_k \sum_{n=n_{2j+1}}^{n_{2j+2}-1} \prod_{k=n_{2j+1}}^{n} \gamma_k \bigg) \bigg)$$
$$\leq \sum_{j=0}^{L(x)} \bigg(\theta_3^j \bigg(\sum_{n=n_{2j}}^{n_{2j+1}-1} \gamma^{n-n_{2j}} + \theta_3' \sum_{n_{2j+1}}^{n_{2j+2}-1} (1 + C(n - n_{2j+1}))^{-p/(p-2)} \bigg) \bigg).$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Because the two sums in the inner parentheses above are both uniformly bounded, there is a $C' > 0$ for which

$$\sum_{n=0}^{\tau(x)} \prod_{k=0}^{n} \gamma_k \leq C'' \sum_{j=0}^{L(x)} \theta_3^j,$$

which gives us the second estimate in the lemma. $\square$

We continue with the proof of the theorem. Observe that

$$\left(\Xi_{\tau(x)}^{-1} \circ \prod_{n=0}^{\tau(x)-1} A_n \circ \Xi_0\right)(v) = D(g^{\tau(x)})_x v \quad \text{for all } v \in T_x M,$$

and

$$(P_{\tau(x)}^{-1} \circ \Xi_{\tau(x)}^{-1} \circ \prod_{n=0}^{\tau(x)-1} B_n \circ \Xi_0 \circ P_0)(v) = D(g^{\tau(x)})_y v \quad \text{for all } \in T_y M.$$

In particular, since both Ξ_n and P_n are linear isometries for all $n \geq 0$, we have

$$\left\| \prod_{n=0}^{\tau(x)-1} A_n \overline{v} \right\| = \| D(g^{\tau(x)})_x v \| \quad \text{for all } v \in T_x M,$$

and

$$\left\| \prod_{n=0}^{\tau(x)-1} B_n \overline{w} \right\| = \| D(g^{\tau(x)})_y w \| \quad \text{for all } w \in T_y M,$$

where $\overline{v} = \Xi_0 v \in \mathbb{R}^2$ and $\overline{w} = (\Xi_0 \circ P_0) w \in \mathbb{R}^2$. Additionally, for $v \in T_{x_n} M$ and $w \in T_{y_n} M$,

$$\angle(Dg_{x_n} v, (P_{n+1} \circ Dg_{y_n}) w) = \angle(A_n \overline{v}, B_n \overline{w}),$$

where $\overline{v} = \Xi_n v$ and $\overline{w} = (\Xi_n \circ P_n) w$.

Now, suppose $v \in K^+(x)$ and $w \in K^+(y)$, and once again denote $\overline{v} = \Xi_0 v$ and $\overline{w} = (\Xi_0 \circ P_0) w$. Since $P_0 w \in K^+(x)$, Lemmas 7.4 and 7.5 yield

$$\begin{aligned} |\log \| D(g^{\tau(x)})_x v \| \| D(g^{\tau(x)})_y w \|| &= \left| \log \frac{\| \prod_{n=0}^{\tau(x)-1} A_n \overline{v} \|}{\| \prod_{n=0}^{\tau(x)-1} B_n \overline{w} \|} \right| \\ &\leq C\widetilde{C}(d(x, y) + \angle(v, P_0 w)), \end{aligned} \tag{7.15}$$

where we are using the fact that $\angle(v, P_0 w) = \angle(\overline{v}, \overline{w})$. Furthermore, for $v \in T_x M$ and $w \in T_y M$, the definition of γ_n and Lemma 7.5 give us

$$\begin{aligned} &\frac{\angle(D(g^{\tau(x)})_x v, (P_{\tau(x)} \circ D(g^{\tau(x)})_y) w)}{\angle(v, P_0 w)} \\ &\quad = \prod_{n=0}^{\tau(x)-1} \frac{\angle(Dg_{x_n}(Dg_x^n v), (P_{n+1} \circ Dg_{y_n})(Dg_y^n w))}{\angle(Dg_x^n v, P_n(Dg_n^n w))} \end{aligned}$$

$$
\begin{aligned}
&= \prod_{n=0}^{\tau(x)-1} \frac{\angle(A_n(\Xi_n(Dg_x^n v)), B_n((\Xi_n \circ P_n)(Dg_y^n w)))}{\angle(\Xi_n(Dg_x^n v), (\Xi_n \circ P_n)(Dg_y^n w))} \\
&\leq \prod_{n=0}^{\tau(x)-1} \gamma_n \leq \theta_2. \qquad (7.15)
\end{aligned}
$$

Denote $\widehat{G} : \Lambda \to \Lambda$ by $\widehat{G}(x) = g^{\tau(x)}(x)$. If $v^n \in E^u(\widehat{G}^n(x))$ and $w^n \in E^u(\widehat{G}^n(y))$, then there are $v \in E^u(x)$ and $w \in E^u(y)$ such that $v^n = D\widehat{G}_x^n v$ and $w^n = D\widehat{G}_y^n w$. By (7.15), (7.15), and condition (Y3),

$$
\begin{aligned}
\left| \log \frac{\|D\widehat{G}_{\widehat{G}^n(x)} v^n\|}{\|D\widehat{G}_{\widehat{G}^n(y)} w^n\|} \right| &\leq C\widetilde{C}(d((g^{\tau(x)})^n(x), (g^{\tau(x)})^n(y)) \\
&\quad + \angle(D(g^{\tau(x)})_x^n v, P_{\tau(x)} D(g^{\tau(x)})_y^n w)) \\
&\leq C\widetilde{C}(a^n d(x, y) + \theta_2^n \angle(v, P_0 w)).
\end{aligned}
$$

Since $0 < a, \theta_2 < 1$, this proves (Y4)(a).

8. *Proof of Theorem 4.1*

We now drop our assumption that the pseudo-Anosov diffeomorphism g admits only one singularity. By Proposition 6.2 and Theorem 7.3, since $g : M \to M$ is a Young diffeomorphism, the geometric potential $\varphi_1(x) = -\log |Dg|_{E^u(x)}|$ admits an equilibrium measure, which is the unique g-invariant SRB measure. This is the same measure as μ_1 introduced in Proposition 3.2, as μ_1 is absolutely continuous along the unstable foliations and thus an SRB measure. (This justifies our use of the notation μ_1 to describe this measure).

By Proposition 3.1, the pseudo-Anosov homeomorphism f and the pseudo-Anosov diffeomorphism g possess the same topological and combinatorial data, including topological entropy. Thus the number S_n of s-sets $\Lambda_i^s \subset \Lambda$ with inducing time $\tau_i = n$ for g is the same for both f and g. Therefore by Lemma 7.2, there is an $h < h_{\text{top}}(g) = h_{\text{top}}(f)$ such that $S_n \leq e^{hn}$.

Recall that ν is the measure on M given locally by the product of lengths of local stable and unstable leaves described in Theorem 2.8, and μ_1 is the measure given by the Riemannian metric ζ described in Proposition 3.2. By Theorem 2.8, ν has a density with respect to μ_1, which vanishes at the singularities. By Proposition 10.13 and Lemma 10.22 of [6], $h_\nu(f) = h_{\text{top}}(f) = \log \lambda$, so in fact $h < h_\nu(f)$. Since $\nu = \mu_1$ on $M \setminus \mathcal{U}_0$, and $\mu_1(\mathcal{U}_0)$ may be made arbitrarily small by shrinking r_0 if necessary, the Pesin entropy formula implies

$$
\begin{aligned}
h_{\mu_1}(g) &= \int_M \log |Dg|_{E^u(x)}| \, d\mu_1(x) \\
&= \int_{M \setminus \mathcal{U}_0} \log \lambda \, d\nu + \int_{\mathcal{U}_0} \log |Dg|_{E^u(x)}| \, d\mu_1(x) < h_\nu(f) + \varepsilon, \qquad (8.1)
\end{aligned}
$$

where $\varepsilon > 0$ is as small as we need. From this we conclude that $h < h_{\mu_1}(g)$. Hence by Proposition 6.2, there is a $t_0 < 0$ for which, for all $t \in (t_0, 1)$, there is a measure μ_t on P that is an equilibrium state for the geometric t-potential φ_t.

Since f is Bernoulli, every power of f is ergodic, so f satisfies the arithmetic condition. Since f and g are topologically conjugate, this is also true for g.

We now prove (6.3). If $x, y \in \Lambda_i^s$ and $y \in \gamma^s(x)$, the distance $d(f^j(x) f^j(y))$ decreases with j. On the other hand, if $y \in \gamma^u(x)$, then $d(f^j(x), f^j(y))$ increases with j, but is bounded by diam P when $j = \tau(x)$. An application of the triangle inequality and hyperbolic product structure of Λ now yields (6.3). It now follows that μ_t has exponential decay of correlations and satisfies the central limit theorem, by Proposition 6.2. Since (M, g, μ_t) has exponential decay of correlations, this dynamical system is mixing. By [**17**, Theorem 2.3], (M, g, μ_t) is Bernoulli.

To show r_0 may be chosen to accommodate any t_0, we show that as $r_0 \to 0$, we may take $t_0 \to -\infty$. Fix $\varepsilon > 0$, and choose $x \in \Lambda_i^s$. Recall that $g = f$ outside of $\widetilde{\mathcal{U}}_0$; in particular, the local stable and unstable leaves are unchanged outside of $\widetilde{\mathcal{U}}_0$. Assume x is a generic point for the SRB measure μ_1. Let $\widetilde{\mathcal{U}}_2 = \bigcup_{k=1}^m \phi_k^{-1}(D_{\tilde{r}_1/4})$, and write τ_i as

$$\tau_i = \sum_{j=1}^{s} n_j,$$

where the integers n_j are chosen as follows:

- the integer n_1 is the first time when $g^{n_1}(x) \in \widetilde{\mathcal{U}}_0 \setminus \widetilde{\mathcal{U}}_2$;
- the integer n_2 is the first time after n_1 when $g^{n_1+n_2}(x) \in \widetilde{\mathcal{U}}_2$;
- the number n_3 is the first time after $n_1 + n_2$ when $g^{n_1+n_2+n_3}(x) \in \widetilde{\mathcal{U}}_0 \setminus \widetilde{\mathcal{U}}_2$;
- the number n_4 is the first time after $n_1 + n_2 + n_3$ when $g^{n_1+n_2+n_3+n_4}(x) \notin \widetilde{\mathcal{U}}_0$;

and so on. It is possible that some n_j may be equal to 0, but this does not change our calculations. Observe $Q \leq n_1$, where Q is the number from (7.1). If r_0 is sufficiently small, Q is large enough to ensure that

$$\log |J^u g^{n_1}(x)| \leq n_1(\log \lambda + \varepsilon). \tag{8.2}$$

By (7.7), for $x \in \widetilde{\mathcal{U}}_0 \setminus \widetilde{\mathcal{U}}_2$, we have $\log |J^u g(x)| \leq \log N$ for some constant N independent of r_0 or of the number of prongs p. Therefore,

$$\log |J^u g^{n_2}(x)| \leq n_2 \log N \quad \text{and} \quad \log |J^u g^{n_4}(x)| \leq n_4 \log N. \tag{8.3}$$

For $x \in \widetilde{\mathcal{U}}_2$, if x is in a neighborhood of a singularity with p prongs, $\Psi_p(u) = (p/2)^{(2p-4)/p} u^{(p-2)/p}$ and $\dot{\Psi}_p(u) = (p-2)/p(p/2)^{(2p-4)/p} u^{-2/p}$. By (7.7), for such points x, $\log |J^u g(x)| \leq \log \lambda$. Therefore,

$$\log |J^u g^{n_3}(x)| \leq n_3 \log \lambda. \tag{8.4}$$

Similar estimates hold for the other n_j. Observe that

$$\log |J^u \widehat{G}(x)| \leq \sum_{j=1}^{s} \log |J^u g^{n_1+\cdots+n_j}(g^{n_1+\cdots+n_{j-1}}(x))|. \tag{8.5}$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Similarly to Lemma 5.2, the number of iterates the orbit of x spends in $\widehat{\mathcal{U}}_0 \setminus \widehat{\mathcal{U}}_2$ is bounded above by a constant T_0' independent of both r_0 and p. It follows from (8.2)–(8.5) and the definition of λ_1 in (6.1) that

$$\log \lambda_1 \leq \log \lambda + \varepsilon + \frac{2T_0' \log N}{Q} \leq \log \lambda + 2\varepsilon.$$

Meanwhile, (8.1) implies that for sufficiently small r_0,

$$\left| \int_M \log |Dg|_{E^u(x)}| \, d\mu_1(x) - \log \lambda \right| < \varepsilon, \tag{8.6}$$

or equivalently,

$$\log \lambda - \varepsilon \leq h_{\mu_1}(g) \leq \log \lambda + \varepsilon.$$

Furthermore, one can show $\log \lambda_1 \geq h_{\mu_1}(g)$ (see [**13**, Remark 3], which is a general statement about Young diffeomorphisms). Therefore,

$$\log \lambda - \varepsilon \leq h_{\mu_1}(g) \leq \log \lambda_1 \leq \log \lambda + 2\varepsilon.$$

It follows that the difference $\log \lambda_1 - h_{\mu_1}(g)$ can be made arbitrarily small if r_0 is chosen to be sufficiently small. By (6.2), this shows that $t_0 \to -\infty$ as $r_0 \to 0$.

We now show how μ_t may be extended to a measure on M, as opposed to a measure only on images of the base of the tower. Suppose we have another element $\widetilde{P}$ of the Markov partition satisfying (7.1). As above, there is a $\widetilde{t}_0 = t_0(\widetilde{P}) < 0$ for which. for every $t \in (\widetilde{t}_0, 1)$, there is a unique equilibrium state $\widetilde{\mu}_t$ for the geometric t-potential among all measures μ for which $\mu(\widetilde{P}) > 0$, and $\widetilde{\mu}_t(U) > 0$ for all open sets $\widetilde{U} \subset P$. Since g is topologically conjugate to a Bernoulli shift, g is topologically transitive. Therefore for any open sets $\widetilde{U} \subset \widetilde{P}$ and $U \subset P$, there is an integer $k \geq 0$ for which $g^k(\widetilde{U}) \cap U \neq \emptyset$. By invariance of $\widetilde{\mu}_t$ and μ_t under g, it follows that $\mu_t = \widetilde{\mu}_t$.

Consider now an element of the Markov partition that does not satisfy (7.1). If r_0 is sufficiently small, the union of all partition elements satisfying (7.1) form a closed set $Z \subset M$, whose complement is a neighborhood of the singular set S with each component containing a single singularity. If ω is a g-invariant probability measure that does not give weight to partition elements in Z, then ω is a convex combination of the δ-measures concentrated at the singularities. If P is our partition element in the proof of Theorem 7.3, we observe $\omega(P) = 0$, so ω is clearly out of consideration as an equilibrium measure for φ_t. So any equilibrium measure for (M, g) must charge partition elements in Z. Therefore, set

$$t_0 = \max_{P \in \mathcal{P},\, P \cap Z \neq \emptyset} t_0(P).$$

Since $t_0 \to -\infty$ as $r_0 \to 0$ and $\mu_t(P) > 0$ for $t_0 < t < 1$, this t_0 suffices for statement (1) of Theorem 4.1.

To prove statement (2) of Theorem 4.1, suppose ω is an invariant ergodic Borel probability measure. By the Margulis–Ruelle inequality,

$$h_\omega(g) \leq \int_M \log |Dg|_{E^u(x)}| \, d\omega(x) = -\int_M \varphi_1 \, d\omega.$$

https://doi.org/10.1017/etds.2021.43 Published online by Cambridge University Press

Hence $h_\omega(f) + \int \varphi_1 \, d\omega \le 0$. If ω has only 0 as a non-negative Lyapunov exponent almost everywhere, then $\log |Dg|_{E^u(x)}| = 0$ ω-almost everywhere. The only point at which $\log |Dg|_{E^u(x)}| = 0$ is at the singularities of g, so ω is a convex combination of the δ-measures at the singularities. In this instance, we have $h_\omega(g) + \int \varphi_1 \, d\omega = 0$, so $P(\varphi_1) = 0$, and ω is an equilibrium state for φ_1.

On the other hand, part (1) of Proposition 6.2 guarantees the existence of an SRB measure μ_1 for g. In particular, μ_1 is a smooth measure, so by the Pesin entropy formula, $h_\mu(f) + \int \varphi_1 \, d\mu = 0$, so μ is also an equilibrium measure. Any other equilibrium measure with positive Lyapunov exponents also satisfies the entropy formula. By [**10**], such a measure is also an SRB measure, and by [**16**] this SRB measure is unique. This proves statement (2).

Finally, to prove statement (3) of Theorem 4.1, fix $t > 1$, and let ω be an ergodic measure for g. Again, by the Margulis–Ruelle inequality,

$$h_\omega(g) \le t \int \log |Dg|_{E^u(x)}| \, d\omega,$$

with equality if and only if $\int \log |Dg|_{E^u(x)}| \, d\omega = 0$. In particular, we have equality if and only if ω has zero Lyapunov exponents ω-almost everywhere. As we saw, the only measures satisfying this are convex combinations of δ-measures at singularities, so $h_\omega(g) + \int \varphi_t \, d\omega \le 0$, with equality only for $\omega = \sum \lambda_i \delta_{x_i}$, with $\sum \lambda_i = 1$. Hence the only equilibrium states for φ_t with $t > 1$ are convex combinations of δ-measures at singularities.

Acknowledgments. I would like to thank Penn State University and the Anatole Katok Center for Dynamical Systems and Geometry where this work was done. I also thank my advisor, Y. Pesin, for introducing me to this problem and for valuable input over the course of my investigation into pseudo-Anosov systems.

References

[1] L. Barreira, Y. Pesin. *Nonuniform Hyperbolicity: Dynamics of systems with nonzero Lyapunov exponents (Encyclopedia of Mathematics and Its Applications, 115)*. Cambridge University Press, 2007.

[2] R. Bowen. *Equilibrium states and ergodic theory of Anosov diffeomorphisms (Lecture Notes in Mathematics, 470)*. Springer, 1975.

[3] J. Buzzi, S. Crovisier, and O. Sarig. Measures of maximal entropy for surface diffeomorphisms. *Preprint*. 2019, arXiv:1811.02240.

[4] V. Climenhaga, Y. Pesin. Building thermodynamics for non-uniformly hyperbolic maps. *Arnold Math. J.* **3**(1) (2017), 37–82.

[5] V. Climenhaga, Y. Pesin, and A. Zelerowicz. Equilibrium measures for some partially hyperbolic systems. *J. Mod. Dyn.* **16** (2020), 155–205.

[6] A. Fathi, F. Laudenbach, and V. Poénaru, Travaux de Thurston sur les surfaces, in Séminaire Orsay. *Astérisque* **66–67** (1979).

[7] F. Faure, S. Gouëzel, E. Lanneau. Ruelle spectrum of linear pseudo-Anosov maps. *J. Éc. Polytech. Math.* **6** (2019), 811–877.

[8] M. Gerber, A. Katok. Smooth models of pseudo-Anosov maps. *Ann. Sci. Éc. Norm. Supér. (4)* **15**(4) (1982), 173–204.

[9] A. Katok. Bernoulli diffeomorphisms on surfaces. *Ann. of Math. (2)*, **110**(3) (1979), 529–547.

[10] F. Ledrappier and L. S. Young, The metric entropy of diffeomorphisms, *Bull. Amer. Math. Soc. (N.S.)*, **11**(2) (1984), 343–346.
[11] R. Penner. A construction of pseudo-Anosov homeomorphisms. *Trans. Amer. Math. Soc.* **310**(1) (1988), 179–197.
[12] Y. Pesin, S. Senti, and K. Zhang. Thermodynamics of towers of hyperbolic type. *Trans. Amer. Math. Soc.* **368**(12) (2016), 8519–8552.
[13] Y. Pesin, S. Senti, and K. Zhang. Thermodynamics of the Katok map (revised version). *Ergod. Th. & Dynam. Sys.* **39**(3) (2019), 764–794.
[14] M. Pollicot, H. Weiss. Multifractal analysis of the Lyapunov exponent for continued fraction and Manneville-Pomeau transformations and applications to Diophantine approximations. *Comm. Math. Phys.* **207** (1999), 145–171.
[15] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.*, **74**(2) (1980), 189–197.
[16] F. Rodriguez-Hertz, M. A. Rodriguez-Hertz, A. Tahzibi, and R. Ures. Uniqueness of SRB measures for transitive diffeomorphisms on surfaces, *Comm. Math. Phys.* **306**(1) (2011), 35–49.
[17] F. Shahidi, A. Zelerowicz. Thermodynamics via inducing, *J. Stat. Phys*, **175**(2) (2019), 351–383.
[18] W. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc. (N.S.)*, **19**(2) (1988), 417–431.
[19] D. Veconi. Equilibrium states of almost Anosov diffeomorphisms. *Discrete Contin. Dyn. Syst.* **40**(2) (2020), 767–780.
[20] R. Venegeroles. Thermodynamic phase transitions for Pomeau-Manneville maps. *Phys. Rev.* **86**(2) (2012), 021114.
[21] T. Wang. Unique equilibrium states, large deviations and Lyapunov spectra for the Katok map. *Ergod. Th. & Dynam.* Sys. doi: 10.1017/etds.2020.26.2020. Published online 20 March 2020.